UNIVERSITAT DE BARCELONA

Dissercting the activities of Capicua, Dorsal and Groucho in *Drosophila* dorsoventral patterning

Aikaterini Papagianni

TESIS DOCTORAL

Dissecting the activities of Capicua, Dorsal and Groucho in *Drosophila* dorsoventral patterning

Aikaterini Papagianni

Barcelona, September 2017

Departamento de Genética
Programa de Doctorado de Genética
Facultad de Biología
Universidad de Barcelona

Dissecting the activities of Capicua, Dorsal and Groucho in *Drosophila* dorsoventral patterning

Memoria presentada por
Aikaterini Papagianni
Para optar al grado de

Doctora en Biología por la Universidad de Barcelona

Esta tesis ha sido realizada en el Departamento de Biologia del Desarrollo del Instituto de Biología Molecular de Barcelona (Consejo Superior de Investigaciones, Parc Científic de Barcelona), bajo la supervisión del Dr. Gerardo Jiménez Cañero.

El director	El alumno	El tutor
Dr. Gerardo Jiménez	Aikaterini Papagianni	Dr. Francesc Cebrià

INDEX

ABBREVIATIONS **1**

INTRODUCTION **3**

1. EARLY DEVELOPMENT OF DROSOPHILA **3**
1.1 AP PATTERNING DURING EMBRYOGENESIS 5
1.2 TERMINAL PATTERNING DURING EMBRYOGENESIS 5
1.3 DV PATTERNING DURING EMBRYOGENESIS 6

2. MATERNAL CONTROL OF DEVELOPMENT **9**

3. AXIS SPECIFICATION DURING OOGENESIS. **10**
3.1 DV POLARITY ESTABLISHMENT DURING OOGENESIS 12
3.2 TERMINAL PATTERNING DURING OOGENESIS 15

4. DORSAL, THE MAIN NUCLEAR EFFECTOR IN DV POLARITY **16**
4.1 STRUCTURE AND FUNCTIONS OF DL 16
4.2 TRANSCRIPTIONAL REGULATION BY DL 17
4.3 DL-MEDIATED ACTIVATION 18
4.4 DL-MEDIATED REPRESSION 19

5. CAPICUA, A DL COREPRESSOR CANDIDATE **24**
5.1 STRUCTURE AND FUNCTIONAL DOMAINS OF CIC 24
5.2 HUMAN HOMOLOGUES OF CIC AND THEIR FUNCTIONS 25
5.3 MECHANISM OF CIC FUNCTION IN *DROSOPHILA* 25
5.4 REGULATION AND FUNCTIONS OF CIC DURING *DROSOPHILA* DEVELOPMENT..... 26

OBJECTIVES **31**

RESULTS **33**

1. FUNCTIONAL ANALYSIS OF CIC IN THE ESTABLISHMENT OF THE DV AXIS **33**
1.1 CIC MUTATIONS REVEAL A DOUBLE FUNCTION IN THE ESTABLISHMENT OF THE DV AXIS 33
1.2 THE MAIN EMBRYONIC CIC ISOFORM IS INVOLVED IN EMBRYONIC DV PATTERNING 34

2. MOLECULAR MECHANISM OF CIC FUNCTION IN EMBRYONIC DV PATTERNING **39**
2.1 CIC REPRESSES TYPE III GENES THROUGH THE AT-RICH SITES OF THE VRE ELEMENTS 39
2.2 CIC AND DL SHOW A SYNERGISTIC FUNCTION FOR REPRESSION OF DORSAL- SPECIFIC GENES 48
2.3 THE CIC N2 MOTIF RECRUITS GRO TO THE DL REPRESSOR COMPLEX 54

3. STUDY OF CIC AS AN RTK SENSOR ELEMENT DURING THE ESTABLISHMENT OF THE EMBRYONIC DV AXIS **61**

DISCUSSION **65**
1. THE HMG-BOX FAMILY CIC REPRESSOR HAS A DUAL FUNCTION IN DV AXIS ESTABLISHMENT 65
2. CIC BINDS, WITH LOW AFFINITY, TO THE AT-RICH SITES IN SILENCER ELEMENTS OF DORSAL-SPECIFIC GENES AND REQUIRES THE PRESENCE OF DL TO REPRESS THEM 68
MECHANISMS OF CIC AND DL TRANSCRIPTIONAL SYNERGY 71
4. INTERACTIONS BETWEEN THE CIC N2 MOTIF AND GRO ARE ESSENTIAL FOR REPRESSION OF DORSAL-SPECIFIC GENES IN THE EARLY EMBRYO 73
5. CIC IS A SENSOR OF RTK SIGNALLING DURING DV PATTERNING 76

CONCLUSIONS **79**

MATERIALS AND METHODS **81**

1. SYNTNETIC DNA CONSTRUCTS **81**
1.1 RECOMBINANT DNA TECHNIQUES 81
1.2 TREATMENT OF DNA 81

2. RNA ANTISENSE PROBE GENERATION AND LABELLING **92**

3. GENETIC ANALYSES..... 94
3.1 FLY CULTURE AND HUSBANDRY.....94
3.2 GENETIC MANIPULATION METHODS.....94
4. OVARY ANALYSIS.....101
4.1 COLLECTION AND FIXATION OF FEMALE OVARIES.....101
4.2 IMMUNOHISTOCHEMISTRY OF FIXED OVARIES.....101
5. EMBRYO ANALYSIS.....102
5.1 PREPARATTION OF EMBRYONIC CUTICLES.....102
5.2 COLLECTION AND FIXATION OF EMBRYOS.....102
5.3 IN SITU HYBRIDIZATION IN FIXED EMBRYOS.....103
5.4 DOUBLE FLUORESCENT IN SITU HYBRIDIZATION IN FIXED EMBRYOS.....105
5.5 IMMUNOHISTOCHEMISTRY IN FIXED EMBRYOS.....107
5.6 PREABSORBING OF PRIMARY ANTIBODIES.....108
6.1 EXPRESSION AND PURIFICATION OF RECOMBINANT PROTEINS.....109
6.2 PROTEIN-PROTEIN INTERACTION ANALYSES- GST PROTEIN PULLDOWN ASSAY.....111
6.3. PROTEIN-DNA INTERACTIONS-EMSA ASSAY.....111
BIBLIOGRAPHY.....115
SUMMARY IN SPANISH.....129

ABBREVIATIONS

Abbreviation (genes)	Full name
aos	*argos*
Atx	Ataxin
bHLH	basic Helix Loop Helix
bcd	*bicoid*
bwk	*bullwinkle*
cact	*cactus*
cic	*capicua*
Cy	Curly
Da	Daughterless
Dl	Dorsal
dpERK	double phospho Extracellular Related Kinase
Dpn	Deadpan
DSP1	Dorsal Switch Protein 1
Dri	Dead-Ringer
EGF	Epidermal Growth Factor
En	Engrailed
Grk	Gurken
Gro	Groucho
gt	*giant*
H	Hairy
hb	*hunchback*
HMG	High Mobility Group
Hox	Homeobox
Hu	Humeral
IκB	Inhibitor of kappa beta
ind	*intermediate neuroblasts defective*
kni	*knirps*
kr	*kruppel*
MAPK	Mitogen Activated Protein Kinase
mirr	*mirror*
nos	*nanos*
Ndl	Nudel
NF-κB	Nuclear Factor Kappa Beta
NTF-1	Nuclear Transcription Factor 1
osk	*oskar*
pip	*pipe*
rho	*rhomboid*
RTK	Receptor Tyrosine Kinase
Sb	Stubble
sna	*snail*
sim	*single-minded*
sog	*short gastrulation*
Spz	Spätzle
sxl	*sex-lethal*
tld	*tolloid*
tll	*tailless*
trk	*trunk*
tsl	*torso-like*
Tor	Torso

Tub	Tubuline
twi	*twist*
zen	*zerknüllt*

Other abbreviations	**Full name**
AP	anteroposterior
ana	*annanasai*
bp	base pairs
Cas9	Caspase 9
CBS	Cic binding sites
ChIP	Chromatin Immunoprecipitation
CRISPR	Clustered Regulatory Interspaced Short Palindromic Repeats
CRM	*cis*-Regulatory Module
CTD	C-terminal domain
DAPI	4',6-diamidino-2-phenylindole
DFS	Dominant Female Sterile
DNA	Deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
dUTP	deoxyuridine-5'-triphosphate
DV	dorsoventral
Flp	Flipase
FRT	Flipase Recognition Target
GLC	Germ-line clones
GST	Glutathionine S transferase
h	hours
His	Histidine
kb	kilobases
KO	knock-out
M	molar
mel	*melanogaster*
Min	minutes
ml	Milliliters
mol	moles
mut	mutant
µg	micrograms
NHEJ	Non homologous end joining
OD	Oligoderndroglioma
Pol	Polymerase
PCR	Polymerase Chain reaction
pse	*pseudooscura*
RHD	Rel Homology Domain
RNA	Ribonucleic acid
UAS	Upstream Activator Sequence
UTR	Untranslated Region
vir	*virilis*

INTRODUCTION

Animal development is a complex process during which genetically identical cells, originating from the division of a single diploid cell, the egg, differentiate and assemble into highly organized tissues and organs. In bilaterian animals, development also involves the formation of two asymmetry axes: the anterior to posterior (AP) and dorsal to ventral (DV). Morphogenesis and determination of the body axes are accomplished through the precise and coordinated function of molecules that constantly receive, transmit and interpret physical and chemical signals, inducing the adequate changes in the cell. Many of these changes are done at the level of the genetic material and result in its differential activation or suppression. A powerful and versatile model used to understand gene regulation is the fruit fly, *Drosophila melanogaster,* which due to its rapid and robust development has been established as the model of choice for genetic studies. In this thesis, we focus on the molecular mechanisms underlying the patterning of the early *Drosophila* embryo, specifically the establishment of its DV body axis, a process which has traditionally been of interest in the field of developmental biology. Breakage of the embryo symmetry requires a number of factors that act soon after fertilization, but also relies on information provided by the mother. For this reason, in the next sections we will introduce both the processes of embryogenesis and oogenesis.

1. Early development of *Drosophila*

Brief overview of embryogenesis

Embryogenesis of *Drosophila* is a process that takes roughly one day and is divided into 17 stages (figure 1). Once the fertilized *Drosophila* egg is laid, it undergoes a series of rapid, synchronous mitotic divisions. During the first cleavages, no cytokinesis or cellular membrane formation occur. Rather, the dividing nuclei are localized inside a large yolk, which serves as a common cytoplasm, and surrounded by a unique membrane, the plasma membrane. This structure, the syncytial blastoderm, is the form as which the embryo exists during the first three hours of development. After the first nine divisions, which take place in the center of the common yolk, the nuclei migrate to the periphery while they continue to divide, at lower rates. A group of them migrates to the posterior region and adopts a distinct fate, forming the precursors of the future germline.

In the meanwhile, the rest of the nuclei, that have reached the periphery of the syncytium divide for a few rounds more, with each cycle taking about 25 minutes to be completed. During nuclear cleavage 14 (stage 5), the oocyte plasma membrane folds inwards, between the nuclei, dividing them to individual cells. At this point, the embryo enters the cellular blastoderm stage, during which mitotic divisions slow down and the genome is activated. Upon cellularization, morphogenesis also starts to occur: the embryo begins to gastrulate, germ layers are formed, body segments appear and differentiation begins to become obvious along the two orthogonal axes of the embryo. After various events of migration, invagination and cell group formation, the embryo is ready to hatch and enter the first stage of the larval phase. Since early embryogenesis, two assymetry axes, which are maintained until the adult life, are distinguished: the anterior to posterior (AP) and the dorsal to ventral (DV). Finally, the extremes of the embryo adopt a different fate, the terminal fate. In the next sections we will discuss how these asymmetries are established in the early embryo.

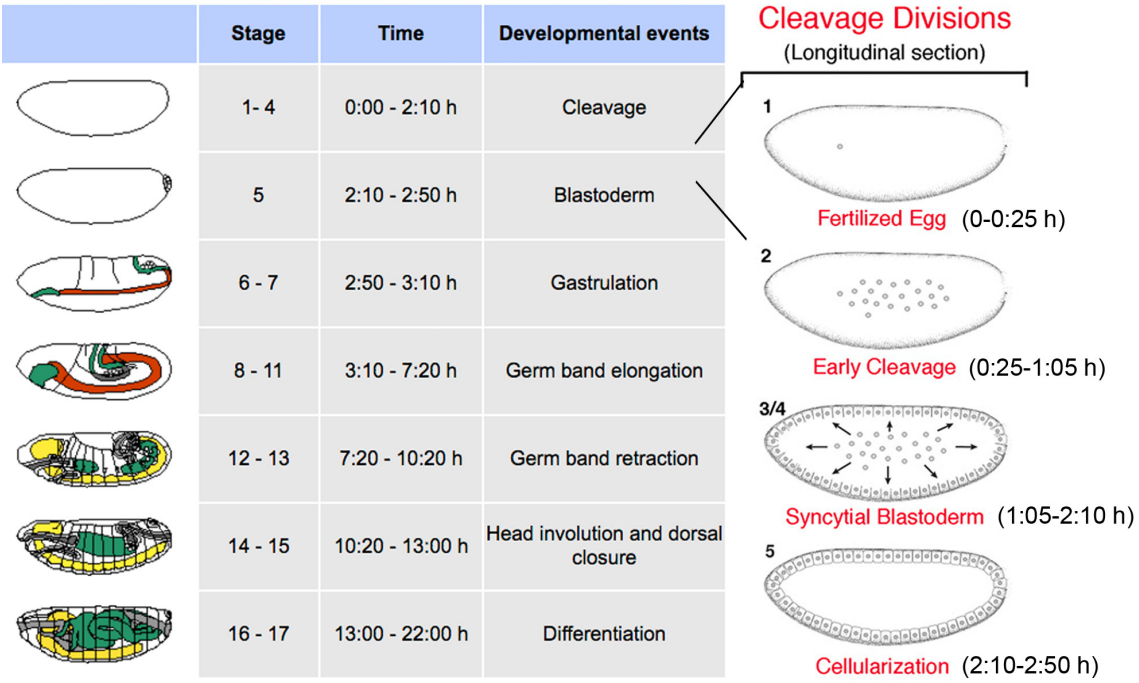


Figure 1. Overview of *Drosophila* embryogenesis stages. Adapted from Fly*Move* and Hales et al., 2015.

1.1 AP patterning during embryogenesis

The head, thorax and abdomen are formed along the embryo AP axis. The thorax and abdomen are segmented to 8 regions, some of which will give rise to structures such as the wings and the legs. The differentiation of these regions depends on complementary protein gradients, which are provided maternally as mRNA´s and translated soon after fertilization; Bicoid (Bcd) anteriorly, and the Hunchback (Hb) and Nos (Nos) proteins posteriorly, turn on the expression of zygotic genes. Of these, the first to be expressed are the gap genes. The boundaries of *giant* (*gt*), *hunchback* (*hb*), *kruppel* (*Kr*) and *knirps* (*kni*) determine where the future embryo sections (head, abdomen etc) will be, and mutations in them cause gaps in the normal body plan (Nüsslein-Volhard and Wieschaus, 1980). During nuclear cycles 11-14 gap genes mutually regulate each other and then, they turn on the expression of a second group, the pair-rule genes (Clyde et al., 2003; Nusslein-Volhard et al., 1987; Rivera-Pomar and Jäckle, 1996). These are expressed as intercalating stripes along the AP axis and their products define the parasegments of the embryo. Similarly to gap genes, the products of pair-rule genes control the expression of other pair-rule genes, as well as the next zygotic genes to be activated: segment polarity genes and homeobox genes. These two last classes of genes pattern the different segments and give rise to the epidermis and anatomical structures such as wings, legs and antennae of the developing individual (Doe et al., 1988; Nüsslein-Volhard and Wieschaus, 1980; Patel et al., 1989, Morata, 1993).

1.2 Terminal patterning during embryogenesis

Cephalic structures at the anterior extreme, and features such as the tail at the posterior extreme, are patterned by the terminal system (revised in Furriols and Casanova, 2003). The key element of this system is Torso, a Tyrosine Kinase receptor, which is present ubiquitously on the plasma membrane of the embryo, but turned on only at the poles, due to the presence of the Torso-like (tsl) protein, which is responsible for the cleavage/activation of the Trunk (Trk) ligand. Activated Torso initiates an intrecellular Ras/Raf/MAPK signalling cascade, which culminates in the phosphorylation of nuclear factors that regulate gene expression.

One of the factors active MAPK phosphorylates and inactivates is the Capicua (Cic) repressor. As a result, Cic targets *tailless* (*tll*) and *huckebein* (*hkb*) are relieved and expressed at the embryo termini. Expression of *tll* and *hkb* excludes the expression of gap genes such as *kni*, giving a terminal fate to these nuclei, and, conversely, exclusion of *tll* and *hkb* from the central regions by active Cic permits patterning of the trunk. (Jimenez et al., 2000). Furthermore, as we will see later on, the localized activation of the Torso RTK pathway at the embryo poles also regulates the expression of genes along the DV axis.

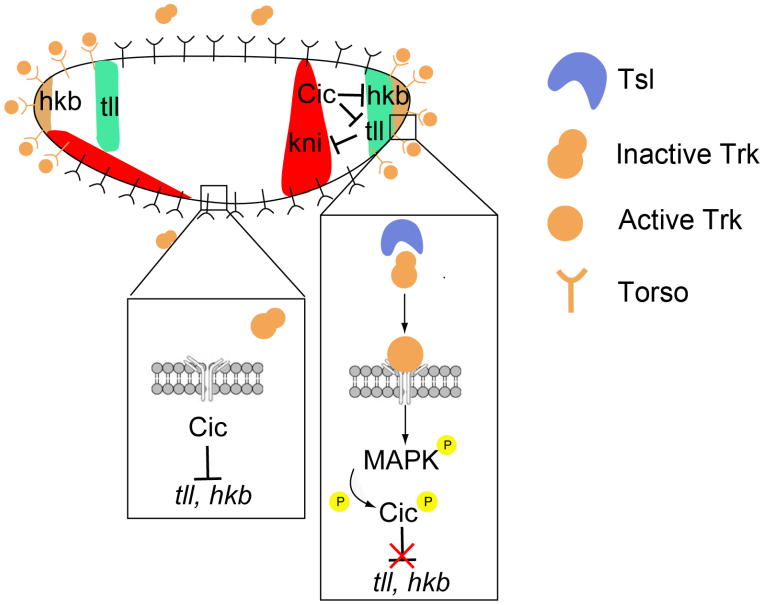


Figure 2. Terminal patterning during embryogenesis. The localized cleavage/ activation of the Trk ligand leads to the localized activation of the Torso receptor. Downstream Torso, the intracellular Ras/Raf/MAPK pathway phosphorylates/inactivates Cic, relieving the repression of the terminal genes *tll* and *hkb*.

1.3 DV patterning during embryogenesis

Considering the dorsal to ventral (DV) axis, at the end of gastrulation the embryo is divided into three regions, the germ layers: the mesoderm, the endoderm and the ectoderm. The mesoderm, at the ventral side of the embryo will give rise to tissues such as muscles, gonads, fat bodies and the heart. The endoderm contains precursors of the intestinal tract. The ectoderm consists of the neurogenic ectoderm, containing neural precursors and the dorsal ectoderm that

gives rise to structures such as the trachea and the larval epidermis. Finally, an extraembryonic membrane on the dorsal side, the amnioserosa, will contribute to dorsal closure, a process needed to form the larvae epidermis. The mesoderm and ectoderm are differentiated during early embryogenesis, while the endoderm emanates later, from the folding of terminal regions. The main factor responsible for the onset of different gene expression programmes within the germ layers is Dorsal (Dl), a morphogen transcription factor that is present in the nuclei of the blastoderm embryo in a ventral to dorsal gradient (Roth, 1989; Roth, 2003; Moussian and Roth, 2005, Stathopoulos and Levine, 2002).

As a definition, morphogens are molecules that can induce at least two different cell types at different concentrations (Ashe and Briscoe, 2006). As such, Dl establishes different tissues by inducing different target genes. At the ventral-most nuclei, high levels of Dl activate two main genes that define the mesoderm: *twi*, which encodes a basic helix-loop-helix (bHLH) factor (Pan 1991, Thisse 1991), and *snail* (*sna*), which encodes a zinc-finger transcriptional repressor, that demarcates the boundaries between the mesoderm and the overlying tissues (Gonzalez-Crespo and Levine, 1993; Kosman et al., 1991; Stathopoulos et al., 2002). Immediately above the sharp borders of *sna*, a transitional region is formed by the expression of *single-minded* (*sim*), a gene activated by Dl and repressed by Sna. This transitional region, the mesectoderm, is a single stripe of cells that separates the mesoderm from the ectoderm, forms the ventral midline of the animal and gives rise to neurons and glial cells.

Within the ectoderm, different regions are established by intermediate Dl levels: at its ventral-most region, the lateral neuroectoderm, *rhomboid* (*rho*) and *ventral neuroblast defective* (*vnd*) are expressed as a lateral stripe of 8-10 nuclei and specify the future peripheral nervous system. Above them, *intermediate neuroblast defective* (*ind*) specifies the intermediate neuroectoderm, that will give rise to the neuroblasts (Ip et al., 1992a; Jiménez et al., 1995; Weiss et al., 1998). The upper-most part of the ectoderm, the dorsal ectoderm, that will surround the neuroectoderm during gastrulation, is specified by the broad expression of genes such as *short gastrulation* (*sog*), which are activated by the lowest Dl levels.

Finally, the low levels of Dl that activate the expression of *sog* also repress a

last set of genes, *zerknüllt* (*zen*), *decapentaplegic* (*dpp*) and *tolloid* (*tld*) restricting them to the dorsal most part of the developing embryo (Doyle et al., 1989; Ferguson and Anderson, 1992; Kirov et al., 1994; Rushlow and Levine, 1990). These specify the amnioserosa, an extraembryonic tissue, which is critical for germ-band extension and dorsal closure (Lacy and Hutson, 2016, Rushlow and Levine 1990). Following this first round of transcriptional regulation, zygotic genes that have been regulated by Dl, such as *dpp* and *sog,* form a second patterning network, that continues the cell differentiation process along the DV axis (Ashe et al., 2000; O'Connor et al., 2006).

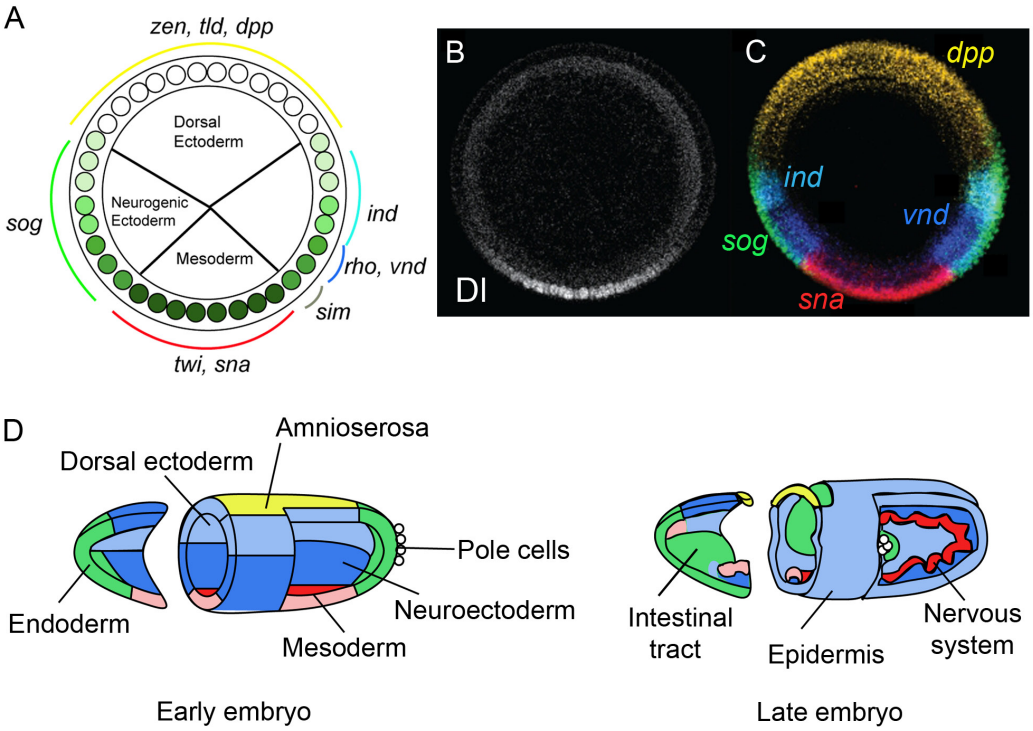


Figure 3. The dorsoventral network of embryogenesis. **(A):** Scheme of cross-section of an early embryo. Green colour indicated the presence of nuclear Dl. **(B)** : cross section of immunostaining of early stage embryos with an anti-Dl antibody **(C):** cross section of in situ hybridization of early stage embryos, showing the expression of Dl targets along the DV axis (Adapted from Reeves and Stathopoulos, 2009). **(D)**: Schematic representation of the formation of ventral and dorsal tissues during embryogenesis.

2. Maternal control of development

Shaping of the *Drosophila* body begins even before the egg is fertilized and is based on information provided by the mother. Maternal control dominates development for a prolonged period, during several cell cleavage cycles after fertilization, and in some cases it persists even after the zygotic genome has been activated.

Four systems of maternal control exist in *Drosophila* (figure 4): the anterior and posterior systems instruct the formation of the perspective head/thorax and abdomen/germline respectively. The dorsoventral system determines the formation of structures along the DV axis. Finally, the terminal system controls both the anterior (acron) and posterior (telson) extremes of the embryo. Although each of these systems regulates a well-defined region of the future animal, crosstalk between them exists. For example, the anterior extremity is determined both by the anterior and terminal systems, while the DV axis is patterned under the control of both dorsoventral and terminal maternal systems. In all of these systems, information held by the mother's germline or somatic cells is provided to its offspring, either as proteins that act during oogenesis, either as mRNA that is translated once fertilization occurs. Another mechanism of maternal control is referred to as ´´late induction´´(LeMosy et al., 2003; Roth, 1998). In this, signalling events that occur in the egg act as positional cues that are interpreted by the embryo, which translates them into distinct cell fates after fertilization.

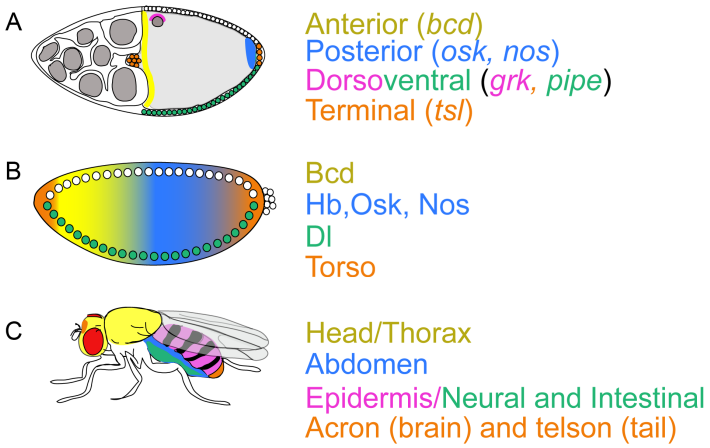


Figure 4. Maternal control of *Drosophila* development. Four maternal systems act in the ovary **(A)** and transmit the necessary information to the embryo **(B)** in order to pattern the different body regions of the future adult **(C)**.

3. Axis specification during oogenesis

One of the functions of the *Drosophila* maternal systems is to provide the bases for the establishment of the embryo body axes. Therefore, to fully understand the molecular events that polarize the embryo, we will take a step back and provide a brief description of oogenesis and the symmetry breaking events occurring at this stage.

Brief overview of oogenesis

The development of the new *Drosophila* adult begins with the formation of the oocyte during oogenesis, a process that takes roughly one week, beginning with an undifferentiated cell that becomes the mature, ready for fertilization egg (reviewed in Bastock and Johnson, 2008). Oogenesis takes place in the female ovary, in a structure called ovariole, where eggs are produced and mature in an assembly line fashion. At the anterior end of the ovariole, the germarium, the stem cells exit and develop to oocytes as they move through the ovariole, so that each ovariole has 6-7 eggs at different stages of development. Initially, the germarium stem cells divide, giving rise to another stem cell and a daughter cell. The daughter cell undergoes 4 mitotic divisions with incomplete cytokinesis, resulting in the formation of a cluster of 16 cells that are interconnected by cytoplasmic bridges, known as ring canals. From these 16 cells, one will differentiate to become the oocyte, which is characterized by its unique ability to undergo meiotic division. The rest 15 cells of the syncytium adopt a polyploid fate and become the so-called nurse cells, that have the role of delivering maternal RNA´s, proteins, nutrients and organells through the ring canals to support the immature oocyte. The cluster of nurse cells and oocyte are surrounded by a somatic follicle cell epithelium and altogether this structure is called the egg chamber, and is the basic unit of *Drosophila* oogenesis. At stage 10 of oogenesis, the nurse cells are eliminated by apoptosis after they are contracted and dump their contents to the oocyte, as the nurse cell cytoplasms are fused to the oocyte cytoplasm (ooplasm). Then, the follicle cells migrate to enclose the oocyte and secrete the vitelline membrane and the eggshell, known as chorion, that both protect the mature egg.

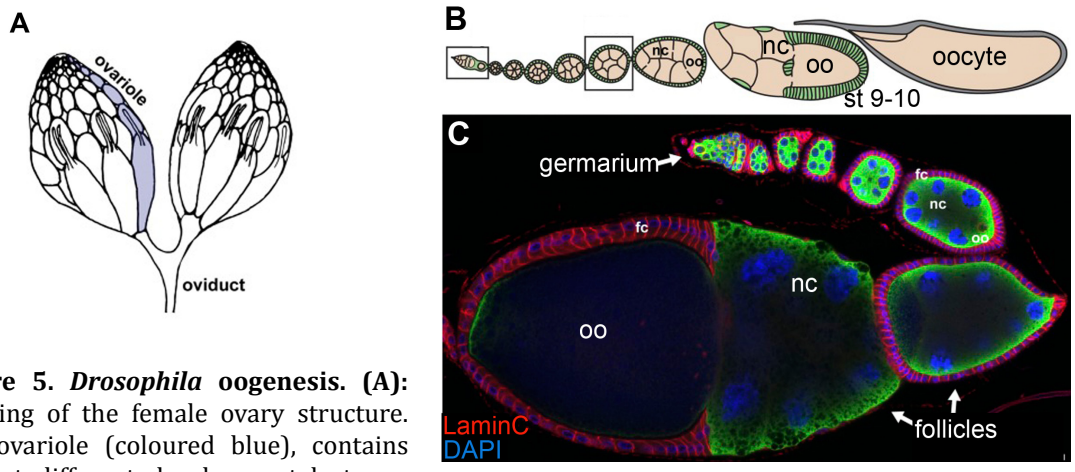


Figure 5. *Drosophila* oogenesis. (A): Drawing of the female ovary structure. The ovariole (coloured blue), contains eggs at different developmental stages. **(B):** Drawing of the different stages of the oocyte, starting from the stem cell at the germarium (left) until the mature oocyte (right). **(C):** Immunocytochemistry of egg chambers at different stages of development. Nuclei are stained with DAPI (Blue) and cell membranes with LaminC (Red). nc; nurse cells, oo; oocyte, fc; follicle cells. Adapted from Ables, 2015.

If we were to divide the egg in orthogonal sections, we would see that, early, since it leaves the germarium, it is not symmetrical, in the sense that distinct subpopulations of follicle cells exist. The asymmetry (polarization) of the egg chamber is a consequence of cell signalling between the germline oocyte and the somatic follicle cells, and is required for the eggshell formation and determination of the embryo body axes. Egg chamber polarization requires the correct localization of the *gurken* (*grk*), *bicoid* (*bcd*), *oskar* (*osk*) and *nanos* (*nos*) mRNA's, which are synthesized in the nurse cells and transported to the oocyte cortex (González-Reyes and St. Johnson, 1998; Beclaska and Clavis, 2009). Then, depending on signalling events, they are deposited at different positions in the egg chamber, where they direct the overlying follicle cells to adopt distinct fates. Two axes of asymmetry exist in the egg chamber and correspond to the two body axes that the individual will have as an adult: the anterior to posterior (AP) and the dorsal to ventral (DV). The AP polarity is the first to be determined during early oogenesis with the anterior localization of *bcd* and posterior localization of *nos* and *osk* mRNA´s. Then, during mid-oogenesis, the egg begins to present DV asymmetry and the terminal maternal system is also turned on. Since DV polarization of the embryo depends on both dorsoventral and terminal maternal control, we will focus on these two processes.

3.1.1 DV polarity establishment during oogenesis

During mid-oogenesis (stage 8), the oocyte, which is localized at the dorsal-anterior region of the egg chamber, secretes the Gurken (Grk) protein, which acts a ligand for Torpedo/DER, a Tyrosine Kinase receptor, in the adjacent overlying follicle cells (Queenan et al., 1999; Peri et al., 1999; Stein and Stevens, 2014). Although Grk is secreted at the dorsal-most region, where the oocyte is localized, it acts as a long-range morphogen and activates Torpedo in a graded manner (Stein and Stevens, 2014). This graded activation of Torpedo initiates two branches of the EGF (Epidermal Growth Factor) signalling cascade, which has a double mission: the formation of the dorsal structures of the eggshell (for example the dorsal appendages) and the generation of the distinct subpopulations of dorsal and ventral follicle cells. This latter asymmetry between the follicle cells is the basis for the dorsal to ventral polarization of the future embryo.

DV patterning of the egg chamber

Upon the activation of Torpedo, the canonical Ras/Raf/MAPK pathway is initiated in the dorsal-most region of the follicular epithelium (Peri et al., 1999). Two main downstream factors respond to this EGF signal and define the expression boundaries of Pipe (Pip), the key factor that transmits the DV asymmetry information from the egg chamber to the embryo. On the one hand, Mirror (Mirr), a homeodomain factor, is activated by the EGF signal, forming a gradient with peak levels at the dorsal anterior follicle cells and lowest levels at the lateral follicle cells (Jordan et al., 2000). In this dorsal-lateral domain, Mirr acts as a cell-autonomous repressor of *pipe*, by directly binding to a specific region in its proximal promoter (Andreu et al., 2012a; Fuchs et al., 2012). On the other hand, the HMG-box (High Mobility Group-box) Cic, downregulated by the EGF signal, forms a ventral to lateral gradient, complementary to the EGF activation domain. At the ventral most follicle cells, Cic acts as a repressor of Mirr, relieving the suppression of *pipe* (Andreu et al., 2012b; Goff et al., 2001; Technau et al., 2012). A scheme of the DV polarity network in the egg chamber can be seen in figure 6.

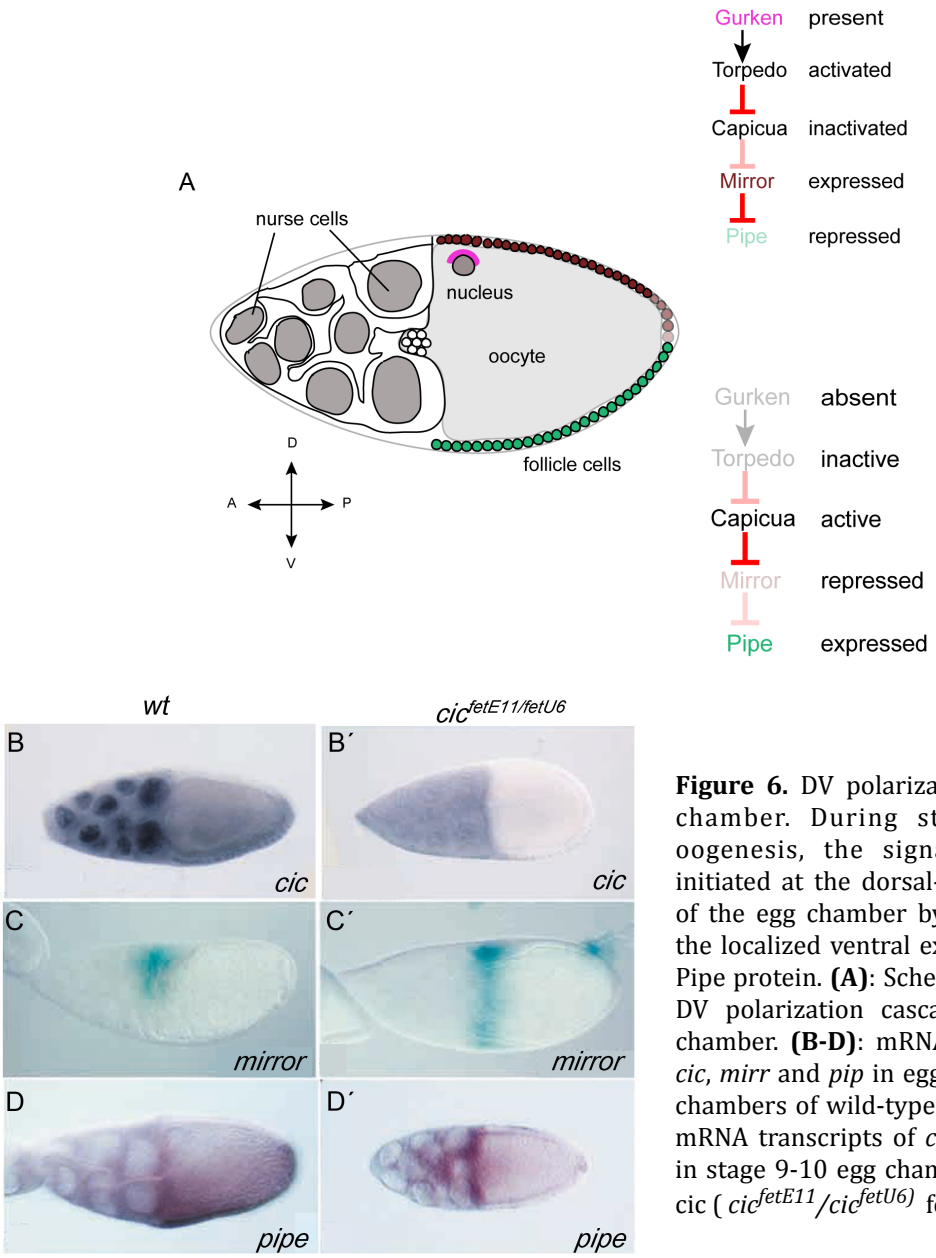


Figure 6. DV polarization of the egg chamber. During stages 9-10 of oogenesis, the signalling cascade initiated at the dorsal-anterior region of the egg chamber by Grk results in the localized ventral expression of the Pipe protein. **(A)**: Scheme showing the DV polarization cascade of the egg chamber. **(B-D)**: mRNA transcripts of *cic*, *mirr* and *pip* in egg stage 9-10 egg chambers of wild-type females.**(B´-D´)** mRNA transcripts of *cic, mirr* and *pip* in stage 9-10 egg chambers of mutant cic ($cic^{fetE11}/cic^{fetU6)}$ females.

