UNIVERSITAS
Miguel
Hernández

UNIVERSITAS
Miguel
Hernández

Molecular mechanisms of oncogenesis mediated by the BTB transcription factor Pipsqueak

Impact of sumoylation and proteolytic processing

Memoria de Tesis Doctoral presentada por

Verónica Miguela Fernández

Thesis Directors:

María Domínguez Castellano, PhD

Esther Caparrós Cayuela, PhD

UNIVERSITAS
Miguel
Hernández

INSTITUTO DE NEUROCIENCIAS

Consejo Superior de Investigaciones Científicas
Universidad Miguel Hernández
Instituto de Neurociencias

Prof. Juan Lerma
DIRECTOR

El Profesor Juan Lerma Gómez, Director del Instituto de Neurociencias de Alicante, Centro Mixto de la Universidad Miguel Hernández (UMH) y de la Agencia Estatal Consejo Superior de Investigaciones Científicas (CSIC),

CERTIFICA:

Que la Tesis Doctoral titulada: "*Molecular mechanisms of oncogenesis mediated by the BTB transcription factor Pipsqueak: impact of sumoylation and proteolytic processing*" ha sido realizada por D.ª Verónica Miguela Fernández, bajo la dirección de D.ª María Domínguez Castellano y Dª Esther Caparrós Cayuela, y da su conformidad para que sea presentada a la Comisión de Doctorado de la Universidad Miguel Hernández.

Y para que conste, a los efectos oportunos, firma el presente certificado en San Juan de Alicante, a veintisiete de octubre de dos mil catorce.

Fdo.: Juan Lerma
Director

INSTITUTO DE NEUROCIENCIAS
CSIC

direccion.in@umh.es
director.in@csic.es
www.ina.umh.es

Tel: +34 965-919221 (d)
+34 965-919220 (s)
Fax: +34 965 919561

Av Ramón y Cajal s/n
Campus de San Juan
03550 SAN JUAN DE ALICANTE– ESPAÑA

UNIVERSITAS
Miguel
Hernández

Consejo Superior de Investigaciones Científicas
Universidad Miguel Hernández
Instituto de Neurociencias

María Domínguez Castellano, Ph.D.
Profesor de Investigación del CSIC

INSTITUTO DE NEUROCIENCIAS

Dª María Domínguez Castellano, Profesor de Investigación del Consejo Superior de Investigaciones Científicas,

AUTORIZA la presentación de la Tesis Doctoral titulada: "Molecular mechanisms of oncogenesis mediated by the BTB transcription factor Pipsqueak: impact of sumoylation and proteolytic processing", realizada por Dª Verónica Miguela Fernández, bajo mi inmediata dirección y supervisión en el Instituto de Neurociencias (CSIC-UMH) y que presenta para la obtención del grado de Doctor por la Universidad Miguel Hernández.

Y para que conste, y a los efectos oportunos, firma el presente Certificado en San Juan de Alicante, a veintisiete de Octubre de dos mil catorce.

Fdo. María Domínguez Castellano

m.dominguez@umh.es
www.ina.umh.es

Tel : +34 965 919390
Fax : +34 965 919561

Av Ramón y Cajal, s/n
03550 SAN JUAN DE ALICANTE – ESPAÑA

UNIVERSITAS
Miguel
Hernández

Consejo Superior de Investigaciones Científicas
Universidad Miguel Hernández
Instituto de Neurociencias

Esther Caparrós Cayuela, Ph.D.
Profesora de Inmunología, UMH

INSTITUTO DE NEUROCIENCIAS

D.ª Esther Caparrós Cayuela, Profesora de Inmunología del Departamento de Medicina Clínica de la Universidad Miguel Hernández,

AUTORIZA la presentación de la Tesis Doctoral titulada: "Molecular mechanisms of oncogenesis mediated by the BTB transcription factor Pipsqueak: impact of sumoylation and proteolytic processing", realizada por D.ª Verónica Miguela Fernández, bajo mi inmediata dirección y supervisión en el Instituto de Neurociencias (CSIC-UMH) y que presenta para la obtención del grado de Doctor por la Universidad Miguel Hernández.

Y para que conste, y a los efectos oportunos, firma el presente Certificado en San Juan de Alicante, a veintisiete de Octubre de dos mil catorce.

Fdo.: Esther Caparrós Cayuela

ecaparros@umh.es
www.ina.umh.es

Tel : +34 965 919447
Fax : +34 965 919561

Av Ramón y Cajal, s/n
03550 SAN JUAN DE ALICANTE – ESPAÑA

UNIVERSITAS
Miguel
Hernández

A mis padres

UNIVERSITAS Miguel Hernández

UNIVERSITAS
Miguel
Hernández

AGRADECIMIENTOS

UNIVERSITAS Miguel Hernández

UNIVERSITAS
Miguel
Hernández

Año 2000. Elche. Clase de Biología de cuarto de la ESO. Mientras Esperanza nos bombardea día tras día con una avalancha de conceptos nuevos sobre los seres vivos nace el germen de una idea, una decisión que marca un punto de inflexión en mi vida: voy a dedicarme a la ciencia. Catorce años, una licenciatura y (casi) un doctorado después me encuentro ante otro punto de inflexión: el final del doctorado. Han sido seis años locos de mucho esfuerzo, risas, momentos de éxtasis absoluto, momentos eureka negativos, pero sobre todo de crecimiento, tanto personal como profesional. Y llegados a este punto de la aventura, y como ocurre en cualquier Final Fantasy, toca organizar el inventario, guardar partida antes de la batalla final, y dar las gracias a todos los que han hecho posible esta experiencia.

En primer lugar gracias a María por darme la oportunidad de hacer la tesis con su grupo. Siempre recordaré cuando aterricé en este laboratorio por casualidad con una beca para estudiantes de penúltimo curso de carrera en 2007. La experiencia fue tan buena que me quedé con ganas de más y eso ha desembocado en esta tesis. Gracias también por no escribir en piedra mi hoja de ruta durante estos años y dejarme tomar algunas decisiones sobre cómo llevar los proyectos. La independencia que he disfrutado en este período me ha servido para madurar como profesional y eso, sin duda, será muy beneficioso en futuras etapas.

Gracias a mis compañeros del laboratorio, tanto a los presentes, como a los que ya no están. Esther Caparrós, la “mamá pato”, que me han enseñado todo lo que ha podido y mucho más desde mi primera estancia en el laboratorio. Gracias por ser esa existencia sólida, cálida y radiante que reconforta incluso en las horas más oscuras. Alisson, siempre cercano, dispuesto a contestar a mis preguntas y a discutir sobre resultados durante horas. Gracias por tus consejos y por contagiarme esa pasión infinita tuya por la ciencia, además de ponerle banda sonora a mi doctorado y amenizar los momentos de tedio con paseos por Brasilia con Google Maps jajaja. Irene (Gutiérrez García!), gracias por inspirarme con tu dedicación y rigor en el trabajo, y también por aquellas charlas a las tantas en el labo en las que descubrimos que, entre otras cosas, también nos une el dibujo. Zeus, empezamos juntos y acabamos juntos. Más viejos y cansados, pero más fuertes y sabios. Gracias por ofrecerme siempre tu ayuda y tu punto de vista. ¡También por las inolvidables sesiones matutinas de natación! Irene (Gutiérrez Pérez!), gracias por ser tan buena compi dentro y fuera del labo. Trabajar contigo y ayudarte en tu progreso me ha hecho darme cuenta de muchas cosas que debía mejorar sobre mí misma, así que gracias por ayudarme a ser mejor. Sergio, el tifón que ha revolucionado el labo. Gracias por los momentos de locura transitoria, risas, “movidas y liadas máximas” que hemos compartido en este

tiempo. ¡Nos vemos en el Empire State! Súper Esther Ballesta, ¿qué habría hecho yo sin ti? Gracias por tu ayuda y por un trabajo impecable todos estos años, también por ese carácter alegre y agradable que hace gustosamente soportables las largas horas en el labo. Andrés, gracias por tu paciencia y por templar mis nervios con tu carácter sereno más de una vez. Ha sido un placer trabajar codo con codo contigo, a pesar de los horarios raros jajaja. Diana, ha sido genial tenerte de vuelta en el labo y poder compartir estos últimos años contigo. Gracias por todos tus consejos y apoyo. Nahuel, gracias por esos momentos dignos del Club de la Comedia y por estar siempre dispuesto a compartir tu experiencia y conocimientos con los pequeños de la casa. Tobi, thank you for your support and for listening when I really needed it. Also, it was really nice having someone in the lab sharing my taste in cinema and being able to talk about videogames! Rosa, gracias por hacerme la vida más fácil con los papeleos y por arrancarme siempre una sonrisa. Javier, el que todo lo sabe, gracias por ayudarme a culturizarme con tus tertulias mañaneras. Irene Oliveira, Laura, Pol y Noe, gracias por vuestro trabajo y por contribuir al buen ambiente del grupo. Y también gracias a otros compañeros que ya no están pero que me han ayudado a lo largo de mi tesis: Dolores, Almudena, Vanina, María Cortina, Jesús, Fran y Miguel.

También quiero dar las gracias a otros miembros del Instituto. En primer lugar a Guille y a su grupo, que siempre me han acogido como a una más de la familia y me han apoyado muchísimo todos estos años. Gracias a los compañeros que compartieron conmigo los cursos de doctorado, ha sido un lujo crecer con vosotros. También a otros investigadores y compañeros que me han ayudado durante el transcurso de esta tesis: Ana Carmena y su grupo, por su buena acogida durante mi estancia, Luis, Joan, José Manuel, Sara, Maribel, Stuart, Ruth y Maite.

Y expandiéndome en el eje temporal, me gustaría dar las gracias a toda la gente que ha estado implicada en mi formación y que ha hecho posible que llegue hasta aquí. En primer lugar al grupo de Jose Luís Micol, en especial a Pedro Robles, por darme la oportunidad de saborear por primera vez qué era trabajar en un laboratorio. También al grupo de Antonio Ferrer, con especial mención a Aarón, con vosotros aprendí a no rendirme ante el fracaso y a analizar las cosas con calma para buscar la mejor solución. A mis profesores de Biología del instituto, Esperanza y Josep Pere, gracias por esas magníficas clases que me condujeron a esta senda. Y gracias en especial a Don Jesús, que me enseñó muchas cosas, desde gramática inglesa hasta arquitectura, pasando por el ajedrez o el dibujo, pero sobre todo me enseñó lo que es ser responsable y a respetar mi trabajo y el de los demás.

Fuera ya del ámbito profesional y académico quiero dar las gracias a mis amigos y familiares por su apoyo incondicional y por ser comprensivos con mis locos horarios, el no poder acudir siempre a los eventos o mis temas de conversación extraños e inquietantes.

Amigos, hermanos. Me apena que nos veamos tan poco, pero me encanta que a pesar de ello seamos capaces de recuperar en un parpadeo la química que nos une, como si hubiéramos dejado una película en pausa y la retomásemos unos días después para seguir viéndola. Alicia, Raquel, Vane, Cristian, Ainhoa, Juanpe, Sabri, Alfonso, Rebeca, Noe, Fani, Víctor, Vicky, Lucía, Martín, Cristina, Carlos, Aurora, Vicente, Ainhoa, Isidro, Joserra, Diana, Juanchu, Luis, Mari. Estos últimos años han sido un "locurón" y puedo decir sin temor a equivocarme que no podría haber tenido mejores compañeros de viaje. Gracias en especial a Alicia y Alfonso, que como yo también saben los sacrificios que entraña esto de hacer un doctorado. Y a Noe y a Víctor, que han hecho que la escritura de esta tesis fuera mucho más ligera. ¡Gracias equipo!

Papá, Mamá, gracias por no pararme los pies cuando os dije que quería estudiar Biología jajaja. Gracias por haberme dejado tomar mis propias decisiones y por apoyarme siempre, aunque a veces tuviéramos problemas de entendimiento, tanto por la temática de las conversaciones como por tener opiniones diferentes. Gracias también por vuestro esfuerzo constante para que nunca me faltase de nada y por enseñarme que nada cae del cielo, que todo hay que ganárselo a pulso dando lo mejor de uno mismo. Gracias por fomentar mi curiosidad y enseñarme que los libros pueden ser tus mejores amigos. Gracias por vuestro cariño. Por ser los mejores ejemplos. En definitiva, gracias por todo.

Y finalmente, Edu. Mi rival, amigo y compañero en este periplo llamado vida. Gracias porque, aunque suene a cliché, sin tu apoyo constante esta tesis no habría sido posible. Y no solo la tesis, llegar a tener la oportunidad de hacerla también. Gracias por las largas horas de discusiones sobre ciencia y resultados, por apaciguar mi ira y consolar mi pena, por compartir mi euforia, por ser siempre honesto y decirme cuando me he equivocado. Te admiro como persona y como profesional y te doy las gracias por inspirarme cada día para que de lo mejor de mí, para llegar más lejos, para ser mejor. ¿Quién nos iba a decir que estaríamos hoy aquí mientras escuchábamos a Esperanza en clase de Biología? ¿Imaginas dónde estaremos en otros catorce años?

UNIVERSITAS Miguel Hernández

INDEX

UNIVERSITAS Miguel Hernández

UNIVERSITAS Miguel Hernández

UNIVERSITAS Miguel Hernández

Abbreviations

UNIVERSITAS Miguel Hernández

UNIVERSITAS
Miguel
Hernández

3-AT	3-aminotriazol
aa	Aminoacid
AD	Activation domain
BCL-6	B cell lymphoma 6
BDSC	Bloomington Drosophila Stock Center (Indiana University)
BTB	Broad-Complex, Tramtrack, and Bric à brac
cDNA	Complementary DNA
Ci	Cubitus interruptus
DAPI	4',6-diamidino-2-phenylindole
DBD	DNA binding domain
DGRC	*Drosophila* Genetic Resource Center
Dl	Delta
DSHB	Developmental Studies Hybridoma Bank
ECD	Extracellular domain
GS	Gene search
HDAC	Histone deacetylase
HRP	Horseradish peroxidase
HTH	Helix-turn-helix
ICD	Intracellular domain
kDa	kiloDalton
MEP-1	MOG interacting and ectopic P-granules
MF	Morphogenetic furrow
NuRD	Nucleosome Remodelling and Deacetylation
PBS	Predicted biological score
Pc-G	Polycomb-group
PCR	Polymerase chain reaction
PH3	Phospho-histone 3
PLZF	Promyelocityc leukaemia zinc finger
POZ	Poxvirus and zinc finger
Psq	Pipsqueak
qPCR	Quantitative polymerase chain reaction
SEM	Standard error of the mean
SID	Smallest interaction domain
SIM	Sumo interacting motif
STUbLs	SUMO-targeted ubiquitin ligases
SUMO	Small ubiquitin-related modifier
UAS	Upstream activating sequence

Ubls	Ubiquitin-like proteins
VDRC	Vienna *Drosophila* RNAi Center
WB	Western blot
Y2H	Yeast two-hybrid

UNIVERSITAS Miguel Hernández

Summary

UNIVERSITAS Miguel Hernández

UNIVERSITAS
Miguel
Hernández

Cancer is a cellular phenomenon that occurs within the context of normal tissues and a normal microenvironment within the organism. Nowadays one of the major challenges regarding cancer is to understand how the oncogenic mechanisms confer malignant properties to the cells by altering the regulation of physiological tissue growth, the cessation of growth, cell differentiation, apoptosis and hormonal control. Over the past years, many works have linked the activity of numerous BTB transcription factors, such as PLZF or BCL-6, with human cancer. However, the molecular mechanisms that regulate their activity during oncogenesis remain poorly understood. In this Thesis I have used the *Drosophila* BTB transcription factor Pipsqueak as a paradigm to unravel new mechanisms of tumorigenesis mediated by BTB transcription factors. More specifically, I have focused on how sumoylation and proteolytic processing modulate Pipsqueak activity during normal development and oncogenesis. On the one hand, we have demonstrated that Pipsqueak sumoylation at Lysine 633 attenuates its transcriptional activity and its pro-apoptotic activity. The results presented in this Thesis suggest that sumoylation mediates the interaction of Pipsqueak with distinct proteins, such as the transcriptional repressor MEP-1 or the component of the basal transcriptional machinery DmTAF3, which are able to modulate the activity of this BTB transcription factor. On the other hand, we have found that the long PipsqueakB isoform is subjected to two mutually exclusive proteolytic events. We show that the two types of processing occur in different compartments of the cell, and specifically Type I processing is necessary for Pipsqueak transcriptional activity. We also demonstrate the responsible of this type of processing is the proteasome. The results of both studies offer new hints that will help us to better understand the complex molecular mechanisms underlying the regulation of the activity of BTB transcription factors.

UNIVERSITAS Miguel Hernández

El cáncer es un fenómeno celular que ocurre en el contexto de tejidos normales y en el microentorno normal de un organismo. En la actualidad, uno de los mayores retos relacionados con el cáncer es entender cómo los mecanismos oncogénicos confieren propiedades malignas a las células a través de alteraciones en la regulación del crecimiento fisiológico de tejidos o del cese de ese crecimiento, diferenciación celular, apoptosis y control hormonal. En los últimos años, varios trabajos han vinculado la actividad de numerosos factores de transcripción de tipo BTB, como PLZF o BCL-6, con cáncer en humanos. Sin embargo, los mecanismos moleculares que regulan su actividad durante la oncogénesis siguen siendo poco conocidos. En esta Tesis he utilizado el factor de transcripción de tipo BTB de *Drosophila* Pipsqueak como paradigma para desentrañar nuevos mecanismos de tumorogénesis mediados por factores de transcripción de tipo BTB. Más específicamente, me he centrado en cómo la sumoilización y el procesamiento proteolítico modulan la actividad de Pipsqueak durante el desarrollo normal y la oncogénesis. Por un lado, hemos demostrado que la sumoilización de Pipsqueak en la Lisina 633 atenúa su actividad transcripcional y su actividad proapoptótica. Los resultados presentados en esta Tesis sugieren que la sumoilización media la interacción de Pipsqueak con distintas proteínas, como por ejemplo el represor de la transcripción dependiente de sumoilización MEP-1 o el componente de la maquinaria basal de transcripción DmTAF3, que son capaces de modular la actividad de este factor de transcripción de tipo BTB. Por otro lado, hemos descubierto que la isoforma larga de Pipsqueak conocida como PipsqueakB es objeto de dos eventos proteolíticos mutuamente excluyentes. En este trabajo mostramos que los dos tipos de procesamiento ocurren en diferentes compartimentos de la célula y que específicamente el procesamiento Tipo I es necesario para la actividad transcripcional de Pipsqueak. También hemos demostrado que el responsable de este tipo de procesamiento es el proteasoma. Los resultados de ambos estudios ofrecen nuevas pistas que nos ayudarán a entender mejor los complejos mecanismos moleculares que subyacen a la regulación de la actividad de los factores de transcripción de tipo BTB.

UNIVERSITAS Miguel Hernández

I. INTRODUCTION

UNIVERSITAS Miguel Hernández

UNIVERSITAS
Miguel
Hernández

Despite the remarkable advances of several decades of research, cancer still remains a pressing health concern. Cancer is a cellular phenomenon that occurs within the context of normal tissues and a normal microenvironment within the organism. As such, one of the major challenges we now face is to understand how the oncogenic mechanisms that confer malignant properties to cells are integrated with, and perturb, the cellular process that regulate physiological tissue growth, the cessation of growth, cell differentiation, apoptosis and hormonal control. In this Thesis I have used the fruit fly *Drosophila melanogaster* to identify genes and mechanisms involved in cancer development. Studies in flies can provide valuable information that could guide us towards a better interpretation and understanding of how oncogenic mutations orchestrate proliferation and migratory behaviour.

1. Development, growth-promoting organizers and cancer

During development, growth control is repeatedly linked to the specification of specialized regions called organizers (Diaz-Benjumea and Cohen, 1995; Irvine and Rauskolb, 2001; Lawrence, 1992), which were first defined at the start of the last century by Spemann and Mangold, in the *Xenopus laevis* embryo. Organizers control embryonic development through the formation of boundaries or compartments that express morphogenes that instruct at long range and in a non-autonomous way, not only cellular specification (as was defined by Spemann and Mangold), but also the patterning of the organ or tissue (Irvine, 1999). For example, a well-studied organizer for growth control is the dorso/ventral organizer in the eye of *Drosophila* (Cho and Choi, 1998; Dominguez and de Celis, 1998; Papayannopoulos *et al.*, 1998), which is mechanistically similar to those in the wing imaginal disc and the apical ectodermal ridge (AER) of the primordium in vertebrate's limbs. Our understanding of how these organizers instruct developing organs when to stop growing is still fragmentary (reviewed in Dominguez, 2014) and the study of tumorigenesis in flies may broaden our knowledge about this important topic.

Several studies in diverse vertebrate and invertebrate models have led to the identification of many of the organizer signals, for example EGF, Wnt (also known as Wingless in *Drosophila*), Hedgehog and Notch. Recently, clinical evidences have involved the aberrant activation of many of these organizing signals in the initiation and progression of numerous types of human cancers. These observations have revived the interest in the association between the mechanisms that regulate the formation and

function of these growth organizers and the processes that promote cancer (Steelman *et al.*, 2008; Vogelstein and Kinzler, 2004).

2. Notch signalling pathway

The Notch signalling pathway is emerging as one of the most potent oncogenic pathways in humans. This evolutionary conserved signalling pathway (Artavanis-Tsakonas *et al.*, 1999; Mumm and Kopan, 2000), was first identified in *Drosophila* and in *Caenorhabditis elegans* nearly a century ago, when the inactivation of just one copy of Notch (haploinsufficiency) was shown to produce notches at the wing margin in flies (Moohr, 1919) and defects in vulval development in worms (Ferguson and Horvitz, 1985; Greenwald *et al.*, 1983). Posteriorly, classic embryonic analyses of lethal loss-of-function mutations revealed a key role of Notch in neurogenesis (Poulson, 1937). Additional studies showed that this function is evolutionary conserved and probably underlies many neurologic alterations associated with mutations in the NOTCH1-4 genes in human. Finally, during the late '90s and early 2000s the role of Notch in growth was uncovered and is now firmly linked to oncogenesis (reviewed in Dominguez and Casares, 2005). Notch has pleiotropic roles in flies and humans, and as such it is not surprising that a tumour suppressor role for Notch has also been deduced in some mice models (reviewed in Radtke and Raj, 2003). In this work, we have focused on Notch as an oncogene.

2.1 Overview of Notch signalling cascade

The mammalian Notch receptor family consists of four type I transmembrane receptors (termed NOTCH1–4) (**Figure 1A**), all of which have been implicated in human cancer. There are five known Notch-receptor ligands, which are collectively referred to as DSL proteins (DELTA-like 1, 3 and 4, and Serrate/JAGGED 1 and 2) (**Figure 1B**). These ligands possess redundant and overlapping functions, thus complicating the study of Notch in vertebrate animal models. The genome of *Drosophila* encodes only one Notch receptor (Kidd *et al.*, 1986; Wharton *et al.*, 1985a) and two ligands, Delta (Dl) and

Serrate (Ser) (Haltiwanger, 2002) (**Figure 1A** and **B**), considerably simplifying the study of Notch signalling functions *in vivo*.

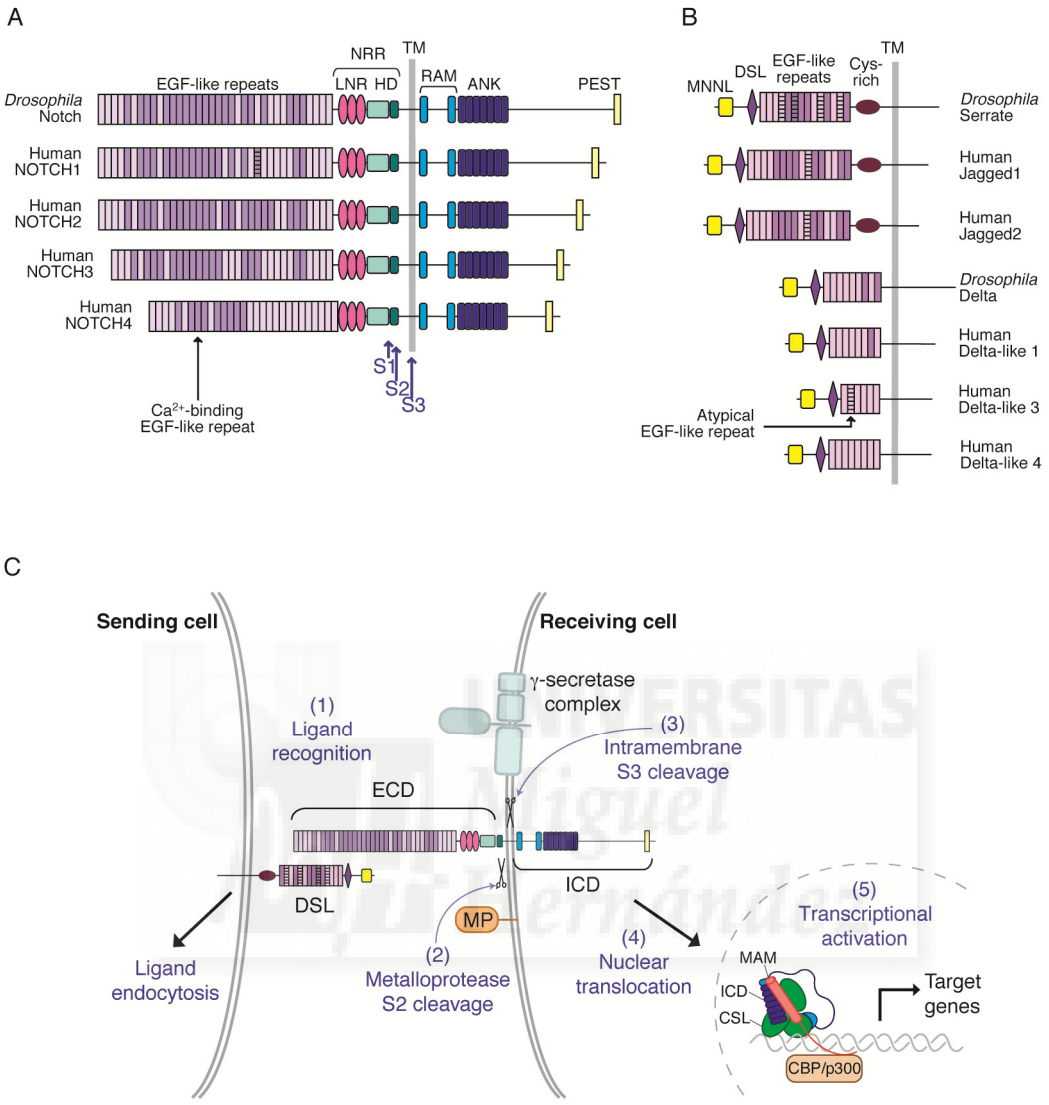


Figure 1. Notch signalling pathway and domain organization of Notch receptors and DSL ligands. Domain organization of Notch receptors (**A**) and DSL-family ligands (**B**) from fly and human. (**C**) Model for the major events in the Notch signalling pathway. Signals initiated by the engagement of ligand (1) lead to metalloprotease (MP) cleavage at site S2 (2). This proteolytic step allows the cleavage of Notch by the γ-secretase complex at site S3 within the transmembrane domain (3) and release of Notch intracellular domain (ICD) from the membrane. ICD translocates to the nucleus (4), where it forms a transcriptional activation complex with CSL and MAM, inducing the transcription of target genes (5). TM, transmembrane domain; MP, metalloprotease; ECD, Notch extracellular domain; ICD, Notch intracellular domain. Adapted from Gordon *et al*. (2008).

Notch receptors are expressed in the cellular surface as heterodimeric proteins. Notch is synthesized as a precursor form that is cleaved by a furin-like convertase (S1 cleavage) to generate the mature receptor, which is composed of an extracellular

domain (ECD), and a intracellular domain (ICD) that are held together by non-covalent interactions. The ECDs of Notch proteins are comprised of 36 epidermal growth factor (EGF)-like repeats that have a role in ligand–receptor interactions. Carboxy-terminal to the EGF-like repeats are three cysteine-rich LIN-12 and Notch repeats (LNRs), which prevent ligand independent signalling, and a C-terminal hydrophobic region that mediates the interaction between the ECDs and the transmembrane domains. The Notch ICD, which is composed of conserved protein domains, such as the ankyrin repeats and the PEST domain, is the active form of the protein and mediates Notch signalling (reviewed in Aster *et al.*, 2008; Gordon *et al.*, 2008).

The DSL ligands, like the Notch receptors, are type I transmembrane proteins. Upon binding to the Notch receptor (**Figure 1C**), the ligand induces a conformational change, exposing the S2 cleavage site in the ECD to the ADAM17/TACE metalloprotease (Brou *et al.*, 2000; Lieber *et al.*, 2002). Following S2 cleavage, truncated Notch undergoes a third cleavage (S3 cleavage) that is mediated by the presenilin–γ-secretase complex, which is composed of presenilin 1 (PSEN1), PSEN2, nicastrin (NCSTN), presenilin enhancer 2 (PEN2) and anterior pharynx-defective 1 (APH1). The S3 cleavage results in the release of the active Notch ICD from the plasma membrane (Brou *et al.*, 2000; De Strooper *et al.*, 1999; Struhl and Basler, 1993) and its subsequent translocation into the nucleus (Kopan and Ilagan, 2009). The major downstream effector of Notch signalling once it reaches the nucleus is the transcription factor CSL (for mammalian CBF1, *Drosophila* Suppressor of Hairless and LAG-1 in *C. elegans*). Notch translocation mediates the conversion of this repressor complex into a transcriptional activation complex, partially through replacement of co-repressors and histone deacetilase (HDAC) complexes and the recruitment of co-activator complexes including proteins of the Mastermind (MAM) family (Artavanis-Tsakonas *et al.*, 1999; Bray, 2006; Kovall, 2008; Lai, 2004; Le Borgne *et al.*, 2005). Activation of transcription at CSL-binding sites also appears to depend on the recruitment of additional co-activators, such as p300 (E1A binding protein p300; EP300) or CREB-binding protein (CBP) (Fryer *et al.*, 2002; Wallberg *et al.*, 2002); these co-activators might constitute the main link between the core Notch-containing complex and the general transcription machinery. Notch signalling is thought to exert its pleiotropic effects by initiating a transcriptional cascade that involves both the activation and the repression of target genes, including transcriptional regulation by epigenetic mechanisms. Although the details of such a transcriptional cascade are not completely understood, several well-characterized target genes have been described. Among these genes are the basic-helix–loop–helix (bHLH) transcriptional repressors

Hairy and Enhancer of split (HES) family, the hairy-related transcription factor (HRT; also known as HEY) family, Notch receptors, Notch ligands, cyclin D1 (CCND1) and MYC. Notch transcriptional activity is terminated by phosphorylation of Notch on the C-terminal PEST domain, which targets it for ubiquitination by ubiquitin ligases and subsequent degradation by the proteasome (reviewed in Fortini, 2009).

2.2 Role of Notch in tumorigenesis

In *Drosophila* and vertebrates, Notch signalling pathway is one of the more broadly used signal transduction pathways in growth control, as well as cell fate specification, differentiation, migration and apoptosis, in many different types of tissues through several developmental stages and in the adult stage (Rizzo *et al.*, 2008) (**Table 1**). Mutations, deletions, translocations and/or viral insertions affecting Notch signalling pathway components are involved in a long list of human diseases, including cancer.

A link between Notch and cancer was first established in 1991 (Ellisen *et al.*, 1991) when an aberrant activation of NOTCH1 was found to occur in rare cases of paediatric and adult T-cell acute lymphoblastic leukaemia (T-ALL). Posteriorly, the group of Jon Aster reported the finding of recurrent truncated, active forms of the NOTCH1 receptor mutations in T-ALL (Aster, 2005; Grabher *et al.*, 2006; Weng *et al.*, 2004). Subsequent studies have significantly expanded our understanding of how Notch participates in cancer, and defined some of its partners in different cancer processes. However, strong activating mutations in the Notch receptor are rare, even in liquid tumours; while in solid tumours, the inappropriate Notch activation arises from overexpression of Notch receptor or ligands, or through mutations in pathway regulators. Thus, a more likely view is that, in human cancers, Notch activation cooperates with other oncogenes or the loss of tumour suppressors to initiate a tumour (reviewed in Dominguez, 2014).

Aberrant activation of Notch signalling pathway has now been associated with numerous solid tumours, including nervous system tumours, melanoma, breast cancer, prostate, lung, cervix, etc. (Miele *et al.*, 2006). In particular, high or aberrant expression of Notch ligands (mainly JAGGED1) is associated with poor prognosis in breast and prostate cancer (Karanu *et al.*, 2000; Leong *et al.*, 2007; Santagata *et al.*, 2004; Zhang *et al.*, 2006). Studies in *Drosophila* have furthered our understanding of how Notch initiates and promotes the progression of tumorigenesis *in vivo*. Thus, it is now known

that Notch induces tumours in cooperation with epigenetic regulators (Ferres-Marco *et al.*, 2006; Liefke *et al.*, 2010), the FGF-R and EGF-R (Hurlbut *et al.*, 2007), the survival pathway Pi3K/AKT (Palomero *et al.*, 2007), or the Hedgehog signalling pathway and the conserved microRNA mir-7 (Da Ros *et al.*, 2013). Despite these advances, the knowledge about the mechanisms by which these oncogenic pathways contribute to the initiation and progression of invasion and metastasis is still limited.

Vertebrates	*Drosophila*	*C. elegans*
Inhibition of neurogenesis	Inhibition of neurogenesis	Regulation of early blastomere specification
Regulation of fate choices in the inner ear	Regulation of gliogenesis, neural lineage fates	Regulation of AC/VU decision
Inhibition of non-neural ectodermal derivates (*Xenopus* ciliated cells, chick feather buds)	Inhibition of wing venation	Regulation of vulval precursor fates
Inhibition of myogenesis, cardiogenesis	Inhibition of myogenesis, cardiogenesis	Induction of left-right asymmetry
Induction of left-right asymmetry	Inhibition of midgut precursors	Induction of germline proliferation
Regulation of limb bud development	Induction of mesectoderm	
Regulation of somitogenesis	Induction of wing margin	
Regulation of lymphopoiesis	Induction of leg segments	
Regulation of vascular development	Induction of dorso/ventral eye polarity	
Regulation of kidney development	Induction of cone cells in the eye	
	Regulation of hematopoiesis	

Table 1. Examples of Notch signalling roles during development in different species. A Non-exhaustive list of developmental processes regulated by Notch signalling in different species. Adapted from Lai (2004).

2.3 Notch and the development of the eye in *Drosophila*

Notch signalling pathway plays pleiotropic roles during the development of the fly retina (Voas and Rebay, 2004), including first the promotion of retinal precursor cell proliferation and subsequently, as cells transit from proliferation to a commitment state, Notch signalling is required for retinal cell fate specification, planar cell polarity and apoptosis.

Drosophila eye development has emerged as a powerful paradigm for *in vivo* genetic and epigenetic studies of cancer (Bier, 2005; Dominguez, 2014; Potter *et al.*,

2000). The eye is a non-vital organ and hence its complete absence or the formation of tumours does not cause lethality or prevent fertility. Therefore, adult flies with tumours can be used for further genetic analysis. Moreover, our laboratory has been the first to establish that metastasis derived from the eye can be easily scored in adult flies by using the conditional ON-OFF *eyeless-Gal4* driver (see below). Thus, the studies of tumorigenesis and the cellular interactions between tumour cells and their microenvironment are greatly facilitated in this tissue. In this Thesis we have made use of the powerful tools of *Drosophila* genetics and the advantages of the fly eye as an experimental model to unravel mechanisms of oncogenesis mediated by the BTB-containing oncogene *pipsqueak*, thus extending our previous knowledge (Ferres-Marco *et al.*, 2006).

2.3.1 Drosophila melanogaster *life cycle, imaginal discs and retinal development*

Compared with other animal models of cancer, *Drosophila* has a relative short life cycle (approximately of 11 days at 25 ºC), which consists of embryonic, larval, pupa and adult stages (**Figure 2A**). During the three larval stages, the imaginal discs grow exponentially, increasing their size approximately 200 fold. Unlike the imaginal discs and the brain, which increase their size by mitotic cell division, the larval tissues (epidermis, salivary glands, and fat body) grow rapidly by endoduplication without cellular division. Through the larval stages, the epidermis is moulted two times, being growth completely arrested during that process. During metamorphosis, the imaginal discs stop proliferating, and start to differentiate and fuse to give rise to the adult fly structures (the head, eye, antenna, thorax, wings, legs and genitalia) (**Figure 2B**), while the majority of the larval tissues undergo histolysis, and the brain undergoes a profound remodelling to prepare for adult tasks (flying, reproduction, vision, etc.). The fly abdomen is structured from nests of histoblasts, which start proliferation during the pupal stage (Madhavan and Madhavan, 1980). At the end of the metamorphosis the adult fly hatches from the pupa and in 2-3 days is competent for reproduction (Lawrence, 1992).

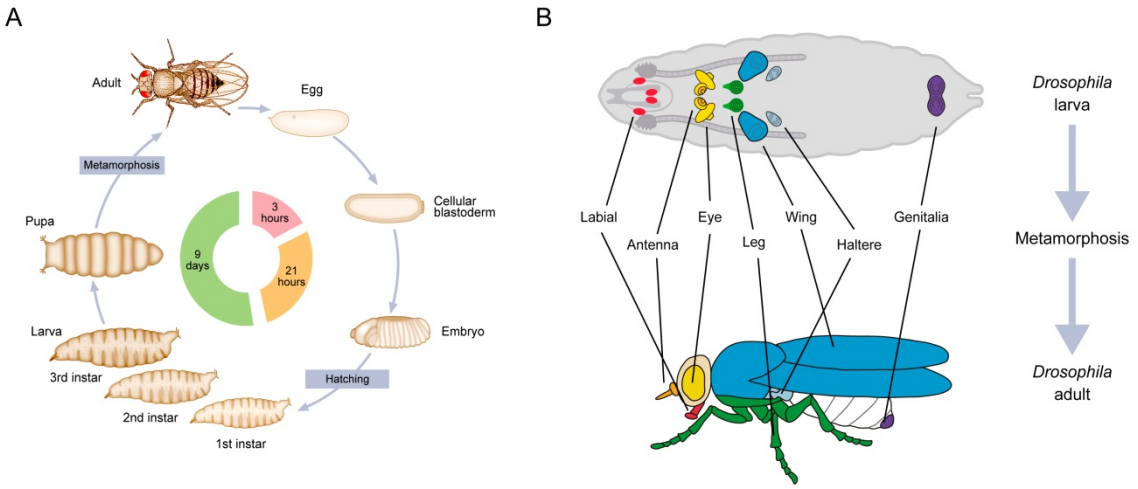


Figure 2. *Drosophila* life cycle and imaginal discs. (**A**) *Drosophila* development starts with the deposition of the fertilized egg, from which the embryo will develop. After the egg hatches, the resulting larva grows by endoduplication, moulting two times and originating three larval stages (LI-LIII). During larval development, imaginal disc cells (**B**) proliferate extensively. By the end of the third larval stage, the larva stops feeding and moving and encapsulates in the pupa and undergo a four-day-long metamorphosis, after which the adult fly emerges (ecloses).

The imaginal discs are a bilayer epithelia composed of a monolayer of cuboidal epithelial cells and an overlying squamous epithelium known as the peripodial membrane (both of which originate from the embrionary ectoderm). Particularly, the pair of eye-antennal discs gives rise to most of the head capsule and the major cephalic sensory organs (the compound eyes, antennae, maxillary palps, and ocelli).

The size and structure of the adult compound eye of *Drosophila* is remarkably robust, with less than 0.2% variations between individuals, making this system particularly useful for the study of size regulation and cancer. Each compound eye consists of 779-800 identical eye units called ommatidia. Each ommatidium contains eight photoreceptor cells, surrounded by four cone cells (the fly eye lenses) and 8 pigment cells and support cells. Ommatidia are oriented with specular symmetry along the dorso/ventral axis or equator of the eye (**Figure 3B**), being this orientation and the growth of the eye dependent on Notch (Dominguez and Casares, 2005).

The eye-antennal disc has a complex embryonic origin (Campos-Ortega and Hartenstein, 1985) and hence I will focus on the part of the disc that will give rise to the compound eye and whose growth and development depend on Notch activity. The eye disc part originates from a group of ~6-20 cells from the dorsal head epidermis at the embryonic stage 12. A period of extensive remodelling and cell death brings together the eye disc part with the parts that will give rise to the antennal, maxillary palps and

labial disc part (Bate and Martinez Arias, 1993). At the end of the embryonic stage 14, all imaginal discs, including the eye-antennal imaginal disc, undergo an invagination process. After that, they remain quiescent for the rest of the embryogenesis, until the second larval stage (LII), when imaginal cells resume growth and proliferation upon larval feeding. In the following paragraphs I will resume briefly the development of the eye disc and the growth-promoting organizer. For details in this process see Dominguez and Casares (2005).

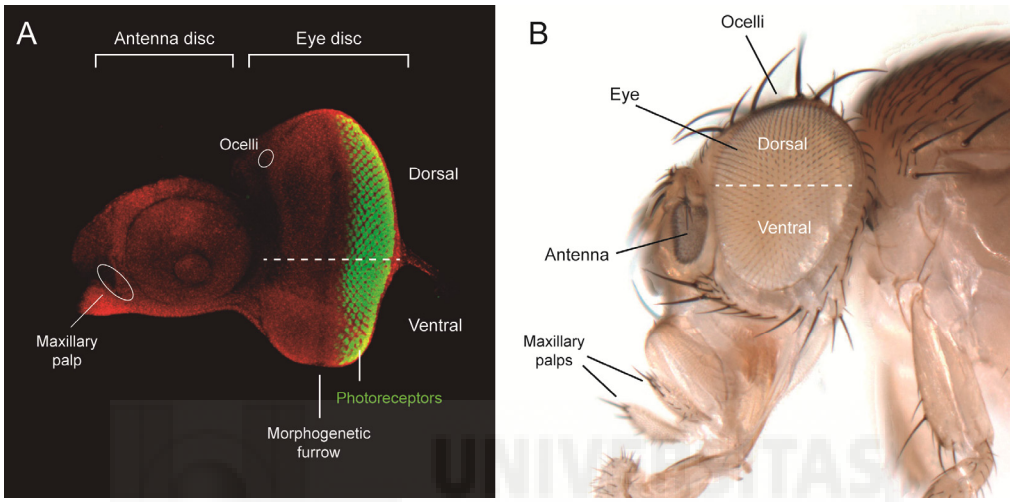


Figure 3. The compound eye-antennal imaginal disc. (**A**) The eye-antennal imaginal disc from an L3 larva gives rise to most of the head capsule and the major cephalic sensory organs: the compound eyes, antennae, maxillary palps, and ocelli. (**B**) Structure of the adult head.

The eye disc part divides asynchronously and extensively during the second and third instar larval stages. However, at the start of the third larval stage, at the most posterior end of the eye disc, a morphogenetic wave of retinal differentiation begins its progression in a posterior-to-anterior direction. This morphogenetic "wave" is associated with a transient furrow in the disc epithelium known as the morphogenetic furrow (MF) (**Figure 3A**). The MF is the site where commitment to photoreceptor fate is initiated. As it moves anteriorly, cells anterior to this wave stop proliferation and become arrested in the cell cycle before some of these cells are recruited to the ommatidial clusters. The cells located more anterior to the arrested cells continue cell proliferation and remain in an undifferentiated state. At the MF, the recruitment of ommatidial cells occurs in a highly stereotyped manner (reviewed in Treisman, 2013). During the pupal stage, the surplus cells that did not participate in the construction of the ommatidium are eliminated by programmed cell death in which Notch is also implicated.

The spatial and temporal distinction between proliferation and retinal differentiation, and particularly the existence of promoter regions that enable the activation (or repression) of genes exclusively during the proliferation period makes the developing eye disc of *Drosophila* a powerful experimental paradigm for the genetic dissection of the mechanisms that control cell cycle progression, survival and apoptosis during normal and tumour growth, with no interference with the mechanisms of cell specification and differentiation.

2.3.2 Notch and the dorso/ventral organizer of the eye imaginal disc of Drosophila

The early growth of the eye imaginal disc depends on the activity of an organizer that controls, not only the global growth of the disc, but also its morphogenesis, the dorso/ventral polarity and the initiation point of the MF. This organizer depends on the localized activation of Notch signalling pathway (Cavodeassi *et al.*, 1999; Cho and Choi, 1998; Dominguez and de Celis, 1998; Papayannopoulos *et al.*, 1998).

The formation of this organizer first implies the establishment of two asymmetric regions known as dorsal and ventral compartments, limited by different cell lineages (Dominguez and de Celis, 1998). The generation of this asymmetry depends on the expression of three related genes: *mirror*, *araucan* and *caupolican*, which form the *Iroquois* gene complex or *Iro-C* (Cavodeassi *et al.*, 1999). This will subdivide the eye disc in two different regions, dorsal and ventral (Dominguez and de Celis, 1998), that later on will allow the establishment of the Notch organizer.

The expression of *Iro-C* complex is regulated early by expression of the gene *pannier* in the dorsal border of the eye. *pannier* controls the expression of the gene *wingless*, which in turn acts synergistically with Hedgehog pathway to control the activation of the *Iro-C* complex in the dorsal compartment (Cavodeassi *et al.*, 1999; Heberlein *et al.*, 1998; Maurel-Zaffran and Treisman, 2000). Restricted expression of *Iro-C* in the dorsal region seems to be reinforced by the expression of the transcription factors Sloopy paired 1 and 2 in the ventral region of the eye, and by the JAK/STAT pathway (Gutierrez-Aviño *et al.*, 2009). Its activity should repress any ventral expression generated by Wingless diffusible signal (Sato and Tomlinson, 2007).

Expression of the *Iro-C* complex in the eye provides cells with dorsal identity, while the ones not expressing it acquire ventral identity (Cavodeassi *et al.*, 1999). The genes *mirror*, *araucan* and *caupolican* encode transcriptional repressors that repress expression of the gene *fringe* in the dorsal region, restricting it to the ventral compartment. Such restriction originates the establishment of a border of *fringe*

expression in the medial dorso/ventral line, essential for the formation of the Notch organizer region (Cho and Choi, 1998; Dominguez and de Celis, 1998; Papayannopoulos *et al.*, 1998) (**Figure 4**).

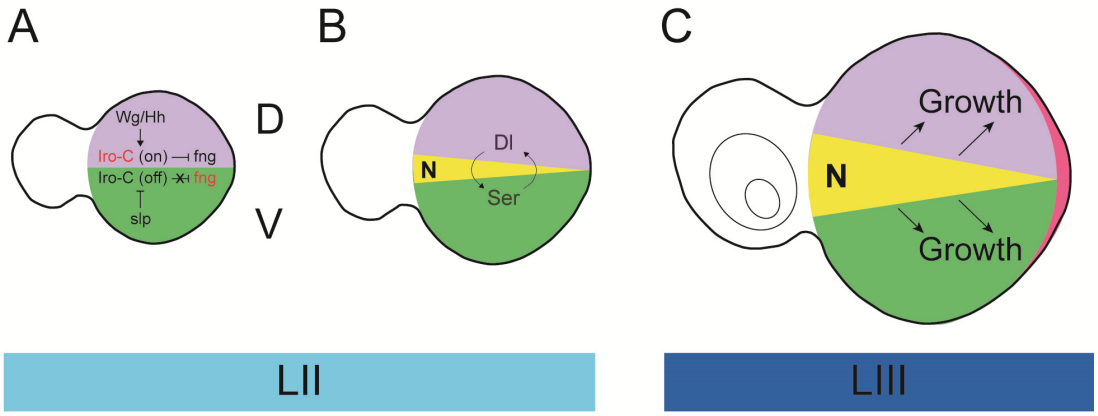


Figure 4. Establishment of the dorso/ventral organizer in the eye disc of *Drosophila*. (**A**) The expression of the complex *Iro-C* (induced by Wingless and Hedgehog) confers dorsal identity to the cells expressing it, while its absence defines the ventral identity of the disc. The activity of Iro-C restricts the expression of *fng* to the ventral region. (**B**) Fringe reduces Notch affinity for its ligand Serrate in the ventral compartment, allowing activation of the receptor only in the dorsal cells adjacent to the dorso/ventral border. This signal activates the transcription of the ligand *Dl* in the dorsal compartment, which in turn activates Notch in the ventral cells located near the dorso/ventral border, and promotes *Serrate* transcription. (**C**) The positive feedback between Dl and Serrate increases Notch activation giving rise to the organizer, which coordinates the growth of the cells in the eye imaginal disc. Adapted from Gutierrez-Aviño *et al.* (2009).

The gene *fringe* encodes an O-fucose specific β-1,3-N-Acetylglucosaminyltransferase whose function is to potentiate the sensibility of the transmembrane receptor Notch for Dl type ligands, while simultaneously decreases Notch sensibility to Serrate type ligands (Haltiwanger, 2002). This differential modulation of Notch ligands mediated by Fringe allows local activation of the Notch receptor along the border of the dorso/ventral compartment of the eye (Cho and Choi, 1998; Dominguez and de Celis, 1998; Papayannopoulos *et al.*, 1998) (**Figure 4**). High activation of Notch in the border of expression of *fringe* is maintained by a positive feedback between the ligands (Cavodeassi *et al.*, 1999; Wu and Rao, 1999).

Complete elimination of *fringe* expression or its generalized expression impedes the activation of the gene *Dl* by Serrate, blocking the feedback between the ligands and the high activation of Notch, resulting in flies with very small eyes or even absent eyes (Dominguez and de Celis, 1998; Dominguez *et al.*, 2004) (**Figure 5B**). On the contrary, *Dl* generalized expression expands the region in which Notch is activated to all the

ventral region, producing an increase in cell proliferation leading to flies with overgrown eyes (Dominguez and de Celis, 1998) (**Figure 5C**).

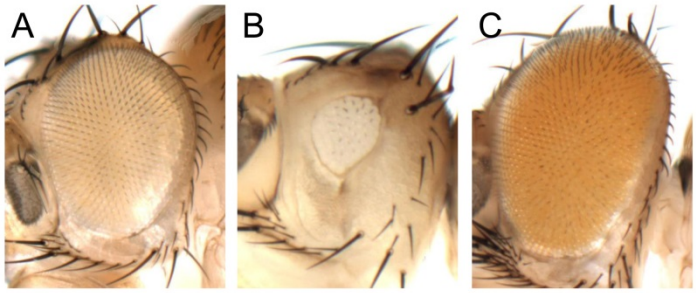


Figure 5. Defects in the formation of the eye induced by alterations in the formation of the dorso/ventral organizer. (**A**) Wild-type eye (w^{1118}). (**B**) Small eye defect produced by generalized expression of the gene *fng* (*ey-Gal>fng*). The presence of fng in the dorsal compartment of the eye hinders the proper establishment of the dorso/ventral border and hence the organizer, resulting in flies with smaller eyes. (**C**) Overgrown eye defect produced by generalized expression of the gene *Dl* (*ey-Gal>Dl*). The presence of the ligand Dl in the ventral compartment of the eye expands the activation domain of Notch receptor towards this area of the disc, increasing proliferation in the ventral part of the eye and originating flies with eyes bigger than the wild-type.

3. A genetic design for unbiased screen for cancer genes

3.1 Misexpression Screen: the Gene Search system

With the aim of identifying new genes that collaborate with Notch signalling pathway in growth and cancer development, our group devised a misexpression screen using the Gene Search (GS) system to screen for genes that provoked tumours when co-expressed with the Notch ligand *Dl* in the proliferating *Drosophila* eye (**Figure 6A**). This method of combining misexpression via GS or EP lines in the "sensitized" Delta-overexpression (i.e. overgrowth) background has now been adopted by many laboratories worldwide to identify and validate genes in cancer and cancer suppression. In our original design, we used the *eyeless-Gal4/UAS* system (Brand and Perrimon, 1993; Dominguez and de Celis, 1998), combined with individual insertions of the original transposable P-element GS (Toba *et al.*, 1999).

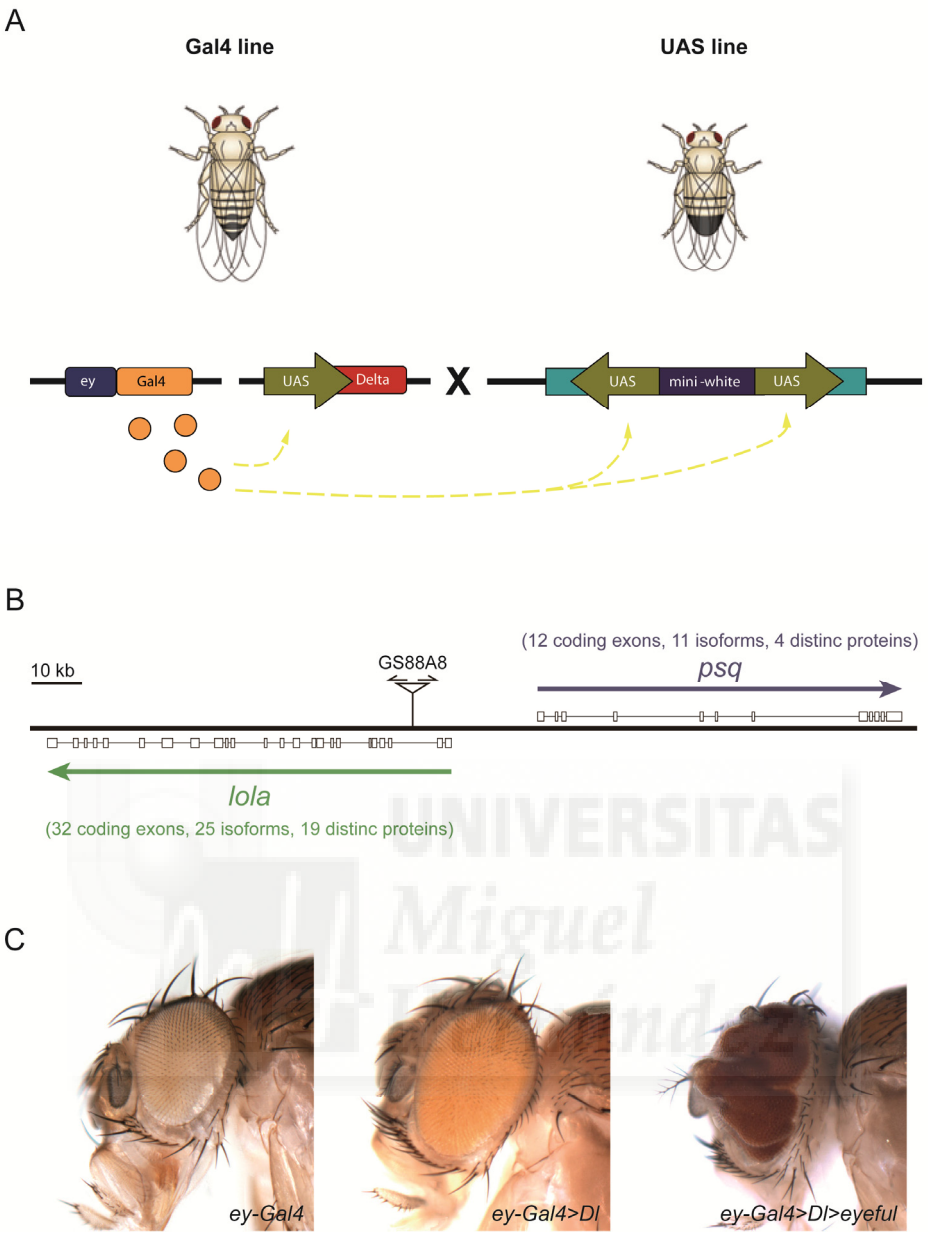


Figure 6. Design of the genetic screen search for genes involved in Notch-mediated tumorigenesis. (**A**) Flies expressing *ey-Gal4* and *UAS-Dl* (*ey-Gal4>Dl*) are crossed with flies carrying a GS-element inserted in their genome. The progeny of these crosses expresses the protein Gal4, which binds to the UAS sequences and induces the co-expression of the ligand *Dl* and the gene or genes adjacent to the GS-element insertion during the eye growth. The GS lines that enhance or suppress the "overgrown eye" defect induced by *Dl* overexpression are isolated and characterized genetically and molecularly, with the aim to identify new genes involved in Notch-mediated tumorigenesis. (**B**) Genomic organization of the *lola* and *psq* genes in the cytological region 47A11-B1 and the *eyeful* GS insertion site. (**C**) Example of the tumour phenotype caused by the mutation *eyeful* (*GS88A8*).