As a combinatorial output of the above regulatory network, during stages 9-10 of oogenesis, *pipe* is expressed in a broad ventral domain at the follicular epithelium of the egg chamber. The product of *pipe* is a sulfonyltransferase enzyme that is localized in the Golgi apparatus and modifies certain components of the vitelline membrane (Zhang et al., 2009; Zhu et al., 2007). These modified components are then secreted to the perivitelline space that surrounds the egg and initiate a serine protease cascade, which transmits the ventralizing cue to the

future embryo (Stein and Stevens, 2014; Zhang et al., 2009). Briefly, the output of this cascade is the processing of the inactive Spätzle (Spz) ligand to its active form. Upon fertilization, active Spz binds to the Toll receptor, which is located on the embryo plasma membrane and initiates an intracellular signalling pathway. Spz is cleaved in a ventral to dorsal graded fashion, following the domain of Pipe expression and activates Toll in the same graded manner (Roth, 1994). In turn, the Toll signalling pathway culminates with the phosphorylation and degradation of an IκB homologue protein, Cactus (Cact) (Belvin et al., 1995; Roth et al., 1991), which retains the Dl in the embryo cytoplasm. As a result of the graded Toll activation and Cact degradation, Dl enters the embryo nuclei in a ventral to dorsal gradient (Anderson et al., 1985; Rushlow et al., 1989; Steward et al., 1988). It is this Dl gradient that initiates the process of DV patterning in the early embryo.

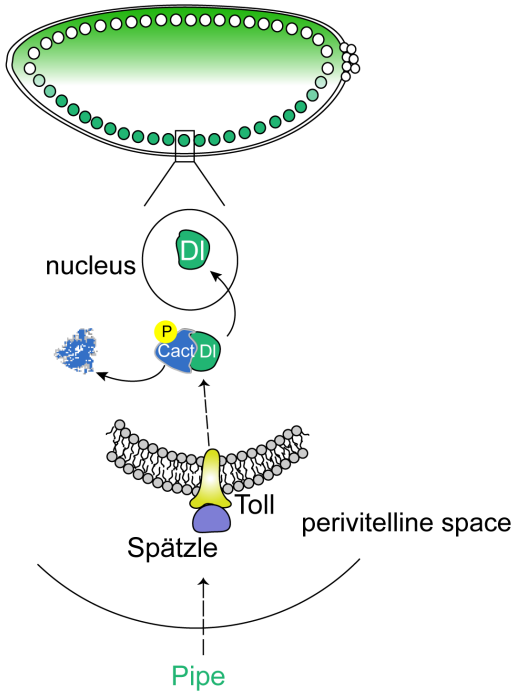


Figure 7. Pipe transmits the ventralizing signal during the egg to embryo transition. Pipe induces a serine protease cascade in the perivitelline space which ultimately result in activation and binding of Spz to the Toll receptor. Toll signalling culminates with phosphorylation and degradation of Cact and release of Dl from the cytoplasm.

3.2.2 Terminal patterning during oogenesis

During mid-oogenesis, the terminal maternal system is also turned on. A complex of germ-line expressed proteins mediate the accumulation of the Torso-like (Tsl) protein at the anterior and posterior follicle cells (Furriols et al., 2007; Mineo et al., 2015; Roch et al., 2002, Stevens et al., 2003). Through a not quite understood mechanism, which involves cleavage of the Trunk (Trk) ligand to its active form, a positional cue is generated at the extremities of the egg chamber (Casali and Casanova, 2001; Henstridge et al., 2014; Johnson et al., 2015). This cue is transferred to the fertilized embryo and results in the localized activation of the Torso RTK receptor, which is maternally provided as a ubiquitous mRNA and translated along the whole circumfence of the plasma membrane of the embryo after egg activation. Likewise the localized activation of Toll by Spz determines the ventral fate, selective activation of the Torso pathway by Trk at the embryo poles confers a terminal fate (Casanova, 1995; Casanova and Struhl, 1989; Perrimon et al., 1995; Schupback and Wierchaus, 1986). Upon fertilization, the anterior structures such as the head skeleton and labium are patterned based on information from the anterior and terminal system. At the other end of the embryo, structures such as the abdominal segment 8, the posterior spiracles and the and the Filzkörper filaments are patterned based on the information provided by the posterior and terminal systems.

The terminal system also participates in DV patterning of the embryo. The same protein complex that is responsible for the restricted translocalization of Tsl, is also involved in the dorsoventral maternal system. Specifically, it ensures the proper activity of the Nudel (Ndl) protein, which forms part of the serine protease cascade downstream Pipe, that initiates the Dl localization signal (Mineo et al., 2017). Furthermore, after fertilization, activation of the Torso pathway interferes with the expression of genes regulated by Dl.

4. Dorsal, the main nuclear effector in DV polarity

As we have seen above, the key factor in embryonic DV patterning is Dl, a transcription factor that belongs to the Rel family, which includes the mammalian NF-κB (nuclear factor kappa-light chain enhancer of B activated cells) proteins. Specifically, it is homologous to the vertebrate proto-oncogene *c-rel* and its corresponding viral oncogene *v-rel* (Steward et al., 1989), which are involved in various cellular processes, including inflammatory response and may cause certain cancer types, such as lymphomas, when they are mutated or amplified. Dl is maternally provided to the *Drosophila* embryo as a uniformly expressed mRNA that is translated approximately 90 minutes after fertilization. Soon after, activation of the Toll pathway triggers its selective nuclear import, by degrading Cact. As the Toll receptor is activated in a graded fashion, Dl enters the nuclei in a gradient as well, with peak concentrations at the ventral most nuclei of the embryo. Recent studies have shown that the nuclear-cytoplasmic localization of Dl is not stable, but results from a highly dynamic equilibrium that involves import, increase and drop of nuclear levels of Dl within each mitotic cleavage (DeLotto et al., 2007; Kanodia et al., 2009). As we move upwards along the presumptive DV axis, the rate of nuclear import decreases and levels gradually drop, with the dorsal most region having essentially cytoplasmic Dl.

4.1 Structure and functions of Dl

As other members of the Rel family, Dl has an N-terminal region of 300 aminoacids, the so-called Rel homology domain (RHD), which is responsible for DNA binding, dimerization and binding to inhibitory regulators, such as Cact (Govind et al., 1996; Isoda et al., 1992; Tatei and Levine, 1995; Tony Ip et al., 1991). On the other hand, its C-terminal domain (CTD), which contains characteristic polyQ stretches, is responsible for the transcriptional activities of Dl, but also for selective nuclear import (Govind et al., 1996; Isoda et al., 1992; Jia et al., 2002; Rushlow et al., 1989). Dl is known to interact with various partner proteins, both activators and repressors, to accomplish transcriptional regulation. These interactions are mediated by both the RHD and carboxyterminal regions (Flores-Saaib and Courey, 2000; Shirokawa and Courey, 1997). Finally, it has been

suggested that inter-protein interactions between the RHD and the CTD also exist and are essential for regulating the targets (Jia et al., 2002). The structural domains of Dl are illustrated in figure 8.

Considering its functions, Dl is the key regulator of a number of zygotic genes which, in turn, determine distinct cell fates along the DV axis of the developing embryo. However, these are only some of the genes regulated by Dl. Recent genome-wide studies have identified approximately 100 targets, including proteins, such as N-Cadherin, and micro-RNAs (Biemar et al., 2006; Markstein et al., 2002; Stathopoulos et al., 2002). Furthermore, whole-genome ChIP-chip analysis has shown that Dl also fine-tunes the expression of the AP axis, such as *kni*, and regulates genes encoding signal transduction components (Zeitlinger et al., 2007).

Similarly to its vertebrate analogues, Dl also has a role in immune response. Upon bacterial challenge in the larvae, Dl, that is localized in the cytoplasm of the fat bodies, responds to Toll signalling and translocates to the nuclei (Lemaitre et al., 1996; Reichhart et al., 1993). Dl homodimers induce the expression of the antifungal peptide *drosomycin*, while Dorsal-Relish heterodimers activate the expression of the *defensin* peptide which protects against bacteria, fungi and some viruses (Ip et al., 1993).

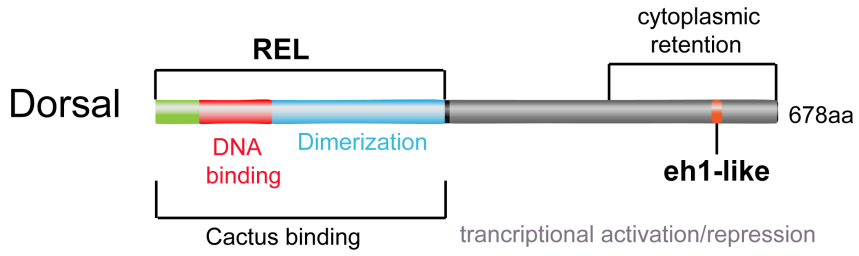


Figure 8. The functional domains of the Dl protein. The first 380 residues are the conserved Rel domain, in which DNA binding, dimerization and binding to the Cact inhibitor motifs are localized. The C-terminal domain is responsible for transcriptional activities and cytoplasmic retention. The eh1-motif has been proposed to interact with the Gro corepressor.

4.2 Transcriptional regulation by Dl

As already mentioned, Dl is a DNA-binding transcription factor, which regulates a large set of target genes in different regions of the developing embryo.

To achieve the multitude of responses required for the formation of the DV axis, Dl acts both in a concentration and context-dependent manner. On the one hand, Dl, due to its graded presence in the nuclei, is considered a morphogen molecule and induces different genes at different concentrations. On the other hand, independently of changes in its concentration, Dl has a dual function: depending on the target gene, it either functions as an activator or as a repressor.

4.3 Dl-mediated activation

Intrinsically, Dl functions as a transcriptional activator and this occurs along the largest portion of the DV axis. Different concentrations of Dl trigger the activation of different genes which, in turn, specify distinct regions, the mesoderm and ectoderm. However, cell fates along the DV axis do not always coincide with the steep changes in Dl concentration, nor can transcriptional outputs can always be explained by differential strength of its binding to regulatory sequences (Liberman et al., 2009; Stathopoulos et al., 2002). Instead, it is well accepted that Dl is only one element of a complex network that integrates broadly distributed activators, localized repressors and modulating inputs of numerous signalling pathways (Liberman et al., 2009; Stathopoulos and Levine, 2005; von Ohlen and Doe, 2000). To turn on its targets, Dl synergizes with activator proteins belonging mainly to the bHLH (basic Helix-Loop-Helix) family, namely Daughterless (Da), T4 and, most importantly Twi, which Dl itself turns on (Gonzalez-Crespo and Levine, 1993; Ip et al., 1992a; Kosman et al., 1991).

In regions such as the mesoderm, where Dl is abundant in the nuclei, its targets, nominated type I, are turned on by the sum of the independent activities of Dl and its cofactors, while no cooperative DNA binding is observed. The low levels of activation that Dl triggers are potentiated by other activators that contribute to the additive recruitment of the basal transcriptional machinery (Hong et al., 2008; Shirokawa and Courey, 1997). This is reflected in the *cis*-regulatory modules (CRM's) of type I genes, which contain low-affinity sites for Dl and somehow unlinked bHLH binding sites, whose composition, arrangement and orientation is flexible. Actually, even deleting the bHLH sites still permits low expression levels of the mesoderm targets (Jiang and Levine, 1993a; Szymanski and Levine, 1995). In the neurogenic ectoderm, where the slope of the Dl gradient is steep and its levels

are intermediate, it activates targets nominated type II. To turn on type II genes, Dl synergizes with proteins of the bHLH family, and other activators as well. In this tissue, the organization of the target CRM's is slightly different; although they still maintain a grade of flexibility in their organization, the need for proximity and correct orientation of the binding sites are indicative of the cooperative interactions between Dl and cofactors. For example, the CRM of *rho* contains various arrays of high and low affinity Dl sites closely linked to bHLH sites, ensuring that the intermediate levels of Dl occupy the promoter sufficiently to activate it in a broad domain (Szymanski and Levine, 1995).

A last group of genes activated are type III genes, which are turned on by low levels of Dl at the dorsal ectoderm. Activation of type III genes, such as *sog*, is achieved through high-affinity Dl sites, which are closely linked to sites for Zelda, an ubiquitously expressed activator which synergizes cooperatively with Dl.

4.4 Dl-mediated repression

Remarkably, the same low levels of Dl that activate *sog*, also repress a different set of genes within the type III group, which include *zen*, *tld* and *dpp.* As a result, these are restricted to the dorsal-most nuclei of the blastoderm embryo, where they will specify the amnioserosa (Ray 1991). At the absence of Dl or when Dl binding sites are mutated, expansion of dorsal-specific genes towards ventral regions of the embryo is observed (Jiang et al., 199b, Kirov et al., 1993; Kirov et al., 1994). How Dl switches from its intrinsic, activator function to a repressor mode in order to restrict genes from the ectoderm has been a long-standing question in the field of *Drosophila* developmental genetics. Extensive studies on the regulation of dorsal-specific type III genes, have shown that their repression depends on short elements found in their enhancers, nominated VRE (Ventral Repression Elements) or VRR (Ventral Repression Regions). These modules of 600-800 bp are sufficient to reproduce the endogenous pattern of the genes and contain all the sequences necessary to drive ventral repression (Doyle et al., 1989; Jiang et al., 1993a; Kirov et al., 1993; Kirov et al., 1994; Valentine et al., 1998). VRE´s include high affinity binding sites for Dl and binding sites for general activators, such as Zelda, which, in contrast to the *sog* CRM, are somehow separeted from the Dl sites (Kirov et al., 1993). Instead, in the VRE´s, the Dl sites are closely neighboured by characteristic

sequences rich in A and T bases (hereafter reffered to as AT-rich sites), that have been suggested to serve as binding sites for one or more repressors that form a multiprotein complex with Dl and contribute to the switch of its activity. Purification of embryonic and cell culture extracts have revealed a number of AT-rich associated repressors, including the Cut/Dead-ringer (also named Retained) complex, NTF-1 (Nuclear Transcription Factor-1) and the Dorsal Switch Protein (DSP1) (Brickman et al., 1999; Valentine et al., 1998; Huang et al., 1995; Ip, 1995). However, phenotypes caused by mutant alleles of these repressors are not very strong, especially during stages prior to cellularization. For example, Dri effects are prominent after stage 5, while Cut has been demonstrated to have a strictly zygotic function (Valentine et al, 1998). Therefore, it has been accepted that additional repressors function through the AT-rich sites. One such repressor has been proposed to be Cic, whose hypomorphic mutations result in ventral expansion of the *zen* transcripts in early embryos (Jimenez et al., 2000). However, the posterior characterization of its function in the DV polarization of the egg chamber, which affects the formation of the Dl gradient itself, has obscured the implication of Cic in Dl-mediated repression.

Another essential element during Dl-mediated repression is the non-DNA binding Groucho (Gro) corepressor, a protein that is ubiquitously expressed in the early embryo and is involved in various developmental processes, such as sex determination, embryo segmentation, neurogenesis and bristle formation (Jennings and Ish-Horowicz, 2008; Paroush et al., 1994, Turki-Judeh and Courey, 2012). Gro belongs to the Enhancer-of-Split complex (E(spl)-C) and consists of five identifiable regions of which the Q rich N-terminal region and a C-terminal WD-repeat domain are conserved (Schrons 1992, Jennings et al., 2006, Buscarlet and Stifani, 2007). A scheme of the structural domains of Gro can be seen in figure 9. Gro is recruited by short peptide motifs of DNA-bound repressors, that fall into two main classes: the WRPW motif and its variations, and the Engrailed motif (eh1) which has the consensus sequence FSISNILS. Although these two motifs adopt different conformations, they both contact overlapping sites in the pore of the β-propeller formed by the Gro WD domain (Buscarlet and Stifani, 2007). Once Gro is recruited by repressors bound to the promoters of target genes, it interferes with

transcription by inhibiting the RNA Poll complex recruitment and chromatin remodeling. Null mutants of Gro cause ventral derepression of *zen* and, interestingly, a motif resembling the eh1 peptide, with the sequence PTLSNLLS (hereafter eh1-like), has been identified in Dl and found to interact with Gro *in vitro* (Dubnicoff et al., 1997; Flores-Saaib et al., 2001). However, this motif lacks the first phenylalaline residue, which is critical for Gro recruitment, making the interaction weak (Jimenez et al., 1999, Flores-Saaib et al., 2001). It has been proposed that interactions between AT-rich bound repressors and Dl might expose a cryptic Gro-recruiting motif (Flores-Saaib et al., 2001; Hong et al., 2008). Other theories postulate that Dl and the AT-repressors function as a high affinity platform that collectively recruits Gro (Dubnicoff et al., 1997; Ratnaparkhi et al., 2006).

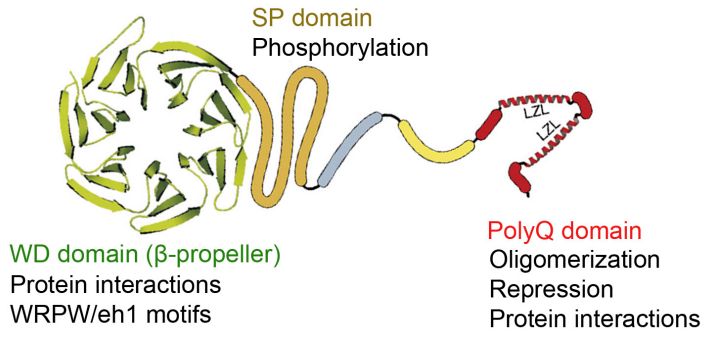


Figure 9. The structural domains of the Gro corepressor. The WD domain interacts with recruiting proteins such as Hairy and Engrailed. The SP domain is phosphorylated, resulting in Gro inactivation. The polyQ stretch is responsible for oligomerization, repression and protein-protein interaction.
Adapted from Buscarlet and Stifani, 2007.

A final aspect of Dl-mediated repression is that it receives inputs from the terminal system. Activation of the Torso RTK pathway at the termini of the embryo modulates the repressor activities of Dl and, consequently, prevents the adoption of a ventral fate at the terminal regions (Casanova, 1991; Rusch and Levine, 1994). Due to the activation of Torso, the expression domains of dorsal-specific type III genes (*zen*, *tld* and *dpp*) at the terminal parts of the embryo extend to ventral nuclei where they would otherwise be repressed by Dl. Mutations that impair Torso activity result in repression of *zen* at the embryo poles, while in gain-of-function Torso mutants *zen* expands towards the central regions, overcoming the presence of Dl. This expansion is similar to derepression of *zen* in *dl* mutants (Figure 10). The mechanism through which the Torso pathway exerts this anti-repression effect is not fully understood. It has been proposed that a mediator of Torso is WntD, a

Drosophila Wnt homologue, which acts as a feedback inhibitor of Dl (Ganguly et al., 2005; Gordon et al., 2005). During stage 4 of embryogenesis, *wntD* is turned on in a ventral domain by Dl, but in the embryo trunk it is inhibited by Cic. Torso-mediated downregulation of Cic permits the expression of WntD, which reduces Dl nuclear levels at the poles (Helman et al., 2012). A different possibile mechanism proposed is that the activation of the Torso pathway disrupts interactions between Dl and its corepressors, or modulates one or more of the corepressors (Rusch and Levine, 1994). The corepressor downregulated by Torso could be the Cut/Dri complex, Gro, or Cic, in case it is indeed a member of the Dl repressosome.

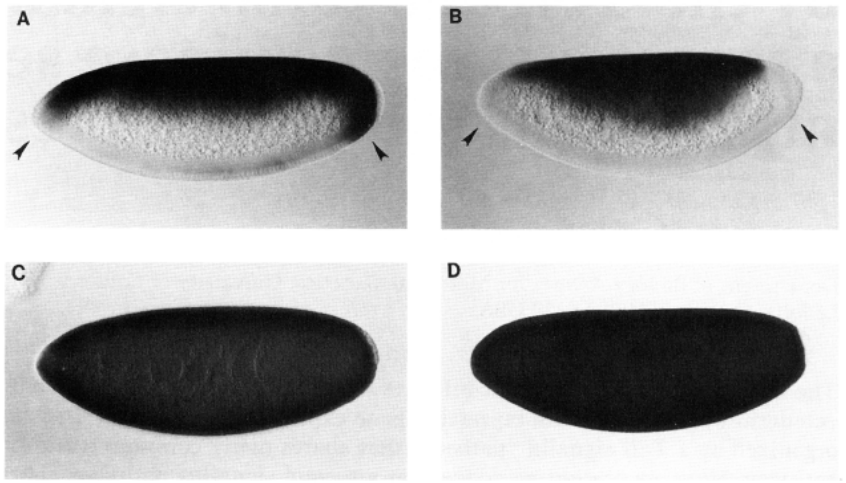


Figure 10. The Torso pathway has anti-repression effects on *zen* at the embryo poles. **(A):** mRNA transcripts of *zen* in wild-type embryos. Expression at the termini extends ventrally, despite the presence of Dl in the nuclei. **(B):** mRNA transcripts of *zen in tor* mutant embryos (tor^{PM51}). Expression is excluded from the poles in the absence of Tor activation. **(C, D):** mRNA transcripts of *zen* in *dl* mutant and *tor* dominant mutants respectively. Lack of Dl and constitutive activation of Tor both result in derepression of *zen*. (Adapted from Rusch and Levine, 1994).

Despite the identification of the AT-rich sites and Gro as essential elements of Dl-mediated repression, the molecular mechanism of how it switches from an activator to a repressor is still not fully understood. Indeed, the fact that the Dl and AT-richsites in VRE's are tightly organized and are not permissive on changes regarding their composition or phasing, suggests that a multiprotein complex (hereafter referred to as Dl represssome or Dl repressor complex) is formed. Following the logic of the regulatory code along the mesoderm and ectoderm, this arrangement also suggests that DNA cooperative interactions between Dl and the

bound cofactors accounts for at least one of the regulatory mechanisms (Valentine et al., 1998, Hong et al., 2008; Jiang et al., 1992). However, full characterization of the Dl repressosome and how it functions remains an open question.

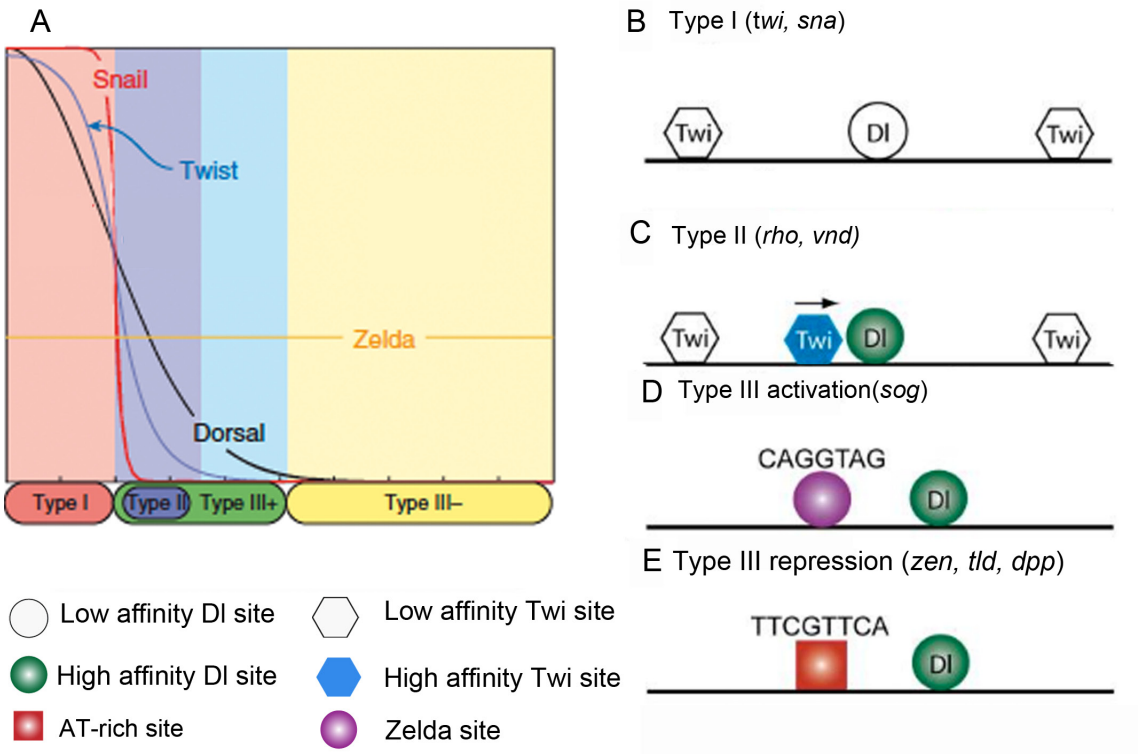


Figure 11. Activation and repression along the DV axis. **(A):** Nuclear levels of Dl, Twi, Sna and Zelda determine gene regulation across the DV axis. **(B):** high levels of Dl activate mesoderm targets such as *twi* and *sna*. Low affinity Dl and Twi sites are separated and act independently. **(C):** Intermediate Dl levels activate type II targets at the neuroectoderm (*rho, vnd*) through cooperative interactions with bHLH factors such as Twi. **(D):** Low Dl levels bind to high affinity sites and interact with Zelda to activate *sog* at the dorsal ectoderm. **(E):** The same low Dl levels that activate *sog* repress type III dorsal specific genes (*zen, tld, dpp*), restricting them to the presumptive amnioserosa. Repression requires high affinity Dl sites and closely linked AT-rich sites. Adapted from Reeves and Stathopoulos, 2009 and Hong et al., 2008).

5. Capicua, a Dl corepressor candidate

One of the factors suggested to contribute to the switch of Dl to a transcriptional repressor is Cic, which, as we have mentioned in previous sections, is involved in the egg chamber polarization and the embryonic terminal system. In the next sections, we will provide a more detailed description of its structural domains and known functions.

5.1 Structure and functional domains of Cic

Cic is a nuclear factor that acts downstream the RTK Torso signalling pathway in *Drosophila* (Jiménez et al., 2000). It is an HMG-box (High Mobility Group) protein expressed in various isoforms that fall into two main classes, the long (Cic-L) and the short (Cic-S), encoded by the *bullwinkle* (*bwk)* and *cic* locus respectively (Goff et al., 2001; Jiménez et al., 2000; Roch et al., 2002b). Of these, the best characterized are the Cic-S isoforms, which have various functions in different developmental stages. Considering their structure, all Cic isoforms share the HMG-box region, and the carboxyterminal C1 and C2 domains. The HMG-box accounts for the DNA-binding domain, characteristic of the family, while C1 domain is involved in repression of Cic target genes and has been recently shown to play an active role in DNA-binding (Forés et al., 2017; Jiménez et al., 2000). The C2 domain is a response element to the RTK pathways, serving as a docking site for active MAPK (*rolled* in *Drosophila*)(Astigarraga et al., 2007). Apart from these shared regions, at least some of the Cic-S isoforms have an aminoterminal motif nominated N2, which, is associates to the Gro corepressor in the context of embryogenesis (Forés et al., 2015). The Cic-L proteins have a specific aminoterminal N1 domain with an unknown function.

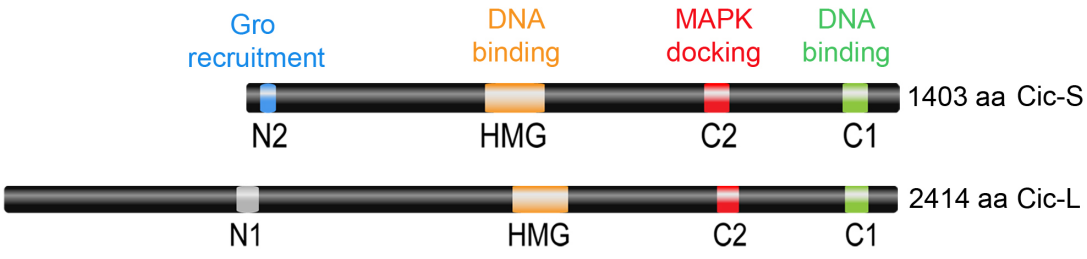


Figure 12. Structure of *Drosophila* Cic proteins. Two main isoforms, Cic-S and Cic-L are expressed and have shared and isoform-specific domains.

5.2 Human homologues of Cic and their functions

In mammals, a homologue of Cic that shows high conservation of the HMG and C1 domains exists. Although the molecular mechanism of its function is not fully understood, it is involved in physiological processes such as lung alveolarization and bile acid homeostasis in the liver (Kim et al., 2015; Lee et al., 2011). Moreover, Cic is involved in a series of pathological conditions. In various malignancies, such as breast cancer, oligodendroglioma (OD) brain tumours and Ewing-like sarcomas, mutations of Cic have been observed (Bettegowda et al., 2011; Kawamura-Saito et al., 2006). These include missense mutations that inactivate the HMG-box and the C1 domain, as well as a chromosomal translocation that results in the fusion with the DUX4 protein. This fusion protein activates a set of oncogenic driver genes, which suggests that, normally, Cic has a role as a tumour suppressor (Kawamura-Saito et al., 2006; Okimoto et al., 2016). Finally, human Cic associates with the Ataxin (ATX) cofactor and mutations that interfere with this interaction have been associated with neurotoxicity observed in spinocerebellar ataxia type 1 (Lam et al., 2006; Lim et al., 2008). All together, these examples of the biomedical functions of Cic, underline the importance of further elucidating its functions.

5.3 Mechanism of Cic function in *Drosophila*

The molecular mechanism through which Cic functions in *Drosophila* has been well characterized for the Cic-S proteins. These isoforms recognize the canonical, octameric **T(G/C)AATG(A/G)A** site in the CRM´s of target genes and contact it through the HMG-box and C1 motifs. Once it is bound to DNA, Cic-S recruits Gro to repress its targets. Although it does not contain a peptide is similar to either WRPW or eh1 motifs, genetic studies have shown that it uses the N2 motif and possibly contacts the Gro β-propeller, adopting a different conformation than other Gro-associated proteins (Ajuria et al., 2011; Forés et al., 2015; Forés et al., 2017a). Notably, Gro-dependent Cic repression only occurs during embryonic patterning, it is therefore conceivable that in other tissues it functions with other, still unidentified corepressors to restrict its targets.

5.4 Regulation and functions of Cic during *Drosophila* development

The first function of Cic that was characterized in *Drosophila* is patterning of the embryo AP axis. At the blastoderm stage, a 1400 aminoacid Cic-S protein that includes the N2 motif is the predominant isoform, which is maternally provided as an ubiquitously expressed mRNA (Jiménez et al., 2000). Immunostaining of embryos with an anti-Cic antibody shows that the protein is distributed at the nuclei of the presumptive trunk region, but gradually drops at the anterior and posterior region (see figure 13). This drop of Cic levels is complementary to the graded activation of the Torso receptor. The intracellular signal transduction of activated Torso culminates in the phosphorylation/activation of MAPK, which binds to the C2 motif of Cic and phosphorylates it at yet unknown residues. Indeed, mutant forms of Cic lacking the C2 motif are expressed in the terminal nuclei of the embryo (Astigarraga et al., 2007, figure 13). The events downstream Cic phosphorylation are still not well understood. It has been suggested that upon activation of the pathway, nuclear-cytoplasmic shuttling rates change, and Cic, spending more time in the cytoplasm, is eventually degraded (Grimm et al., 2012). However, posterior studies showed that ectopic activation of Torso did not result in ectopic degradation of Cic, which remains nuclear at the central regions of the embryo (de las Heras and Casanova, 2006). The final result is that Cic repressor activities over the terminal genes *tll* and *hkb* are relieved at the embryo poles, contributing to the formation of terminal structures. Actually, the name Capicua (from *cap i cua*, which means head and tail in Catalan) originated from the observation that embryos proceeding from mutant *cic* mothers, lacked any segments, and only consisted of anterior (head) and posterior (tail) rests, due to the expansion of terminal genes at the expense of central gap genes.

Cic is also involved DV patterning of the embryo. As we have seen above, it is a key element downstream EGFR signalling during the establishment of the ovary DV polarity, and, thus, the generation of the positional cue that will determine the dorsal and ventral cell fates in the embryo. Cic downregulation by the EGFR pathway also involves C2 motif docking and Cic phosphorylation. However, in contrast to the embryo, phosphorylation results in translocation of Cic to the

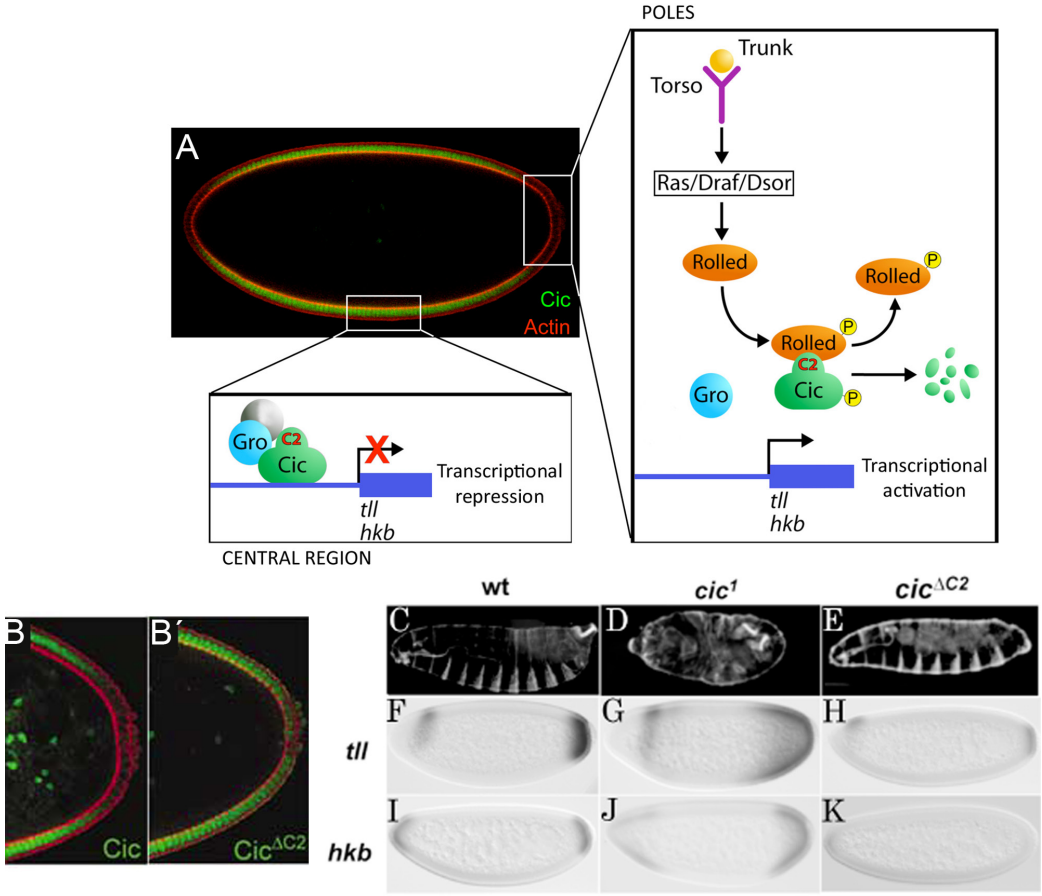


Figure 13. Regulation and functions of Cic in embryonic AP patterning. **(A):** Immunostaining of a blastoderm showing the nuclear distribution of Cic. Schematic representation of regulation of Cic in the embryo. Activation of the Torso pathway at the poles results in activated MAPK binding to the Cic C2 motif, leading to Cic downregulation and relief of its Gro-dependent repressor activities over *tll* and *hkb*. **(B, B´):** Immunostaining of embryos shows nuclear localization of Cic in wild-type embryos and embryos expressing $Cic^{\Delta C2}$ (MAPK insensitive). Note that $Cic^{\Delta C2}$ perists in the terminal nuclei. Cuticle phenotypes, tll and hkb expression in wild-type **(C,F,I)**, cic^{1} **(D,G,J)**, or embryos expressing the $Cic^{\Delta C2}$ protein under maternal control **(E,H,K)**. Adapted from Astigarraga et al., 2007.

cytoplasm, but not degradation (Astigarraga et al., 2007). In the ventral follicular epithelium, Cic represses *mirr* and this is required for the formation of the Dl gradient and for the formation of the dorsal appendages of the eggshell (Andreu et al., 2012b, Atkey et al., 2006).

Furthermore, the gene products of the *bwk* locus have been shown to be necessary for the formation of the dorsal appendages (Dorman et al., 2004; Rittenhouse and Berg, 1995), suggesting that this feature could also require the action of Cic.

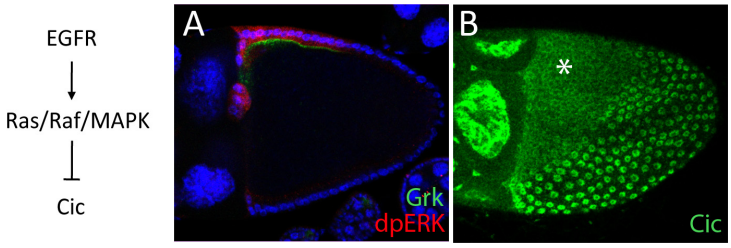


Figure 14. Cic responds to the EGFR signal during late oogenesis. Grk activates the EGFR cascade which culminatesin phosphorylation of Cic by dpERK. Ventrally restricted Cic participates in the formation of the Dl gradient.

Cic contributes to DV patterning later during development as well. During stage 5 of embryogenesis, Cic represses the homeobox gene *ind* at the dorsal-most cells, by binding to a canonical TGAATGAA sequence in its promoter, identified as A-box. This repression is relieved at the ventrolateral domain of the embryo due to the activation of an RTK pathway. Activation of the Torpedo receptor initiates a EGF signal which downregulates Cic, permitting the expression of *ind* and the formation of the neuroectoderm (Ajuria et al., 2011, see figure 15).

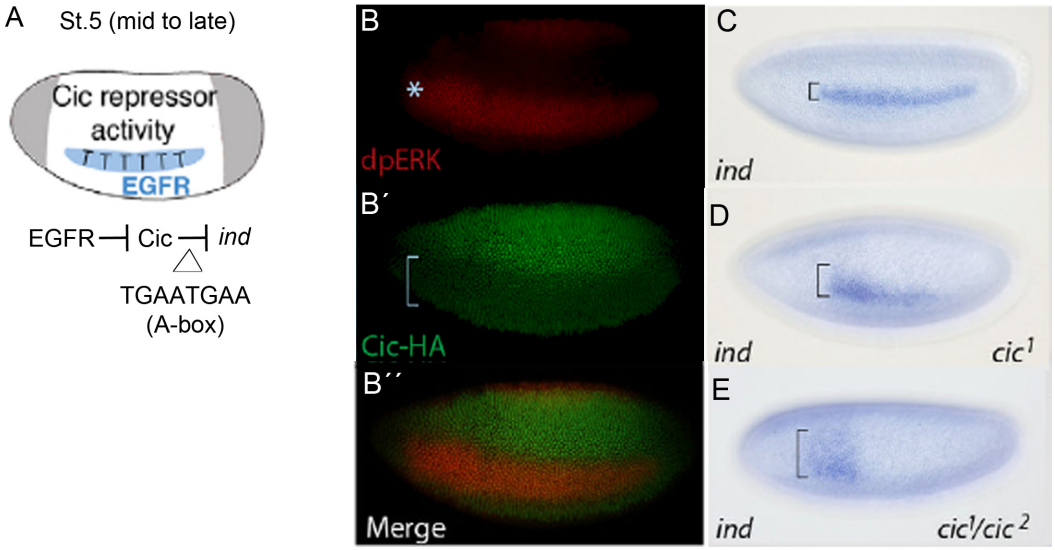


Figure 15. Cic is involved in the formation of the neuroectoderm during embryo cellularization. **(A):** representation of how Cic regulates the expression of *ind*, under the control of EGFR signalling. **(B-B´´)**: Immunostainings of stage 5 embryos showing the domains of EGFR activation (dpERK), Cic expression (Cic-HA) and overlapping domains. **(C-E)**: the expression of *ind* in wild-type, cic^1 and cic^1/cic^2 mutant embryos. Adapted from Ajuria *et al.*, 2011.