Dolors Ferres-Marco and Maria Dominguez (Ferres-Marco, 2010) carried out a high throughput genetic screen in a GS line inserted in the second chromosome. The

GS element was mobilized and the individual new insertions generated were tested for their capacity to form tumours in combination with the gain of Dl. The GS element contains two inverted copies of the *Gal4* transcription factor binding site called UAS (Upstream Activating Sequences) adjacent to the core promoter from the *Hsp70Bb* gene. Each copy is located near the terminal inverted repeats at each end of the vector, and oriented to direct transcription outward (**Figure 6A**). Thus, upon Gal4 activation the UAS repeats will induce transcription toward the flanking DNA, both 5' and 3' of its integration site. This is different from the EP element (Rørth, 1996), which only contains UAS sequences in the 5' end. The GS-element also carries a mini-*white* gene between the UAS sequences as marker of the insertion.

In the original screen, 1514 GS lines were generated with independent insertions in the genome via the mobilization of the original GS-element in the second chromosome (Ferres-Marco *et al.*, 2006; Toba *et al.*, 1999). Among these lines, approximately 0.2% induced metastatic tumours when combined with *Dl* overexpression (**Figure 6C**). One of the GS mutations identified was named *eyeful* (*GS88A8*, Ferres-Marco *et al.*, 2006). The *GS88A8* is inserted in an intron of the gene called *longitudinals lacking* (*lola*), and the next gene located at approximately 40 kb in the opposite direction is *pipsqueak* (*psq*) (**Figure 6B**). We found that eyeful phenotype is caused by the co-expression of both genes, but *pipsqueak* has the predominant role (particularly in metastatic behaviour) (Ferres-Marco *et al.*, 2006). Hence, we focused on studying the molecular mechanisms underlying Pipsqueak-mediated oncogenesis.

3.2 Identification of Pipsqueak as a new oncogene of the BTB/POZ transcription factor family

3.2.1 Gene structure

psq is a large and complex locus, which encodes eleven mRNA variants produced by alternative splicing and alternative promoter use (http://flybase.org/, (Ferres-Marco *et al.*, 2006; Grillo *et al.*, 2011; Horowitz and Berg, 1996; Weber *et al.*, 1995) (**Figure 7A**).

There are four long mRNA isoforms described with alternative 5' ends, *psq-RA*, *psq-RB, psq-RC* and *psq-RM* (following Flybase nomenclature [http://flybase.org/]), being the first three transcripts around 5 kb long and the last one 6.3 kb long. This corresponds to the 5.1 kb *psq-1* and a 6 kb unnamed message transcript reported by Horowitz and Berg (1996). They observed that *psq-1* is below the level of detection in

males but highly abundant in females, specifically in the ovaries; and in 0-2 hours embryos, decreasing to levels below detection in older embryos, larvae and pupa. This is a typical pattern observed for maternal mRNAs that are deposited into the developing oocyte. On the contrary, the 6 kb transcript appears to be present equally in both sexes and throughout development, though its levels can vary.

The remaining mRNA variants are shorter, varying between 4.4 and 4.8 kb, which is similar to the 4.4 kb *psq-2* isoform also reported by Horowitz and Berg (1996). The main difference between *psq-1* and *psq-2* is that the second has a novel sequence in the 5' end and a short extra sequence at the 3' end.

3.2.2 Protein structure

The different *psq* mRNA variants generate long and short protein isoforms (**Figure 7B**). The long isoforms are 1064 to 1123 aminoacids (aa) long, and are produced from the long mRNA variants (*psq-RA*, *psq-RB, psq-RC* and *psq-RM*). The short isoforms come from the short mRNA variants (Horowitz and Berg, 1996) and are predicted to be between 639 to 645 aa long, lacking the BTB domain of Psq.

Long and short isoforms share a number of repetitive aa sequences, including multiple regions enriched with serine/glycine, proline or alanine residues, several OPA repeats (Wharton *et al.*, 1985b), and a domain with histidine residues alternating with other aa (mainly glutamic acid or glycine), followed by four tandem copies of a conserved helix-turn-helix (HTH) motif, which is a sequence-specific DNA-binding domain called HTH_Psq domain (Lehmann *et al.*, 1998). The Psq domain binds to GAGA sequence repetitions that are present in many elements that respond to Polycomb protein and that regulate homeotic genes, as well as in numerous genes that regulate cell cycle and development (Huang *et al.*, 2002; Lehmann *et al.*, 1998).

The long isoforms contain in the amino-terminal end a protein-protein interaction domain called BTB/POZ (Broad-Complex, Tramtrack and Bric à brac/Poxvirus and Zinc finger) (Bardwell and Treisman, 1994; Godt *et al.*, 1993) (**Figure 7B**). Since its discovery as a sequence motif in genes of DNA virus (Koonin *et al.*, 1992), it has been found in proteins of a variety of species, as diverse as slime moulds (Escalante *et al.*, 1997) and humans (Albagli *et al.*, 1995), being many of them DNA-binding C2H2 zinc finger (ZF) proteins. In BTB/POZ proteins that contain a zinc finger DNA-binding motif, DNA-binding is strongly inhibited by the BTB/POZ domain (Bardwell and Treisman, 1994). This inhibition appears to be the result of oligomerization through protein-protein interactions mediated by the BTB/POZ domain. The tendency of BTB/POZ proteins to

oligomerize in solution and their localization in distinct nuclear substructures (Bardwell and Treisman, 1994; Horowitz and Berg, 1996; Koonin *et al.*, 1992; Raff *et al.*, 1994) suggests that they might act by modifying chromatin structure (Albagli *et al.*, 1995; Croston *et al.*, 1991; Dorn *et al.*, 1993; Kerrigan *et al.*, 1991; Tsukiyama *et al.*, 1994).

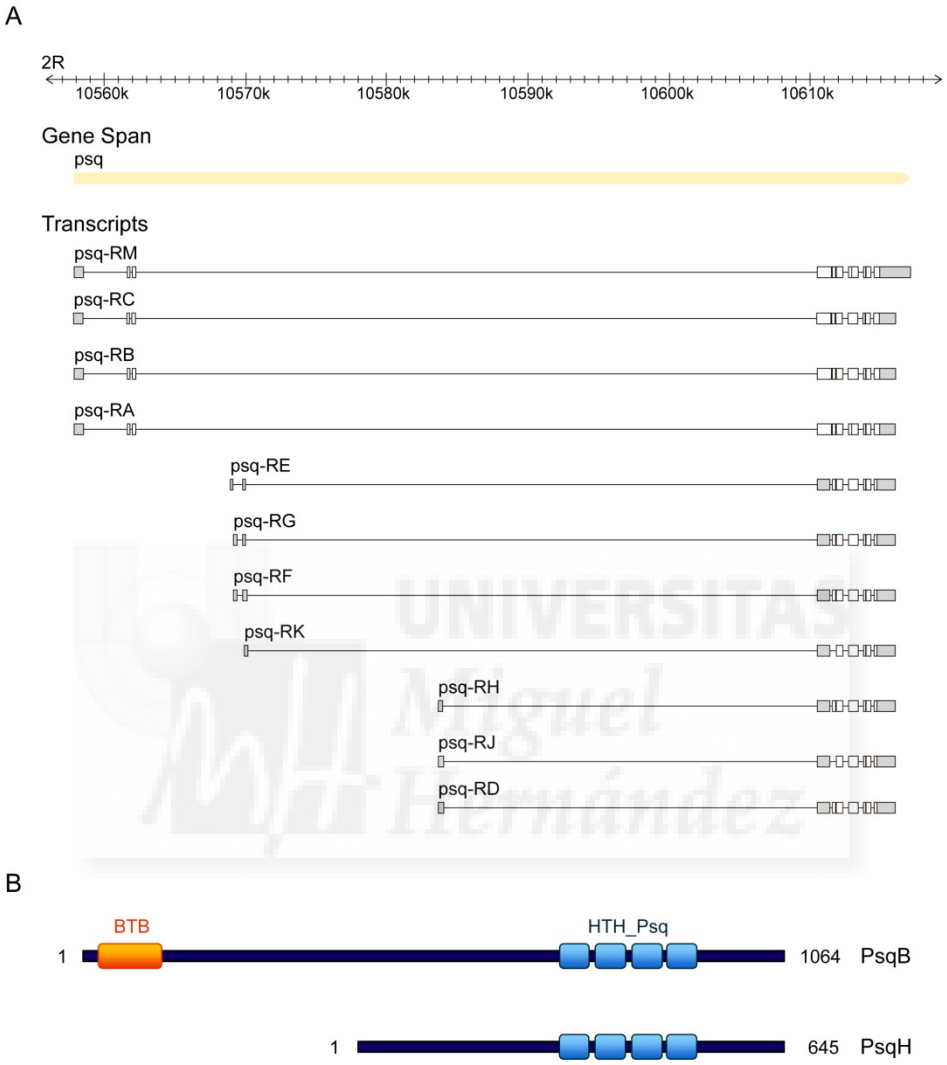


Figure 7. Pipsqueak gene model and products. (**A**) Schematic representation of the transcripts encoded by *psq*. White boxes indicate exons. Grey boxes indicate 5' and 3'-UTR. Adapted from Flybase. (**B**) Examples of the Psq long isoform protein structure (PsqB showed as example) and Psq short isoform (PsqH showed as example).

3.2.3 Known roles of Pipsqueak during development

Loss of maternal *psq* function revealed a role early in oogenesis (Horowitz and Berg, 1996; Siegel *et al.*, 1993), and defined *psq* as one of the posterior group genes responsible for pole cell formation and proper abdominal segmentation of the embryo (Siegel *et al.*, 1993). Moreover, *psq* directs correct localization of the *gurken* product, which is involved in establishment of the dorso/ventral axis of the embryo (Horowitz and Berg, 1996). Recently, a novel role for *psq* has been described in embryonic terminal patterning through the transcriptional regulation of *torso* (*tor*) (Grillo *et al.*, 2011). During metamorphosis, *psq* has been postulated to have a role in retinal cell fate determination of photoreceptors R3/R4 and for proper differentiation of other adult structures, such as wings and legs (Weber *et al.*, 1995). Of note, transcripts encoding isoforms containing the BTB domain are required mostly during oogenesis, with not known role during embryonic or larval stages. Regarding Psq activity at the molecular level, Psq binds to the GAGA sequence (Lehmann *et al.*, 1998; Schwendemann and Lehmann, 2002), which is present in many Hox genes and in hundreds of other chromosomal sites (Ringrose and Paro, 2004). Additionally, Psq protein has been found in a complex with Polycomb, which suggests that Psq may act in sequence-specific targeting of a Polycomb-group (Pc-G) complex that contains HDAC activity (Huang *et al.*, 2002; Ringrose and Paro, 2004).

3.2.4 Pipsqueak acts as an epigenetic regulator

Results from our laboratory unveiled a new function for Psq in tumour development (Ferres-Marco *et al.*, 2006). In that work, a GS line was identified that when coupled to Dl overexpression, produced massive overgrowths and tumours. Some of these mutant flies display macroscopically visible secondary eye tumours within the head, thorax, and abdomen, hence the line was named ‘‘eyeful’’ (**Figure 8A** and **B**). The *eyeful* GS-element (*GS88A8*) causes the deregulation of two neighbouring genes, *lola* and *psq*. It was shown that transcription of both genes contributes to tumorigenesis, although Psq seems to be more important for tumour growth and metastasis (**Figure 8C-E**). Both the BTB and the DNA-binding domain of Psq are essential for its oncogenic function, as mutations in either domain rescue the tumorigenic phenotype (**Figure 8D** and **E**).

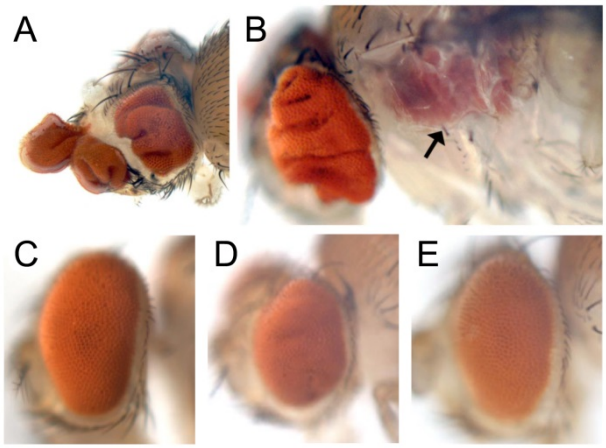


Figure 8. The *eyeful* phenotype is rescued by mutations in *lola* and *psq*. (**A-B**) The combination of *ey-Gal4>Dl* with the GS insertion *eyeful* (*GS88A8*) causes tumour growths in the eye and secondary growths (black arrow in **B**). Mutations in either *lola* (*GS88A8*^{rev10}) (**C**) or *psq* (**D, E**), are able to rescue the tumorigenic phenotype. Additionally, mutations in the BTB (*GS88A8*^{rev9}) (**D**) or in the DNA domain of Psq (*GS88A8*^{rev12}) (**E**) revert GS88A8-mediated tumorigenic phenotype, indicating that both domains are important for Psq oncogenic activity.

Several human oncogenic proteins, like PLZF (promyelocityc leukaemia zinc finger) and BCL-6 (B cell lymphoma 6), associate with HDACs, co-repressors, and Pc-G repressors through their BTB domain (Melnick *et al.*, 2002). It is possible that Psq and/or Lola (which also possesses a BTB domain), interact through their BTB with HDACs and Pc-G repressors to recruit this complexes to particular genes, for example homeotic genes with GAGA sequences. So deregulation of *psq* and *lola* might induce tumorigenesis through aberrant epigenetic silencing of genes that contribute to the uncontrolled growth of tumour cells. In line with this, it was found a loss or strong reduction of the open chromatin mark H3K4 trimethylation (H3K4me3) in the developing eye tissue from which the tumour arises, suggesting that the chromatin in the mutant tissue has been condensed or silenced. Additionally, reducing the dosage of genes related to gene silencing and chromatin condensation, like *Rpd3/HDAC*, *E(z)*, *Su(var)3-9*, *Pc*, and *Esc*, impeded tumour development. Finally, it was demonstrated that the increase in epigenetic silencing reduced the expression of the retinoblastoma family protein (*Rbf*) gene, a well-known tumour suppressor gene, and this down-regulation was shown to be necessary for tumour development.

To sum up, results from our laboratory confirmed that Psq and the nearby gene called *lola*, both combined with *Dl* overexpression, act as epigenetic regulators of Pc-G family in the formation of highly invasive tumours (Ferres-Marco *et al.*, 2006).

3.2.5 BTB transcription factors in cancer

Psq does not have a homolog outside of the Arthropod group, but its general domain architecture is well conserved through evolution (**Figure 9**). Actually, the BTB domain is often found in proteins in combination with zinc finger domains (Perez-Torrado *et al.*, 2006), which are structurally related to Psq DNA-binding domain. Until now there have been identified at least 49 genes that encode BTB-ZF proteins and they commonly serve as sequence-specific silencers of gene expression through the recruitment of transcriptional co-repressors (Ahmad *et al.*, 1998; Ahmad *et al.*, 2003; Li *et al.*, 1999; Siggs and Beutler, 2012).

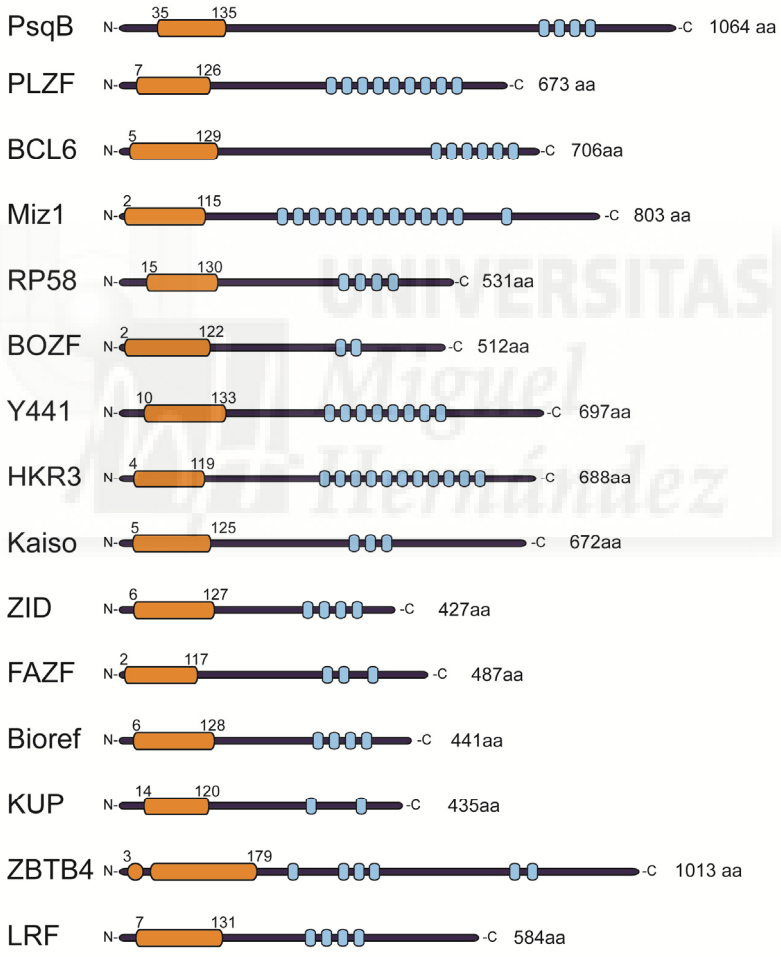


Figure 9. Comparison of the general architecture of several BTB transcription factors. The overall architecture of BTB-ZF proteins is highly conserved. N-terminal BTB domains (orange) are followed by multiple copies of zinc finger motifs (light blue). The start and end sites of BTB domains are also marked. The BTB domain in ZBTB4 contains an internal gap. Adapted from Wang (2009).

Interestingly, the subfamily of BTB transcriptional repressors includes the human oncogenes BCL-6 (B cell lymphoma 6), PLZF (promyelocityc leukaemia zinc finger), LRF (leukaemia/lymphoma related factor/Pokemon), FAZF (Fanconi anemia zinc finger) and ZID (zinc finger protein with interaction domain). In these human oncogenes, the BTB domain is a key element for the oncogenesis, apparently due to the recruitment of co-repressor nuclear proteins as well as HDAC complexes (Melnick *et al.*, 2002). In this Thesis work we focus on understanding the molecular mechanisms that regulate Psq activity during tumorigenesis as a paradigm to unravel new mechanisms of oncogenesis mediated by BTB transcription factors.

4. Sumoylation pathway and its implication in cancer

A multitude of mechanisms determine the *in vivo* function of proteins. Among them are the regulation of protein levels via control of expression levels and turnover, in addition to regulation of protein activity, together with localization and/or interactions by constitutive or reversible post-translational modifications. These modifications, usually accomplished via enzymatic reactions, result for example in acetylation, methylation, phosphorylation, ADP ribosylation, carboxylation, adenylylation, and glycosylation or prenylation of aa side chains. Among these is also ubiquitination, a post-translational modification that involves attachment of the 76-residue protein ubiquitin to other proteins. Ubiquitination often targets the substrate protein for degradation by the proteasome, but it can also have several other functions such as controlled activation of the protein via limited proteolysis (Chen *et al.*, 1999; Hershko and Ciechanover, 1998; Hoppe *et al.*, 2001; Jiang and Struhl, 1998; Pickart, 2001; Rape and Jentsch, 2002) or receptor mediated endocitosis via monoubiquitination (Hicke, 1997).

In the past 30 years, several small ubiquitin-like proteins (Ubls) that also act as post-translational modifications on other proteins have been discovered. Ubls have a variety of different functions, but they do not target their substrates directly for proteasome-dependent proteolysis. These Ubls vary widely in their degree of sequence similarity to ubiquitin and can be divided in two groups: proteins that are not available for conjugation (e.g. Rad23, Dsk2p, Elongin B), and proteins that, like ubiquitin, are attached to other proteins (reviewed in Jentsch and Pyrowolakis, 2000). To this second

group belongs the Small Ubiquitin-related Modifier (SUMO) family, the Ubl with the widest range of functions and the most known substrates.

SUMOs constitute a highly conserved protein family found in all eukaryotes and are required for viability of most eukaryotic cells, including budding yeast, nematodes, fruit flies, and vertebrate cells in culture (Apionishev *et al.*, 2001; Epps and Tanda, 1998; Fraser *et al.*, 2000; Hayashi *et al.*, 2002; Jones *et al.*, 2002). Since its discovery in the late '90s (Mahajan *et al.*, 1997), SUMO has been found covalently attached to more than 50 proteins, which include the androgen receptor, IκBα, c-jun, HDACs, p53, and other proteins that participate in transcription, DNA repair, nuclear transport, signal transduction, and the cell cycle. This huge amount of targets leads to sumoylation being involved in a long list of biological functions and also makes it difficult to predict the outcome of SUMO attachment. With the data available until now, only one general conclusion about sumoylation effect is possible: downstream consequences of sumoylation are target specific, but are usually caused by altered interactions of the modified protein with other macromolecules, including proteins, DNA, or RNA.

4.1 Sumoylation pathway components

Sumoylation is an evolutionarily ancient and highly conserved reversible post-translational modification. Like ubiquitination, sumoylation results in the formation of an isopeptide bond between the carboxy-terminal glycine residue of SUMO and the ε-amino group of a lysine residue in the acceptor protein. A three-step enzyme pathway attaches SUMO to specific substrates, and SUMO-specific proteases cleave SUMO off its targets (**Figure 10**). The enzymes of the SUMO pathway, although analogous to those of the ubiquitination pathway, are specific for SUMO and have no role in conjugating ubiquitin or any of the other Ubls.

Most targets seem to undergo rapid cycles of sumoylation and desumoylation, often resulting in a very low steady-state level (frequently less than 1%), of the modified species. That is one of the reasons for the late discovery of SUMO. Another is that SUMO-specific proteases rapidly desumoylate all conjugates instantly upon cell lysis, unless cells are lysed under denaturing conditions or SUMO-specific proteases are inhibited. A third reason is that, for some sumoylated proteins, eliminating the SUMO attachment site has fairly subtle effects on protein function, so that functional domains containing the attachment sites were not immediately apparent.

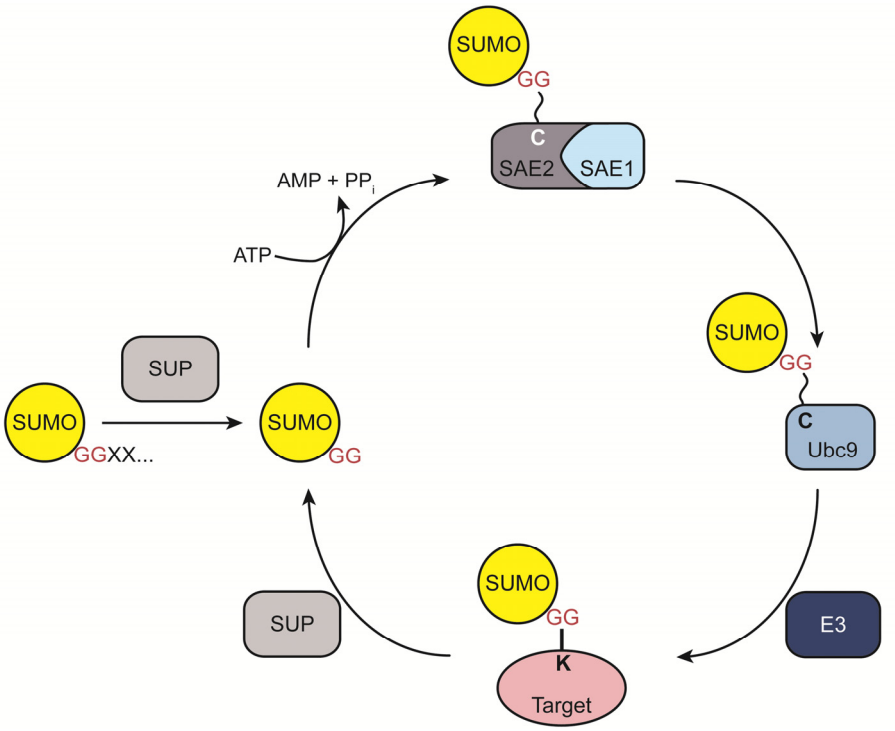


Figure 10. The cycle of reversible sumoylation. Prior to the first conjugation cycle, a nascent SUMO has to be proteolytically cleaved to expose the C-terminal glycine-glycine (GG) motif. This is achieved by SUMO-specific proteases (SUPs in the figure) of the Ulp/SENP family. A mature SUMO is activated by the SUMO E1 heterodimer SAE1/SAE2 (SUMO-activating enzyme subunits 1 and 2) in an ATP-consuming step, resulting in formation of a thioester bond between the C-terminal glycine of SUMO and the catalytic cysteine (C) of SAE2. SUMO is then transferred to the catalytic cysteine of the E2 enzyme Ubc9, again forming a thioester bond. Ubc9 catalyzes formation of an isopeptide bond between the C-terminal glycine of SUMO and a lysine (K) residue in the substrate, usually together with a specific SUMO E3 ligase. The target lysine residues usually fall within a SUMO conjugation consensus motif ΨKXE. Modification with SUMO is reversible owing to SUMO-specific proteases that cleave the isopeptide bond and release SUMO for further cycles. Abbreviation: XX, aa C-terminal of the GG motif in nascent SUMO. Adapted from Flotho and Melchior (2013).

Regarding the SUMO-acceptor site it was shown to be the short sequence ΨKXE, where Ψ is a large hydrophobic aa, generally isoleucine, leucine, or valine; K is the lysine residue that is modified; X is any residue; and E is a glutamic acid. Other more extended motifs have been proposed and are summarized in **Table 2**. These extended motifs are thought to provide further specificity and enhanced intrinsic propensities for sumoylation. However, having one of these motifs does not necessarily mean that a protein will be a target for sumoylation. Since all SUMO consensus sites are not modified and certain modification sites clearly deviate from the consensus, other factors, such as subcellular environment due to localization, appropriate presentation of the sequence on the substrate or phosphorylation of neighbouring aa

residues, must affect the target site specificity (Bossis and Melchior, 2006; Hietakangas *et al.*, 2006).

Motif	Consensus sequence	Regulatory mechanisms
Consensus motif	Ψ**K**XE	-
Inverted consensus motif	(E/D)X**K**Ψ	-
Hydrophobic consensus motif	ΨΨΨ**K**XE	-
Phosphorylation-dependent sumoylation motif (PDSM)	Ψ**K**XEXX(pS)P	Phosphorylation by proline-directed kinases
Negatively charged aa–dependent sumoylation motif (NDSM)	Ψ**K**XEXXEEEE	-
Phosphorylated sumoylation motif (pSuM)	Ψ**K**X(pS)P[a]	Phosphorylation by proline-directed kinases
Extended pSuM	Ψ**K**X(pS)P[a](pS) XXX(pS)P	Phosphorylation by proline-directed kinases and GSK3 (carboxy-terminal extension)

Table 2. List of known SUMO acceptor sites. [a] Demonstrated only in the context of a proline-directed kinase phosphorylation motif; however, other kinase recognition sites are conceivable. Adapted from Flotho and Melchior (2013).

In *Drosophila*, all of the components of the sumoylation pathway have been identified, and their function has been studied (**Table 3**). The *in vivo* analysis of sumoylation in the fruit fly presents a number of advantages. On the one hand, gene redundancy is lower in *Drosophila* when compared with vertebrate models, which simplifies functional analysis. On the other hand, and related to the genetic accessibility of this organism, the various components of the pathway have been implicated in multiple cellular and physiological processes by means of genome wide genetic screens performed *in vivo*.

Gene	CG accession number	Homologues	Function	Biological processes
smt3	CG4494	SUMO3	SUMO moiety	Cell proliferation or survival; chromatin modification; embryogenesis; EGFR signalling; immune response; lysosomal transport; oogenesis; wing morphogenesis
Ulp1	CG12359	Sentrin/SUMO-specific protease 1	Isopeptidase	Smt3-conjugates nucleocytoplasmic shuttling
Aos1	CG12276	Activating enzyme subunit 1	E1A-activating	-
Uba2	CG7528	Activating enzyme subunit 2	E1B-activating	-
lesswright	CG3018	UBC9; ubiquitin-conjugating enzyme E2	E2-conjugating	Cell proliferation; chromatin modification; embryogenesis; immune response; wing morphogenesis
tonalli	CG7958	Zimp7 and Zimpo10; retinoic acid-induced 17	E3 ligase	Chromatin modification
Suppressor of variegation 2-10	CG8068	PIAS3	E3 ligase	Chromatin modification and chromosomal inheritance; negative regulation of JAK/STAT signalling; wing morphogenesis

Table 3. Sumoylation pathway components in *Drosophila*. List of the known homologs of sumoylation pathway components identified in *Drosophila* until now. Adapted from Talamillo *et al.* (2008).

4.1.1 SUMO proteins

SUMO proteins belong to the small family of protein modifiers known as Ubls (van der Veen and Ploegh, 2012), which are characterized by the ubiquitin fold (globular β-grasp fold) and a characteristic C-terminal Gly-Gly motif that is exposed after proteolytic maturation. They are ~11 kDa proteins, but they appear larger on SDS-PAGE and add ~20 kDa to the apparent molecular weight of most substrates. Primary sequence and

surface charge distributions differ significantly between Ubls and as a consequence, each Ubl modification pathway requires distinct sets of enzymes and binding partners. A SUMO-specific characteristic is the flexible amino-terminal of about 20 aa, which seems to primarily serve as an acceptor in SUMO chain formation (**Figure 11**).

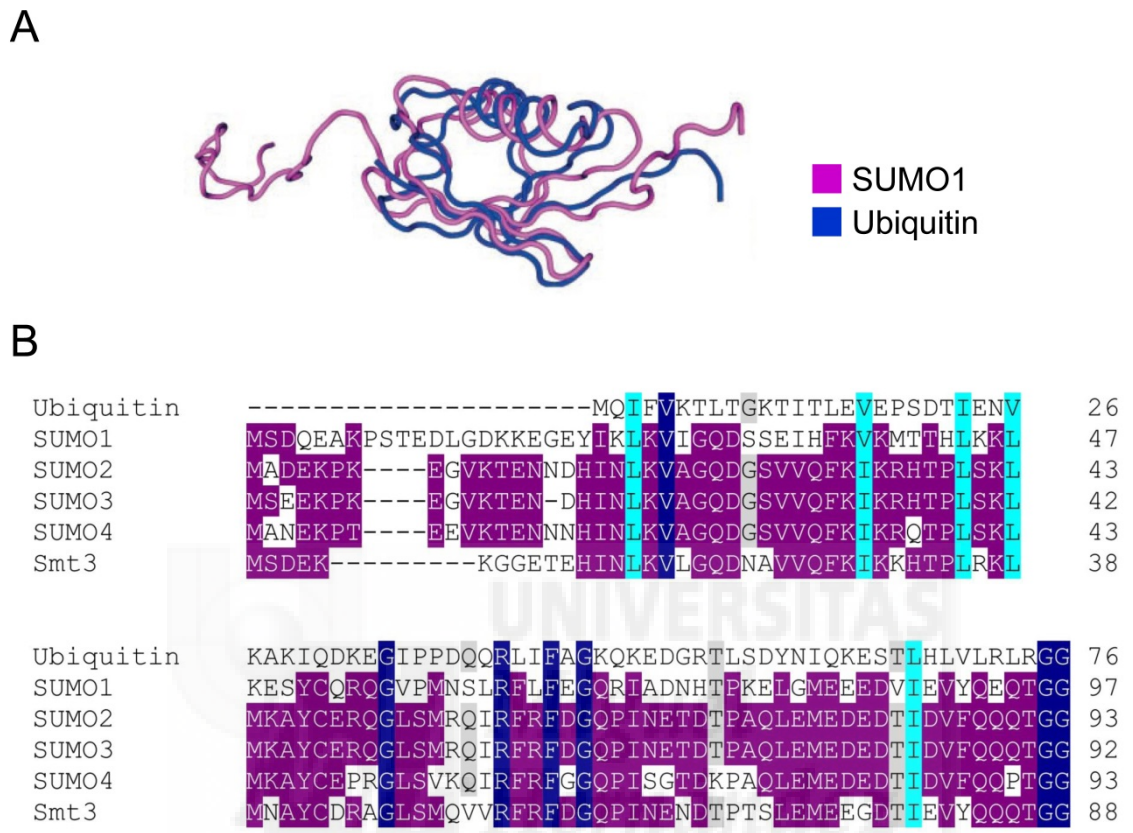


Figure 11. Comparison of SUMO and ubiquitin protein structures. (**A**) Alignment of the protein backbone of SUMO1 (pink) and ubiquitin (blue). The amino-terminal ends are on the left and the carboxy-terminal on the right. SUMO structure is of the precursor and includes the C-terminal tetrapeptide that is cleaved off for maturation of the protein. Adapted from Johnson (2004) (**B**) Sequence alignment of *Homo sapiens* ubiquitin, SUMO1, SUMO2, SUMO3 and SUMO4 and the *Drosophila* SUMO protein Smt3. Positions that are identical in all sequences are shaded dark blue, and conserved positions are light blue. Positions that are identical in at least three of the SUMO proteins, but not in ubiquitin, are shaded dark pink. Positions that are identical in at least three of the SUMO proteins and in ubiquitin are shaded grey.

The SUMO protein was initially identified in *Saccharomyces cerevisiae* and all eukaryotes tested to date express at least one SUMO precursor protein. Among the species expressing only a single SUMO protein are yeast, *C. elegans*, and *Drosophila*. *Drosophila* Smt3 is closely related to the vertebrate homologue SUMO3 and is expressed throughout development, being more prominent during embryogenesis and in adult females (Huang *et al.*, 1998; Lehembre *et al.*, 2000; Long and Griffith, 2000;

Ohsako and Takamatsu, 1999). Other eukaryotes, including higher plants and vertebrates, express several SUMO proteins. For example, mammals express three SUMO proteins that can be divided into two families, SUMO1 and SUMO2/3. Although SUMO2 and -3 are very similar (97% sequence identity for human proteins) and currently cannot be distinguished by antibodies, SUMO1 is quite divergent (47% sequence identity between mature human SUMO1 and SUMO2). Additionally, a fourth SUMO paralog, very similar to SUMO2/3, has been described in human cells (Guo *et al.*, 2004), but its role is still uncertain.

Although SUMO1 and SUMO2/3 proteins are activated and conjugated by the same enzymes, SUMO1 appears to have a partially distinct function from SUMO2/3, which are assumed, at present, to be functionally identical. Cells contain a large pool of free, unconjugated SUMO2/3, but there is virtually no pool of free SUMO1; at any given time, the vast majority of SUMO1 is conjugated to other proteins (Matunis *et al.*, 1996; Saitoh and Hinchey, 2000). Furthermore, conjugation of SUMO2/3 is strongly induced in response to various stresses, but SUMO1 conjugation is not (Saitoh and Hinchey, 2000). Thus, one function of SUMO2/3 may be to provide a reservoir of free SUMO for stress responses. There is also evidence that different SUMOs are used preferentially for different substrates. Another difference between SUMO1 and SUMO2/3 is that SUMO2/3 contain ΨKXE sequences in their N-terminal extensions, which can serve as SUMO attachment sites, thereby allowing formation of poly-SUMO chains (Tatham *et al.*, 2001).

Sumoylation is essential for most organisms with the notable exceptions of *Schizosaccharomyces pombe* (Tanaka *et al.*, 1999) and the filamentous fungus *Aspergillus nidulans* (Wong *et al.*, 2008), for which SUMO deletion mutants display severe growth defects. Whether individual SUMO proteins are essential depends on whether their loss can be compensated by others. SUMO1 is not essential in mouse, suggesting that most functions can be fulfilled by SUMO2/3; however, its deletion can lead to severe phenotypes that seem to be background specific (Alkuraya *et al.*, 2006; Evdokimov *et al.*, 2008; Wang *et al.*, 2011; Zhang *et al.*, 2008). Although sumoylation is indispensable for early zebrafish development, all three SUMO paralogs are functionally redundant (Yuan *et al.*, 2010). By contrast, in *Arabidopsis thaliana*, which encodes eight SUMO genes (Miura *et al.*, 2007), simultaneous disruption of AtSUMO1 and -2 (orthologs of human SUMO2 and -3) is embryonically lethal (Saracco *et al.*, 2007).

4.1.2 The SUMO E1 enzyme

The SUMO E1 enzyme is composed of two subunits, SUMO-activating enzyme subunit 1 (SAE1; also known as Aos1) and SUMO-activating enzyme subunit 2 (SAE2; also known as Uba2). Both subunits are conserved from yeast to human, including *Drosophila* (**Table 3**). Similar to other E1s, the SAE1/SAE2 E1 activates SUMO's carboxy-terminal part in a two-step reaction that involves ATP hydrolysis. Formation of a SUMO adenylate is followed by dramatic E1 active-site remodelling and subsequent SAE2-SUMO thioester bond formation. Upon interaction of the thioester-charged E1 enzyme with Ubc9, which binds to SAE2's ubiquitin fold domain, SUMO is transferred to the E2 enzyme.

Most organisms contain a single SUMO-activating enzyme, which is required for conjugation of all SUMO variants to all substrates. Interestingly, the SUMO E1 is a heterodimer, whereas the ubiquitin E1 is a monomer, but both components of the SUMO enzyme are related to the ubiquitin enzyme. SAE1 resembles the amino-terminal part of the ubiquitin E1, while SAE2 corresponds to the carboxy-terminal part and contains the active site cysteine (Desterro *et al.*, 1999; Johnson *et al.*, 1997; Okuma *et al.*, 1999). Although the two-subunit structure of the SUMO E1 suggests that SAE1 and SAE2 might function or be regulated separately, all cellular SAE2 and SAE1 is found in the heterodimer (Azuma *et al.*, 2001).

4.1.3 The SUMO E2 enzyme

The single E2 conjugating enzyme, called Ubc9 in yeast and vertebrates, is highly conserved from yeast to human and plays an essential role in sumoylation that exceeds the role of most ubiquitin E2 enzymes. In addition to providing activated SUMO, Ubc9 is directly involved in the selection of many specific SUMO targets. These targets carry a sumoylation consensus site (described in section 4.1 of this *Introduction* and **Table 2**) to which Ubc9 can bind directly, albeit with low affinity. As a consequence, the targets can be sumoylated with E1 and high concentrations of Ubc9 *in vitro*. However, very few proteins are known whose sumoylation is efficient in the absence of E3 ligases; a well-known exception is RanGAP1, which binds Ubc9 with high affinity owing to an additional interaction surface (Bernier-Villamor *et al.*, 2002).

In *Drosophila*, the ortholog of the E2 enzyme is called Lesswright (Lwr) and it has high homology to Ubc9s identified in human (83% identical) (Long and Griffith, 2000). Like other orthologs of the vertebrate sumoylation machinery, such as Aos1 or

Uba2, Lwr and the SUMO protease Ulp1 (**Table 3**), they are enriched during embryogenesis, preferentially in the CNS and gonads (Donaghue *et al.*, 2001; Lehembre *et al.*, 2000; Shigenobu *et al.*, 2006; Shih *et al.*, 2002). They are all also preferentially expressed in females (Proschel *et al.*, 2006) and in undifferentiated tissues (Jasper *et al.*, 2002).

4.1.4 SUMO E3 ligases

E3 ligases catalyse the transfer of ubiquitin or ubiquitin-related proteins from an E2 enzyme to a target lysine residue (Hershko and Ciechanover, 1998). They mediate or stabilize the interaction of the target with the charged E2, and they lock the flexible Ubl-E2 thioester bond in an orientation that is favourable for nucleophilic attack by the target lysine. Up to now, SUMO E3 ligase activity has been clearly demonstrated for two types of proteins, Siz/PIAS (SP) E3 ligases and the nucleoporin RanBP2. Several other proteins have been identified that enhance sumoylation, like the human Polycomb protein Pc2/CBX4, the topoisomerase I–binding RING finger protein Topors or the tumour suppressor p14 Arf, but further investigation is needed to define them as *bona fide* E3 ligases.

PIAS proteins are evolutionarily conserved in eukaryotes, and in mammals they are encoded by four genes, PIAS1, PIASx (PIAS2), PIAS3 and PIASy (PIAS4). Other more distantly related mammalian homologues of PIAS proteins are the human proteins Zimp7 and Zimp10 (Beliakoff and Sun, 2006). Homologues of mammalian PIAS proteins are found in non-vertebrate animal species, plants and yeast, like proteins Siz1 and Siz2/Nfi1 *S. cerevisiae* (Hari *et al.*, 2001; Johnson and Gupta, 2001). In *Drosophila* there is one PIAS ortholog called Su(var)2–10 (suppressor of variegation 2–10) (**Table 3**), also known as dPIAS, which is involved in the maintenance of the proper chromosomal structure and chromosomal inheritance (Hari *et al.*, 2001; Le *et al.*, 2004), heterochromatin-induced gene silencing. (Hari *et al.*, 2001; Mohr and Boswell, 1999; Reuter and Wolff, 1981), and like its vertebrate homologues, acts as a negative regulator of the JAK (Janus kinase)/STAT pathway (Hombria and Brown, 2002). Sequence homology analyses have allowed the identification of an additional E3 ligase called Tonalli (Tna), ortholog of Zimp7 and Zimp10, that contains a SP-RING domain characteristic of this family of proteins (Gutiérrez *et al.*, 2003).

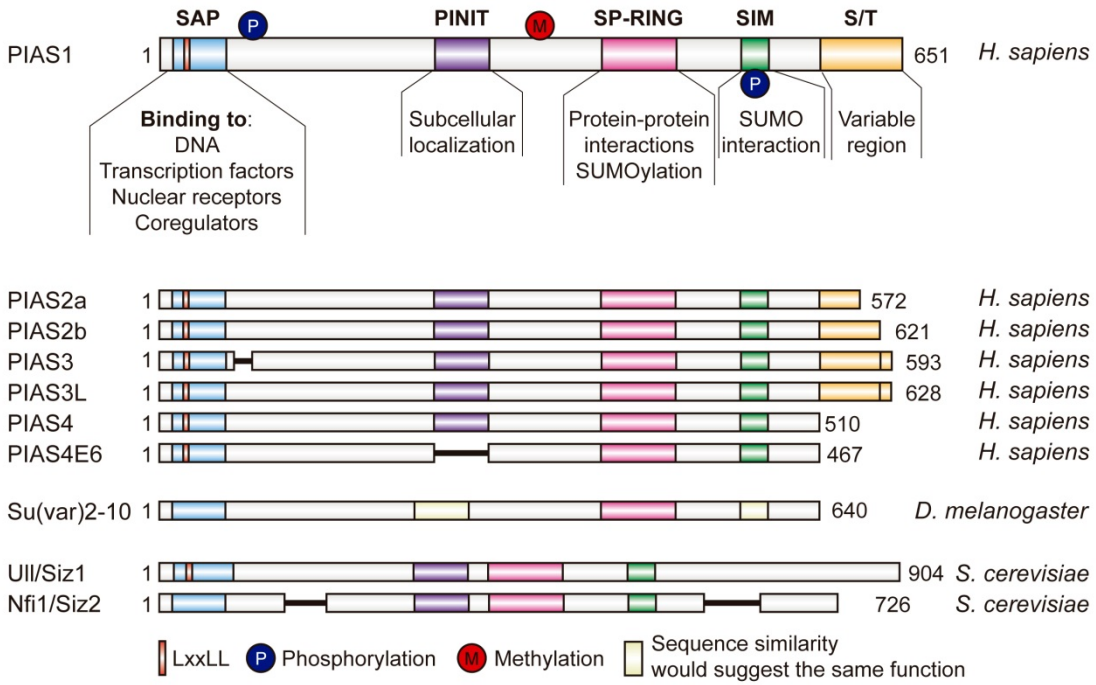


Figure 12. Schematic structures of PIAS proteins. The domain structures and post-translational modifications of the different human PIAS proteins are illustrated. Orthologs PIAS proteins from *S. cerevisiae* and *Drosophila* are also shown. Numbers on the right depict the aa in each PIAS protein. Adapted from Rytinki *et al*. (2009).

Mammalian PIAS proteins share a high degree of sequence homology within their ~430 amino-terminal aa. Overall, five different domains or motifs have been identified in their sequence: an amino-terminal SAP (scaffold attachment factor-A/B, acinus and PIAS) domain, a PINIT motif, a RING-type zinc-binding structure, a SIM (SUMO interacting motif) and a serine/threonine-rich C-terminal region (S/T) (Duval *et al.*, 2003; Hochstrasser, 2001; Minty *et al.*, 2000; Tan *et al.*, 2002). The most conserved of these regions are the SAP domain and the central RING-type zinc finger. Non-vertebrate PIAS orthologs contain essentially the same motifs and domains as their mammalian counterparts (**Figure 12**). The SAP domain is found in many chromatin-associated proteins and is involved in sequence- or structure-specific DNA-binding (Aravind and Koonin, 2000). Between the SAP and RING domain, there is the PINIT motif that, at least for PIAS3L, has been shown to be essential for the nuclear retention (Duval *et al.*, 2003). In the central part of PIAS proteins, there is a cysteine rich region forming a putative RING-type of zinc finger termed Siz/PIAS RING (SP-RING) (Hochstrasser, 2001). The SP-RING domain interacts with the SUMO E2 enzyme Ubc9 and is essential for the PIAS E3 function both *in vivo* and *in vitro* (Johnson and Gupta, 2001; Kotaja *et al.*, 2002; Sachdev *et al.*, 2001; Takahashi *et al.*,

2001a; Takahashi *et al.*, 2001b). They are thought to promote sumoylation by acting as adaptors that stabilize the interaction between the SUMO thioester-loaded Ubc9 and the acceptor protein. Between the SP-RING and the S/T-rich regions, PIAS proteins harbour the SIM. This motif allows non-covalent interaction with SUMO proteins and its structure and function will be discussed later on in this *Introduction*. SIMs are found in a large number of nuclear proteins, typically in those with known roles in sumoylation. The main differences between PIAS proteins lie in their 100–450 residue carboxy-terminal tails, which share no sequence similarity with each other or with other known proteins. It is likely that these carboxy-terminal domains interact with specific substrates.

By analogy with the ubiquitin system, the purpose of the different PIAS proteins may be to sumoylate different substrates, but currently the only clear example of this is the specificity of Siz1 for septins and PCNA. Sumoylation of many vertebrate-derived substrates can be stimulated by several different PIAS proteins, upon overexpression both in cells and *in vitro*. This may suggest that either PIAS proteins have overlapping substrate specificities or that *in vitro* assays do not faithfully reproduce physiological substrate selection mechanisms. Another function of the different PIAS proteins may be to promote attachment of the different SUMO isoforms. PIASy preferentially conjugates SUMO2, rather than SUMO1, to the transcription factors LEF1 and GATA-2, and it strongly enhances overall SUMO2 conjugation (Chun *et al.*, 2003; Sachdev *et al.*, 2001). It is also not clear that all PIAS effects are mediated by SUMO conjugation. In particular, PIAS proteins inhibit binding of STAT transcription factors to DNA *in vitro*, and there is no evidence that this effect involves SUMO (Chung *et al.*, 1997; Liu *et al.*, 1998; Rogers *et al.*, 2003; Ungureanu *et al.*, 2003).

A second type of SUMO E3 ligases consists of the 358-kDa RanBP2 (also known as Nup358) (Pichler *et al.*, 2002; Pichler *et al.*, 2004), which localizes to the cytoplasmic fibrils of the nuclear pore and contains several types of functional domains (Pichler *et al.*, 2002; Wu *et al.*, 1995; Yokoyama *et al.*, 1995). Its vertebrate specific E3 ligase region, which is natively unfolded, shows no obvious homology to other proteins. In addition to having the capacity to act as an E3 in the sumoylation of several proteins, including RanGAP1, RanBP2 can form a stable trimeric complex with SUMO-RanGAP1 and Ubc9, and thus it is responsible for the localization of SUMO-RanGAP1 to the nuclear pore (Matunis *et al.*, 1998; Saitoh *et al.*, 1997).

4.1.5 SUMO proteases

SUMO proteases are required for two tasks, maturation of SUMO proteins (C-terminal hydrolase activity) and removal of SUMO from targets (isopeptidase activity). Even though SUMO maturation is essential for sumoylation, there is currently no clear evidence that this is a rate-limiting or regulated event. By contrast, desumoylation rates are key determinants for steady-state sumoylation of individual target proteins.

All known SUMO-cleaving enzymes contain a ~200 aa C-terminal domain (the Ulp domain), which has the SUMO cleaving activity (Mossessova and Lima, 2000). The Ulp domain does not share sequence similarity with the enzymes that cleave ubiquitin. Instead, it is distantly related to a number of viral proteases (Li and Hochstrasser, 1999; Strunnikov *et al.*, 2001). The different SUMO-cleaving enzymes have varying N-terminal domains, which promote interactions with specific targets and mediate distinct intracellular localizations (Hang and Dasso, 2002; Li and Hochstrasser, 2003; Panse *et al.*, 2003; Zhang *et al.*, 2002).

So far, a single gene family that encodes SUMO-specific Cys proteases has been identified. The corresponding proteins are Ulp1 and Ulp2 in yeast (Li and Hochstrasser, 1999, 2000), and the six Ulp homologues in humans are called sentrin specific proteases (SENP1, -2, -3, -5, -6, and -7). Members of the SENP family differ in their activity in maturation and isopeptide cleavage and also in their activity towards different SUMO paralogs: for example, SENP3 and SENP5 preferentially deconjugate SUMO2/3 from substrates (Di Bacco *et al.*, 2006; Gong and Yeh, 2006). Finally, SENPs vary in their predominant *in vivo* localization: *S. cerevisiae* Ulp1 and mammalian SENP2 are enriched at nuclear pore complexes (Hang and Dasso, 2002; Li and Hochstrasser, 2003; Zhang *et al.*, 2002); SENP5 is enriched in the nucleolus (Di Bacco *et al.*, 2006; Nishida *et al.*, 2000); SENP1 appears to shuttle between the cytoplasm and the nucleus (Bailey and O'Hare, 2004; Gong *et al.*, 2000); and SENP6 has been reported to be both in the nucleus (Mukhopadhyay *et al.*, 2006) and the cytoplasm (Kim *et al.*, 2000). In addition to these SUMO-specific proteases, there are others like DeSI-1 (desumoylating isopeptidase-1) and DeSI-2, or USPL1 that belong to protein families that also include ubiquitin proteases (Schulz *et al.*, 2012; Shin *et al.*, 2012).

The *Drosophila* genome encodes at least two SUMO proteases, termed Ulp1 and Ulp2. The first one, Ulp1 (**Table 3**), has been extensively characterized, (Smith *et al.*, 2004). It is primarily associated with the nucleoplasmic face of the nuclear pore complex where it may deconjugate proteins as they exit the nucleus thus serving as a

molecular switch to control the biochemical properties of a protein as a function of its subcellular localization.

4.2 Non-covalent interactions with SUMO

Sumoylation frequently modifies the range of interactions that a target protein can establish in three different ways. First, because sumoylation targets lysine residues of substrates, it can potentially compete with other lysine-directed post-translational modifications like acetylation or ubiquitination. Secondly, because SUMO is a bulky modification, sumoylation can also potentially interfere with protein-protein interactions through steric hindrance. The third mechanism of SUMO action, and by far the most commonly used, is the promotion of protein-protein interactions (Bergink and Jentsch, 2009; Geiss-Friedlander and Melchior, 2007). Several studies indicate that the SUMO modification is frequently used to strengthen the interaction of proteins that exhibit already weak affinities for each other on their own. The molecular basis for this property is that a particular surface on SUMO (comprising the aa 35–55 of *S. cerevisiae* SUMO) has moderate affinities to sequence motifs called SIMs, which are found on partner proteins (Kerscher, 2007). This property of SUMO is not only employed to foster intermolecular interactions in protein complexes or networks but also promotes intramolecular interactions in cases in which a SUMO modification of a substrate binds to a SIM on the same protein.

The SIM consists of a short stretch of branched hydrophobic residues [typically (Val/Ile)-X-(Val/Ile)-(Val/Ile) or (Val/Ile)-(Val/Ile)-X-(Val/Ile)] that is flanked N- or C-terminally by serine residues and/or by a stretch of acidic residues (Hecker *et al.*, 2006; Song *et al.*, 2004; Song *et al.*, 2005). The interaction between SUMO and SIMs involves both hydrophobic and hydrophilic interactions. The latter comprises lysine residues in SUMO and acidic or serine residues flanking the hydrophobic SIM core. However, similar to other motifs that sense post-translational modifications (Husnjak and Dikic, 2012; Seet *et al.*, 2006), the affinity between SUMO and SIMs is typically rather low (in the low-micromolar range). Then, considering such low affinity, how can a SIM contribute to specific interactions of sumoylated proteins? One mechanism is that the SIM cooperates with a second low-affinity interaction site, one that is specific for the given target. Indeed, this is how sumoylated PCNA recruits the DNA helicase Srs2. Srs2 harbours two closely spaced motifs in its C-terminal domain, one that interacts with SUMO (a SIM), and one that interacts with PCNA; both are required to specifically

recognize sumoylated PCNA (Armstrong *et al.*, 2012). Additionally, several works have shown that phosphorylation of serine residues within the SIM stimulate SUMO-SIM interactions (Hecker *et al.*, 2006; Stehmeier and Muller, 2009), while SUMO acetylation abolishes binding to some SIM-containing proteins, including PML, Daxx, and PIAS family members (Ullmann *et al.*, 2012). These two novel means of SUMO-SIM interaction regulation greatly expand the possibilities for target-specific effector interactions.

4.3 Biological functions of sumoylation

Since the identification of the first SUMO-modified protein, RanGAP, (Matunis *et al.*, 1996), a large number of proteins have been shown to be post-translationally modified by SUMO, and new substrates of SUMO modification continue to be identified at a rapid pace. Many of the known targets of sumoylation are nuclear proteins with important roles in regulating transcription, chromatin structure, and DNA repair. Furthermore, the nuclear targets of many signalling pathways including TGF-β, wnt, and cytokines are post-translationally modified by SUMO. Some of the many proteins now known to be modified by SUMO are shown in **Figure 13**.

4.3.1 Transcriptional regulation

Many of the SUMO-modified proteins identified to date are promoter-specific transcription factors, co-activators, or co-repressors. Although the effects of SUMO modification on transcription factor activity are varied, in the majority of cases that have been described to date, attachment of SUMO appears to repress the activity of transcriptional activators (reviewed in Gill, 2003; Verger *et al.*, 2003). The transcription factor Sp3, for example, has been shown to be SUMO modified *in vivo* and removal of SUMO by mutation of the Sp3 acceptor lysines or co-transfection with a SUMO protease dramatically increased transcriptional activity of Sp3 (Ross *et al.*, 2002; Sapetschnig *et al.*, 2002). The major site of SUMO modification in Sp3 lies within a previously defined inhibitory domain. Similarly, the sites of SUMO attachment in many transcription factors, including C/EBP proteins, Elk-1, c-Myb, and steroid hormone receptors, have also been mapped to regions previously shown to function as inhibitory domains (Abdel-Hafiz *et al.*, 2002; Bies *et al.*, 2002; Kim *et al.*, 2002a; Poukka *et al.*,

2000; Subramanian *et al.*, 2003; Tian *et al.*, 2002; Yang *et al.*, 2003). Thus, SUMO modification may function to repress activation by many transcription factors.

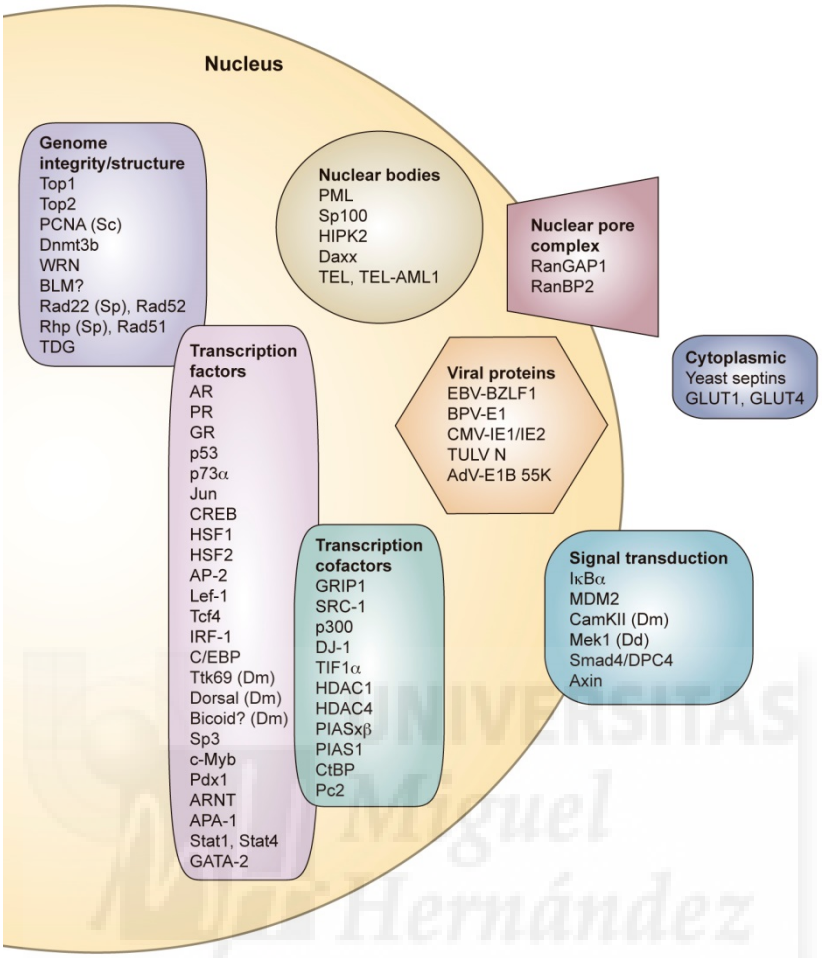


Figure 13. SUMO substrates grouped by function. Budding yeast, fission yeast, *Drosophila* and *Dictyostelium* proteins are indicated by (Sc), (Sp), (Dm) and (Dd), respectively; all other are mammalian proteins. Adapted from Seeler and Dejean (2003).

One mechanism to negatively regulate transcription factor activity is by altering their interaction with DNA, chromatin, or other proteins. Sumoylation has been suggested to stimulate DNA-binding by HSF (Goodson *et al.*, 2001; Hong *et al.*, 2001), while in the case of SATB2 sumoylation may reduce chromatin association of this factor (Dobreva *et al.*, 2003). For several transcription factors, however, including LEF, Sp3, and Ttk69, SUMO modification does not alter DNA-binding activity *in vitro* (Lehembre *et al.*, 2000; Sachdev *et al.*, 2001; Sapetschnig *et al.*, 2002). So for the moment, the effects of SUMO modification on association of transcription factors with DNA *in vivo* remain largely unknown.

Current findings support the view that, in many cases, SUMO modification inhibits transcription factor activity by promoting interactions with proteins that repress transcription. In several contexts SUMO-dependent repression of transcription has been observed, suggesting that SUMO-modified transcription factors bound at the promoter actively participate in repression (Holmstrom *et al.*, 2003; Kim *et al.*, 2002a; Ross *et al.*, 2002; Yang *et al.*, 2003). Furthermore, recruitment of SUMO to a promoter, in the context of a Gal4-SUMO fusion, is sufficient for repression, suggesting that SUMO itself can directly bind transcriptional co-repressors (Ross *et al.*, 2002; Yang *et al.*, 2003). Consistent with this hypothesis, SUMO modification of the transcription factor Elk-1 has been shown to increase association with HDAC2 and this has been correlated with a decrease in histone acetylation at the repressed promoter (Yang and Sharrocks, 2004). Similarly, SUMO-modified p300, but not unmodified p300, has been shown to bind HDAC6 (Girdwood *et al.*, 2003). Other co-repressors, in addition to the HDACs, may also be recruited to promoters dependent on SUMO.

Studies of the mechanisms by which SUMO modification regulates transcription-factor activity are complicated by the fact that in many cases not only the activator but also its associated co-activators, and/or co-repressors, are modified by SUMO. It is possible that association of multiple SUMO-modified proteins in a complex may amplify effects of experimental alterations in the sumoylation machinery or mask effects of mutation of individual SUMO acceptor lysines. If SUMO functions to promote interactions important for transcriptional repression, it may not matter exactly which subunit in a transcription factor complex is SUMO modified, so much as that SUMO is conjugated to one or more subunits in the complex. According to this view, it is modification of the complex *per se* that is important for regulation (Jackson, 2001).

But even if it is sometimes believed that the prominent effect of sumoylation on transcription is repression, growing lists of transcription repressors that are inhibited by sumoylation and transcription activators that are activated by sumoylation indicate a more complicated role of sumoylation in the regulation of transcription. One example of this are the heat shock transcription factors HSF1 and HSF2, which increase their DNA-binding activity upon sumoylation, and modification by mutation decreases HSF1 transcriptional activity (Goodson *et al.*, 2001; Hong *et al.*, 2001).

4.3.2 DNA repair

The genome in the cell is constantly damaged by extrinsic and intrinsic factors. To survive, eukaryotic organisms have evolved highly conserved DNA damage repair

mechanisms to ensure that the genome is copied faithfully during each cycle of cell division. SUMO modification has been shown to regulate the subnuclear localization, protein–protein interactions, and activity of many factors involved in maintenance of the genome, including yeast PCNA, the helicase WRN, Topoisomerases I and II, and the thymine-DNA glycosylase enzyme TDG (reviewed in Muller *et al.*, 2004).

Studies of the yeast proliferating cell nuclear antigen (PCNA) highlight one strategy by which SUMO modification can impact protein function: competition with other modifications for a common lysine. PCNA serves as a sliding processivity clamp for replicative DNA polymerases and plays a key role during post-replication repair. Modification of yeast PCNA at Lys164 by ubiquitin or SUMO has been suggested to direct PCNA for alternative functions (Hoege et al. 2002; Stelter and Ulrich 2003; Haracska et al. 2004). Ubiquitination of PCNA is induced by DNA damage and monoubiquitinated PCNA is important for translesion synthesis repair, whereas polyubiquitinated PCNA promotes error-free DNA repair (Hoege et al. 2002; Stelter and Ulrich 2003; Kannouche et al. 2004). SUMO competes for attachment to this lysine and genetic evidence indicates that SUMO conjugation inhibits damage-induced DNA repair and mutagenesis (Hoege *et al.*, 2002; Stelter and Ulrich, 2003). PCNA is sumoylated most heavily during the S phase of the cell cycle; this may suggest that sumoylation prevents inappropriate recruitment of post-replication repair enzymes during the wrong phase of the cell cycle. This seemingly controversial cross-talk between ubiquitination and sumoylation of PCNA suggests that modification of PCNA is critically fine-tuned and that the cross-talk appears to ensure the completion of post-replication repairs without yielding abortive recombination events.

Another mechanism by which SUMO modification can impact protein function is exemplified by the thymine-DNA glycosylase enzyme (TDG), which plays an important role in base excision repair in response to certain forms of DNA damage. TDG removes thymine and uracil from mismatched G-T and G-U base pairs. After removal of the mismatch, the TDG must be released from the apurinic (G:_) site for the downstream enzymes to restore G:C pairs. Sumoylation of TDG has been shown to help with this release by reducing TDG binding affinity to DNA (Baba *et al.*, 2005; Hardeland *et al.*, 2002; Steinacher and Schar, 2005).

4.3.3 Chromosome organization and function

Perfect copy of genetic materials to the daughter cells during cell division depends on precisely orchestrated chromosome dynamics including sister chromatid cohesion,

chromosome condensation, and segregation. Genetic studies of SUMO pathway function in model organisms have demonstrated the role of sumoylation in chromosome condensation, cohesion or mitotic (or meiotic) chromosome separation. For example, the budding yeast SUMO (Smt3) and desumoylase (Smt4) were initially identified as high-copy suppressors of the centromere-binding protein Mif2p/Cenp-C (Meluh and Koshland, 1995). Consistently, Smt3 was later identified as a chromosome cohesion defect gene (Andrews *et al.*, 2005). Similarly, disruption of the SUMOE2, E3 (SIZ1 or Mms21p), or desumoylase results in spindle defects in fruit flies, chromosome segregation defects in mice, and chromosome segregation, condensation or telomere defects in budding yeast (Apionishev *et al.*, 2001; Hari *et al.*, 2001; Nacerddine *et al.*, 2005; Strunnikov *et al.*, 2001; Zhao and Blobel, 2005). Mutation of the *Drosophila* PIAS, Su(var)2-10, leads to aberrant chromosome structure and mitotic defects (Hari *et al.*, 2001).

Among known substrates of sumoylation that are involved in these regulations are Cenp-C, topoisomerase II (top2), the cohesion protein Pds5 and nuclear pore complex protein RanGAP1. Recent studies have confirmed that Cenp-C is a target of SUMO1 and this protein plays a key role at centromeres for mitotic progression in human cell lines (Chung *et al.*, 2004; Everett *et al.*, 1999). Desumoylation of Top2p has been shown to play an active role in maintaining centromere cohesion in budding yeast, suggesting that its sumoylation inhibits the cohesion (Bachant *et al.*, 2002). Similarly, desumoylation of Pds5 appears to be required for cohesion maintenance, whereas its sumoylation peaks at anaphase and seems to be necessary for dissolution of cohesion during mitosis in budding yeast (Stead *et al.*, 2003). Finally, sumoylated RanGAP1 is targeted to the microtubule spindle and kinetochores to guide their attachment during mitosis in HeLa cells (Joseph *et al.*, 2004; Joseph *et al.*, 2002). Taken together, it seems that sumoylation promotes chromosome separation whereas desumoylation helps with cohesion.

4.3.4 Nuclear transport

Post-translational modification by SUMO has been shown to regulate subcellular localization of many targets including RanGAP, the first identified SUMO substrate (Matunis *et al.*, 1996). RanGAP is a small GTPase activating protein that plays a role in nuclear import. Unmodified RanGAP is cytoplasmic, whereas SUMO modified RanGAP is associated with the nuclear pore (Mahajan *et al.*, 1997; Matunis *et al.*, 1996). SUMO modification of RanGAP greatly increases its interaction with RanBP2, a component of

the nuclear pore complex and a SUMO E3 ligase (Pichler *et al.*, 2002). Although the RanGAP homolog in budding yeast is not sumoylated, the SUMO conjugation pathway has also been shown to play a role in nuclear trafficking in this organism (Stade *et al.*, 2002). In mammalian cells, SUMO modification of several substrates has been linked to nuclear import. In the case of the IκB kinase regulator NEMO, for example, fusion of NEMO to SUMO was sufficient for translocation to the nucleus (Huang *et al.*, 2003). Location of the RanBP2 SUMO E3 ligase at the nuclear pore may contribute to a broad role for SUMO in regulation of nuclear trafficking (reviewed in Melchior *et al.*, 2003).

Another well-characterized example of SUMO-dependent changes in subcellular localization comes from studies of the tumour suppressor PML. PML was originally identified as part of a fusion protein with RARα resulting from a chromosomal translocation associated with acute promyelocytic leukaemia. Wild-type PML, which is post-translationally modified by SUMO, is present in a subnuclear structure called the PML nuclear body. Many other SUMO-modified proteins including transcription factors, chromatin modifiers, and proteins involved in genomic maintenance have also been found in PML nuclear bodies (Zhong *et al.*, 2000b). Notably, both SUMO E3 ligases and SUMO-specific proteases have also been found to localize in PML nuclear bodies (Best *et al.*, 2002; Kotaja *et al.*, 2002; Sachdev *et al.*, 2001). Mutation of the SUMO acceptor lysines in PML or overexpression of a SUMO protease causes nuclear body components such as CBP or Sp100 to relocalize in the nucleus (Best *et al.*, 2002; Zhong *et al.*, 2000a). These findings suggest that SUMO-modified PML supports some protein–protein interactions important for assembly or stability of this subnuclear domain. Although the function of the PML nuclear body remains a subject of active investigation, disruption of PML nuclear bodies is associated with changes in cell proliferation, differentiation, and survival (Zhong *et al.*, 2000b).

Sumoylation also helps send nuclear proteins to the cytoplasm. The primary functional site of *Dictyostelium* MEK1 (dMEK1) is in the cytoplasm where it is required for aggregation and chemotaxis. Interestingly, the cytoplasmic localization of dMEK1 depends on its sumoylation, and its non-sumoylated form is predominantly present in the nucleus (Sobko *et al.*, 2002). Chemoattractant stimulation induces rapid sumoylation of dMEK1, its translocation from the nucleus to the cytosol and the leading edge of migrating cells. Sumoylation has also been shown to promote the transport of preribosomes from the nucleolus to the cytoplasm and it seems that dynamic sumoylation as well as desumoylation is required for this whole process (Panse *et al.*, 2006). Finally, Mdm2 has been recently linked to sumoylation-dependent p53 nuclear export (Carter *et al.*, 2007).

4.3.5 Non-nuclear functions

The large body of published work that focuses on the role of SUMO in transcription, DNA repair, nuclear bodies and nucleocytoplasmic transport might give the impression that sumoylation is restricted to the nuclear compartment. However, over the last years the work of several groups points to many roles of SUMO in the soluble phase of the cytoplasm, the plasma membrane, mitochondria and the endoplasmic reticulum.

Sumoylation has been related to mitochondrial dynamics through DRP1, a cytosolic dynamin-like GTPase that is involved in mitochondrial fission. Overexpression of SUMO1 or depletion of SENP5 by small interfering RNA results in fragmented mitochondria, which implies that reversible SUMO modification is necessary for maintaining the balance between mitochondrial fission and fusion (Harder *et al.*, 2004; Zunino *et al.*, 2007).

Another example of non-nuclear sumoylation target is the protein-tyrosine phosphatase-1B (PTP1B). PTP1B is a ubiquitously expressed enzyme that localizes to the cytoplasmic face of the endoplasmic reticulum and the nuclear envelope. PTP1B negatively regulates growth-factor signalling and cell proliferation by dephosphorylating key receptor tyrosine kinases. Insulin treatment stimulates the sumoylation of PTP1B, which in turn impairs its activity, suggesting a positive role for SUMO in receptor tyrosine kinase signalling (Dadke *et al.*, 2007).

Finally, SUMO was linked to channel and receptor regulation at the plasma membrane. Sumoylation appears to serve negative regulatory functions. However, molecular details are still scarce. The plasma membrane voltage-gated potassium channel Kv1.5 was subsequently found to be regulated through reversible sumoylation (Benson *et al.*, 2007). As Kv channels play crucial roles in the highly regulated electrical responses throughout the cardiovascular system, these findings might have far-reaching medical implications. Additional SUMO targets at the plasma membrane are the metabotropic glutamate receptor-8 (mGluR8) (Tang *et al.*, 2005) and the GluR6 subunit of kainate receptor (Martin *et al.*, 2007). The *in vivo* function of mGluR8 sumoylation is currently unknown. By contrast, sumoylation of GluR6 is induced in response to kainite, and this modification appears to be a prerequisite for kainite-induced endocytosis of the receptor (Martin *et al.*, 2007).

4.4 Role of sumoylation in the initiation and progression of cancer

As described earlier, SUMO substrates include many proteins with important roles in regulating cell proliferation and differentiation. This suggests that, like ubiquitination, altered sumoylation may contribute to disease onset or progression. In fact, over the past few years it has been shown that several diseases including pathogenic infection, cancer, and neurodegenerative disorders are associated with alterations in sumoylation.

Regarding cancer, the sumoylation system appears to be involved in many aspects related the oncogenic process, from initiation to metastasis, directly or indirectly. It has been shown that sumoylation is able to regulate the activity of several oncogenes and tumour suppressor genes implicated in cancer initiation and progression. For example, sumoylation of c-Jun antagonizes its phosphorylation and negatively regulates its activity, thereby limiting its oncogenic capacity (Muller *et al.*, 2000). In the case of Mdm2, sumoylation attenuates its ubiquitination increasing its E3 ligase activity towards p53, leading to degradation of the tumour suppressor (Buschmann *et al.*, 2000). p53 itself is among the better studied tumour suppressor proteins that are sumoylated, but the nature and significance of its sumoylation still remains unclear. There are studies indicating that p53 sumoylation enhances its transcriptional activity and apoptotic response, resulting in tumour suppression (Bischof *et al.*, 2006; Gostissa *et al.*, 1999; Muller *et al.*, 2000; Rodriguez *et al.*, 1999), while there are other works showing that PIAS proteins repress the transcriptional activity of p53 *in vitro* and *in vivo*, suggesting that sumoylation of p53 limits its transcriptional output (Schmidt and Muller, 2002), thereby limiting its possible contribution to the control of cell growth and/or death. Other oncogenic signalling pathways regulated by sumoylation include Wnt, NF-κB, nuclear receptor transcription factors and their co-regulators.

In addition, it has been shown that abolition of sumoylation leads to genomic instability, thus providing a cue to initiate malignant cell transformation. One well known example is the RanBP2-knockout mice model (Dawlaty *et al.*, 2008; Nacerddine *et al.*, 2005); that shows a dramatic defect in chromosome segregation. This defect is caused by a decrease of Topoisomerase IIa (Topo IIa) sumoylation by RanBP2, crucial for targeting the protein to inner centromeres where Topo IIa decatenates DNA for proper separation of sister chromatids in mitosis. RanBP2 hypomorphic mice demonstrated much higher incidence of tumour formation compared to wild-type when treated with a carcinogen. Also, spontaneous tumour occurred with higher frequency in the mutants and with shorter latency.

Sumoylation also plays a major role in establishment and maintenance of favourable environments for tumour development through the regulation of hypoxia responses and inflammatory status (Cheng *et al.*, 2007; Kim and Baek, 2006; Liu *et al.*, 2007a). On example is sumoylation of hypoxia-inducible factor 1α (HIF-1α), which provides this molecule with a new interacting motif for binding to a ubiquitin E3 ligase complex in the nucleus upon hypoxic challenges. Sumoylated HIF-1α is subjected to ubiquitination and thereby degraded by the proteasome. In the presence of SENP1, HIF-1α is protected from degradation by a desumoylation process and generates a hypoxic response. The work of (Cheng *et al.*, 2007) showed that SENP1 deficiency caused a critical defect in response to developmentally induced hypoxia. There is the possibility that SENP1 deficiency can similarly restrict tumour-induced hypoxic responses, such as vascular genesis and metastasis.

Regarding inflammatory status, sumoylation appears to be involved in the restriction of overinflammation by stabilizing IκB protein (Desterro *et al.*, 1998), by potentiating transrepression of certain nuclear receptors (Ghisletti *et al.*, 2007; Pascual *et al.*, 2005), or by negatively regulating cytokine signalling via PIAS1 protein (Liu *et al.*, 2004; Liu *et al.*, 2005), although sumoylation mediates NF-κB activation under several stress conditions (Huang *et al.*, 2003; Mabb *et al.*, 2006; Wuerzberger-Davis *et al.*, 2007). Thus, it is possible that the malfunction of the sumoylation system or overexpression of SENPs in the cancer microenvironment may positively influence cancer growth and propagation.

Finally, several studies also link sumoylation to tumour metastasis. In one example, the chromatin-remodelling protein reptin helps recruit the co-activator Tip60 to facilitate the transcription of the tumour metastasis suppressor KAI1. When sumoylated, reptin loses this function and instead facilitates β-catenin-mediated repression of the KAI1 promoter (Kim *et al.*, 2006). In the second example, sumoylation of the TGF-β receptor 1 (TβR1) enhances its affinity to Smad proteins, allowing more efficient phosphorylation and activation of Smad2/3 in response to TGF-β (Kang *et al.*, 2008). In these two cases, sumoylation appears to enhance metastatic potentials of malignant cells via the repression of metastatic suppressor KAI1 expression and the elevation of metastatic activity of TGF-β. Thus, the shift to excess sumoylation in certain cellular contexts might contribute to the gain of metastatic ability in malignant cells.

5. DmTAF3, a component of the basal transcriptional machinery

bip2 (or *DmTAF3*) was first identified as *Bric à brac Interacting Protein 2*, in a two-hybrid screen for interacting partners of *Bric à brac* 1 (BAB1) and *Bric à brac 2* (BAB2) (Pointud *et al.*, 2001) and demonstrated to be dTAFII155, the *Drosophila* homologue of yeast TAFII47 (Gangloff *et al.*, 2001a). From now on I will use the name *DmTAF3* instead of *bip2*, according to the unified nomenclature proposed in the work of Tora (2002).

The *DmTAF3* gene is 6.5 kb in length, and encodes a single transcript of 4.5 kb that is detected throughout *Drosophila* development. High levels of the *DmTAF3* transcript are detected in the early embryo (Pointud *et al.*, 2001), suggesting that it might correspond to maternally deposited transcript. The *DmTAF3* transcript is also detected in 2-24 hours old embryos, third instar larvae, prepupa, pupa, and adult males and females (Pointud *et al.*, 2001).

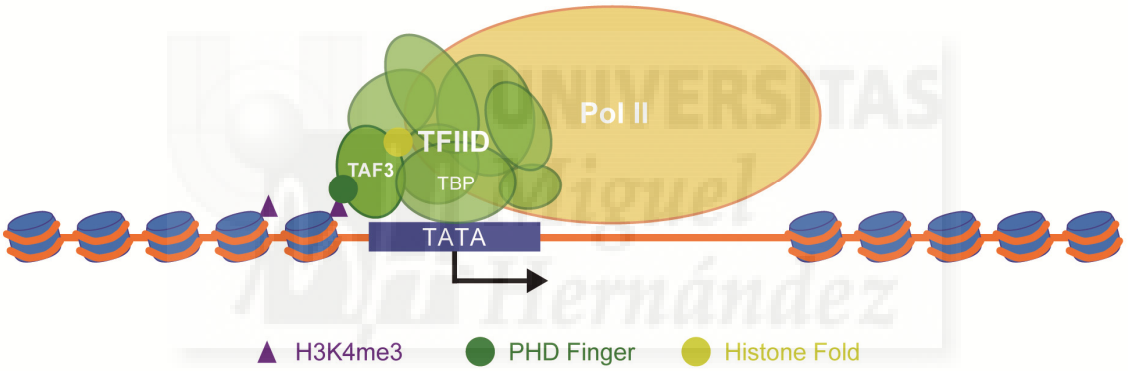


Figure 14. TAF3 and the TFIID complex are required for RNA polymerase II initiation of transcription. TAF3 belongs to the TFIID complex and, through recognition of the open chromatin mark H3K4me3 mediated by its PHD domain, directs this complex to chromatin regions that are ready to be transcribed, aiding in the initiation of RNA polymerase II-mediated transcription.

The *DmTAF3* gene encodes a putative protein of 1406 aa residues with a predicted molecular weight of 155 kDa that appears at 160 kDa in SDS-PAGE. The protein can be detected in extracts from 0-2 hours old embryos, 2-24 hours old embryos, third instar larvae, and adult ovaries, with a nuclear distribution (Pointud *et al.*, 2001). *DmTAF3* contains several known motifs implicated in chromatin mediated transcriptional regulation (**Figure 17B**). In its amino-terminal region presents a Histone fold domain (aa 1–75), a motif involved in the heterodimerization of several TAFIIs that can form histone-like pairs (Birck *et al.*, 1998; Gangloff *et al.*, 2001b; Gangloff *et al.*, 2000; Nakatani *et al.*, 1996; Xie *et al.*, 1996). In the carboxy-terminal part, it contains a

PHD finger (Plant Homeo Domain-finger) (aa 1342–1392), an atypical class of zinc fingers that have been suggested to be involved in DNA-binding and/or protein-protein interaction (Linder *et al.*, 2000). DmTAF3 also contains a putative A/T hook motif (aa 574–586) (Aravind and Landsman, 1998).

DmTAF3 belongs to the TAFII group, which are TATA binding protein (TBP) associated proteins that also comprise TFIID, one of the general factors required for initiation of transcription by RNA polymerase II (**Figure 14**). TAFIIs are thought to contribute to TFIID function through contacts with other transcription factors, histones and/or DNA (Chen and Hampsey, 2002). In the case of mammalian TAF3, it has been shown that TFIID directly binds to the H3K4me3 mark through the PHD-finger of TAF3 (Vermeulen *et al.*, 2007). H3K4me3 is associated not only with actively transcribed genes, but also with silent developmental genes that are poised for activation upon embryonic stem cell differentiation (Bernstein *et al.*, 2006; Mikkelsen *et al.*, 2007). Later on, Robert Tjian's group found that, while other TFIID subunits are destroyed during myogenesis, TAF3 is selectively retained in myotubes in a specialized complex with TBP-related factor 3, TRF3 (Deato *et al.*, 2008). Thus, these studies establish that TAF3, either as a subunit of TFIID or in association with other potential partners (e.g., TRF3) may regulate transcription by targeting cell-type-specific complexes to core promoters, including those that are marked by H3K4me3. Additionally, TAF3 binds the architectural protein CTCF to mediate regulatory interactions between distal CTCF/cohesin bound regions and proximal promoters, being this a crucial element for early lineage segregation during stem cell differentiation (Liu *et al.*, 2011). Regarding *Drosophila* TAF3, it binds on polytene chromosomes to many sites representing mainly decondensed transcriptionally active chromatin, partially colocalizing with BAB2 (Gangloff *et al.*, 2001a). It also can act as an Antennapedia co-factor for ectopic wing formation, linking *Antennapedia* to an activating TFIID complex and to the basal transcriptional machinery (Prince *et al.*, 2008). Finally, it has been shown that both human and *Drosophila* TAF3 act as a negative regulators of p53 transcription activation function (Bereczki *et al.*, 2008).

6. MEP-1 and its role as sumoylation-dependent transcriptional repressor

The *MEP-1* gene was first identified in *C. elegans* as having an essential role in the repression of genes involved in sexual determination (Belfiore *et al.*, 2002; Kasturi *et al.*; Leight *et al.*, 2005; Unhavaithaya *et al.*, 2002). Later on, a *Drosophila* homolog was found (gene product CG1244, protein accession number AAF47669), which shares only 20% overall aa identity with *C. elegans* MEP-1, but has conserved residues throughout its entire length, and retains the general domain architecture of the worm protein (Belfiore *et al.*, 2002).

In *Drosophila*, *MEP-1* encodes 7 isoforms that differ in their 5'UTR. MEP-1 mRNA expression is high in the embryo and in the early and middle stages of the metamorphosis. In the adult, it's expressed at high levels in the brain, in the ovary and in the larval CNS. On the contrary, the expression on the testis is very low, hinting a possible role in germline specification as occurs in *C. elegans*. All MEP-1 isoforms give rise to the same 1152 aa polypeptide. The protein has 7 putative zinc fingers, grouped in two clusters of three and four zinc fingers respectively, with a glutamine-rich region between both of them (see **Figure 22** in the *Results* section). The protein shows a nuclear distribution (Belfiore *et al.*, 2002).

After its discovery in *C. elegans*, this protein was termed MEP-1 due to its ability to bind the MOG proteins, and because P-granules (RNA-rich cytoplasmic granules present in germ cells), accumulate ectopically in *mep-1* deficient animals (MOG interacting and ectopic P-granules) (Unhavaithaya *et al.*, 2002). MEP-1 acts together with the MOG proteins to repress *fem-3* mRNA (Belfiore *et al.*, 2002), a sex-determining gene that promotes male development and its involved in the switch from spermatogenesis to oogenesis during *C. elegans* development (Barton *et al.*, 1987; Hodgkin *et al.*, 1985). MEP-1 also associates with the protein Mi-2, the ATPase subunit of the conserved Nucleosome Remodelling and Deacetylation (NuRD) complex (Unhavaithaya *et al.*, 2002), and this association is conserved in *Drosophila* (Kunert *et al.*, 2009; Stielow *et al.*, 2008). Both in *C. elegans* and in *Drosophila* these proteins have been shown to form a repressor complex recruited by DNA-bound SUMO modified transcription factors that likely promote the formation of repressive chromatin structures (Kunert *et al.*, 2009; Leight *et al.*, 2005; Stielow *et al.*, 2008). However, the work of Kunert *et al*., shows that this MEP-1/Mi-2 complex is distinct from NuRD and is able to promote these changes in chromatin structure without the NuRD deacetylase subunit activity. In support of the role of MEP-1 as a sumoylation dependent transcriptional repressor, in the work from Leight *et al*., (2005), it was shown that

sumoylation of the *C. elegans* transcription factor LIN-1 mediates interaction with MEP-1, allowing its genetic repression. Additionally, MEP-1 has been identified in a screen looking for genes involved in sumoylation dependent transcriptional repression (Stielow *et al.*, 2008). This screen was done using *Drosophila* Kc167 cells and, due to the lack of MEP-1 mutations in *Drosophila*, the biological function of this gene is still speculative.

7. The proteasome: structure, function and role in the cell

The ubiquitin-proteasome system (UPS) is the major cellular system for the regulated degradation of proteins in the cell. Substrate degradation by the UPS is rigorously controlled by tagging of the 76-residue ubiquitin to surface-exposed lysine residues of redundant or misfolded proteins. This process is carried out by a cascade of enzymes called E1, E2, and E3, (**Figure 15**) that activate free ubiquitin and carry it to the target protein, similar to what was already described for the sumoylation pathway in a previous section of this *Introduction*. Once the initial ubiquitin is attached, polyubiquitination of the substrate occurs through the sequential transfer of additional ubiquitin molecules, forming ubiquitin chains that ‘‘flag’’ the protein for destruction by the 26S proteasome, often called ‘‘the proteasome’’, a multicatalytic enzyme complex expressed in the nucleus and cytoplasm of all eukaryotic cells (Pickart, 2001; Wilkinson, 1999).

To date, only one known E1 enzyme exists, but about 20 E2 enzymes have been identified in yeast and humans, and an even larger number (perhaps hundreds) of E3 ligases exist (Pickart, 2001). One of the main groups of E3 ligases is the SCF complex (Skp1, Cullin, F-box) (reviewed in Dye and Schulman, 2007), which promotes ubiquitination by positioning the activated E2 in close proximity to the substrate. The SCF complex is the prototype of an emerging family of cullin-based ligases including, among others, the Cul3-based ligases, that have been shown to use BTB proteins as substrate-specific adaptors (Furukawa *et al.*, 2003; Geyer *et al.*, 2003; Pintard *et al.*, 2003; Xu *et al.*, 2003 and reviewed in Pintard *et al.*, 2004).

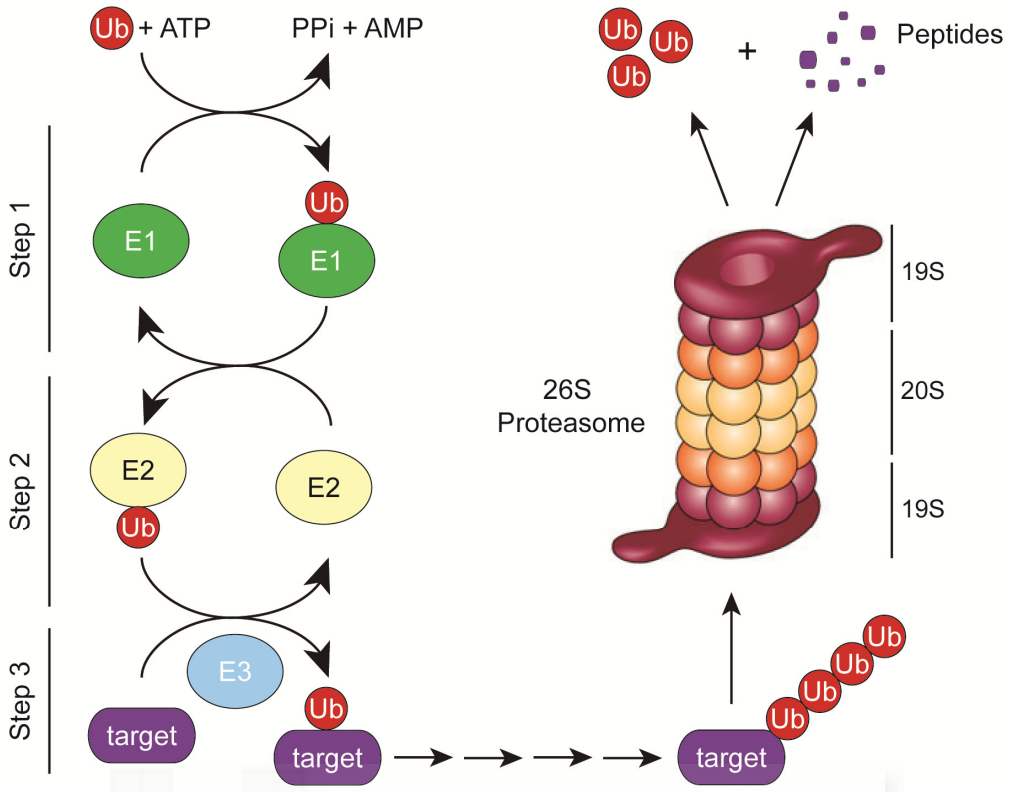


Figure 15. Ubiquitination of proteins targeted to the proteasome. A ubiquitin-activating enzyme (E1) binds ubiquitin, which is then transferred to a ubiquitin-conjugating enzyme (E2); a ubiquitin-protein ligase (E3) helps transfer ubiquitin to the target substrate. Multiple ubiquitin molecules are attached to proteins before recognition and degradation by the 26S proteasome. Before degradation, the ubiquitin chain is removed, allowing free ubiquitin molecules to be recycled. Adapted from Adams (2003 and 2004).

Regarding the structure of the 26S proteasome, it is a multiprotein complex consisting of a 20S core particle that is associated with one or two 19S regulatory particles (**Figure 15**). Each 19S subunit is capable of binding the polyubiquitin chain and cleaving it from the protein substrate. The substrate is then denatured, or ‘‘unfolded’’, and fed into the proteasome proteolytic core. The 20S core particle is a cylinder composed of four stacked rings and has little relevant proteolytic activity in its cellular milieu. The two outer rings (called α rings) complex with the 19S regulatory particles, forming a narrow channel through which only denatured proteins may pass. The catalytic chamber is formed by the two inner β rings, each of which contains three active sites. These sites differ in their substrate specificity and activity and have been named after enzymes that show similar proteolytic activity or specificity. These active sites are thus termed chymotrypsin-like, trypsin- like, and post-glutamyl peptide hydrolase-like (PGPH) (Groll *et al.*, 2001). Proteins are degraded by the core particle in

a progressive manner, generating peptides of 3-25 aa in length (Nussbaum *et al.*, 1998).

Ubiquitin-mediated proteolysis of a variety of cellular proteins plays an important role in many basic cellular processes. Among these we find regulation of cell cycle and cell division, differentiation and development, involvement in the cellular response to stress and extracellular effectors, morphogenesis of neuronal networks, modulation of cell surface receptors, ion channels and the secretory pathway, DNA repair, transcriptional regulation, transcriptional silencing, long-term memory, circadian rhythms, regulation of the immune and inflammatory responses, and biogenesis of organelles (reviewed in Ciechanover and Iwai, 2004). The list of cellular proteins that are targeted by ubiquitin is growing rapidly. Among them are cell cycle regulators such as cyclins, cyclin-dependent kinase inhibitors, and proteins involved in sister chromatid separation, tumour suppressors, transcriptional activators and their inhibitors (Ciechanover and Iwai, 2004). Cell surface receptors and endoplasmic reticulum proteins are also targeted by the system. Finally, mutated and denatured/misfolded proteins are recognized specifically, and are removed efficiently. By selectively removing these proteins, sparing their normal functioning counterparts, the system plays a key role in the cellular quality control and defence mechanisms. The products of the proteasome can play an important role in the immune response. In the case of degradation of foreign proteins, such as those of viral origin, the resulting short peptides are presented by the major histocompatibility complex class I molecules to the cytotoxic T cell that lyse the presenting cell (reviewed in Pamer and Cresswell, 1998; Rock and Goldberg, 1999). Additionally, the proteasome is responsible for activating NF-κB translocation to the nucleus after IκB degradation (Alkalay *et al.*, 1995; Chen *et al.*, 1995; Palombella *et al.*, 1994), an event necessary for the expression of various proteins critical in immune and inflammatory responses, including many cytokines and cell adhesion molecules (Read *et al.*, 1995). The proteasome is also involved in controlling the activity of NF-κB pathway through limited proteolysis of the NF-κB precursor protein p105 (Palombella *et al.*, 1994). Another likely example for ubiquitin/proteasome-dependent processing is the processing of the transcriptional regulatory protein Cubitus interruptus (Ci), which turns this component of the Hedgehog signalling pathway from a transcriptional activator into a repressor (Aza-Blanc *et al.*, 1997). These examples suggest that the proteasome may play a more general role in the proteolytic maturation of other soluble or membrane-associated proteins or in the partial cleavage of mature cell proteins. Finally, although protein degradation is the most well-established role of the proteasome, in the last years it has

been shown that it also has non-proteolytic functions, with an active role in multiple aspects of gene expression, including the recruitment of co-activators to promoters, the initiation of transcription, and also gene elongation (reviewed in Baker and Grant, 2005).

UNIVERSITAS Miguel Hernández

II. OBJECTIVES

UNIVERSITAS Miguel Hernández

UNIVERSITAS
Miguel
Hernández

The general aim of this work is to understand the molecular mechanisms that regulate Pipsqueak activity in tumorigenesis as a paradigm to unravel new mechanisms of oncogenesis mediated by BTB transcription factors. To this end, we focused in the following specific objectives:

1. To study the impact of sumoylation and SUMO-mediated interactions over Pipsqueak activity *in vivo* and in the context of the *Drosophila melanogaster* Notch-driven tumorigenesis paradigm.

2. To decipher the role of proteolytic processing on Pipsqueak transcriptional activity and how this mechanism modifies Pipsqueak oncogenic capacity.

UNIVERSITAS Miguel Hernández

UNIVERSITAS Miguel Hernández

III. RESULTS

UNIVERSITAS Miguel Hernández

UNIVERSITAS
Miguel
Hernández

Part 1. Sumoylation-mediated regulation of the tumorigenic activity of the BTB transcription factor Pipsqueak

Numerous human proto-oncogenes encode proteins of the large superfamily of BTB-containing proteins that include transcriptional regulators. BTB-bearing transcription factors act as transcriptional repressors as well as activators, and hence the mechanisms underlying tumorigenesis mediated by these proteins remain poorly understood. In this work, we show that the proto-oncogene Pipsqueak, encoding a BTB-containing transcriptional repressor, is sumoylated *in vitro* and *in vivo* and that, like its mammalian counterparts and other BTB factors in *Drosophila*, it can also act as a transcriptional activator. Using a yeast two-hybrid screen approach, we identified that Pipsqueak interacts with the SUMO ligase Su(var)2-10 and the sumoylation dependent transcriptional repressor MEP-1. Interestingly, Pipsqueak also interacts with a key component of the basal transcriptional machinery called DmTAF3, which we show is required positively for Pipsqueak-driven tumorigenesis. Moreover, we show that Pipsqueak contains two classical sumoylation sites and that Pipsqueak sumoylation occurs, and is required, exclusively at one of them, Lysine 633, located within the domain that interacts with the SUMO-related proteins identified in the Y2H screen. Blocking Pipsqueak sumoylation by the point mutation K633R boosts its transcriptional activity both *in vivo* and in cell culture, and also increases its pro-apoptotic activity, suggesting that Pipsqueak may act as transcriptional activator of a subset of genes that may participate in its tumorigenic programme. Intriguingly, sumoylation of Pipsqueak may, on the other hand, enhance or facilitate its binding to the component of the basal transcriptional machinery DmTAF3. Importantly, we show that both MEP-1 and DmTAF3 have a putative SUMO interaction motif (SIM) in their region of interaction with Pipsqueak, suggesting that MEP-1, DmTAF3 or both recognize sumoylated Pipsqueak through their SIM. In the first part of this work, I will present and discuss the *in vivo* consequences of Pipsqueak sumoylation and SUMO-mediated interactions in the *Drosophila* Notch-driven tumorigenesis paradigm, and its role in normal development.

UNIVERSITAS Miguel Hernández

1. Identification of Pipsqueak partners by a yeast two-hybrid approach

The BTB protein-protein interaction domain allows homo-, hetero-, or oligomerization of a protein. This can enhance the affinity and specificity of DNA recognition, as well as provide a simple regulatory mechanism rendering the factors inactive under a certain concentration threshold. Apart from providing such a dimerization interface, the recruitment of transcriptional co-regulators via the BTB domain has been shown to be relevant during BTB transcription factor-mediated genetic deregulation during oncogenesis (He *et al.*, 1998; Hong *et al.*, 1997; Lin *et al.*, 1998). With the aim of contributing to the molecular knowledge of this mechanism, we used the yeast two-hybrid technique (Y2H) to identify proteins that interact with the BTB-containing isoform of Psq. We designed the experiment to include 720 aa of Psq (from Met1 to Gln720), which spans the BTB domain and the central part of Psq located between the BTB and the Psq helix-turn-helix domains (**Figure 16A**). This central region has been shown to be of crucial importance for Psq-BTB regulation and function, possibly with general relevance for numerous BTB proteins, including the oncogenic transcription factors PLZF and BCL-6 (Kang *et al.*, 2003; Suliman *et al.*, 2012).

The Y2H technique allows detection of the interaction between two proteins through the activation of a reporter gene (**Figure 16B**). Classically, an eukaryotic transcriptional activator contains a domain that specifically binds to DNA (binding domain) and a domain that recruits the transcription machinery (activating domain). In the Y2H system, these domains are separated in two different polypeptides, each of them fused to different proteins that will act as bait or prey. The basis of this assay is that the transcription of the reporter gene only occurs if bait and prey interact, bringing together both parts of the transcriptional activator and reconstituting its functionality.

We used as prey a *Drosophila* embryo library from Hybrigenics Services, which is an equimolar pool of two cDNA libraries prepared from 0-12 hours and 12-24 hours embryo mRNA (**Figure 16A**). Each cDNA was cloned in frame with the Gal4 transcription factor activation domain. As bait we used a 720 aa Psq fragment including the BTB/POZ and the central region of the protein (Met1 to Gln720), fused to the LexA DNA-binding domain.

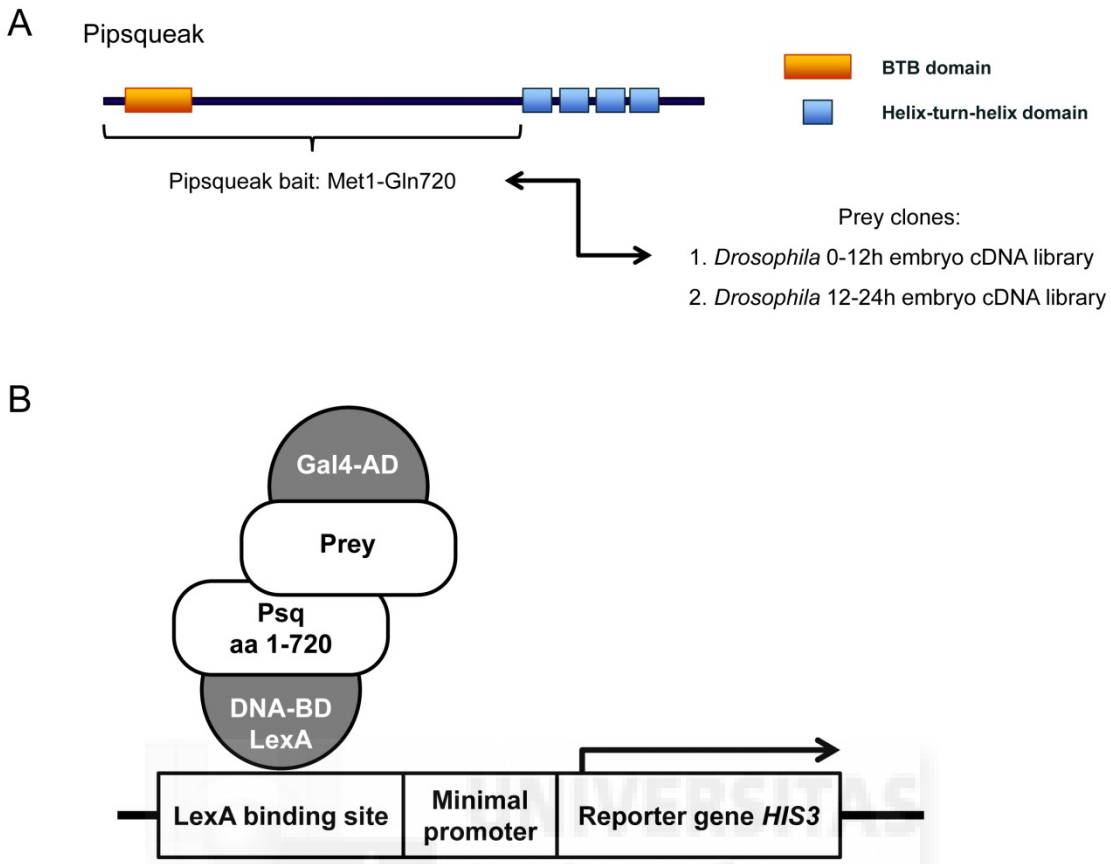


Figure 16. Description of the yeast two-hybrid experiment performed to identify new Pipsqueak interacting partners. (**A**) General schema of Psq protein. A fragment spanning the BTB and the central part of Psq was used as bait for the Y2H. The preys were cDNA from a *Drosophila* embryo library, which is an equimolar pool of two cDNA libraries prepared from 0-12 hours and 12-24 hours embryo mRNA. (**B**) When a fragment of a given protein fused to the Gal4 transcription factor activating domain interacts with the Psq fused to the LexA DNA-binding domain, this reconstitutes the functionality of the transcription factor, activating the transcription of the gene *HIS3*. For a more detailed explanation see *Materials and Methods*.

The results of these experiments were partially generated before the beginning of this Thesis, but as this information is essential for this work's description, the complete list of positive clones resulted from this experiment is included in *Appendix II*, as well as a more detailed explanation of the Y2H protocol in *Materials and Methods* section.

Of the 154 positive clones identified in the Y2H experiment, 2 clones identify Psq itself and 11 identify lola-like, a protein containing a BTB domain similar to that of Psq and its structurally related neighbour gene *lola*. These results validate the design

of the experiment, as it is known that BTB domains are able to form homo- or heterodimers.

Comparison of the aa sequence of all the fragments of a given protein that interact with Psq in the Y2H allowed us to identify a common sequence, called Smallest Interaction Domain (SID), that very likely contains all the structural determinants required for the interaction with Psq. In this Thesis, I have focused on analysing the interaction of Psq with the product of three different genes identified in the Y2H assay, and also how they modify Psq transcriptional and tumorigenic activity. These genes are *DmTAF3*, also known in *Drosophila* as *bip2*, a component of basal transcriptional machinery; *MEP-1* (or *CG1244)*, involved in sumoylation dependent transcriptional repression; and Su(var)2-10, *Drosophila* homolog of the SUMO ligase PIAS3. For the complete list of Psq interacting partners resulting from the Y2H assay see *Appendix II*.

2. Study of the interaction between Pipsqueak and DmTAF3

In the Y2H assay we found 22 positive clones corresponding to *DmTAF3/bip2* (*Appendix II*), a component of the basal transcriptional machinery conserved in humans (Gangloff *et al.*, 2001a). From now on I will use the name *DmTAF3* instead of *bip2*, according to the unified nomenclature proposed in the work of Tora (2002). Sequence comparison of the 12 fully sequenced clones corresponding to *DmTAF3* defined a SID with Psq located between DmTAF3 aa Gln809 and His878 (**Figure 17**).

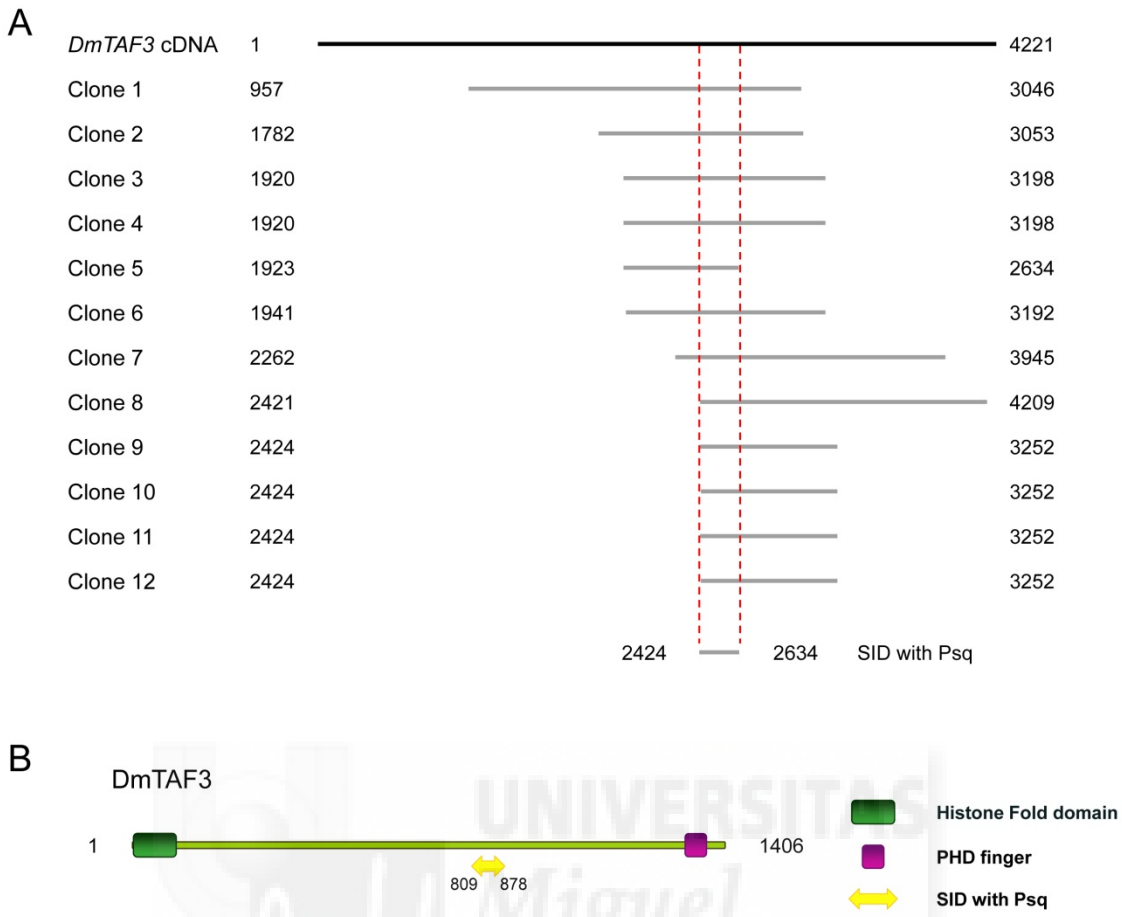


Figure 17. Identification of DmTAF3 Smallest Interaction Domain (SID) with Pipsqueak. (**A**) Sequence comparison of all *DmTAF3* fully sequenced clones identified in the Y2H screen allowed the identification of the SID with Psq. The numbers refer to the coding sequence, being 1 the first base of the first codon of the protein. (**B**) Representation of DmTAF3 protein with the SID and its main known domains depicted. The numbers refer to the aa sequence, being 1 the first aa of the protein.