Later during development, contributes to wing patterning, by repressing wing vein specific genes (Roch et al., 2002). In the larval imaginal disc Cic is expressed in the presumptive wing pouch region (figure 16). There it represses genes such as *argos* (*aos*), by binding to typical octamerical binding sites in their regulatory sequences (Ajuria et al., 2011). This repression is relieved by an RTK signal. The EGFR pathway, which is active in the presumptive wing vein regions, downregulates Cic activity, allowing expression of wing patterning genes (Roch et al., 2002). Remarkably, this function of Cic is independent of the Gro corepressor (Forés et al., 2015).

Other functions of Cic, that are beyond the scope of our study, exist, so they will be briefly mentioned. In the eye imaginal disc of the larva, Cic inhibits growth, and is under the control of Ras signalling (Tseng et al., 2007). Again under the control of EGFR downregulation, Cic inhibits stem cell proliferation in the midgut by repressing cell cycle genes, such as CyclinE and other effectors, such as Yan (Jin et al., 2015). Finally, a recent study has demonstrated that downregulation by Cic by a different kinase, Minibrain, again downstream ERK signalling, is necessary for the development of the wings, eyes and brain (Yang et al., 2016). Thus, Cic generally senses RTK signals, restricting organ growth and tissue patterning.

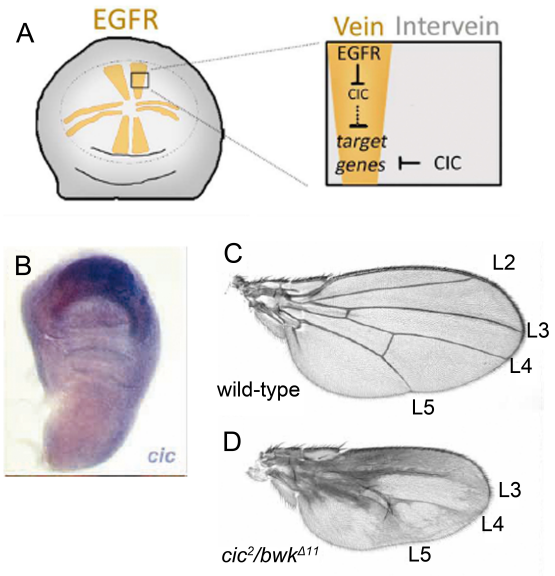


Figure 16. Cic has a function in wing patterning. **(A):** Scheme of Cic-mediated downregulation of targets in the wing imaginal disc downstream EGFR signalling. **(B):** Immunostaining of larval wing imaginal disc showing distribution of Cic in the wing pouch. **(C, D):** wing vein pattern in wild-type and *cic* mutant flies respectively. L2-L5 indicate presumptive wing veins. Extra vein tissue and deformed wings are seen in the absence of Cic. Adapted from Roch et al., 2002.

OBJECTIVES

Patterning of the *Drosophila* embryo and establishment of its body axes has been a subject of studies through many years, however, despite the extensive knowledge gained, still much remains to be discovered. Roughly 30 years after the characterization of the Dl protein as the main effector of DV patterning, the exact mechanism through which it functions is not fully understood. Rather, studies constantly demonstrate that many pieces of the puzzle were missed until now. Our study aims to focus on the repression facet of DV patterning and unravel the molecular mechanisms underlying it. Specifically, the aims of this work are the following:

1. Investigate the possible role of Cic as a repressor during early embryo DV patterning and correlate it to its known function in the DV polarization of the egg chamber.

2. Study the possible relationship between Cic and Dl, as well as with other components of the Dl repressosome, such as the Gro corepressor.

3. Elucidate the mechanism of RTK-mediated modulation of Dl repressor activities over type III dorsal-specific genes and study the possible involvement of Cic as an RTK signaling sensor in this process.

RESULTS

1. Functional analysis of Cic in the establishment of the DV axis

1. Cic mutations reveal a double function in the establishment of the DV axis

Cic has a known role in establishing the DV polarity of the egg chamber, by responding to an EGFR signal and regulating the ventral expression of *pipe* in the follicular epithelium. Pipe is the determining factor for the formation of the Dl gradient, which, in turn, is the main effector of embryonic DV patterning. Therefore, Cic is a factor that participates in the Dl gradient formation and required to transmit the DV polarity positional cue to the embryo. In fact, mutations such as cic^{fetE11} and cic^{fetU6} have been shown to produce dorsalized eggs due to this requirement of Cic in the follicular epithelium (Goff et al., 2001).

Previous to the characterization of the oogenic function of Cic in DV polarity, studies done in our group had suggested that it is involved in embryonic DV patterning, particularly in the repression of dorsal-specific (type III) genes. This hypothesis was based on the observation that embryos deriving from cic^{1} mutant females (hereafter called cic^{1} embryos), showed ventral derepression of *zen*, which is normally restricted to the dorsal-most nuclei by a complex formed by Dl. However, the characterization of the function of Cic in the follicle cells raised the question whether the effects seen on *zen* repression could be reflecting a deregulation of the egg chamber DV polarity, rather than a direct implication of Cic at the level of embryonic patterning. To address this question, we monitored *pipe* expression as a marker to measure the activity of Cic in egg chambers from cic^{fetE11}, cic^{fetU6} and cic^{1} mutant females. To visualize *pipe*, we have used a previously described fusion transgene in which a minimal enhancer module of *pipe*, sufficient to recapitulate its endogenous pattern, drives the expression of *lacZ* (Andreu et al., 2012a), and immunostained stage 9-10 egg chambers. As Goff and colleagues had previously stated, the cic^{fetE11} and cic^{fetU6} mutations resulted in almost complete loss of the *pipe* domain in the ventral follicle cells, and this was reflected in the fertilized egg; embryos deriving from cic^{fetE11}/cic^{fetU6} females, showed only a residual expression of *twi* at the poles, corresponding to the residual expression of *pipe* at the egg chamber, while the transcripts of *zen,* were expanded

towards the ventral region (see figure 17). This alteration of the Dl targets indicated that both its activator and repressor properties were impaired and was in accord with the lack of Pipe. Intriguingly, in trans-heterozygous cic^{1}/cic^{fetU6} females, *pipe* expression was indistinguishable from the wild-type pattern and, as a consequence, the Dl gradient was correctly formed. This Dl gradient was functional considering its activator properties, as it could turn on the expression of *twi* along the ventral circumfence of the embryos. However, in embryos laid by these females, *zen* transcripts were expanded ventrally, indicating abnormality in the repressor functions of Dl. Therefore, the cic^{1} mutation selectively affects repression along the DV axis of the embryo, without impairing DV polarity of the egg chamber. To further confirm this observation, we used the Dominant Female Sterile technique (DFS) to generate cic^{1} mutant clones specifically in the female germline (GLC). We observed that in embryos deriving from cic^{1} clones, *zen* transcripts were clearly derepressed. This defect could only be attributed to the mutant maternal transcripts of Cic produced by the germline nurse cells and deposited to the embryo, since the follicular epithelium of these females was wild- type.

1.2 The main embryonic Cic isoform is involved in embryonic DV patterning

Our finding that Cic has an embryonic function in DV patterning, which is independent from its oogenic function, led us to characterize the isoform that exerts this function. Upon cloning of Cic, the main isoform expressed in the embryo was described (Jiménez et al., 2000). This is a short class Cic, that contains the Gro-associated N2 motif and is maternally provided as an ubiquitous mRNA. As a first assay to test whether this same isoform is involved in repression along the DV axis, we have used full length and truncated transgenes of the embryonic Cic and measured their ability to restore the repression of *zen* in an otherwise *cic* mutant background (cic^{1}/cic^{Q474X}). Indeed, a full-length transgene of the embryonic Cic isoform (herafter named Cic^{A}), which we expressed under the endogenous promoter of Cic, therefore at physiological levels, was capable of repressing *zen* in the absence of endogenous Cic.

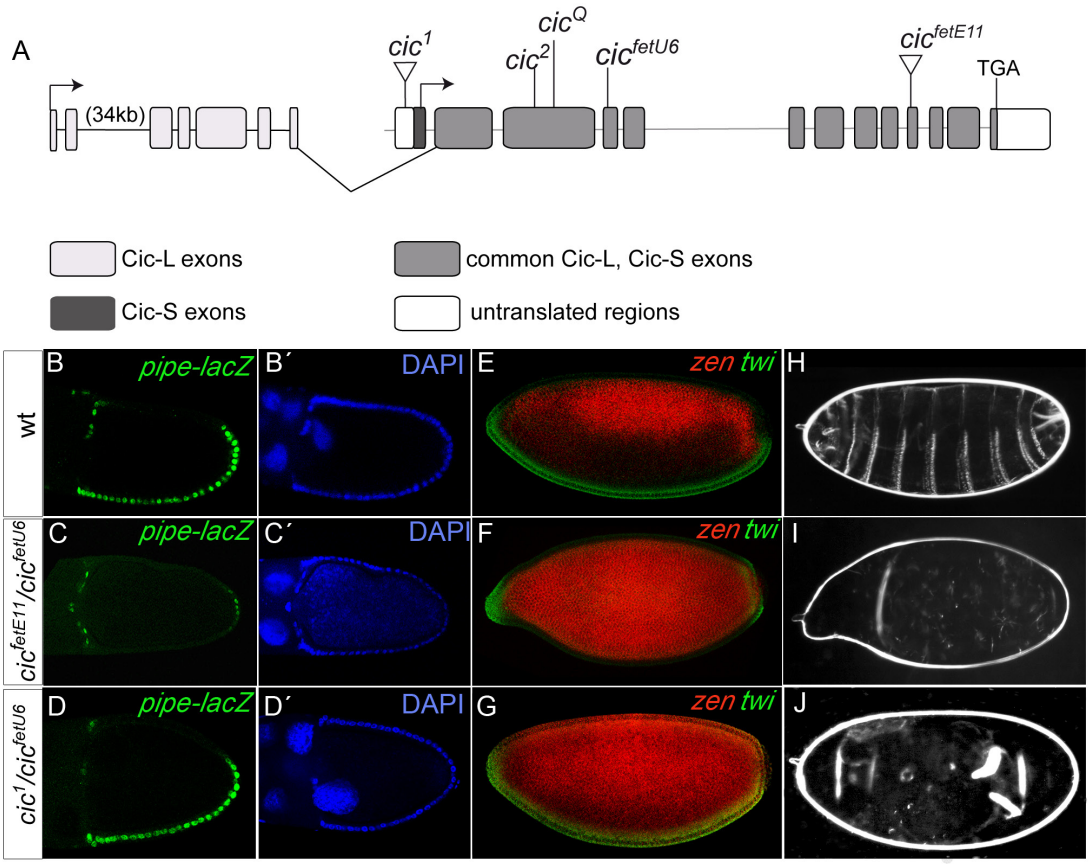


Figure 17. Different mutations of the Cic locus cause different phenotypes in DV patterning **(A):** Schematic representation of the Cic locus and mutant alleles. Coloured boxes represent exon regions and connecting lines intronic sequences. Triangles represent mutations caused by transposon elements while vertical lines point mutations. cic^{Q} (an abbreviation for cic^{Q474X} and cic^{fetU6} are missense mutations that cause truncation of the translated protein. **(B-D´)**: Expression of the *pipe-lacZ* protein in stage 9-10 egg chambers: expression is visualized by staining with an anti-βgal antibody and nuclear staining of the respective ovarian nuclei with DAPI. **(E-G):** mRNA transcripts of *zerknullt* (*zen*) and *twist* (*twi*) in wild-type and *cic* mutant mbryos (stage late 4-early 5). **(H-J):** Cuticle phenotypes of wild-type and *cic* mutant embryos In overall, we observed that the cic^{1} mutation left unaffected the function of Cic in the ovarian follicle cells, however disrupted embryonic DV patterning, in particular the repression of the dorsal-specific gene *zen*. This observation indicated that Cic has two distinct functions in DV patterning, one at the stage of oogenesis embryogenesis.

A smaller form of the protein containing essentially the conserved motifs N2, HMG-box, C2 and C1 (Cic^{mini}) (see Astigarraga et al., 2007), was able to partially rescue *zen* repression in a cic^{1}/cic^{Q474X} background. However, a transgene lacking 60 aminoacids of the N-terminal region ($Cic^{\Delta N60}$) was uncapable of doing so (see figure 18). These findings suggest that the Cic isoform involved in the repression of genes along the DV axis in the embryo, is the main embryonic, Gro-associated Cic-S, which is also involved in AP patterning.

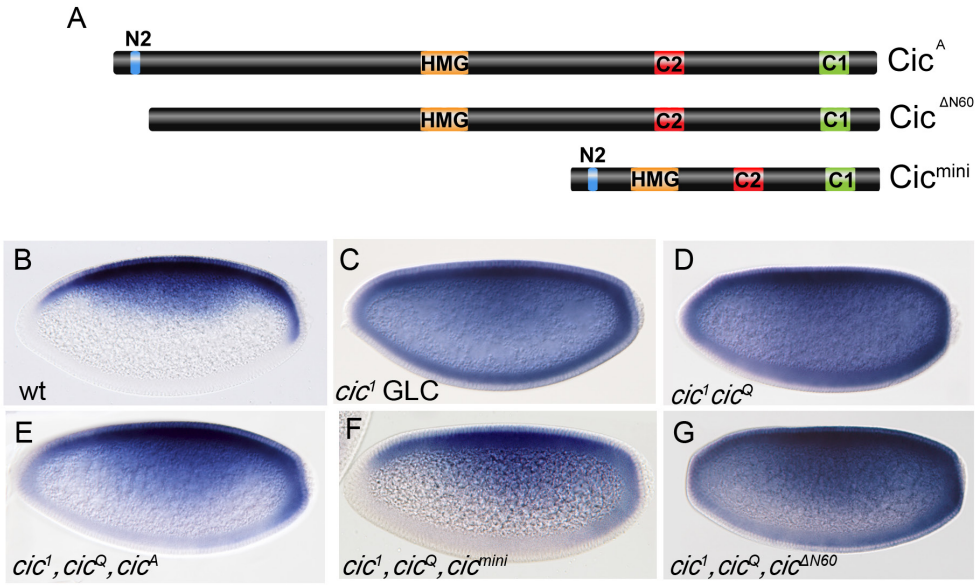


Figure 18. Integrity of the Cic-S protein is required for *zen* repression. **(A):** Diagram of the Cic transgenes used to rescue the DV defects caused by Cic mutant alleles. Cic^{A} is the full-length main embryonic Cic-S isoform. $Cic^{\Delta N60}$ is a variant of Cic^{A}, lacking the aminoterminal first 60 residues, while Cic^{mini} is a synthetic Cic^{A} form in which the most conserved regions have been assembled. **(B-G)**: expression of *zen* mRNA transcripts in embryos that proceed from females that are either wild-type (B), carry germline clones for the cic^{1} allele **(C)** or are trans-heterozygous for the cic^{1} and cic^{Q474X} (cic^{Q}) alleles **(D)**. *zen* transcripts in embryos that proceed from females that are cic^{1}/ cic^{Q474X} and express as rescue transgenes the Cic^{A}, Cic^{mini} or $Cic^{\Delta N60}$ proteins **(E-G).**

Although the effects of cic^{1} and rescue experiments by the Cic variants strongly indicated that the Cic^{A} isoform functions in embryo DV patterning, the nature of the cic^{1} allele led us to an alternative approach to confirm our hypothesis. The cic^{1} mutation is caused by the insertion of a hobo element in the 5' UTR region shared by all possible Cic-S proteins in the follicular epithelium and the embryo (see figure 17). Although it theoretically disrupts their expression, we could not exclude the remote possibility that the inserted element generates cryptic splicing sites or alternative promoter elements that permit the expression of some functional Cic isoform in the follicular epithelium. In this case, the wild-type expression of Pipe would be due to a rescue of the cic^{fetU6} effects by this functional isoform, and not because cic^{1} does not affect the functions of Cic in the egg chamber. To rule out the above possibility, we have used a reverse genetics approach, with the purpose to completely knock down the N2 containing Cic form.

Using the CRISPR/Cas9 system, we introduced a frameshift deletion of 11 bp in the Cic locus, directly downstream the ATG triplet that is the start codon for the main embryonic Cic-S isoforms. This creates a premature stop codon and as a product it gives a short polypeptide, whose sequence is irrelevant to Cic-S and lacks the N2 motif. After establishing a line for the Cic-S frameshift mutation, aka named cic^5, we have further characterized it, examining its effects considering N2-dependent and N2-independent Cic functions (figure 19). Homozygous cic^5 females were completely sterile and laid eggs lacking all their segments and posterior spiracle, as in other strong Cic mutations. This cuticle phenotype indicated that the repressor activities of Cic in the early embryo were highly affected. As markers for Cic repressor activity, we monitored the expression of the genes *tll* and *kni*. In cic^5 mutants, *tll* expression expanded towards the embryo trunk, at the expense of the posterior band of *kni*, indicating that cic^5 behaves as a null allele of Cic-S, considering its functions in embryonic AP patterning. However, wing vein formation, which is independent of the N2 motif, was normal, indicating that only the N2-associated functions of Cic were affected by the cic^5 mutation.

Considering DV patterning, we examined *pipe* in egg chambers from cic^5 homozygous females, and observed that it was expressed as in a wild-type animal, indicating that cic^5 leaves the oogenic functions of Cic unaffected. In the embryo, we monitored *twi* as a Dl- activation marker and *zen* as a repression marker. As in cic^1 mutant embryos, *twi* was activated normally along the ventral circumfence, but *zen* was expanded throughout almost the entire embryo. These effects were more prominent than in cic^1 embryos, and were also observed in females carrying germline clones of cic^5. In overall, the cic^5 allele invalidates the set of Cic-S variants that include the N2 motif and behaves as a null allele in the embryo, while it does not affect other Cic functions, including the ones in the ovary follicle cells.

In conclusion, we have selectively targeted the functions of the major embryonic Cic-S isoform and demonstrated that it is involved in repression events required for the formation of the embryo DV axis, independently of its upstream role in DV patterning during oogenesis, which is possibly accomplished through a different, N2-independent isoform of the protein.

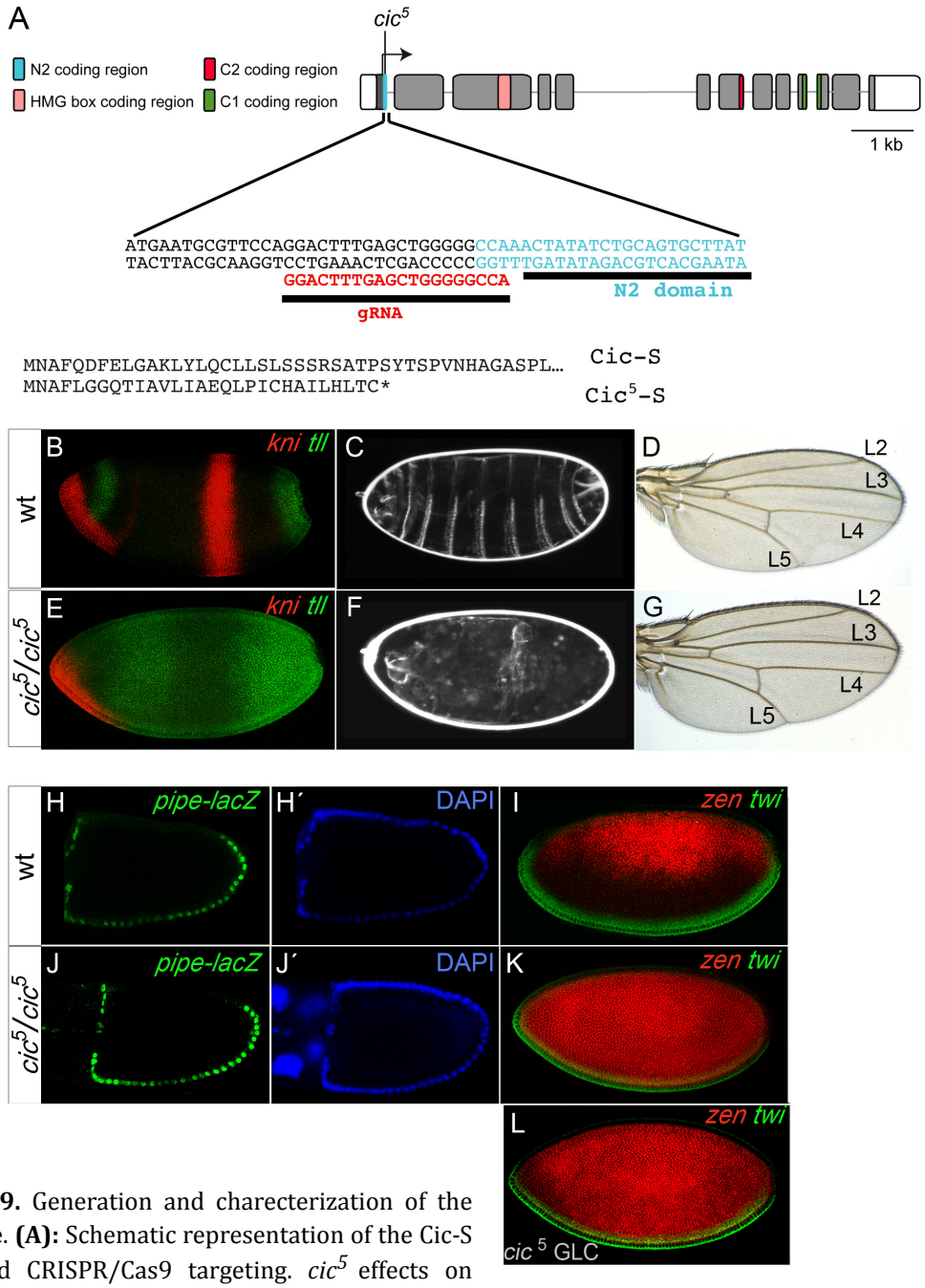


Figure 19. Generation and charecterization of the *cic*^{5} allele. **(A):** Schematic representation of the Cic-S locus and CRISPR/Cas9 targeting. *cic*^{5} effects on known Cic functions: mRNA transcripts of *tll* and *kni* **(B, E)**, cuticle phenotype **(C, F)** and wing vein formation **(D, G)** in embryos deriving from wild-type or homozygous *cic*^{5} females. *cic*^{5} effects on DV patterning: expression of *pipe-lacZ* protein is monitored by staining with an anti-βgal antibody in stage 9-10 egg chambers of wild-type **(H)** and homozygous *cic*^{5} females **(J)**. DAPI staining is used to visualize of the nuclei of the follicle cells of the egg chambers **(H´, J´)**. mRNA transcripts of *zen* and *twi* in embryos deriving from females that are wild-type **(I)**, *cic*^{5} homozygous (K), or carry germline clones of the *cic*^{5} mutation **(L)**.

2. Molecular mechanism of Cic function in embryonic DV patterning

In section 1 of the Results we have uncoupled the functions of Cic in DV patterning during oogenesis and embryogenesis. Once we had shown that Cic has a strictly embryonic function in the establishment of the DV axis, we set off to study the molecular mechanism of this function. We had shown that the main embryonic Cic-S isoform is involved in repression of genes along the DV axis during early embryogenesis (stage late 4-early 5). This same isoform represses the terminal genes *tll* and *hkb* during the same stage through a well-described mechanism: it contacts canonical octameric sites in regulatory sequences of its target genes through its HMG-box and C1 domains and recruits the Gro corepressor through its aminoterminal N2 motif (Ajuria et al., 2011a; Forés et al., 2015; Forés et al., 2017a). Furthermore, this repressor function is subjected to the control of the Torso RTK pathway through the carboxyterminal C2 motif of Cic, that serves as a docking site for the activated MAPK (dpERK), which inactivates Cic and relieves its repressor activities. Based on this knowledge, we set off to study whether any of the above mechanisms apply for Cic repression of type III dorsal-specific genes.

2.1 Cic represses type III genes through the AT-rich sites of the VRE elements

Repression of the dorsal-specific type III genes in blastoderm embryos is controlled by short elements within their enhancers, nominated Ventral Repression Elements (VRE). Interestingly, the AT-rich regions of *zen* and *tld* VRE´s contain sequences that highly resemble the consensus binding site for Cic (CBS) in its other embryonic targets. In fact, in all AT-rich regions of both genes we identified octameric sites that differ from the CBS by only one nucleotide. This similarity led us to the hypothesis that Cic could be binding to them and act as one of the long-seeked components of the Dl repressosome. To explore this possibility, we have analyzed the interactions between Cic and the AT-rich elements *in vitro* and *in vivo*.

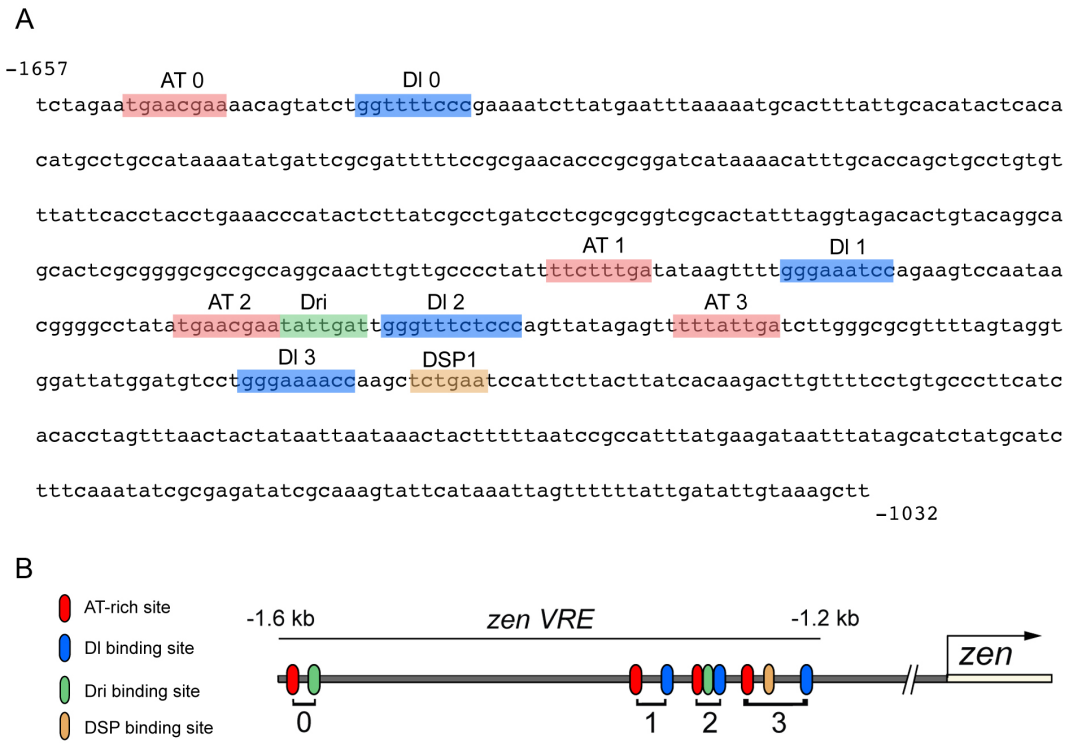


Figure 20. Sequence **(A)** and schematic representation **(B)** of the VRE element at the *zen* promoter.

Cic binds *in vitro* to the AT-rich sites of *zen* and *tld*

The VRE element of *zen* includes four AT-rich sites, nominated AT0, AT1, AT2 and AT3. From these, AT0 and AT1 are 10 bp upstream the respective Dl sites Dl0 and Dl1, while spacing between AT2 and AT3 from their Dl site pair is 8 and 38 bp respectively (Kirov et al., 1993). As mentioned above, the AT-rich site sequences highly resemble the Cic Binding Sites (CBS) of its known targets such as *tll*, *hkb* and *ind*. Specifically, the AT0 and AT2 sites, that both contain the sequence TGAACGAA, present a single base change in position 5 respectively to the *hkb/ind* CBS site (TGAATGAA). The AT1 and AT3 sites, TTCTTTGA and TTTATTGA resemble the reverse complement sequence of the *tll* CBS (TCAATGAA), with a base change in position 4 and 3 respectively.

To test a direct interaction between Cic and the AT-rich sequences of the *zen* VRE, we performed Electrophoretic Mobility Shift Assay (EMSA) experiments, using a synthetic form of Cic, containing the HMG-box and C1 domains (figure 21).

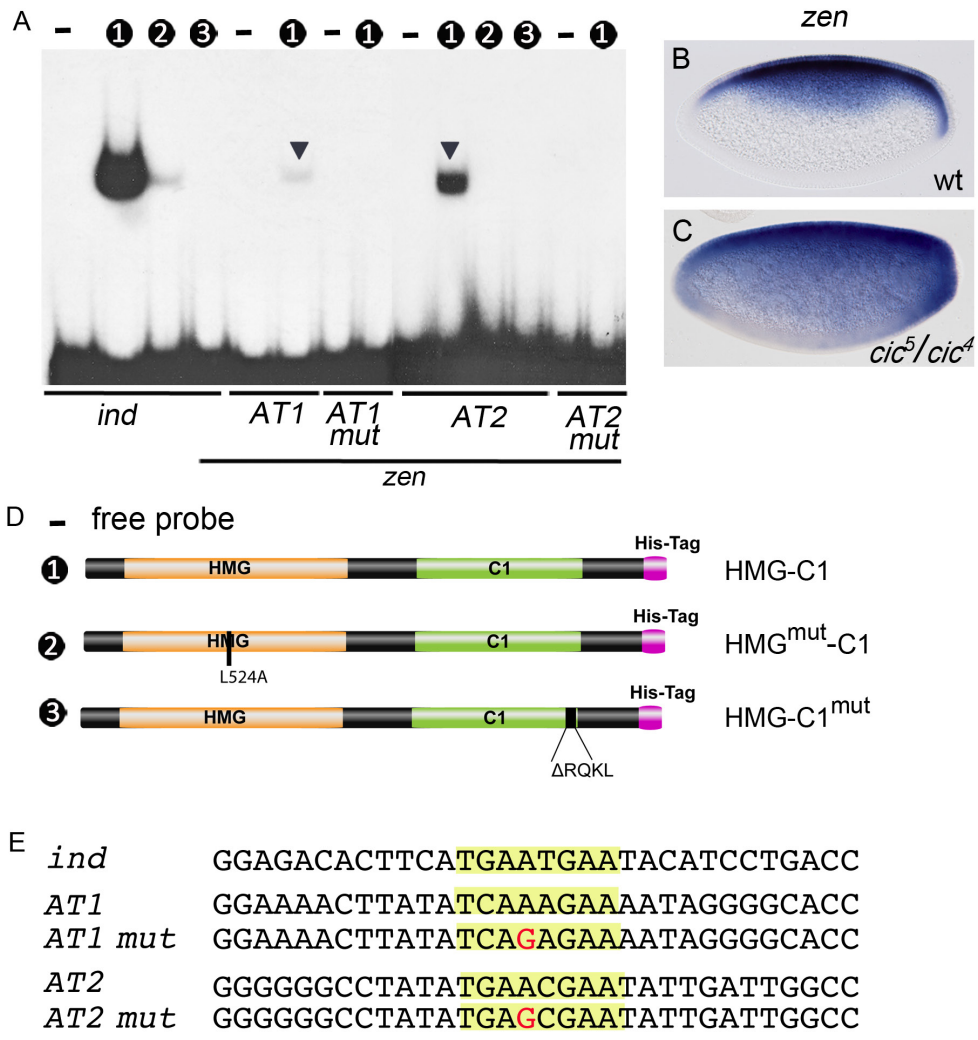


Figure 21. (A): Cic binds *in vitro* to the AT-rich sequences in the VRE element of *zen* and this is linked to *in vivo* repression. Numbers above gel indicate fusion proteins used in the binding reactions. **(B,C):** mRNA transcripts of *zen* in wild-type and cic^5/cic^4 embryos. **(D):** Diagram of Cic-His-tag fusion proteins used for *in vitro* binding assays. **(E):** wild-type and mutant labelled DNA probes containing binding sites from the promoters of *ind* and *zen.*

As a positive control to validate the binding reactions, we have used a DNA probe that contains the CBS for *ind*, incubated with the same HMG-C1 protein. As negative controls, we have used two variants of the HMG-C1 protein: one bearing a single aminoacid change in the HMG-box domain (L524A) (HMG^{mut}-C1) and one bearing a deletion of four aminoacids of the C1 domain (ΔRQKL) ($HMG\text{-}C1^{mut}$). Both of these mutations have been shown to abolish binding to known CBS sites (Forés et al., 2017). We observed that, the HMG-C1 protein, expressed as a His-tag

fusion, bound *in vitro* to short DNA probes containing the AT1 and AT2 sites of *zen*, with affinities that were 12,5 fold and 7,5 fold less respectively to the *ind* CBS. Albeit weak, interaction between Cic and the *zen* AT´s, was specific, since either mutation of the binding site or the protein, disrupted the interaction. When the AT1 site was changed to TCAGAGAA (G<A in position 4 of the octamer), or the AT2 site was changed to TGAGCGAA (G<A in position 4), binding was completely abolished. On the other hand, nor the HMG^{mut}-C1 or HMG-$C1^{mut}$ variants were capable of binding to the native AT2 site (figure 21).

Next, we raised the question whether this weak interaction between Cic and the VRE´s is indeed relevant for *in vivo* repression. As a first assay to answer this question, we have examined the expression of *zen* in embryos that carried *in vivo* the same mutation of the C1 domain that inhibited binding of the protein to the AT2 site *in vitro*. Indeed, this allele, hereafter referred to as cic^4 (see Forés et al., 2017), was not capable of rescuing *zen* derepression caused by the cic^5 mutation (fig. 18). This showed that the integrity of the C1 motif, apart from being necessary for Cic to bind to the AT2 site, is also essential for repression.

In the *tld* VRE (which from now on referred to as *tld* CRM to avoid confusions with the *zen* VRE), two previously described AT-rich sites exist (Kirov et al., 1994), each 10 bp downstream the respective Dl site. The AT1 (TTCGTTCA) resembles the CBS of *hkb/ind* with a single base change in position 4 while the AT2 site of *tld* contains two putative suboptimal CBS sites that overlap: TCCATTGA and TGAATGCA. The first putative site resembles the reverse complement *tll* CBS, with a change in position 2, while the second putative site resembles the *hkb/ind* CBS, with a change in position 7. A detailed description of the *tld* CRM sequence and the functional conservation of the AT-rich sites along various *Drosophilidae* species can be seen in figure 22. Considering EMSA assays for *tld*, the HMG-C1 protein bound with low affinity to the AT1, and with intermediate affinity to the AT2 site. Similarly to the approach we followed for *zen*, we introduced a single-base mutation in the *tld* sites. When mutating the octamers AT1 to TGAGCGAA and AT2 to TTCATGGATGC, binding was completely lost.

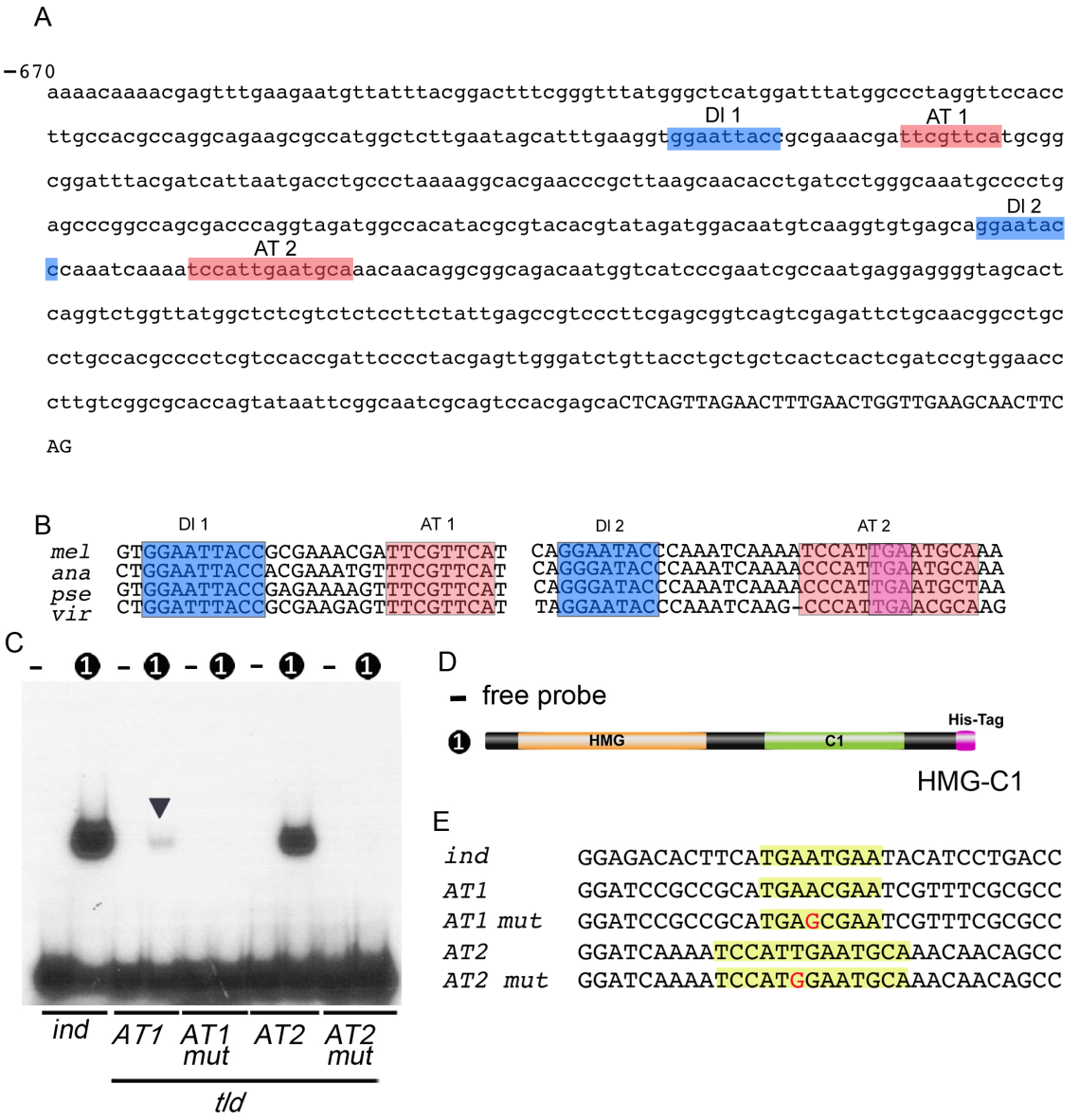


Figure 22. **(A):** The *tld* CRM sequence. Dl and AT-rich sites are highlighted. **(B):** alignment of sequences flanking the Dl-AT pairs 1 and 2 in the *tld* CRM in different *Drosophilidae* species. *mel*; *melanogaster*, *ana*;*ananasai*, *pse*; *pseudooscura*, *vir*; *virilis*. **(C):** Cic binds *in vitro* to the AT-rich sequences included in the CRM element of *tld*. EMSA analyses of Cic fusion proteins binding to different wild-type and mutant labelled DNA probes containing binding sites from the promoters of *ind* and *tld*. Numbers indicate fusion proteins used in the binding reactions. **(D, E):** Scheme of the HMG-C1 His-tag fusion protein and the labelled DNA probes used in the *in vitro* binding assay.

All together, our results demonstrate that Cic directly binds to the octameric AT-rich sites of the *zen* and *tld* promoters. This interaction occurs through the same mechanism by which Cic binds to its other known targets, by contacting DNA with the HMG-box and C1 domains, and is sequence-specific, since it is disrupted by mutations in the AT sites. Finally, compared to the consensus CBS sites, Cic shows low/intermediate affinities for the suboptimal AT-rich sites.