2.1 DmTAF3 is specifically required for Pipsqueak-driven tumorigenesis

We first verified that the interaction between Psq and DmTAF3 detected in the Y2H experiment has functional relevance *in vivo*. We used endogenous *DmTAF3* loss-of-function mutants, RNA interference transgenes to knockdown *DmTAF3* tissue-specifically, and gain of expression of *DmTAF3* using UAS transgenic lines (see *Appendix I*). To investigate the requirement of DmTAF3 for Psq-mediated functions, we assayed both the capacity of *DmTAF3* mutations to suppress or enhance epigenetic silencing induced by *psq* overexpression, as described in (Ferres-Marco *et al.*, 2006).

We next examined the suppression/enhancement of Psq-driven tumorigenesis by introducing *DmTAF3* loss or gain-of-function mutations. In these experiments, we used the UAS/Gal4 binary system (Brand and Perrimon, 1993) to overexpress (or downregulate via RNAi) *DmTAF3*, specifically in the developing eye of *Drosophila* using the *eyeless* (ey)-Gal4 driver.

psq overexpression promotes epigenetic silencing of the *mini-white* gene contained in the *ey-Gal4* transposons (Ferres-Marco *et al.*, 2006). Given that Psq is thought to promote tumorigenesis by aberrant epigenetic repression of target genes (Ferres-Marco *et al.*, 2006), epigenetic silencing of the *mini-white* gene is used here as a proxy assay to directly monitor epigenetic silencing activity mediated by Psq. In this epigenetic silencing assay, the simple overexpression of *psq* and its neighbouring gene *lola* by the GS line *88A8* induced silencing of the *mini-white* gene in patches of eye cells in approximately 50% of the animals. Ferres-Marco found that mutations in *psq* prevented tumorigenesis and epigenetic silencing (also known as position effect variegation). Hence, this functional assay will be referred to as "variegation assay". This change is conspicuously manifested as a loss of red pigmentation in the adult fly eyes, and is easily identifiable even in living flies under the scope (**Figure 18A** and **B**). Using this assay we tested whether loss and/or gain of *DmTAF3* influenced Psq-driven epigenetic silencing in the variegation assay.

As control, the different *DmTAF3* lines were crossed to the *ey-Gal4* flies to evaluate the impact of *DmTAF3* alone in epigenetic silencing and/or eye growth or patterning. The histogram in **Figure 18C** shows the impact of loss and gain of *DmTAF3* in Psq-mediated epigenetic silencing of the *mini-white* gene. The *DmTAF3* mutant alleles (loss and gain-of-function) used had alone no impact on *Drosophila* eye pigmentation, size or patterning (data not shown).

psq overexpression using the *GS88A8* line (*eyeful*) and *ey-Gal4* (*ey-Gal4>eyeful*, see *Introduction*), causes eye colour variegation in approximately 50% of the flies. Reducing *DmTAF3* levels (down to ~30% of its normal mRNA) in this background using the $TAF3^{Fa4a}$ mutant allele has no effect in Psq-induced variegation (**Figure 18C**), so we decided to induce stronger downregulation of *DmTAF3* using eye-specific expression of RNAi transgenes against *DmTAF3* (*Materials and Methods*). One *DmTAF3* RNAi line (*#107591*) completely abolished eye variegation caused by Psq (**Figure 18C**), while the other RNAi *DmTAF3* lines (*#43174*, *#48037* and *#48036)* resulted in a strong-to mild reduction of the number of variegated eyes observed (14.5%, 17.2% and 30.5% respectively, and compared to the control). This rules out that reduction of variegation is a side-effect of the RNAi, as different independently

generated constructs reduced the variegation induced by *eyeful*. We suggest, given the interaction between DmTAF3 and Psq unveiled in the Y2H, that this suppression is likely mediated by the interaction of DmTAF3 with Psq in epigenetic silencing *in vivo*. In agreement with this, we found that *DmTAF3* overexpression using two different transgenic lines (*UAS-Myc::TAF3-Bfm* and *UAS-Myc::TAF3-Dmm1*), decreased variegation (down to 52.1% and 33.2% respectively). This would suggest that *DmTAF3* activity is necessary and sufficient for Psq-mediated epigenetic silencing.

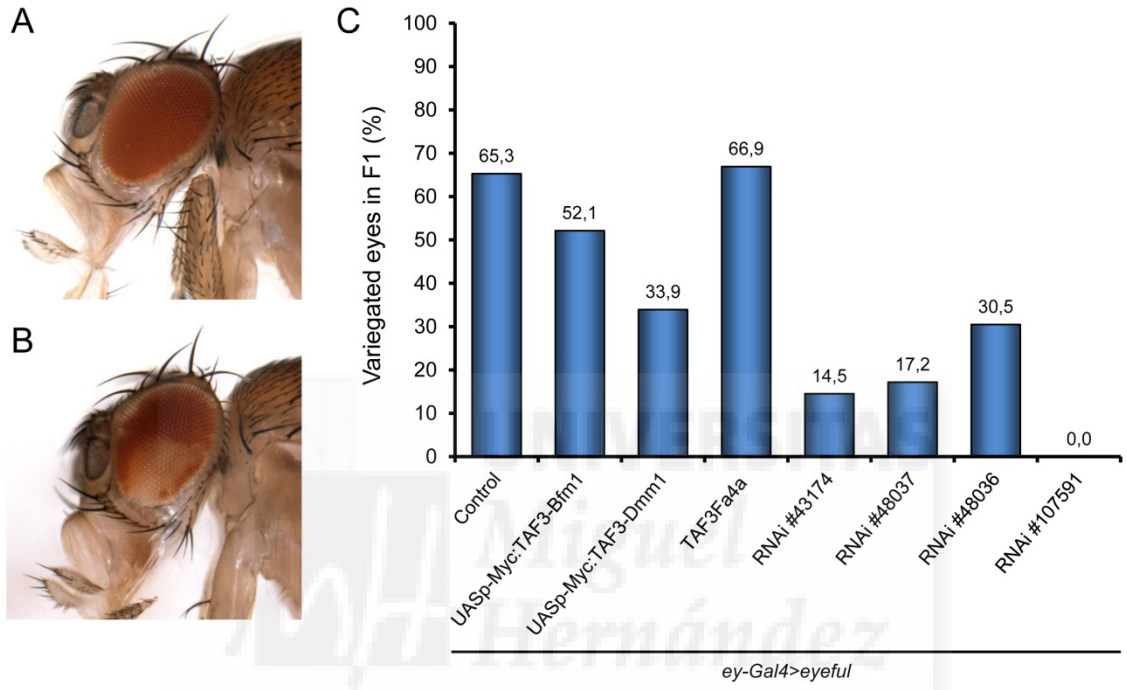


Figure 18. *DmTAF3* loss-of-function decreases position effect variegation caused by *pipsqueak* overexpression. Representative examples of the non-variegated (**A**) and variegated eye phenotype (**B**) in adult eyes of *ey-Gal4>eyeful*. Photographs were taken from female flies. (**C**) Quantification of the penetrance of position effect variegation phenotype in the eyes of adult flies. Several gain-of-function and loss-of-function stocks were crossed with the *ey-Gal4>eyeful* line to see the effect of modifying *DmTAF3* levels over Psq induced variegation. As control we crossed *ey-Gal4>eyeful* with w^{1118} flies. n ≥ 200 eyes per condition. Crosses were maintained at 25ºC.

Next, we investigated whether introducing mutations of *DmTAF3* or the expression of RNAi modulated Psq-driven tumorigenesis *in vivo*. We used a strain of flies that is a triple mutant carrying the *ey-Gal4* driver, the *GS88A8* line that drives *psq* and *lola* overexpression, and also a *Delta* transgene (Ferres-Marco *et al.*, 2006). This strain develops metastatic tumours due mainly to the co-expression of *psq* and the Notch ligand *Dl* (see *Introduction* and Ferres-Marco *et al.*, 2006). In this strain of flies, referred hereafter to as *ey-Gal4>Dl>eyeful*, approximately 55% of the flies develop eye

tumours, and around 12% of these flies display distant metastases (Ferres-Marco *et al.*, 2006; Gutierrez-Aviño *et al.*, 2009; Liefke *et al.*, 2010; Palomero *et al.*, 2007). As control, the different *DmTAF3* mutants and transgenes were crossed with *ey-Gal4* flies to determine the effect of increasing or reducing *DmTAF3* levels in the normal development of the eye. We also assayed the impact of *DmTAF3* mutations and transgenes on *Dl* overexpression driven by *ey-Gal4*. These assays unveiled that loss or gain of *DmTAF3* has no discernible impact on eye development in a context of normal *psq* expression. While *DmTAF3* overexpression (*UAS-Myc::TAF3-Bfm1* and *UAS-Myc::TAF3-Dmm1)* did not further enhance Psq-mediated tumorigenesis in the *ey-Gal4>Dl>eyeful* background (**Figure 19E**), the partial loss of *DmTAF3* by introducing a mutant copy of the gene ($TAF3^{Fa4a}$ mutation) or its eye-specific downregulation via RNAi expression, decreased the penetrance of the eye tumour phenotype (**Figure 19E**). Moreover, this suppression is specific because reducing *DmTAF3* activity in a genetically different tumour paradigm did not modify tumorigenesis (**Figure 19G**). Thus, reducing the levels of *DmTAF3*, a component of the basal transcriptional machinery, specifically reduced Psq (and or Lola)-driven tumorigenesis *in vivo*.

To sum up, these observations validate that the interaction between Psq and DmTAF3 identified in the Y2H screen has biological function. More specifically, the interaction between both proteins is necessary for Psq-mediated epigenetic silencing *in vivo* and for Pipsqueak-induced oncogenesis *in vivo*. In following sections of this Thesis I will describe the biochemistry studies performed to unravel the underlying mechanism responsible for this interaction.

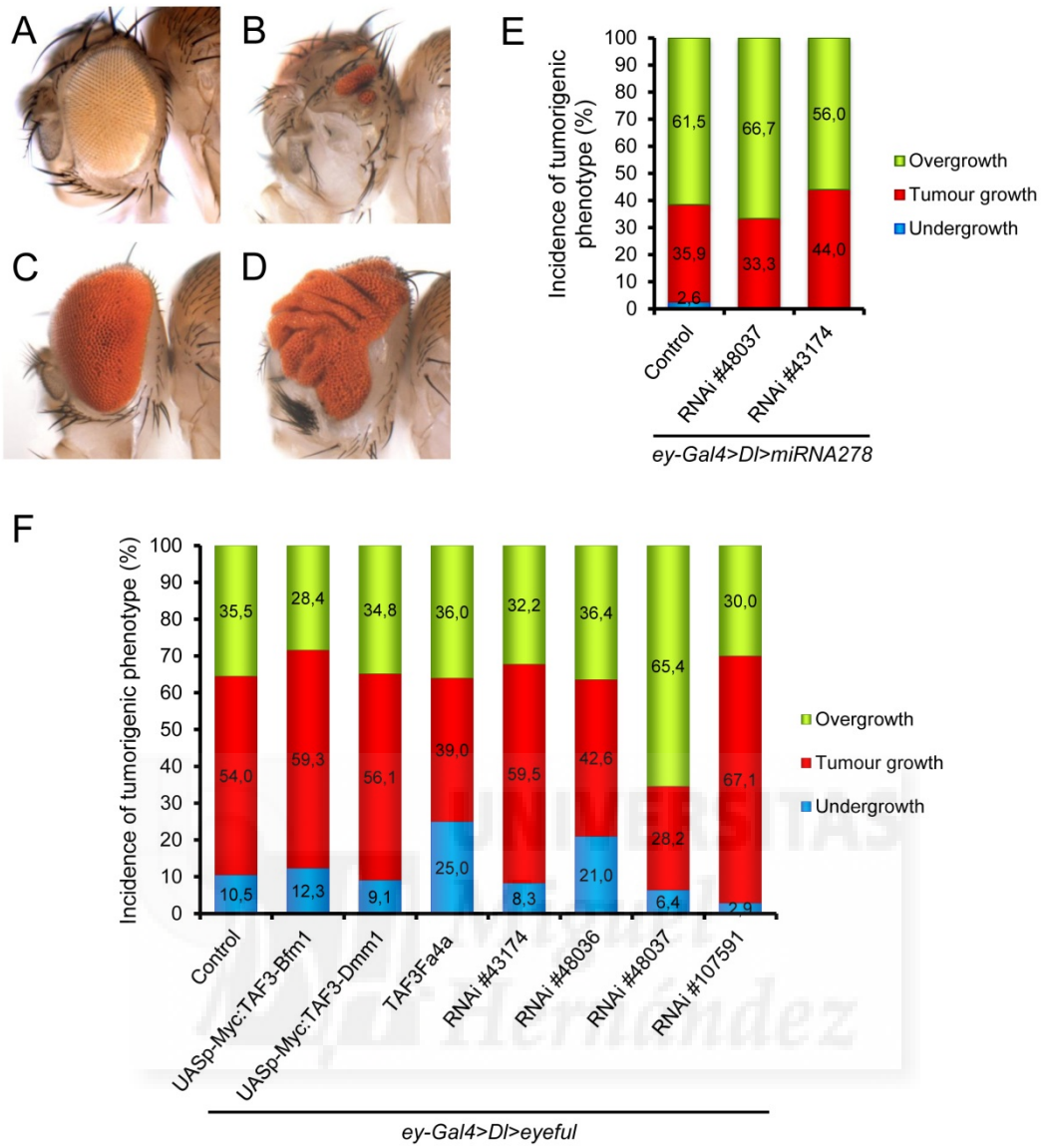


Figure 19. *DmTAF3* loss-of-function decreases the penetrance of eye tumour phenotype caused by *pipsqueak* overexpression. Representative examples of the different phenotypes observed in the *ey-Gal4>Dl>eyeful* background: (**A**) wild-type eye, (**B**) undergrown eye, (**C**) overgrown eye and (**D**) tumoural eye. Photographs were taken from female flies. (**E, F**) Quantification of the penetrance of tumorigenic, overgrowth and undergrowth phenotypes. Several gain and loss-of-function stocks were crossed with the *ey-Gal4>Dl>eyeful* line (**F**) to see the effect of varying *DmTAF3* levels over Psq induced tumorigenesis. Some of the stocks were re-tested with an additional tumour paradigm (*ey-Gal4>Dl>miRNA278*) (**E**) to assess if the effect observed was specific of Psq-mediated tumorigenesis. As control we crossed *ey-Gal4>Dl>eyeful* or *ey-Gal4>Dl>miRNA278* with w^{1118} flies. n ≥ 200 eyes per condition. Crosses were maintained at 27ºC.

2.2 The central region of Pipsqueak, not the BTB domain, mediates the interaction with DmTAF3

Having verified the functional relevance of the interaction between Psq and DmTAF3, we next wondered how they interact: is it direct? Or is the interaction mediated by a third unknown element? Do they form a complex with several proteins? To address these questions we started performing a detailed analysis of previously published Y2H data related to DmTAF3 (available in www.theBIOGRID.org). We found out that DmTAF3 had been previously identified in several Y2H assays in *Drosophila* as an interacting partner of several BTB containing proteins like Bric à brac 1 and 2 (Pointud *et al.*, 2001), Trithorax-like (Chopra *et al.*, 2008), Lola, Lola-like and Tramtrack among others (www.theBIOGRID.org). Interestingly, the data available indicates that these proteins, including Pipsqueak in our Y2H, bind to the same region of DmTAF3 (**Figure 20**). Upon this discovery, we wondered how all these different proteins could be interacting with the same region of DmTAF3. First we hypothesized that the BTB could be mediating the interaction of these proteins with that specific DmTAF3 region. This, however, does not explain how proteins without a BTB domain, like p53 (Bereczki *et al.*, 2008), Antennapedia (Prince *et al.*, 2008) and the protein Rad51 of *Drosophila* (www.theBIOGRID.org and http://pim.hybrigenics.com) are also able to interact with that same region of DmTAF3 (**Figure 20**). It is also a possibility that the interactions with non-BTB proteins are actually artefacts from the different Y2H experiments analysed, and only the BTB proteins are true positives. A third possible explanation is that there is some region or domain common to all the BTB and non-BTB proteins that mediates DmTAF3 recognition. Finally, the last explanation could be that all these proteins are post-translationally modified in some way and DmTAF3 is able to recognize that post-translational modification instead, rather than the protein *per se*.

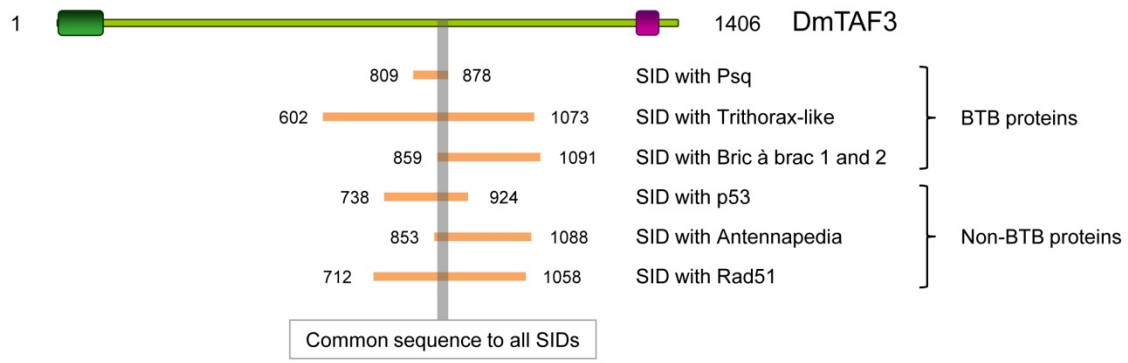


Figure 20. Structurally different proteins interact with the same region of DmTAF3. Representation of several SIDs from proteins with or without BTB, sharing a common region of interaction with DmTAF3.

To rule out the possibility that the BTB domain was mediating the interaction with DmTAF3, at least in the case of Pipsqueak, we tried to localize more precisely the exact region of Psq that is recognized by DmTAF3 by using 1-by-1 Y2H. For this experiment we used as baits three different fragments of Psq protein (**Figure 21A**): the fragment used in the original Y2H screen, which contains the BTB domain and the middle section of Psq (Met1-Gln720), the BTB domain (Met1-Asp131) and the central part of the protein (Ala132-Gln720). As prey for this experiment we used a *DmTAF3* clone spanning from Glu841 to Asp1067, fused to the Gal4 transcription factor activating sequence (**Figure 21A**). This region of DmTAF3 contains part of the SID with Psq found in the Y2H (Gln809-His878). As a result of the 1-by-1 experiment we found that this DmTAF3 fragment interacts with the region of Psq located between Ala132-Gln720 (**Figure 21C**).

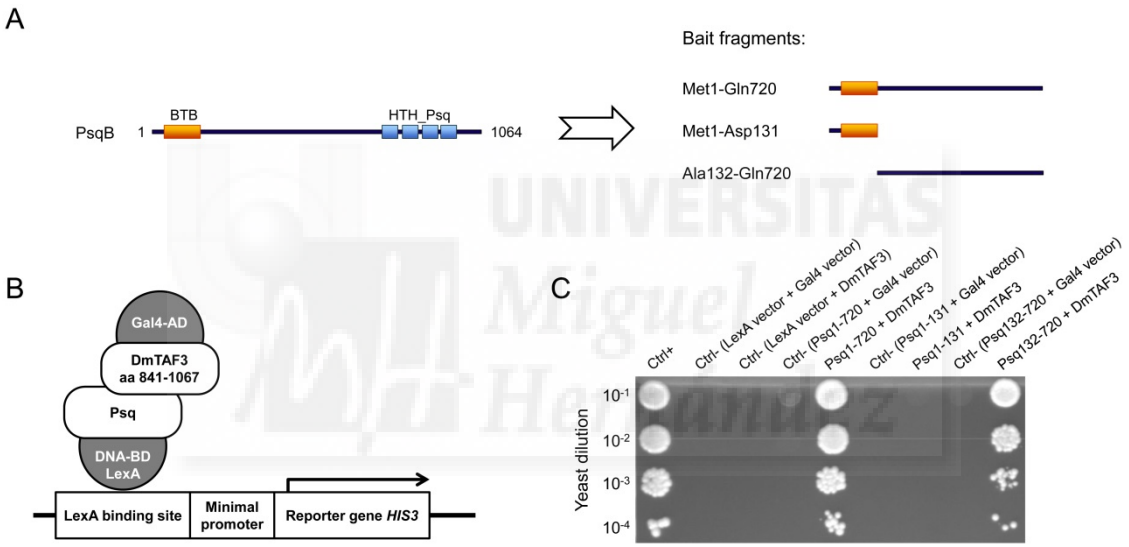


Figure 21. DmTAF3 interacts with the central region of Pipsqueak. (**A**) Representation of PsqB isoform and the different Psq fragments used as baits for the assay. (**B**) Schematic representation of the 1-by-1 Y2H assay to test the capacity of DmTAF3 to bind to different fragments of Psq. The interaction assay is based on the reporter gene *HIS3*. (**C**) For each interaction tested in the 1-by-1 Y2H assay, four different dilutions of yeast strains expressing the indicated constructs were spotted in selective medium without histidine (DO-3). 3-aminotriazole (3-AT), an inhibitor of the *HIS3* gene product, was added at a concentration of 10 mM to the DO-3 plates to increase stringency and reduce possible autoactivation by the bait proteins.

With this experiment we rule out that Psq interaction with DmTAF3 is mediated by the BTB domain of Psq. The result also suggests that interactions with non-BTB proteins might not be false positives, as the BTB domain is not essential to interact with that specific DmTAF3 region. Then, we still have two possibilities: is there a specific

domain common to all this proteins that mediates direct interaction with DmTAF3 or is there a protein modification that allows indirect binding? To explore these options we carried out a deep analysis of DmTAF3 protein sequence, and also of its interacting proteins, looking for motifs, domains or post-translational modification sites that could explain this interaction. The result of such analysis will be described in following sections of this Thesis, but before that it is necessary to talk about the other two new Psq interacting partners found in the Y2H, MEP-1 and Su(var)2-10/PIAS3. The reason is that the analysis of Pipsqueak interaction with these two proteins gave us valuable hints that led us to formulate a hypothesis that explains, not only how Pipsqueak interacts with DmTAF3, MEP-1 and Su(var)2-10/PIAS3, but also how DmTAF3 is able to recognize proteins with a huge structural variety.

3. Study of the interaction between Pipsqueak and MEP-1

Another interesting hit found in the Y2H assay was MEP-1, a sumoylation-dependent transcriptional repressor. We obtained 16 positive clones, being 7 of them fully sequenced (*Appendix II*). A comparison of the common sequence between these 7 fragments revealed that there are two possible SID regions that interact with Psq in MEP-1 protein (**Figure 22**). The first one is located close to the amino-terminal part of MEP-1, between aa Asn211 and Pro480, and the second is at the carboxy-terminal part, spanning from aa Pro909 to Ala1082. After analysing other Y2H results with PIMRider software (https://pimr.hybrigenics.com/) we found out that other proteins interacting with the SID located close to the carboxy-terminal region of MEP-1, like Merlin or Ptc, do it with a low confidence degree, while p53, RAD51 and CG17227 interact with a high confidence degree with the SID located close to the amino-terminal part of MEP-1. Moreover, the carboxy-terminal SID contains several zinc fingers (**Figure 22B**) and the classical function of this protein domain is to mediate DNA-binding, so the likelihood that in this same region there is a protein binding domain is low. All these observations suggest that MEP-1 interaction with Merlin and Ptc could be an artefact of the Y2H experiments. For all these reasons we propose that MEP-1 SID with Psq is the one located close to the amino-terminal end of MEP-1, between aa Asn211 and Pro480.

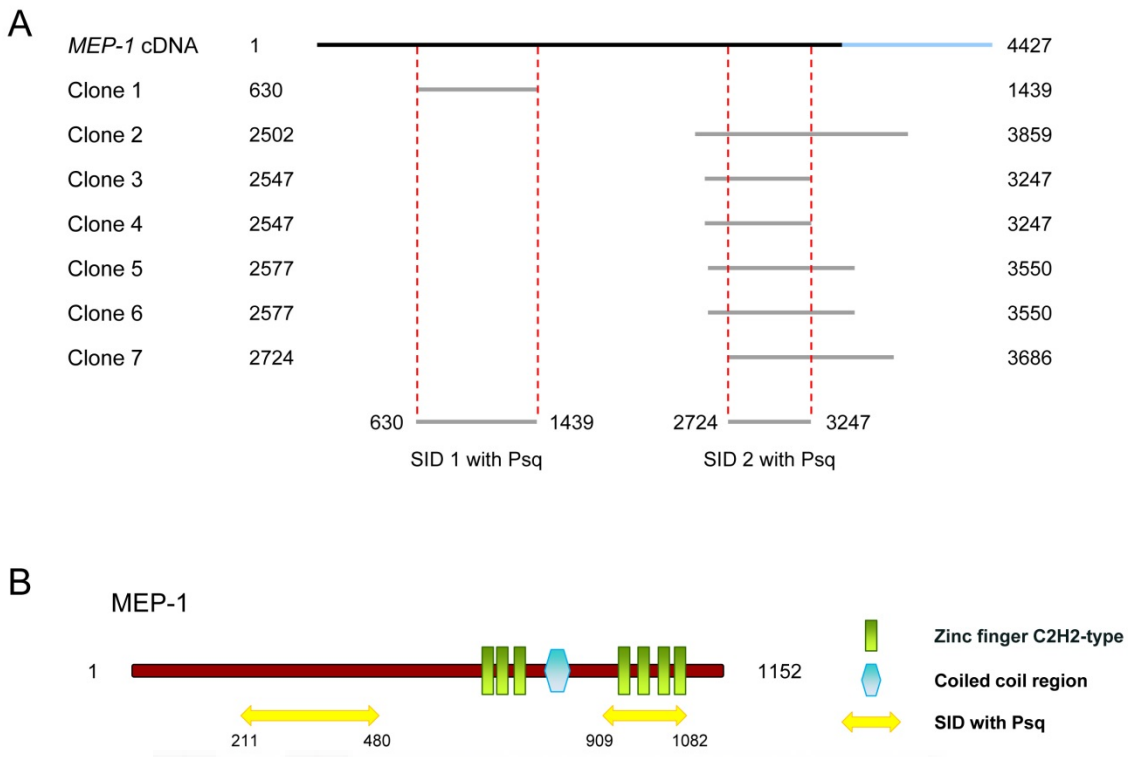


Figure 22. Identification of MEP-1 SID with Pipsqueak. (**A**) Sequence comparison of all *MEP-1* fully sequenced clones identified in the Y2H allowed the identification of two SIDs with Psq. The numbers refer to the coding sequence, being 1 the first base of the first codon of the protein. The blue line represents the 3'-UTR. (**B**) Representation of MEP-1 protein with the two SIDs and its main known domains depicted. The numbers refer to the aa sequence, being 1 the first aa of the protein.

The link between MEP-1 and sumoylation in *Drosophila* (see *Introduction*) and its possible conservation through evolution was attractive for us, taking into account that other positive clones from the Y2H identified a conserved component of the sumoylation pathway, the SUMO ligase PIAS3 also known as Su(var)2-10 in *Drosophila* (Johnson, 2004), (*Appendix II*). Furthermore, up to that moment no BTB protein had been involved in sumoylation or had been described as a possible target of sumoylation. The strong connection between SUMO and cancer (see *Introduction*), opens the possibility of sumoylation being a mechanism involved in regulating Psq activity and/or its oncogenic function. In this sense, it is important to emphasise that since the beginning of this Thesis, several recent works have now described connections between sumoylation and several *Drosophila* and other species' BTB proteins (Kang *et al.*, 2003; Reddy *et al.*, 2010; Stankovic-Valentin *et al.*, 2007), although the biological function of this regulation has not been described yet.

3.1 MEP-1 dampens Pipsqueak oncogenic function

To determine if the interaction between Psq and MEP-1 detected in the Y2H experiment has biological relevance and, more specifically, any effect in Psq oncogenic function *in vivo*, we followed the same strategy as in the case of *DmTAF3*. We used RNAi transgenes to knockdown *MEP-1* (see *Appendix I*) and assess what is the effect of varying *MEP-1* expression levels in Psq-mediated epigenetic silencing and tumorigenesis. Unfortunately, we could not combine this strategy with the use of loss-of-function mutants for MEP-1, as there are none available.

To test if reducing *MEP-1* levels had any effect in eye growth and development independent of its interaction with Psq, we crossed *MEP-1* RNAi flies with *ey-Gal4* (**Figure 23A**). Of the four RNAi lines tested, lines *#24533* and *#24534* showed a slight reduction in the size of the eye that affected 30.2% and 90.3% of the progeny respectively. On the other hand, lines *#35399* and *#33676* did not show any phenotype when compared to the control. In addition to this control, we crossed the different RNAi fly lines with a fly stock that overexpresses *ey-Gal4>Dl* to determine if there is any Psq-independent functional interaction between *MEP-1* and Notch signalling pathway. In these crosses we found that lines *#24533* and *#24534* again produce smaller eyes compared to the control in 29% and 40.4% of the progeny, respectively. Line *#33676* also presents a 23.2% of smaller eyes, being the line *#35399* the only one that does not have differences with the control. The results from this control crosses tell us that *MEP-1* is relevant for the proper development of the eye. This function could be specific to the eye development, but also MEP-1 may have a more general role in growth. Additionally, somehow it can be related to Notch signalling pathway as there is a recovery of the phenotype when we cross these lines with *ey-Gal4>Dl*.

Next, we crossed the different *MEP-1* RNAi lines with *ey-Gal4>eyeful* and *ey-Gal4>Dl>eyeful* to test how reducing *MEP-1* levels affects Psq-mediated variegation and Psq-driven tumorigenesis, respectively. In the variegation assay we observed that reducing MEP-1 expression levels using four different RNAi lines caused always 100% lethality in the pupal stage. These flies, dissected manually from the pupa, presented severe deformations in the head or had no head at all (**Figure 23A**). In the case of the tumorigenesis assay, the four lines displayed again the same phenotype: 100% of the flies die inside the pupa and when opened, they had no head. Additionally, we can say that the lethality caused by reducing *MEP-1* levels is specific of the *psq* overexpression background, as combining *MEP-1* RNAi with *psq* RNAi in the *ey-Gal4>eyeful* or *ey-Gal4>Dl>eyeful* backgrounds rescues the lethality (**Figure 23B**). Moreover, the

phenotype observed in the variegation assay and in the tumorigenic assay when MEP-1 levels are reduced mimics *psq* RNAi phenotype, which consists in lack of variegation and decrease in the penetrance of the tumorigenic phenotype (**Figure 23C**).

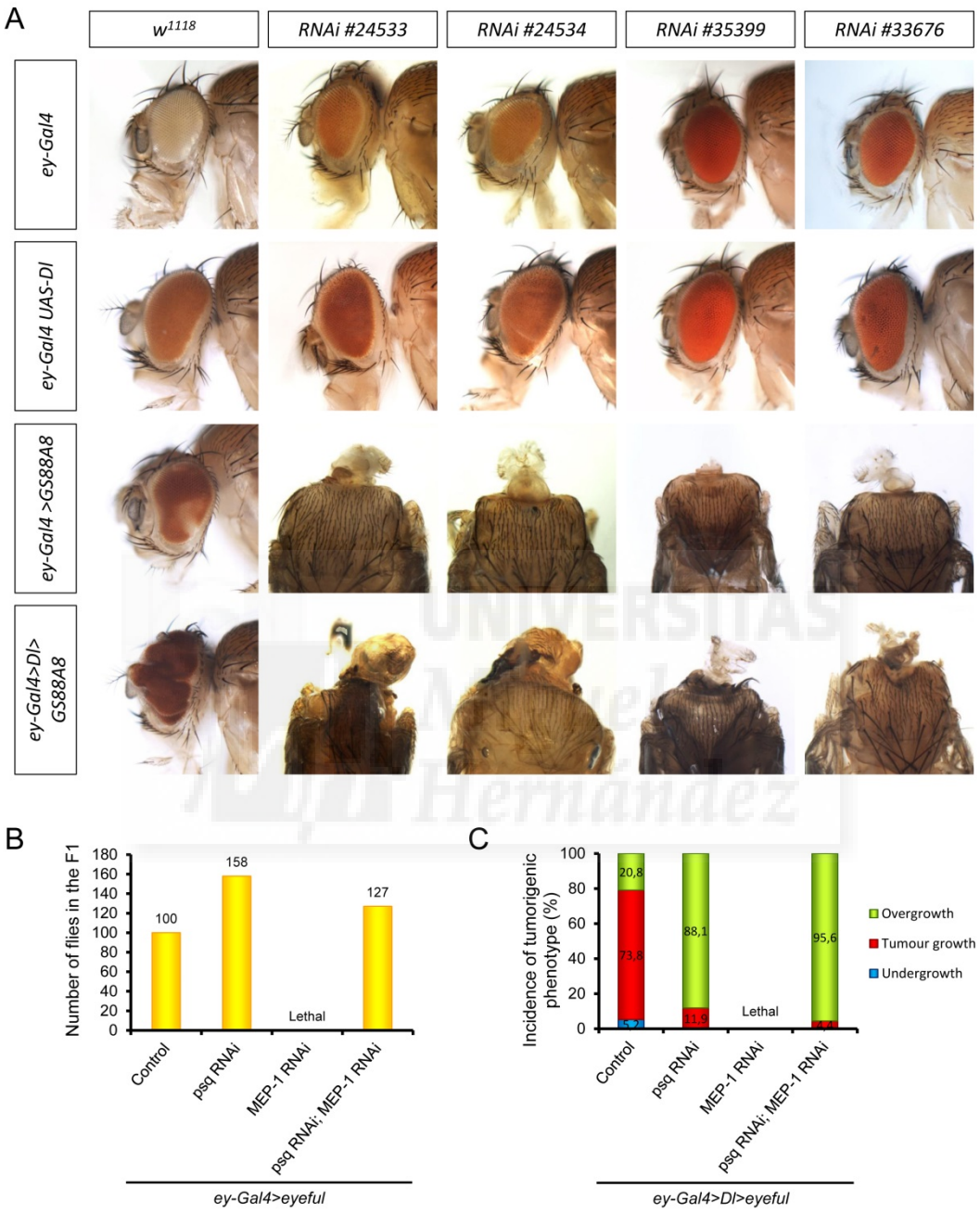


Figure 23. *MEP-1* loss-of-function causes lethality in a *pipsqueak* overexpression background. (**A**) Representative examples showing the adult eye phenotype obtained when crossing *ey-Gal4*, *ey-Gal4>Dl*, *ey-Gal4>eyeful* or *ey-Gal4>Dl>eyeful* with different *DmTAF3* RNAi lines. Photographs were taken from female flies. (**B, C**) Quantifications showing that reducing *psq* levels with RNAi line VDRC *#106404* rescues the lethality caused by *MEP-1* RNAi (line *#24534*) in a *psq* overexpressing background, both in the variegation assay and in the tumorigenesis assay, respectively. As control we crossed the four MEP-1 RNAi lines with w^{1118} flies. n ≥ 200 eyes per condition. Crosses were maintained at 25ºC, except the crosses with *ey-Gal4>Dl>eyeful*, which were maintained at 27ºC.

In conclusion, although reduction of *MEP-1* levels has an effect in the development of the eye and there seems to be some interaction with Notch signalling pathway independent of Psq activity, these facts alone are not enough to explain the strong lethal phenotype observed in the interaction with *ey-Gal4>eyeful* and *ey-Gal4>Dl>eyeful*. It seems that a reduction in *MEP-1* levels acts as a "brake unleash" increasing Psq activity, leading to severe defects that cause lethality during development.

Following this idea, MEP-1 overexpression should rescue the lethality observed in these crosses. To test this hypothesis we used four different stocks carrying GS insertions in *MEP-1* (see *Appendix I*). In the crosses with *ey-Gal4* and *ey-Gal4>Dl*, none of the lines showed differences with the controls, meaning that endogenous MEP-1 levels are enough to carry on properly with its function, and overexpression does not affect this balance. On the contrary, *MEP-1* overexpression in a *psq* overexpression background shows strong effects. First, in the variegation assay we can see that *MEP-1* overexpression with three of the four lines tested (*#206592*, *#202015* and *#203392*) completely abolished Psq-mediated variegation, while line *#201714* only showed a reduction of the incidence of the variegation phenotype of approximately 10% respect to the control (**Figure 24A**). In the case of the tumorigenic assay, when we crossed these lines with *ey-Gal4>Dl>eyeful* we see a strong rescue of the tumorigenic phenotype with lines *#206592* and *#202015* (**Figure 24B**). However, line *#201714* showed a similar phenotype to the control, as it occurred in the variegation assay, indicating that this line may not be actually overexpressing MEP-1 as strongly as the other lines, explaining why we see a weaker effect when we use it. Finally, line *#203392* showed a 15% reduction in the incidence of the tumour phenotype compared to the control, but to our surprise it also promoted a huge increase in the incidence of the undergrowth eye phenotype (**Figure 24B**). This could be due to the low number of flies obtained for this cross. Usually we count >200 eyes per condition to be really sure of the phenotypes observed and be able to run solid statistical analysis. This number is achieved by four to six repetitions of the same cross, but there are stocks with reduced viability, as is the case for line *#203392*. Four repetitions of the cross *ey-Gal4>Dl>eyeful* x *#203392* resulted in only 18 flies of the desired genotype. In the case of the variegation assay, the number was higher (32 flies) but still very far from the numbers obtained with the other stocks. Therefore, by raising the numbers we may observe a similar phenotype. In conclusion, two of the MEP-1 overexpressing lines, *#206592* and *#202015*, were able to rescue almost completely the tumorigenic phenotype caused by *psq* overexpression, while the other

two showed a tendency towards a decrease, in addition to the abolishment of Psq mediated epigenetic silencing.

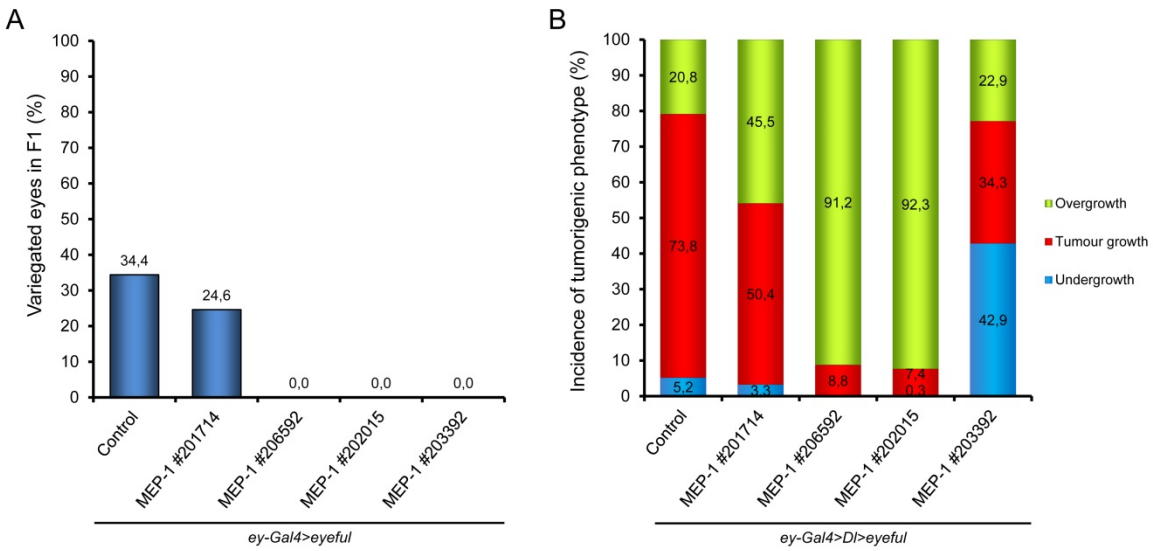


Figure 24. *MEP-1* gain-of-function rescues the lethality caused by MEP-1 RNAi lines in a *pipsqueak* overexpression background. (**A**) Quantification of the penetrance of position effect variegation phenotype in the eyes of adult flies. Several gain-of-function stocks were crossed with the *ey-Gal4>eyeful* line to see the effect of varying *MEP-1* levels over Psq induced variegation. (**B**) Quantification of the penetrance of tumorigenic, overgrowth and undergrowth phenotypes. Several gain and loss-of-function stocks were crossed with the *ey-Gal4>Dl>eyeful* line to see the effect of increasing *MEP-1* levels over Psq induced tumorigenesis. As control we crossed *ey-Gal4>eyeful* and *ey-Gal4>Dl>eyeful* with w^{1118} flies. n ≥ 200 eyes per condition. All crosses were maintained at 25ºC, but for the crosses with *ey-Gal4>Dl>eyeful* that were maintained at 27ºC.

From these experiments we infer that reduction of *MEP-1* levels dramatically affects flies overexpressing *psq*. However, *MEP-1* overexpression is able to rescue these effects and the variegation and tumorigenesis mediated by *psq* overexpression. According to these results, in normal conditions MEP-1 acts as a brake limiting Psq capacity to induce invasive growth. Reduction of *MEP-1* levels increases that oncogenic capacity, leading to premature death. Taking into account that MEP-1 is a transcriptional repressor, it is possible that MEP-1 interaction with Psq is necessary to modulate Psq transcriptional activity. But how this interaction occurs remains an unanswered question. In following sections of this Thesis we assess which is the underlying mechanism of MEP-1 and Psq interaction from a molecular point of view.

4. Study of the interaction between Pipsqueak and Su(var)2-10

Among the Psq protein partners identified in the Y2H experiment we found a second protein related to sumoylation: Su(var)2-10, the homolog of the human SUMO ligase PIAS3. As I previously mentioned in the *Introduction*, there are several SUMO ligases in mammals, but Su(var)2-10 is the only homolog of PIAS SUMO ligases discovered until now in *Drosophila* (Talamillo *et al.*, 2008).

While DmTAF3 and MEP-1 were classified as positive hits with the highest confidence index, Su(var)2-10 was assigned to the second category of this classification, meaning that the probability of Su(var)2-10 being a true hit was lower (see *Appendix II*). This lower confidence index could also be related to the high dynamic nature of the interaction between Su(var)2-10 and the sumoylation target, which makes it difficult to detect the interaction in assays such as the Y2H. The transient character of these interactions could additionally explain why we did not detect the conjugating enzyme Ubc9 in our screen. We nevertheless decided to further explore the interaction with Su(var)2-10 because it could serve us to establish a link between Psq and MEP-1 through sumoylation. As MEP-1 is a transcriptional repressor that is able to recognise sumoylated proteins, could it be possible that Su(var)2-10 interacts with Psq to sumoylate it? Is it possible that MEP-1 is interacting with sumoylated Psq?

To address these questions, we analysed our Y2H results to identify which part of Su(var)2-10 was interacting with Psq. In this case we obtained 4 positive clones for the Su(var)2-10-Psq interaction in the Y2H. Unfortunately, only one of them was fully sequenced (*Appendix II*), so we could not define the SID as precisely as we did in the case of DmTAF3 or MEP-1. The region of interaction with Psq is located between aa Ala362 and Glu602 of Su(var)2-10 (**Figure 25**). This fragment includes the SP-RING domain of Su(var)2-10 (from aa Val368 to Pro417), which is responsible for the SUMO ligase activity (see *Introduction*). It has been proposed that this domain recruits the substrate and the E2 conjugating enzyme and also serves as scaffold to position and orient them optimally for SUMO transfer (Johnson and Gupta, 2001; Kotaja *et al.*, 2002; Sachdev *et al.*, 2001; Takahashi *et al.*, 2001a; Yunus and Lima, 2009). The fact that Psq interacts with the SP-RING of Su(var)2-10 suggests that this SUMO ligase could be interacting with Psq to sumoylate it. This possibility resulted very intriguing at the time, as up to that point there were no evidences of Psq sumoylation in the literature.

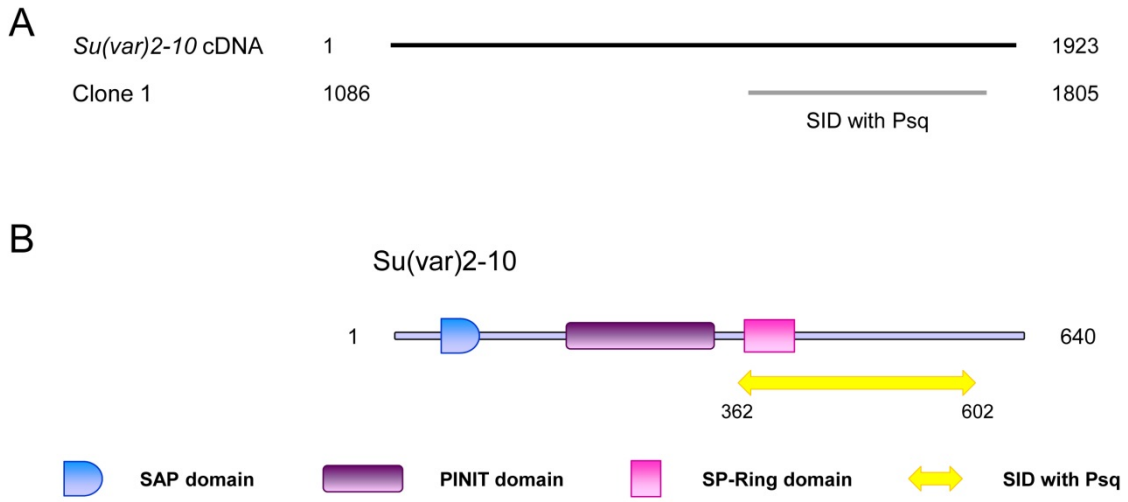


Figure 25. Identification of the Su(var)2-10 SID with Pipsqueak. (**A**) Sequence comparison of the only *Su(var)2-10* fully sequenced clone identified in the Y2H pointing the SID with Psq. The numbers refer to the coding sequence, being 1 the first base of the first codon of the protein. (**B**) Representation of Su(var)2-10 protein with the SID and its main known domains depicted. The numbers refer to the aa sequence, being 1 the first aa of the protein.

4.1 Su(var)2-10 binds to the central region of Pipsqueak, as does DmTAF3

In order to determine which region of Psq was recognized by Su(var)2-10 we performed a 1-by-1 Y2H assay. For this experiment we used as baits the same Psq fragments employed in the DmTAF3 1-by-1 (**Figure 21A and 26A**). As prey we used a cDNA clone that spanned from aa Thr363 to Leu601 of Su(var)2-10 protein, which contains the full region of interaction with Psq (**Figure 25B**). Similarly to what happened with DmTAF3, Su(var)2-10 only interacts with the Psq fragments that contain the region located between aa Ala132 and Gln720 (**Figure 26C**); therefore, again, the BTB domain is not required for the interaction between Psq and Su(var)2-10.

As I previously mentioned in the *Introduction*, Psq has a modular structure with a BTB domain at the amino-terminal part and four helix-turn-helix domains at the carboxy-terminal connected by a central region. There is a lot of data regarding how the BTB domain mediates interaction between proteins and also how the helix-turn-helix domain allows DNA recognition. But little is known about the motifs present in the central region of Psq protein and their relevance for its function. Su(var)2-10 interaction with this part of the protein suggests that Psq could be sumoylated, but are there

sumoylation sites in Pipsqueak? Our attempts to test this hypothesis are described in the following sections of this Thesis.

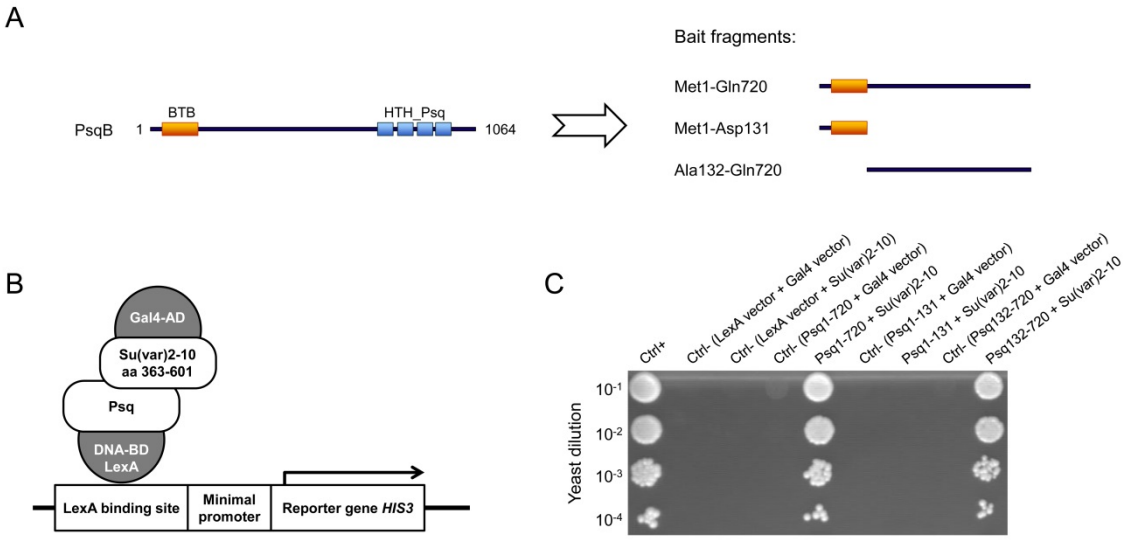


Figure 26. Su(var)2-10 interacts with the central region of Pipsqueak. (**A**) Representation of PsqB isoform and the different Pipsqueak fragments used as baits for the assay. (**B**) Schematic representation of the 1-by-1 Y2H assay to test the capacity of Su(var)2-10 to bind to different fragments of Psq. The interaction assay is based on the reporter gene *HIS3*. (**C**) As in the case of DmTAF3, for each interaction tested in the 1-by-1 Y2H assay, four different dilutions of yeast strains expressing the indicated constructs were spotted in the selective medium without histidine (DO-3). 3-aminotriazole (3-AT), an inhibitor of the HIS3 gene product, was added at a concentration of 10 mM to the DO-3 plates to increase stringency and reduce possible autoactivation by the bait proteins.

4.2 Mutants of *Su(var)2-10* increase Pipsqueak oncogenic function

To functionally validate the interaction between Psq and Su(var)2-10 observed in our Y2H, we used loss and gain-of-function approaches to assess the effect of *Su(var)2-10* on Psq-mediated epigenetic silencing and tumorigenesis.

In the case of the variegation assay, we used four different loss-of-function stocks: a P-element insertion that interrupts *Su(var)2-10* (*#11344*), two lines carrying ethyl methanesulfonate mutant alleles of *Su(var)2-10* (*#3245* and *#3246*) and one RNAi transgenic line (*#24998*). The three mutants behave in a similar fashion; all of them enhance the variegation mediated by Psq respect to the control (**Figure 27A**). However, the *Su(var)2-10* RNAi line is almost completely lethal. On that cross most of the flies die during pupal stages without developing a head and the few flies that reach

adulthood have small eyes . It is difficult to explain how an RNAi transgenic line could be more potent than, for example, the *Su(var)2-10 #6235* line that carries the *Su(var)2-10*^{2} mutant allele, in which a G to A mutation yields a tryptophan to STOP change in aa Val260, presumably making this a null allele as it lacks the SP-RING domain responsible for the SUMO ligase activity. *Su(var)2-10* expression levels have not been measured in this different stocks so we cannot know if this difference in phenotype is due to differences in expression levels. In any case, our data suggest that a reduction in *Su(var)2-10* levels increases Psq activity, measured here as an increase on variegation, i.e. epigenetic silencing.

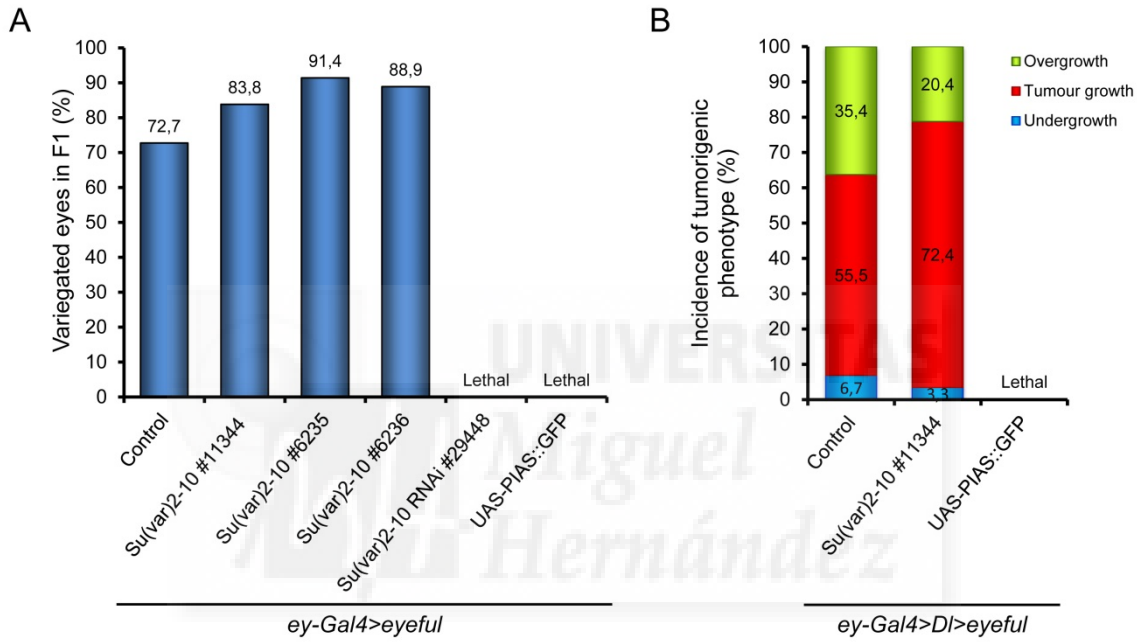


Figure 27. *Su(var)2-10* gain-of-function rescues the lethality caused by MEP-1 RNAi lines in a *pipsqueak* overexpression background. (**A**) Quantification of the penetrance of position effect variegation phenotype in the eyes of adult flies. Several gain-of-function stocks were crossed with the *ey-Gal4>eyeful* line to see the effect of varying *Su(var)2-10* levels over *psq* induced variegation. (**B**) Quantification of the penetrance of tumorigenic, overgrowth and undergrowth phenotypes. Several gain and loss-of-function stocks were crossed with the *ey-Gal4>Dl>eyeful* line to see the effect of varying *Su(var)2-10* levels over *psq* induced tumorigenesis. As control we crossed *ey-Gal4>eyeful* and *ey-Gal4>Dl>eyeful* with w^{1118} flies. n ≥ 200 eyes per condition. All crosses were maintained at 25ºC, except the crosses with *ey-Gal4>Dl>eyeful* that were maintained at 27ºC.

On the other hand, overexpression of *Su(var)2-10* with the stock *yw;(UAS-PIAS::GFP)26b.3* causes 100% lethality (**Figure 27A**). Flies dissected from the pupa do not develop a head, as we observed also with the *Su(var)2-10* RNAi line. Taking into consideration the concerns raised in the previous paragraph about *Su(var)2-10*

RNAi line and the fact that has the same phenotype as the gain-of-function, it may be possible that this line is not working properly, and the phenotype we see is not due to the effect over *Su(var)2-10*. Also, the lethality caused by the overexpression stock could be due not only to an increase in Psq sumoylation, as increasing *Su(var)2-10* levels and hence its activity leads to a general increase in the levels of sumoylated proteins, making it difficult to predict the outcome of this perturbation in the system.

In this variegation assay the percentage of variegated eyes in the control is higher than in previously shown experiments (compare **Figure 24A** and **Figure 27A**). This is explained because eventually the *ey-Gal4>eyeful* stock loses activity and we need to do a new recombinant every few years in order to not lose completely the phenotype. These crosses were performed years ago with a newly made stock, while the crosses shown in DmTAF3 and MEP-1 sections are more recent. That is why every time we perform this assay we count again the percentage of variegated eyes in a control cross (*ey-Gal4>eyeful* x w^{1118}), to monitor the strength of the phenotype and have a reliable value to relate our experimental crosses.

In the tumorigenesis assay, reducing *Su(var)2-10* levels with line #11344 increased the penetrance of the tumorigenic phenotype from 55.5% to 72.4% respect to the control, while *Su(var)2-10* overexpression produced 100% lethality of the flies, and the animals dissected from the pupa had no head (**Figure 27B**). If we do not take into consideration the results obtained with the RNAi line, these results are in agreement with the ones obtained in the variegation assay: reducing *Su(var)2-10* levels increases Psq activity and its overexpression is lethal to the organism, meaning that a fine regulation of *Su(var)2-10* levels is necessary for the proper development of the organism. This is not surprising as this is the only known homolog of PIAS SUMO ligases working in *Drosophila,* and sumoylation modulates the activity of key transcription factors and other relevant molecules. In any case, the fact that the epigenetic silencing and the tumorigenic Psq activities are boosted when *Su(var)2-10* levels are reduced suggests that Psq might be sumoylated and that this post-translational modification may attenuate Psq activity.

4.3 Mutations in *smt3* and the E2 conjugating enzyme *lesswright* do not provide solid evidence of sumoylation influence over Pipsqueak activity *in vivo*

While we were assessing the effect of varying Su(var)2-10 levels in Psq activity, we thought that testing the effect of mutations in other members of the sumoylation pathway would also provide valuable information. It could serve us as another evidence to support the idea of Psq sumoylation, independently of the interaction with Su(var)2-10 discovered in the Y2H assay. In the case of the variegation assay, reducing SUMO levels using different alleles of the gene *smt3* (the only *SUMO* homolog in *Drosophila*), produced a slight increase in Psq-mediated variegation (**Figure 28A**). At the same time we tested four different *lesswright* mutant alleles (the *Drosophila* homolog of the *Ubc9* gene that encodes the E2 conjugating enzyme) and three of them showed a similar phenotype, an increase in variegation compared to the control condition. Unfortunately, we did not have any lines that allowed overexpression of these genes, so we could not test the effect of a gain-of-function of the sumoylation pathway on Psq-mediated epigenetic silencing.

In the tumorigenesis assay, reduction in *SUMO* levels using different alleles of *smt3* in a tumour background had opposite results depending on the mutant allele tested. Mutation $smt3^{04493}$ induced a reduction in the incidence of the tumour phenotype of 9% respect to the control, while $smt3^{K06307}$ induced an increase of 22% in the incidence of tumours (**Figure 28B**). We also had similar contradictions in the case of *lwr* mutants. lwr^{05486} and lwr^{5} mutations reduced the apparition of tumours between 30 and 12% respectively, while other mutations (lwr^{4-3} y lwr^{13}) increased the incidence of tumours 22 and 14% respectively.

The results obtained with some of the mutants go along with our findings for *Su(var)2-10* while other completely contradict them. This disparity is likely due to the pleiotropic effects of altering the sumoylation pathway. Reducing sumoylation in a global way, like we did in these experiments, could affect not only proto-oncogenes such as Psq, but also proteins with opposite functions that are also sumoylated, like the tumour suppressor p53 (Wu and Chiang, 2009b). This could hide the specific effects over Psq function, producing contradictory results.

To sum up, in this case, the genetic experiments altering the endogenous expression of different sumoylation pathway components have not allowed us to establish a clear causal link between sumoylation and Psq tumorigenic activity *in vivo*.

For this reason we decided to perform experiments that allowed us a more detailed analysis of Psq and SUMO interaction using new mutants that I will describe below.

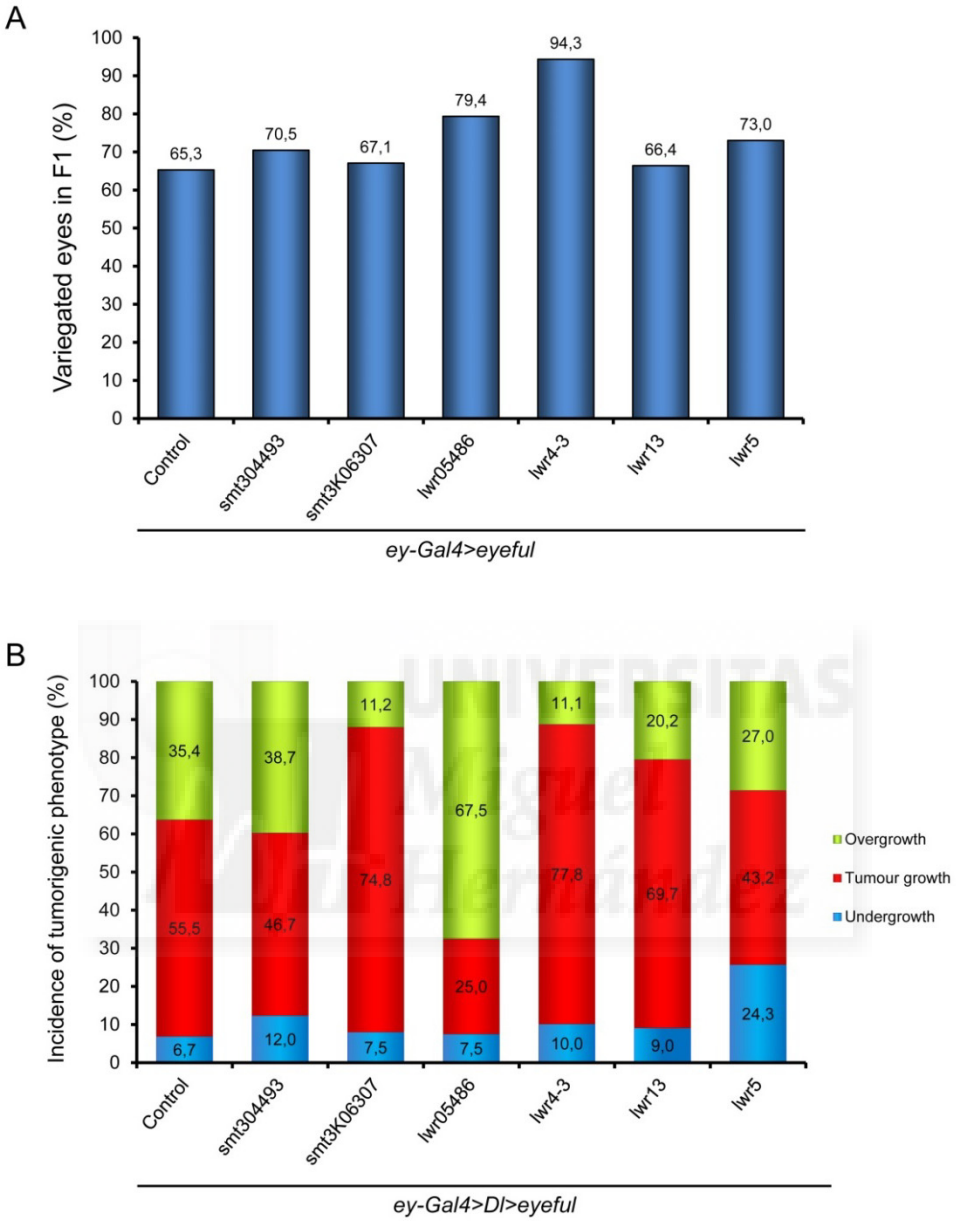


Figure 28. Mutations in several members of the sumoylation pathway affect differently Pipsqueak activity. (**A**) Quantification of the penetrance of position effect variegation phenotype in the eyes of adult flies. Several stocks carrying loss-of-function mutant alleles for *smt3* and *lwr* were crossed with the *ey-Gal4>eyeful* line to see the effect of interfering with sumoylation over Psq-induced variegation. (**B**) Quantification of the penetrance of tumorigenic, overgrowth and undergrowth phenotypes. Several stocks carrying loss-of-function mutant alleles for *smt3* and *lwr* were crossed with the *ey-Gal4>Dl>eyeful* line to see the effect of interfering with sumoylation over Psq-driven tumorigenesis. As control we crossed *ey-Gal4>eyeful* and *ey-Gal4>Dl>eyeful* with w^{1118} flies. n ≥ 200 eyes per condition. All crosses were maintained at 25ºC, except the crosses with *ey-Gal4>Dl>eyeful* that were maintained at 27ºC.

5. Sumoylation-dependent recognition of Pipsqueak

As a result of the Y2H experiment performed to discover new partners that could modify Psq activity, we have found several proteins that are able to interact with it and have a functional impact on its oncogenic function, like DmTAF3, Su(var)2-10 and MEP-1. In the case of Su(var)2-10 and MEP-1, both of these proteins are related to the sumoylation pathway and the experiments conducted up until now led us to hypothesize that Psq can be a target of sumoylation. The experiments shown in this section are aimed to confirm whether Psq is indeed sumoylated and, in parallel, to investigate whether MEP-1 and DmTAF3 are able to recognize and interact with a SUMO protein attached to a substrate.

5.1 Pipsqueak has two putative sumoylation sites in the region that interacts with Su(var)2-10 and DmTAF3, but is only sumoylated at Lysine 633

In order to prove the hypothesis of Psq sumoylation we started looking for sumoylation sites in Psq. We analysed the entire sequence of the protein focusing on the region that interacts with Su(var)2-10 in the 1-by-1 experiment (Ala132-Gln720). As I described in the *Introduction*, the linkage between SUMO and its substrates is an isopeptide bond between the C-terminal carboxyl group of SUMO and the ε-amino group of a lysine residue in the substrate. Many of the lysine residues where SUMO becomes attached are in the short consensus sequence ΨKXE, where Ψ is a large hydrophobic aa, generally isoleucine, leucine, or valine; K is the lysine residue that is modified; X is any residue; and E is a glutamic acid (see *Introduction*). This motif is bound directly by Ubc9 and E3 conjugating enzymes, like Su(var)2-10, and this molecules probably enhance specificity by interacting with other features of the substrate.

Using different bioinformatics tools, like the ELM resource (http://elm.eu.org/) or SUMOplot™ Analysis program (http://www.abgent.com/sumoplot), we were able to identify several sumoylation sites in Psq that coincide with the canonical sequence described above (**Table 4**). Among them, the sites IKSE (residues 517-520), IKHE (residues 632-635) and QKEP (residues 675-678) are located in the central region of Psq recognized by Su(var)2-10, making them good candidates to be functional sumoylation sites. The QKEP site was assigned a low score probably due to its unusual aa composition, while sites IKSE and IKHE have the highest score. The

remaining sumoylation sites are located outside of the region of interaction with Su(var)2-10 and/or were assigned lower scores by the different algorithms used for the prediction. For this reasons we decided to focus first in the sites containing Lys518 and Lys633.

SUMO site	Position	Sequence	Score
1	K633	IDPSQ**IKHE**PGMII	0.94
2	K518	NLVQH**IKSE**VIEAK	0.94
3	K926	SLYGR**FKRG**KYDVV	0.68
4	K929	GRFKR**GKYD**VVANT	0.67
5	K773	MGIET**PKKE**GGTKS	0.61
6	K676	GSPHD**QKEP**HYTNL	0.39
7	K774	GIETP**KKEG**GTKSW	0.31

Table 4. List of Pipsqueak sumoylation sites. Predicted Sumoylation sites found in Pipsqueak after analysing its aa sequence with SUMOplot™ Analysis program. This program predicts the probability for the SUMO consensus sequence (ΨKXE) to be engaged in SUMO attachment. The sites are sorted by score.

Our first approach to understand the effect of sumoylation on Psq function was very general; we just tested the consequences of mutating several molecules of the sumoylation pathway, making it difficult to draw any specific conclusion. As a more specific approach, we decided to mutate the sumoylation sites in Psq and analyse the impact of these mutations on its oncogenic capacity *in vivo*, as well as on its ability to interact with other proteins, such as MEP-1 or DmTAF3. For this we generated three Psq mutants: one in which both sumoylation sites are mutated and two with only one site mutated (**Figure 29**). Like that, we could discern which sumoylation site is functional: Lys518, Lys633, both of them, or none.

The mutation consists in replacing the lysine residue of each of the sumoylation sites for an arginine. Although this aa is also basic and has a similar size, this mutation completely abolishes the sumoylation process (Krumova *et al.*, 2011; Liu *et al.*, 2007b; Riquelme *et al.*, 2006; Spengler *et al.*, 2002). To generate these Psq mutants, first a *psq* isoform B transgene was synthetized (GeneArt). This transgene was cloned in the *Drosophila* vector *pUAST* that allows transient controlled expression in cell culture experiments and also generation of transgenic fly lines. To produce the three mutant versions of Psq protein we tried to perform directed mutagenesis but it did not work out due to the big size of the Psq plasmid used as template (13300bp). Our alternative

strategy was conducting directed mutagenesis on a smaller fragment of Psq that contained both sumoylation sites and then subclone these fragments in the original plasmid by substituting the wild-type sequence (see *Materials and Methods* for more details).

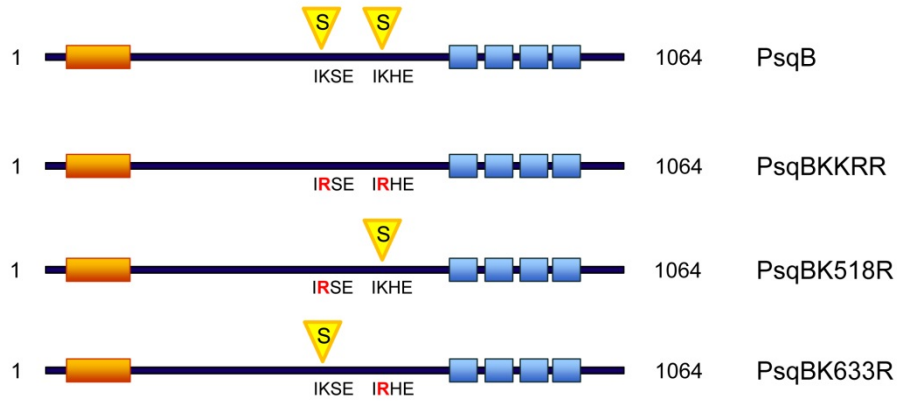


Figure 29. Mutagenesis of Pipsqueak sumoylation sites. Schema showing PsqB protein and its main domains. The two sumoylation sites represented were predicted with the highest probability by SUMOplot™ Analysis program, and are located within the SID with the SUMO ligase Su(var)2-10. In the mutant versions of Psq, the lysine (K) residue where SUMO is attached was substituted by an arginine (R).

To facilitate the biochemistry assays, these PsqB constructs were tagged on its amino-terminal part with 3 repetitions of the FLAG tag and in the carboxy-terminal part with a variant of the Red Fluorescent Protein (RFP) called Tag-RFP-T. Tagging both ends of the protein is also important regarding the second part of the results of this Thesis, as at the time of designing these constructs we already had evidences of Psq protein being processed (*Results* section part 2).

First of all we tested if Psq was a target of sumoylation by performing co-immunoprecipitation assays in S2 cells followed by Western-Blot (WB). We overexpressed SUMO using the construct *UAS-HA::SUMO* in combination with our Psq wild-type construct, *UAS-3xFLAG::psq::Tag-RFP-T* (hereafter abbreviated as $psqB^{Tag}$) (**Figure 30A** and **B**, lanes 1-4). On **Figure 30A**, in lane 4 we can observe the co-immunoprecipitation of SUMO-HA with Psq (immunoprecipitated with an anti-FLAG antibody); while in **Figure 30B** lane 4 shows the co-immunoprecipitation of Psq when SUMO-HA is immunoprecipitated using an anti-HA antibody. These results indicate that Psq is sumoylated *in vivo* in an heterologous system.

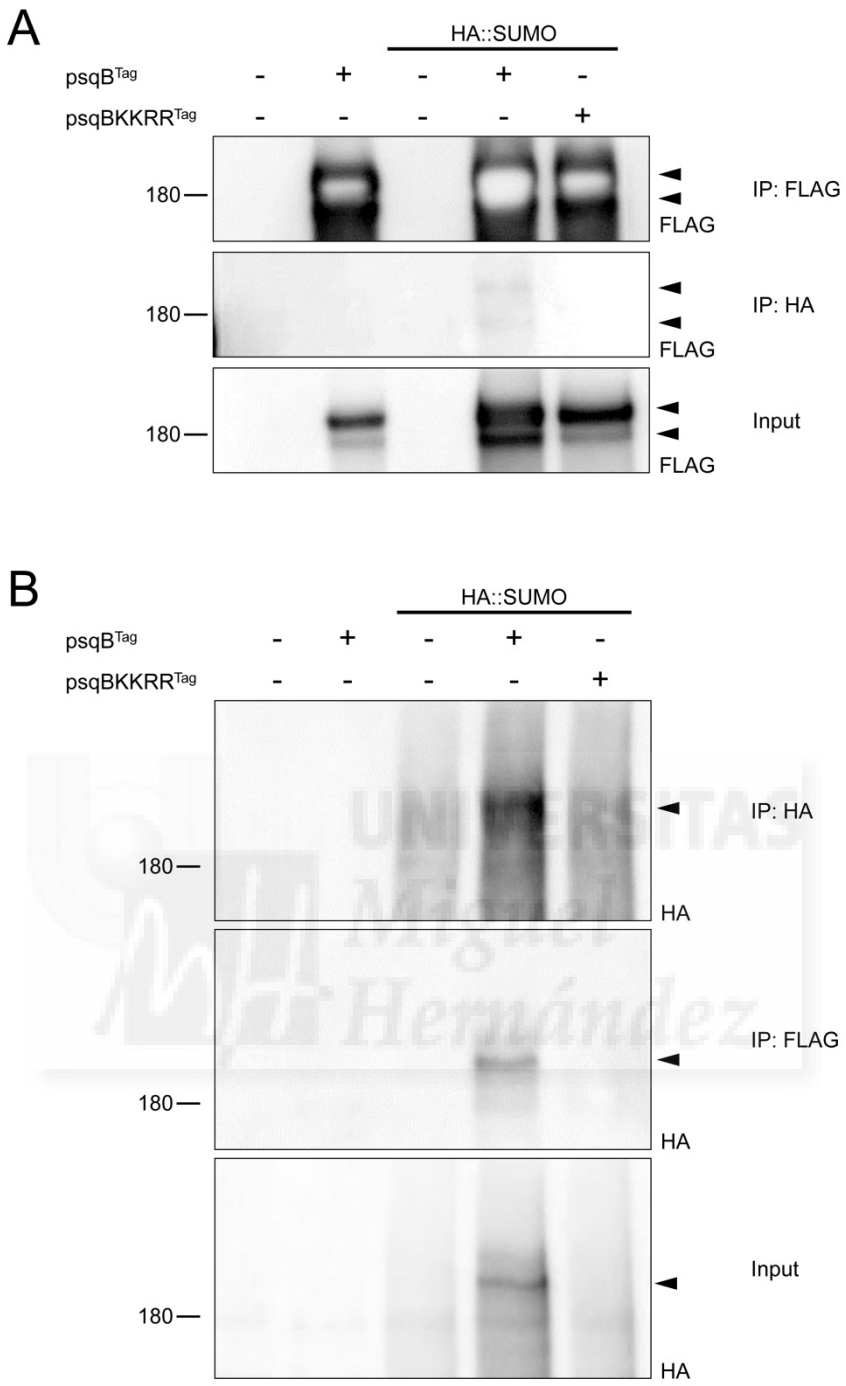


Figure 30. Pipsqueak is sumoylated in S2 cells. Psq-SUMO co-inmunoprecipitation in S2 cells. As indicated, the constructs *psqB^{Tag}*, *psqBKKRR^{Tag}*, and *UAS-HA::SUMO* were transiently transfected into S2 cells. The protein lysates were immunoprecipitated with anti-FLAG or anti-HA antibodies, and the immunoprecipitates were analysed in WB probed with antibodies against FLAG (**A**) and HA (**B**). The input panel shows Psq in the total cell lysates (**A**) and the sumoylated proteins as a smear (**B**). In all panels arrows indicate sumoylated Psq.

In this experiment we also wanted to test the functionality of the chosen sumoylation site candidates, Lys518 and Lys633. For that, we overexpressed SUMO in combination with the mutant construct *3xFLAG::psqKKRR::Tag-RFP-T* (hereafter *psqBKKRR^{Tag}*) (**Figure 30A** and **B**, lane 5). As it is shown in the blot, mutation of both sumoylation sites blocks the interaction between SUMO and Psq, while the interaction still occurs with PsqB^{Tag}. This result demonstrates that the only functional sumoylation sites in Psq are Lys518, Lys633 or both.

As an additional confirmation of Psq sumoylation, we co-expressed the construct *psqB^{Tag}* together with either *UAS-HA::SUMO* or *UAS-HA::SUMOΔGG*. As mentioned in the *Introduction*, the SUMO molecule needs to be proteolytically processed to be mature and active. This processing reveals a carboxy-terminal Gly-Gly motif that enables conjugation with the E1 activating enzyme AOS1–UBA2 and the start of the sumoylation pathway. The mutant *UAS-HA::SUMOΔGG* lacks this motif, so it cannot be bound to target proteins. In **Figure 31A** and **B** we can observe that wild-type HA::SUMO is able to conjugate to Psq (line 4), whereas the interaction is abolished when the mutant *HA::SUMOΔGG* construct is expressed (line 5), as expected.

The results shown in **Figure 30** and **31** demonstrate that Psq is sumoylated. Now, we wondered if both Lys518 and Lys633 were functional SUMO acceptors. Our experiments have already discarded 5 candidates (**Table 4**). To assess the functionality of the two remaining sumoylation sites we performed new co-inmunoprecipitation assays in S2 cells using this time the Psq constructs containing only one mutated sumoylation site: *3xFLAG::psqK518R::Tag-RFP-T* and *3xFLAG::psqK633R::Tag-RFP-T* (abbreviated as *psqBK518R^{Tag}* and *psqBK633R^{Tag}* from now on, respectively). While mutation at Lys518 does not affect SUMO binding to Psq (**Figure 32**, lane 5), mutation at Lys633 completely abolishes the interaction between both proteins (**Figure 32**, lane 6), demonstrating that Lys633 is the only functional sumoylation site in Psq.

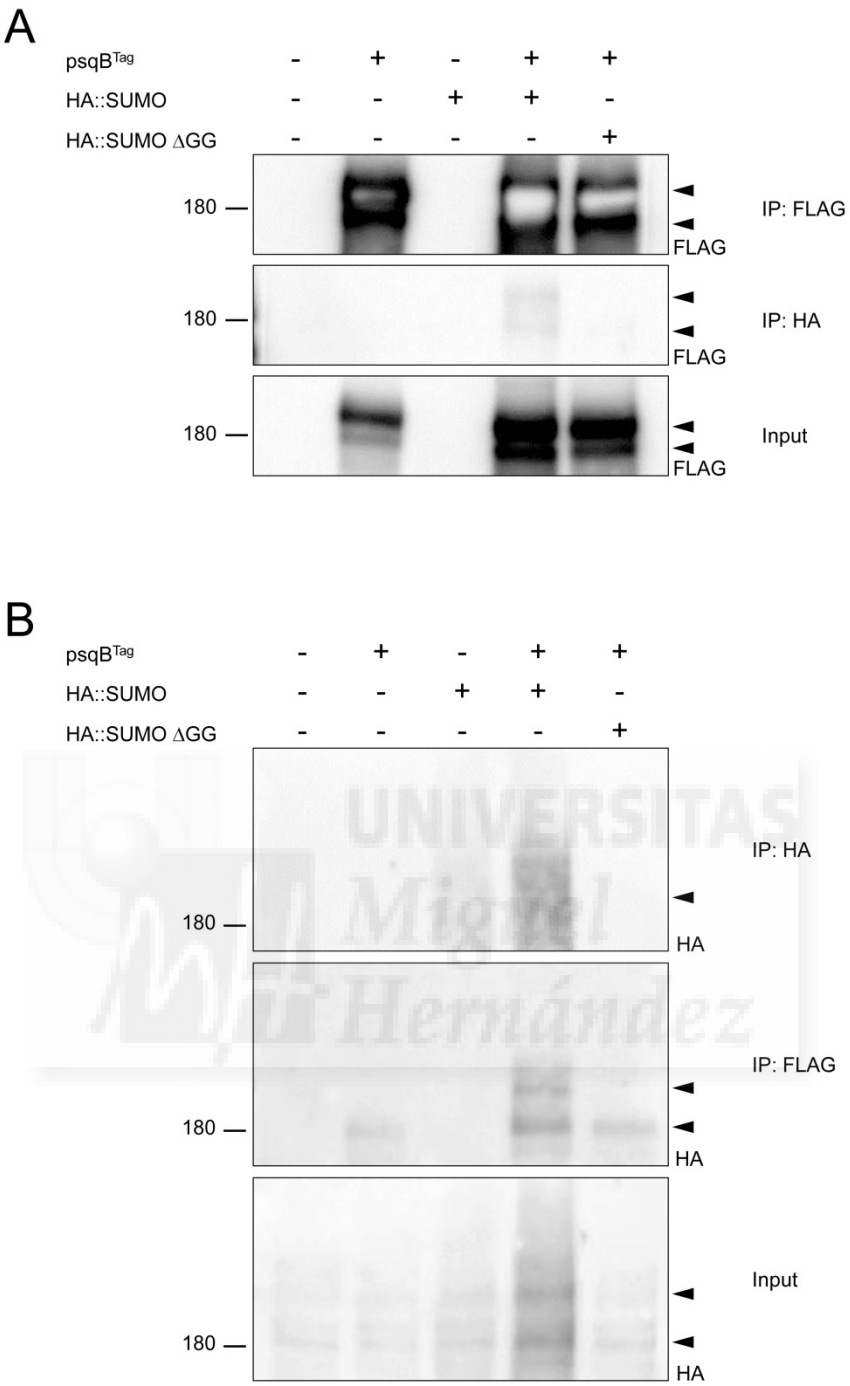


Figure 31. Pipsqueak is sumoylated in S2 cells. Psq-SUMO co-inmunoprecipitation in S2 cells. As indicated, the constructs *UAS-psqB^{Tag}*, *UAS-HA::SUMO* and *UAS-HA::SUMO⊿GG* were transiently transfected into S2 cells. The protein lysates were immunoprecipitated with anti-FLAG or anti-HA antibodies and the immunoprecipitates were analysed in WB probed with antibodies against FLAG (**A**) and HA (**B**). The input panel shows Psq in the total cell lysates (**A**) and the sumoylated proteins as a smear (**B**). In all panels arrows indicate sumoylated Psq.

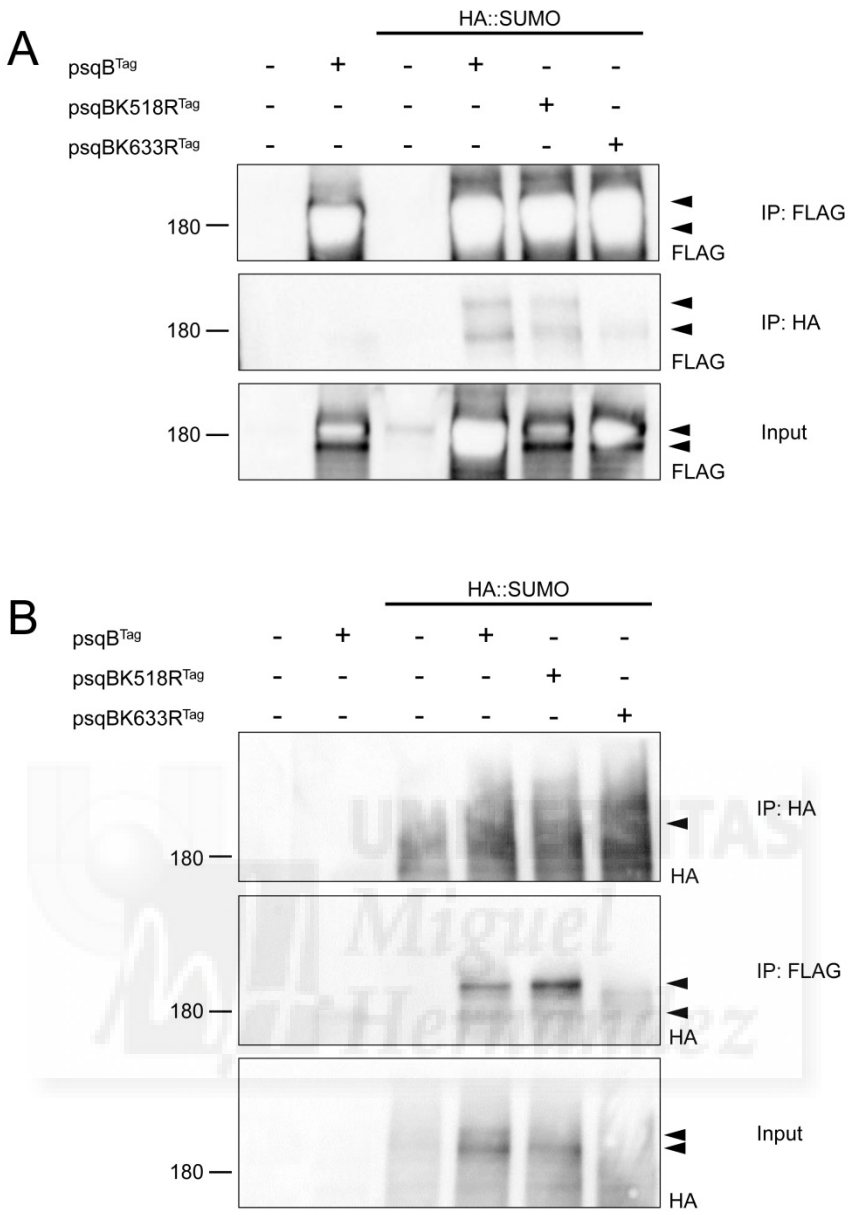


Figure 32. Pipsqueak is sumoylated at Lysine 633. Psq-SUMO co-inmunoprecipitation in S2 cells. As indicated, the constructs *psqB^{Tag}*, *psqBK518R^{Tag}, psqBK633R^{Tag}* and *UAS-HA::SUMO* were transiently transfected into S2 cells. The protein lysates were immunoprecipitated with anti-FLAG or anti-HA antibodies and the immunoprecipitates were analysed in WB probed with antibodies against FLAG (**A**) and HA (**B**). The input panel shows Psq in the total cell lysates (**A**) and the sumoylated proteins as a smear (**B**). In all panels arrows indicate sumoylated Psq.

In conclusion, we have found that Psq is subjected to sumoylation specifically at Lys633. All Psq isoforms, with or without BTB, share the region that contains the functional sumoylation site (**Figure 7B**). Thus, in addition to having different isoforms, sumoylation could serve as an additional mechanism to modulate Psq activities. There are evidences that sumoylation of transcription factors generally translates into

transcriptional repression (see *Introduction*), but how sumoylation contributes to Psq function still remains unclear. To address this question we used the *pUAST* plasmids to generate transgenic flies and study the effect of specifically blocking Psq sumoylation on flies *in vivo*. The results derived from these experiments will be shown in following sections of this Thesis.

5.2 MEP-1 has a putative SIM located in the region that recognizes Pipsqueak

MEP-1 is a sumoylation dependent transcriptional repressor, meaning that it can bind to sumoylated proteins to exert its functions. Having demonstrated that Psq can be sumoylated at Lys633, we wondered if the Psq-MEP-1 interaction detected in our Y2H and genetic assays was mediated by sumoylation at Psq Lys633. For this hypothesis to be true, MEP-1 should possess a specific motif called SIM that allows SUMO recognition by non-covalent interaction (see *Introduction* and (Hannich *et al.*, 2005; Hecker *et al.*, 2006; Minty *et al.*, 2000; Song *et al.*, 2004).

In general, a SIM has a hydrophobic nucleus, Val/Ile-Val/Ile-X-Val/Ile, that is surrounded by several acidic and/or serine residues that flank the core sequence either on the amino or carboxyl part. The third position in the hydrophobic nucleus (X) is frequently occupied by an acidic aa (Hannich *et al.*, 2005; Hecker *et al.*, 2006; Minty *et al.*, 2000; Sekiyama *et al.*, 2008; Song *et al.*, 2004). Using the ScanProsite software (http://prosite.expasy.org/scanprotiste/), we analysed MEP-1 aa sequence and we found a short sequence inside the region of interaction with Psq that possesses these characteristics (VVNLDEDSDEE, from aa 241 to 251, **Figure 33A**). It is important to note that this sequence is contained in the MEP-1 region interacting with Psq that we defined thanks to the Y2H (aa 211-480). In addition to this, it is imperative that a SIM fulfils certain criteria, such as being localized outside of globular domains or in protein areas with known domains (Vogt and Hofmann, 2012). Furthermore, all the SIM domains described up to date in the literature are found in protein regions with low structural complexity. The analysis of the MEP-1 protein sequence using web tools such as Globprot (http://globplot.embl.de/cgiDict.py) or IUPred (http://iupred.enzim.hu/pred.php) shows that the predicted SIM is accordingly located in a disorganized region (**Figure 33A**).

Another characteristic shared by all SIMs is that they are highly conserved sequences flanked by non-conserved sequences. Using BLAST (http://blast.ncbi.nlm.nih.gov/) and UCSC Genome Bioinformatics Site

(https://genome.ucsc.edu/) we saw that among related *Drosophila* species, the putative MEP-1 SIM sequence is highly conserved (**Figure 33B**). Additionally, in some of them (*e.g. D. willinstoni*, *D. mojavensis*) the SIM motif is flanked by non-conserved sequences, meeting the criteria. Therefore, after these *in silico* studies we concluded that the region of MEP-1 interacting with Psq, defined thanks to the Y2H assay, contains a putative SIM that meets all the criteria required to be functional. Putting together our biochemical data showing Psq sumoylation and the *in silico* defined MEP-1 SIM, we propose that the interaction between Psq and the transcriptional repressor MEP-1 occurs through the recognition of sumoylated Psq by the SIM found in MEP-1.

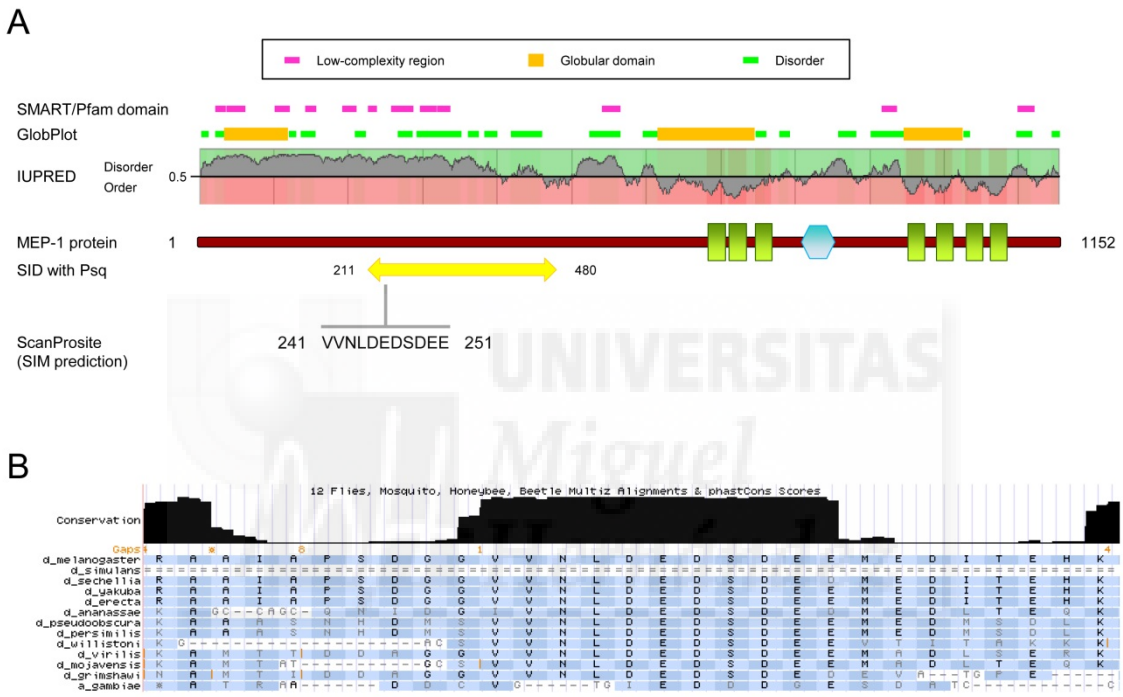


Figure 33. MEP-1 has a SIM in the region that interacts with Pipsqueak. *In silico* analysis of MEP-1 aa sequence. Several bioinformatics resources such as ScanProsite, GlobPlot, IUPRED, Blast or UCSC Genome Browser were used to identify putative SIM domains in the sequence of MEP-1 (**A**) and check their conservation (**B**). Additional analysis of MEP-1 domain composition and ordered/disordered regions was necessary to define the SIM. The alignment depicts *Drosophila melanogaster* MEP-1 protein sequence compared to homologous sequences in other species.

5.3 DmTAF3 has also a putative SIM located in the region that binds to Pipsqueak

In the previous section we have described a sumoylation-dependent mechanism of interaction between Psq and MEP-1 that also explains the interaction with Su(var)2-10. However, the mechanism underlying Psq and DmTAF3 interaction still remains unknown. After the 1-by-1 experiment we found that Psq BTB domain is not necessary for the interaction with DmTAF3. Actually, and very interestingly, DmTAF3 binds to the same region that interacts with Su(var)2-10 and contains the functional Psq sumoylation site (**Figures 26**, **29** and **32**). Previous results from our and other groups have shown that there are several proteins with different structures which are able to interact with the same region of DmTAF3 (**Figure 20**). Once we ruled out the possibility of the BTB being the key element to interact with this region with our 1-by-1 experiments (**Figure 21**), we wondered if there is a specific domain or post-translational modification site common to all of these proteins that allows this interaction. To explore these options we carried a careful sequence analysis of DmTAF3 protein, and also of the proteins that interact with it, looking for motifs, domains or post-translational modification sites that could explain this interaction.

We started by defining more specifically which is the common DmTAF3 sequence recognized by all the DmTAF3 partners known to date (www.theBIOGRID.org). Combining our results with the Y2H data already published for other proteins (https://pimr.hybrigenics.com/) we were able to define a small DmTAF3 region of 20 aa present in all the interactions studied (**Figure 20**). Surprisingly, our previously acquired knowledge in the field of sumoylation came in handy as we realized that in this small sequence there is a SIM. Actually, using again the ScanProsite software (http://prosite.expasy.org/scanprotiste/), we identified three different small motifs in DmTAF3 that possess the canonical SIM sequence (**Figure 34A**). Out of the three, only one of them (PIEVSDDSDES, from aa 866 to 878) is located in the SID with Psq (**Figure 17**). The analysis of DmTAF3 protein sequence with Globprot (http://globplot.embl.de/cgiDict.py) and IUPred (http://iupred.enzim.hu/pred.php) revealed that two of the three putative SIM were localized in highly structured globular domains of the protein or in ordered regions, being the SIM present in the interaction area with Psq (PIEVSDDSDES), the only located in a disorganized region. Finally, using BLAST (http://blast.ncbi.nlm.nih.gov/) and UCSC Genome Bioinformatics Site (https://genome.ucsc.edu/) we did a sequence comparison of different DmTAF3 homologs present in related animal species to check the conservation of the three SIMs (**Figure 34B**). We observed that only the SIM

located in the region of interaction with Psq was conserved. Thus, after these *in silico* studies we concluded that, of the three SIMs initially identified, only the one contained in the region of interaction with Psq, defined thanks to the Y2H, meets all the criteria needed to be functional.

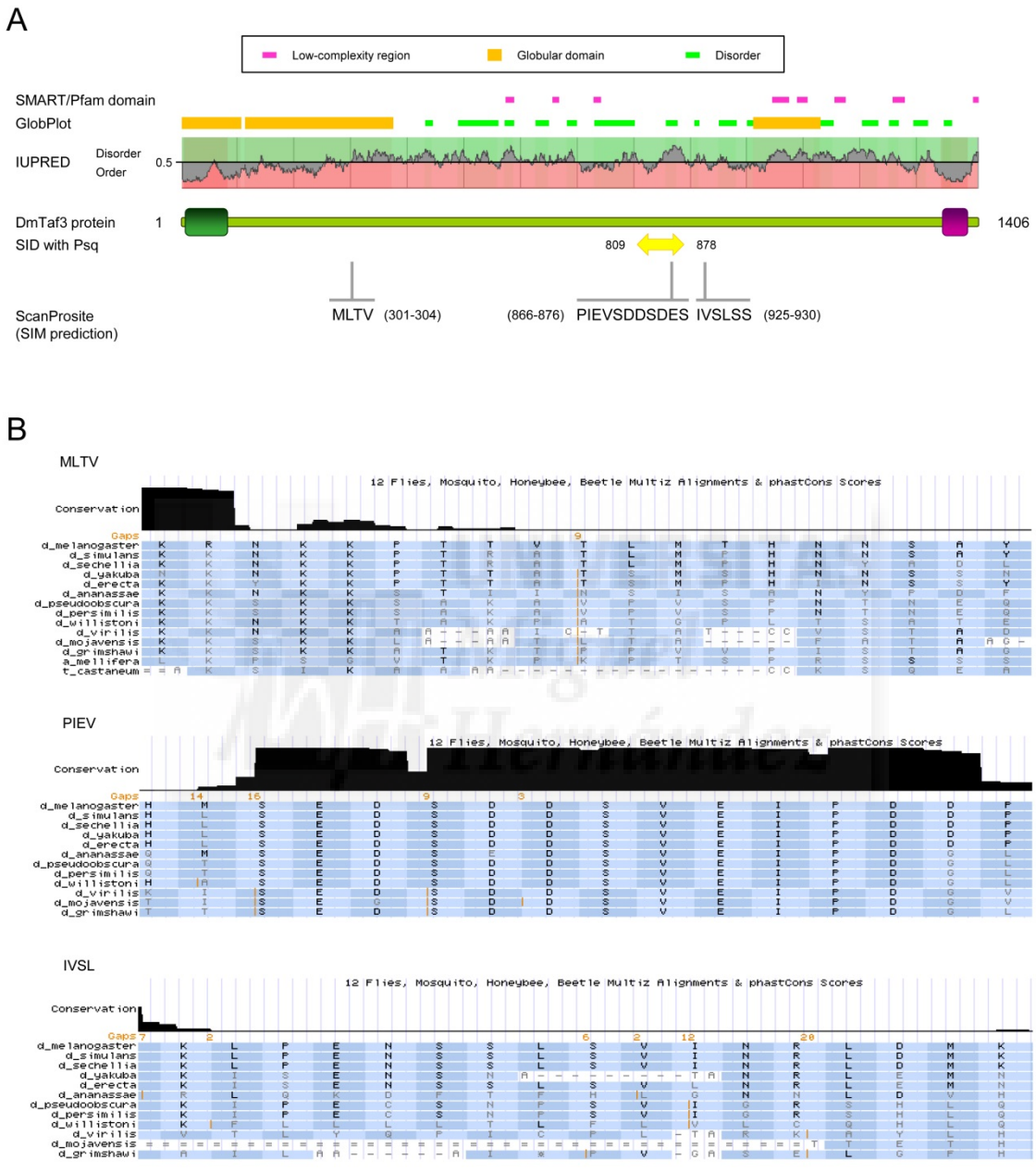


Figure 34. DmTAF3 has a SIM in the region that interacts with Pipsqueak. *In silico* analysis of DmTAF3 aa sequence. Several bioinformatics resources such as ScanProsite, GlobPlot, IUPRED, Blast or UCSC Genome Browser were used to identify putative SIM domains in the sequence of DmTAF3 (**A**) and check their conservation (**B**). Additional analysis of DmTAF3 domain composition and ordered/disordered regions was necessary to define the SIM. The alignment depicts *Drosophila melanogaster* DmTAF3 protein sequence compared to homologous sequences in other species.

Again, combined with our biochemical data indicating Psq sumoylation, the *in silico* definition of a proper DmTAF3 SIM leads us to propose that the interaction between Psq and DmTAF3 depends on SUMO, as it would occur with Psq-MEP-1 interaction. We have shown that Psq has a functional sumoylation site and two interacting partners of Psq, MEP-1 and DmTAF3, have putative SIMs able to recognize sumoylated proteins. Still, a definitive functional validation of these *in silico* defined SIMs should be performed to confirm their biological significance (see *Discussion*). Nevertheless, it is of utmost importance to unveil the relevance of such interactions for Psq activity, both in physiological conditions and in an oncogenic context.

This SIM could explain the interaction between Psq and DmTAF3 but, what about the rest of the partners already found for DmTAF3 in previous studies? Analysing the protein sequence of these proteins we were surprised again when we found that most of them have putative sumoylation sites and some of them have already been proved to be functional, like in the case of p53 (Wu and Chiang, 2009a). Could it be that DmTAF3 uses the SIM to recognize the SUMO groups attached to these proteins allowing the interaction with partners with different structures? Then, is SUMO the common feature that explains all of these interactions?

In parallel to the study of DmTAF3 sequence, Dr. Alisson Gontijo and Dr. Esther Caparrós performed an analysis of the protein sequences of all BTB transcription factors present in *Drosophila*, trying to unveil how common this kind of SUMO-dependent modulation could be. They found that many of these proteins, including Broad (K273, K325, K497), Fruitless (K546), Trithorax-like (K268, K373) or Lola (K127, K385, K506, K872), contain putative sumoylation sites and these are frequently located in the central region of the molecule, between the BTB and the DNA-binding domains, as it occurs with Psq. This finding suggests that this kind of regulation could not only be specific for Psq, but far more general than what we thought at the beginning.

6. Role of sumoylation on Pipsqueak activity *in vivo*

In this section, we take advantage of the Psq constructs to answer several questions: is sumoylation necessary for Psq function? If this is the case, does it positively contribute to Psq function or does it act as a brake? Moreover, we can use them to study how Psq sumoylation affects its tumorigenic capacity and its ability to interact with the basal transcriptional machinery through DmTAF3. Given that both the sumoylation sites and DmTAF3 interaction with BTB proteins are conserved, our results can be used as a general model for interactions between BTB proteins and DmTAF3 *in vivo* and, particularly, as a model of an oncogenic mechanism mediated by BTB proteins. The next results also represent a starting point to infer the biological significance of the interaction between Psq and MEP-1.

6.1 Sumoylation attenuates Pipsqueak capacity as transcriptional activator

To determine the functional significance of Pipsqueak sumoylation, we started by studying the effect of blocking Psq sumoylation on its transcriptional activity. For that we transfected the different Psq constructs carrying the mutations in the sumoylation sites in S2 cells and measured the relative mRNA levels of two read-outs of Psq activity by quantitative polymerase chain reaction (qPCR). The first one is *CG9925*, a gene expressed in the germline of both male and female flies early during gametogenesis. Microarray and qPCR experiments performed in our lab demonstrate that there is an increase in this gene's mRNA levels in response to *psq* expression, specifically the BTB containing isoforms (unpublished data). In the case of the second read-out, *torso*, it shows increased mRNA levels in response to *psq* expression, independently of the isoform (Grillo *et al.*, 2011). In our case we used the BTB containing isoform B of Psq to make the different constructs, so both genes should be equally good read-outs.

Mutation in Lys518 does not have a significant effect over Psq activity, as *CG9925* levels are very similar to the wild-type construct condition. On the other hand, mutation in Lys633 and the construct with mutations in both sumoylation sites promote a 6.5 and 6 fold change increase in *CG9925* expression, respectively (**Figure 35A**). In the case of *torso* (**Figure 35B**), the tendency is similar. When Lys633 is mutated there is a 27.3 (*psqBK633R*^{Tag}) and 39.7 (*psqBKKRR*^{Tag}) fold change increase in *torso* expression. These results indicate that Psq sumoylation somehow attenuates Psq

transcriptional activity, at least in the case of *CG9925* and *torso*. Taking into account that MEP-1 is involved in sumoylation-dependent transcriptional repression, these observations begin to shed light in the biological significance of Psq-MEP-1 interaction. Moreover, together with the co-immunoprecipitation assays, these data support that the only functional sumoylation site in Psq is located at Lys633.

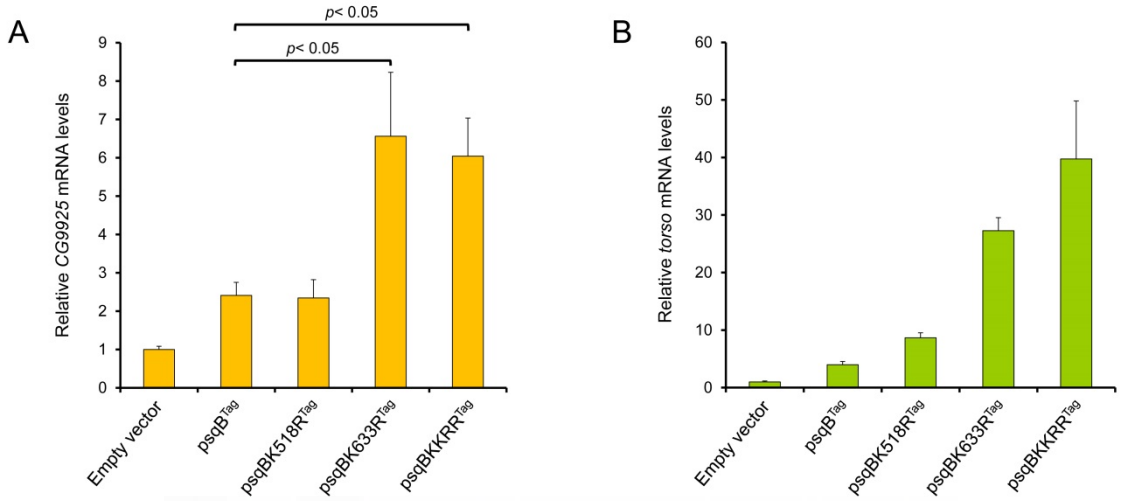


Figure 35. Blocking Pipsqueak sumoylation increases its transcriptional activity. Quantification of the relative change in mRNA levels of *CG9925* (**A**) or *torso* (**B**) by qPCR in S2 cells transfected with the *psq* constructs carrying mutations in the sumoylation sites. Mutation in Lys633 but not in Lys518 increases the relative mRNA levels of both *CG9925* and *torso*, compared to $psqB^{Tag}$. In the case of *CG9925,* n = 5 independent biological replicates. In the case of *torso*, the differences were not statistically significant ($psqB^{Tag}$ vs $psqBK633R^{Tag}$, p = 0.12; $psqB^{Tag}$ vs $psqBKKRR^{Tag}$, p = 0.15). n = 3 independent biological replicates. The *p*-value was calculated using two-tailed Student's *t*-test. The data are presented as the mean + SEM (standard error of the mean).

6.2 Overexpression of sumoylation-defective Pipsqueak mutants causes loss of eye tissue due to an increase in apoptosis

In parallel to the cell culture assays, we generated transgenic flies carrying the constructs $psqB^{Tag}$, $psqBK518R^{Tag}$, $psqBK633R^{Tag}$ and $psqBKKRR^{Tag}$. we have used these transgenic flies to assay the consequences of blocking Psq sumoylation *in vivo* and study how this affects Psq tumorigenic capacity. With this objective in mind, we used again the UAS/Gal4 binary system to overexpress the different *psq* mutants using the *ey-Gal4* line to drive their expression in the eye of *Drosophila*. Under these conditions we observed that overexpression of the wild-type construct $psqB^{Tag}$ decreases the viability of the flies (approximately ~50% of the flies overexpressing the

construct die). In addition to this, around 12.5% of the *psqB*^{Tag} overexpressing flies that reach adulthood present smaller eyes compared to the control (**Figure 36A-B**). These phenotypic traits are also observed when we overexpress a construct containing the endogenous *psqB* without tags (data not shown). In the case of the flies overexpressing the mutant *psqBK518R*^{Tag}, 26.1% of the eyes observed in the F1 presented smaller eyes respect to the control, as it can be seen in **Figure 36C**, being the rest of them wild-type size, similar to what occurs when we overexpress *psqB*^{Tag}. However, mutation *psqBK633R*^{Tag} causes 100% lethality and the flies that reached the pupal stage (manually dissected from the pupa) had no head (**Figure 36D**). Overexpression of the mutant for both sumoylation sites (*psqBKKRR*^{Tag}) also caused 100% lethality during pupal stage, and flies did not develop the head (**Figure 36E**).