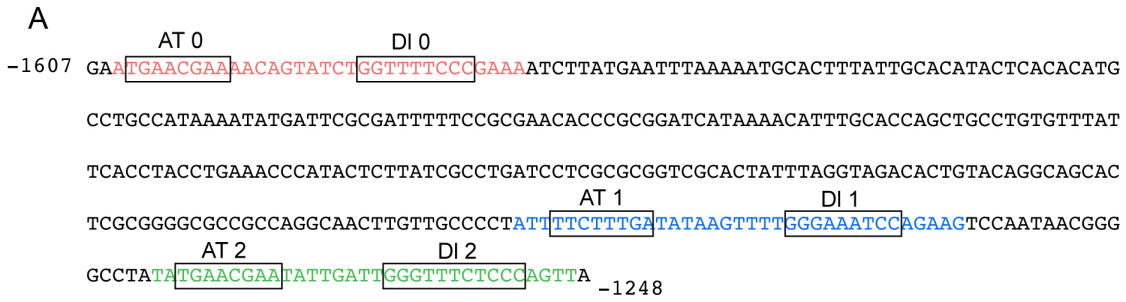
Cic binding to AT-rich sites is related to *in vivo* repression

To test whether the weak interaction between Cic and the AT-rich octamers in the VRE´s of *zen* and *tld* correlates with the repression of these genes in the early embryo, we have used a reporter gene based approach in which fused have endogenous and synthetic VRE elements to the β-galactosidase coding region. After stably inserting these reporters in the genome of flies, we have monitored the *lacZ* transcripts as a readout of the VRE regulatory activity.

First, we have generated a *zen*-VRE-*lacZ* reporter gene. The endogenous *zen* VRE extends between sequences -1,6 kb to -1,0 kb upstream the transcription site of the gene (Kirov et al., 1993). We have opted to include the sequences -1,6 kb to -1,2 kb, which encompass the pairs AT0/Dl0 – AT3/Dl3 and had been previously shown to drive ventral repression in a heterologous assay (Jiang et al., 1993). When this 432 bp fragment was placed three times in tandem, it drove strong expression of *lacZ* that was indistinguishable from the endogenous pattern of *zen*. We have then examined the expression of this reporter in various mutant backgrounds: in embryos derived from homozygous cic^5 females, transcripts were found through the whole circumfence during stages 4-5 embryos, similarly to the ventral derepression observed in *dl* and *gro* mutant backgrounds (see figure 24). These results indicated that the VRE fragment of *zen* we had selected recapitulated ventral repression, for which the presence of Cic is as important as Dl and Gro.

The above experiment suggested a direct relationship between *in vitro* binding of Cic to the AT-rich sites and repression through the *zen* VRE element. To delimit this relationship to the AT-rich sites rather than the whole VRE module, we have generated a reporter gene in which *lacZ* expression is driven by isolated clusters of AT-rich/Dl site pairs. This synthetic enhancer, which we nominated AT/Dl(0-2) contained the naturally occurring AT-rich/Dl sites 0, 1 and 2 of the *zen* VRE, while the sequences flanking the AT/Dl clusters were substituted by random sequences, maintaining the naturally occurring distances. Since the AT/Dl(0-2) element lacked the sequences that normally drive transcriptional activation of *zen*, such as the Zelda binding sites, we have coupled the AT/Dl(0-2) fragment to the promoter of the *sex-lethal* (*sxl)* gene, which drives ubiquitous activation in blastoderm female embryos (figure 23). When combining the *sxl* enhancer with the

synthetic element containing the AT/Dl$^{(0-2)}$ clusters to drive the expression of the *lacZ* reporter, the pattern seen was almost identical to the endogenous *zen* pattern.



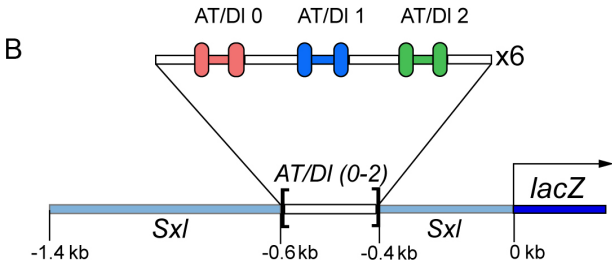


Figure 23. Sequence and diagram of the *sxl*-$^{AT/Dl(0-2)}$-*lacZ* reporter. Endogenous sequences from the *zen* VRE are marked in pink, blue and green colour. Sequences flanking the AT/Dl pairs (black colour) were substituted by synthetic, irrelevant sequences (white bars). This fragment was places six times in tandem substituting a fragment of the *sxl* promoter. Numbers indicate positions relative to the *sxl* transcription site.

Therefore, the AT/Dl site pairs can stand alone as *cis*-repressor elements and contain all the sequences necessary for the Dl repressor complex to be recruited and function. As expected, ventral repression of the *Sxl*-$^{AT/Dl(0-2)}$ reporter was impaired in *dl* and *gro* mutant backgrounds (figure 24). Similar derepression occurred in homozygous *cic*^{5} embryos, providing a direct correlation between the *in vitro* binding and *in vivo* repression of Cic through the AT-rich sites.

We then followed a similar approach to validate the binding of Cic to the AT-rich sites of *tld in vivo* (see figure 25). We have included the sequences -670 to -28 bp upstream the transcription site, which encompass two AT-rich/Dl site pairs, and placed this 642 bp fragment, which we nominated *tld*^{CRM} upstream the *lacZ* gene-coding region. This fragment was sufficient to drive expression similar to that of endogenous *tld* in blastoderm stage embryos. In embryos deriving from *cic*^{5} homozygous females, *lacZ* transcripts were expressed ubiquitously, indicating that the same embryonic Cic isoform that represses *zen* is also involved in *tld* repression through its CRM (VRE) element. To test whether it acts through the

same sequences to which it binds with intermediate affinity *in vitro*, we have generated variations of the *tld*^{CRM} reporter, bearing the same mutations of the AT-rich sites that abolished *in vitro* binding of Cic. In the *tld*$^{CRMmut1+2}$-*lacZ* embryos, which carried single base mutations in both AT1 and AT2 sites, we observed strong derepression, while in the *tld*^{CRMmut2}-*lacZ*, which carried only the mutation of the AT2 site, derepression was partial. These results indicate that Cic indeed, binds to the AT sites of the *tld* CRM to drive its ventral repression. Furthermore, it seems that Cic acts on both AT-rich sites redundantly, since the *in vivo* mutation of the AT2 site, for which it shows higher affinity *in vitro*, is not so deleterious as mutating both AT-rich sites.

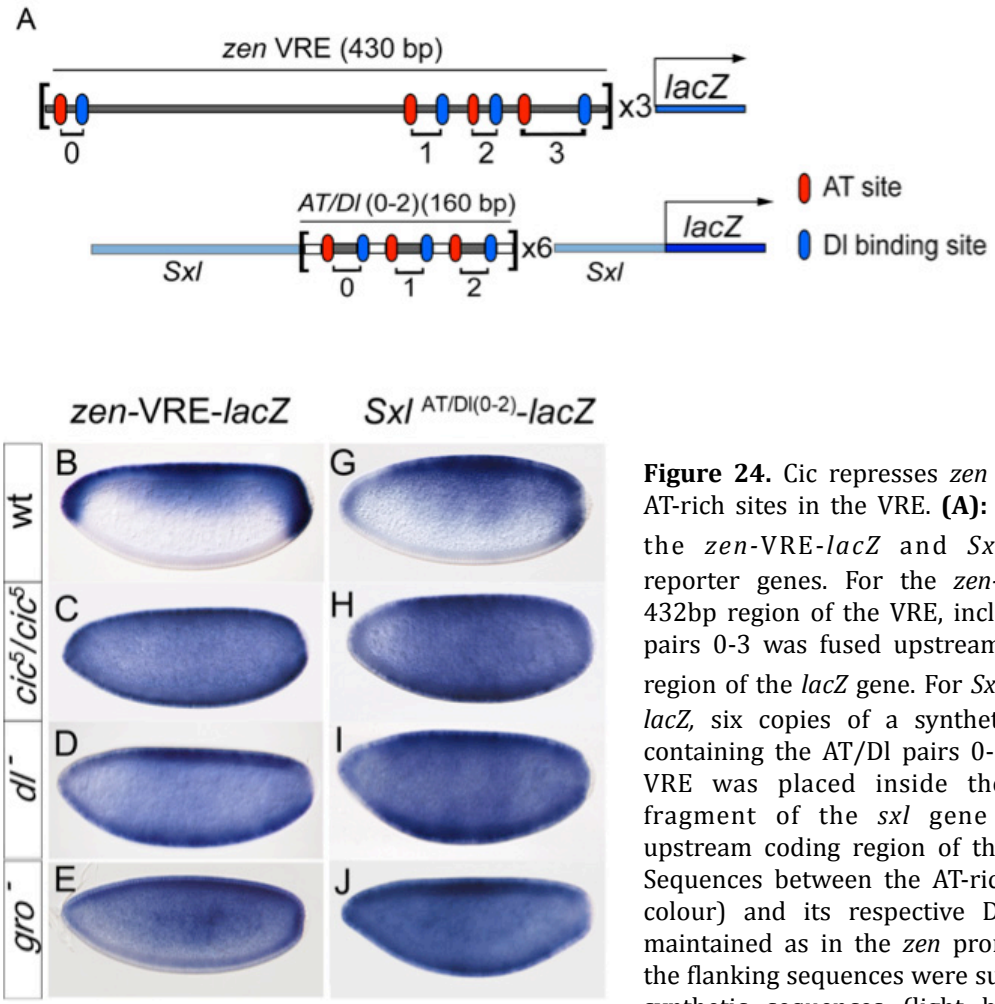


Figure 24. Cic represses *zen* through the AT-rich sites in the VRE. **(A):** Diagrams of the *zen*-VRE-*lacZ* and *SxI*$^{AT/Dl(0-2)}$ reporter genes. For the *zen*-VRE-*lacZ,* a 432bp region of the VRE, including AT/Dl pairs 0-3 was fused upstream the coding region of the *lacZ* gene. For *SxI*$^{AT/Dl\ (0-2)}$-*lacZ,* six copies of a synthetic fragment containing the AT/Dl pairs 0-2 of the *zen* VRE was placed inside the promoter fragment of the *sxl* gene and fused upstream coding region of the *lacZ* gene. Sequences between the AT-rich site (grey colour) and its respective Dl site were maintained as in the *zen* promoter, while the flanking sequences were substituted by synthetic sequences (light blue colour), respecting the distances as seen in the endogenous promoter. **(B-F):** mRNA transcripts of the *zen*-VRE-*lacZ* gene in wild-type, *cic*^{5} homozygous, *dl* mutant, *gro* mutant and *cic*^{5}/*cic*^{4} embryos. **(G-J)** mRNA transcripts of the *SxI*$^{AT/Dl\ (0-2)}$-*lacZ* gene in wild-type, *cic*^{5} homozygous, *dl* mutant, *gro* mutant embryos.

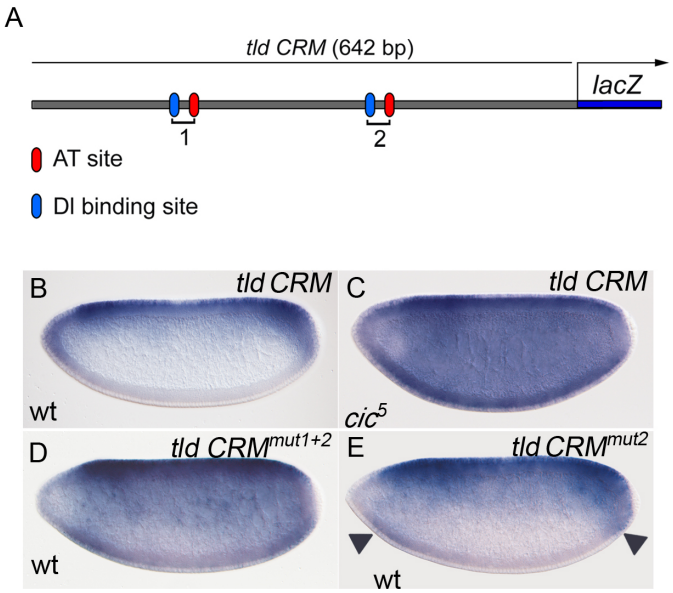


Figure 25. Cic represses *tld* through the AT-rich sites of its CRM element. **(A):** Diagram of the *tld* CRM-*lacZ* reporter gene. A 642 bp region immediately upstream the transciption site of *tld*, containing two pairs of AT/Dl sites was fused upstream the coding region of the *lacZ* gene. The variant *tld* CRM^{mut12}-*lacZ* carries point mutations of the AT1 and AT2 sites, while the variant *tld* CRM^{mut2}-*lacZ* has only the AT2 site mutated. Both point mutations are the same as the ones used above, in the EMSA assay. **(B-C):** mRNA transcripts of *lacZ* in wild-type and homozygous *cic*^{5} embryos carrying the *tld* CRM reporter. **(D-E):** mRNA transcripts of *lacZ* in wild type embryos carrying the *tld* CRM^{mut1+2}-*lacZ* and *tld* CRM^{mut2}-*lacZ* reporter.

2.2 Cic and Dorsal show a synergistic function for repression of dorsal-specific genes

Although Cic binds directly to the AT-rich regions of the VRE modules of *zen* and *tld* and this interaction is relevant with *in vivo* repression, these genes are expressed in the dorsal nuclei, despite the presence of Cic along the whole DV axis. Therefore, Cic-mediated repression must be aided by other factors that are only active in the ventral parts of the embryo. Since it has been long established that repression of type III genes depends on Dl, which is present in a dorsal-to-ventral gradient and binds to sites closely linked to the AT-rich octamers, we set off to explore a possible functional relationship with Cic.

Cic and Dl are mutually necessary for repression of *zen*

As an approximation to correlate the activities of Cic and Dl, we have modified their nuclear gradients in blastoderm embryos and observed how this affects the repression of *zen*, as a representative of type III genes. Normally, at the dorsal-most regions Dl is retained at the cytoplasm by its inhibitor, Cact. By targeting *cact* (see methods), we have generated embryos in which the nuclear gradient of Dl was flat, having the same levels in ventral and dorsal regions. In these embryos, the *zen*-VRE-*lacZ* reporter was repressed along the whole DV axis and transcripts were detected only at the poles (see figure 26). Since Cic repressor activities are downregulated at the poles, the pattern we observed indicated that repression was occuring in the nuclei where both Cic and Dl were active. Furthermore, when $cact^-$ embryos were also devoid of Cic, transcripts of the *zen*-VRE-*lacZ* were expressed ubiquitously. Therefore, providing Dl to the dorsal-most nuclei made them competent of repressing *zen*, while the presence of Dl was not enough to promote repression in the nuclei where Cic activities were downregulated (at the poles of a $cact^-$ embryo), or in the nuclei that are mutant for Cic ($cact^-$, cic^- embryos). This suggests that Dl assists Cic repression in some manner. Of course, this does not mean that Cic and Dl alone are sufficient to mediate repression, since other, ubiquitously expressed factors, such as Gro and Dri have been shown to be essential for the formation of the Dl repressosome.

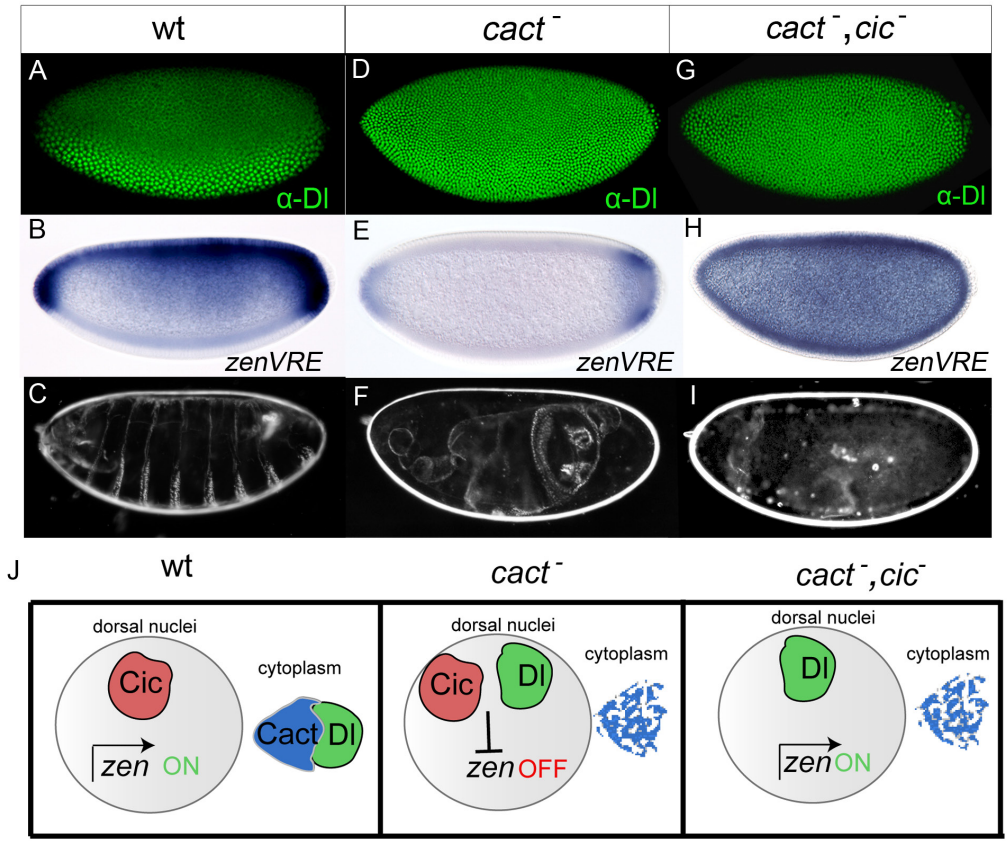


Figure 26. The presence of Dl supports Cic-mediated repression. **(A):** Immunostaining of wild-type stage 4-5 embryos with an anti-Dl antibody. Dl forms a ventral to dorsal nuclear gradient. **(B-C):** *zen*-VRE-*lacZ* mRNA transcripts and cuticle of a wild-type embryo. **(D):** Immunostaining of stage 4-5 embryos expressing and RNAi against *cact*, with an anti-Dl antibody. Dl is nuclear along the DV axis. **(E-F):** *zen*-VRE-*lacZ* mRNA transcripts and cuticle preparation of a *cact*$^{-}$ embryo. *lacZ* expression occurs only at the poles, while the cuticle has a strong ventralized phenotype. **(G):** Immunostaining of stage 4-5 embryos expressing and RNAi against *cact*, that are also devoid of maternal Cic, with an anti-Dl antibody. Dl is nuclear along the DV axis. **(E-F):** *zen*-VRE-*lacZ* mRNA transcripts and cuticle of a *cact*$^{-}$ embryo. *lacZ* expression is uniform while the cuticle lacks any signs of AP or DV polarity. **(J):** Schematic representation of the above experiment.

Mechanism of Cic and Dl transcriptional synergy

As the experiment of the *cact* knockdown showed, Cic and Dl collaborate to repress the transcription of dorsal-specific type III genes. Actually, we could say that, in a qualitative sense, the two proteins show interdependence and synergy considering transcriptional repression. This synergy could be occuring at various levels and through different possible mechanisms. One possibility is that the close proximity of Dl to Cic in the context of the VRE modules serves to increase the affinity of Cic for the suboptimal AT-rich sites. Increase of affinity can be due to a conformational change produced at the level of the DNA, to favour binding of Cic to

the non-perfect octamers. This change can be produced by Dl and possibly other factors, such as Dri, which bind proximaly to the AT-rich octamers. Alternatively, interactions between Cic and Dl could be inducing a conformational change in Cic, making it capable of recognizing a different binding site than the usual. Another possibility is that protein interactions between Cic and Dl do not result in any significant structural change, but, rather, interaction with Dl, which is firmly bound to a high-affinity site, might serve to ´anchor´ Cic to the non-perfect AT-rich site and prevent its dissociation from it. Finally, we could contemplate a different model in which synergy is not due to protein-protein interactions or cooperative DNA binding, but is the sum of interactions between both Cic and Dl with the repressor machinery, which, in an additive manner serve to create a high-affinity platform for the recruitment of corepressor proteins.

To investigate the mechanism through which Cic and Dl collaborate to repress their target genes, we first assayed possible interactions between the two proteins (figure 27). We have used a bacterially expressed GST-fusion of the N-terminal region of Dl (aminoacids 1-380), which spans the conserved Rel Homology Domain (RHD) and various *in vitro* expressed and labelled forms of Cic (see methods). As a positive control we have used an *in vitro* expressed Twi, which is known to interact with the RHD, while as negative controls we have used *in vitro* expressed Luciferase, an unrelated protein deriving from the firefly *Photinus pyralis*. As an additional negative control, we have then incubated *in vitro* expressed Cic forms with GST fusions of the embryonic proteins Huckebein and Hairy, which have no functional relationship with Cic. We were able to detect direct interaction between the RHD domain and all of the Cic variants tested, including the most minimal one, that contained essentially the sequences of the HMG-box and C1 domain, however the unrelated GST-fusions also interacted with the same minimal form of Cic, therefore we were not able to confirm a specific *in vitro* interaction.

We next inquired whether cooperative DNA binding is involved in Cic-Dl synergy. To test this possibility, we have performed Electrophoretic Mobility Shift Assays (EMSA), using bacterially expressed His-tag fusions of full-length and truncated Cic forms incubated with GST-fusions or *in vitro* expressed forms of Dl.

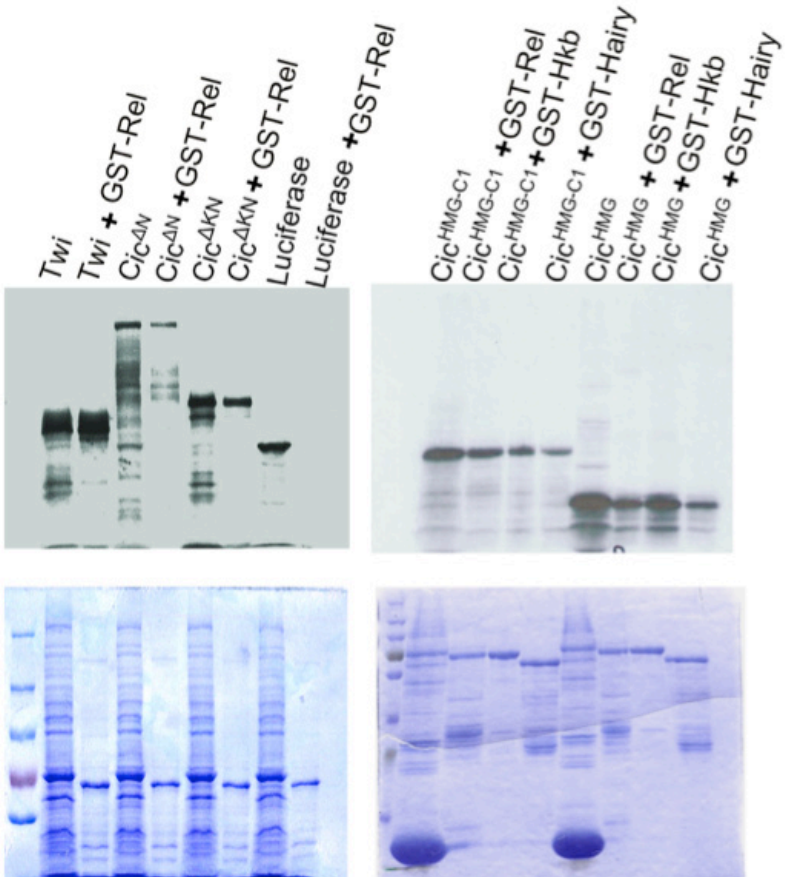


Figure 27. Dl and Cic do not interact *in vitro*. Pulldown assays using GST-fusions of the Dl RHD region and various *in vitro* expressed and radioactively labelled synthetic forms of Cic (see methods). In vitro expressed and labelled Twi and Luciferase proteins were used as a positive and negative control respectively. A GST-fusion of Hkb, which is unrelated to Cic was used to further validate the interactions. Panels on top are gels exposed by autoradiography and below the respective coomassie staining of the gel to assure the same amount of GST-fusion protein was loaded.

As DNA probes, we have used small fragments containing the region spanning each individual AT-rich/Dl site of both *zen* and *tld* promoters. Additionally, we have used a longer fragment encompassing AT-rich/Dl sites 1-3 from the *zen* VRE. Although we could detect binding of Cic and Dl separately on the probes, when incubating both proteins with the target DNA, we were not capable of visualizing a shift in the mobility of the Cic- AT-rich complex, which would indicate the formation of a tripartite complex with Dl. In a further attempt to visualize cooperative DNA

binding *in vitro*, we included a fragment of the Dri protein, which binds to TATTGAT site, which is between the AT2 and Dl2 site in the *zen* VRE (Valentine et al., 1998). Nevertheless, we observed no changes in the mobility of the AT-rich/Cic complex. Similarly, including Gro, which is a well-established partner of the Dl repressosome, did not induce any change in the nature of the interaction. In overall, in our hands, Cic and Dl do not interact *in vitro* and they do not form a complex on a DNA fragment containing their binding sites. Other factors of the repressor complex, such as Dri or Gro do not contribute to the formation of a complex, nor do they increase the intrinsically low affinity of Cic for the AT-rich sites *in vitro*. A possible explanation is that cooperativity at the level of DNA binding exists *in vivo*, but is achieved through a mechanism that involves chromatin and/or nucleosome modifications, which cannot be recapitulated in an *in vitro* system. Also, we cannot rule out the possibility that the formation of a multiprotein complex with the promoter does not require unknown additional factors that may function as scaffold proteins.

The sequence of the AT-rich sites dictates the necessity of Dl for repression

Although the mechanism that underlies the synergy between Cic and Dl is not clear, it is obvious that Dl is required for Cic-dependent repression. Given the low-affinity nature of the AT-rich sites, due to the deviation of one base respectively to the consensus CBS, we questioned whether it is this discrepancy that brings the necessity of Dl. To examine this possibility, we have optimized the AT-rich sites in the context of the *zen* VRE promoter. By site-specific mutagenesis, we introduced four point mutations that converted the natural TGAACGAA (AT0 and AT2) to TGAATGAA (CBS for *hkb/ind*), while TTCTTTGA (AT1) and TTTATTGA (AT3) were converted to TTCATTGA (CBS for *tll*). The reporter gene we created by these changes (*zen*-VRE^{op}-*lacZ*), showed a dramatically altered pattern, being expressed only at the embryo poles, resembling the expression pattern of *hkb* and the expression of the *zen*-VRE-*lacZ* in the *cact*$^{-}$ background. Therefore, optimization of the AT-rich sites resulted in repression throughout the entire DV axis, despite of the absence of Dl at the dorsal most nuclei (figure 28).

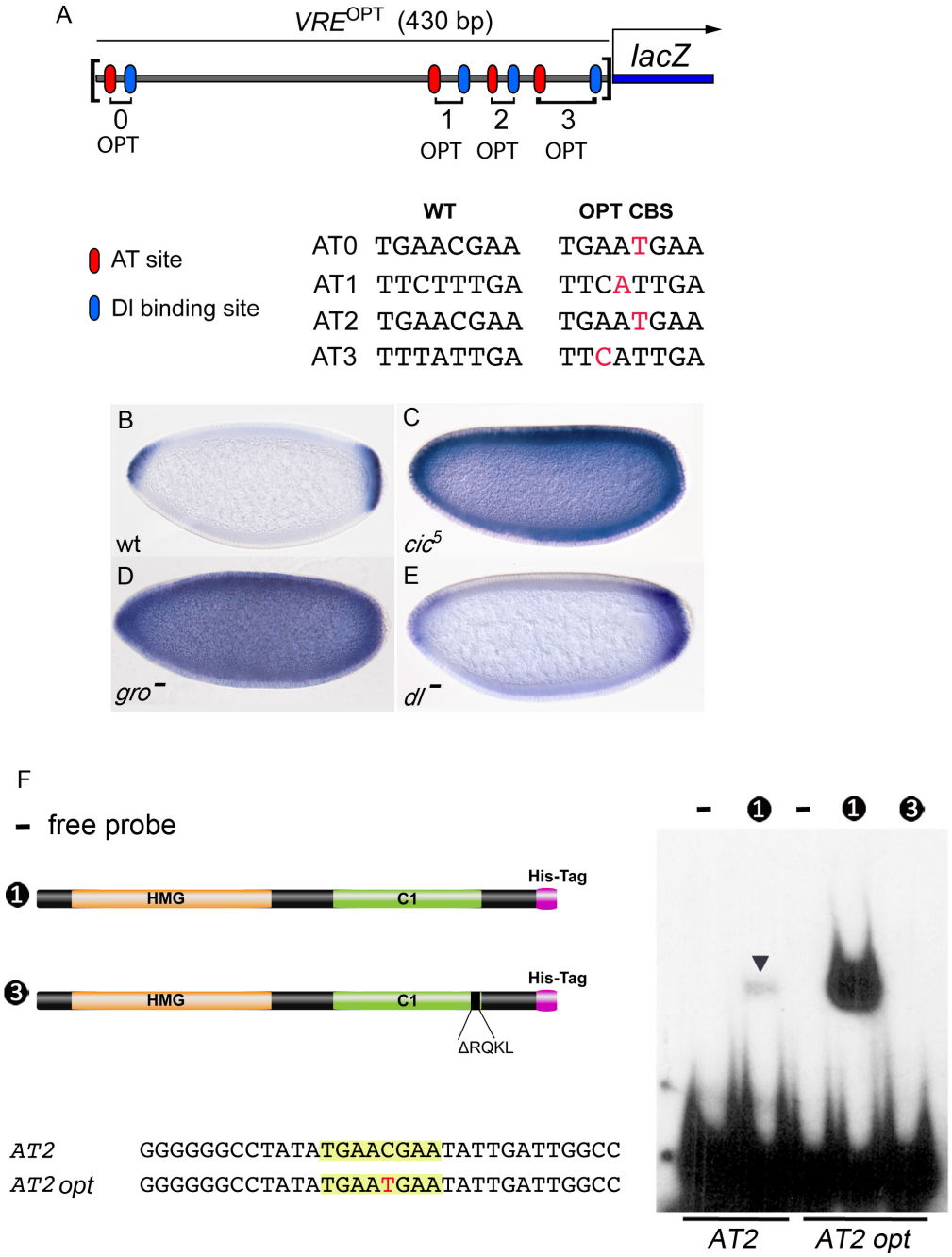


Figure 28. Cic-mediated repression depends on Dl, due to the low-affinity nature of the AT-rich sites. **(A):** Schematic representation of the *zen*-VRE^{op}-*lacZ* reporter, in which the native AT sites have been changed to high-affinity CBS, by change of one nucleotide. When the native AT-rich sites of the *zen* VRE are converted to high-affinity sites, repression depends on the presence of Cic and Gro, but not on Dl. mRNA transcripts of *zen*-VRE^{op}-*lacZ* in embryos that are wild-type **(B)**, *cic*^{5} homozygous **(C)**, carry the *gro*^{MB36} mutation and also express an RNAi against *gro* **(D)**, or express an RNAi against *dl* **(E).** Conversion of the AT2 site to a high-affinity site through a point mutation increases binding affinity for Cic *in vitro*. **(F):** Schematic representation of the Cic His-tag fusion proteins and the labelled DNA probes used and EMSA assay. Numbers indicate fusion proteins used in the binding reactions.

Furthermore, we observed that the repression of the *zen*-VRE^{op}-*lacZ* reporter required Cic and Gro, but not Dl. Consistent with this observation, when we converted the endogenous *zen* AT2 site to the same 'optimal' AT2 site, binding affinity of Cic for the promoter *in vitro* increased dramatically, and this binding was specific, since it did not occur when the HMG-C1^{mut} protein was used. Seemingly, either providing Dl to the nuclei (see *cact* knockdown experiment), either optimizing the AT-rich sites for Cic, lead to the same result, repression of type III targets in the dorsal most regions. Therefore, we conclude that the essentiality of Dl in the repressor complex formed at the VRE promoters arises from the nature of the AT sites, which compromises the binding affinity of Cic to low levels. It seems that the role of Dl must be to render possible an interaction that is a priori unfavourable.

2.3 The Cic N2 motif recruits Gro to the Dl repressor complex

Previous research has established that an indispensable element of the repressosome formed at the *zen* VRE is the Gro corepressor (Dubnicoff et al., 1997). However, interactions between Gro and Dl are weak and dependent on the presence of the AT-rich sites. Based this observation researchers proposed a model according to which a repressor that binds adjacent to Dl, collaborates to create a high-affinity platform for the recruitment of Gro, something that Dl alone is uncapable of doing (Jiang et al., 1992; Ratnaparkhi et al., 2006). This repressor protein was initially suggested to be the ARID-family Dead-Ringer (Retained) protein, which, as mentioned above, has been found to bind to the AT-rich region adjacent to the Dl2 site of the *zen* VRE (Valentine et al., 1998). However, the effects of Dri mutants on *zen* expression are strictly zygotic and prominent after stage 4 of embryonic development (Valentine et al., 1998). Therefore, it is possible that during earlier stages of embryogenesis, a different repressor is collaborating with Dl to recruit Gro and mediate repression.

We have previously shown that Cic mediates repression of dorsal-specific genes by binding *in vivo* adjacent to Dl, to the AT-rich sequences. We have also shown that this function is carried by an isoform of Cic that includes the aminoterminal N2 motif, which is Gro-related in the context of AP patterning in the

early embryo. Therefore, we set off to explore whether Cic is involved in the recruitment of Gro for repression of Dl targets. To answer this question, we decided to compare the effects caused by mutations in the Cic N2 motif with mutations of the eh1-like motif of Dl. The N2 motif is comprised of the sequence KLYLQCLL, with the tripeptide LYL being critical for Gro-dependent repression of the Cic targets *tll* and *hkb* (Forés et al., 2015, unpublished data). To test if this applies for DV targets, we have generated short in-frame deletions within the N2, using the CRISPR/Cas9 system (figure 29). The smallest deletion we have recovered is that of the LY aminoacids. This new mutation of Cic, which we named cic^{6}, caused a very strong phenotype; homozygous females were completely sterile and laid eggs with cuticles lacking all thoracic and abdominal segments. In these embryos the expression of endogenous *zen*, as well as the *zen*-VRE-*lacZ* reporter were strongly derepressed, however, *twi* was indistinguishable from a wild-type animal, indicating that Dl-dependent repression but not activation was impaired. To compare these effects with those of the eh1-like motif loss, we generated a mutant of Dl, using again the CRISPR/Cas9 system. We have recovered a novel mutant (dl^{28}) in which an in-frame stop codon was generated in the eh1-like motif, truncating Dl after the PT sequence. This mutation resulted in a cuticle phenotype with disrupted DV polarity. Surprisingly, in spite of lacking an integrate eh1-motif, in dl^{28} mutants neither *zen* or the *zen*-VRE reporter were derepressed. On the contrary, their expression was even more restricted dorsally, compared to a wild-type embryo. A possible explanation for this discrepancy is that the deletion of the carboxy-terminal region of Dl, which is related to cytoplasmic retention, alters its nuclear gradient. A similar change has been observed with Dl alleles in which larger portions of the C-terminal region were deleted; in these, Dl expands to dorsal nuclei where it would not be normally found (Isoda et al., 1992). This expansion of the Dl gradient would also explain the altered *twi* expression, which was expanded, although attenuated in dl^{28} mutants. The attenuation in the strength of *twi* expression indicates that activation properties of Dl could be compromised in this mutant, which, however, did not seem to affect repression.

In overall, we observed that a deletion of merely two aminoacids in the Gro-associated N2 motif of Cic is deleterious for *zen* repression, while truncating almost the entire Dl eh1-like motif had no effect. On the contrary, repression was shifted dorsally. Given that in dl^{28} embryos the first two residues of the eh1-motif are maintained intact, we cannot exclude the remote possibility that these interact with Gro. However this is quite difficult, since the PT sequence is considerably different from the FS of the eh1 motif, in which the F residue is known to be especially critical (Flores-Saaib et al., 2001; Jiménez et al., 1999), while the more conserved part of the motif has been removed.

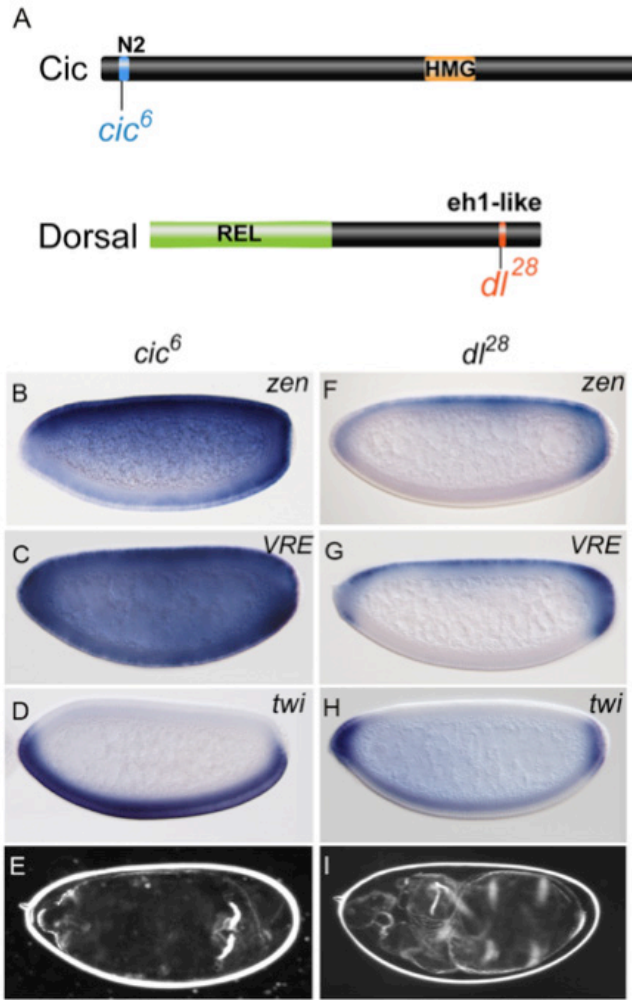


Figure 29. The N2 motif of Cic is more important than the eh1-like motif of Dl for ventral repression of *zen*. **(A):** Scheme of the Cic and Dl proteins indicating the position of the cic^6 (ΔLY) and the dl^{28} (S640X) mutations. *zen* **(B,F)**, *zen*-VRE-*lacZ* **(C,G)** and *twi* **(D,H)** mRNA transcripts in cic^6 and dl^{28} homozygous embryos. cic^6 embryos secrete a cuticle lacking any polarity or segments **(E)**, while dl^{28} embryos show a cuticle phenotype with disrupted DV polarity **(I)**.

The unexpected observation that integrity of the Dl eh1-like motif is not as important as that of the Cic N2, lead us to question which of them eventually acts as a Gro-dependent repressor. To address this question, we have set up an assay to measure the capability of these motifs to function Gro-dependent repressor

elements *in vivo*, in a heterologous context (figure 30). We have based our approach on a previously described experiment that uses the Hairy protein (Parkhurst et al., 1990). Hairy is a bHLH family repressor, which recruits Gro through its C-terminal WRPW motif to represses genes for segmentation in embryogenesis and bristle formation during larval development. When the Hairy protein is, artificially, expressed in the early embryo, it interferes with a process it normally does not regulate: sex determination. The key factor in sex determination is the *sex-lethal* (*sxl*) gene. In males, *sxl* is repressed throughout the whole blastoderm by the Deadpan (Dpn) bHLH family protein that also has a C-terminal WRPW domain, while in females it is expressed ubiquitously. When Hairy is misexpressed in early embryos, it can mimic the repressor activity of Dpn and restrict the expression of *sxl*. For example, when the coding sequence for Hairy is fused to the promoter region of the anteriorly expressed *hb* gene, Hairy is expressed at the anterior half of the embryo, and there it can repress the *sxl* gene in females embryos. Therefore, the hunchback-hairy assay, as it was named, provides a tool to monitor Gro dependent repression, at a timepoint and space it does not naturally occur.