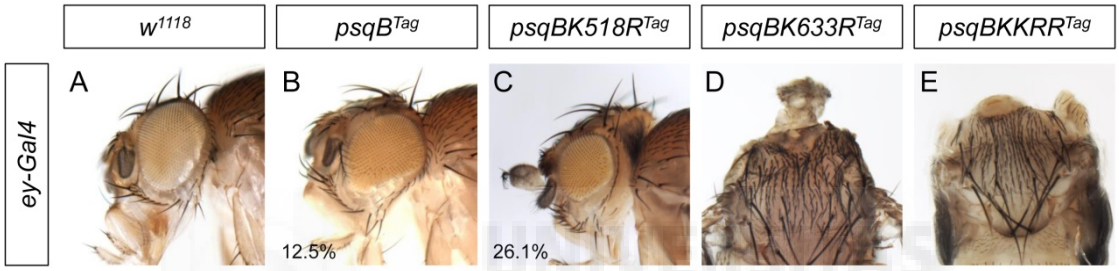


Figure 36. Overexpression of sumoylation-defective Pipsqueak mutants is lethal. Representative examples showing the adult eye phenotype obtained when overexpressing different wild-type or mutant *psq* constructs under the promoter *ey-Gal4*. Photographs were taken from female flies. Crosses were maintained at 25ºC.

Taken together, and added to previous results obtained in co-immunoprecipitation and qPCR experiments, these *in vivo* results support that the only functional sumoylation site in Psq is located at Lys633. In this way, it is logical that lines carrying the mutations K633R and KKRR have a similar phenotype, while the K518R mutant is more similar to the wild-type.

The principal aim of these *in vivo* approaches is to monitor Psq activity with the variegation and tumorigenesis assays to test how blocking Psq sumoylation affects its capacity to induce epigenetic silencing and tumorigenic growth. As single overexpression of these constructs with *ey-Gal4* is lethal, we cannot proceed as we did in our previous *in vivo* studies, that is, crossing these lines with *ey-Gal4>eyeful* and *ey-Gal4>Dl>eyeful* and quantifying changes in the percentage of variegated and tumorigenic eyes, respectively. To be able to perform these experiments, we should first find the cause of the lethality and rescue it.

As I already mentioned in the *Introduction*, *Drosophila* eye-antennal imaginal disc originates the majority of the structures present in the adult fly head. Abnormalities during the development of this imaginal disc can produce strong defects in the head that, in extreme cases, can lead to phenotypes similar to what we obtain when overexpressing Psq sumoylation mutants (with only the fly proboscis remaining intact). To look for morphological defects that could explain the "no head" phenotype we dissected L3 larval eye-antennal discs and performed immunohistochemistry stainings. Overexpression of *psqB*^{Tag} or the mutant *psqBK518R*^{Tag} using the *ey-Gal4* driver did not alter the normal pattern of expression of the morphological marker *wingless* or the post-mitotic cell marker *elav* compared to the control (**Figure 37A**-**C**). Contrary to this, *psqBK633R*^{Tag} and *psqBKKRR*^{Tag} overexpression produce a massive reduction in the eye territory of the eye-antennal disc (**Figure 37D**-**E**).

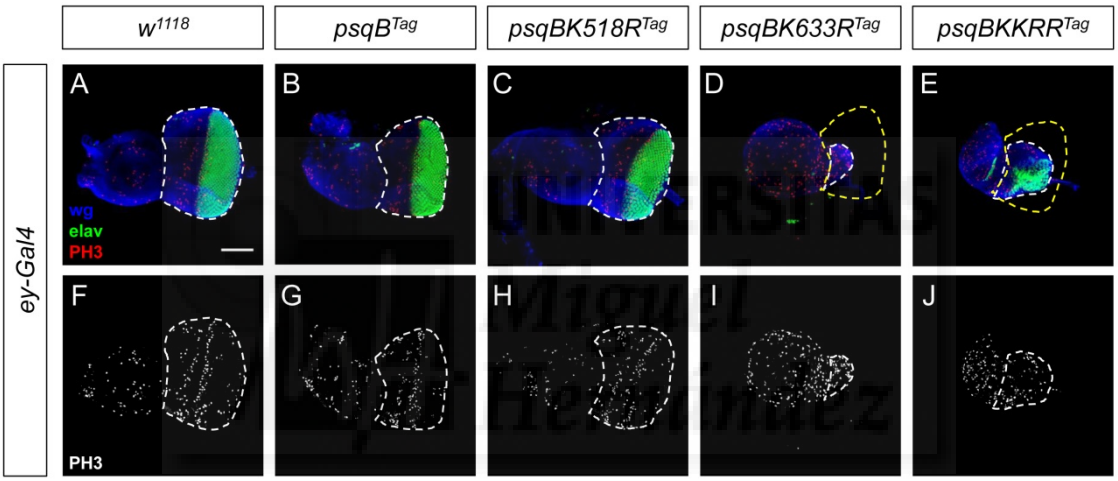


Figure 37. Overexpression of sumoylation-defective Pipsqueak mutants produces loss of eye tissue. Immunostaining of L3 eye-antennal discs. (**A**-**E**) Staining with anti-wingless (blue), anti-elav (green) and anti-PH3 (red) antibodies. The white dashed line delimitates the eye part of the eye-antennal disc. The yellow dashed line in panels **D** and **E** delineates a WT size disc for the sake of comparison. (**F**-**J**) Staining with anti-PH3 (grey) antibody. Scale bar in (**A**) represents 100 μm.

There are several explanations for the cause of this tissue loss: defects in cell proliferation, defects in cell differentiation or cell death. Initially there is no reduction in cell proliferation, as staining with phospho-histone H3 (PH3) did not show any defect in cell proliferation in the remaining eye tissue compared to the control (**Figure 37I**-**J**). A cell differentiation defect would be identified by a decrease on *elav* positive cells in a wild-type-like disc size, which is not the case. Finally, regarding cell death, evidences from our lab indicate that Psq has a pro-apoptotic function through the activation of the

gene *reaper* (unpublished data). There is a possibility that blocking Psq sumoylation boosts its pro-apoptotic activity leading to cell death. Overexpression of the *psqB*^{Tag} or the mutant *UAS- psqBK518R*^{Tag} did not induce apoptosis in the tissue, shown by lack of anti-cleaved caspase 3 staining (**Figure 38A**-**C** and **F-H**). However, in the case of *psqBK633R*^{Tag} or *psqBKKRR*^{Tag} overexpression, there is a clear increase of caspase 3 staining in the remaining eye tissue (**Figure 38D**, **E**, **I** and **J**). This result suggests that cell death is the underlying cause of the tissue loss observed when overexpressing sumoylation-defective Psq mutants.

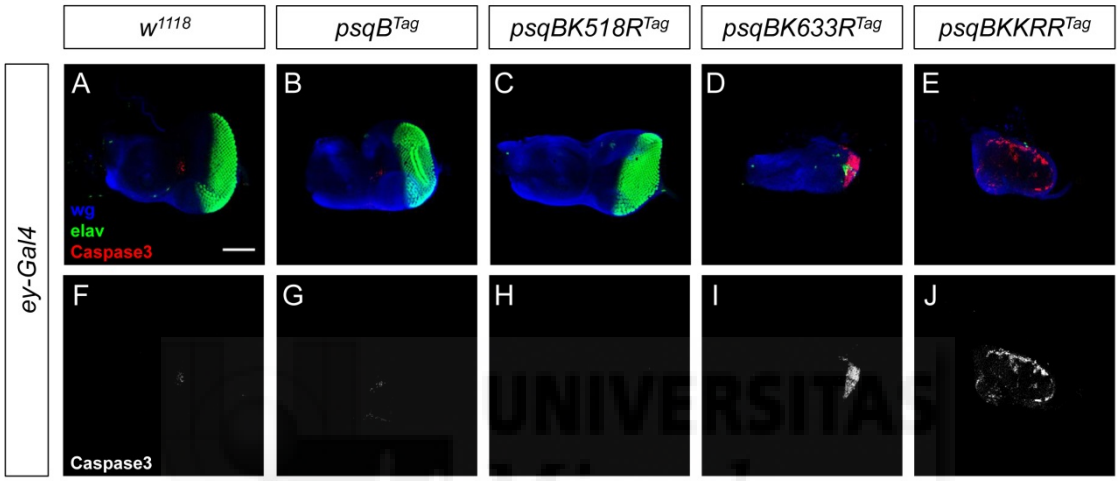


Figure 38. Overexpression of sumoylation-defective Pipsqueak mutants promotes apoptosis. Immunostaining of L3 eye-antennal discs. (**A**-**E**) Staining with anti-wingless (blue), anti-elav (green) and anti-cleaved caspase 3 (red) antibodies. (**F**-**J**) Staining with anti-cleaved caspase 3 (grey) antibody. Scale bar in (**A**) represents 100 μm.

The results shown in the first part of *Results* of this Thesis indicate that Psq sumoylation is of utmost importance for Psq proper function in the context of *Drosophila* eye development. Whether this phenotype is related to Psq interaction with MEP-1 or DmTAF3 remains unsolved and further research is required.

Part 2. Gene activation by a BTB transcription factor requires proteasome-dependent partial proteolysis

BTB transcription factors are encoded by a large and diverse family of genes that typically contain an amino-terminal BTB/POZ domain and a caboxy-terminal DNA-binding domain. The molecular relevance of this architectural combination is not fully understood. The elucidation of the molecular mode of action of these proteins may reveal important insights into the great success of this family in evolution. Alternative promoter usage at the *Drosophila pipsqueak* (*psq*) locus produces messenger RNAs that encode for DNA-binding proteins containing or not an N-terminal BTB domain. Psq short isoforms (without BTB) are essential for viability whereas Psq long isoforms (with BTB) are required maternally for fertility. Using transgenes containing one of the long Psq isoforms (PsqB) that was tagged both at the N- and C-termini, we have found that distinct Psq fragments originate directly from PsqB by partial proteolysis at two distinct sites located in the central unstructured region of Psq, hereafter named Type I and Type II partial proteolysis. Type I occurs in the insoluble fraction of the cell and generates N-terminal ~30-50 kDa + C-terminal ~120-150 kDa fragments, whereas Type II takes place in the soluble fraction and generates N-terminal 70 kDa + C-terminal 100 kDa fragments. Using mutated PsqB transgenes containing overlapping deletions, we show that Type I processing, but not Type II, is required for PsqB mediated *CG9925* transcriptional derepression. Strikingly, we find that chemical inhibition of proteasome activity abolishes PsqB Type I processing and consequentially its ability to derepress *CG9925*. These results suggest a model where the proteasome-dependent partial proteolysis of PsqB is required for its derepression activity, providing a new molecular mechanism of gene activation by a BTB transcription factor. Additionally, our experiments demonstrate that the sequence located between His241 and Gly428 is critical for PsqB nuclear localization, and suggest that both *CG9925* derepression and Psq-driven tumorigenesis depend on PsqB presence in the nucleus.

UNIVERSITAS Miguel Hernández

1. Pipsqueak is subjected to two mutually exclusive forms of processing

psq has eleven mRNA variants produced by alternative splicing and alternative promoter use (see *Introduction*). During the course of our experiments to determine if Psq was target of sumoylation, we repeatedly observed in several WB gels a few bands that could be either unspecific bands or smaller forms of Psq corresponding to the different isoforms. Initial pilot experiments investigating these putative Psq isoforms lead to the development of a parallel research line aimed at studying the impact of Psq proteolytic processing on its tumorigenic activity.

To study Psq processing in more detail we generated a new polyclonal antibody against a fragment located between Val453 and Gln552 of Psq protein (**Figure 39A**), which is common to all Psq isoforms (see *Introduction*). We then used this antibody to study the accumulation of the different bands corresponding to Psq protein in an overexpression paradigm. We crossed *hs-Gal4* flies with flies carrying the *UAS-psqB* transgene constructed by Jorge Bolívar (*see Materials and Methods*), to ectopically express *psq* isoform B (*psqB*) using the heat shock inducible promoter *hsp-70* in *Drosophila* larvae and adults. When we analysed the lysates in a WB, we saw the accumulation over time of bands with a smaller size (90 and ~130 kDa, white arrow heads, **Figure 39B**) than the expected full length PsqB (160 kDa, black arrow head, **Figure 39B**). We also noticed that the ratio between both increasing smaller isoforms was slightly different between larvae and adults, but we have not explored this difference further. Importantly, both the 90 and the ~130 kDa bands show an increase in their intensities over time after the heat shock, suggesting that the *UAS-psqB* transgene expression is linked with the generation of the smaller products. There are two possible explanations for this. One is that PsqB expression upregulates the expression of endogenous smaller *psq* isoforms. The other explanation is that that full length PsqB is processed into the smaller forms by some unknown mechanism. To test the latter hypothesis, we decided to generate a double tagged version of PsqB, with 3 repetitions of the FLAG tag at the amino-terminal part and the monomeric red (orange) fluorescent protein Tag-RFP-T at the carboxy-terminal part, generating the construct *3xFLAG::psqB::Tag-RFP-T* (abbreviated as $psqB^{Tag}$) (**Figure 40D**).

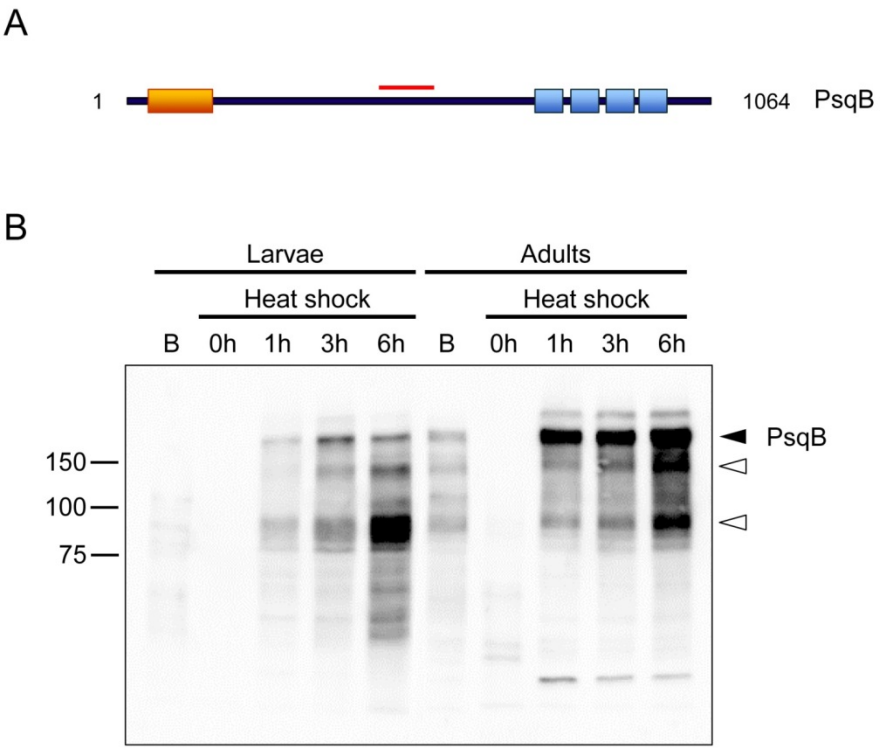


Figure 39. Pipsqueak protein profile in WB using anti-Psq antibody. (**A**) Schematic representation of Psq isoform B protein. The epitope recognized by anti-Psq antibody (from Val453 to Gln552) is depicted in red. (**B**) The presence of PsqB protein was detected with anti-Psq antibody in lysates of *hs-Gal4/UAS-psqB Drosophila* larvae and adults analysed by WB. The levels of Psq were measured before (B) and after heat shocking the samples for 10 minutes at 37ºC at the indicated times (0, 1, 3 and 6 hours after heat shock).

Transient transfection of *psqB*^{Tag} under the control of the inducible *pMT* promoter in S2 cells followed by WB showed that PsqB is processed giving rise to two different fragments, each of them with one of the tags (**Figure 40A-C**). This result rules out that the smaller bands detected in the heat shock experiment with the anti-Psq antibody are products of different endogenous isoforms and strongly supports that full length PsqB can be processed into smaller fragments through a still unknown mechanism. The size of the fragments suggested that there could be a post-translational mechanism that uncoupled the amino-terminal BTB domain from the carboxy-terminal DNA-binding domain of PsqB. If this were true, maybe the processed amino-terminal fragments were soluble, while the carboxy-terminal were DNA bound and/or insoluble.

To learn more about the subcellular distribution of the processed Psq-PB fragments we fractionated *psqB*^{Tag} transiently transfected S2 cells into soluble and insoluble fractions. Our results show that the processing of Psq-PB is different

depending on the cell fraction analysed. In the insoluble fraction (including cytoskeleton, intermediate filaments, nuclear matrix) we can observe the full length PsqB with tags (180 kDa), plus a 140 kDa band recognized by the anti-Tag-RFP-T (**Figure 40B**) and the anti-Psq antibodies (**Figure 40C**) corresponding to the carboxy-terminal part of the protein, and a 40 kDa band recognized by the anti-FLAG antibody and not by anti-Psq or anti-Tag-RFP-T corresponding to the amino-terminal part (**Figure 40A**). On the other hand, in the soluble fraction (including the citoplasm, DNA), we can detect a 110 kDa Tag-RFP-T-tagged band (**Figure 40B**) and a 70 kDa FLAG-tagged band (**Figure 40A**) in addition to the full length $PsqB^{Tag}$ (**Figure 40A** y **B**). We can also detect the 110 kDa Tag-RFP-T-tagged band (**Figure 40B**) in the insoluble fraction, probably due to a small contamination of the insoluble fraction with soluble material. We decided to call the processing in the insoluble fraction "Type I processing" and the one taking place in the soluble fraction "Type II processing" (**Figure 40D**). Note that the size of the $PsqB^{Tag}$ fusion protein (180 kDa, PsqB plus tags) is bigger than endogenous PsqB (160 kDa) and, therefore, the proteolytic fragments shown in **Figure 40** present a different size compared with their endogenous counterparts. Namely, it is possible that the ~130 kDa and ~90 kDa fragments in **Figure 39** correspond to the carboxy-terminal 140 kDa and 110 kDa Tag-RFP-T-tagged bands showed in **Figure 40** as result of Type I and II processing, respectively, while the amino-terminal fragments are not detected by the anti-Psq antibody in **Figure 39**). The fact that there was a 40 kDa amino-terminal fragment with no possible DNA-binding domain in the insoluble fraction raised the possibility that it was covalently bound to something. Also, we noticed that the sum of the apparent molecular weight of the respective amino- and carboxy-terminal parts of the soluble or insoluble fractions are approximately equivalent to the weight of the full-length Psq-PB proteins (**Figure 40D**). These results suggested that, contrary to our expectations, the PsqB processed fragments were not solubilized, but rather retained in the same subcellular compartments where the full-length PsqB once was. Hence, maybe the PsqB fragments were still covalently bonded despite being proteolytically processed.

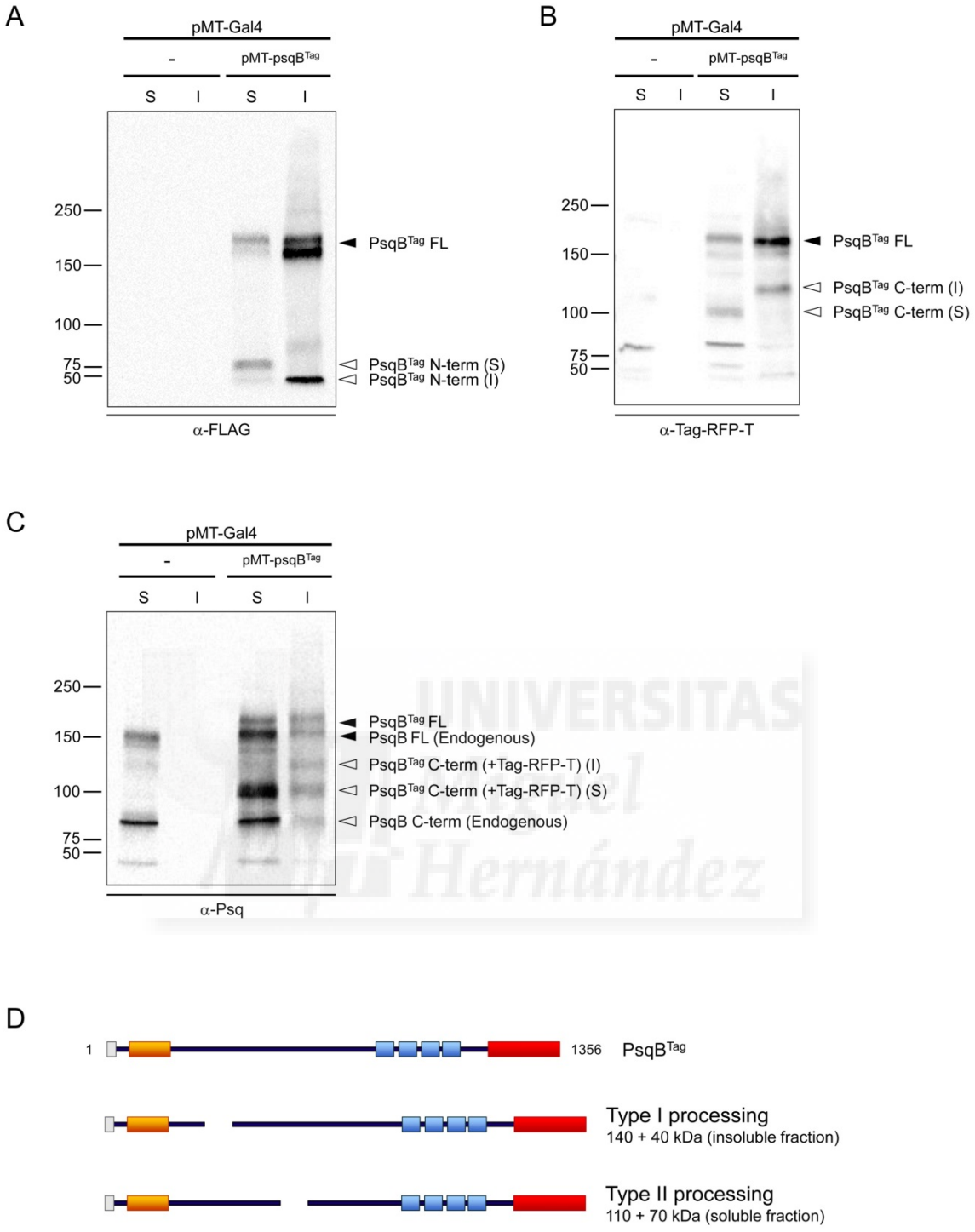


Figure 40. Pipsqueak is subjected to two different types of processing. WB of lysates from S2 cells transfected with *pMT-psqB^{Tag}* or empty vector. Soluble (S) and insoluble (I) cell fractions were separated by centrifugation for their analysis. The protein lysates were analysed by WB with antibodies against FLAG (**A**), Tag-RFP-T (**B**) or Psq (**C**). (**D**) Schematic representation of PsqB^{Tag} protein and the fragments derived from each type of processing.

To test the hypothesis that the PsqB^{Tag} fragments remained attached to each other via disulphide bonds after processing, we made lysates from S2 cells transfected with the *pMT-psqB^{Tag}* construct in the presence or absence of the reducing agent β-mercaptoethanol. As we can see in **Figure 41**, addition of β-mercaptoethanol had no effect over the soluble Type II processing (70 kDa + 110 kDa bands), whereas it liberates the 140 and 40 kDa-fragments produced by Type I processing. Hence, the attachment of the insoluble PsqB fragments generated by Type I processing is mediated by disulphide bonds. The meaning of this differentiating feature between both processing types remains unclear, but it suggests that these two proteolytic events are probably mediated by different molecular mechanisms.

Interestingly, and probably due to a better resolution in these particular gels, we were able to observe that Type I processing does not produce 2 Psq fragments, but rather a ladder of amino- and carboxy-terminal fragments (~30-50 kDa and ~120-150 kDa, respectively). This result will be discussed in following sections of this Thesis, but it suggests that Type I processing might not have a specific cleavage site.

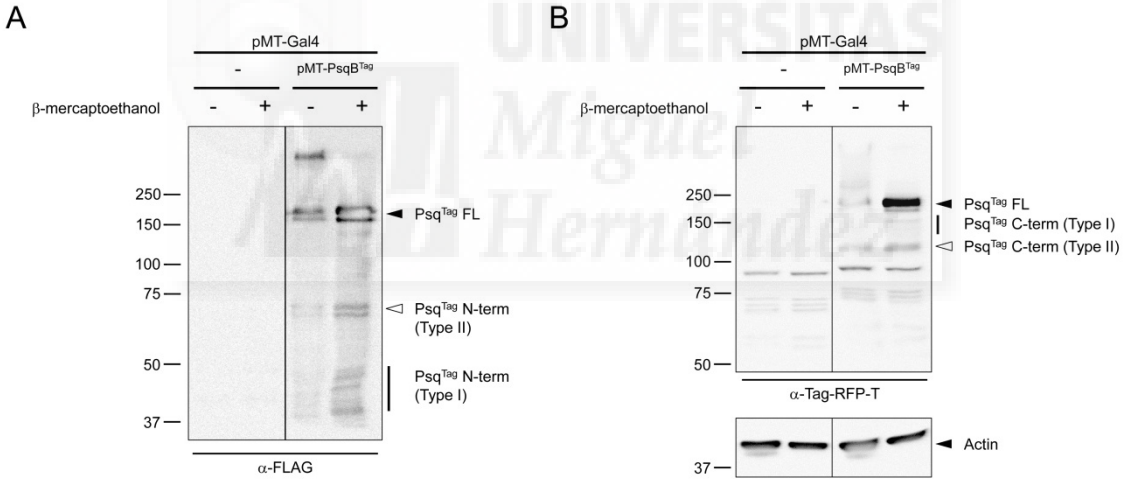


Figure 41. β-mercaptoethanol differentially affects PipsqueakB Type I and Type II processing. As indicated, S2 cells were transiently transfected with the construct *pMT-psqB^{Tag}* or the *pMT* empty vector. The protein lysates were analysed in WB probed with antibodies against FLAG (**A**), Tag-RFP-T or Actin (**B**).

PsqB processing was a very interesting finding for us because, attending to the fragment's sizes observed in the WB gels, it could be a way to separate the BTB domain from the DNA-binding domain, at least in the case of Type II processing. This result could be suggesting a new mechanism to modulate the activity of this transcription factor. Indeed, physically separating the ability to interact with other proteins and the ability to bind DNA could have a severe impact in PsqB function. Also,

the size of the processed fragments is similar to the predicted weigh of some of the products of the shorter BTB-less Psq isoforms, to date involved in zygotic functions including eye development (see *Introduction* and Weber *et al.*, 1995). With these ideas in mind, we hypothesized that PsqB processing could serve as a mechanism to modulate the balance between distinct PsqB "forms", which may have different roles, therefore controlling PsqB function in a broader sense. For example, could PsqB processing be a parallel mechanism to increase the amount of "short Psq"? For all these reasons we decided to tackle the issue of PsqB processing as a secondary research project for this Thesis. Understanding how Psq sumoylation and processing modulate its activity will deepen our knowledge about its functions both in normal and oncogenic conditions, and it might serve as a model on how other BTB proteins with similar structure function in the tumorigenic context.

2. Localization of PipsqueakB processing sites

The first step in our study of Psq proteolytic processing was the localization of Psq processing sites. The products of Psq processing observed in previous experiments suggested that the processing sites should be located between the BTB and the DNA-binding domains of PsqB protein, towards the amino-terminal part. In order to start mapping the PsqB proteolytic processing sites, we designed several overlapping deletions spanning that part of PsqB protein (**Figure 42A**): *3xFLAG::psqBΔ2::Tag-RFP-T* (from now on *psqBΔ2^{Tag}*), *3xFLAG::psqBΔ3::Tag-RFP-T* (*psqBΔ3^{Tag}*), *3xFLAG::psqBΔ4::Tag-RFP-T* (*psqBΔ4^{Tag}*), *3xFLAG::psqBΔ5::Tag-RFP-T* (*psqBΔ5^{Tag}*). To generate these deletions we followed the protocol developed by Stoynova *et al.*, (2004), in which they describe how to generate large deletions using primers with modified nucleotides and an inverse PCR (see *Materials and Methods*).

We then transfected S2 cells with these constructs under the control of the inducible *pMT* promoter and performed WB experiments, which allowed us to determine if any of the deletions contained the processing sites and thus abolished Psq processing. In **Figure 42B** we observe WB gels of total protein samples of S2 cells transfected with the different *psqB* deletion mutants. We used total lysates instead of separating the soluble and the insoluble fraction to detect both Type I and Type II PsqB

processing. Also, due to technical problems, we could not clone deletion 1, which is why it is not present in the gel. To avoid confusions with the names of the deletions we decided to preserve the original numbers. All deletion constructs produced a full-length PsqB protein at similar levels as the control $PsqB^{Tag}$ protein (**Figure 42B-C**), although they migrated slightly faster in the gel due to the difference in size caused by the deletions. In these samples we observed that constructs carrying deletion 2 and 3 still showed Type II processing (70 kDa + 110 kDa), but the pattern of the bands was different compared to what we observed for the wild-type fusion protein $PsqB^{Tag}$. Both deletions reduce the size of the amino-terminal processed fragment of PsqB (**Figure 42B**), while the carboxy-terminal part remains unchanged (**Figure 42C**). This indicates that Type II processing site is located posterior to deletions 2 and 3, towards the carboxy-terminal end. On the other hand, deletions 4 and 5 completely eliminate Type II processing, suggesting that Type II processing site is located somewhere in the sequence spanned by deletions 4 and 5, between Asn415 and Ala636 (**Figure 42A**). One possibility is that the processing site is located in the overlapping region between deletions 4 and 5, explaining why both deletions have the same effect over Psq processing. Another possibility is that the proteolytic site is located in one of the deletions, but the surrounding region of the protein may be vital to form the tertiary protein structure needed for proper interaction with the still unknown molecule responsible for Psq Type II processing.

In the case of Type I processing (~30-50 kDa + ~120-150 kDa), analysis of total S2 cell lysates by WB suggests that it is blocked by deletion 2 (**Figure 43A-B**). Regarding deletion 3, we have WB results indicating that it may behave as deletion 2, but the images are not as conclusive. Posterior results show that both $psqB\Delta2^{Tag}$ and $psqB\Delta3^{Tag}$ are not able to reach the nucleus. This probably explains why we do not see Psq Type I processing in these mutants, as this processing occurs in the insoluble fraction of the cell, most likely in the nucleus. Therefore, these mutant constructs do not allow us to unambiguously conclude that the Type I processing site is located in the region spanned by deletions 2 and 3.

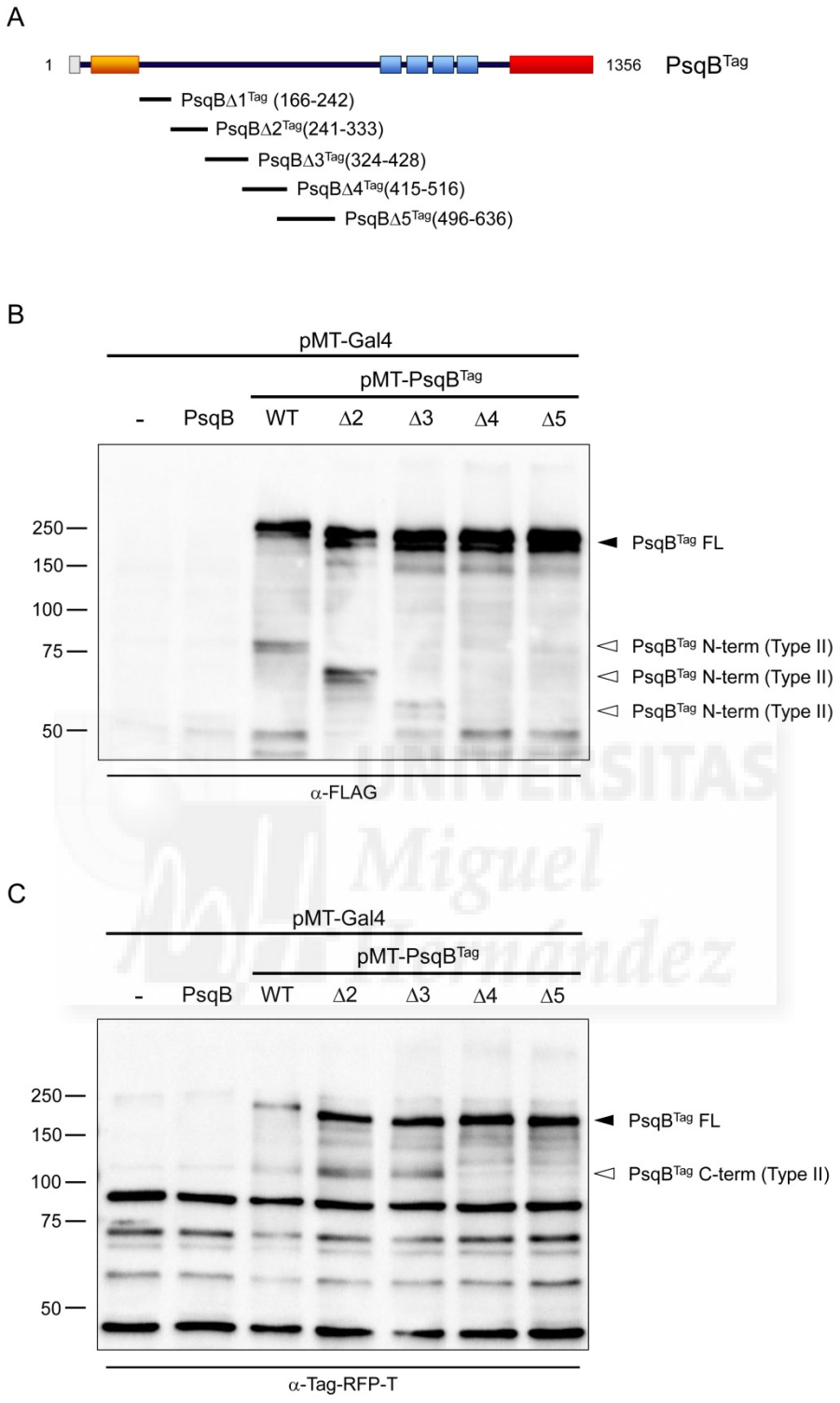


Figure 42. Deletions located between Asn415 and Ala636 block PipsqueakB Type II processing. WB of lysates from S2 cells transfected with *pMT* empty vector, endogenous *psqB*, *psqB*^{Tag} or the mutant constructs *psqBΔ2*^{Tag}, *psqBΔ3*^{Tag}, *psqBΔ4*^{Tag} or *psqBΔ5*^{Tag}. Protein lysates were analysed in WB probed with antibodies against FLAG (**B**) and Tag-RFP-T (**C**).

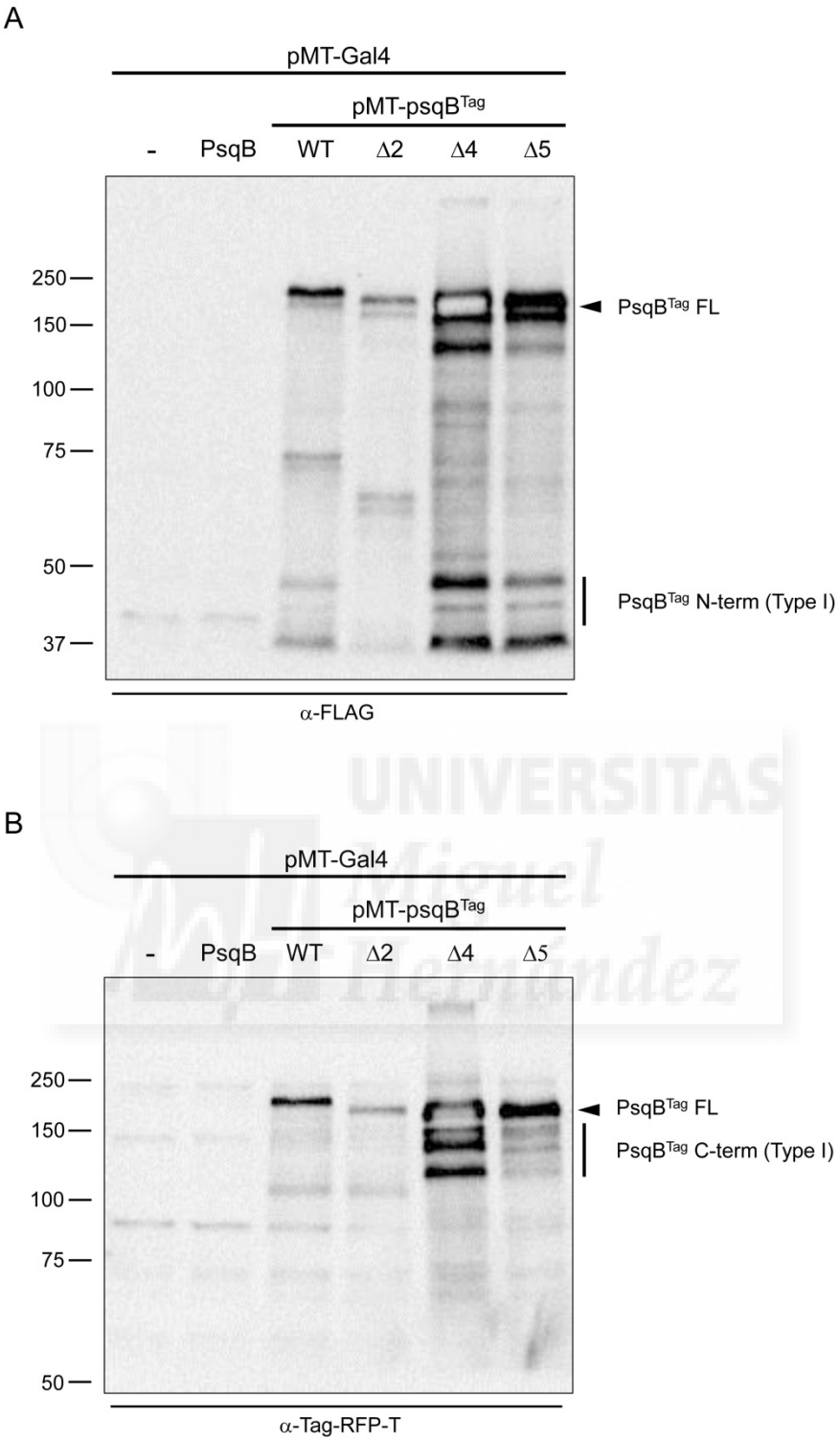


Figure 43. Type I processing is absent in PsqBΔ2^{Tag}. WB of lysates from S2 cells transfected with *pMT* empty vector, endogenous *psqB*, *psqB^{Tag}* or the mutant constructs *psqBΔ2^{Tag}*, *psqBΔ3^{Tag}*, *psqBΔ4^{Tag}* or *psqBΔ5^{Tag}*. Protein lysates were analysed in WB probed with antibodies against FLAG (**B**) and Tag-RFP-T (**C**).

Up to this point we have demonstrated that PsqB is processed in two specific independent sites and these two events are associated with different compartments of the cell (soluble vs. insoluble fraction). The fact that Type I and Type II processing sites are located in different protein regions further supports the idea of two different mechanisms for PsqB processing. With the different PsqB overlapping deletions we have been able to narrow the search for these processing sites to smaller areas of Psq, but we have not found the specific processing sites yet. For this reason we decided to perform new smaller overlapping deletions that span the areas containing the processing sites. In parallel to this, we decided to look for conserved proteolytic sites for known proteases in these regions. Additionally, we looked again at our Y2H results to see if there was some protease interacting with PsqB. Unfortunately, none of these approaches has given positive results at the moment of writing this Thesis.

3. Effect of proteolytic processing on PipsqueakB transcriptional activity

As I have mentioned in previous sections of this Thesis, microarray and qPCR experiments performed in our lab demonstrate that the gene *CG9925* acts as a read-out of PsqB activity; *CG9925* mRNA levels increase in response to the expression of BTB-containing Psq isoforms.

To further validate *CG9925* as a good read-out of PsqB activity, we compared its relative mRNA levels in samples of Kc167 cells transfected with constructs carrying endogenous *psqB*, *psqB^{Tag}*, or *3xFLAG::psq⊿6::Tag-RFP-T* (from now on *psqB⊿6^{Tag}*), a mutation that completely removes the Psq DNA-binding domains (**Figure 44A**). As a result we can see that both PsqB and PsqB^{Tag} derepress *CG9925 in vitro* and that Psq DNA-binding domains are critical for its transcriptional activity (**Figure 44B**).

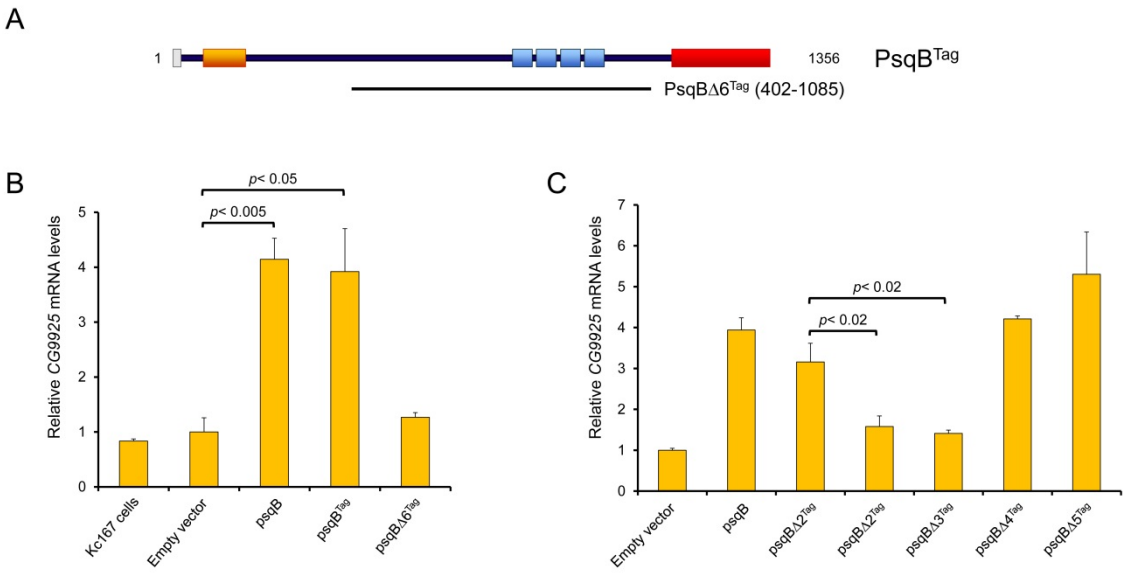


Figure 44. Pipsqueak deletions 2 and 3 abolish *CG9925* Pipsqueak-dependent response. (**A**) Schema representing PsqB^{Tag} protein and the relative position of deletion 6. (**B-C**) Quantification of the relative change in mRNA levels of *CG9925* by qPCR in Kc167 cells transfected with full length *psqB* (with or without tags) or the constructs carrying different deletions. In (**B**) n = 4 independent biological replicates; Empty vector vs *psqB*, p = 0.004; Empty vector vs *psqB^{Tag}*, p = 0.03. In (**C**) n = 3 independent biological replicates; *psqB^{Tag}* vs *psqBΔ2^{Tag}*, p = 0.018; *psqB^{Tag}* vs *psqBΔ3^{Tag}*, p = 0.016). The p-value was calculated using two-tailed Student's *t*-test. The data are presented as the mean + SEM.

Next we characterized the effect of PsqB processing over its transcriptional activity. For that we transfected the constructs carrying the different Psq deletions in Kc167 cells and measured the relative mRNA levels of *CG9925* by qPCR (**Figure 44C**). We observed that the ability of PsqB^{Tag} to derepress *CG9925* transcription is almost completely lost in PsqBΔ2^{Tag} and PsqBΔ3^{Tag}. Contrary to this, PsqBΔ4^{Tag} and PsqBΔ5^{Tag} were even slightly more effective in derepressing *CG9925* expression than the wild-type PsqB^{Tag} version, albeit the difference was not statistically significant. These results suggest that Type II processing (**Figure 40**) does not have a significant effect over PsqB^{Tag} transcriptional activity in this particular case. Regarding Type I, deletions 2 and 3 may not affect this processing but, as we will see in the following sections, they abolish PsqB^{Tag} nuclear localization, indicating that Psq presence in the nucleus is required for *CG9925* expression.

4. Role of PipsqueakB processing *in vivo*

We next studied how Psq processing affects its tumorigenic activity using our *in vivo* Notch-driven tumorigenesis paradigm. For that we generated transgenic flies carrying the different Psq deletions (*psqBΔ2^{Tag}*, *psqBΔ3^{Tag}*, *psqBΔ4^{Tag}* and *psqBΔ5^{Tag}*). In the next sections I will present the results derived from the analysis of the expression of these transgenic fly lines.

4.1 Type I processing requirement in PipsqueakB tumorigenic capacity *in vivo*

As a first step for the genetic assays, we started by crossing *psqBΔ2^{Tag}* and *psqBΔ3^{Tag}* transgenic lines with the *ey-Gal4* stock, which allowed us to check the expression of the protein and the possible effect of overexpressing these PsqB mutants in the growth and development of *Drosophila* eye. For each of these constructs, *psqBΔ2^{Tag}* and *psqBΔ3^{Tag}*, 9 different transgenic fly lines were obtained and analysed. All crosses were viable and produced adult flies with eyes identical to the control (**Figure 45A, C** and **D**). Strikingly, these results differ from our wild-type construct. As I already mentioned in the first part of the *Results* section, *psqB^{Tag}* overexpression decreases the viability of the flies (~50% of the progeny overexpressing the construct dies), and among the flies that survive, around 12.5% have smaller eyes respect to the control (**Figure 45B**). Both the improvement in viability and the lack of phenotype in the eye already suggests that PsqB deletions 2 and 3 somehow negatively affect PsqB regular activity. Additional data from *psqB* overexpression in salivary glands supports this hypothesis. In addition to driving expression in the *Drosophila* eye, the *eyeless* promoter also drives expression in the salivary gland of L3 larvae. *psqB^{Tag}* overexpression using the *eyeless* promoter causes a reduction in the size of salivary glands (**Figure 45E-F**). However, overexpression of *psqBΔ2^{Tag}* and *psqBΔ3^{Tag}* does not affect salivary gland development (**Figure 45G-H**), suggesting that both deletions reduce *psqB* function.

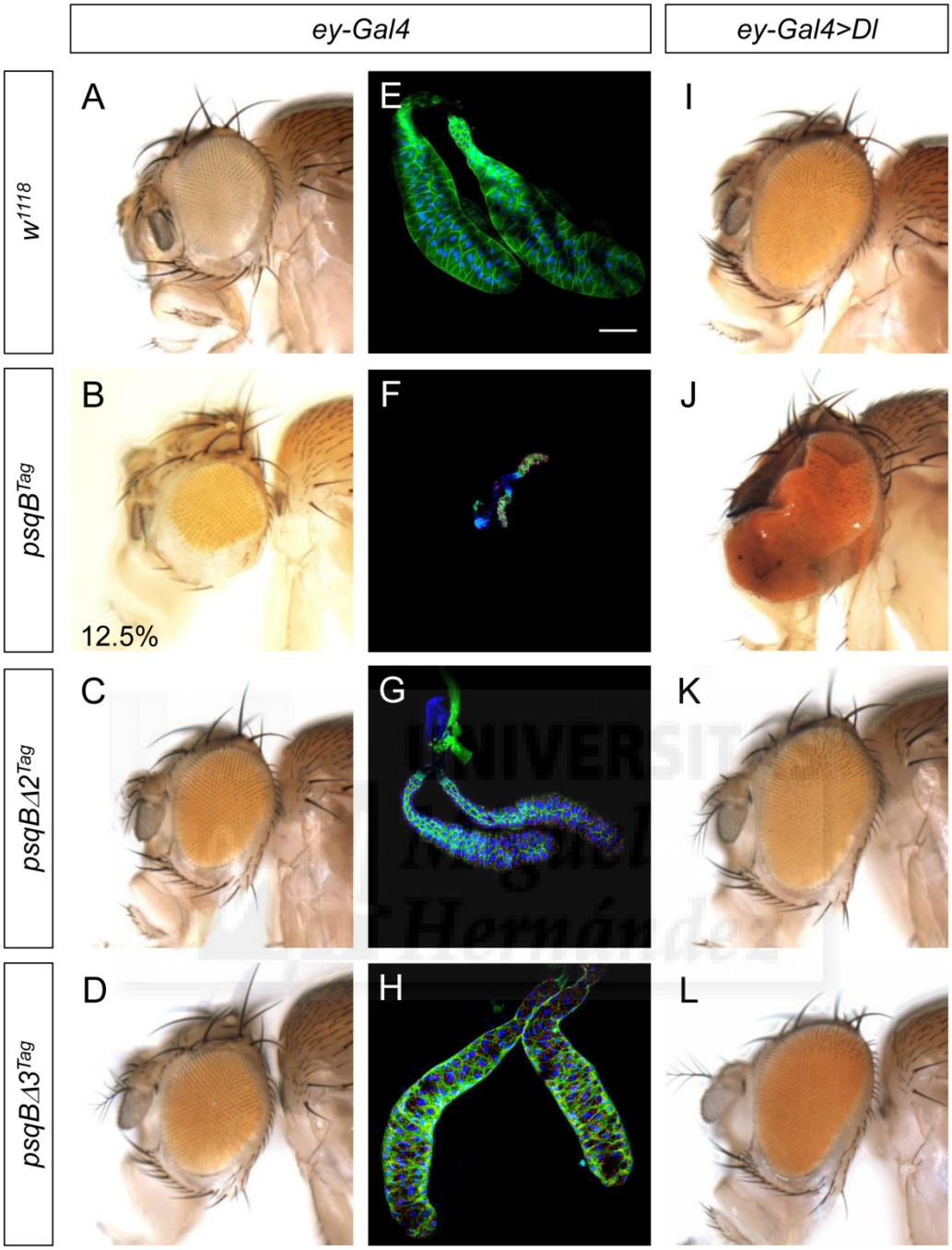


Figure 45. ***In vivo*** **effects of PipsqueakB Type I processing elimination.** Representative examples showing the adult eye phenotype obtained when overexpressing different wild-type or mutant *psqB*^{Tag} constructs under the *ey-Gal4* promoter (**A-D**) or in the Notch overexpression background *ey-Gal4>Dl* (**I-L**). All photos were taken from female flies. The crosses were maintained at 25ºC in the case of *ey-Gal4* or at 27ºC in the case of *ey-Gal4>Dl*. (**E-H**) Confocal images of L3 larvae salivary glands immunostained with DAPI to mark the nucleus (blue) and discs large to mark the cell membrane (green). Tag-RFP-T is shown in red. Scale bar in (**E**) represents 100 μm.

The next step was to combine the overexpression of our PsqB mutants with a Notch overexpression background via its ligand *Dl* (*ey-Gal4>Dl*), to find out the effect of these deletions over PsqB oncogenic capacity. *psqB^{Tag}* overexpression causes 21.54% of the flies to develop eye tumours (**Figure 45J**), which is a lower percentage compared to the stock *ey-Gal4>Dl>eyeful*, used until now as control in the tumorigenic assays (**Figure 20F** and **27B**). This difference in oncogenic capacity can be due to the place of insertion of the transgene and/or to the lack of *lola* overexpression. In any case, these experiments show that PsqB^{Tag} ectopic expression alone is sufficient to promote tumour growth in an *ey-Gal4>Dl* background. When we cross *psqBΔ2^{Tag}* and *psqBΔ3^{Tag}* with *ey-Gal4>Dl*, independently of which of the 9 transgenic lines we used, we observe that the tumorigenic activity is completely lost and all eyes show an overgrowth phenotype indistinguishable from the control *ey-Gal4>Dl* condition (**Figure 45I**, **45K** and **45L**). This result supports the hypothesis that deletions 2 and 3 negatively affect PsqB tumorigenic activity.

4.2 Type II processing requirement in PipsqueakB tumorigenic capacity *in vivo*

Injection of the constructs *psqBΔ4^{Tag}* or *psqBΔ5^{Tag}* generated 8 and 5 different transgenic fly lines, respectively. As in the previous section, we first crossed these transgenic fly lines with the *ey-Gal4* stock. As occurs in the case of wild-type *psqB^{Tag}* overexpression, *psqBΔ4^{Tag}* overexpression promotes a reduction in the viability of the progeny of these crosses (the percentage varies between the different transgenic lines, being some of them even more lethal than PsqB^{Tag} overexpression). Moreover, and contrary to what happened in the case of *psqBΔ2^{Tag}* and *psqBΔ3^{Tag}*, a percentage of the flies that reach adulthood presents smaller eyes respect to the control (this percentage varies depending on the transgenic line) (**Figure 46A-C**). Finally, in flies overexpressing *psqBΔ4^{Tag}* we also observe a reduction in the size of the salivary glands in L3 larvae, as occurs when *psqB^{Tag}* is overexpressed (**Figure 46E-G**). In the case of *psqBΔ5^{Tag}* transgenic fly lines, the viability is similar to what we observe in *psqB^{Tag}* and *psqBΔ4^{Tag}* overexpression, and the effect over the size of the eyes in the adult flies and the salivary glands is also the same as in *psqB^{Tag}* and *psqBΔ4^{Tag}* (**Figure 46D** and **46H**): the L3 larvae showed smaller salivary glands and a percentage of the progeny, that varied between the different transgenic lines, had smaller eyes. The results derived from these crosses indicate that eliminating the sequence spanning

psqBΔ4^{Tag} or *psqBΔ5*^{Tag} does not hinder Psq oncogenic function, suggesting that Psq Type II processing is not relevant in this context.