To adapt the above assay to our experiment, we have substituted the WRPW of Hairy with a 13 aminoacid sequence that includes the Cic N2 motif and a flanking region (AKLYLQCLLSLSS) (h^{N2}) or a 13 aminoacid sequence that includes the Dl eh1-like motif (NGPTLSNLLSFDS) ($h^{eh1\text{-}like}$). By using a fragment of the Hb promoter sequence, we have directed the expression of these transgenes to the anterior half of the embryo. Both proteins were fused to a triple-HA tag at their C-terminal, to enable *in vivo* visualization of the protein. After confirming that both hb-h^{N2} and hb-h-$^{eh1\text{-}like}$ transgenes showed nuclear expression at the anterior half of the embryos, we monitored the expression of the *sxl* transcripts as a readout of their repressor activity. In female embryos expressing the h^{N2} chimera, *sxl* expression was restricted to the posterior half, indicating that the N2 motif was acting as a repressor. On the contrary, female embryos carrying the $h^{eh1\text{-}like}$ chimera showed almost uniform expression of the *sxl* gene product, indicating very low, or absent repressor activity. Therefore, the isolated N2 motif of Cic is a stronger Gro-dependent repressor element than the Dl eh1-like motif in a heterologous context.

This explained the observations of the CRISPR deletions and supported a model in which the Gro-dependent repressor of the repressosomes of type III genes is Cic.

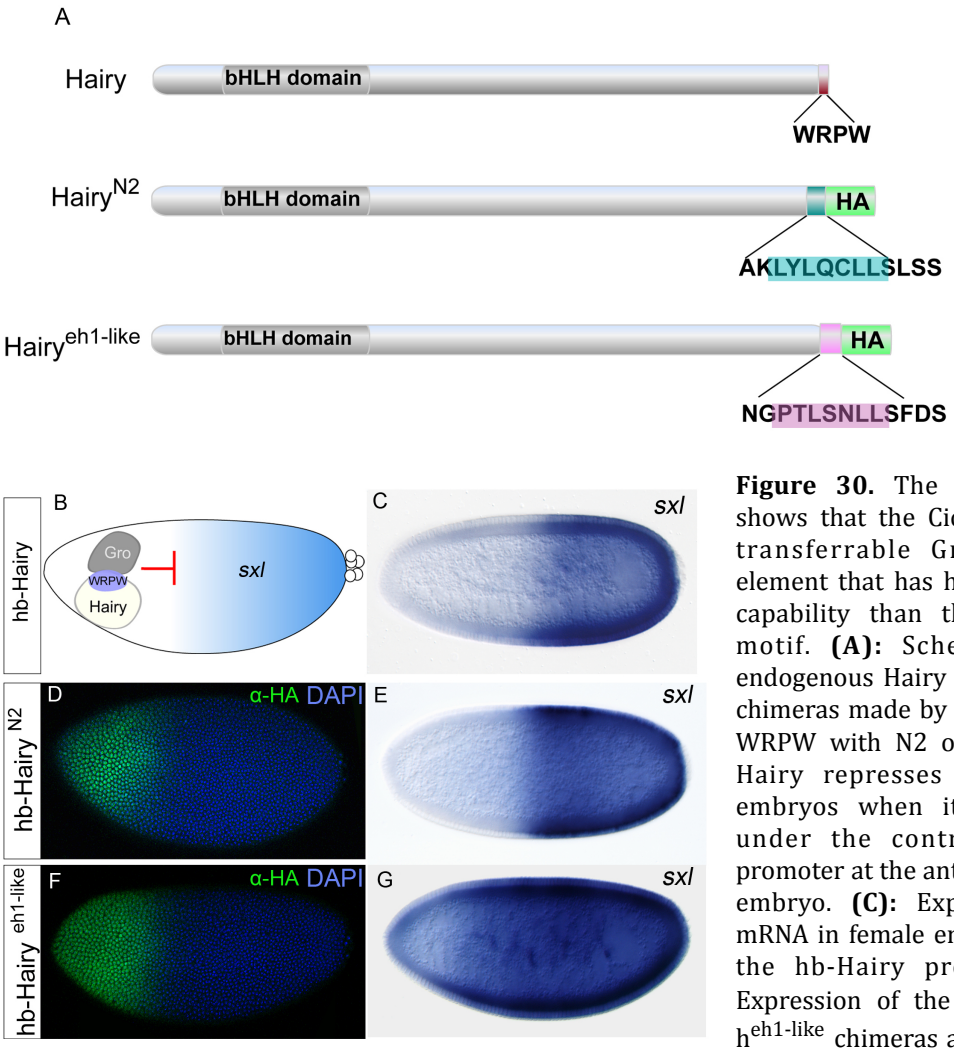


Figure 30. The hb-h/*sxl* assay shows that the Cic N2 motif is a transferrable Gro-associating element that has higher repressor capability than the Dl eh1-like motif. **(A):** Schemes showing endogenous Hairy protein and the chimeras made by substituting the WRPW with N2 or eh1-like. **(B):** Hairy represses *sxl* in female embryos when it is expressed under the control of the *hb* promoter at the anterior half of the embryo. **(C):** Expression of *sxl* mRNA in female embryos carrying the hb-Hairy protein. **(D, G):** Expression of the hb-h^{N2} or hb-h$^{eh1\text{-}like}$ chimeras and *sxl* mRNA in the same embryos. Expression of the protein is visualized by immunostaining with an anti-HA antibody, while DAPI is used for nuclei staining.

To further test our model, we have taken advantage of different Gro mutants. A previous study that analyzed 7 missense mutations of the WD-repeat domain, revealed a mutant allele, Gro^{MB41} (R483H) that is peculiar, in the sense that it severely disrupts the binding of both Hairy (WRPW) and Engrailed eh1 (FxIxxIL) motifs, but leaves unaffected the repressor activities of Cic, as seen by the pattern of *tll* (Jennings et al., 2006). Given that Dl recruits Gro through an eh1-like motif, it is conceivable that this interaction would be affected in a Gro^{MB41} mutant background, unlike the Gro-dependent activities of Cic. We, therefore, considered

the Gro^{MB41} mutation as a good tool to distinguish Dl and Cic mediated repression and used the Dominant-Female-Sterile (DFS) technique to generate germ-line clones (GLC) of this allele. Strikingly, unlike other Gro missense mutations, such as Gro^{e48} (Dubnicoff et al., 1997), Gro^{MB41} had no effect on the repression of *zen*, which showed wild-type expression. This result indicated that at least one of the repressors of the Dl complex is not affected by the Gro^{MB41} mutation. We reasoned that if this protein is Cic, converting it sensitive to the Gro^{MB41} mutation would disrupt the pattern of *zen*. To achieve this, we have used a chimeric form of Cic (Cic^{eh1}), in which a broad region flanking the N2 motif (FQDFELGAKLYLQCLL), was substituted by the eh1 motif of Engrailed (VPLAFSISNIL) (Forés et al., 2015). This transgene, expressed under the endogenous promoter of Cic, was placed in a cic^{1}/cic^{Q} mutant background, and *zen* expression was monitored. The Cic^{eh1} chimera was able to completely rescue the derepression of *zen* in the absence of endogenous Cic, demonstrating that the substitution of the N2 motif does not interfere with the functionality of the protein, instead it converts it to a repressor that recruits Gro in a mode similar to that of Engrailed. Interestingly, when we replaced all endogenous Cic protein with the Cic^{eh1} form, Gro^{MB41} clones had a severe effect, causing complete derepression of *zen*, indicating that Cic was now uncapable of restoring repression. Therefore, we conclude that Cic is the component of the Dl complex that acts as Gro-associated repressor, and is not affected by the Gro^{MB41} mutation.

Our results delineate a model in which weak interactions between Dl and Gro are not involved in repression, but might rather serve to stabilize the complex that could be unstable due to the poor binding of Cic to the AT-rich sites. This model is supported by our observation that, the *zen*-VRE^{op} reporter where Cic binding is strong to the AT-rich sites, repression is independent of Dl, but dependent on Cic and Gro. Thus, Cic recruits Gro to repress *zen* and this recruitment is possibly supported by Dl, that assists both the binding of Cic to the AT sites and the stable recruitment of Gro. Alternatively, the interaction between Dl and Gro could be somewhat transient and serve as a primary recruitment of Gro while binding of Cic to the AT-rich sites is secured.

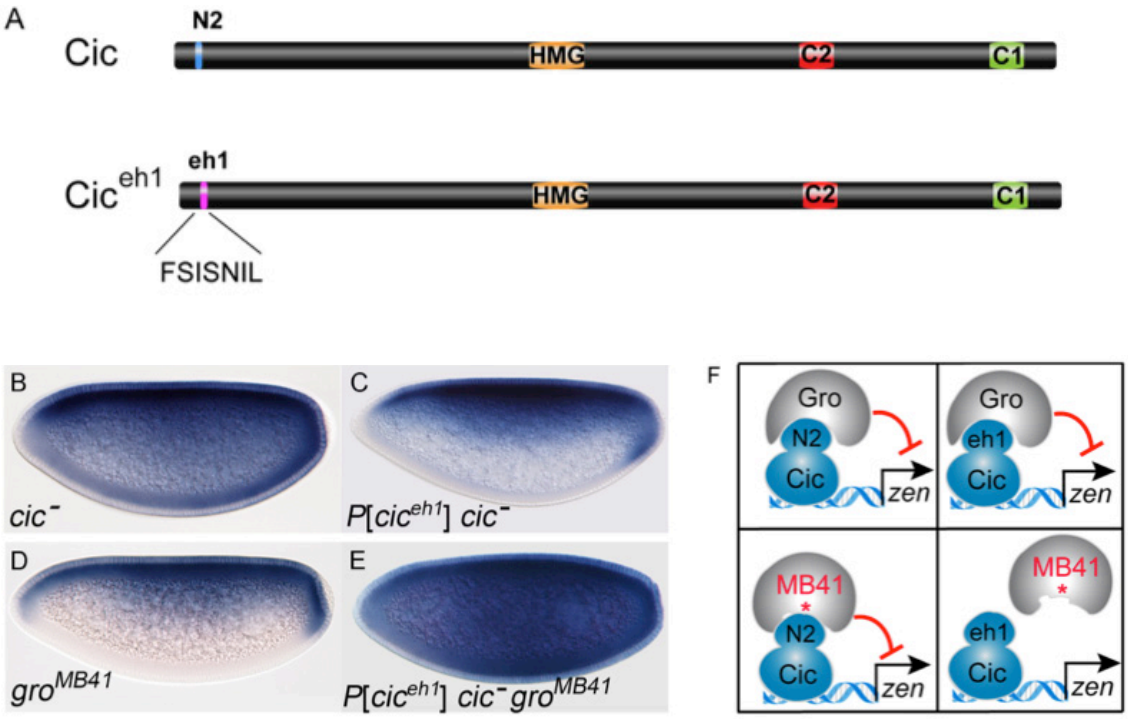


Figure 31. The *gro*^{MB41} mutation affects the repression of *zen* only when Cic is sensitive to it. **(A):** Schematic representation of the endogenous Cic protein and the chimeric Cic^{eh1} in which the N2 motif is substituted by the eh1 motif of Engrailed. **(B):** *zen* mRNA transcripts are derepressed in *cic*^{1}/*cic*^{Q} embryos. **(C):** *zen* mRNA transcripts in *cic*^{1}/*cic*^{Q} embryos that carry one copy of the Cic^{eh1} rescue transgene. Note that ventral repression is restored. **(D):** *zen* mRNA transcripts in embryos proceeding from females carrying *gro*^{MB41} germ-line clones. Ventral repression is not affected. (E) *zen* mRNA transcripts are strongly derepressed in embryos carrying *gro*^{MB41} germ-line clones and all copies of the endogenous Cic protein substituted by the *gro*^{MB41}-sensitive Cic^{eh1} protein. **(F):** Schematic representation of *zen* repression by *gro*^{MB41} insensitive (native) and sensitive (chimeric) Cic forms.

3. Study of Cic as an RTK sensor element during the establishment of the embryonic DV axis

It is noteworthy that the expression of dorsal specific genes such as *zen* is permitted at the embryo poles despite the presence of the Dl protein. This apparent discrepancy has been explained by the activation of the Torso pathway, which has a negative effect on the repressor functions related to Dl (Casanova, 1991; Flores-Saaib et al., 2001; Rusch and Levine, 1994). Different mechanisms could account for the relief of Dl-dependent repression at the embryo termini. One of these is that the concentration of Dl changes in response to Torso activation. Initiation of the MAPK signalling cascade downstream the Torso receptor downregulates Cic, which is a repressor of the *wntD* gene (Helman et al., 2012). This induces the expression of WntD that blocks Toll signalling and limits Dl nuclear localization. To test if this reduction of the Dl levels at the poles accounts for the derepression of type III targets, we have examined the expression of *zen* in embryos proceeding from *wntD* knock-out females ($wntD^{KO}$). In these, the levels of Dl are homogenous along the AP axis (see Helman et al., 2012), so we would expect *zen* to be repressed at the embryo poles. However, we observed that the pattern was indistinguishable from the wild-type, indicating that the increase in Dl levels is not sufficient to trigger repression at the poles. On the other hand, when *wntD* embryos were also devoid of Cic (cic^1/cic^2), *zen* was obviously expanded towards the embryo trunk (figure 32), indicating that relief of repression could be related to Cic.

A different scenario that has been previously suggested is that the anti-repression effect of the Torso pathway is accomplished by disrupting the interactions between Dl and repressor elements. This disruption could be due to the modification of either Dl or its partner repressor (Flores-Saaib et al., 2001; Rusch and Levine, 1994). Since Cic is a part of the Dl silenceosome and it is downregulated by Torso in other contexts, we questioned whether the same mechanism applies for DV genes. To test this scenario, we have activated ectopically the Torso receptor and have monitored the expression of the *zen*-VRE-*lacZ* reporter.

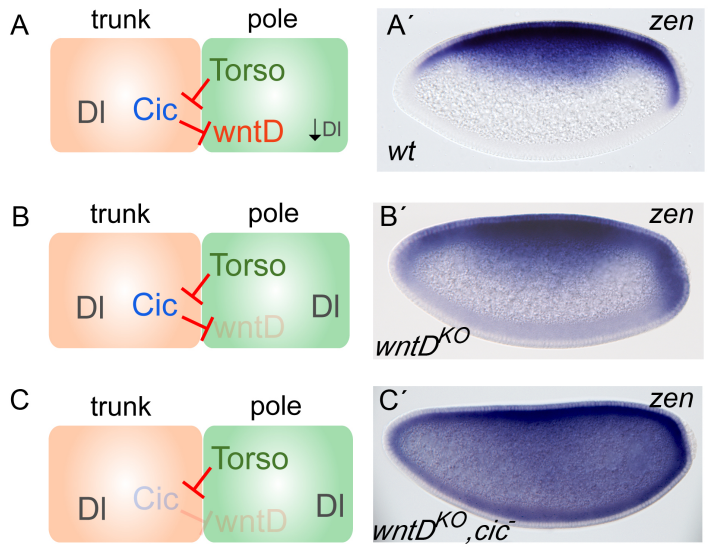


Figure 32. The antirepression effect of Torso on type III genes depends on Cic, but not on WntD. Panels **(A-C)** show the Torso-WntD regulation circuit at the trunk and poles of wild-type, *wntD* knock-out and double *wntD* knock-out/Cic mutants respectively. Panels **A´-C´** show the mRNA transcripts of *zen* in stage 4-5 embryos proceeding from the respective mutant females. Increase of Dl levels in a *wntD* knock-out embryo has no effect in *zen* repression at the poles **(B´)**. Lack of Cic results in *zen* derepression towards the trunk despite the stable levels of Dl due to WntD knock-down **(C´)**.

By driving the ubiquitous expression of the *tsl* gene in the follicular epithelium of female ovaries (TubGal4<UAS-*tsl* genotype-see methods), we obtained embryos in which the Torso receptor was activated throughout the whole circumfence. This resulted in a strong phenotype with 100% embryonic lethality and cuticles that lacked the presence of any segments or polarity signs. In this background, transcripts of *tll* were expanded towards the trunk of the embryo, at the expense of the posterior band of *kni*, indicating that Cic repressor activities were downregulated. To reverse this effect, we used mutant forms of Cic that are insensitive to Torso-mediated downregulation. An insensitive variant is the $Cic^{\Delta C2}$ transgene, in which the entire C2 motif has been depleted. When inserted randomly in the genome by P-element transformation, and expressed under the control of the endogenous Cic promoter, $Cic^{\Delta C2}$ acts as a dominant allele that escapes downregulation at the embryo poles (Astigarraga et al., 2007). Interestingly, when we maternally expressed this dominant form of Cic in a TubGal4<UAS-*tsl* background, it was capable of overriding the ubiquitous activation of the Torso pathway, as seen from the pattern of terminal and gap genes (figure 33). In a TubGal4<UAS-*tsl; P[w+,*$Cic^{\Delta C2}$*]* background, *tll* expression was almost lost, with only some residual expression at the dorsal-anterior nuclei. Accordingly, the posterior band of *kni*, no longer repressed by *tll*, was recovered and expanded posteriorly (fig. 33). Collectively, activating the Torso pathway by

providing Tsl indeed impaired Cic repressor activities and this effect was reversed by expressing the Torso-insensitive form $Cic^{\Delta C2}$.

Considering the DV axis, in the TubGal4<UAS-*tsl* background, the expression of *zen*-VRE-*lacZ* was strongly expanded, however the domain of *twi* extended along the whole AP axis, as in a wild-type embryo, indicating that Dl was still competent of activating type I targets. Since activator and repressor activities of Dl are mediated by the same motifs of the protein (Flores-Saaib et al., 2001; Jia et al., 2002), we suggested that Torso activation does not impair the Dl region involved in repression. Rather, we proposed that the element affected is Cic. To test this hypothesis, we examined *zen*-VRE-*lacZ* in embryos that were TubGal4<UAS-*tsl*; P[w^+,$Cic^{\Delta C2}$]. In this background, the reporter expression was not only restored, but also retracted from the poles. This was accord with our previous observations that, in nuclei where both Cic and Dl are active, repression of *zen* occurs. Furthermore, it excluded the possibility that modification of other components of the repressosome, for example Gro, by the Torso pathway applies for the anti-repression effect. On the other hand, *twi* was slightly shifted dorsally. One possible explanation for the expansion of *twi* is that the $Cic^{\Delta C2}$ form causes a slight expansion of the Dl gradient. This is expected, since the ovarian Cic protein is downregulated through the same C2 motif (Astigarraga et al., 2007). Therefore, by deleting the EGFR response element of Cic, its activities are extended to a dorsal region of the egg chamber, where it normally would not act and this Cic repression drives broader Pipe and, subsequently Dl expression domain.

In conclusion, our findings indicate that the repressor complex at the promoter of *zen* is inactivated through the downregulation of Cic and this effect can be reversed by removing the Cic Torso response element C2. Downregulation of Cic prevents both the recruitment of Gro and possibly its interactions with the AT-rich sites. In overall, Cic acts as a sensor of RTK signalling during embryonic DV patterning.

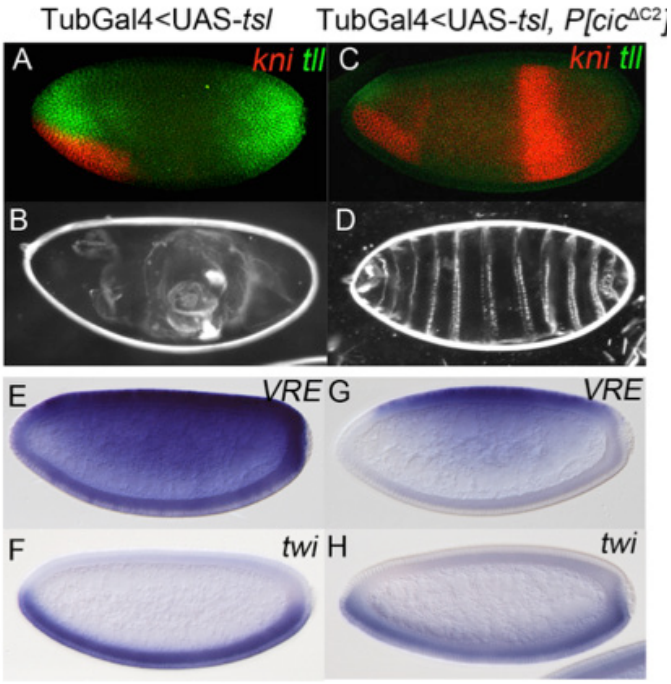


Figure 33. AP and DV effects caused by the ubiquitous activation of the Torso pathway in early embryos. **(A)** mRNA transcripts of *tll* and *kn*i in embryos proceeding from mothers expressing ubiquitously the Tsl protein. The posterior domain of *tll* is expanded towards the trunk, at the expense of the posterior domain of *kni*. These embryos secrete a cuticle that has no segments or posterior spiracle **(B)** and in their DV axis show derepression of the *zen*-VRE-*lacZ* transgene **(E)**, yet normal expression of *twi*. When the Torso-insensitive $Cic^{\Delta C2}$ transgene is expressed in the same background, *tll* is completely repressed and, subsequently, the posterior band of *kni* is recovered **(C)**. Segments are also recovered in the cuticle **(D).** At the DV axis, ventral repression of *zen*-VRE-*lacZ* is restored and it is retracted from the poles **(G)**. *twi* is slightly expanded dorsally **(H)**.

DISCUSSION

General Discussion

How body axes of bilaterian animals are established is a fundamental question in the field of developmental biology. An excellent model organism to study the molecular mechanisms underlying asymmetry establishment is the *Drosophila melanogaster* embryo, thanks to some of its unique features. Unlike other insects, the *Drosophila* embryo is fully developed and ready to hatch in barely 15 hours after the fertilized egg is laid. Therefore, gene regulation and cell communication should be extremely precise and rapid. Due to this demand, *Drosophila* embryogenesis is under prolonged maternal control, which is involved in cell fate decisions even after the zygotic genome has been activated, ensuring that the fatemap of the body axes is present before segmentation and gastrulation begin (Lynch and Roth, 2011). Therefore, it allows to study how maternal information is transmitted to the offspring and how it is integrated in complex networks during early development. Furthermore, the striking similarity between factors and pathways in *Drosophila* and higher organisms, gives the opportunity to obtain knowledge on molecules and processes involved in physiological functions and disease in humans. For the means of this work, we have used the process of embryonic dorsoventral patterning as a model to study gene regulation by the NF-κB holomogue factor Dl, the HMG-box Cic protein and other related factors.

1. The HMG-box family Cic repressor has a dual function in DV axis establishment

In contrast to AP patterning, that integrates inputs from various networks, DV patterning of the *Drosophila* embryo depends on essentially one system, that must orchestrate differential cell fate along the whole axis. Through the dorsoventral maternal system, the female instructs the embryo to create a ventral-to-dorsal gradient of a single factor, Dl. This positional information is generated in the ovary follicle cells and passed on as an asymmetry signal after these are degraded during late oogenesis, and involves activation of an EGFR signalling pathway and various downstream transcription factors. Amongst others, the homeobox Mirr, the HMG-box protein Cic and the sulfonyltransferase enzyme Pipe, create the signal for the ventral-to-dorsal Dl gradient formation through selective

activation of the Toll pathway (Andreu et al., 2012a; Andreu et al., 2012b; Goff et al., 2001; Jordan et al., 2000). In turn, Dl is in charge of coordinating the expression of genes that are hallmarks of different presumptive tissues along the whole DV axis. This means that Dl should regulate the transcription of targets even in regions where it is absent from the nuclei. Consequently, it acts both as an activator and a repressor of zygotic genes. How Dl switches from its, intrinsic, activator mode, to a repressor, has been a long-asked question. After numerous studies carried in various groups, it is accepted that this turnover depends on the contribution of DNA-bound repressors, including the Cut/Dri complex and DSP1 proteins (Decoville et al., 2000; Huang et al., 1995; Ip, 1995a; Jiang et al., 1993; Kirov et al., 1993; Valentine et al., 1998). These factors might act by modifying the chromatin template and assisting the recruitment of non-DNA binding repressor factors, such as Gro, which inhibits the basal transcription machinery. However, phenotypes caused by either *dri* or *dsp1* mutant alleles, or mutations of their binding sites in the target gene promoters, are not always sufficient to explain repression in blastoderm embryos, suggesting the existence of additional factors in the repressor complex formed at dorsal-specific gene promoters.

In this thesis, we have investigated the role of the HMG-group protein Cic as a possible factor involved in Dl-related repression during early embryogenesis. To this end, we have analyzed available mutations of the Cic locus and have used a reverse genetic approach to generate a novel mutation, which selectively knocks down the major embryonic Cic protein. Our analysis has shown that, indeed, Cic is implicated in Dl-dependent repression, but not in Dl-dependent activation during early embryogenesis (stages 4-5). This function is independent of its upstream implication in the maternal dorsoventral system, which affects both Dl-dependent activation and repression. The function we describe is carried by the major embryonic isoform, a Cic-S class protein that bears the aminoterminal N2 motif, a Gro-associated repressor module implicated in other activities of Cic during early embryogenesis, such as the repression of the terminal genes *tll* and *hkb.*

Based on our mutant analysis, we also suggest that the upstream function of Cic in DV patterning of the egg chamber is exerted by a distinct, uncharacterized isoform of the protein that presumably lacks a functional N2 motif. The notion that distinct ovarian and embryonic isoforms of Cic exist and that each one of these has

a separate role in DV patterning explains why different mutations of the Cic locus selectively affect one or other function. Furthermore, it explains why DV polarity of the egg chamber is independent of the Gro corepressor (Technau et al., 2012). In any case, some Cic functions, such as wing vein patterning, are independent of Gro, while others, such as embryonic AP patterning depend on Gro (Ajuria et al., 2011a; Astigarraga et al., 2007; Forés et al., 2015; Roch et al., 2002). This suggests that Cic forms lacking the N2 motif exist. To date, little is known about the structure and activities of N2-independent isoforms, so the molecular characterization of these is a research line of interest that will shed light to possibly unknown functions and targets of Cic, both in the ovary and in other tissues.

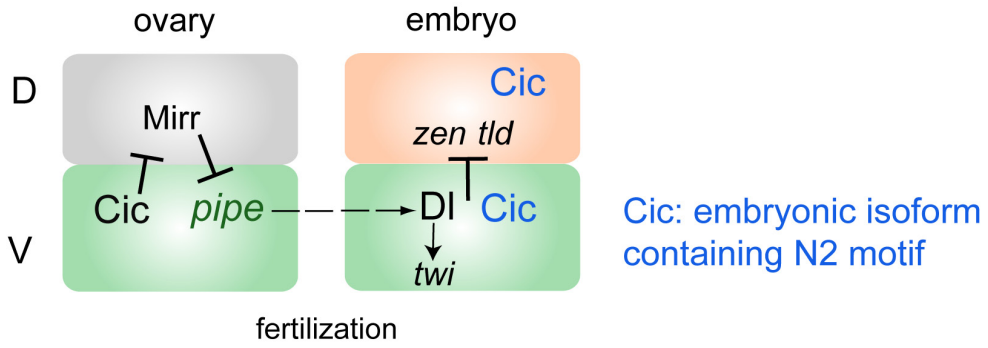


Figure 34. Schematic representation of the double Cic function in DV axis establishment during early *Drosophila* development. D;dorsal region, V; ventral region. An uncharacterized isoform of Cic which is nuclear in the ovary ventral, but not dorsal follicle cells restricts *mirror* to the dorsal follicle cells, permitting the expression of *pipe* ventrally. The product of *pipe* mediates the nuclear localization of Dl to the ventral nuclei of the embryo upon fertlilization. Dl activates ventrally *twi* , and, in concert with the embryonic form of Cic (marked in blue), restricts the expression of *zen* and *tld* to the dorsal-most nuclei. The embryonic Cic is active both in dorsal and ventral embryo nuclei, but is unable to repress *zen* and *tld* at the absence of nuclear Dl.

2. Cic binds, with low affinity, to the AT-rich sites in silencer elements of dorsal-specific genes and requires the presence of Dl to repress them.

Upon the finding that the embryonic Cic-S isoform participates in Dl-mediated repression of type III dorsal-specific genes, we set off to study the molecular mechanism of this function, taking advantage of our knowledge about how the same Cic-S acts in other contexts. We have found that, as for the repression of the terminal genes *tll* and *hkb*, Cic acts as a DNA-bound factor for the repression of *zen* and *tld*, recognizing AT-rich octamers, which it contacts through its HMG-box and C1 motifs. The AT-rich sites in the promoters of *zen* and *tld* resemble the canonical, octameric consensus Cic binding site (CBS), with a discrepancy of one nucleotide and as a result, behave as low-affinity or intermediate-affinity binding sites. We note that have not included the Dl-repression target *dpp* in our study, since the AT-rich sites in its CRM show a higher discrepancy compared to the consensus CBS, and it is known that, at least *in vitro*, this deviation is deleterious for DNA binding (Forés at al., 2017). Furthermore, bona fide binding sites for the NF-T1 factor in the *dpp* Ventral Repression Region, suggest that it might be the major repressor factor for this target (Huang et al., 1995).

In contrast to the AP axis, where Cic binds strongly to its targets, in the DV axis the suboptimal nature of the AT-rich sites compromises the repressor activity of Cic to the ventral nuclei, where Dl is present. Through mutant and reporter gene analysis, we have shown that, in wild-type embryos, the presence of Dl is the limiting factor for Cic-dependent repression to occur, while when the AT-rich sites are converted to high-affinity CBS, Dl is dispensable. Notably, the suboptimal nature is not the same for all of the AT-rich sites. For example, the AT2 site in the *tld* CRM is a better Cic binding site than other AT-sites of *tld* or *zen* and this is possibly correlated with the nature of the Dl2 *tld* site, which shows low score-affinity, similar to the Dl sites of the type I target *sna*. It has been proposed that the low quality of a Dl binding site in a region where Dl levels are low, is compensated by optimal corepressor sites, which foster cooperative occupancy of Dl (Papatsenko and Levine, 2005). In turn, we propose that the intermediate affinity AT2 site is a compensatory mechanism to enhance Cic binding, overcoming low availability of DNA-bound Dl, due to low concentration and low binding site quality.

As our results show, the expression pattern of *zen* and *tld* is a combinatorial readout of the low/intermediate-affinity binding site nature and the dependence on Dl. One could raise the question, in a system where precision is so important, why would gene expression be controlled by poor binding to regulatory sequences? A possible explanation for this seemingly non-rational mechanism could be found at the evolutionary history of the regulated genes. The ancestral *zen* is a member of the Hox cluster, which determines AP polarity. In higher dipterans, such as *Drosophila*, *bcd* emerged through a duplication in the *zen* locus, and was established as the anterior determinant, while *zen* gained a function in the extraembryonic region of the egg (Brown et al., 2002; Stauber et al., 1999). This change of the *zen* function, along with the fusion of the amnion and the serosa in one tissue, the amnioserosa, a feature also obtained in higher dipterans, (Rafiqi et al., 2008; Schmidt-Ott, 2000), must have implied a spatial and temporal shift in its expression pattern (Rafiqi et al., 2008). For example, in the more ancestral, short-germ insect *Tribolium castaneum, zen* is maternally expressed at the anterior most part of the embryo, while in *Drosophila* it is a rapidly turned on zygotic gene that encompasses the dorsal-most circumfence. Considering *tld,* its ancestral form has a similar function with the *Drosophila* homologue, however its expression pattern is different. For example, in *Tribolium*, it is expressed uniformly, but excluded from the egg poles during early embryogenesis (Nunes da Fonseca et al., 2010). On the other hand, the role of Dl as the key regulator of the DV axis is also a feature gained recently during evolution. Immune response could be the more ancestral role of the Toll pathway, while DV axis specification evolved as the amnion tissue has moved to the dorsal part and fused to the serosa in holometabolous dipterans like *Drosophila* (Chen et al., 2000; Lynch and Roth, 2011; Moussian and Roth, 2005). In this line of evidence, the *Tc*-Dl, which is expressed in the serosa, is translocated to the nuclei upon pathogen challenge, but does not seem to regulate *Tc-zen* or *Tc-tld* (Chen et al., 2000; Nunes da Fonseca et al., 2010; Schmidt-Ott, 2005; Stauber et al., 2000). On the contrary, ancestral Cic homologues seem to be involved in *zen* repression, as seen by RNAi knockdown phenotypes in *Tribolium castaneum* (Pridöhl et al., 2017). Possibly, the use of low-affinity sites for Cic and the interplay with Dl is an adaptive mechanism *Drosophila* used to establish a novel expression pattern for *zen* and *tld,* involving additional factors, such as Dl, in their regulation.

It is also possible that the high-affinity Dl sites and low-affinity Cic sites have emerged through subtle changes in the already existing CRM´s of the genes, providing a flexible way to modify gene expression patterns. Indeed, this is a mechanism that is seen more often than *de novo* evolution of developmental enhancers (Perry et al., 2009). Studies on the evolution of DV patterning among insects have given various such paradigms. For example, the dorsal ectoderm of the closely related *Anopheles gambiae* is broader compared to the one of *Drosophila* and this is attributed to two evolutionary modifications in its CRM elements; on the one hand, binding sites for Dl in the *sog* enhancer have changed from high-affinity to low quality, while on the other hand, the AT-rich sites (intermediate-affinity Cic sites) in the *tld* VRE module are lost (Goltsev et al., 2007).

Other examples of this seemingly not ideal mechanism of gene regulation exist. For instance, various genes, namely *dpp* and *wg* (*wingless*), that pattern the wings and legs are driven by a factors binding to sites that deviate considerably from their consensus sequences (Ramos and Barolo, 2013). Of course, this patterning mechanism is not limited to *Drosophila*. At the anterior neural plate and the dorsal neural cord of the invertebrate *Ciona intestinalis*, the *Ortodenticle homeobox* (*Otx*) gene is activated by GATA and ETS (E-twenty six) factors, that bind to suboptimal sites in its enhancer (Farley et al., 2015). In both cases, optimization of the low-affinity sites leads to unappropriate, ectopic activation of the target genes and patterning defects.

Considering the dependence of Cic on Dl for the binding to suboptimal sites, it resembles the so-called ´Hox paradox´ case. Although all members of the Homeobox (Hox) family recognize with high-affinity the same site *in vitro*, they execute highly divergent functions *in vivo*, depending on cofactors that change their binding preferences, a phenomenon known as ´´latent specificity´´(Mann et al., 2009; Merabet and Lohmann, 2015; Slattery et al., 2011). Interaction with different cofactors directs distinct *in vivo* functions through a variety of mechanisms, such as enhancement of the affinity for a suboptimal binding site, changes in binding preferences via allosteric conformation of the protein or the DNA template, and increase in chromatin accessibility by outcompetition of nucleosomes (Beh et al., 2016; Merabet and Mann, 2016; Zandvakili and Gebelein, 2016).

Finally, our finding that the *Drosophila* Cic depends on Dl to regulate targets through non-canonical sites might also have future implications in understanding the function of its human homologue. It has been reported that in certain cancers, such as oligodendrogliomas, Cic mutations result in the upregulation of a set of genes, making the tumors more aggressive and resistant to chemotherapy (Gleize et al., 2015; Okimoto et al., 2016). Recently, ChIP-qPCR experiments have shown that in these tumors, for at least two of the genes derepressed, Cic binds to sites containing a single mismatch compared to the canonical CBS (LeBlanc et al., 2017). This observation has raised the hypothesis that Cic binding affinity is possibly influenced by contextual elements, such as the surrounding sequences and the presence of cofactors. Interestingly, in a different context, interaction between another member of the HMG-box family (HMG-I) and the Dl homologue NF-κB p50 changes DNA-binding affinities and is essential for the regulation of the interferon-β (IFN-β) gene (Thanos and Maniatis, 1992; Zhang and Verdine, 1999). Further studies could reveal if Cic interacts with other factors, including NF-κB proteins during regulation of oncogenic driver genes.

3. Mechanisms of Cic and Dl transcriptional synergy

Based on our findings that Cic is involved in the repression of *zen* and *tld* through low-affinity binding sites and that this function is supported by the presence of Dl, we set off to investigate the molecular interactions between the two proteins and how they influence interactions with the *cis*-elements of the promoters. Under the conditions we have used, we have not been able of detecting a specific protein-protein interaction between Cic and the N-terminal of Dl, which has been shown to interact with other HMG proteins, such as DSP1 (Martin et al., 2003). Under the conditions we have used, we have neither detected cooperative binding between Cic and Dl on a DNA template containing AT/Dl site pairs *in vitro*, nor have we observed an increase in the specificity or affinity of Cic for the AT-rich sites in the presence of Dl. Finally, including other previously identified factors of the repressor complex, such as Dri or Gro, have not been sufficient to visualize the multiprotein complex formation on the promoter *in vitro*. An possible explanation is that the complex formation depends on either unknown additional facors

excluded from our approach, for example scaffold proteins/ nuclear components, or post-translational modifications we cannot reproduce *in vitro.* An alternative explanation is that DNA binding cooperativity involves chromatin and/or nucleosome modifications, which cannot be recapitulated in an *in vitro* system. Similar cases in which DNA cooperativity involves nucleosome competition and bypasses protein-protein interactions have been reported for other factors, including mammalian NF-κB proteins (Adams and Workman, 1995; Mirny, 2010; Vashee et al., 1998).

A line of evidence supporting that Dl and Cic interact cooperatively, is the regulatory code that has been established for Dl targets. Although Dl synergizes with its cofactors through a variety of mechanisms, ranging from cooperative interactions to independent contact of rate-limiting components of the transcription complex, a context-dependent tendency for one mechanism or other exists. At target promoters with tight organization and low flexibility in composition, arrangement and spacing of the binding sites, cooperative DNA binding accounts for at least one of the regulatory mechanisms (Hong et al., 2008; Jiang and Levine, 1993; Szymanski and Levine, 1995). Remarkably, binding sites in both *zen* and *tld* VRE´s are closely linked and small changes in the Dl and AT-rich sites strongly disrupt repression (Valentine et al., 1998, our results).

It is noteworthy that the distribution of the AT/Dl pairs is not the same, even in the VRE of the same gene, and this might indicate the existence of more than one cooperativity mechanisms. For example, the AT0/Dl0 and AT1/Dl1 sites in the *zen* VRE, and both AT/Dl pairs in the *tld* CRM show helical phasing (10bp) which favours protein interactions on the chromatin template (Papatsenko et al., 2009), while the AT2 and Dl2 sites of the same enhancer are separated by 8 bp. One possibility is that in this latter case interactions are facilitated by the presence of Dri, which has been shown to bind to the TATTGAT site located between the AT2 and Dl2 sites and could be modifying the conformation of the DNA template (Iwahara et al., 2002; Valentine et al., 1998). This would also explain the previous observation that crude embryo nuclear extracts show differential binding to the AT1 and AT2 *zen* sites (Valentine et al., 1998), indicating differences in the complexes formed on each site. Finally, the Dl3 site is more closely linked to a downstream putative site for DSP1 than to the CBS-resembling AT3 site, suggesting

that for this pair Dl might not function cooperatively with Cic, but with DSP1. This is a scenario that has been tested in cell culture assays, but not *in vivo* (Ip, 1995). However, we should note that if a complex with DSP1 is formed on the Dl3 site, it probably has a pattern refinement role, or is significant after the cellularization stage, since our synthetic *Sxl-*$^{AT/Dl(0-2)}$ enhancer that lacks this module, is still able to repress ventrally *in vivo* in early embryos. In overall, although we favour a cooperative binding model, more elaborated *in vitro* experiments and further *in vivo* assays, should help confirm this hypothesis and distinguish between the different complexes formed at each AT-rich/Dl cluster.

4. Interactions between the Cic N2 motif and Gro are essential for repression of dorsal-specific genes in the early embryo

An essential component of the repressosome formed at the enhancers of dorsal-specific genes such as *zen*, is the Gro corepressor, which is expressed ubiquitously in the early embryo and inhibits transcription by interacting with the histone deacetylase Rpd3. Females carrying germ-line clones of *gro* alleles that eliminate all maternal products of the gene, such as *gro*^{E48} and *gro*^{BX22}, lay embryos in which the expression of *zen* is strongly derepressed (Dubnicoff et al., 1997). According to previous models, Gro is recruited to the silencing complex of dorsal-specific genes by interacting with Dl and its corepressors. In fact, *in vitro* interactions have been detected between Gro and both Dl and Dri (Flores-Saaib et al., 2001; Valentine et al., 1998). Particularly for Dl, the Gro-interacting region has been mapped to its carboxyterminal, which contains a motif (PTLSNLLS) that shows some homology with the eh1 Gro recruiting motif (FSISNILS) of Engrailed. However, this eh1-like motif of Dl lacks the phenylalaline residue, which has been shown to be critical for Gro association, making the interaction weak (Flores-Saaib et al., 2001; Jiménez et al., 1999). It was initially suggested that this ensures the bifunctionality of Dl as an activator and a repressor (Flores-Saaib et al., 2001; Ratnaparkhi et al., 2006). This would be similar to how Runt proteins are capable of both activating and repressing, thanks to the single amino acid change in their Gro recruiting motif respectively to WRPW peptide of Hairy , which is a dedicated

repressor (Aronson et al., 1997; Levanon et al., 1998). However, repression along the DV axis is also dependent on the context of the Dl binding sites; when intact high affinity Dl binding sites of are removed from the *zen* VRE context, they activate transcription (Jiang et al., 1992; Pan and Courey, 1992). So it was then proposed that factors binding adjacently to Dl assist recruitment, by creating a high affinity platform for Gro, or by exposing a cryptic Gro interacting motif (Dubnicoff et al., 1997; Ratnaparkhi et al., 2006; Hong et al., 2008). Based on our finding that Cic binds to the AT-rich sites adjacent to Dl in the *zenVRE*, we questioned whether the requirement of these sequences reflects the requirement of Cic for Gro recruitment. Since in other contexts Cic acts as a Gro-dependent repressor through its aminoterminal N2 motif (Forés et al., 2015), we set off to examine whether this motif is important for the repression of *zen*. Through the isolation of a novel Cic mutant, in which two aminoacids of the N2 core are deleted (cic^{6}), we confirmed that its integrity is essential for the ventral repression of *zen*. We have then compared this effect with the loss of the Dl eh1-like, by generating a new dl mutant (dl^{28}), which entirely lacks this motif. We were surprised to find that, in contrast to the Cic N2 motif, deletion of the Dl eh1-like motif was not deleterious for repression of *zen*. On the contrary, in dl^{28} mutant embryos, *zen* was even more restricted to the dorsal nuclei compared to wild-type embryos, possibly due to an expansion of the Dl gradient, which has been previously reported for alleles that lack the C-terminal of the protein, and is attributed to reduced cytoplasmic retention (Isoda et al., 1992). Notably, the activation target *twi* seemed to be more affected in dl^{28} embryos, since its levels were much lower than in wild-type embryos, but expanded dorsally. Therefore, our observations indicate that the C-terminal region of Dl is more important considering activator functions than repressor functions, in contrast to what had been believed until now (Flores-Saaib et al., 2001).

The unexpected result that the Dl eh1-like peptide is dispensable for repression of *zen*, raised the question if it is indeed a repressor motif. For this we set up a heterologous assay to measure its Gro-dependent repressor capability. By substituting the WRPW motif of Hairy with the eh1-like or N2, we observed that , indeed, the N2 is a transferrable, Gro-associated repressor motif, while the eh1-like

motif has a poorly ability of acting as such. Furthermore, we have observed that a Gro mutation that disrupts the association between Gro and the eh1-like motif, but not with the N2 motif, leaves *zen* repression unaffected. Interestingly, when Cic was rendered sensible to this mutation repression was severely disrupted. In overall, our results indicate that the main Gro-recruiting element of the repressosome of *zen* is Cic and precicely its N2 motif.

Although we have seen a dispensability of the eh1-like motif for repression, we do not exclude the possibility that interactions between Dl and Gro exist, and that these might take place in sequences outside the eh1-like motif, including the Rel homology domain (Flores-Saaib et al., 2001; Isoda et al., 1992). These interactions could have a role beyond repression, for example stabilizing the protein complex. Another possibility is that weak or transient interactions between Dl and Gro serve as a primary recruitment mechanism of the corepressor to the template, but are not sufficient to trigger Gro-dependent repression. This hypothesis is supported by a recent genome-wide study that has shown that Gro is transiently found in thousands of sites in the genome of the developing *Drosophila* embryo, incuding regions where Dl-mediated activation occurs (Chambers et al., 2017). Recruitment of inactive Gro would be an intelligent strategy to regulate transcription in the context of early embryogenesis, where nuclear cycles are fast and *de novo* transcripts arise within each nuclear division. Similarly, it has been found that different levels of paused RNA Pol II or inhibitory nucleosomes are found genome-wide in the promoters of developmental genes (Lagha et al., 2013; Gilchrist et al., 2010). For example, the general activator Zelda recruits poised Pol II to the proximal promoters of two neuroectoderm specific genes, *brinker* (*brk*) and *sog*, before they are turned on, and this contributes to the rapid and sharp-bounadry expression, that would be naive if it were to be newly recruited after each nuclear division (Boettiger and Levine, 2009; Boija and Mannervik, 2016). Finally, we note that our study has focused on *zen* to study the mechanism of Gro recruitment in Dl-mediated repression. It is yet unknown if Gro is also required for repression of *tld,* so it would be interesting to clarify whether recruitment of Gro by the Cic N2 is implied in this target as well. In overall, we propose a different model regarding the function of Dl. Until now it was believed that low affinity binding between Gro and Dl is what makes Dl bifunctional (Ratnaparki et al., 2006).

In contrast, we propose that this flexibility considering Dl function arises from the presence of its neighbouring cofactors: bHLH proteins strengthen basal activation mediated by Dl, while the low-affinity, context dependent binding of Cic to the AT-rich sites triggers repression. In a manner, we could say that rather than switching from an activator to a repressor by recruiting cofactors, Dl is the cofactor that facilitates the assembly of multi-protein complexes and the action of distinct transcriptional factors.

5. Cic is a sensor of RTK signalling during DV patterning

An aspect of Dl-mediated repression is its regulation by the Torso RTK signaling pathway. At the early embryo, the Torso receptor is selectively activated at the poles, where it initiates the canonical Ras/Raf/MAPK pathway. This modulates the activities of Dl, giving the terminal cells a dorsal fate and setting the anterior and posterior boundaries of the presumptive mesoderm (Casanova, 1991; Helman et al., 2012; Rusch and Levine, 1994). Activation of the Torso pathway permits the expression of *wntD*, a feedback inhibitor that reduces the nuclear levels of Dl (Gordon et al., 2005; Helman et al., 2012). Various hypotheses have been raised on how Dl-mediated repression is relieved by RTK signaling at the embryo poles. One of these could be changes in Dl concentration in response to WntD. However, as it was previously suggested (Liberman et al., 2009; Reeves and Stathopoulos, 2009), and as we have seen with the *wntD* knockout mutant, this low decrease of the Dl levels at the poles is not sufficient to permit the expression of dorsal-specific type III genes. Furthermore, the slight decrease in Dl concentration at the poles cannot explain the asymmetry between anterior and posterior expression of *zen* (Kim et al., 2011). We therefore turned to the scenario that the Torso pathway modifies one or more components of the Dl repressosome, disrupting by this way the complex. Based on the fact that Cic, which is a principal element of the Dl repressor complex, is downregulated in response to the Torso pathway during early embryogenesis in the context of AP patterning, we have questioned whether the same mechanism applies for Dl-repression targets. By activating the Torso receptor ubiquitously in blastoderm embryos and using Cic mutants that are insensitive to RTK-mediated downregulation, we have shown that

the anti-repression effect of the Torso pathway for *zen* is achieved through downregulation of Cic. Since making Cic insensitive to Torso downregulation (with the $Cic^{\Delta C2}$ mutation) restores repression at the poles, we argue that modifications, such as phosphorylation of Dl or Gro downstream Torso activation do not impair repression, as it was previously suggested (Helman et al., 2011; Flores-Saaib et al., 2001). On the contrary, we support that the main sensor of RTK-mediated relief of type III genes at the poles is Cic. Of course, we should not oversimplify this mechanism and keep in mind that it includes other factors as well. For example, the anterior and posterior pole cannot be considered identical; anteriorly, Bcd, which outcompetes Cic as a MAPK substrate, refines the precise expression pattern of *zen* and contributes to the assymetrical expression of the gene considering its anteroposterior pattern (Kim et al., 2011).

To summarize, we have found that the main embryonic Cic-S isoform, containing the aminoterminal N2 motif is involved in DV patterning of the early *Drosophila* embryo, by forming part of the repressor complex that inhibits the transcription of dorsal-specific type III Dl targets. Specifically, Cic is involved in the differentiation of the dorsal ectoderm by mediating the ventral repression of *zen* and *tld*. To exert repression, Cic binds to octameric AT-rich sites in the Ventral Repressor Elements of these genes, recognizing them with low affinity. Consequently, Cic recruits the Gro corepressor through its N2 motif and this results in repression through the inhibition of the basal transcription machinery. Because Cic recognizes the AT-rich sites with low affinity, its action is compromised by the presence of Dl in the nucleus, which in turn depends on Toll signaling. At the ventral nuclei where the Toll receptor is triggered by its ligand, Spz, signaling leads to degradation of Cactus, permitting the entrance of Dl in the nuclei and repression to occur. On the contrary, Dl is retained by Cact at the cytoplasm of the dorsal-most regions where Toll signaling is inactive, and as a result, Cic is incapable of repressing. Furthermore, repression of type III genes is subjected to RTK signaling. At the embryo poles, activation of the Torso receptor initiates the canonical RTK/ Ras/Raf/MAPK pathway. As a final step, activated MAPK downregulates Cic, inhibiting the formation of the repressor complex.

Finally, our results delineate a mechanism during which suboptimal Cic binding sites are used in early *Drosophila* embryogenesis to simultaneously, though

precisely pattern the AP and DV axes. Therefore, we add more to the existing evidence that affinities for transcription factors can have instructive roles during development. A schematic representation of this repression mechanism is seen below.

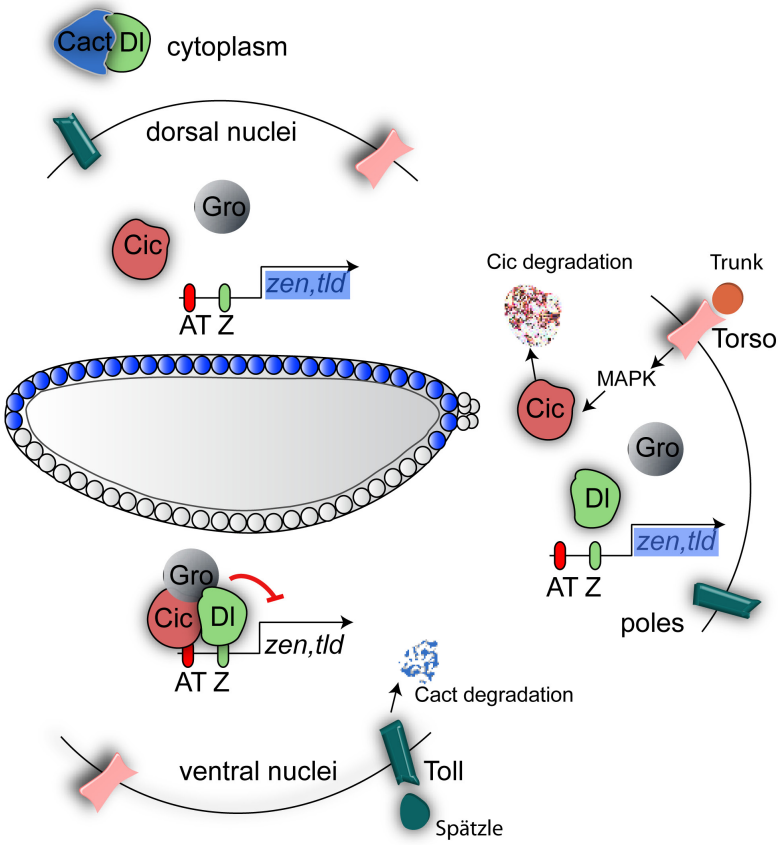


Figure 35. Regulation of type III genes *zen* and *tld* in blastoderm stage embryos. At the ventral nuclei, activation of the Toll pathway results in degradation of Cact, releasing Dl from the cytoplasm. Nuclear Dl forms a repressor complex with Cic and Gro, and gene expression is inhibited. At the terminal nuclei activation of the Torso pathway results in Cic degradation, preventing the repressor complex formation and relieving the expression of the target genes. At the dorsal-most nuclei, Dl is cytoplasmic and, at its lack, Cic is unable to bind to the suboptimal AT sites and repress type III genes.

CONCLUSIONS

- The Capicua repressor participates in the establishment of the DV axis of the *Drosophila melanogaster* embryo independently of its upstream function in the DV polarization of the egg chamber. Precisely, it represses the type III dorsal-specific genes *zen* and *tld*.
- The embryonic function of Cic in DV patterning is carried by the main embryonic isoform (isoform A in Flybase), which is a Cic-S protein that controls the establishment of the AP axis through the repression of terminal genes.
- The embryonic function of Cic in DV patterning requires its N2 motif.
- The oogenic function of Cic in the establishment of DV polarity is exerted through an uncharacterized isoform that does not require the activity of the N2 motif.
- The embryonic isoform of Cic binds directly, with low affinity to the AT-rich sites of the VRE elements in the promoters of *zen* and *tld*.
- Repression through the AT-rich sites depends on the presence of the Dl morphogen, which is present in the nuclei of the ventral-most regions.
- Optimization of the AT-rich sites results in inappropriate repression of *zen* in the dorsal-most nuclei. Thus, association of Cic to high versus low affinity binding sites in the same nuclei permits the simultaneous specification of AP and DV fates.
- Cic is the main factor that recruits the Gro corepressor at the VRE element of the *zen* promoter. The Cic N2 motif is a transferable Gro-associated element, in contrast to the Dl eh1-like motif which is uncapable of repressing through Gro.
- Relief of repression of *zen* and *tld* at the terminal nuclei is independent of the Torso-*wntD* pathway.
- The anti-repression effect of Torso on *zen* is due to Cic downregulation and not due to Dl or Gro modification.

MATERIALS AND METHODS

1. Synthetic DNA constructs

In order to introduce genetic modifications, accomplish *in vitro* protein-protein/ protein-DNA interaction assays, and generate probes to visualize gene expression, we have used a series of molecular biology techniques to assemble, isolate and manipulate synthetic DNA sequences, such as plasmid vectors and short DNA/RNA probes. As bacterial hosts for the molecular cloning experiments, we have used cells of the DH5a *E.coli* strain that had been previously rendered competent by chemical treatment.

1.1 Recombinant DNA techniques

Standard PCR (Polymerase Chain Reaction) recombinant PCR

To amplify fragments from a given template, the following mix was used:

Buffer TaqPol 10X	2,5
dNTPs mix 10mM each	1 µl
Forward primer 250ng/µl	1 µl
Reverse primer 250 ng/µl	1 µl
Template DNA	0,2-0,5ng
Taq polymerase (Thermofisher Scientific)	0,17 µl
H_20	Up to 25µl

Denature	95°C	30 sec
Annealing	48-56°C	35 sec
Amplification	72 °C	1 min/ kb
Repeat cycles		24-29 times

If amplifying the DNA involved breaking cell membranes, (ex PCR on fly preps or bacterial colonies) an initial cycle of 98°C for 2 min was included. Annealing temperature was adjusted according to the overlapping sequence of the primers with the template DNA.

Recombinant PCR

In this PCR variation, two different templates were used at the same time with a pair of primers, to generate a fragment consisting of the combination of the two templates in tandem. During the first cycles the templates are amplified separately. In a next round, the two products are mixed and amplified with two primers, resulting in a unique product consisting of both fragments ligated.

Inverse PCR

For site directed mutagenesis (SDM) on circular DNA templates, we used the inverse PCR method. Two back-to-back primers containing the desired point mutations or insertion/deletion were designed. Then a PCR reaction using the high-fidelity Phusion Taq was set up:

High fidelity buffer 5X	10 µl
dNTPs mix 10mM each	1 µl
Forward primer 250ng/µl	1 µl
Reverse primer 250 ng/µl	1 µl
Template DNA	0,2-0,5ng
Phusion polymerase (New England Biolabs)	0,5 µl
H_20	Up to 50µl

Denature	95°C	30 sec
Annealing	48-56°C	35 sec
Amplification	72°C	20-30 sec/kb
Repeat cycles		24-29 times
Final elongation	72°C	4min

Annealing temperature was adjusted according to the portion of the primer that anneals to the template. Once the PCR product was size-selected and purified to a final volume of 40 µl, its 5´ ends are phosphorylated and ligated to circularize it, using the following procedure:

The following mix was incubated for 30 min at 37°C .

Buffer T4 ligase 10X	2 µl
PNK enzyme (ThermoFisher Scientific)	1 µl
Purified PCR product	4 µl
H_2O	12 µl

Afterwards, 1 µl of T4 ligase was added to the mix and ligation is carried at 18°C for 1h.

RT PCR

With this PCR variant, we have created a cDNA transcript from RNA (in our case extracted from female ovaries).

RNA extracted from female ovaries	0,1-5 µg
Primer-1	15-20 pmol
Nuclease-free H_2O	Up to 12 µl

Incubate 5 min at 65 °C and keep on ice.

Add the following mix to reverse transcribe from the RNA template:

Reaction buffer 5X	4 µl
Riboblock RNAse inhibitor	1 µl
dNTP's 10mM	2 µl
Retrotranscriptase	1 µl

Cycles

39 °C	5 min
42 °C	1 h
70 °C	5 min

At this point we had obtained the single-stranded cDNA retrotranscribed from the RNA template. Subsequently, we amplified the cDNA with primers 3 and 4 and the standard PCR reaction, using high-fidelity Taq enzyme.

1.2 Treatment of DNA

When necessary, plasmids or digested DNA fragments were treated with specific enzymes to introduce modifications on the backbone of the deoxyribonucleotides or fill single stranded overhangs.

Enzymatic digestion for insert liberation and transformant cell test

The digestion mix containing the appropriate buffer at 1X final concentration (unless otherwise required), the restriction enzyme(s) and approximately 1µg of DNA was incubated at 37 °C for 1-2h. For reactions requiring more than one enzymes, the mix was set up according to the guidelines given by the companies providing the reactives (Thermo Fisher Scientific, New England Biolabs). Then, the digested DNA was run on an agarose gel to confirm the digestion efficiency, size-select and clean up.

Dephosphorylation of 5´ends with Alkaline Phosphatase

Following digestion with a unique enzyme, or enzymes that recognize sequences with compatible ends, if it was considered convenient, the 5´ends were dephosphorylated by adding 4 µl of FAST AP buffer 1X and 1 µl AP enzyme to the buffer mix and incubating at 37 °C for 15-20 min.

Phosphorylation of 3´ends with T4 Phosphonucleotide kinase

Buffer T4 ligase 10X	2 µl
T4 PNK (ThermoFisher Scientific)	1 µl
DNA	300 pmol
H_2O	Up to 20 µl

Incubate at 37 °C for 30 min and heat-inactivate at 65 °C for 20 min.

5´ overhang filling with Klenow 3´ 5´exo-fragment

DNA probes used for EMSA experiments were labelled by filling the overhangs with the Klenow exo fragment from the DNA Polymerase that has polymerase, but not exonuclease activity. For procedure see below.

T7 endonuclease treatment

To detect mismatches in amplified DNA (for example when screening for a CRISPR mutant), we used the T7 endonuclease, which cleaves non-perfectly matched DNA recognizes at the double strand at the first, second, or third phosphodiester bond 5 ´of the mismatch. For this, we have amplified the desired region (ie a chromosome region) with standard PCR. Then we have used the following mix:

PCR product	12,5 µl
Buffer type 2 (NEB) 10X	2 µl
T7 (New England Biolabs)	0,5 µl
H_2O	Up tp 20 µl

Incubate the at 37 °C for 15 min and then placed on ice for 3 min. Run the PCR treated with T7 and the non-treated product on a gel.

Ligation

Ligation reactions were generally set up in a final volume of 10 µl, using a ratio of insert: vector 4:1 and were incubated at 18°C for 1 h.

Oligo annealing

Single-stranded synthesized oligos were annealed by bringing them to a concentration of 1µg/µl, heating to 95-98 °C for 5 min to denature and allowing to anneal cooling down.

Gel electrophoresis and purification

DNA and RNA fragment size selection and isolation was done in 1%-2% agarose gels containing BrEt to visualize the nucleic acid. Purification was done with the Illustra™ GFX™ PCR DNA and Gel purification band kit, following the instructions provided by the manufacturer.

DH5a culture

Bacterial colonies were inoculated in 3ml LB (Luria-Bertani) medium (10gr tryptone, 5gr yeast extract, 10gr NaCl in 950ml destilled H_2O, pH adjusted to 7.0 with NaOH 5N) for minipreps and 150 ml LB medium for midipreps, in presence

of 2% antibiotic that guarantees the selective growth of cells that have incorporated the plasmid. The culture was incubated O/N at 37°C under strong agitation.

1.4 DNA preparation and analysis

Once a positive clone was selected and grown, the plasmidic DNA was prepared with the commercial kits Nucleospin (MACHERY-NAGEL) (for minipreps) and Quiagen (Quiagen) (for midipreps) following the manufacturer's instructions. Quantification was done with the use of NanoDrop 2000c UV-vis spectrophotometer and samples were sequenced by GATC Biotech AG, following the company's instructions for sample preparation.

Amplification of genomic DNA from fly preps

For single fly preparations, the fly was anaesthesized on ice and smashed with 40 µl squishing buffer + 0,8 µl Proteinase K from the 20X stock and incubated at 37 °C for 30min-1h. Proteinase K was then inactivated by heating at 95°C for 4min. Fly preps were stored at 4°C for several weeks.

Squishing buffer: Tris-HCl (pH 8,2) 10 mM, EDTA 1 mM, NaCl 25 mM

Constructs

To assemble and manipulate sequences, we have used the pBluescript (pSK^+) and pUC57 vectors. As transformant vectors to generate transgenic animals, we have used the pCaSpeR_hs43_lacZ, pCaSpeR4, placZattB and pCFD3 vectors. For protein expression, we have used the pET17b, pET29b, pGEX and pZEX vectors.

Transgenic constructs

***zen*-VRE$_{(x3)}$ –*lacZ*:** This reporter has three tandem copies of a fragment of the minimal *zen* VRE promoter that has been described to be sufficient to drive endogenous-like expression.A pUC57-T vector with a 467bp fragment of the -1657 to -1226 (containing sites ATZ0-ATZ3) sequence upstream the *zen* start codon, flanked by EcoRI and BglII sites at each of its extremes was synthesized by GenScript®. This synthetical fragment that also contained a NotI site at its 3' site, downstream the ATZ3 site,

was subcloned sequentially in the pCaSpeRhs43_*lacZ* vector as EcoRI/NotI, EcoRI and BamHI fragments.

zen*-VRE^{op}-*lacZ: This reporter consists of the VRE fragme

nt that has the three AT sites (AT0-AT3) mutated by one base pair to consensus octameric Cic binding sites (zopAT).

Using the plasmid pUC57-T-*zen* (see VRE-*lacZ* construct) as a template, the mutation of the AT0 site was introduced by amplifying with primers zop1 and zop2.

Then, using as a template the plasmid carrying the AT0 mutation (zopAT0), the VRE fragment was amplified with primers zop1 and zop3 that introduce the AT1 mutation. In parallel, the same template was amplified with primers zop2 and zop4 that introduce the AT1 mutation as well. Then by recombinant PCR, using as templates fragments zop13 and zop42 the mutations zopAT0 and zop AT1 were combined (zopAT01). This fragment was subcloned in pUC57-T.

The AT2 mutation was introduced by amplifying the zopAT01 fragment with primers zop1/zop5 and zop6/zop2 and then amplifying both products with recombinant PCR with primers zop1 and zop2. The product of the recombinant PCR had mutations in sites AT0 AT1 and AT2 (zopAT02). The AT3 mutation was introduced using the same strategy, amplifying the, subcloned in pUC57-T zopAT02, with primers zop1/zop7 and zop8/zop2 and the products were amplified by recombinant PCR with primers zop1/zop2. The recombinant product was subcloned in the pUC57-T vector and then cloned to the pCaSpeRhs34_*lacZ* vector as a EcoRI/BamHI fragment.

sxl- zen*-ATZ02s-*lacZ: This construct consists of a synthetic fragment of the *zen* VRE including sites ATZ0, ATZ1 and ATZ2. The sequences of the AT and Z sites as well as the sequences between them are identical as in the endogenous VRE, however the sequences flanking the ATZ sites are modified.

A pUC57-T plasmid containing the synthetic VRE of sites ATZ0-2 (ATZ02s), flanked by sites NotI, BglII, EcoRI and BamHI at 5´and 3´ends was synthesized by GenScript®. This 348bp fragment containing two tandem copies of the VRE

ATZ02s was cloned at the EcoRI, BamHI and BglII sites of a modified pSK+ vector that had this sequence in its multiple cloning site. Then, the 1kb NotI fragment containing six copies of the ATZ02s was subcloned in the pCaSepeRhs43lacZ vector that contained the –1,4 kb upstream the *sxl* promoter, substituting the sequences -0,6 kb to -0,4 kb, which are dispensable for the regulation of *sxl.*

***tld*-CRM-*lacZ* and mutants:** A 774 bp fragment containing the sequences -824 to -52 upstream the transcription site of the *tolloid* gene, was amplified from a fly prep of a W^{III8} adult with tld-1 and tld-3 primers and subcloned in a pUC57-T vector. The BamHI insert containing the *tld*-CRM was subcloned at the BglII site of the p*lacZ*attB transformation vector. A variant of the *tld*-CRM promoter containing a point mutation of the AT2 site was created by inverse PCR using primers tldmut-1 and tldmut-2 (*tld*-CRM$^{mut2)}$. Using the resulting plasmid as a template, another variant containing point mutations on sites AT1 and AT2 was created by inverse PCR using primers tldmut-3 and tldmut-4 (*tld*-CRM$^{mut1+2)}$.

Hb-h^{N2S}: This construct encodes a chimeric protein in which the C-terminal aminoacids (268-337) of Hairy was substituted by the sequence AK**LYLQCLL**SLSS of the N-terminal of Cic (aminoacids 11 to 23 of the isoform A of Cic in Flybase). The oligos hbN2S-1 and hbNS2-2 were annealed and cloned as a BamHI/XbaI fragment in a pSK+ vector that contained the sequence coding aminoacids 1-267 of Hairy. Then, the BsteII/XbaI fragment of Hairy-N2S was subcloned in a modified pCaSpeR4 vector that contained the sequence upstream the *hunchback* coding sequence. Finally, a triple HA tag was inserted as an NheI fragment at the C-terminal of the fusion protein coding sequence.

Hb-h^{eh1}: This construct encodes a chimeric protein in which the C-terminal aminoacids (268-337) of Hairy was substituted by the sequence the sequence NG**PTLSNLL**SFDS (aa 635-647) of the C-terminal of Dorsal. The same strategy was used as in the hb-h^{N2S}, using the oligos hb-dor5 and hb-dor6.

CRISPR/Cas9 gRNA plasmids: For generation of mutants with the CRISPR/Cas9 method, gRNA coding sequences were generating by annealing two complement oligos, which were cloned at the BbsI site of the pCFD3 vector.

Protein expression constructs

pET17b-cic$^{\Delta KN}$: In the pCaSpeR4-701 vector, which codes for the full length cDNA of Cic-S, including the 3´ UTR region, the Acc65/NotI fragment at the N terminal was substituted by and the KN adaptor. Then, Cic$^{\Delta KN}$ was cloned as an EcoRI/XhoI fragment in the pET17b vector.

pET17b-cic$^{\Delta N}$: This construct expresses a truncated form of Cic lacking the N terminal sequences encoded between the KpnI sites. The KpnI fragment of the pCaSpeR4-701 vector was substituted by the KN adaptor. Then, Cic$^{\Delta N}$ was cloned as an EcoRI/XhoI fragment in the pET17b vector.

pET17b-cic$^{HMG\text{-}C1}$: This plasmid expresses a minimal form of Cic in which the HMG and C1 domains are assembled (see Forés 2017).

pET17b-cic$^{HMG\text{-}C1}$: This plasmid expresses the HMG box of Cic (see Forés 2017).

pET29b- cic$^{HMG\text{-}C1}$: This construct expresses a C-terminal His-tagged fusion of the Cic $^{HMG\text{-}C1}$ protein (see Forés 2017).

pET17b-Dl: this plasmid encodes the full length form of Dl. To construct the pET17b-Dl plasmid, first the Dl_EN adaptor, that contains the sequences of the Dl C-terminal 2007 to 2034 bp, followed by a NotI site. Subsequently, the NdeI/EcoRV fragment containing the sequences 1-2006 of Dl was obtained from the pAR-Dl plasmid and cloned to the pET17b_Dl_EN plasmid to reconstruct full length Dl.

pET17b-DlΔC: this plasmids encodes for a truncated form of Dl including aminoacids 1-401. For this the NarI/NotI fragment of Dl in the pET17b_Dl vector was substituted by the Dl_ΔC adaptor that creates an in-frame stop codon.

pET29b-DlΔC: this plasmid encodes for a N-terminal His-tag fusion of the Rel homology domain. The Dl coding region from pET17b-Dl was subcloned to the pET29b vector as a NdeI/EcoRV fragment. Subsequently, the C-terminal region (aa 400-670)was deleted by digesting with EheI/EcoRV and religating the vector.

pGEX-RHD: This plasmid encodes a N-terminal GST-tagged fusion of a protein coding for aminoacids 1-387 of Dl (the Rel homology domain) and was kindly provided by Sergio González-Crespo.

pZEX-Hkb: This plasmid encodes a N-terminal GST-tag fusion of the Hkb protein and was kindly provided by Ze´ev Paroush.

pGEX-Hairy: This plasmid encodes a N-terminal GST-tag fusion of the Hairy protein, cloned as described in Paroush 1994.

pET29b-Dri ARID: this plasmid contains the coding sequence for the ARID (DNA-binding domain) of Dead-Ringer (aminoacids of isoform flybase). To generate this fragment, the ARID region was retrotranscribed with primer Dri-1 from RNA extract of wild-type female ovaries, so that the maternally provided cDNA would be obtained. Subsequently, the cDNA was amplified with primers Dri-2 and Dri-3 and cloned in the pUC57-T vector. To discard the in-frame stop codon that was created, the ARID region was reamplified with primers Dri-4 and Dri-5 and was isolated as a BamHI/XhoI fragment, which was then cloned in the pET29b vector.

pAR-twist: this plasmid encodes for the Twi protein and has been kindly provided by Sergio González-Crespo.

Primers used for transgenic and protein expression constructs

primer	5’ to 3’ sequence
zen-1 (forward)	TTTCCGTCCCTTTCTAGAATG
zen-2 (reverse)	GTAAGAATGGATTCAGAGCTTG
zop-1 (forward)	ATGAATTCAATGAATGAAAACAGTATCTG
zop-2 (reverse)	CCGGGATCCGATTGTAAGAATG
zop-3 (reverse)	CTTATATCAATGAAAATAGGGGCAACAAG
zop-4 (forward)	CCTATTTTCATTGATATAAGTTTTGGG
zop-5 (reverse)	CAATATTCATTCATATAGGCCCCG
zop-6 (forward)	CCTATATGAATGAATATTGATTGGGTTTC
zop-7 (reverse)	CAAGATCAATGAAAACTCTATAACTGG
zop-8 (forward)	GAGTTTTCATTGATCTTGGGCGCG
tld-1 (forward)	GTGGAGGATCCTTACAGATG
tld-3 (reverse)	CTGAAGTTGCTTCAACCAGTTC
tldmut-1 (forward)	CATGGATTTTGATTTGGGTATTC
tldmut-2 (reverse)	GAATGCAAACAACAGGCGGCAG
tldmut-3 (forward)	GCGAATCGTTTCGCGGTAATTC
tldmut-4 (reverse)	TCATGCGGCGGATTTACGATC
hb-N2S1 (forward)	GATCCGCCAAACTATATCTGCAGTGCTTATTATCGCTGAG CAGCT
hb-N2S2 (reverse)	CTAGAGCTGCTCAGCGATAATAAGCACTGCAGATATAGTT TGGCG
hb-dor5 (forward)	GATCCAATGGGCCAACGCTCAGCAATCTGCTTAGCTTCGA TAGCT
hb-dor6 (reverse)	CTAGAGCTATCGAAGCTAAGCAGATTGCTGAGCGTTGGCC CATTG
dri-1	GCGGCGACATTTGCTGCATG
dri-2	AGGGTCGCGAAATGGGCGTC
dri-3	GAGGCTTCCAATTCATCGCAG
dri-4	ATGGATCCGATTGAGGCTTC
dri-5	ATCTCGAGGGGTCGCGAAATGGGC

Oligos used to generate CRISPR/Cas9 gRNA's

Mutation generated	Forward oligo	Reverse oligo
cic^{5}	GTCGTGGCCCCAGCTCAAAGTCC	AAACGGACTTTGAGCTGGGGGCCA
cic^{6}	GTCGTAAGCACTGCAGATATAGTT	AAACAACTATATCTGCAGTGCTTA
dl^{28}	GTCGCTAAGCAGATTGCTGAGCGT	AAACACGCTCAGCAATCTGCTTAG

2. RNA antisense probe generation and labelling

To generate antisense RNA probes, we have used plasmids containing a sequence encoding for fragment of the gene transcript and the T7 or T3 promoter, which allows its transcription. The probes were generated with the following procedure:

1. Digestion of the plasmid to obtain a linear template:

Digestion buffer 10X	5 µl
Plasmid (midiprep)	10 µl
Restriction enzyme	2 µl
H_2O	33 µl

Incubate the mix at 37°C for 1-2 h. Run 5 µl on an agarose gel to confirm digestion, including a lane of the non-digested plasmid as a negative control.

Incubate the mix at 37°C for 1-2 h. Run 5 µl on an agarose gel to confirm digestion, including a lane of the non-digested plasmid as a negative control.

2. Extraction and precipitation of the digested plasmid:

Add 55 µl of H_2O and 100 µl chloroform to the digestion mix. Mix well and centrifuge at 11000 rpm for 5 min. Transfer the aqueous solution (upper) and precipitate by adding 10 µl AcNa 3 M and 250 µl absolute EtOH and incubating at -20 °C for 20-20 min. Centrifuge at 13000 rpm for 5 min and discard the supernatant. Wash the pellet with 500 µl EtOH 70% and resuspend in 10 µl H_2O.

4. Transcription

Linearized template (step 2)	5 µl
Transcription buffer 10X	1 µl
T7, or T3 RNA polymerase	1 µl
RNAsin	1 µl
dUTP mix (labelled with digoxigenin, or biotin)	1 µl
H_2O	1 µl

Incubate at 37°C for 2h. After transcription is over, add 17 µl H_2O and run 2 µl of the mix on an agarose gel to confirm transcription.

5. Pausing of transcription and RNA precipitation:

Add 25 µl carbonate buffer 2X and incubate at 65°C for 40min, to produce a controlled degradation of the RNA. Add the following mix to precipitate the RNA:

Stop solution	50 µl
LiCl 4M	10 µl
tRNA(carrier) 20mg/ml	5 µl
EtOH, absolute	300 µl

Incubate at -20 °C for 15 min and centrifuge at 13000 rpm, 4 °C for 15 min.

Eliminate the supernatant and wash the pellet with EtoH 70%.

Resuspend the RNA in 150 µl hybridization solution.

Transcription buffer 10X: Tris-HCl (pH 7,5) 0,4 M, $MgCl_2$ 0,06 M, NaCl 0,1 M, Sperdimin- HCl 0,02 M.

Carbonate buffer 2X: Na_2CO_3 120 mM, $NaHCO_3$ 80 mM. Adjust pH to 10,2

Stop solution: NaAc 0,2M pH 6,0 (adjusted with acetic acid).

3. Genetic analyses

3.1 Fly culture and husbandry

Stocks were maintained at 18°C in vials containing baker´s yeast paste. Mating and rescue experiments were done at 25°C, while experiments using the UAS-Gal4 system were done at 29°C.

3.2 Genome engineering methods

Transgenic fly generation through germline transformation

To study biological pathways, cell processes and gene functions in a determined time point during development, we have modified, overexpressed or misexpressed genes and created reporter genes to study enhancer elements. To achieve this, we have generated transgenic *Drosophila* lines by germline transformation with the P-element and ΦC31 phage integrase systems. In staged preblastoderm embryos (1h after egg laying) that are mutant for the visible marker contained in the vector plasmid (ex if the vector carries the *mini-white*$^{+}$ gene, we inject in w embryos), the following mix was injected:

High quality Vector DNA (midiprep)	6 µg
High quality helper plasmid (midiprep)	2 µg
Injection buffer	2 µl
H_2O	Up to 20 µl

Spin the mix at maximum speed before injecting.

Injection buffer 10X

KCl	5 mM
Phosphatase buffer pH 6,8	0,1 mM
H_2O	Up to 200 mL

Stocks used for germline transformation

Integration	Stock genotype	Position inserted	Selection marker
P-element mediated	w[1118]	Aleatory	*white*
ΦC31 phage integrase	y [1] M{vas-int.Dm}ZH-2A w[*]; M{3xP3-RFP.attP}ZH-86Fb	86B	*white*
ΦC31 phage integrase	y[1] v[1] P{y[+t7.7]=nos-phiC31\int.NLS}X; P{y[+t7.7]=CaryP}attP40	25C6	*vermillion*

Recovery, mapping and establishment of transgenic lines

After the recovery period, first instar larvae were collected and cultured at 25 °C. Eclosed adults (F0) are individually mated with adults of the same mutant strain as the injection embryos and offspring were screened for transformants by eye colour rescue. The transgenic individuals recovered (F1) were crossed again individually with the mutant stain of injection and the heterozygous offspring of each line are then crossed with their heterozygous siblings. Subsequently, homozygous progeny are selected to establish the transgenic line.

Mapping of transgenic lines

For P-element transformation, integration of the transgene is random, therefore, after establishing a homozygous transgenic line, mapping crosses were performed to identify on which chromosome the transgene has landed. Typically, transgenes land on chromosomes X, II or III, since the IV chromosome is rather small and essentially heterochromatic. Once we have obtained male transformants of a given transgenic line carrying the transgene P {*transgene*, w^+}, we have crossed them individually to 3-4 females carrying a balancer chromosome with a visible marker for chromosome II (w; sna^{sco}/SM6A, CyO; +/+) or III (yw; +/+; nmo/TM3,Sb). In the next generation, we selected inidividuals that carry the balancer chromosome and intercrossed them. The portion of the progeny carrying pigmented eyes and balancer chromosome indicated the landing chromosome for the transgene.

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

To introduce mutations to different loci of the *Drosophila* genome and subsequently establish mutant lines, we took advantage of the recently developed CRISPR/Cas9 method, that allows site-specific genome modification (Port et al., 2014). For the means of our studies, we have used the Non-Homologous End Join (NHEJ) repair method to recover mutations in protein coding sequences.