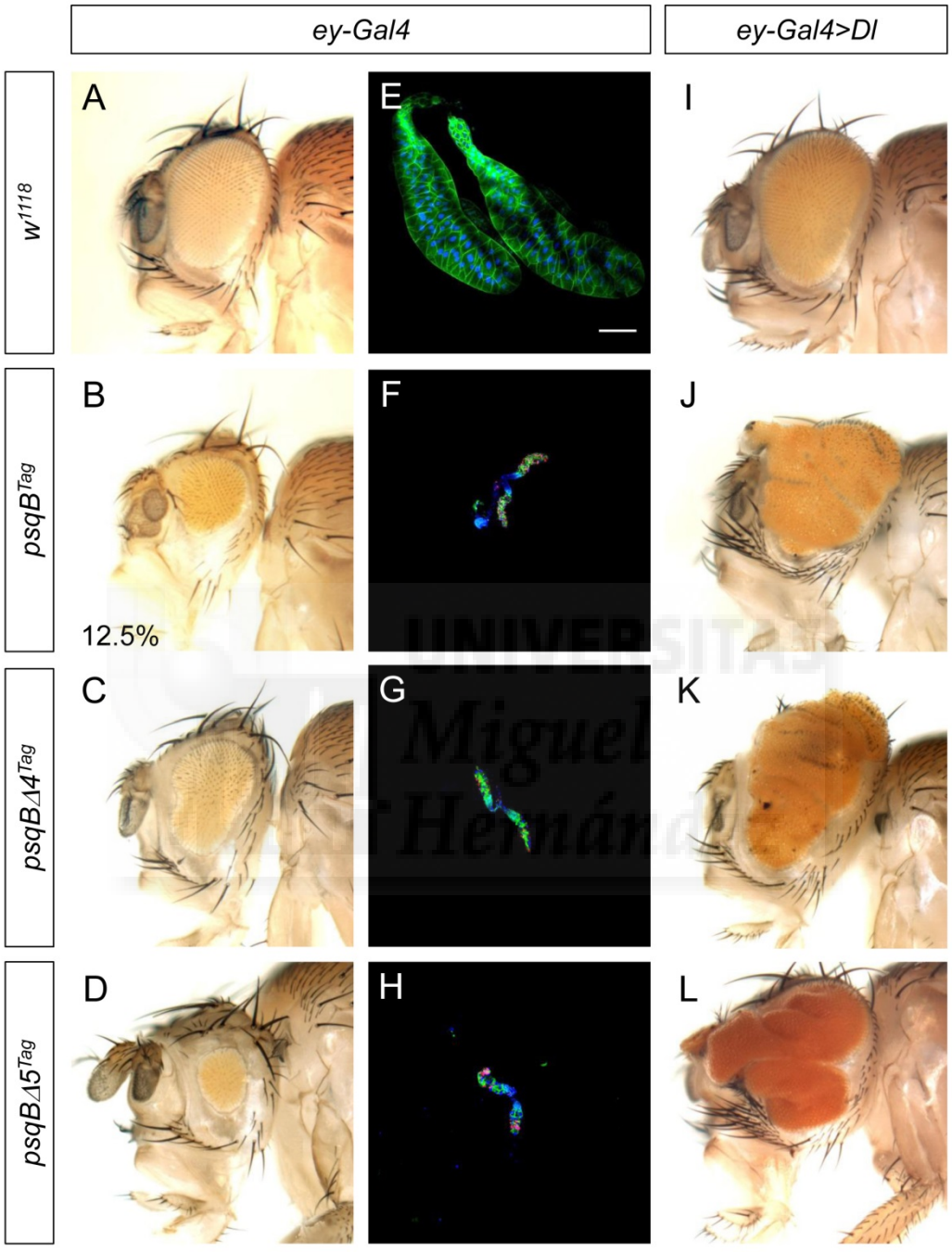


Figure 46. ***In vivo*** **effects of PipsqueakB Type II processing elimination.** Representative examples showing the adult eye phenotype obtained when overexpressing different wild-type or mutant *psqB*^{Tag} constructs under the *ey-Gal4* promoter (**A-D**) or in the Notch overexpression background *ey-Gal4>Dl* (**I-L**). All photos were taken from female flies. The crosses were maintained at 25ºC in the case of *ey-Gal4* or at 27ºC in the case of *ey-Gal4>Dl*. (**E-H**) Confocal images of L3 larvae salivary glands immunostained with DAPI to mark the nucleus (blue) and discs large to mark the cell membrane (green). Tag-RFP-T is shown in red. Scale bar in (**E**) represents 100 μm.

Next we assayed how PsqB^{Tag} deletions 4 and 5 affect its oncogenic capacity using the stock *ey-Gal4>Dl*. Six out of eight *psqBΔ4^{Tag}* transgenic lines produce tumour growth in 10-30% of the flies, depending on the line, which is comparable to the control *psqB^{Tag}* overexpression (21.54% of the *ey-Gal4>Dl>psqB^{Tag}* flies show tumoural growth), (**Figure 46J** and **46K**). Among the other two lines, one of them showed an identical phenotype to our Notch overexpression control, *ey-Gal>Dl* (100% overgrown eyes); while the other line was 100% lethal. These phenotypical differences between the transgenic lines are probably due to the place of insertion of these constructs. In agreement with these results, overexpression of *psqBΔ5^{Tag}* in the context of Notch overexpression behaves in a similar way. Four out of five transgenic lines retain the capacity to promote tumour growth (**Figure 46L**), although the penetrance of the tumorigenic phenotype is even lower than in the case of *psqΔ4^{Tag}* (below 10%). The remaining line showed no phenotype with any of the drivers tested, so we concluded that the place of insertion did not allow the transgene expression and we decided not to use it in further experiments. These results suggest that deletions 4 and 5 do not have a strong effect over PsqB tumorigenic capacity *in vivo*, supporting the hypothesis that PsqB Type II processing has not a critical role in this particular context.

4.3 Effect of PipsqueakB deletions over its subcellular distribution

In addition to looking at the macroscopic morphology of the salivary glands, we used our immunohistochemistry preparations to study the subcellular localization of Psq and see if the different mutations induced changes in the distribution of the protein. Unfortunately, due to technical problems we could not follow de distribution of the 3xFLAG::N-terminal processed fragment of PsqB, which would have been very informative. In **Figure 47B** and **H** we can see that PsqB^{Tag} colocalizes with DAPI in the nucleus, and also accumulates in a few cytoplasmic bodies. The pattern of distribution is similar in *psqBΔ4^{Tag}* and *psqBΔ5^{Tag}* mutants (**Figure 47E**, **F**, **K** and **L**), which is consistent with the results from the previous figure, showing that these mutants behave as wild-type *psqB^{Tag}*. In the case of *psqBΔ2^{Tag}* and *psqBΔ3^{Tag}* (**Figure 47C**, **D**, **I** and **J**), these mutant versions of PsqB are not present in the nucleus and remain in the cytoplasm inside the cytoplasmic bodies, which have increased in number respect to *psqB^{Tag}*, *psqBΔ4^{Tag}* and *psqBΔ5^{Tag}* preparations.

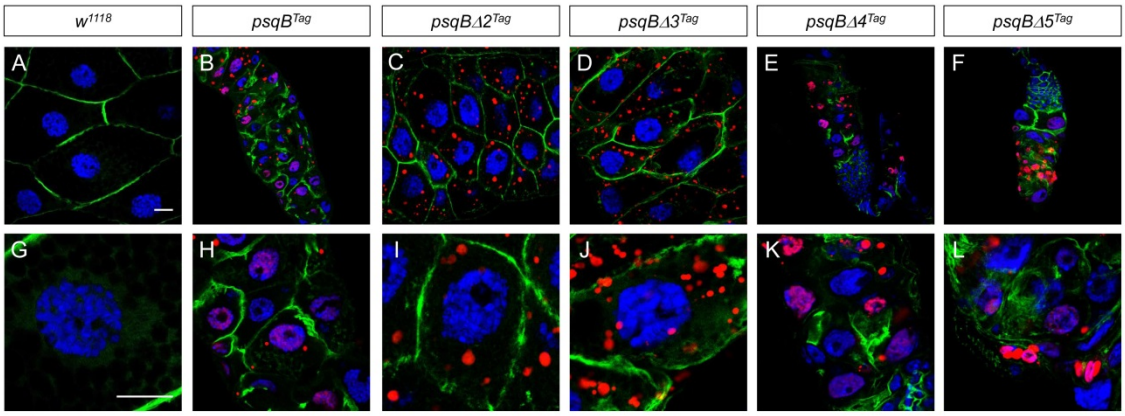


Figure 47. Pipsqueak protein subcellular distribution. Confocal images of L3 larvae salivary glands from animals overexpressing the construct *psqB*^{Tag}, *psqBΔ2*^{Tag}, *psqBΔ3*^{Tag}, *psqBΔ4*^{Tag} or *psqBΔ5*^{Tag} under the *ey-Gal4* promoter. The tissue was immunostained with DAPI to mark the nucleus (blue) and discs large to mark the cell membrane (green). Tag-RFP-T is shown in red. Scale bar in (**A**) and (**G**) represents 25 μm.

If we take into consideration that Psq Type I processing only occurs in the insoluble fraction of the cell, which includes the nucleus, we cannot consider that *psqBΔ2*^{Tag} and *psqBΔ3*^{Tag} mutants are useful to address the effect of blocking Psq Type I processing. The localization of these proteins exclusively in the cytoplasmic bodies would explain why these mutants fail to activate *CG9925*, are not able to produce tumour growth in combination with *Dl* overexpression and do not affect the size of salivary glands upon overexpression. The mislocalization of the protein could be due to deletions 2 and 3 totally or partially deleting a sequence necessary for PsqB nuclear localization, and thus preventing PsqB from executing its functions.

5. Proteasome activity is responsible for PipsqueakB Type I processing

The Type I processing does not seem to always generate fragments of the exact same size, such as Type II processing does. Rather there seems to be a ladder of amino-terminal fragments, and possibly a ladder of carboxy-terminal fragments, which is harder to detect in our gels. Which enzyme could perform this variable cleavage? We considered the possibility that the enzyme might not have a specific cleavage site, but instead it would cleave the protein in a less specific manner towards the very amino-

terminus. A candidate enzymatic activity that could explain these results would be the proteasome. Apart from being involved in the ubiquitin-dependent protein degradation pathway, the proteasome is also involved in the activation of certain proteins by limited proteolysis, such as the mammalian NF-κB, the *Drosophila* Ci and the yeast Def1 (Palombella *et al.*, 1994; Tian *et al.*, 2005; Wilson *et al.*, 2013). To test if PsqB was subject to proteasome-dependent partial proteolysis, we transiently transfected S2 cells with *psqB*^{Tag} or empty vector in the presence of different proteasome inhibitors: lactacystin, which is a permanent proteasome inhibitor that has previously been shown to work in S2 cells (Matsubayashi *et al.*, 2004), the widely used and reversible MG132, and Z-LLF-CHO, another reversible inhibitor without bibliographic evidence of its use in *Drosophila* cell lines up until now. We treated the cells with the different inhibitors for 6 hours and tested the effect over PsqB processing by WB. For this experiment we used total cell lysates instead of separating the soluble and insoluble fractions. Using an anti-ubiquitin antibody we detected the accumulation of ubiquitinated proteins as a smear at high molecular weights in the lines of cell lysates treated with Z-LLF-CHO and MG132 (**Figure 48**, first panel, lanes 6, 8, 10 and 12). In the case of lactacystin the smear is not as clear (**Figure 48**, first panel, lanes 2 and 4). This control demonstrates that two of the three inhibitors are working properly in these conditions, and that we are inhibiting degradation of ubiquitinated proteins by proteasome activity. Lack of inhibition with lactacystin might be due to a low inhibitor concentration. Regarding PsqB processing, anti-FLAG antibody detection showed that proteasome inhibition does not affect Type II PsqB processing, as we still detect the 70 kDa 3xFLAG::PsqB amino-terminal band in the conditions treated with the inhibitors (**Figure 48**, lanes 4, 8 and 12). On the contrary, treatment with Z-LLF-CHO and MG132 decreased the amount of the ~30-50 kDa 3xFLAG::PsqB amino-terminal bands that are typical of Type I processing. In agreement with this, anti-Tag-RFP-T showed that proteasome inhibition with Z-LLF-CHO and MG132 blocked PsqB processing Type I but not Type II, as we still detect the 110 kDa C-terminal PsqB::Tag-RFP-T band, but not the ~120-150 kDa band typical of Type I processing (**Figure 48**, lanes 8 and 12). Therefore, these data show that proteasome activity underlies PsqB Type I processing.

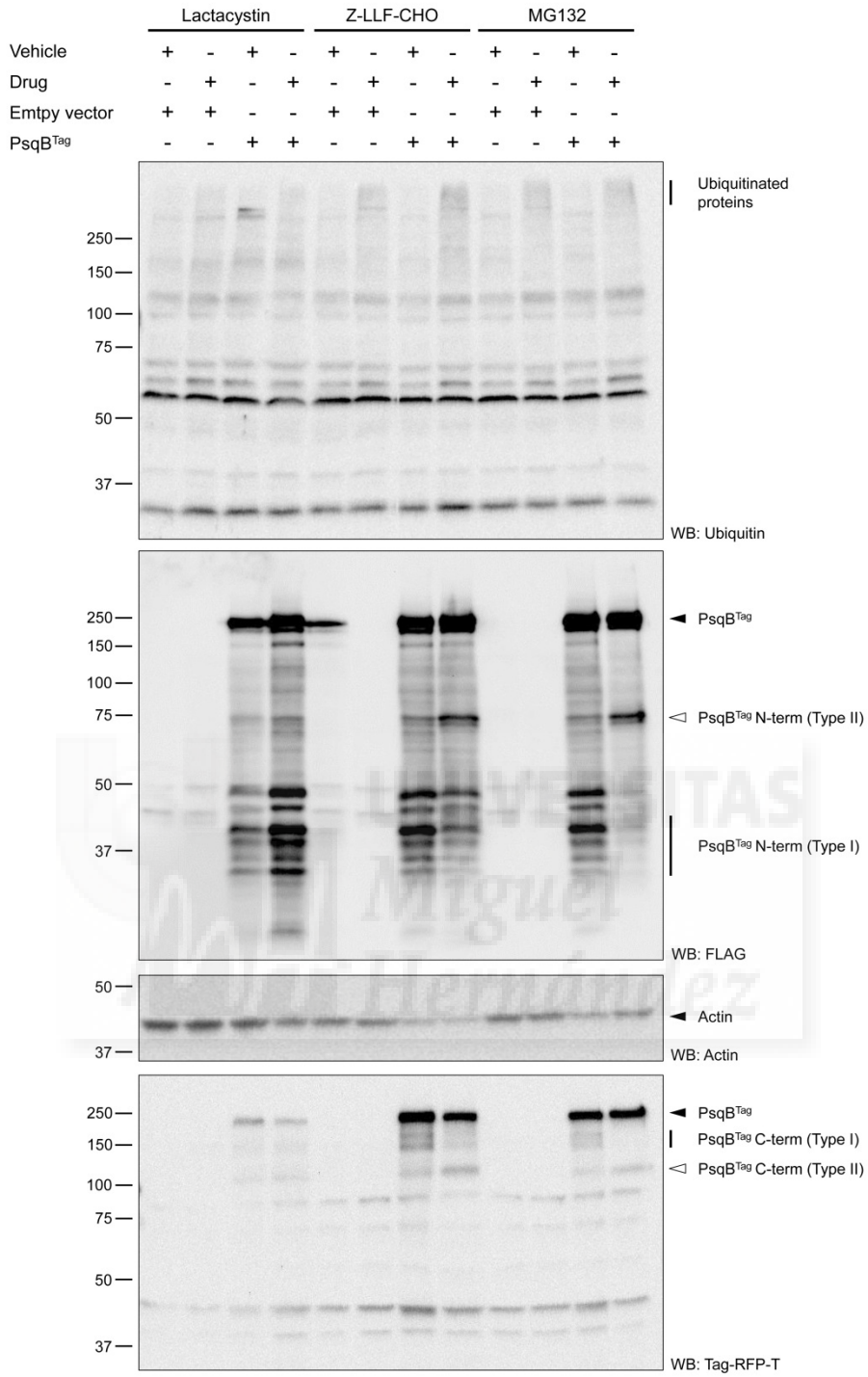
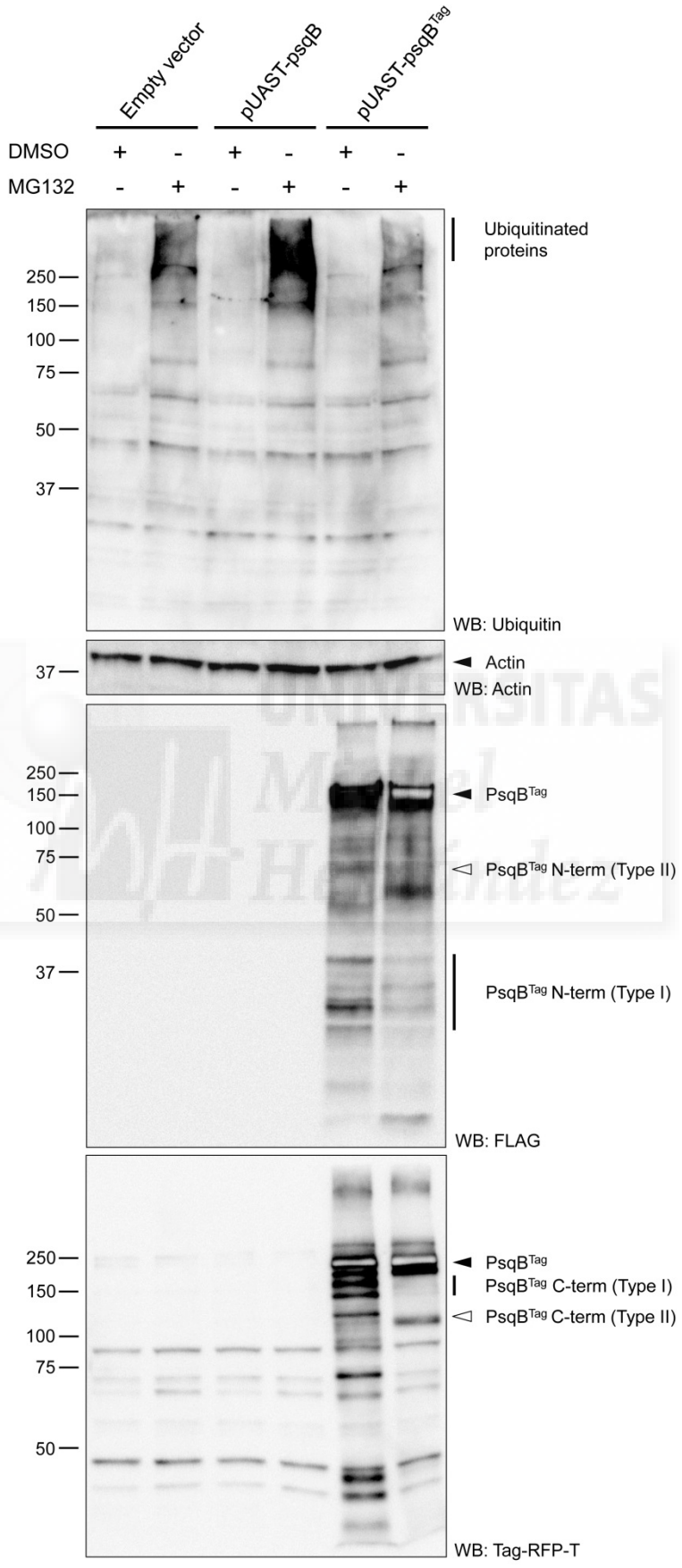


Figure 48. Proteasome inhibition blocks PipsqueakB Type I processing. WB of lysates from S2 cells transfected with *pMT* empty vector or *psqB^{Tag}*. Lactacystin (40 μM), Z-LLF-CHO (40 μM) and MG132 (50 μM) were added when indicated to the cell medium and incubated for 6 hours prior to the cell lysis. The vehicles for the drugs were water (for lactacystin), or DMSO (for Z-LFF-CHO and MG132). The samples were analysed in WB and probed with antibodies against ubiquitin (top panel), FLAG (second panel), actin (third panel) and Tag-RFP-T (bottom panel).

Next we wondered if by blocking Type I processing of wild-type PsqB via proteasome inhibition we would also affect its ability to derepress *CG9925*. To test this hypothesis we transfected S2 cells with *psqB*^{Tag} or empty vector and treated them with MG132. The lysates were then divided to analyse them by WB and qPCR to monitor PsqB^{Tag} processing and *CG9925* expression, respectively. The experiment shows again that proteasome inhibition prevents PsqB Type I processing (**Figure 49A**). In addition to this, measuring *CG9925* expression levels shows that MG132 addition abolishes the *CG9925* response to PsqB (**Figure 49B**), mimicking the effect of *psqB* deletion mutants 2 and 3 that prevented PsqB translocation to the nucleus and hence, its Type I processing. Interestingly, there is also a significant reduction of *CG9925* levels following MG132 treatment of control cells transfected with empty vector. This could indicate that endogenous PsqB nuclear localization, which should likewise be affected by the MG132 treatment, is also necessary for maintaining a basal level of *CG9925* transcriptional derepression in S2 cells, but this requires further tests.

To sum up, these experiments show that the proteasome is responsible for PsqB Type I processing, and chemical inhibition of its activity results in a decrease in *CG9925* expression levels. These results indicate that, not only Psq nuclear localization, but proteasome mediated Type I processing are both needed for PsqB transcriptional activity. In the case of PsqB processing Type II, our experiments clearly demonstrate that the proteasome is not directly responsible for it. Even so, we cannot rule out the possibility that proteasome inhibition has also an effect over this processing because Z-LLF-CHO and MG132 addition increase the amount of Psq Type II processed fragments observed in the gels (**Figure 48**, panels 2 and 4, lanes 7 and 11 compared to lanes 8 and 12). Still, this effect could merely be a reflection of an unbalance on the different forms of PsqB processing; lack of Type I processing may increase the amount of Type II fragments as more PsqB protein would be available for this process.

A


Empty vector
pUAST-psqB
pUAST-psqB^Tag
DMSO + - + - + -
MG132 - + - + - +
Ubiquitinated proteins
250
150
100
75
50
37
WB: Ubiquitin
37
Actin
WB: Actin
250
150
100
75
50
37
PsqB^Tag
PsqB^Tag N-term (Type II)
PsqB^Tag N-term (Type I)
WB: FLAG
250
150
100
75
50
PsqB^Tag
PsqB^Tag C-term (Type I)
PsqB^Tag C-term (Type II)
WB: Tag-RFP-T

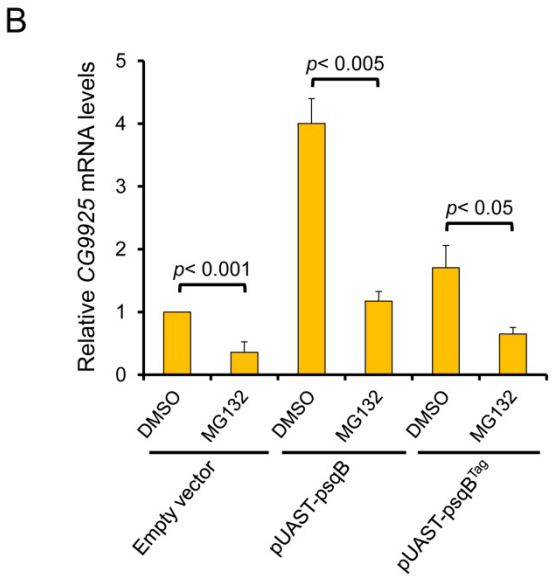


Figure 49. Proteasome inhibition abolishes PipsqueakB-dependent CG9925 derepression. (**A**) WB analysis of lysates from S2 cells transfected with *pUAST* empty vector, *pUAST-psqB* or *pUAST-psqB^{Tag}.* MG132 (50 μM) was added to the cell medium when indicated and incubated for 6 hours prior to the cell lysis. DMSO was used as control. The samples were analysed in WB and probed with antibodies against ubiquitin (top panel), actin (second panel), FLAG (third panel) and Tag-RFP-T (bottom panel). (**B**) Quantification of the relative change in mRNA levels of *CG9925* by qPCR in S2 cells transfected with empty vector or full length *psq* (with or without tags) treated with MG132 or DMSO as control. n = 3 independent biological replicates. Empty vector DMSO vs empty vector MG132, p = 0.00006; *pUAST-psqB* DMSO vs *pUAST-psqB* MG132, p = 0.003; *pUAST-psqB^{Tag}* DMSO vs *pUAST-psqB^{Tag}* MG132, p = 0.046. The p-value was calculated using two-tailed Student's *t*-test. The data are presented as the mean + SEM.

The results presented in this second part of the *Results* section suggest that both PsqB presence in the nucleus and a specific proteolytic processing are crucial for PsqB transcription factor function. We have demonstrated that this specific processing, Type I, occurs only in the insoluble fraction, most likely in the nucleus where PsqB exerts its transcriptional activity. Our experiments show that Type I processing is mediated by the proteasome and is required for PsqB transcriptional activity. Regarding the effect of processing over Psq tumorigenic activity *in vivo*, our experiments show that Type II processing is not related to this process. Unfortunately we were not able to assess the impact of Type I processing, as *psqBΔ2^{Tag}* and *psqBΔ3^{Tag}* mutants are unable to reach the nucleus. Further experiments will be needed to clarify Type I processing contribution to PsqB oncogenic activity. How this proteolytic event could affect transcription factor activity will be further considered in the *Discussion* section, but these data opens the intriguing question of whether this is a

general mechanism for BTB transcription factor function both in *Drosophila* and in other species.

UNIVERSITAS Miguel Hernández

UNIVERSITAS Miguel Hernández

IV. DISCUSSION

UNIVERSITAS Miguel Hernández

UNIVERSITAS
Miguel
Hernández

BTB transcription factors are encoded by a large and diverse family of genes and they have been shown to have crucial roles in many human diseases including cancer. With the aim of elucidating the molecular mechanisms underlying the regulation of the activity of these factors, we have used the *Drosophila* BTB transcription factor Pipsqueak as a paradigm to study how sumoylation and partial proteolytic processing can modulate BTB transcription factor activity during normal development and oncogenesis. Throughout this section, we will discuss the results comparing and contrasting them with previous findings in the fields of BTB transcription factor activity regulation, sumoylation and proteasome-dependent proteolytic processing, as well as proposing future experiments.

1. SUMO-mediated regulation of the BTB transcription factor Pipsqueak: a whole new world of possibilities

Our detailed examination of the molecular mechanisms of oncogenesis mediated by Psq led us to the discovery of Psq as a target of sumoylation, a post-translational modification strongly linked with cancer in the last years (Bettermann *et al.*, 2012; Kim and Baek, 2006, 2009; Zhao, 2007). The whole set of results presented in the first part of this Thesis show that Psq sumoylation has a crucial role in fine tuning Pipsqueak activity. Psq has several putative sumoylation sites, and our co-immunoprecipitation experiments in S2 cells and *in vivo* genetic assays have validated that Lysine 633 is the only functional sumoylation site in Psq. We have used the gene *CG9925* as a read-out of Psq transcriptional activity, as *CG9925* increases its expression levels in response to Psq overexpression. qPCR experiments show that blocking Psq sumoylation increases even more *CG9925* expression, suggesting that Psq sumoylation somehow attenuates Psq transcriptional activity. Blocking Psq sumoylation specifically in the eye of *Drosophila* produces strong defects during development leading to 100% lethality of the flies during metamorphosis. This is probably due to the loss of eye tissue observed in the eye-antennal imaginal discs in L3 stage larvae. Immunostaining of these imaginal discs with PH3 revealed that there is not a reduction in cell proliferation in the remaining eye tissue. However, we did observe an increase of anti-cleaved caspase 3 immunostaining, suggesting that the loss of eye tissue was caused by apoptosis. Evidence from our lab indicates that Psq has a pro-apoptotic function through the activation of the gene *reaper*. A construct containing 4kb 5' of *reaper* promoter has shown increased expression in response to

Pipsqueak overexpression (Ferres-Marco, 2010). Additionally, *reaper* has GAGA binding sites in its promoter (Farkas *et al.*, 1994), which have also been described as Psq binding sites (Schwendemann and Lehmann, 2002). As we have found that sumoylation attenuates Psq transcriptional activity and that Psq has a pro-apoptotic role through the activation of *reaper* expression, blocking sumoylation in the eye could lead to an increase of *reaper* levels resulting in cell death.

As mentioned in the *Introduction*, most sumoylation targets seem to undergo rapid cycles of sumoylation and desumoylation that result in only a small portion of the protein being sumoylated at a given point (frequently less than 1%). Moreover, there is a tightly controlled balance between the amount of sumoylated and non-sumoylated protein. In the case of Psq, sumoylation could serve as a mechanism to rapidly modify the transcriptional activity of this BTB protein. We have shown that, in the context of *CG9925* expression, sumoylated and non-sumoylated Psq have different transcriptional activities, with non-sumoylated Psq being a stronger transcriptional activator. There are at least two hypothetical models to explain these observations:

i) Psq sumoylation allows the recruitment of transcriptional repressors (**Figure 50A**). As the amount of sumoylated Psq increases in the cell there is a decrease in the expression of genes activated by Psq activity (see following *Discussion* section). Conversely, blocking Psq sumoylation impedes the interaction with repressor factors through SUMO recognition, causing an increase in transcriptional activity.

ii) Another less straight forward explanation is that non-sumoylated Psq is able to interact with transcriptional co-activators and this interaction is disrupted upon Psq sumoylation, leading to a decrease in transcriptional activity mediated by this transcription factor (**Figure 50B**).

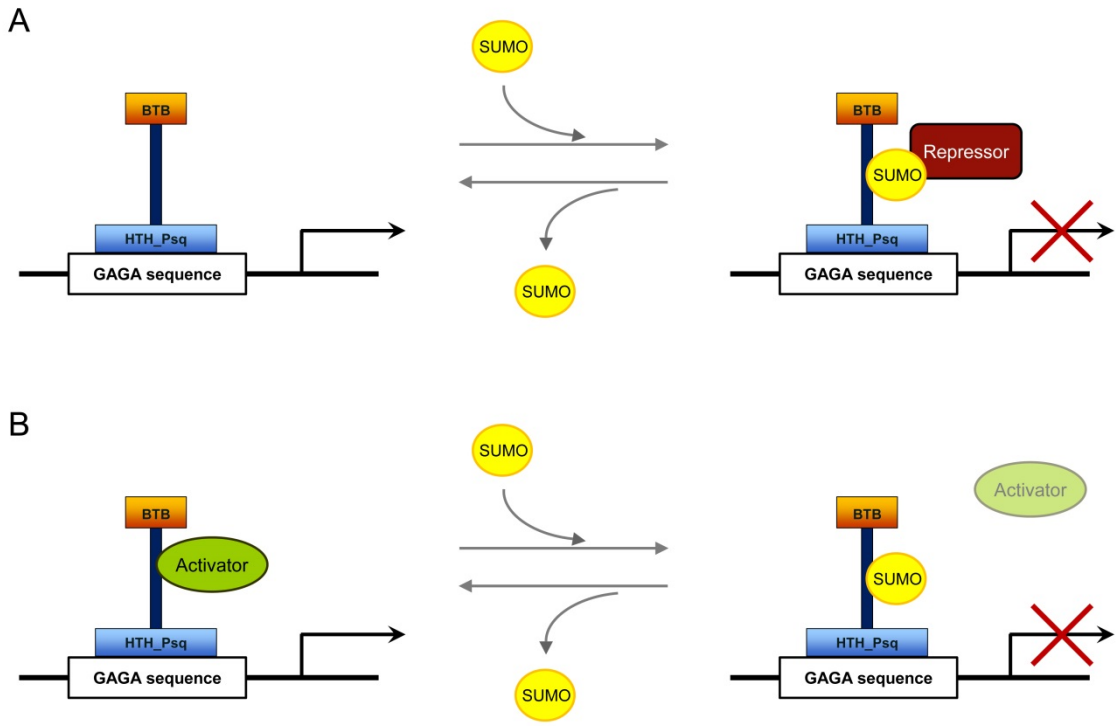


Figure 50. Modulation of Pipsqueak transcriptional activity via sumoylation. (**A**) Recruitment of transcriptional repressors through SUMO interaction inhibits Psq transcriptional activity. (**B**) Psq sumoylation hinders the interaction with transcriptional activators resulting in inhibition of Psq transcriptional activity.

Indeed, it has been previously suggested that Psq may have roles in both transcriptional activation and silencing (Huang *et al.*, 2002; Schwendemann and Lehmann, 2002). In the first work, genetic evidence indicates that Psq could be involved in the targeting of the major Pc-G protein complex CHRASCH to the GAGA sequences that are commonly found in Pc-G response elements), while the second showed that Psq and Trithorax-like bind to the same chromosome regions and suggested that both proteins cooperate in the transcriptional activation and silencing of homeotic genes. The hypothesis of sumoylation-dependent modulation of Psq transcriptional activity could be a general occurrence among target genes activated by Psq, like homeotic genes (Huang *et al.*, 2002), and not be limited to *CG9925* regulation. However, this model could also explain how Psq mediates gene silencing through the recruitment of repressors in a sumoylation-dependent manner. Analysis of expression of additional genes regulated by Psq will be necessary to confirm this hypothesis. Therefore, the main conclusion will be that Psq sumoylation modifies its transcriptional activity by allowing the interaction with many new proteins, which would modulate its transcriptional profile. As starting point for a new study, it would be very informative to perform ChIP-seq or microarray experiments to detect changes in the

binding to target genes or in gene expression promoted by blocking Psq sumoylation. Knowing the identity of genes differentially regulated by wild-type and sumoylation-defective Psq will allow us to formulate more reliable hypothesis regarding Psq sumoylation function.

In this Thesis, our Y2H assay has identified three new Psq interacting partners: DmTAF3, MEP-1 and Su(var)2-10. In the following sections we will discuss how these SUMO-mediated interactions fit in the proposed model.

2. Recruitment of transcriptional repressors through SUMO-SIM interaction: regulating gene transcription in a highly dynamic way

One of the new Psq partners identified in our Y2H assay is MEP-1, a sumoylation dependent transcriptional repressor. MEP-1 loss-of-function experiments show that a decrease in MEP-1 in a context of *psq* overexpression leads to lethality, probably due to severe defects during the development of the eye disc that result in flies with no head. In agreement with this, MEP-1 gain-of-function experiments show that MEP-1 overexpression is able to rescue, at least partially, the phenotypes of variegation and tumorigenesis mediated by *psq* overexpression. Therefore, our genetic *in vivo* experiments indicate that MEP-1 restricts Psq-mediated epigenetic silencing and oncogenic function. Interestingly, MEP-1 loss-of-function in a context of *psq* overexpression mimics the lethal phenotype obtained when overexpressing sumoylation-defective Psq mutants, suggesting that MEP-1 interaction with Psq through SUMO is important for Psq activity *in vivo*.

Considering all this evidence, MEP-1 fits the first situation of the model proposed in the previous section: Psq sumoylation allows the recruitment of transcriptional repressors. However several questions arise from this: How does MEP-1 repress Psq activity? And how does MEP-1 interact with Psq?

Regarding the first question, MEP-1 has been previously involved in transcriptional repression. First in *C. elegans*, where MEP-1 has been found to act together with the MOG proteins to repress the sex determining factor *fem-3* mRNA (Belfiore *et al.*, 2002), which promotes male development and is involved in the switch from spermatogenesis to oogenesis during *C. elegans* development (Barton *et al.*,

1987; Hodgkin *et al.*, 1985). Later on, MEP-1 has been shown to interact with Mi-2, a component of the Nucleosome Remodelling and Deacetylation (NuRD) complex (Unhavaithaya *et al.*, 2002). However, the work of Kunert *et al*., (2008) shows that this MEP-1/Mi-2 complex is distinct from NuRD and is able to promote changes in chromatin structure without the NuRD associated histone deacetylase activity. Both proteins form a repressor complex that is recruited by DNA-bound sumoylated transcription factors, which likely promote gene silencing through chromatin remodelling (Kunert *et al.*, 2009; Leight *et al.*, 2005; Stielow *et al.*, 2008). Our experiments show that Psq sumoylation decreases its transcriptional activity, at least in the context of *CG9925* expression. Sumoylated Psq could recruit a MEP-1 containing repressor complex, thus modifying the transcriptional outcome of Psq binding to the *CG9925* locus. In support of this hypothesis, loss-of-function experiments using Mi-2 RNAi transgenic flies show the same phenotype as MEP-1 RNAi: decreasing the expression of either MEP-1 or Mi-2 in a *psq* overexpression background produces lethality in the flies (data not shown). Even if the amount of sumoylated Psq is less than 1% of the total amount of the protein, association with complexes able to modify chromatin structure can translate into huge and lasting changes in gene expression. The recruitment of such a complex could explain how Psq represses the expression of its target genes.

Concerning how MEP-1 interacts with Psq to mediate sumoylation-dependent transcriptional repression, *in silico* analysis of MEP-1 protein sequence have revealed that MEP-1 possesses a putative SIM in the region that interacts with Pipsqueak. SIMs allow SUMO recognition by non-covalent interactions. However, the affinity between SUMO and SIMs is typically rather low, and standard techniques to validate protein interaction (such as co-immunoprecipitation) are not useful, complicating the validation of such interaction. Mutation of the hydrophobic core residues within the SIM to alanine has been shown to abolish the interaction with SUMO (Parker and Ulrich, 2012). To validate the functionality of the putative MEP-1 SIM, we will use directed mutagenesis to introduce the mutations V241A, V242A, N243A and L244A in the SIM hydrophobic core (MEP-1 SIM*). Then, we will generate plasmids to allow expression in bacteria of HIS-tagged versions of wild-type MEP-1 and MEP-1 SIM*. We will perform pull-down assays confronting purified wild-type MEP-1 and MEP-1 SIM* to lysates of cells overexpressing SUMO. If MEP-1, but not MEP-1 SIM*, is able to interact with SUMO, this would mean that MEP-1 has a functional SIM. Still we would lack biochemical evidence that MEP-1 binds to sumoylated Psq. Although not definite, a good evidence supporting this hypothesis would have been a 1-by-1 Y2H assay between MEP-1 and

Psq, showing that MEP-1 binds to the Psq region containing the sumoylation site. Unfortunately, we have performed this experiment with DmTAF3 and Su(var)2-10, but not with MEP-1. So to prove that MEP-1 SIM is able to interact with sumoylated Psq we should perform a pull-down assay confronting purified wild-type MEP-1 or MEP-1 SIM* to lysates of cells overexpressing SUMO and Psq (wild-type or sumoylation-defective). The results of these experiments will serve to validate or discard the hypothesis of MEP-1 being recruited by Psq via SUMO recognition. Additional experiments will be necessary to prove that MEP-1 is the responsible for Psq transcriptional activity attenuation upon sumoylation.

Regarding Su(var)2-10, our *in vivo* loss-of-function experiments show that a decrease in the expression levels of this E3 SUMO ligase enhances Psq-mediated epigenetic silencing and tumorigenesis, while the overexpression is lethal. These results suggest that reducing *Su(var)2-10* levels increases Psq activity, which is consistent with the results obtained with MEP-1 gain-of-function experiments. This fits a model in which Su(var)2-10 sumoylates Psq, allowing the recruitment of a MEP-1 containing repressor complex that will result in transcriptional repression of Psq target genes. However, it would be interesting to perform an *in vitro* sumoylation assay, to prove that Su(var)2-10 acts as a E3 SUMO ligase, as there is not biochemical evidence of this in the bibliography, and also to show that Su(var)2-10 is responsible for Psq sumoylation.

3. New role for an old acquaintance: DmTAF3 recruitment via SUMO-SIM interactions and its role in Pipsqueak activity

Another interesting hit found in the Y2H assay was DmTAF3, an integral subunit of the transcription factor TFIID. As was previously mentioned in the *Introduction*, metazoan TAF3 contains a C-terminally located PHD finger, which binds with high affinity and specificity to H3K4me3 (Vermeulen *et al.*, 2007). This interaction may serve to anchor TFIID to activated promoters or to promoters of silent genes that are poised for activation, as has been demonstrated for some genes required for embryonic stem cell differentiation (Bernstein *et al.*, 2006; Mikkelsen *et al.*, 2007). Our genetic experiments show that decreasing DmTAF3 levels reduces Psq-mediated variegation and tumorigenesis, meaning that DmTAF3 activity is required for Psq-mediated epigenetic

silencing and oncogenesis *in vivo*. This effect was shown to be specific of a *psq* overexpression context, as decreasing DmTAF3 levels in an oncogenic condition caused by overexpression of *Dl* and *miRNA278* did not alter in any way the tumour phenotype observed in this tumour paradigm.

These results and the previous knowledge on mammalian TAF3 suggest that DmTAF3 interaction with Psq could be directing the TFIID transcriptional activator complex to specific regions in the DNA, resulting in DmTAF3 acting as a transcriptional co-activator and boosting Psq activity. However, this explanation is contradicted by our observations that Psq sumoylation diminishes its transcriptional activity. If DmTAF3 is actually interacting with Psq in a SUMO-dependent manner, how could this translate into transcriptional repression? One possibility is that DmTAF3 is indeed acting as a co-activator, as sumoylated Psq can interact with many different proteins via SUMO-SIM interaction. Therefore, in a given context either activating or repressing factors can interact with Psq influencing its transcriptional activity. The balance between these interactions will decide the final response. Another possibility is that DmTAF3 is actually acting as a transcriptional repressor. In agreement with this hypothesis, the work of Bereczki *et al.,* (2008), has shown that both human and *Drosophila* TAF3 act as a negative regulator of p53 transcription activation function. Binding of DmTAF3 to p53 decreases p53 transcriptional activity partially due to a decrease in p53 protein levels. The authors suggest that DmTAF3 could recruit an ubiquitin ligase to promote p53 degradation. If we incorporate this into our model, sumoylated Psq would be recognized by DmTAF3, leading to the recruitment of a ubiquitin ligase that would promote Psq degradation, resulting in a decrease in Psq transcriptional activity. Moreover, there is yet another example in the literature that explains how DmTAF3 could be attenuating Psq transcriptional activity. The PHD domain of the KAP1 co-repressor functions as an intramolecular E3 ligase for sumoylation of an adjacent domain (Ivanov *et al.*, 2007). DmTAF3 PHD domain has been shown to directly bind to the H3K4me3, but could have an additional role as SUMO ligase. If this is true, instead of being an interacting partner of sumoylated Psq as MEP-1, DmTAF3 may be sumoylating Psq, like Su(var)2-10 does (as implied by our results). Then, sumoylated Psq could be recognized by a MEP-1 containing repressor complex leading to transcriptional repression.

Recently it has been demonstrated that TAF3 has an additional role in mediating long-range chromatin regulatory interactions that safeguard the finely-balanced transcriptional programs underlying pluripotency (Liu *et al.*, 2011). This work shows that TAF3 binds the architectural protein CTCF to mediate regulatory

interactions between distal CTCF/cohesin bound regions and proximal promoters, thus facilitating long-distance DNA looping. Though this function is mediated by a vertebrate-specific TAF3 domain, there is the possibility that DmTAF3 interaction could be promoting changes in chromatin conformation leading to changes in gene expression. Therefore, DmTAF3 ability to mediate chromatin conformation changes and its interaction with Psq upon sumoylation may additionally explain the variations in Psq transcriptional activity upon sumoylation.

Finally, and similar to what occurs with MEP-1, we still need to validate the functionality of the SIM found after *in silico* analysis of DmTAF3 sequence. The result of the 1-by-1 Y2H experiment showing that DmTAF3 interacts with the central region of Psq, which contains the sumoylation site, supports that DmTAF3 could be interacting with Psq through SUMO. As in the case of MEP-1, it would be necessary to perform pull-down experiments comparing the ability to recognise SUMO and sumoylated Psq of wild-type versus mutated DmTAF3. In this sense, we have already generated a DmTAF3 carrying the mutations P866A, I867A, E868A and V869A in the hydrophobic core of the putative SIM. Both the mutant and the wild-type DmTAF3 have been subcloned into plasmids for bacterial expression of HIS-tagged versions of these proteins. We are currently setting up the conditions to perform the pull-down assays confronting purified wild-type DmTAF3 or mutant DmTAF3 SIM* to lysates of cells overexpressing SUMO and Psq (wild-type or sumoylation-defective).

4. Conservation of sumoylation-dependent regulation of BTB transcription factor activity

Psq is a member of the BTB transcription factor family, which comprises proteins present in viruses and throughout eukaryotes, from fungi to plants to metazoans. Some proteins are made up of just a BTB; however, it is more frequent for the BTB to be combined with other domains. Over two dozen different domains are found associated with the BTB in proteins, of which five are much more frequent than the others. They are the MATH, Kelch, NPH3, Ion transport and zinc finger domains (reviewed in Perez-Torrado *et al.*, 2006). Psq belongs to the last group and, although this protein does not have a homolog outside of the Arthropod group, the architecture of BTB-ZF proteins is highly stereotypical and is well conserved through evolution (see **Figure 9** of the

Introduction). Psq DNA-binding domains are not zinc finger, but instead four tandem copies of the Psq-type helix-turn-helix (HTH) DNA-binding domain. This domain is structurally related to known DNA-binding domains, both in its repetitive character and in the putative three-alpha-helix structure of the Psq motif, but it lacks the conserved sequence signatures of the classical eukaryotic DNA-binding motifs.

The BTB-ZF transcription factors use the BTB domain as a protein-protein interface, either for oligomerization or for interaction with other proteins, while DNA-binding zinc finger motifs determine sequence specificity (Bardwell and Treisman, 1994). However, little is known about the central region located between both domains. In the last years, a repressor domain called RD2 has been identified in the central region of the BTB transcription factor PLZF (Melnick *et al.*, 2000b). Although the RD2 domain is less characterized and understood than the BTB/POZ domain, it has been shown that mutations in this domain modulate the transcriptional activity of PLZF (Kang *et al.*, 2003; Melnick *et al.*, 2000b). Moreover, the work of Kang *et al.*, demonstrated that Lys242, a residue contained in the RD2 domain, is target of sumoylation *in vivo*, which abolishes the DNA-binding ability of PLZF. Another recent study has shown that Metadherin protein binds to PLZF sumoylation site, hindering PLZF ability to bind the *c-myc* promoter and abolishing its transcriptional repression. Additionally, Metadherin binding to PLZF increases its complex formation with HDAC4 instead of HDAC1, two well-known HDACs with which PLZF interacts to achieve its inhibitory effect through chromatin remodelling (Thirkettle *et al.*, 2009).

Our experiments have revealed Lys633 as a functional sumoylation site located in the central region of Psq, between the BTB and the DNA-binding domains. Although there is no repressive domain defined in that part of the protein, we have demonstrated that Lys633 sumoylation attenuates Psq transcriptional activity, probably due to the recruitment via SUMO-SIM interaction of transcriptional repressors, like MEP-1. These results, combined with the information regarding PLZF sumoylation, suggest that sumoylation could be a general mechanism to regulate the transcriptional activity of BTB-ZF transcription factors. In support of this hypothesis, many BTB-ZF transcription factors in *Drosophila* possess sumoylation sites in the central region of their protein sequence; in addition to BTB proteins from other subfamilies. The same occurs in human, where in addition to PLZF, many members of the BTB-ZF family which have been characterized as important transcriptional factors involved in development and cancer also have putative sumoylation sites: BCL6 (Lys391) (Albagli-Curiel, 2003; Polo *et al.*, 2004), leukaemia/lymphoma related factor (LRF)/Pokemon (Lys539) (Maeda *et al.*, 2005), hypermethylated in cancer (HIC)1 (Lys314) (Chen *et al.*, 2004; Pinte *et al.*,

2004) and Myc interacting zinc finger (MIZ)1 (Lys138, Lys166, Lys251 and Lys696) (Kim *et al.*, 2002b). Further experiments will be needed to prove the functionality of all these putative sumoylation sites and the effect of this post-translational modification over the function of these proteins.

5. Pipsqueak activity is modulated via proteolytic processing

The results presented in this Thesis show that Psq transcriptional activity requires a specific proteolytic processing. This event takes place only in the insoluble fraction of the cell, most likely in the nucleus, and our experiments indicate that Psq must be in that compartment to be processed in this manner. We have demonstrated that this specific processing, called Type I, generates two groups of fragments: ~30-50 kDa N-terminal fragments, and ~120-150 kDa C-terminal fragments. Chemical inhibition experiments have shown that proteasome activity is responsible for Psq Type I processing, and this process is required for Psq transcriptional activity, measured as CG9925 derepression. Finally, we still do not know the effect of Type I processing over Psq tumorigenic activity *in vivo*, as deletions 2 and 3 abolish Psq nuclear localization and hence impede Psq processing. To assess the relevance of this processing event for Psq-driven oncogenesis we should first locate more precisely the processing site and generate new mutants that only affect Psq processing, and not its subcellular distribution.

But how does the proteasome interact with Psq? And how does it promote limited proteolysis instead of complete degradation? As it has been previously mentioned in this Thesis (see *Introduction*), Cul3-based ligases ubiquitinate target proteins to mark them to be processed or degraded by the proteasome (Xu *et al.*, 2003 and reviewed in Pintard *et al.*, 2004). In the case of Psq, genetic evidence from our group indicates that there is an interaction between Psq and Guftagu (Gft), the *Drosophila* homolog of Cul3, relevant for Psq-mediated tumorigenesis *in vivo* (Gontijo A., Caparrós E., *et al*., in preparation). Psq ubiquitination via Cul3/Gft activity could serve as a mark for proteasome-mediated proteolytic processing, as it has been proposed for other SCF-type ubiquitin ligase substrates, such as NF-κB and *Drosophila* Ci (Jia *et al.*, 2005; Orian *et al.*, 2000). Despite being structurally unrelated, both transcription factors need proteasome limited proteolysis to exert their functions.

Additionally, it has been shown that homodimerization (NF-κB) or interaction with heterologous partners (Ci) (reviewed in Rape and Jentsch, 2002), play a role in protecting these substrates against total degradation. One function of the BTB domain is to mediate the formation of homo- or heterodimers. In the case of the BTB-ZF transcription factors PLZF and BCL-6, dimerization through this domain is an architectural feature necessary for their normal function (Ahmad *et al.*, 1998; Ahmad *et al.*, 2003; Melnick *et al.*, 2000a). In the case of Psq, at the moment of writing this Thesis there was no evidence suggesting the formation of dimers or higher order associations through BTB interaction. However, as Psq Type I processing is required for its transcriptional activity, dimerization could be necessary to protect Psq from total degradation while is activated through proteasome-dependent limited proteolysis.

Another very interesting explanation of how the proteasome processes Psq arises from the work of Andreas Matouschek's group. Using *in vitro* and cell culture assays, both with mammalian and *Drosophila* cells, they have demonstrated that proteasome-mediated proteolytic processing occurs when the progression of the proteasome along its substrate's polypeptide chain is stopped by a signal consisting of two components: a sequence of simple aa composition ("simple sequence") followed by a tightly folded domain in the direction of proteasome movement (Tian *et al.*, 2005). Other groups had previously shown that the proteasome stops degradation of a polypeptide chain when it reaches a tightly folded domain (reviewed in Rape and Jentsch, 2002). Matouschek's group has demonstrated that in the case of Ci, the "simple sequence" is rich in asparagine, serine and glutamine residues and the folded domain is a C2H2 zinc finger, whereas in NF-κB precursor protein p105 the "simple sequence" is rich in glycine residues and the folded domain is the Rel-homology domain. The strength of the signal depends on the complexity of the "simple sequence" and the resistance of the folded domain to unravelling, and both components of the signal are transferable. This work shows that the processing mechanism, but not the sequence of the signal, is conserved between flies and humans, and Psq could be a new example of this processing mechanism.

The proteasome can also initiate degradation at internal unstructured loops within large proteins (Fishbain *et al.*, 2011; Kraut and Matouschek, 2011; Liu *et al.*, 2003; Prakash *et al.*, 2004). Our experiments show that Psq Type I processing likely occurs between the BTB domain and Type II processing site. The analysis of the region of Psq spanned by deletions 2 and 3 shows that it does not contain any known domain and has low complexity, suggesting that is not tightly folded and could be accessible. There is the possibility that the central unstructured region of Psq forms a

loop that fits in the proteasome and serves as starting point for degradation in a processive manner, resulting in two groups of processed Psq fragments instead of two fragments. Actually, there is evidence indicating that the proteasome channel that leads to the degradation chamber can accommodate more than one polypeptide chain at once (Lee *et al.*, 2002; Liu *et al.*, 2003), but crystal structures of the proteasome core particles show that the channel will be a tight fit for two chains (Groll *et al.*, 2000). This tight fit or the load put on the proteasome by the simultaneous presence of two polypeptide chains and their folded domains could reduce the effectiveness of the unfolding and degradation machinery, and thus the proteasome's processivity. In view of all this data, a model in which the proteasome mediates partial proteolysis of Psq through the recognition of an internal degradation site located in the central unstructured region of the protein appears as an interesting possibility that will require further consideration.

Another important issue is how this proteolytic event could affect Psq transcriptional activity. An interesting option could be that proteolytic processing Type I is used as a mean to separate the BTB domain from the DNA binding domain of Psq. We will consider this possibility, although our experiments suggest that Psq BTB domain remains attached to the DNA-binding fragment via disulphide bond/s. One of the functions of the BTB domain is the recruitment of transcriptional repression complexes, such as N-CoR and SMRT, which in turn draw HDACs to target promoters (Melnick *et al.*, 2002). In the case of Psq, it has been previously involved in sequence-specific targeting of a Pc-G complex that contains HDAC activity (Huang *et al.*, 2002). However, a specific role of the BTB domain in this interaction remains to be confirmed. BTB-mediated recruitment of repressor complexes could be needed to inhibit expression of certain Psq targets. Upon Type I Psq processing, the inhibiting factors would be physically removed, BTB-dependent repression would stop and Psq may function now as an activator, explaining why we do not observe *CG9925* derepression when we inhibit proteasome activity (i.e. Type I processing). An interesting experiment could be to express a mutant Psq protein lacking the 40 kDa N-terminal fragment generated by Type I processing and test its capacity to activate transcription by itself. A curious fact is that this fragment is similar to the short Psq isoforms required mostly during eye development.