Design and cloning of the gRNA that targets the desired genomic sequence

In the desired coding region, on either DNA strand, a 20 or 21 nucleotide sequence containing a PAM (NGG) sequence was selected. To ensure specificity, we used the BLAST (NCBI) and E-CRISP online tools to screen for off-target sequences in the genome. Then, to express a gRNA that binds with complementarity to the PAM containing strand, we cloned a double strand BbsI fragment to the site of the pCFD3:U6:3 vector (Addgene plasmid 49410).

Generation the transgenic line expressing the gRNA

To generate transgenic lines that stably express the gRNA under the control of the U6:3 promoter it in their somatic and germ line cells, we used the ΦC31-mediated site-specific integration system. The pCFD3:U6:3 plasmid containing the gRNA coding sequence was injected in blastoderm embryos of the strain y^1 v^1 P{nos-phiC31}X; P{CaryP}attP40 (Bloomington stock 25709), (Bischof et al., 2007). G0 hatched flies were crossed individually to the vermillion mutant strain (v^1) and transgenics were identified by eye colour rescue in the F1 generation. To confirm the transgene insertion and establish a homozygous gRNA line, the positive transformants were crossed individually to flies of the genotype v^1; sna^{Sco}/SM6a, a line which was generated in our group.

Mutant generation through NHEJ, screening and mutant line establishment

Once a homozygous line expressing the gRNA was established, it was crossed to a different transgenic line, expressing the Cas9 nuclease, under the promoter of *nanos*: y^2, cho^2 v^1; attP40 {*nos*-Cas9}/CyO. In the next generation, we recovered males that were transheterozygous for the *nos*-Cas9 and the gRNA transgenes. In

the germline of these individuals, different mutations due to the Cas9 induced Double Strand Breaks (DSB) occured. The transheterozygous flies (preferably males, as the recombination rate is higher in their germline) were individually crossed with flies carrying a balancer for the same chromosome the targeted gene is located on. Then, the offspring of this cross were screened for the desired mutation and a transgenic line was established.

To screen the offspring for a CRISPR induced mutations, we have used the following methods:

- Cross the possible CRISPR mutant, that is transheterozygous for a balancer chromosome (+; *mut*/SM6A;+/+ or +; *mut*/TM3Sb), with a known mutation of the same locus (ie. deficiency of the locus) that has known deleterious effects and screen the progeny for rescue of phenotype. For example, possible CRISPR mutant males for the Cic locus were crossed to w; e, cic^{fetE11}/TM6BHu females. A male whose female transheterozygous progeny w; e, cic^{fetE11}/*mut*? is sterile, possibly carries a mutation that cannot rescue the cic^{fetE11} effects. Therefore the w; *mut*?/ TM6BHu progeny of this cross were intercrossed to establish a mutant stock, while the parent was sacrificed and sequenced.

- Once the possible mutant had been crossed to a fly carrying a balancer chromosome, we sacrificed the mutant and do a fly prep. We amplified the targeted region with adequate primers. Optionally, the amplification product was treated with the T7 endonuclease that cleaves double stranded DNA that has mismatches. Finally, the amplification product was sequenced with adequate primers to confirm the mutation. If sequencing revealed the desired mutation, the *mut*?/Balancer progeny were intercrossed to establish the mutant stock.

Generation of Germinal line mutant clones (GLC) with the female sterile dominant technique

To study the maternal function of genes, we have generated females carrying germ-line clones of mutations in the gene of interest, using the Autosomal Dominant Female Sterile (DFS) technique, combined with the Flp/FRT system (Chou and Perrimon, 1996; Chou et al., 1993), that takes advantage of dominant mutations in

the *ovo* gene that block oogenesis on the one hand, and recombination between FRT sequences catalyzed by the Flp recombinase on the other hand. We have selectively produced females that carry clones for a mutation of interest in their germline, allowing us to study the effect of mutations that are lethal when present in the somatic cells. The procedure used to produce GLC is the following:

1. Cross females carrying the mutation of interest (*mut*) and a proximal FRT sequence on the same chromosome arm, with males that carry the ovo^{D1} mutation and an FRT sequence on the homologous chromosome arm of the mutation, as well as an X-linked *hsp70*-Flp transgene. All female progeny of this cross carry the hsp70-Flp transgene and, in addition, some of them are transheterozygous for the P [ovo^{D1}]-FRT chromosome and the *mut*-FRT chromosome.

2. Produce a heat-shock to the progeny of this cross, by incubating them at 37 °C for 30 min-2 h when they are at the second larval stage of development. This produces active Flp recombinase that catalyzes the exchange between the homologous FRT sites of the transheterozygous females.

3. When the heat-shocked larvae eclose, select the transheterozygous females, by lack of balancer chromosome or chromosome with a visible marker. In these females mitotic recombination catalyzed by Flp might have occurred in the germ line. Recombination gives rise to germ-line cells that are homozygous either for the P [ovo^{D1}] mutation (wild-type twin-spot clone), either for the *mut* mutation of interest (homozygous mutant clone). When the heat-shock is done during early development, recombination is very efficient and clones are big, often populating almost the whole germline. The germ line cells in which recombination has not occurred, remain transheterozygous for the P [ovo^{D1}]-FRT chromosome and the *mut*-FRT chromosome, as inherited from their parents. Since the ovo^{D1} mutation blocks oogenesis when present even in one copy, the only eggs laid proceed from germ cells homozygous for the *mut*-FRT chromosome. The somatic follicle cells that surround the oocyte remain transheterozygous for the ovo^{D1} and the mutant chromosome, allowing the study of strictly maternal effects of the given mutation.

In this study, we have generated germline clones of *cic* mutations using w^+; FRT82B, cic^1 and w; FRT82B, cic^5 recombinant lines, and *gro* mutations using the

w^+; FRT82B, gro^{MB41} recombinant lines. For double *cic* and *gro* mutants the like w^+; cic^1, FRT82B, gro^{MB41} line was used (Forés et al., 2015).

UAS-Gal4 ectopic expression

To drive the expression of genes in developmental stages and tissues where it normally doesn't take place, we have used the UAS-Gal4 (Duffy, 2002), a bipartite system that takes advantage of the yeast Gal4 protein, a transcriptional activator, and the UAS sequence (Upstream Activation Sequence) to which Gal4 binds to activate transcription. To apply the UAS/Gal4 system, we have used two different transgenic lines. One of them expresses the Gal4 protein (driver) under the control of an endogenous specific promoter that determines the timing and site of the Gal4 expression. The other carries a genomic construct that contains a cluster of UAS sites placed upstream a gene of interest (target). By crossing the two lines, the Gal4 protein binds to the UAS sites and activates the transcription of the adjacent gene.

For the means of our studies we have used the following Gal4 lines:

Line	Genotype	Bloomington Stock
MVD-Gal4	$w^{[1118]}$;P{w[+mC]=GAL4::VP16 nos.UTR}CG6325[MVD1]	4937
Tub-Gal4	$y^{[1]}$,$w^{[+mC]}$;P{w[+mC]=tubP GAL4}LL7/TM3,$Sb^{[1]}$ $Ser^{[1]}$	5138

To ectopically express the product of the *tsl* gene in the ovary follicle cells, we have crossed females UAS-*tsl*; tsl^{604}/TM6BHu (de las Heras et al., 2006) with Tub-Gal4 males and selected female progeny UAS-tsl/+;TubGal4/TM6BHu, which we cultured at 29 °C. To introduce the $Cic^{\Delta C2HA}$ transgene in this background, we have replaced the tsl^{604} with a chromosome III that has the P*[*w^+*;* $cic^{\Delta C2HA}$*]* and again used the Tub-Gal4 line to drive ubiquitous expression in the follicular epithelium.

RNAi gene knock down in embryos (TRiP)

To target gene expression selectively in a given tissue and timing, we have used the heritable and inducible RNA interference (RNAi) knockdown technique. In this variation of the UAS/Gal4 system, Gal4 binds to the UAS sites to activate the

transcription of a 300-400bp inverted repeat of a coding region of the targeted gene, linked by an intronic region. This inverted repeat gives rise to a double-stranded hairpin RNA (hpRNA). For our studies, we have targeted mRNA expression during early stages of development using the MVD-Gal4 driver , which maternally provides the the RNA. Stocks expressing RNAi's against Dl, Gro and Cact, were obtained by Bloomigton Stock center (TriP project). These stocks are transgenic lines that carry the inverted repeats in synthetic vectors VALIUM 20 and VALIUM 22.

RNAi	Genotype	Bloomington Stock (TriP)
$Cact_i$	$y^{[1]}$ $sc^{[*]}$ $v^{[1]}$; P{y[+t7.7] v[+t1.8]=TRiP.HMS00084} attP2	34775
Dl_i	$y^{[1]}$ $sc^{[*]}$ $v^{[1]}$; P{y[+t7.7] v[+t1.8]=TRiP.GL00676}attP2	38905
Gro_i	$y_{[1]}$ $sc^{[*]}$ $v^{[1]}$; P{y[+t7.7] v[+t1.8]=TRiP.HMS01506} attP2	35759

For gene knock-down, we have crossed the MVD-Gal4 lines with the line expressing the desired RNAi. The trans-heterozygous female progeny was collected and cultured at 29 °C. Note that when knocking down Gro, we have used a recombinant line that carries the gro^{MB36} mutation additionally to the *gro* RNAi.

Other mutant backgrounds

For rescue analysis of *zen,* the recombinant lines w^+; $cic^{1,}$ cic^{Q} w^+; cic^1, cic^{mini} w^+; cic^1cic^{HA} and w^+; $cic^Qcic^{\Delta N60}$ were used.

$cact^-$, cic^- backgrounds were obtained by crossing the w^+; MVDGal4, cic^5/TM3Sb and w;UAS-$cact_i$, cic^5/TM3Sb and culturing the transheterozygous females at 29°C.

For the study of the tld-CRM-lacZ in a mutant cic background, the w^+; tld^{CRM}-*lacZ*, cic^5 line was used.

The *P [Cic^{eh1}],cic^{-}* background was obtained by crossing w;*cic^{1}*, P[w^{+}, *cic*^{eh1}] and *w^{+}; cic^{Q}* flies.

For the double *wntD*, *cic* mutants, the recombinant lines *w^{+}*; *wntD^{KO}, cic^{1}* and *w^{+}*; *wntD^{KO}*, *cic^{2}* were used.

3. Ovary analysis

3.1 Collection and fixation of female ovaries

To analyze female ovaries, newborn to one-day females are collected and placed in a vial with excess of baker's yeast and 2-3 males. The presence of males stimulates the process of egg production.

1. Dissect and clean ovaries in ice-cold PBS 1X.

2. Fix in a solution of PBS-formaldehyde 4%.

3. Do 3 washes with PBS-Tx and disaggregate the ovarioles by pipetting.

4. Store the ovaries in Methanol at -20 °C or continue directly with analysis.

3.2 Immunohistochemistry of fixed ovaries

1. Use approximately 20 µl of fixed ovaries (8-10 females). For immunostaining, use ovaries dissected on the same or previous day.

2. Block unspecific binding by incubating 30 min-1 h with PBS-Tx/BSA at 4 °C, under rotation.

3. Incubate with primary antibody, diluted to the appropriate concentration in PBS-Tx/BSA at 4 °C O/N, under rotation.

4. Perform 3-4 washes of 20 min with PBS-Tx/BSA

5. Incubate with the secondary antibody diluted to the appropriate concentration in PBS-Tx/BSA and incubate under rotation at RT for 1-2 h. Do not exceed this timing, as it can increase background signal. If DAPI is required for nuclei staining, it is added during the last 10-15min of secondary antibody incubation, at a final concentration of 1:50000. The secondary antibodies are coupled to

fluorescent molecules, therefore protection from light is required in subsequent steps.

6. Perform 3-4 washes of 20 min with PBS-Tx under rotation at RT.

7. Mount the embryos in 60 µl Fluromount (Sigma).

Analysis and imaging was done with standard epifluorescence microscopes and confocal SP5 microscopes (Leica Geosystems S.L.).

Fixing solution: PBS 1X, Formaldehyde 4%

PBS-Tx: PBS 1X, Triton-X 0,1%

PBS-Tx/BSA: PBS 1X, Triton-X 0,3%, BSA 1%

4. Embryo analysis

4.1 Preparation of embryonic cuticles

For analyzing cuticles of lethal mutant phenotypes, embryos were collected 3-4 days after being layed (so that they are dehydrated enough to be collected with forceps), dechorionated in 100% bleach for 1 min, washed with distilled water and mounted in Hoyer's solution: lactic acid 1:1. The sample was allowed to dry on a thermal block at 60°C for 16-24 hours.

4.2 Collection and fixation of embryos

For fixation and further analysis, embryos were collected on plates containing a mix of agar and apple juice. For analysis of maternal and early zygotic gene expression, we have collected embryos during 90 min at 25-29 °C and allowed development for 75-90 min at 25°C. After collected, the embryos are dechorionated in bleach 100% for 1min and washed with distilled H_2O. For fixation, embryos are placed in vials containing a solution of 4% Formaldehyde-PBS: Heptane 1:1 for 20 min under strong agitation. Subsequently, the fixed embryos are precipitated in a Heptane: Methanol 1:1 and washed 3-4 times with Methanol 100%. If they were to be used for mRNA analysis, embryos can be stored

at -20°C up to several months, while embryos fixed for protein analysis may be stored at -20 °C up to 24 h.

4.3 *In situ* mRNA hybridization in fixed embryos

In situ hybridization allows us to visualize mRNA expression in fixed tissues, using as a probe a labelled antisense RNA fragment complementary to the mRNA transcribed. After collecting and fixing the embryos, the procedure is the following:

1. Collect and fix staged embryos of the desired genotype. For each hybridization, approximately 10-15 µl of embryos are used.
2. Wash the embryos stored in Methanol 4 times with Ethanol 100%
3. Wash the embryos with Ethanol: Xylenes 1:1
4. Wash the embryos twice with Xylenes 100% and incubate at room temperature for 1-1,5h.
5. Wash with Ethanol: Xylenes 1:1
6. Wash 2 times with Ethanol 100% and 3 times with Methanol 100%.
7. Incubate in PBT-F/M for 5 min.
8. Incubate in fixation solution (PBS-Tween 1%-Formaldehyde 10%) for 20 min.
9. Wash four times with PBT and incubate with PBT containing 0,2 µl Proteinase K for 4 min.
10. Wash four times with PBT and incubate in fixation solution for 20 min.
11. Wash five times with PBT and incubate in PBT: Hybridization solution 1:1 for 1 min.
12. Prehybridization: Wash twice with hybridization solution (80 µl) and incubate at 55 °C for 2 h.
13. Probe preparation: Dilute and denature the RNA probe: For each sample to be hybridized, dilute 0,5µl of the probe in 60 µl of hybridization solution. Heat at

80 °C for 5min, to break down secondary stuctures and cool down by placing 1min on ice.

14. Hybridization: Add the denatured probe solution to the embryos and incubate at 55 °C for 16-20 h.

15. Remove the probe solution and incubate with hybridization solution for 10 min at 55 °C.

16. Incubate with PBT: Hybridization solution 1:1 for 10min at 55 °C.

17. Wash three times with PBT under rotation for 20 min, at room temperature.

18. Incubate with the antibody against the probe: the RNA probes we use are generally synthesized with dUTPs that are labelled with digoxigenin, so as an antibody we use anti-digoxigenin. Dilute the anti-dig antibody in PBT and incubate under rotation at room temperature for 1-2 h. Timing is important, as excessive incubation with antibody may increase background signal.

19. Remove the antibody solution and wash three times with PBS-Tween 1% under rotation for 20 min, at room temperature.

20. Remove the PBT solution and add the development solution

21. Remove the development solution plus NBT and BCIP substrates. These are aromatic compounds that catalyze the colorimetric reaction by the Alkaline Phosphatase that is conjugated to the anti-Digoxigenin antibody. Allow the reaction to take place until the color reaches (several minutes to hours, depending on the probe and the expression of the gene monitored).

22. Stop the colorimetric reaction by adding PBT.

23. Wash for 5-10 min with PBS-EtOH solutions, gradually increasing the amount of Ethanol (30%, 50%, 70%).

24. Wash with Ethanol 100% for approx. 30 min, replacing with fresh Ethanol every 5 min.

25. Wash with Ethanol: Xylenes 1:1.

26. Wash twice with Xylenes 100%.

27. Mount embryos in a microscope slide using Permount medium (ThermoFisher Scientific). About 85-90 µl of Permount are enough to cover a 24 mm x 24 mm coverslip.

PBT: PBS 1X, Tween 20%

PBT/F: PBS 1X, Tween 20%, formaldehyde 4%

PBT-F/M: PBT-F: Methanol 1:1

4.4 Double fluorescent *in situ* mRNA hybridization in fixed embryos

Stage embryos and fix as mentioned above.

1. Wash 5 times with absolute EtOH.

2. Incubate for 1h in Xylene/EtOH 10% at RT.

3. Wash 5 times with EtOH and 5 times with MetOH.

4. Wash with PBT-F/M for 5 min.

5. Post-fixation: Wash 2 times with PBT-F and incubate in PBT-F for 25 min at RT.

6. Incubate in PBT-acetone 80% for 10 min at -20 °C.

7. Wash 5 times with PBT.

8. Post-fixation: Wash 2 times with PBT-F and incubate in PBT-F for 25 min at RT.

9. Wash 5 times with PBT.

10. Incubate in PBT: hybridization solution 1:1 for 10min at RT.

11. Pre-hybridization: Wash 2 times with hybridization solution and incubate for minimum 2 h at 55 °C.

12. Probe preparation: Dilute and denature the RNA probe: For each sample to be hybridized, dilute 0,5 µl of the probe in 60 µl of hybridization solution. Heat at 80 °C for 5min, to break down secondary stuctures and cool down by placing 1min on ice.

13. Hybridization: Add the denatured probe solution to the embryos and incubate at 55 °C for 16-20 h.

14. Wash twice with hybridization solution. Wash with PBT: hybridization solution 1:1.

15. Wash with PBT: hybridization solution 1:1.

16. Wash 2 times with PBT.

17. Blocking: Incubate with PBT-Western Blocking Reagent 20% for 30 min at RT, under rotation.

18. Add primary antibodies, according to the labels of the probes: the probe labelled with digoxigenin is detected with sheep α-digoxigenin and the probe labelled with biotin is detected with mouse α-biotin. Each of the antibodies is used at a final concentration of 1:125 and diluted in PBT-Western Blocking Reagent 20%. Incubate under rotation for 2 h at RT or O/N at 4°C.

19. Perform 2 washes of 15 min with PBT at RT under rotation.

20. Blocking: Incubate with PBT-Western Blocking Reagent 20% for 30 min at RT, under rotation.

21. Incubate with secondary antibodies, diluted in PBT-Western Blocking Reagent 20%, at final concentration of 1:500 each. Sheep α-digoxigenin is detected with α-sheep coupled to $Alexa^{555}$ and mouse α-biotin is detected with α-mouse coupled to $Alexa^{488}$. Incubate for 1-2 h under rotation at RT, protecting from light in subsequent steps. If DAPI is required for nuclei staining, it is added during the last 10-15 min of secondary antibody incubation, at a final concentration of 1:10000.

22. Perform 2 washes of 25 min with PBT, under rotation at RT.

23. Mount the embryos in 60 µl Fluoromount.

Analysis and imaging was done with standard epifluorescence microscopes and confocal SP5 microscopes.

Hybridization solution (50mL)

Formamide	50 %
SCC 20 X	5 X
Salmon sperm	100 µg/ml
Heparin	50 µg/ml
Tween	0,1 %
DI	23 %

4.5 Immunohistochemistry in fixed embryos

To visualize protein expression we have used staged embryos that were fixed as mentioned above. After fixing, embryos are either used directly, or stored at -20 °C for up to 24 h.

1. Wash the embryos with PBS: Methanol 1:1.

2. Wash the embryos with PBS-Tx.

3. Blocking: to avoid non-specific binding of the primary antibody, wash the embryos with PBS-Tx/BSA-2 at 4 °C under rotation for at least 30 min.

4. Add the primary antibody at the appropriate concentration in PBS-Tx/BSA-2 and incubate at 4 °C under rotation O/N.

5. Perform 3-4 washes of 20min with PBS-Tx/BSA-2 under rotation at 4 °C.

6. Add the secondary antibody at the appropriate concentration in PBS-Tx/BSA-2 and incubate under rotation at RT for 1-2 h. Do not exceed this timing, as it can increase background signal. If DAPI is required for nuclei staining, it is added during the last 10-15min of secondary antibody incubation, at a final concentration

of 1:10000. The secondary antibodies are coupled to fluorescent molecules, therefore protection from light is required in subsequent steps.

7. Perform 3-4 washes of 20 min with PBS-Tx under rotation at RT.

8. Mount the embryos in 60 µl Fluromount.

Analysis and imaging was done with standard epifluorescence microscopes and confocal SP5 microscopes (Leica Geosystems, S.L.).

PBS-TX/BSA-2: PBS 1X, Triton-X 0,3%, BSA 2 %

Primary antibody concentrations

α-Dl	1:200
α-beta galactosidase	1:200
α-HA	1:1000

Alexa-Fluor antibodies were used at a standard concentration of 1:500.

4.6 Preabsorbing of primary antibodies

To avoid non-specific signal some primary antibodies were plecleared, by incubating with embryos that are devoid of the protein. For example, α-dig is preabsorbed in w^{III8} embryos.

1. Fix embryos as mentioned above.

2. Wash with PBS 1X: MetOH 1:1.

3. Block by incubating with PBS-Tx/ BSA-2 for 30 min-1 h under rotation at 4 °C.

4. Add primary antibody diluted in PBS-Tx/BSA-2 and incubate at 4 °C under rotation for 8-16 h.

α-digoxigenin	1:200
α-HA	1:100

5. Biochemical *in vitro* assays

5.1 Expression and purification of recombinant proteins

Expression of GST-tag fusion proteins

To express GST tag fusion proteins, the ORF of the full length, or desired domain of the protein, was cloned in frame with the GST coding sequence in the pGEX vector (see constructs). A batch of competent E.coli cells of the BL21 strain is transfected with the pGEX plasmid and grown on a petri dish at 37 °C. Colonies grown on this plate can be used up to 15 days to inoculate. The procedure for expressing and purifying the GST-fusion proteins is the following:

1. Pick a single colony of transformed BL21 cells and inoculate a 3mL LB culture in the presence of 0,2% of ampicillin (6ul) and 0,4% glucose. Grow O/N at 37 °C.
2. Dilute the culture 1:100 in fresh LB, adding antibiotic to a final concentration of 0,2% and grow at 37 °C until OD_{600}=0,6-0,8.
3. Induce protein expression by adding IPTG (Isopropyl-β-D-1-thiogalactopyranoside) at final concentration of 1mM. Incubate at 28-30 °C for 3-4 h under agitation.
4. Centrifuge the culture at 4000 rpm for 10min at 4 °C and resuspend the pellet in 2 ml MTPBS solution with PMSF (phenylmethylsulfonyl fluoride) 1%.
5. Sonicate the cells with 4 pulses of 10 sec and intervals of 15 sec.
6. Add Triton-X 100 to a final concentration of 1% and centrifuge at 13000 rpm for 30 min at 4 °C.
7. Bind the GST fusion protein to the Glutathionine Sepharose™ bead slurry (GE HEalthcare) by adding 75 ul of slurry for each ml of fusion protein and incubating under rotation at 4 °C for at least 15 min. To prepare the bead slurry wash 3-4 times with 5 bed volumes of MTBPS, letting the beads decant and centrifuging for 1 min at 3000 rpm after each wash. After the last wash leave 1:1 bead volume:MTPBS +PMSF 1%. The protein bound to the beads can be stored at 4 °C for up to two weeks.

MTPBS: Na_2HPO_4 1 M, NaH_2PO_4 1M, NaCl 5 M, EDTA 0,5 M, β-mercaptoethanol 0,1%

Expression and purification of His-tag fusion proteins

To express His tag fusion proteins, the ORF of the full length, or desired domain of the protein, was cloned in frame with the GST coding sequence in the pET29b vector (see constructs). A batch of competent E.coli cells of the BL21 strain is transfected with the plasmid and grown on a petri dish at 37 °C. Colonies grown on this plate can be used up to 15 days to inoculate. The procedure for expressing and purifying the GST-fusion proteins is the following:

1. Pick a single colony of transformed BL21 cells and inoculate a 3mL LB culture in the presence of 0,2% kanamycin (6 µl) and 0,4% glucose. Grow O/N at 37 °C.
2. Dilute the culture 1:100 in fresh LB, adding antibiotic to a final concentration of 0,2% and grow at 37 °C until OD_{600}=0,6-0,8.
3. Induce protein expression by adding IPTG at final concentration of 1mM. Incubate at 28-30 °C for 3-4 h under agitation.
4. Centrifuge the culture at 4000 rpm for 10min at 4 °C and resuspend the pellet in 50 ml STE buffer.
5. Centrifuge the culture at 4000 rpm for 10min at 4 °C and resuspend the pellet in 5 ml lysis buffer.
6. Sonicate the cells with 4 pulses of 10 sec and intervals of 15 sec. Centrifuge at 15000 rpm for 10min at 4 °C. Confirm the expression of the intact protein by SDS-PAGE.
7. Bind the His-tag fusion protein to the His-select nickel bead slurry (Signa), by adding 50 µl of slurry for each ml of fusion protein and incubating under rotation at 4 °C for at least 30 min. The protein bound to the beads can be stored at 4 °C for up to two weeks.
8. To purify the protein from the nickel beads, wash twice with 500 µl of wash buffer, and twice with 500 µl of elution buffer.

STE buffer: NaCl 0,1 M, Tris-HCl (pH 8) 10 mM, EDTA 1 mM
Lysis buffer: NaCl 0,5 M, Glycerol 20%, HEPES (pH 7,9) 20 mM, EDTA 1 mM, β-mercaptoethanol 20 mM, PMSF 1 mM, NP40 0,1 %.
Wash buffer: PBS 50 mM, NaCl 300 mM.

Elution buffer: PBS 50 mM, NaCl 300 mM, Imidazol 30 mM

5.2 Protein-protein interaction analyses- GST Protein pulldown assay

To analyse the *in vitro* interaction between proteins, we have used the pulldown method. Briefly, a recombinant protein fused to the GST tag was expressed in E.coli cells and purified by affinity to Glutathionine Sepharose 4B beads. On the other hand, the possible partner was expressed *in vitro,* with the cell free $T_NT^®$ Quick coupled Transcription/Translation System (Promega). Proteins were incubated and interactions were analysed by SDS PAGE and autoradiography.

1. Incubate the GST-fusion protein bound to the Gluthathionine Sepharose beads (bait protein) with 5 ul of the *in vitro* expressed protein of interest (prey) in binding buffer O/N under rotation at 4°C. The prey protein is expressed with the cell free $T_NT^®$ T7 Quick system following the manufacturer´s instructions. Instead of using the Methionine included in the kit, we use 1ul of
2. Methionine that has a radioactive S isotope (^{35}S) (Perkin Elmer).
3. On the next day do 6-7 washes with RIPA buffer letting the beads decant slowly and then centrifuging at 1000rpm for 1 min.
4. After the last wash add sample buffer+ DTT to a portion 1:1 and denature at 100°C for 1min. Centrifuge at maximum speed for 5 min.
5. Resolve the complex by SDS PAGE electrophoresis, loading 20-25 ul of the complex.
6. Transfer the gel to a Whatmann paper and dry at 80 °C for 1 h and expose by autoradiography.

Binding buffer: HEPES-KOH (pH 7,9) 20 mM, KCl 50 mM, $MgCl_2$ 2,5mM, Glycerol 10%, DTT 1 mM, NP40 0,2%, BSA 2%, PMSF 1 mM.

RIPA buffer: Tris-HCl (pH 7,5) 10 mM, NaCl 150 mM, EDTA 1 mM, NP40 0,2%.

5.3 Protein-DNA interactions-EMSA assay

To analyse in vitro interactions between proteins with fragments of promoters, we have used the Electrophoretic Mobility Shift Assay (EMSA). This technique is based on the principle that a protein-DNA complex has different mobility from the free DNA fragment.

Radioactive labelling of DNA by fill-in of 5´-overhangs

This DNA labelling method is based on the activity of the Klenow exo-fragment, an enzyme of E.coli that proceeds from the cleavage of DNA polymerase. This enzyme retains the 5´-3´ polymerase activity and the 3´-5´exonuclease activity, but not the 5 ´-3´ exonuclease activity of DNA polymerase. To take advantage of this activity, we have designed DNA probes with 5´overhangs. These overhangs were present either due to the design of the annealed oligos, either because of the enzymatic cleavage of the NotI site. Labelling was done with α-^{32}P-dCTP (Perkin Elmer), using the illustra™ ProbeQuant™ G-50 Micro Columns kit (GE Healthcare).

Prepare the labelling mix:

Linear DNA (digestion fragment, or annealed oligos)	0,1-4 µg
Buffer BamHI 10X	2 µl
dNTP mix, 2mM each	2,5 µl
α-^{32}P-dCTP	2 µl
Klenow exo fragment	0,5 µl
Nuclease free H_2O	Up to 20 µl

1. Incubate at 30 °C for 15 min.
2. Stop the reaction by adding 30 µl buffer probe type A.
3. Preparation of the column: Once the reaction is stopped, resuspend the resin of the column by tapping. Loosen the lid and replace the closure. Centrifuge at 3000 rcf for 1min. Use immediately, to avoid drying of the resin.
4. Apply the labellig mix on the center of the resin.
5. Centrifuge at 3500 rcf for 2 min.

Electrophoretic Mobility Shift Assay (EMSA)

1. Prepare a native 4% polyacrylamide gel. Wash the wells and pre-run at 4 °C in TBE buffer 0,5X at 120V-20 mA for 30min. This eliminates the excess of Ammonium Persulfate (APS) and unpolymerized acrylic acid from the gel, that could interfere with complex stability.
2. Prepare the binding reaction:

Binding Buffer 5X	4 µl
Poly (dI-dC) BSA 1 µg/µl	1 µl
BSA 1 µg/µl	1 µl
Labelled DNA probe	1-2 ng
In vitro expressed protein (TNT system)	1-3 µl
Nuclease free H_2O	Up to 20 µl

Incubate at RT for 20 min.

1. Add 5 µl EMSA loading buffer to each reaction and load on the gel. For each probe used, include a lane with 2 µl of the free probe. Also include a lane with loading buffer in which Bromophenol blue has been added, so that running of the gel can be monitored. Run the gel at 4 °C at 200V-25 mA.
2. Transfer the gel to a whatmann paper and dry at 80°C for 1h.
3. Expose at -80 °C if an intensifying screen is used, or at RT if the cassette does not have an intensifying screen. Exposure can vary from several hours to several days.
4. Signals were quantified using the Phosphorimager (Molecular Dynamics) equipment.

Binding buffer 5X: HEPES NaOH (pH 7,9) 50 mM, Tris-HCl (pH 7,9) 20 mM, KCl 300 mM, EDTA 5 mM, DTT 5 mM, Glycerol 12%

Loading buffer: TBE 0,25X, Glycerol 34%

BIBLIOGRAPHY

Ables, E. T. (2015). *Drosophila* Oocytes as a model for understanding meiosis: An educational primer to accompany. Corolla is a novel protein that contributes to the architecture of the synaptonemal complex of *Drosophila*. *Genetics* **199**, 17–23.

Adams, C. C. and Workman, J. L. (1995). Binding of disparate transcriptional activators to nucleosomal DNA is inherently cooperative. *Mol. Cell. Biol.* **15**, 1405–21.

Ajuria, L., Nieva, C., Winkler, C., Kuo, D., Samper, N., Andreu, M. J., 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**, 915–24.

Anderson, K. V., Bokla, L. and Nüsslein-Volhard, C. (1985). Establishment of dorsal-ventral polarity in the *Drosophila* embryo: The induction of polarity by the Toll gene product. *Cell* **42**, 791–798.

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

Andreu, M. J., Ajuria, L., Samper, N., González-Pérez, E., Campuzano, S., González-Crespo, S. and Jiménez, G. (2012b). EGFR-Dependent downregulation of capicua and the establishment of drosophila dorsoventral polarity. *Fly (Austin).* **6**, 234–239.

Aronson, B. D., Fisher, A. L., Blechman, K., Caudy, M. and Gergen, J. P. (1997). Groucho-dependent and -independent repression activities of Runt domain proteins. *Mol. Cell. Biol.* **17**, 5581–7.

Ashe, H. L. and Briscoe, J. (2006). The interpretation of morphogen gradients. *Development* **133**, 385–94.

Ashe, H. L., Mannervik, M. and Levine, M. (2000). Dpp signaling thresholds in the dorsal ectoderm of the *Drosophila* embryo. *Development* **127**, 3305–12.

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**, 668–77.

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.

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

Becalska, A.N., Gavis, E.R. Lighting up mRNA localization in *Drosophila* oogenesis. (2009) *Development*. **136**, 2493-503.

Beh, C. Y., El-Sharnouby, S., Chatzipli, A., Russell, S., Choo, S. W. and White, R. (2016). Roles of cofactors and chromatin accessibility in Hox protein target specificity. *Epigenetics Chromatin* **9**, 1.

Belvin, M. P., Jin, Y. and Anderson, K. V. (1995). Cactus protein degradation mediates *Drosophila* dorsal-ventral signaling. *Genes Dev.* **9**, 783–793.

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* **333**, 1453–5.

Biemar, F., Nix, D. A., Piel, J., Peterson, B., Ronshaugen, M., Sementchenko, V., Bell, I., Manak, J. R. and Levine, M. S. (2006). Comprehensive identification of *Drosophila* dorsal-ventral patterning genes using a whole-genome tiling array. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 12763–8.

Bischof, J., Maeda, R. K., Hediger, M., Karch, F. and Basler, K. (2007). An optimized transgenesis system for *Drosophila* using germ-line-specific phiC31 integrases. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 3312–7.

Boettiger, A. N. and Levine, M. (2009). Synchronous and stochastic patterns of gene activation in the *Drosophila* embryo. *Science* **325**, 471–3.

Boija, A. and Mannervik, M. (2016). Initiation of diverse epigenetic states during nuclear programming of the *Drosophila* body plan. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 8735–40.

Brickman, J. M., Adam, M. and Ptashne, M. (1999). Interactions between an HMG-1 protein and members of the Rel family. *Proc Natl Acad Sci U S A* **96**, 10679–10683.

Brown, S. J., Fellers, J. P., Shippy, T. D., Richardson, E. A., Maxwell, M., Stuart, J. J. and Denell, R. E. (2002). Sequence of the *Tribolium castaneum* homeotic complex: the region corresponding to the *Drosophila melanogaster* antennapedia complex. *Genetics* **160**, 1067–74.

Buscarlet, M. and Stifani, S. (2007). The "Marx" of Groucho on development and disease. *Trends Cell Biol.* **17**, 353–361.

Casali, A. and Casanova, J. (2001). The spatial control of Torso RTK activation: a C-terminal fragment of the Trunk protein acts as a signal for Torso receptor in the *Drosophila* embryo. *Development* **128**, 1709–1715.

Casanova, J. (1991). Interaction between torso and dorsal, two elements of different transduction pathways in the *Drosophila* embryo. *Mech. Dev.* **36**, 41–5.

Casanova, J. , Furriols M, McCormick CA, Struhl G, (1995). Similarities between trunk and spätzle, putative extracellular ligands specifying body pattern in Drosophila. *Genes Dev.* **20.** 2539-44

Casanova, J. and Struhl, G. (1989). Localized surface activity of torso, a receptor tyrosine kinase, specifies terminal body pattern in *Drosophila. Genes Dev.* **3**, 2025–2038.

Chambers, M., Turki-Judeh, W., Kim, M. W., Chen, K., Gallaher, S. D. and Courey, A. J. (2017). Mechanisms of Groucho-mediated repression revealed by genome-wide analysis of Groucho binding and activity. *BMC Genomics* **18**, 215.

Chen, G., Handel, K. and Roth, S. (2000). The maternal NF-kappaB/dorsal gradient of *Tribolium castaneum*: dynamics of early dorsoventral patterning in a short-germ beetle. *Development* **127**, 5145–5156.

Chou, T. B. 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[ovoD1] dominant female-sterile insertions in *Drosophila* and their use in generating germ-line chimeras. *Development* **119**, 1359–69.

Clyde, D.E., Corado, M.S., Wu, X., Paré, A., Papatsenko, D., Small, S.(2003).A self-organizing system of repressor gradients establishes segmental complexity in *Drosophila*. *Nature*. **426**, 849-53.

de las Heras, J. M. and Casanova, J. (2006). Spatially distinct downregulation of Capicua repression and tailless activation by the Torso RTK pathway in the *Drosophila* embryo. *Mech. Dev.* **123**, 481–6.

Decoville, M., Giraud-Panis, M. J., Mosrin-Huaman, C., Leng, M. and Locker, D. (2000). HMG boxes of DSP1 protein interact with the rel homology domain of transcription factors. *Nucleic Acids Res.* **28**, 454–62.

DeLotto, R., DeLotto, Y., Steward, R. and Lippincott-Schwartz, J. (2007). Nucleocytoplasmic shuttling mediates the dynamic maintenance of nuclear Dorsal levels during *Drosophila* embryogenesis. *Development* **134**, 4233–4241.

Doe, C. Q., Smouse, D. and Goodman, C. S. (1988). Control of neuronal fate by the *Drosophila* segmentation gene even-skipped. *Nature* **333**, 376–378.

Dorman, J. B., James, K. E., Fraser, S. E., Kiehart, D. P. and Berg, C. a (2004). bullwinkle is required for epithelial morphogenesis during *Drosophila* oogenesis. *Dev. Biol.* **267**, 320–41.

Doyle, H. J., Kraut, R. and Levine, M. (1989). Spatial regulation of zerknullt: a dorsal-ventral patterning gene in *Drosophila. Genes Dev.* **3**, 1518–1533.

Dubnicoff, T., Valentine, S. A., Chen, G., Shi, T., Lengyel, J. A., Paroush, Z. and Courey, A. J. (1997). Conversion of dorsal from an activator to a repressor by the global corepressor Groucho. *Genes Dev.* **11**, 2952–7.

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

Farley, E. K., Olson, K. M., Zhang, W., Brandt, A. J., Rokhsar, D. S. and Levine, M. S. (2015). Suboptimization of developmental enhancers. *Science* **350**, 325–8.

Ferguson, E. L. and Anderson, K. V. (1992). *decapentaplegic* acts as a morphogen to organize dorsal-ventral pattern in the *Drosophila* embryo. *Cell* **71**, 451–461.

Flores-Saaib, R. D. and Courey, a J. (2000). Regulation of dorso/ventral patterning in the *Drosophila* embryo by multiple dorsal-interacting proteins. *Cell Biochem. Biophys.* **33**, 1–17.

Flores-Saaib, R. D., Jia, S. and Courey, a J. (2001). Activation and repression by the C-terminal domain of Dorsal. *Development* **128**, 1869–79.

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**, e1004902.

Forés, M., Simón-Carrasco, L., Ajuria, L., Samper, N., González-Crespo, S., Drosten, M., Barbacid, M. and Jiménez, G. (2017). A new mode of DNA binding distinguishes Capicua from other HMG-box factors and explains its mutation patterns in cancer. *PLoS Genet.* **13**, e1006622.

Fuchs, A., Cheung, L. S., Charbonnier, E., Shvartsman, S. Y. and Pyrowolakis, G. (2012). Transcriptional interpretation of the EGF receptor signaling gradient. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 1572–1577.

Furriols, M. and Casanova, J. (2003). In and out of Torso RTK signalling. *EMBO J.* **22**, 1947–1952.

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**, 11660–5.

Ganguly, A., Jiang, J. and Ip, Y. T. (2005). *Drosophila* WntD is a target and an inhibitor of the Dorsal/Twist/Snail network in the gastrulating embryo. *Development* **132**, 3419–29.

Gilchrist, D. A., Dos Santos, G., Fargo, D. C., Xie, B., Gao, Y., Li, L. and Adelman, K. (2010). Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell* **143**, 540–51.

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.

Goltsev, Y., Fuse, N., Frasch, M., Zinzen, R. P., Lanzaro, G. and Levine, M. (2007). Evolution of the dorsal-ventral patterning network in the mosquito, *Anopheles gambiae*. *Development* **134**, 2415–24.

Gonzalez-Crespo, S. and Levine, M. (1993). Interactions between dorsal and helix-loop-helix proteins initiate the differentiation of the embryonic mesoderm and neuroectoderm in *Drosophila*. *Genes Dev.* **7**, 1703–1713.

González-Reyes, a and St Johnston, D. (1998). Patterning of the follicle cell epithelium along the anterior-posterior axis during *Drosophila* oogenesis. *Development* **125**, 2837–2846.

Gordon, M. D., Dionne, M. S., Schneider, D. S. and Nusse, R. (2005). WntD is a feedback inhibitor of Dorsal/NF-kappaB in *Drosophila* development and immunity. *Nature* **437**, 746–9.

Govind, S., Whalen, A. M. and Steward, R. (1992). *In vivo* self-association of the *Drosophila* Rel-protein Dorsal. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 7861–7865.

Govind, S., Drier, E., Huang, L. H. and Steward, R. (1996). Regulated nuclear import of the *Drosophila* rel protein dorsal: structure-function analysis. *Mol. Cell. Biol.* **16**, 1103–14.

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. *Development* **139**, 3962–8.

Hales KG, Korey CA, Larracuente AM, Roberts DM. (2015) Genetics on the Fly: A Primer on the *Drosophila* Model System. *Genetics*. **3**, 815-42.

Helman A, Cinnamon E, Mezuman S, Hayouka Z, Von Ohlen T, Orian A, Jiménez G, Paroush Z. (2011). Phosphorylation of Groucho mediates RTK feedback inhibition and prolonged pathway target gene expression. *Curr Biol*. **21,** 1102-10.

Helman, A., Lim, B., Andreu, M. J., Kim, Y., Shestkin, T., Lu, H., Jiménez, G., Shvartsman, S. Y. and Paroush, Z. (2012). RTK signaling modulates the Dorsal gradient. *Development* **139**, 3032–9.

Henstridge, M. a, Johnson, T. K., Warr, C. G. and Whisstock, J. C. (2014). Trunk cleavage is essential for *Drosophila* terminal patterning and can occur independently of Torso-like. *Nat. Commun.* **5**, 3419.

Hong, J.-W., Hendrix, D. a, Papatsenko, D. and Levine, M. S. (2008). How the Dorsal gradient works: insights from postgenome technologies. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 20072–20076.

Huang, J. D., Schwyter, D. H., Shirokawa, J. M. and Courey, A. J. (1993). The interplay between multiple enhancer and silencer elements defines the pattern of d*ecapentaplegic* expression. *Genes Dev.* **7**, 694–704..

Huang, J. D., Dubnicoff, T., Liaw, G. J., Bai, Y., Valentine, S. A., Shirokawa, J. M., Lengyel, J. A. and Courey, A. J. (1995). Binding sites for transcription factor

NTF-1/Elf-1 contribute to the ventral repression of *decapentaplegic*. *Genes Dev.* **9**, 3177–89.

Ip, Y. T. (1995). Transcriptional regulation. Converting an activator into a repressor. *Curr. Biol.* **5**, 1–3.

Ip, Y. T., Park, R. E., Kosman, D., Yazdanbakhsh, K. and Levine, M. (1992a). dorsal-twist interactions establish snail expression in the presumptive mesoderm of the *Drosophila* embryo. *Genes Dev.* **6**, 1518–30.

Ip, Y. T., Park, R. E., Kosman, D., Bier, E. and Levine, M. (1992b). The dorsal gradient morphogen regulates stripes of rhomboid expression in the presumptive neuroectoderm of the *Drosophila* embryo. *Genes Dev.* **6**, 1728–39.

Ip, Y. T., Reach, M., Engstrom, Y., Kadalayil, L., Cai, H., González-Crespo, S., Tatei, K. and Levine, M. (1993). Dif, a dorsal-related gene that mediates an immune response in *Drosophila*. *Cell* **75**, 753–63.

Isoda, K., Roth, S. and Nüsslein-Volhard, C. (1992). The functional domains of the *Drosophila* morphogen dorsal: evidence from the analysis of mutants. *Genes Dev.* **6**, 619–30.

Iwahara, J. and Clubb, R. T. (1999). Solution structure of the DNA binding domain from Dead ringer, a sequence-specific AT-rich interaction domain (ARID). *EMBO J.* **18**, 6084–94.

Jennings, B. H. and Ish-Horowicz, D. (2008). The Groucho/TLE/Grg family of transcriptional co-repressors. *Genome Biol.* **9**, 205.

Jennings, B. H., Pickles, L. M., Wainwright, S. M., Roe, S. M., Pearl, L. H. and Ish-Horowicz, D. (2006). Molecular recognition of transcriptional repressor motifs by the WD domain of the Groucho/TLE corepressor. *Mol. Cell* **22**, 645–55.

Jia, S., Flores-Saaib, R. D. and Courey, A. J. (2002). The Dorsal Rel homology domain plays an active role in transcriptional regulation. *Mol. Cell. Biol.* **22**, 5089–99.

Jiang, J. and Levine, M. (1993a). Binding affinities and cooperative interactions with bHLH activators delimit threshold responses to the dorsal gradient morphogen. *Cell* **72**, 741–752.

Jiang, J. and Levine, M. (1993b). Binding affinities and cooperative interactions with bHLH activators delimit threshold responses to the dorsal gradient morphogen. *Cell* **72**, 741–52.

Jiang, J., Rushlow, C. A., Zhou, Q., Small, S. and Levine, M. (1992). Individual dorsal morphogen binding. **1**, 3147–3154.

Jiang, J., Cai, H., Zhou, Q. and Levine, M. (1993a). Conversion of a dorsal-dependent silencer into an enhancer: evidence for dorsal corepressors. *EMBO J.* **12**, 3201–3209.

Jiménez, F., Martin-Morris, L. E., Velasco, L., Chu, H., Sierra, J., Rosen, D. R. and White, K. (1995). *vnd,* a gene required for early neurogenesis of *Drosophila*, encodes a homeodomain protein. *EMBO J.* **14**, 3487–3495.

Jiménez, G., Verrijzer, C. P. and Ish-Horowicz, D. (1999). A conserved motif in goosecoid mediates groucho-dependent repression in *Drosophila* embryos. *Mol. Cell. Biol.* **19**, 2080–7.

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. (2012). The Capicua repressor--a general sensor of RTK signaling in development and disease. *J. Cell Sci.* **125**, 1383–91.

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**, e1005634.

Johnson, T. K., Henstridge, M. A., Herr, A., Moore, K. A., Whisstock, J. C. and Warr, C. G. (2015). Torso-like mediates extracellular accumulation of Furin-cleaved Trunk to pattern the *Drosophila* embryo termini. *Nat. Commun.* **6**, 8759.

Jordan, K. C., Clegg, N. J., Blasi, J. a, Morimoto, a M., Sen, J., Stein, D., McNeill, H., Deng, W. M., Tworoger, M. and Ruohola-Baker, H. (2000). The homeobox gene mirror links EGF signalling to embryonic dorso-ventral axis formation through notch activation. *Nat. Genet.* **24**, 429–433.

Kanodia, J. S., Rikhy, R., Kim, Y., Lund, V. K., DeLotto, R., Lippincott-Schwartz, J. and Shvartsman, S. Y. (2009). Dynamics of the Dorsal morphogen gradient. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 21707–12.

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.

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**, 880–7.

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**, 8272.

Kirov, N., Zhelnin, L., Shah, J. and Rushlow, C. (1993). Conversion of a silencer into an enhancer: evidence for a co-repressor in dorsal-mediated repression in *Drosophila. EMBO J.* **12**, 3193–9.

Kirov, N., Childs, S., O'Connor, M. and Rushlow, C. (1994). The *Drosophila* dorsal morphogen represses the tolloid gene by interacting with a silencer element. *Mol. Cell. Biol.* **14**, 713–22.

Kosman, D., Ip, Y. T., Levine, M. and Arora, K. (1991). Establishment of the mesoderm-neuroectoderm boundary in the *Drosophila* embryo. *Science* **254**, 118–22.

Lacy, M. E. and Hutson, M. S. (2016). Amnioserosa development and function in *Drosophila* embryogenesis: Critical mechanical roles for an extraembryonic tissue. *Dev. Dyn.* **245**, 558–68.

Lagha, M., Bothma, J. P., Esposito, E., Ng, S., Stefanik, L., Tsui, C., Johnston, J., Chen, K., Gilmour, D. S., Zeitlinger, J., et al. (2013). Paused Pol II coordinates tissue morphogenesis in the *Drosophila* embryo. *Cell* **153**, 976–87.

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.

LeBlanc, V. G., Firme, M., Song, J., Chan, S. Y., Lee, M. H., Yip, S., Chittaranjan, S. and Marra, M. A. (2017). Comparative transcriptome analysis of isogenic cell line models and primary cancers links capicua (CIC) loss to activation of the MAPK signalling cascade. *J. Pathol.*

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**, 746–57.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J. M. and Hoffmann, J. A. (1996). The dorsoventral regulatory gene cassette spatzle/Toll/Cactus controls the potent antifungal response in *Drosophila* adults. *Cell* **86**, 973–983.

LeMosy, E. K. (2003). Pattern Formation: The Eggshell Holds the Cue. *Curr. Biol.* **13**, R508–R510.

Levanon, D., Eisenstein, M. and Groner, Y. (1998). Site-directed mutagenesis supports a three-dimensional model of the runt domain. *J. Mol. Biol.* **277**, 509–12.

Liberman, L. M., Reeves, G. T. and Stathopoulos, A. (2009). Quantitative imaging of the Dorsal nuclear gradient reveals limitations to threshold-dependent patterning in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 22317–22.

Lim, J., Crespo-Barreto, J., Jafar-Nejad, P., Bowman, A. B., Richman, R., Hill, D. E., Orr, H. T. and Zoghbi, H. Y. (2008). Opposing effects of polyglutamine expansion on native protein complexes contribute to SCA1. *Nature* **452**, 713–8.

Lynch, J. A. and Roth, S. (2011). The evolution of dorsal-ventral patterning mechanisms in insects. *Genes Dev.* **25**, 107–118.

Mann, R. S., Lelli, K. M. and Joshi, R. (2009). Hox specificity unique roles for cofactors and collaborators. *Curr. Top. Dev. Biol.* **88**, 63–101.

Markstein, M., Markstein, P., Markstein, V. and Levine, M. S. (2002). Genome-wide analysis of clustered Dorsal binding sites identifies putative target genes in the *Drosophila* embryo. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 763–8.

Martin, D., Daulny, A., Decoville, M. and Locker, D. (2003). Mutagenesis analysis of the interaction between the dorsal rel homology domain and HMG boxes of DSP1 protein. *J. Biochem.* **134**, 583–9.

Merabet, S. and Lohmann, I. (2015). Toward a new twist in Hox and TALE DNA-binding specificity. *Dev. Cell* **32**, 259–61.

Merabet, S. and Mann, R. S. (2016). To Be Specific or Not: The Critical Relationship Between Hox And TALE Proteins. *Trends Genet.* **32**, 334–347.

Mineo, A., Furriols, M. and Casanova, J. (2015). Accumulation of the *Drosophila* Torso-like protein at the blastoderm plasma membrane suggests that it translocates from the eggshell. *Development* **142**, 1–6.

Mineo A, Furriols M, Casanova J. (2017). Transfer of Dorsoventral and Terminal Information from the Ovary to the Embryo by a Common Group of Eggshell Proteins in *Drosophila*. *Genetics*. **4**, 1529-1536.

Mirny, L. A (2010). Nucleosome-mediated cooperativity between transcription factors. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 22534–22539.

Morata, G. (1993). Homeotic genes of Drosophila. *Curr. Opin. Genet. Dev.* **3**, 606–14.

Moussian, B. and Roth, S. (2005). Dorsoventral axis formation in the Drosophila embryo--shaping and transducing a morphogen gradient. *Curr. Biol.* **15**, R887-99.

Nunes da Fonseca, R., van der Zee, M. and Roth, S. (2010). Evolution of extracellular Dpp modulators in insects: The roles of tolloid and twisted-gastrulation in dorsoventral patterning of the Tribolium embryo. *Dev. Biol.* **345**, 80–93.

Nusslein-Volhard, C., Frohnhofer, H. and Lehmann, R. (1987). Determination of anteroposterior polarity in *Drosophila*. *Science (80-.).* **238**,.

Nüsslein-Volhard, C. and Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. *Nature* **287**, 795–801.

O'Connor, M. B., Umulis, D., Othmer, H. G. and Blair, S. S. (2006). Shaping BMP morphogen gradients in the *Drosophila* embryo and pupal wing. *Development* **133**, 183–93.

Okimoto, R. A., Breitenbuecher, F., Olivas, V. R., Wu, W., Gini, B., Hofree, M., Asthana, S., Hrustanovic, G., Flanagan, J., Tulpule, A., et al. (2016). Inactivation of Capicua drives cancer metastasis. *Nat. Genet.* 1–14.

Pan, D. and Courey, A. J. (1992). The same dorsal binding site mediates both activation and repression in a context-dependent manner. *EMBO J.* **11**, 1837–42.

Pan, D., Huang, J. D. and Courey, A. J. (1991). Functional analyses of the *Drosophila t*wist promoter reveals a dorsal binding ventral activator region. *Genes Dev.* **5**, 1892–1901.

Papatsenko, D. and Levine, M. (2005). Quantitative analysis of binding motifs mediating diverse spatial readouts of the Dorsal gradient in the *Drosophila* embryo. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 4966–71.

Papatsenko, D., Goltsev, Y. and Levine, M. (2009). Organization of developmental enhancers in the *Drosophila* embryo. *Nucleic Acids Res.* **37**, 5665–77.

Paroush, Z., Finley, R. L., Kidd, T., Wainwright, S. M., Ingham, P. W., Brent, R. and Ish-Horowicz, D. (1994). Groucho is required for Drosophila neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. *Cell* **79**, 805–815.

Patel, N. H., Kornberg, T. B. and Goodman, C. S. (1989). Expression of engrailed during segmentation in grasshopper and crayfish. *Development* **107**, 201–12.

Peri, F., Bökel, C. and Roth, S. (1999). Local Gurken signaling and dynamic MAPK activation during *Drosophila* oogenesis. *Mech. Dev.* **81**, 75–88.

Perrimon, N., Lu, X., Hou, X. S., Hsu, J. C., Melnick, M. B., Chou, T. B. and Perkins, L. A. (1995). Dissection of the Torso signal transduction pathway in *Drosophila*. *Mol Reprod Dev* **42**, 515–522.

Perry, M. W., Cande, J. D. and Boettiger, A. N. (2009). Evolution of Insect Dorsoventral Patterning Mechanisms Evolution of Insect Dorsoventral Patterning Mechanisms. *October* **LXXIV**,.

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**, E2967-2976.

Pridöhl, F., Weißkopf, M., Koniszewski, N., Sulzmaier, A., Uebe, S., Ekici, A. B. and Schoppmeier, M. (2017). Transcriptome sequencing reveals *maelstrom* as a novel target gene of the terminal-system in the red flour beetle *Tribolium castaneum*. *Development* dev.136853.

Queenan, A. M., Barcelo, G., Van Buskirk, C. and Schüpbach, T. (1999). The transmembrane region of Gurken is not required for biological activity, but is necessary for transport to the oocyte membrane in *Drosophila*. *Mech. Dev.* **89**, 35–42.

Rafiqi, A. M., Lemke, S., Ferguson, S., Stauber, M. and Schmidt-Ott, U. (2008). Evolutionary origin of the amnioserosa in cyclorrhaphan flies correlates with spatial and temporal expression changes of zen. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 234–9.

Ramos, A. I. and Barolo, S. (2013). Low-affinity transcription factor binding sites shape morphogen responses and enhancer evolution. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **368**, 20130018.

Ratnaparkhi, G. S., Jia, S. and Courey, A. J. (2006). Uncoupling dorsal-mediated activation from dorsal-mediated repression in the *Drosophila* embryo. *Development* **133**, 4409–14.

Ray, R. P., Arora, K., Nüsslein-Volhard, C. and Gelbart, W. M. (1991). The control of cell fate along the dorsal-ventral axis of the *Drosophila* embryo. *Development* **113**, 35–54.

Reeves, G. and Stathopoulos, A. (2009). Graded dorsal and differential gene regulation in the *Drosophila* embryo. *Cold Spring Harb. Perspect. …* 1–17.

Reichhart, J. M., Georgel, P., Meister, M., Lemaitre, B., Kappler, C. and Hoffmann, J. A. (1993). Expression and nuclear translocation of the rel/NF-kappa B-related morphogen dorsal during the immune response of *Drosophila*. *C. R. Acad. Sci. III.* **316**, 1218–24.

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**, 3023–33.

Rivera-Pomar, R. and Jäckle, H. (1996). From gradients to stripes in *Drosophila* embryogenesis: filling in the gaps. *Trends Genet.* **12**, 478–83.

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

Roth, S. (1994). Proteolytic generation of a morphogen. **4**, 755–757.

Roth, S. (1998). *Drosophila* development: the secrets of delayed induction. *Curr. Biol.* **8**, R906-10.

Roth, S. (2003). The origin of dorsoventral polarity in *Drosophila*. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **358**, 1317–29; discussion 1329.

Roth, S., Stein, D. and Nüsslein-Volhard, C. (1989). A gradient of nuclear localization of the dorsal protein determines dorsoventral pattern in the *Drosophila* embryo. *Cell* **59**, 1189–1202.

Roth, S., Hiromi, Y., Godt, D. and Nüsslein-Volhard, C. (1991). *cactus*, a maternal gene required for proper formation of the dorsoventral morphogen gradient in *Drosophila* embryos. *Development* **112**, 371–388.

Rusch, J. and Levine, M. (1994). Regulation of the dorsal morphogen by the Toll and torso signaling pathways: a receptor tyrosine kinase selectively masks transcriptional repression. *Genes Dev.* **8**, 1247–57.

Rushlow, C. and Levine, M. (1990). Role of the zerknüllt gene in dorsal-ventral pattern formation in *Drosophila. Adv. Genet.* **27**, 277–307.

Rushlow, C. A., Han, K., Manley, J. L. and Levine, M. (1989). The graded distribution of the dorsal morphogen is initiated by selective nuclear transport in *Drosophila*. *Cell* **59**, 1165–1177.

Schmidt-Ott, U. (2000). The amnioserosa is an apomorphic character of cyclorrhaphan flies. *Dev. Genes Evol.* **210**, 373–6.

Schmidt-Ott, U. (2005). Insect serosa: A head line in comparative developmental genetics. *Curr. Biol.* **15**, 245–247.

Schrons H, Knust E, Campos-Ortega JA. (1992)The Enhancer of split complex and adjacent genes in the 96F region of Drosophila melanogaster are required for segregation of neural and epidermal progenitor cells. *Genetics*. Oct;132, 481-503

Shirokawa, J. M. and Courey, a J. (1997). A direct contact between the dorsal rel homology domain and Twist may mediate transcriptional synergy. *Mol. Cell. Biol.* **17**, 3345–55.

Schüpbach, T., Wieschaus, E. (1986). Maternal-effect mutations altering the anterior-posterior pattern of the *Drosophila* embryo. *Roux Arch Dev Biol*. **195**,302-317.

Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I., Zhou, T., Rohs, R., Honig, B., Bussemaker, H. J., et al. (2011). Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. *Cell* **147**, 1270–82.

Stathopoulos, A. and Levine, M. (2002). Dorsal gradient networks in the *Drosophila* embryo. *Dev. Biol.* **246**, 57–67.

Stathopoulos, A. and Levine, M. (2005). Localized repressors delineate the neurogenic ectoderm in the early *Drosophila* embryo. *Dev. Biol.* **280**, 482–493.

Stathopoulos, A., Van Drenth, M., Erives, A., Markstein, M. and Levine, M. (2002). Whole-genome analysis of dorsal-ventral patterning in the Drosophila embryo. *Cell* **111**, 687–701.

Stauber, M., Jäckle, H., Schmidt-Ott, U. and Gehring, W. J. (1999). The anterior determinant bicoid of *Drosophila* is a derived Hox class 3 gene. *Dev. Biol.* **96**, 3786–3789.

Stauber, M., Taubert, H. and Schmidt-Ott, U. (2000). Function of *bicoid* and *hunchback* homologs in the basal cyclorrhaphan fly *Megaselia* (Phoridae). *Proc. Natl. Acad. Sci. U. S. A.* **97**, 10844–9.

Stein, D. S. and Stevens, L. M. (2014). Maternal control of the *Drosophila* dorsal-ventral body axis. *Wiley Interdiscip. Rev. Dev. Biol.* **3**, 301–330.

Stevens, L.M., Beuchle, D., Jurcsak, J., Tong, X., Stein, D. (2003). The *Drosophila* embryonic patterning determinant torsolike is a component of the eggshell. *Curr Biol*. **13**,1058-63.

Steward, R. (1989). Relocalization of the dorsal protein from the cytoplasm to the nucleus correlates with its function. *Cell* **59**, 1179–1188.

Steward, R., Zusman, S. B., Huang, L. H. and Schedl, P. (1988). The dorsal protein is distributed in a gradient in early *Drosophila* embryos. *Cell* **55**, 487–495.

Szymanski, P. and Levine, M. (1995). Multiple modes of dorsal-bHLH transcriptional synergy in the *Drosophila* embryo. *EMBO J.* **14**, 2229–38.

Tatei, K. and Levine, M. (1995). Specificity of Rel-inhibitor interactions in *Drosophila* embryos. *Mol. Cell. Biol.* **15**, 3627–34.

Technau, M., Knispel, M. and Roth, S. (2012). Molecular mechanisms of EGF signaling-dependent regulation of pipe, a gene crucial for dorsoventral axis formation in *Drosophila*. *Dev. Genes Evol.* **222**, 1–17.

Thanos, D. and Maniatis, T. (1992). The high mobility group protein HMG I(Y) is required for NF-kappa B-dependent virus induction of the human IFN-beta gene. *Cell* **71**, 777–89.

Thisse, C., Perrin-Schmitt, F., Stoetzel, C. and Thisse, B. (1991). Sequence-specific transactivation of the *Drosophila twist* gene by the *dorsal* gene product. *Cell* **65**, 1191–1201.

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**, 728–33.

Turki-Judeh, W. and Courey, A. J. (2012). *Groucho. A Corepressor with Instructive Roles in Development*. Curr Top Dev Biol. **98**, 65-96.

Valentine, S. A., Chen, G., Shandala, T., Fernandez, J., Mische, S., Saint, R. and Courey, A. J. (1998). Dorsal-mediated repression requires the formation of a multiprotein repression complex at the ventral silencer. *Mol. Cell. Biol.* **18**, 6584–94.

Vashee, S., Melcher, K., Ding, W. V, Johnston, S. A. and Kodadek, T. (1998). Evidence for two modes of cooperative DNA binding in vivo that do not involve direct protein-protein interactions. *Curr. Biol.* **8**, 452–8.

von Ohlen, T. and Doe, C. Q. (2000). Convergence of dorsal, dpp, and egfr signaling pathways subdivides the *drosophila* neuroectoderm into three dorsal-ventral columns. *Dev. Biol.* **224**, 362–72.

Weiss, J. B., Von Ohlen, T., Mellerick, D. M., Dressler, G., Doe, C. Q. and Scott, M. P. (1998). Dorsoventral patterning in the *Drosophila* central nervous system: the intermediate neuroblasts defective homeobox gene specifies intermediate column identity. *Genes Dev.* **12**, 3591–602.

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**, 10583–8.

Zandvakili, A. and Gebelein, B. (2016). Mechanisms of Specificity for Hox Factor Activity. *J. Dev. Biol.* **4**,.

Zeitlinger, J., Zinzen, R. P., Stark, A., Kellis, M., Zhang, H., Young, R. A. and Levine, M. (2007). Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the *Drosophila* embryo. *Genes Dev.* **21**, 385–90.

Zhang, X. M. and Verdine, G. L. (1999). A small region in HMG I(Y) is critical for cooperation with NF-kappaB on DNA. *J. Biol. Chem.* **274**, 20235–43.

Zhang, Z., Zhu, X., Stevens, L. M. and Stein, D. (2009). Distinct functional specificities are associated with protein isoforms encoded by the *Drosophila* dorsal-ventral patterning gene pipe. *Development* **136**, 2779–89.

Zhu, X., Stevens, L. M. and Stein, D. (2007). Synthesis of the sulfate donor PAPS in either the *Drosophila* germline or somatic follicle cells can support embryonic dorsal-ventral axis formation. *Development* **134**, 1465–1469.

SUMMARY IN SPANISH

INTRODUCCIÓN

El desarrollo de los animales es un proceso de gran complejidad y precisión, mediante el cual células que proceden de la división de una única célula madre, el cigoto, y que son genéticamente idénticas, se diferencian y se asemblan en distintos tejidos y órganos. Este proceso requiere el funcionamiento coordinado de moléculas que constantemente reciben, transmiten e interpretan señales y producen los cambios adecuados dentro de las células, frecuentemente a nivel del genoma, activando o suprimiéndolo.

La señalización celular y el control transcripcional son procesos que controlan tanto funciones fisiológicas, como condiciones patológicas. Durante el desarrollo, las vías de señalización y los factores de transcripción controlan la morfogénesis y el establecimiento de los ejes corporales, pasos necesarios para la formación del nuevo individuo. En animales con asimetría bilateral simple, existen dos ejes de asimetría corporal: el eje anteroposterior (AP) y el eje dorsoventral (DV). Para entender como las vías de señalización y los factores de transcripción controlan el establecimiento de dichos ejes, se han usado varios modelos animales. Uno de los modelos más utilizados para estos estudios genéticos es la mosca de la fruta, *Drosophila melanogaster*. En esta tesis, hemos utilizado *Drosophila* como modelo para entender el funcionamiento molecular de los factores que controlan el establecimiento del eje DV durante las etapas tempranas del desarrollo.

El establecimiento del eje corporal DV de *Drosophila* se controla mayoritariamente por la proteína Dorsal (Dl), un morfógeno que tiene una función doble como factor de transcripción. Cumpliendo la definición de un morfógeno, Dl está presente en los nucleos del embrión temprano de *Drosophila* en forma de un gradiente, siendo la concentración más elevada en los nucleos de la parte ventral y disminuyendo gradualmente hacia las partes más dorsales. Distintas concentraciones de Dl en el nucleo generan respuestas diferentes a nivel genético, estableciendo así un programa de expresión génica distinta en cada tejido del eje DV. Aunque Dl actua en la embriogénesis temprana, las bases para la formación del gradiente nuclear se establecen durante la oogenesis, mediante el sistema dorsoventral materno. A través de este sistema, una señal de la vía EGF (Epidermal Growth Factor) que se activa en el epitelio folicular del ovario y los factores Mirror

(Mirr), Capicua (Cic) y Pipe (Pipe), que responden a esta señal, controlan la entrada graduada de Dl en los nucleos del embrión.

Por otro lado, independientemente de su concentración nuclear, Dl tiene capacidad tanto de activar como de reprimir sus genes dianas. Intrínsicamente, Dl funciona como activador transcripcional, y solo para una pequeña fraccion de sus dianas actua como repressor. Los genes reprimidos por Dl están representados por *zerknüllt* (*zen*), *tolloid* (*tld*) y *decapentaplegic* (*dpp*) y su expresión se restringe a las partes más dorsales del embrión, que darán lugar a un tejido extraembrionario llamado amnioserosa, importante para la formación de la epidermis.

El cambio de modalidad de Dl de activador a represor se ha estudiado exhaustivamente. Se ha visto que depende de la presencia de secuencias ricas en A y T en los promotores de los genes diana, a las cuales se ha sugerido que se unen proteínas represoras que asisten al cambio de modalidad de Dl. Experimentos previos han revelado la identidad de algunos de los factores unidos a las secuencias AT, pero no han sido suficientes para resolver del todo el modelo de la represión por Dl. Por otro lado estudios previos en nuestro grupo han sugerido que un factor adicional que se podría estar uniendo a las secuencias AT es Cic. No obstante, ha sido difícil sacar una conclusión clara , dada la función de Cic en el establecimiento del gradiente de Dl durante la oogenesis. Otros elementos claves de la represión de genes mediante Dl son el corepressor Groucho (Gro) y la vía de señalización RTK (Receptor Tyrosine Kinase) Torso. Gro se ha identificado como elemento necesario para que se forme un represosoma funcional en los promotores de los genes reprimidos por Dl, y, por otro lado, la vía de Torso que se activa en los polos del embrión modula la represión mediada por Dl, permitiendo la expresión de *zen*, *tld* y *dpp* en las partes ventro-terminales.

OBJETIVOS

Dada la importancia de la represión por Dl para el establecimiento del eje DV del embrión de *Drosophila,* como objetivo de esta tesis nos hemos planteado estudiar el mechanismo de dicha represión, centrándonos en el papel de la proteína Cic, y concretamente:

- Estudiar el posible papel de Cic como represor durante el establecimiento del eje corporal DV en la embriogénesis, y relacionar esta función con su ya descrito papel en la polarización DV durante la oogenesis.

- Estudiar la relación entre Cic y los componentes de los represosomas que se forman en las secuencias reguladoras de los genes reprimidos en el eje DV, es decir Dl y Gro.

- Analizar el mecanismo de regulación de la represión dependiente de Dl por la vía de Torso, y el posible papel de Cic como sensor de la vía durante este proceso.

RESULTADOS Y DISCUSIÓN

- Análisis funcional de Cic en la polarización DV: Cic tiene una doble función el establecimiento del eje corporal DV

Como se ha comentado en la introducción, Cic tiene un papel estudiado en la polarización DV del ovario de *Drosophila.* Siendo un componente del sistema dorsoventral materno, Cic controla indirectamente la formación del gradiente de la proteína Dl. Estudios previos han demostrado que mutantes *cic* presentan un fenotipo dorsalizado, debido a esta función de Cic en el ovario. En estos estudios, se ha visto que tanto genes activados por Dl, como genes reprimidos por Dl se desregulan en mutantes *cic*. Por lo tanto, es difícil atribuir a Cic un papel represor durante la embriogénesis, ya que esta función se podría estar emascarando por su función previa durante la oogenesis.

Para aclarar si los fenotipos de desregulación de genes como *zen* en mutantes *cic* se deben a la pérdida de la función de Cic en el ovario, o a una posible función como represor en el embrión, hemos caracterizado varios mutantes Cic analizando la expresión de *pipe* en el ovario como marcador de su actividad oogénica. En el

embrión hemos utilizado *twist* (*twi*) como marcador de la activación mediada por Dl, y *zen* como marcador de la represión mediada por Dl. Hemos observado que, mientras en algunos mutantes de Cic se desregulaba de la expresión de *pipe* en el epitelio folicular, y como consecuéncia desregulación de tanto *twi* como *zen*, había una mutación concreta (cic^1), que no afectaba la expresion de *pipe.* Como sería de esperar, *twi* se activaba normalmente en los mutantes cic^1, a pesar de esto, *zen* estaba expandido hacia las partes ventrales del embrión, es decir, no se estaba reprimiendo. En base a esta observación, hemos generado un mutante nuevo de Cic que anula solo una de las isoformas de la proteína. Hemos confirmado que, efectivamente, esta isoforma está involucrada en la represión mediada por Dl en el embrión, pero no en el sistema materno dorsoventral. En conclusión, nuestro estudio ha demostrado que Cic tiene dos funciones independientes en la polarización DV durante el desarrollo; una durante la oogenesis y otra durante la embriogénesis. Aunque la isoforma que actua durante la oogenesis no está aún caracterizada, es distinta de la que actua durante la embriogénesis.

- Mecanismo molecular de la represión mediada por Cic durante el establecimiento del eje DV del embrión

Una vez encontramos que Cic participa directamente en la represión de genes en el eje DV del embrión, nos hemos planteado caracterizar el mecanismo molecular de esta función. Basándonos en como la misma isoforma embrionaria que está implicada en la represión del eje DV, reprime genes dianas en el eje corporal anteroposterior (AP), hemos buscado una interacción directa entre Cic y las secuecias reguladoras de *zen* y *tld*. Efectivamente, hemos visto una unión directa *in vitro* de Cic a las secuencias AT en los promotores de los genes diana. Dicha unión es específica, pero mucho más débil que la que se ha observado previamente en los genes diana de Cic en el eje AP, debida a la discrepancia que presentan los sitios AT respecto al sitio de unión consenso de Cic. Mediante experimentos funcionales y genéticos hemos observado que, justamente esta pequeña discrepancia (de un nucleotido) entre los sitios AT y el sitio consenso de Cic genera una dependencia de la presencia de Dl para la represión de genes DV. Como consecuencia, los genes como *zen* y *tld* se reprimen solamente en las regiones ventrales, donde tanto Cic como Dl están activos.

Nuestros resultados son importantes porque no solo identifican a Cic como factor necesario para la represión mediada por Dl, si no que muestran cómo una única proteína (Cic) regula simultaneamente genes en dos ejes ortogonales, uniéndose a sitios con alta o baja afinidad. La existencia de sitios ´´óptimos´´ y ´´subóptimos ´´para Cic en sus genes diana es importante para la distinción y el establecimiento diferencial de los dos ejes corporales.

Otro elemento de la represión mediada por Dl es el corepresor Gro. Dado que la forma embrionaria de Cic interacciona con Gro para reprimir sus dianas en el eje AP, nos hemos preguntado si lo mismo ocurre durante la represión de dianas en el eje DV. Mediante la generación de un nuevo alelo mutante de Cic y utilizando alelos mutantes de Gro existentes, hemos comprobado que Cic es el elemento que recluta al corepresor Gro en el promotor de *zen*. Además, utilizando un ensayo para medir la actividad represora *in vivo,* hemos visto que, el motivo de la proteína Cic que recluta a Gro (motivo N2) es un elemento represor autónomo y transferible. En cambio, el motivo de la proteína Dl que hasta ahora se ha pensado que recluta a Gro (motivo eh1-like), es incapaz de actuar como elemento represor en un contexto heterólogo y, además, dispensable para la represión de genes diana en el eje DV.

- Cic es un sensor de la vía de Torso durante el establecimiento del eje corporal DV del embrión

Un aspecto final de la represión mediada por Dl es que se regula por la vía de señalización RTK Torso. Una vez activado el receptor Torso en los polos del embrión temprano, inicia una vía de señalización intracelular, la cual modula las actividades represoras de Dl, permitiendo la expresión de genes como *zen* y *tld* en las partes extremas, a pesar de la presencia de Dl. La transmisión de la señal de la vía de Torso al complejo represor que restringe a genes como *zen* y *tld* se hace a través de un mecanismo desconocido. Se ha propuesto que la activación de Torso permite la expresión de WntD, que es un regulador negativo de Dl. Otra posibilidad es que la activación de la vía inhibe la interacción entre Dl y otros elementos del represosoma. Para explorar estas posibilidades hemos hecho una serie de experimentos genéticos. Utilizando un mutante nulo (knock-out) de *wntD*, hemos visto que no afecta a la represión mediada por Dl. En cambio, en un doble mutante *wntD-cic*, la expresión de *zen* sí se vió afectada.

Entonces, nos hemos centrado en la posibilidad que la vía de Torso afecte la formación del complejo represor inabilitando alguno de sus componentes. Hemos favorecido la idea que el componente inhabilitado podría ser Cic, ya que se sabe que la misma vía lo inactiva durante el establecimiento del eje AP. Sabiendo el mecanismo a través del cual la vía de Torso inactiva a Cic, hemos generado fondos mutantes de activación ectópica de Torso, en los que después hemos introducido formas mutantes de Cic, insensibles a la activación de la vía. Cumpliendo nuestra hipótesis, hemos visto que la derepresión de genes diana como *zen* que ocurre en un fondo de Torso ectópico, se puede impedir haciendo Cic resistente a esta señal. En un fondo donde Torso es ectópico pero Cic insensible, *zen* se reprimió incluso en los polos del embrión, donde normalmente se expresaría. Por lo tanto, concluimos que el elemento afectado del complejo represor de Dl por la vía de Torso, es Cic. Además, la inactivación de otros factores como Dl o Gro, parece no ser relevante en este contexto, ya que solo con convertiendo a Cic en una forma insensible a la vía de Torso es suficiente para generar represión, incluso en nucleos donde estaría aliviada.

En global, nuestros resultados describen un modelo nuevo de represión de genes diana en el eje DV del embrión temprano de *Drosophila*. Según nuestro modelo, un factor clave para la represión de genes específicos de la amnioserosa, como *zen* y *tld,* es el represor Cic. Cic se une a los elementos reguladores de estos genes, a secuencias que se parecen a sus secuencias canónicas de unión. A pesar de esto, una diferencia sutil en la composición de estas secuencias resulta en una unión débil, haciendo la represión dependiente del factor de transcripción Dl, el cual ha sido antes implicado en este proceso. La dependencia de tanto Cic como Dl resulta en un patrón de expresión que se localiza en los nucleos dorsales y terminales, diferenciándolo del patrón de genes que se regulan por sitios de alta afinidad de Cic, y por lo tanto solo se expresan en los nucleos terminales. Los otros aspectos de la represión por Cic, como el reclutamiento de Gro y la respuesta a la via RTK Torso, son iguales en la regulación de los ejes AP y DV. Finalmente, nuestro modelo puede ser la base de investigación biomédica adicional. Se sabe que Cic, que actua como supresor tumoral, en condiciones patológicas se une a secuencias

reguladoras que desvían de sus sitios de unión normales, posiblemente asistido por otros factores.

CONCLUSIONES

- El factor de transcripción Capicua (Cic) tiene un papel represor durante el establecimiento del eje corporal dorsoventral (DV) en la embriogénesis temprana de *Drosophila*. Este papel es independiente de su función el la polarización del epitelio folicular del ovario.
- La isoforma de Cic que actua como represor en el eje DV del embrión es la misma que previamente se había descrito como la isoforma embrionaria principal y que también actua como represor de genes diana en el eje anteroposterior (AP) del embrión. Esta isoforma incluye el motivo aminoterminal N2.
- La función de Cic en el sistema materno dorsoventral durante la oogenesis se lleva a cabo a través de una isoforma distinta de la embrionaria, y no depende de la actividad del motivo N2.
- La isoforma embrionaria de Cic se une directamente a secuencias ricas en AT en los elementos reguladores de los genes *zen* y *tld* con baja afinidad.
- La represión de los genes diana depende de la presencia del morfógeno Dorsal (Dl) y, por lo tanto, ocurre solamente en los nucleos ventrales.
- La optimización de los sitios AT a sitios de alta afinidad de Cic dirije la represión de los genes DV en los nucleos dorsales, dando un patrón de expresión semejante a los genes diana de Cic en el eje AP. Por lo tanto, la unión de Cic a sitios de alta y baja afinidad es un mecanismo de regulación simultánea en el espacio y tiempo, de los dos ejes ortogonales del embrión.
- Cic es el elemento que recluta al corepresor Groucho (Gro) en el complejo represor en el promotor de *zen*.
- La represión mediada por Dl se modula por la activación de la vía de Torso, a través de la inactivación de Cic. Dicha inactivación ocurre de la misma manera como en el eje AP.