On the other hand, we have shown that Psq can undergo an additional independent proteolytic processing, Type II, which occurs between Asn415 and Ala636 and generates two Psq fragments of 70 and 110 kDa. Our qPCR experiments show that, contrary to Type I processing, Type II is not required for Psq transcriptional

activity, suggesting that it must be involved in a different process. This result is in agreement with the observation that Type II processing only takes place in the soluble fraction of the cell, while Psq needs to be in the nucleus to execute its function as a transcription factor. Additionally, the results derived from our *in vivo* experiments indicate that abolishing Type II processing (using mutant *psqΔ4* or *psqΔ5*) does not hinder Psq oncogenic function, suggesting that Psq Type II processing is not relevant for this process. Regarding the molecule causing Type II processing, we can only conclude that the proteasome is not directly responsible for this type of processing. Our experiments clearly demonstrate that Psq Type II processing is not blocked upon chemical proteasome inhibition, however there is an increase in the amount of Psq Type II processed fragments, suggesting the existence of a balance between both types of processing. Another striking difference between Type I and Type II is their different behaviour upon β-mercaptoethanol addition to protein samples. While β-mercaptoethanol had no effect over Type II processing (70 kDa + 110 kDa bands) it releases the ~30-50 and ~120-150 kDa fragments produced by Type I processing (**Figure 41**). This suggests that the fragments generated upon Type I processing remain together thanks to disulphide bonds, while Type II fragments are separated, highlighting again that the mechanisms underlying both types of processing are different. Additional experiments will be required to unveil the reason why these Psq fragments remain together after Type I processing.

One of the current priorities regarding Psq processing is to precisely define where the processing events occur. In the case of Type II processing, our deletion mutants have allowed us to stablish that the processing site is located between Asn415 and Ala636. However, the overlapping mutations have failed to provide information about the location of Type I processing site. In any case, taking into account the size of the fragments generated, we suggest that Type I processing occurs in a sequence located between the BTB domain and Type II processing site. In order to identify these sites we could design even smaller overlapping deletions and test their effect on Psq processing, while checking that they do not affect proper protein distribution. However, the most informative and straightforward approach would be to perform mass spectrometry experiments to define the exact sequence of the processing sites. Moreover, knowing this sequence would facilitate the search for candidate molecules responsible for Psq processing Type II. Once the processing sites have been identified we should check the subcellular distribution and effect over Psq activity of the mutants for these sites. It would also be interesting to perform chromosome squashing assays to compare if the mutants bind to the same chromosomic regions as wild-type Psq.

Differences in binding would suggest that Psq processing is critical for proper regulation of Psq target genes. To study this in further detail, we could also perform ChIP-seq analysis to see quantitative differences in the ability to bind to Psq target genes of wild-type versus non-processable Psq (Type I and Type II). Regarding Type I, identifying the sequence required for this specific processing would be useful to do comparative analysis of Psq homologs or other BTB-ZF proteins from *Drosophila* or other organisms attending to their protein sequence. These comparisons will help to find out if proteasome-mediated Psq processing could serve as a model to understand how other BTB-ZF transcription factors perform their functions.

Finally, during the course of this Thesis, we have treated Psq sumoylation and Psq processing as two separate processes that contribute to regulating the activity and function of this transcription factor. However, it is possible that they are able to influence each other. There is a group of RING-type ubiquitin ligases, the SUMO-targeted ubiquitin ligases (STUbLs), that possess affinities for SUMO conjugates due to the presence of SIMs in their sequence (reviewed in Jentsch and Psakhye, 2013; Sriramachandran and Dohmen, 2014). They are able to recognize sumoylated substrates in order to polyubiquitinate them and lead to proteasomal degradation. This could be an additional mechanism to regulate the amount of Psq in the cell. The E3 ubiquitin ligase Gft, shown to interact genetically with Psq in experiments performed by our group, possesses a putative SIM in its sequence, suggesting that it could also act as a STUbLs; unfortunately this SIM it is located in a highly ordered region of the protein organized in a globular domain, suggesting that this SIM is not functional. Moreover, preliminary experiments performed during the course of this Thesis indicate that the mutant PsqK633R protein is processed as efficiently as the wild-type (data not shown), suggesting that Psq sumoylation and processing are independent events.

6. Functional relevance of Pipsqueak sumoylation and proteolytic processing over its tumorigenic capacity *in vivo*

In 2006, the work of Ferres-Marco *et al*., unveiled a new function for Psq in tumour development. Overexpression of the genes *lola* and *psq* in combination with *Dl* overexpression, produced massive overgrowths and tumours. Genetic experiments using mutants for either *lola* or *psq* showed that *psq* was the main contributor to the

tumorigenic phenotype. Moreover, both the BTB and the DNA-binding domain of Psq were shown to be essential for its oncogenic function. Regarding the mechanisms underlying tumour development, the eye tissue from which the tumour arises showed a strong reduction of the open chromatin mark H3K4 trimethylation (H3K4me3), suggesting an increase in epigenetic silencing. Additionally, reducing the dosage of genes related to gene silencing and chromatin condensation, like *Rpd3/HDAC*, *E(z)*, *Su(var)3-9*, *Pc*, and *Esc*, impeded tumour development. Also, the increase in epigenetic silencing reduced the expression of the tumour suppressor retinoblastoma family protein (*Rbf*) gene, and this down-regulation was shown to be necessary for tumour development. These results indicate that deregulation of *psq* and *lola* might induce tumorigenesis through aberrant epigenetic silencing of genes that contribute to the uncontrolled growth of tumour cells. Recruitment of Pc-G repressors through the BTB domain of both Psq and Lola might also be involved in this process.

In addition to these findings, the work presented in this Thesis provides new evidence showing that two different processes, Psq sumoylation and Psq limited proteolysis, influence Psq transcriptional activity and its oncogenic capacity *in vivo*. Regarding Psq sumoylation, this work demonstrates that this post-translational modification attenuates Psq transcriptional activity. This could be due to the recruitment via SUMO-SIM interaction of transcriptional repressors, such as MEP-1, but further experimentation is needed to fully confirm this. Sumoylation-dependent transcriptional repression could be a mechanism working in parallel with BTB-mediated recruitment of co-repressors to inhibit transcription of target genes, leading to an increase in tumorigenesis. However *in vivo* evidence coming from genetic experiments with MEP-1 and Su(var)2-10 indirectly indicate that Psq sumoylation hinders tumour formation: MEP-1 gain-of-function (probably equivalent to increased MEP-1 interaction with Psq) reduces epigenetic silencing and tumorigenesis; while Su(var)2-10 loss-of-function (equivalent to reduced Psq sumoylation) increases epigenetic silencing and tumorigenesis. Thus, we can envision two alternative scenarios:

i) Psq sumoylation may favour tumour formation through the recruitment of additional transcriptional repressors, similarly to the mechanism proposed in Ferres-Marco *et al*., (2006).

ii) Psq sumoylation may prevent tumour development according to MEP-1 and Su(var)2-10 *in vivo* data, although these experiments do not specifically target Psq sumoylation. Supporting this second scenario, the increase observed in Psq transcriptional activity when Psq sumoylation is blocked might indicate that Psq oncogenic activity is also increased.

Unfortunately, we have not been able yet to study the role of Psq sumoylation in the context of Psq mediated tumorigenesis, as blocking Psq sumoylation in the eye of *Drosophila* is lethal for the flies. This phenotype is associated with a loss of eye tissue during development promoted by an increase in apoptosis, which leads to strong defects in the formation of the head. So, in these conditions, the pro-apoptotic activity of Psq obscures its proliferative role. To avoid this and be able to study if blocking Psq sumoylation also enhances its oncogenic activity we should first block apoptosis. If under these conditions, we end up with a decrease in tumour growth it would mean that Psq sumoylation is necessary for its tumorigenic function, suggesting that SUMO-dependent recruitment of transcriptional repressors and/or other molecules might be critical for this process. In the case that the tumorigenic phenotype does not change (or even if it increases), sumoylation would reduce Psq tumorigenic function, and Psq interactions via SUMO should prevent tumour formation.

Regarding the pro-apoptotic role of Psq, evidence from our lab indicates that Psq has a pro-apoptotic function through the activation of the gene *reaper* (Ferres-Marco, 2010). This work shows that apoptosis is required for tumour invasion, as decreasing apoptosis promotes tumour growth while abolishing invasiveness. The experiments presented in this Thesis also suggest a pro-apoptotic function for Psq, since the abolishment of Psq sumoylation in the eye leads to an increase in cell death. We have demonstrated that sumoylation attenuates Psq transcriptional activity, thus, blocking Psq sumoylation in the eye could lead to an increase of *reaper* transcription levels, resulting in cell death. Cell death contribution to the tumorigenic phenotype observed in the *psq* and *Dl* combined overexpression model is unclear, and further experiments are required to clarify the impact of apoptosis in tumour invasion. The idea of apoptosis positively contributing to the invasive process contradicts the established dogma in which development of strategies to avoid apoptosis is a critical step during tumour progression (Hanahan and Weinberg, 2011). Therefore, the possibility of Psq promoting tumour progression combining the repression of tumour suppressor genes and an increase in apoptosis is a challenging hypothesis.

Another interesting question would be the effect of sumoylation over Psq cooperation with Notch signalling to promote tumour growth. The work of Ferres-Marco *et al*., (2006), demonstrated that either overexpression of *psq* or Notch alone was not sufficient to produce a tumour; they need to cooperate with each other. Blocking Psq sumoylation could boost Psq activity to a level where Notch activity is not necessary to promote tumour growth anymore. Cancer is a multistep process in which cancer cells need to acquire different traits that will enable them to become tumorigenic and

ultimately malignant (Hanahan and Weinberg, 2011). In our tumorigenic model, a deregulation in cell growth due to Notch overexpression is combined with epigenetic silencing of tumour suppressor genes, such as *Rbf*, likely mediated mainly by the activity of Psq, probably in addition to other still unknown processes. Taking this into account, it seems unlikely that high levels of only one of these contributions would be enough to promote oncogenesis.

Concerning Psq limited proteolysis, Type I processing is necessary for Psq transcriptional activity, while Type II is not required for this activity and is not affecting Psq tumorigenic capacity. Unfortunately, we were unable to determine the effect of Type I processing over Psq-mediated tumorigenesis. Initially we performed *in vitro* and *in vivo* experiments using *psqΔ2* and *psqΔ3* mutants, as these deletions appeared to block Type I processing. Overexpression of those mutants results in the abolishment of *CG9925* transcriptional activation and absence of tumorigenic phenotype. However, examination of mutant PsqΔ2 and PsqΔ3 subcellular distribution reveals that those mutants remain in the cytoplasm, most likely because these deletions eliminate or negatively affect a specific sequence needed for Psq nuclear localization, which explains the phenotypes observed. In any case, chemical inhibition of proteasome activity shows that Psq Type I processing is indeed required for Psq transcriptional activity. Consistent with this, we hypothesize that Psq Type I processing will be required for Psq tumorigenic activity.

In summary, the work presented in this Thesis broadens our knowledge on Psq activity respect to the work of Ferres-Marco *et al*., (2006), unveiling new mechanisms underlying Psq-mediated tumorigenic function and highlighting the relevance of combining different regulation processes to fine tune the activity of this transcription factor. However, further experiments are still needed to prove some of the hypothesis suggested in this manuscript and to verify if Psq could serve as a model to understand how other BTB-ZF transcription factors carry out their functions during oncogenesis.

UNIVERSITAS Miguel Hernández

V. MATERIALS AND METHODS

UNIVERSITAS Miguel Hernández

UNIVERSITAS
Miguel
Hernández

1. Fly genetics

1.1 List of mutant and transgenic fly lines

The complete list of mutant and transgenic fly lines used in this Thesis, as well as a short description of each line, can be found in *Appendix I*.

The following transgenes were designed by our group and injected to generate transgenic fly lines by BestGene injection Services: *UAS-3xFLAG::psqB::Tag-RFP-T*, *UAS-3xFLAG::psqBK518R::Tag-RFP-T*, *UAS-3xFLAG::psqBK633R::Tag-RFP-T*, *UAS-3xFLAG::psqBKKRR::Tag-RFP-T*, *UAS-3xFLAG::psqBΔ2::Tag-RFP-T*, *UAS-3xFLAG::psqBΔ3::Tag-RFP-T*, *UAS-3xFLAG::psqBΔ4::Tag-RFP-T* and *UAS-3xFLAG::psqBΔ5::Tag-RFP-T*.

1.2 Image acquisition

Drosophila images (eyes, pharates) were captured on an optical microscope ZEISS Axiophot, using a MicroPublisher 5.0 camera (QImaging) and the QCapture software (QImaging). All pictures were taken using a 5X objective with 1.5X zoom. Each image is a composite of 15 to 25 images of the same sample focused at different heights of the specimen. The in-focus composites were generated using the software AutoMontage Essentials 5.0.

1.3 Heat shock experiments

hs-Gal4;UAS-psqB second instar larvae and adult flies were heat shocked during 10 minutes at 37ºC. After this, the animals were homogenized at the indicated times with RIPA buffer (50 mM Tris-HCl [pH 8], 150 mM NaCl, 1% Nonidet P-40, 0.5% Sodium deoxycholate, 0.1% SDS), supplemented with phosphatase inhibitors (1 mM Sodium orthovanadate [Na_3VO_4] and 1mM Sodium fluoride [NaF]), and protease inhibitors (2 mM Pefabloc from Alexis Biochemicals [ref. #270-022-M050] and 1X cOmplete Mini EDTA-free protease inhibitor cocktail from Roche [ref. #11836170001]). Total cell lysates were incubated for 30 minutes at 4ºC, and stored at -80ºC until they were analysed by Western Blot, as described below.

2. Yeast two-hybrid experiments

2.1 Yeast two-hybrid screen

The coding sequence for aa Met1-Gln720, which contains the BTB and the central region of the *Drosophila* Psq protein (GenBank accession number gi: 24652499, FlyBase ID FBgn0004399), specifically the isoform B, was used as bait for the assay. The sequence was PCR-amplified and cloned in frame with the LexA DNA-binding domain (DBD) into plasmid pB29 (orientation N-bait-LexA-C). pB29 bait plasmid derives from the original pBTM116 (Vojtek and Hollenberg, 1995). The DBD constructs were checked by sequencing the entire inserts. The bait plasmid was transformed in the yeast strain L40ΔGal4 (mata) (Fromont-Racine *et al.*, 1997).

As prey we used a *Drosophila* embryo library from Hybrigenics Services, which is an equimolar pool of two cDNA libraries prepared from 0-12 hours (zygotic + maternal mRNA) and 12-24 hours embryo mRNA. The different cDNA were cloned in frame with the Gal4 activation domain (AD) into plasmid pP6, derived from the original pGADGH (Bartel *et al.*, 1993). The AD constructs were checked by sequencing the insert at its 5' and 3' ends.

Interaction assays are based on the reporter gene *HIS3* (growth assay without histidine). The general logic of the assay is as follows: upon physical binding of protein X with protein Y, the DNA Binding Domain (DBD) of a transcriptional activator is brought in close proximity to its Activation Domain (AD) counterpart. Reconstitution of a functional transcription factor activates the production of an auxotrophy marker (HIS3 in this assay), which in turn allows His^- yeast cells to grow on a selective medium lacking histidine. The DBD constructs were transformed in L40ΔGal4 (mata) yeast cells and the AD constructs in Y187 (matα) yeast strain. The interactions were then tested using a mating approach to generate diploid yeast that will express both fusion proteins, as previously described (Fromont-Racine *et al.*, 1997). The screen was first performed on a small scale to test the autoactivation of the bait, its toxicity and to select the most appropriate selective medium for the assay. The selective medium DO-2, which lacks tryptophan and leucine, was used as a growth control. DO-3 medium, which lacks tryptophan, leucine and histidine, was used to test protein interaction. The full-size screen was performed using 50mM of 3-aminotriazol (3-AT), an inhibitor of imidazole glycerol phosphate dehydratase, the product of the *HIS3* reporter gene. This increases stringency and reduces possible autoactivation by the bait proteins.

Following the procedure described in (Formstecher *et al.*, 2005), 78.2 million interactions were tested, from which 154 positive clones were selected on DO-3 selective medium plus 3-AT. The corresponding prey fragments were amplified by PCR and sequenced at their 5' and 3' ends. They were identified by sequence comparison with the release 3.1 of Berkeley *Drosophila* Genome Project (BDGP) using BLASTN (Altschul *et al.*, 1997). For the complete list of positive clones see *Appendix II*.

For each interaction, a Predicted Biological Score (PBS) was computed to assess interaction reliability. This score represents the probability of an interaction being non-specific. The scores are divided into four categories, from A (lowest probability) to D (highest probability). A fifth category, E, specifically tags interactions involving highly connected prey domains. This category represents highly likely two-hybrid artefacts. The PBS has been shown to positively correlate with the biological significance of protein interactions (Colland *et al.*, 2004; Formstecher *et al.*, 2005; Rain *et al.*, 2001; Terradot *et al.*, 2004; Wojcik *et al.*, 2002).

2.2 One-by-one yeast two-hybrid experiments

The coding sequence for aa Met1-Gln720, Met1-Asp131 and Ala132-Gln720 of the *Drosophila* Psq protein (GenBank accession number gi: 24652499, FlyBase ID FBgn0004399), specifically isoform B, were PCR-amplified and cloned in frame with the LexA DBD into plasmid pB29 (orientation N-bait-LexA-C). The DBD constructs were checked by sequencing the entire inserts, and then transformed in L40ΔGal4 (mata) yeast cells, as in the case of the yeast two-hybrid screen.

Prey fragments, corresponding to aa Glu841-Asp1067 of the *Drosophila* DmTAF3 protein (GenBank accession number gi: 24638624, FlyBase ID FBgn0026262), and to aa Thr363-Leu601 of the *Drosophila* Su(var)2-10 protein (FlyBase ID FBgn0003612), were extracted from the yeast two-hybrid screening of Psq (aa Met1-Gln720) against Hybrigenics Services *Drosophila* embryo cDNA library. These fragments were cloned in frame with the Gal4 AD into plasmid pP6. The AD constructs were checked by sequencing the insert at its 5' and 3' ends, and then transformed in Y187 (matα) yeast strain.

The interaction pairs were tested using a mating approach as previously described (Fromont-Racine *et al.*, 1997). Interaction pairs were tested in duplicate, as two independent clones from each mating reaction were picked for the growth assay. For each interaction, several dilutions (10^{-1}, 10^{-2}, 10^{-3} and 10^{-4}) of the diploid yeast cell culture normalized at $5x10^4$ cells and expressing both bait (DBD fusion) and prey (AD

fusion) constructs were spotted on several selective media (DO-2 and DO-3). Six different concentrations of the inhibitor 3-AT were added to the DO-3 plates to reduce the background generated by baits that activate transcription alone (so-called autoactivating baits). The following 3-AT concentrations were tested: 1, 5, 10, 50, 100 and 200 mM.

3. Generation of plasmid constructs

The sequence of SUMO and SUMOΔGG transgenes are optimized versions of *Drosophila smt3* (FlyBase ID FBgn0264922) for enhanced expression. A hemagglutinin (HA) tag was added in the 5' end of the transgenes to facilitate detection of the protein. These transgenes were synthetized (GeneArt® Gene Synthesis) and cloned in the *Drosophila pUAST* vector using EcoRI/NotI restriction sites.

To generate the *pUAST-psqB* construct, the full length cDNA of *psqB* was subcloned from a *pOT2* plasmid containing *psq* EST LD33470, and inserted in the *pUAST* vector. For that, the *pOT2* plasmid was digested with DraI/XbaI restriction enzymes, and the DraI end was turned into a blunt end to insert the fragment in the XbaI restriction site of *pUAST*. The generation of this construct was performed by Jorge Bolívar in his laboratory.

The transgene *UAS-3xFLAG::psqB::Tag-RFP-T* (abbreviated as $psqB^{Tag}$) was generated optimizing the sequence of *Drosophila psq* isoform B (FlyBase ID FBtr0088275) for enhanced expression. Two different tags were added at the ends of the transgene: a 3xFLAG tag was added in the 5' end, and a Tag-RFP-T (a monomeric red fluorescent protein generated from the wild-type RFP from sea anemone *Entacmaea quadricolor* (Merzlyak *et al.*, 2007) in the 3' end. The transgene was synthetized (GeneArt® Gene Synthesis) and cloned in the *Drosophila pUAST* vector using EcoRI/NotI restriction sites, or in *Drosophila pMT* vector using EcoRI/BamHI restriction sites.

The transgenes *UAS-3xFLAG::psqBK518R::Tag-RFP-T* (abbreviated as $psqBK518R^{Tag}$), *UAS-3xFLAG::psqBK633R::Tag-RFP-T* (abbreviated as $psqBK633R^{Tag}$) and *UAS-3xFLAG::psqBKKRR::Tag-RFP-T* (abbreviated as $psqBKKRR^{Tag}$), were generated from the original $psqB^{Tag}$ transgene. Three different fragments containing the coding sequence for PsqB aa Thr537 to Gly722 were generated, each of them carrying the mutation K550R, K665R or K550R+K665R. Note

that K550 and K665 are the positions of the putative sumoylation sites in the PsqB^{Tag} protein, while K518 and K633 are the positions in the endogenous PsqB protein sequence, respectively. To avoid confusions we decided to name the mutations after the positions in the endogenous protein. Restriction sites for AgeI and AvrII were added to the 5' and 3' ends of the transgenes for subcloning them into *pUAST-psqB^{Tag}* and *pMT-psqB^{Tag}*, substituting the wild-type sequence.

Psq deletion mutants *UAS-3xFLAG::psqBΔ2::Tag-RFP-T* (abbreviated as *psqBΔ2^{Tag}*), *UAS-3xFLAG::psqBΔ3::Tag-RFP-T* (abbreviated as *psqBΔ3^{Tag}*), *UAS-3xFLAG::psqBΔ4::Tag-RFP-T* (abbreviated as *psqBΔ4^{Tag}*) and *UAS-3xFLAG::psqBΔ5::Tag-RFP-T* (abbreviated as *psqBΔ5^{Tag}*), were generated using the plasmid *pMA-psqB^{Tag}* as template in an inverted PCR, following the protocol detailed in Stoynova *et al*., (2004) (**Figure 51**). Specific primers flanking the sequence to be deleted were designed (**Table 5**). These primers incorporate four consecutive phosphorothioate residues, located 12 nucleotides from the 5′ ends of the primers. The product of the inverse PCR is a molecule with blunt ends lacking the fragment between both primers. PCR products were digested with Dpn for 1 hour at 37ºC to eliminate the methylated DNA used as template. Phosphorothioate residues are known to be resistant to the activity of exonucleases. The specific 5′ to 3′ double-stranded exonuclease activity of T7 gene 6 exonuclease was used to hydrolyse the unprotected DNA strand, leaving the phosphorothioate protected strand intact. By doing this, the blunt ends of the PCR products were turned into overhanging ends. The 6 outermost nucleotides of each modified primer were designed to be complementary to the 6 nucleotides of the opposite strand primer located immediately to the 5′ end of its phosphorothioate residues. This allows self-circularization and ligation of the molecule after T7 gene exonuclease digestion. The mutant constructs were then subcloned from the *pMA* plasmid into *pUAST* and *pMT* vectors using EcoRI/XbaI or EcoRI/BamHI restriction sites, respectively.

The transgene *UAS-3xFLAG::psqBΔ6::Tag-RFP-T* (abbreviated as *psqBΔ6^{Tag}*), was generated by enzymatic digestion of *pUAST-psqB^{Tag}* with BglII. This digestion generated two cohesive ends allowing self-circularization and ligation of the molecule. The deleted fragment encoded aa Ile402 to Gln1085, thus generating a PsqB protein lacking part of the central region and the DNA-binding domains, but still retaining the Tag-RFP-T tag (**Figure 44A**).

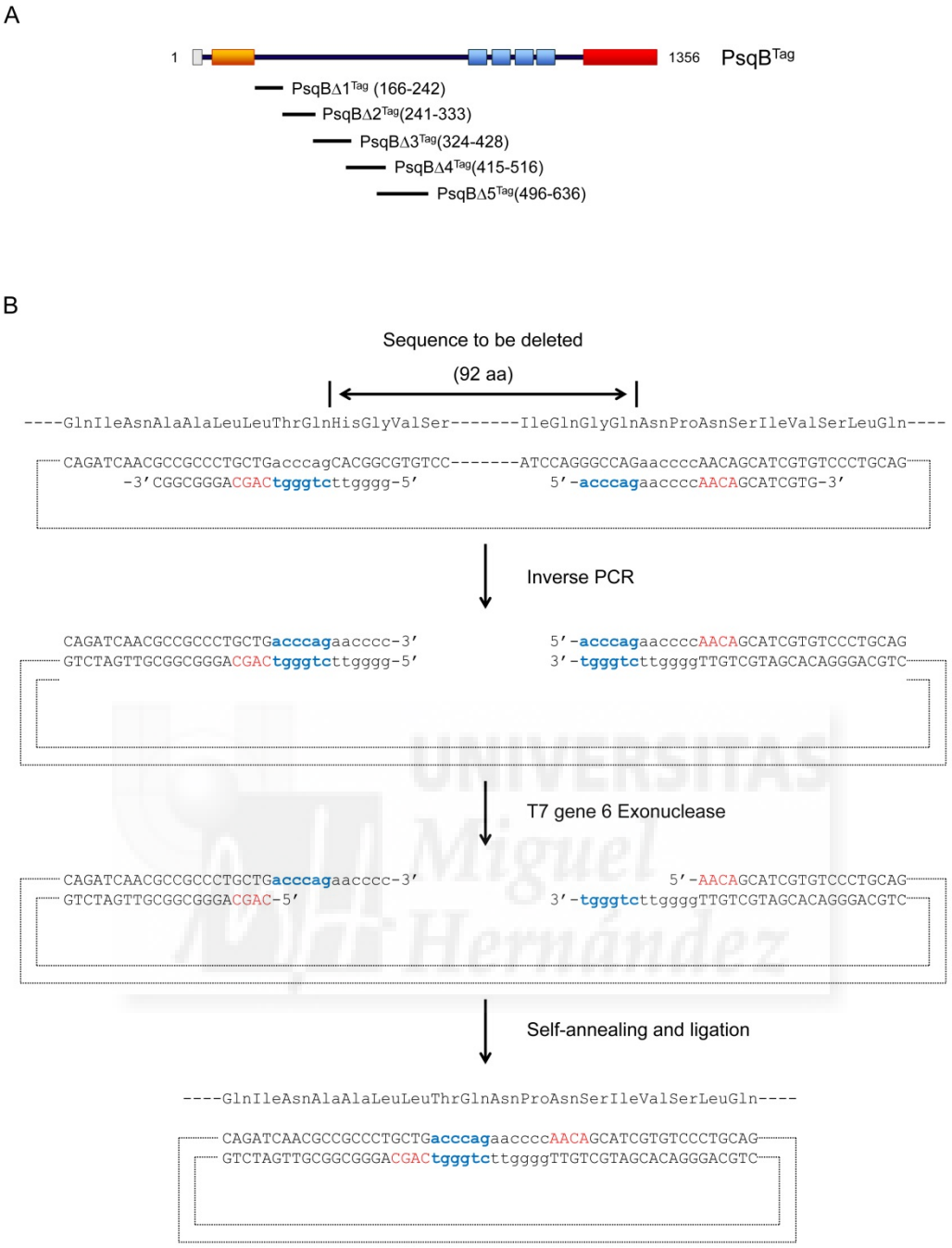


Figure 51. Relative position of Psq overlapping deletions and description of the methodology used for its generation. (**A**) Schematic representation showing the relative positions of the overlapping deletions performed on Psq. (**B**) Representation of the phosphorothioate method for generating large deletion mutants from a plasmid DNA template. The aa sequence shown at the top is part of the PsqB protein, and the fragment that was to be deleted is shown. The corresponding cDNA sequence as present in the *pMA-3xFLAG::psqB::Tag-RFP-T* plasmid is shown, together with the primers used in the inverse PCR to generate Psq deletion 2. Phosphorothioate-modified nucleotides are depicted in red. Note the 6 base extensions at the 5′ ends of each primer designed to generate complementary 3′ overhangs after T7 gene 6 exonuclease treatment. The final deletion mutant retains the correct reading frame of the original protein. Adapted from the work of (Stoynova *et al.*, 2004).

Primer name	Primer sequence
psqΔ1 forward	5'-GCCGCCGTGTCCA*G*T*G*GTTCGGTG-3'
psqΔ1 reverse	5'-GACACGGGCGGCA*T*C*G*TTCAGATC-3'
psqΔ2 forward	5'-ACCCAGAACCCCA*A*C*A*GCATCGTG-3'
psqΔ2 reverse	5'-GGGGTTCTGGGTC*A*G*C*AGGGCGGC-3'
psqΔ3 forward	5'-CACCACGTGACCA*G*T*A*CCGGCCAG-3'
psqΔ3 reverse	5'-GGTCACGTGGTGG*T*G*A*GCTGGTTG-3'
psqΔ4 forward	5'-GCCAGCCTGGCTC*A*G*C*AGATCCAC-3'
psqΔ4 reverse	5'-AGCCAGGCTGGCG*G*G*G*CTGGCCAC-3'
psqΔ5 forward	5'-AGCTGCGGTGTTC*A*T*G*GACAGCAT-3'
psqΔ5 reverse	5'-AACACCGCAGCTC*A*G*G*TCCACGTG-3'

Table 5. List of primers used for generating PsqB deletion mutants. Asterisk indicates phosphorothioate-modified nucleotides.

4. Cell culture and transfections

Schneider 2 (S2) cells (Invitrogen, ref. #10831-014) were maintained in Express Five serum free cell culture medium (Invitrogen, ref. #10486-025), supplemented with L-Glutamine (LabClinics, ref. #M11-004) and penicillin/streptomycin stock of antibiotics (Sigma, ref. #P4333-100ML). Kc167 cells (DGRC cat. no. 1) were maintained in Schneider's *Drosophila* Medium (Gibco, ref. #21720-024), supplemented with 10% inactivated fetal bovine serum (Invitrogen, ref. #10108-165) and penicillin/streptomycin stock of antibiotics (Sigma, ref. #P4333-100ML).

S2 and Kc167 cells were grown in an incubator at 25ºC without CO_2. For transient transfection experiments, 6 well plates were used in which 8 x 10^5 cells were placed per well in 2 ml of Express Five serum free medium or Schneider's *Drosophila* Medium, depending on the cell line used, supplemented with L-Glutamine and no antibiotics or serum. 1μg of total DNA was added per well. The following plasmids were used: *pMT*, *pUAST*, *pMT-Gal4*, *pAc-Gal4*, *pUAST-psqB*, *pUAST-HA::SUMO*, *pUAST-HA::SUMOΔGG*, *pMT-psqB^{Tag}*, *pUAST-psqB^{Tag}*, *pUAST-psqBK518R^{Tag}*, *pUAST-psqBK633R^{Tag}*, *pUAST-psqBKKRR^{Tag}*, *pMT-psqBΔ2^{Tag}*, *pMT-psqBΔ3^{Tag}*, *pMT-psqBΔ4^{Tag}*, *pMT-psqBΔ5^{Tag}*, *pMT-psqBΔ6^{Tag}*, *pUAST-psqBΔ2^{Tag}*, *pUAST-psqBΔ3^{Tag}*, *pUAST-psqBΔ4^{Tag}* and *pUAST-psqBΔ5^{Tag}*. The amount of each plasmid was adjusted to obtain equimolar concentrations. Cells were transfected using Cellfectin II Reagent

(Invitrogen, ref. #10362-100), following the manufacturer's instructions. In transfections including plasmids with the metallothionein promoter (*pMT*), activation of the promoter was induced by adding 1.4 mM $CuSO_4$ to the medium 24 hours after transfection. Cells were lysed 24 hours after $CuSO_4$ addition (48 hours after transfection). In transfections using the *pAc-Gal4* plasmid to promote transcription, cells were lysed 24 hours after transfection. In the proteasome chemical inhibition experiments, cells were incubated with Lactacystin (40 μM) (Sigma, ref. #L6785), Z-LLF-CHO (40 μM) (Sigma, ref. #C0358) or MG132 (50 μM) (Calbiochem, ref. #474790) for 6 hours prior to cells lysis. As control for the inhibitors treatment we used water (for Lactacystin), or DMSO (for Z-LFF-CHO and MG132).

5. Co-immunoprecipitation assays

To detect the interaction between Psq and SUMO, S2 cells were grown in 10 cm plates with Express Five serum free medium supplemented with L-Glutamine and no antibiotics. For each condition, one plate with $5x10^6$ cells in 10 ml medium was used. The cells were transiently transfected according to the protocol described above with the following plasmids: *pUAST*, *pAc-Gal4*, *pUAST-HA::SUMO*, *pUAST-HA::SUMOΔGG*, *pUAST-3FLAG::psqB::Tag-RFP-T*, *pUAST-3FLAG::psqBK518R::Tag-RFP-T*, *pUAST-3FLAG::psqBK633R::Tag-RFP-T* or *pUAST-3FLAG::psqBKKRR::Tag-RFP-T*. Cell lysis was performed adapting the protocol described in (Brandl *et al.*, 2012). Cells were collected from the plate using a cell scrapper and centrifuged for 5 minutes at 1100 r.p.m. in a centrifuge cooled to 4ºC to separate them from the medium. Then, 200 μl of denaturing buffer was added to the cell pellet (50 mM Tris-HCl [pH 7.4], 150 mM NaCl, 1% Nonidet P-40, 1% Sodium deoxycholate, 1% SDS, and 1mM EDTA) supplemented with protease inhibitors (2 mM Pefabloc and 1X cOmplete Mini EDTA-free protease inhibitor cocktail) and 80mM of the desumoylation inhibitor N-Ethylmaleimide (Sigma, ref. # E3876-5G). Under these conditions proteins are immediately denatured, blocking the activity of the SUMO proteases that rapidly desumoylate proteins upon cell lysis. As it was already mentioned in the *Introduction*, the attachment of SUMO to its target proteins is through a covalent bond, so the denaturing treatment will not destroy this interaction. After addition of the cell lysis buffer, the samples were sonicated using a Biorruptor sonicator (Diagenode). In order to completely break the cells, especially the nuclear membrane, the samples underwent 7 cycles of 30 seconds ON/OFF at maximum power. After sonication, the

lysates were incubated in ice for 30 minutes and then, 1800 μl of non-denaturing buffer was added to each sample (50 mM Tris-HCl [pH 7.4], 150 mM NaCl, 1% Nonidet P-40, 1% Sodium deoxycholate, and 1mM EDTA) supplemented with protease inhibitors (2 mM Pefabloc and 1X cOmplete Mini EDTA-free protease inhibitor cocktail) and 80mM of N-Ethylmaleimide. The addition of this buffer to the samples restores native conditions (0.1% SDS), which are needed during the immunoprecipitation for the antibodies to maintain their native structures and be functional. The lysates were cleared adding magnetic beads conjugated with Protein A (Millipore, ref. #16-661) and incubated overnight at 4ºC in a rotating shaker (slow rotation). The cleared lysates were incubated with 1μg of the primary antibodies anti-HA (Abcam, ref. #ab9110) or anti-FLAG (Sigma, ref. #F3165), for 3 hours at 4ºC in a rotating shaker (slow rotation). After the incubation, 30 μl of magnetic beads conjugated with Protein A were added to each sample and incubated for 2 hours at 4ºC in a rotating shaker (slow rotation). The samples were then washed three times with native buffer without inhibitors (50 mM Tris-HCl [pH 7.4], 150 mM NaCl, 1% Nonidet P-40, 1% Sodium deoxycholate, 0.1% SDS and 1mM EDTA). The samples were finally resuspended in 50 μl of 3X SDS loading buffer (6X SDS loading buffer diluted by half, see recipe below) with β-mercaptoethanol (1 μl for each 50 μl of 6X SDS buffer) and boiled for 5 minutes at 95ºC. Using a magnet, the magnetic beads were separated from the sample to be analysed by Western Blot as it will be described below.

6. Western blot

To prepare the cell lysates, cells were collected from the plate using a cell scrapper and mechanically lysed in 3X SDS loading buffer (New England Biolabs, ref. #B7709S) or RIPA buffer (50 mM Tris-HCl [pH 8], 150 mM NaCl, 1% Nonidet P-40, 0.5% Sodium deoxycholate, 0.1% SDS), supplemented with protease inhibitors (2 mM Pefabloc and 1X cOmplete Mini EDTA-free protease inhibitor cocktail). The lysates were incubated for 30 minutes at 4ºC. Total cell lysates were stored at -80ºC after the incubation. For experiments in which soluble and insoluble fraction needed to be separated, samples were centrifuged after lysis for 15 minutes at 13000 r.p.m. in a centrifuge cooled to 4ºC. Supernatant and pellet were separated and 300 μl of RIPA buffer were added to the pellet samples. Protein concentration of the samples was determined using BCA Protein Assay Kit (Pierce, ref. #23227). 25 μg of protein sample were resuspended in

6X SDS loading buffer (300 mM Tris-HCl [pH 8.8], 12% SDS, 0.6% bromophenol blue and 30% glycerol) with β-mercaptoethanol (1 μl for each 50 μl of 6X SDS buffer), and boiled for 5 minutes at 95ºC. In the β-mercaptoethanol experiment (**Figure 41**), this component of the buffer was added to the samples as indicated in the *Results* section. Protein samples were separated in 8% SDS-PAGE gels (heat shock experiment) or 10% gels, and transferred to a PVDF membrane (Inmovilon-P Transfer membranes, Millipore, ref. #IPVH00010). Membranes were blocked in PBS with 0.1% Tween-20 and 5% non-fat dry milk for 1 hour at room temperature. After that, membranes were incubated with the primary antibodies: polyclonal rabbit anti-Psq (1:500), designed in the laboratory of María Domínguez and synthetized by SDIX company using SDIX Genomic Antibody Technology®; anti-HA (Abcam, ref. #ab9110, 1:1000), anti-FLAG (Sigma, ref. #F3165, 1:1000), anti-ubiquitin (Abgent, ref. #ab8134-50, 1:1000), anti-tag-RFP-T (Evrogen, ref. #AB234, 1:3000), or anti-actin (Sigma, ref. #A2066, 1:500); all diluted in PBS with 0.1% Tween-20 and 3% non-fat dry milk. After overnight incubation at 4ºC, membranes were incubated during 1 hour at room temperature with secondary antibodies: HRP-conjugated rabbit anti-IgG (Sigma, ref. #A9169, 1:10000) or HRP-conjugated mouse anti-IgG (Jackson, ref. #115-035-062, 1:5000); all diluted in PBS with 0.1% Tween-20 and 3% non-fat dry milk. Proteins were detected using the chemiluminescent substrate ECL (Pierce, ref. #32209), the detector LAS-100 (Fujifilm) and the Image Reader LAS-1000 software (FujiFilm).

7. Immunohistochemistry and microscopy analysis

For immunohistochemistry experiments, imaginal eye-antennal discs or salivary glands from L3 *Drosophila* larvae were dissected in PBS and collected in ice-chilled PBS (for a maximum period of 30 minutes each round of dissection). The tissue was fixed in 4% PFA at room temperature for 30 minutes and then washed three times with PBT (PBS buffer and 0.3% Triton). The blocking incubation was performed at room temperature for 1 hour with PBT-BSA (PBS buffer, 0.3% Triton and 1% BSA). The primary antibodies were diluted in PBT-BSA and incubated overnight at 4ºC: anti-discs large (Developmental Studies Hybridoma Bank [DSHB], ref. #4F3, 1:100), anti-wingless (DSHB, ref. 4D4, 1:200), anti-elav (DSHB, ref. 9F8A9, 1:50) or anti-cleaved caspase 3 (Cell Signalling, ref. 9661L, 1:2000). After three washes with PBT, the tissue was incubated with the secondary antibodies for 1 hour at room temperature: anti-mouse A405 (Molecular Probes, ref. #A31553, 1:200), anti-mouse A488 (Invitrogen, ref.

#A21202, 1:200), anti-rat A488 (Molecular Probes, ref. #A21208, 1:200) or anti-rabbit A647 (Invitrogen, ref. #A31573, 1:200). Finally, the tissue was washed another three times with PBT and a final wash with PBS. DAPI staining was performed before PBS addition. The tissue was incubated with DAPI (Sigma, ref. D9542-10MG, 1500 nM) for 15 minutes at room temperature, and washed again three times with PBT and a final wash with PBS. Fluoromount-G (Southern Biothech, ref. # 0100-01) was used to mount the tissue.

8. RNA extraction, retrotranscription, and quantitative PCR

Total RNA from S2 or Kc167 cells was isolated using the RNeasy Mini Kit (Quiagen, ref. #74106) according to the manufacturer's protocol. RNA samples were treated with DNase (TURBO DNA-free Kit, Applied Biosystems, ref. #AM1907) to eliminate the remaining DNA from the samples, as indicated in the manufacturer's protocol. 1 μg of RNA was reverse transcribed using SuperScript III Reverse Transcriptase (Invitrogen, ref. #18080-093) and Oligo(dT) Primers (Invitrogen, ref. #18418-020). Quantitative PCRs were performed using Power SYBR Green PCR Master Mix (Applied Biosystems, ref. #4367659), 10 ng of template cDNA, and gene-specific primers (222 nM) (**Table 6**), under the following conditions: 10 minutes at 95°C, and then 40 cycles of 15 seconds at 95°C and 40 seconds at 60°C. Real-time PCR reactions were performed using a 7500 Real-Time PCR system (Applied Biosystems), according to the manufacturer's recommendations. The results were normalized to endogenous *Rp49* expression levels. Three separate samples were collected from each condition and triplicate measurements were conducted. Primers were designed using the Primer3 online tool (http://bioinfo.ut.ee/primer3-0.4.0/primer3/). Data are presented as mean ± standard error of the mean; statistical analyses were performed using two-tailed Student's *t*-test.

Primer name	Primer sequence
RP49-1S	5'-TGTCCTTCCAGCTTCAAGATGACCATC-3'
RP49-1AS	5'-CTTGGGCTTGCGCCATTTGTG-3'
CG9925 qPCR sense	5'-TGCCGTCAATGTCACAAAGT-3'
CG9925 qPCR antisense	5'-AGAGCCGCAAAATTCTTTGA-3'
Torso For	5’-CAGTATCACGTTGGCGTAAGA-3’
Torso Rev	5’-GTGGGCAGAGTGTGTAGTTATAG-3’

Table 6. List of primers used for qPCR. The sequences of the primers were designed using the Primer3 online tool.

UNIVERSITAS Miguel Hernández

VI. CONCLUSIONS

UNIVERSITAS Miguel Hernández

UNIVERSITAS
Miguel
Hernández

1. The BTB transcription factor Pipsqueak interacts with a broad array of proteins. Our Y2H assay identifies dozens of new partners and, among them, we have identified three proteins relevant for Pipsqueak function: DmTAF3, a component of basal transcriptional machinery; MEP-1, involved in sumoylation dependent transcriptional repression and Su(var)2-10, *Drosophila*'s homolog of the SUMO ligase PIAS3.

2. DmTAF3 SID with Pipsqueak is located between DmTAF3 aminoacids Gln809 and His878. Our genetic experiments show that DmTAF3 is necessary for Pipsqueak-mediated epigenetic silencing and Pipsqueak-induced oncogenesis *in vivo*.

3. MEP-1 SID with Pipsqueak is located between MEP-1 aminoacids Asn211 and Pro480. Our genetic *in vivo* experiments indicate that MEP-1 restricts Pipsqueak-mediated epigenetic silencing and oncogenic function.

4. Su(var)2-10 SID with Pipsqueak is located between Su(var)2-10 aminoacids Ala362 and Glu602, which includes the domain responsible for the SUMO ligase activity. Our genetic *in vivo* experiments suggest that Su(var)2-10 restricts Pipsqueak-mediated epigenetic silencing and oncogenic function.

5. DmTAF3 and Su(var)2-10 interact with the central region of Pipsqueak, located between aminoacids Ala132 and Gln720, and not with the BTB domain.

6. Pipsqueak is sumoylated at Lysine 633, located in the region that interacts with DmTAF3 and Su(var)2-10.

7. MEP-1 and DmTAF3 have a putative SIM located in the region that interacts with Pipsqueak, defined by *in silico* analysis.

8. Sumoylation attenuates Pipsqueak capacity as transcriptional activator in the context of *CG9925* expression, and decreases Pipsqueak pro-apoptotic activity, which contributes to proper *Drosophila* eye development.

9. Pipsqueak is subjected to two mutually exclusive forms of processing. Type I occurs in the insoluble fraction of the cell and produces two sets of fragments of

~30-50 kDa and ~120-150 kDa. Type II takes place in the soluble fraction and produces two fragments of 70 and 110 kDa.

10. Type I Pipsqueak processing site is located between the BTB and Type II processing site, while Type II Pipsqueak processing site is positioned between Asn415 and Ala636.

11. Type I processing, but not Type II, is required for Pipsqueak-mediated *CG9925* transcriptional derepression. Also, Type II processing is not required for Pipsqueak tumorigenic capacity *in vivo*.

12. The region of Pipsqueak located between aa His241 and Gly428 is necessary for its nuclear localization. Pipsqueak presence in the nucleus is required for *CG9925* derepression, salivary gland phenotype and tumorigenic activity.

13. Proteasome activity is responsible for Pipsqueak Type I processing and is required for *CG9925* transcriptional derepression.

UNIVERSITAS Miguel Hernández

1. El factor de transcripción de tipo BTB Pipsqueak interacciona con una amplia variedad de proteínas. El ensayo de doble híbrido en levaduras realizado en nuestro laboratorio ha permitido la identificación de decenas de nuevas proteínas capaces de interaccionar con Pipsqueak. Entre ellas, hemos identificado tres proteínas relevantes para su función: DmTAF3, un componente de la maquinaria basal de la transcripción; MEP-1, implicado en represión transcripcional dependiente de sumoilización, y Su(var)2-10, el homólogo en *Drosophila* de la SUMO ligasa PIAS3.

2. El dominio mínimo de interacción de DmTAF3 con Pipsqueak se sitúa entre los aminoácidos Gln809 e His878 de DmTAF3. Nuestros experimentos genéticos muestran que DmTAF3 es necesario para el silenciamiento epigenético y la oncogénesis mediados por Pipsqueak *in vivo*.

3. El dominio mínimo de interacción de MEP-1 con Pipsqueak se sitúa entre los aminoácidos Asn211 y Pro480 de MEP-1. Nuestros experimentos genéticos *in vivo* indican que MEP-1 restringe el silenciamiento epigenético y la función tumorogénica mediada por Pipsqueak.

4. El dominio mínimo de interacción de Su(var)2-10 con Pipsqueak se sitúa entre los aminoácidos Ala362 y Glu602 de Su(var)2-10, región en la que se encuentra el dominio responsable de la actividad SUMO ligasa de esta proteína. Nuestros experimentos genéticos *in vivo* sugieren que Su(var)2-10 limita el silenciamiento epigenético y la función oncogénica mediada por Pipsqueak.

5. DmTAF3 y Su(var)2-10 interaccionan con la región central de Pipsqueak, concretamente con la región comprendida entre los aminoácidos Ala132 y Gln720, y no con el dominio BTB.

6. Pipsqueak se sumoiliza en la Lisina 633, localizada en la región de la proteína que interacciona con DmTAF3 y Su(var)2-10.

7. MEP-1 y DmTAF3 tienen un dominio predicho de interacción con SUMO en la región que interacciona con Pipsqueak, definido gracias a análisis *in silico*.

8. La sumoilización de Pipsqueak atenúa su capacidad como activador transcripcional en el contexto de la expresión de *CG9925*, y disminuye su actividad proapoptótica, contribuyendo así al correcto desarrollo del ojo de *Drosophila*.

9. Pipsqueak está sujeto a dos tipos de procesamiento proteolítico diferentes y mutuamente excluyentes. El procesamiento Tipo I ocurre en la fracción insoluble de la célula y produce dos conjuntos de fragmentos de ~30-50 kDa y ~120-150 kDa. El Tipo II ocurre en la fracción soluble y produce dos fragmentos de 70 y 110 kDa.

10. El sitio de corte del procesamiento Tipo I de Pipsqueak se sitúa entre el dominio BTB y el sitio de corte del procesamiento Tipo II, mientras que el sitio de corte del procesamiento Tipo II se localiza entre los aminoácidos Asn415 y Ala636.

11. El procesamiento Tipo I de Pipsqueak, pero no el Tipo II, se requiere para la desrepresión transcripcional de *CG9925* mediada por Pipsqueak. Además, el procesamiento Tipo II no se requiere para la capacidad tumorogénica de Pipsqueak *in vivo*.

12. La región de Pipsqueak comprendida entre los aminoácidos His241 y Gly428 es necesaria para su localización nuclear. La presencia de Pipsqueak en el núcleo es requerida para la desrepresión de *CG9925*, el fenotipo observado en las glándulas salivares y su actividad tumorogénica.

13. La actividad del proteasoma es responsable del procesamiento de Pipsqueak Tipo I y se requiere para la desrepresión transcripcional de *CG9925*.

APPENDICES

UNIVERSITAS Miguel Hernández

UNIVERSITAS
Miguel
Hernández

Appendix I. List of mutant and transgenic fly lines

Stock	Description	Origin
w^{1118}	Strain with the w^{1118} mutation in the gene *white*.	-
ey-Gal4	Expresses Gal4 under the *eyeless* promoter	María Domínguez's Lab.
ey-Gal4 UAS-Dl	Expresses Gal4 and the Notch ligand *Dl* under the *eyeless* promoter.	María Domínguez's Lab.
ey-Gal4 GS88A8	Expresses Gal4 and the GS insertion 88A8 (inserted between genes *psq* and *lola*), under the *eyeless* promoter.	María Domínguez's Lab.
ey-Gal4 UAS-Dl GS88A8	Expresses Gal4, the Notch ligand *Dl* and the GS insertion 88A8 (inserted between genes *psq* and *lola*), under the *eyeless* promoter.	María Domínguez's Lab.
hs-Gal4	Expresses Gal4 under the heat shock inducible promoter *hsp70.*	BDSC (#1799)
UAS-psqB	Expresses *psq* isoform B*.*	María Domínguez's Lab.
ey-Gal4 miRNA-278;UAS-Dl	Expresses Gal4, the Notch ligand *Dl* and the GS insertion 63F1 under the *eyeless* promoter.	María Domínguez's Lab.
UASp-Myc::TAF3-Dmm1	Expression of *Myc-TAF3-Dmm1*	Daniel Pauli's Lab.
UASp-Myc::TAF3-Bfm1	Expression of *Myc-TAF3-Bfm1*.	Daniel Pauli's Lab.
$TAF3^{Fa4a}$	Loss of function allele.	Daniel Pauli's Lab.
UAS-RNAi DmTAF3	Expresses RNAi against *DmTAF3.* TRIP line.	BDSC (#43174)
UAS-RNAi DmTAF3	Expresses RNAi against *DmTAF3.* GD line.	VDRC (#48036)
UAS-RNAi DmTAF3	Expresses RNAi against *DmTAF3.* GD line.	VDRC (#48037)
UAS-RNAi DmTAF3	Expresses RNAi against *DmTAF3.* KK line.	VDRC (#107591)
MEP-1 GS9354	GS insertion in MEP-1, in position 62F2.	DGRC (#201714)
MEP-1 GS16341	GS insertion in MEP-1, in position 62E8.	DGRC (#206592)

Stock	Description	Origin
MEP-1 GS20908	GS insertion in MEP-1, in position 62E8.	DGRC (#202015)
MEP-1 GS12243	GS insertion in MEP-1, in position 62F2.	DGRC (#203932)
UAS-RNAi MEP-1	Expresses RNAi against *MEP-1.* GD line.	VDRC (#24533)
UAS-RNAi MEP-1	Expresses RNAi against *MEP-1.* GD line.	VDRC (#24534)
UAS-RNAi MEP-1	Expresses RNAi against *MEP-1.* TRIP line.	BDSC (#35399)
UAS-RNAi MEP-1	Expresses RNAi against *MEP-1.* TRIP line.	BDSC (#33676)
UAS-PIAS::GFP26b.3	Expression of *dPIAS::GFP*.	-
Su(var)2-10^{03697}	P{PZ} insertion into *Su(var)2-10.*	BDSC (#11344)
Su(var)2-10^{1}	Ethyl methanesulfonate mutant allele. Causes aa replacement L327M.	BDSC (#6236)
Su(var)2-10^{2}	Ethyl methanesulfonate mutant allele. Causes aa replacement W260 to STOP.	BDSC (#6235)
UAS-RNAi Su(var)2-10	Expresses RNAi against *Su(var)2-10*. GD line.	VDRC (#29448)
smt3^{04493}	Lethal allele. LacZ enhancer trap inserted in *smt3*.	BDSC (#11378)
smt3^{k06307}	Lethal allele. LacZ activity.	BDSC (#10419)
lwr^{05486}	Loss of function allele. Imprecise excision of the P{PZ} element. 34bp of P-element sequences remain, together with 42bp of inserted material, repeats of TAACA.	BDSC (#11410)
lwr$^{4-3}$	Loss of function allele. Imprecise excision of the *P{PZ}* element. 34bp of P-element sequences remain, together with 42bp of inserted material, repeats of TAACA.	BDSC (#9321)
lwr^{13}	Loss of function allele. Imprecise excision: results in a 1603bp deletion from 18bp upstream of the 5' end of the *lwr* gene to 299bp downstream of the poly(A) site of *Spp.*	BDSC (#9323)
lwr^{5}	Loss of function allele. Ethylmethanesulfonate allele. Causes aa replacement R104H.	BDSC (#9317)

Stock	Description	Origin
UAS-3xFLAG::psqB::Tag-RFP-T (#6753-3-3)	P-element insertion in chromosome II.	María Domínguez's Lab.
UAS-3xFLAG::psqBK518R::Tag-RFP-T (#7087-6-1)	P-element insertion in chromosome III.	María Domínguez's Lab.
UAS-3xFLAG::psqBK633R::Tag-RFP-T (#7179-1-2)	P-element insertion in chromosome III.	María Domínguez's Lab.
UAS-3xFLAG::psqBKKRR::Tag-RFP-T(#7087-5-3)	P-element insertion in chromosome III.	María Domínguez's Lab.
UAS-3xFLAG::psqBΔ2::Tag-RFP-T (9 transgenic lines: 13443-1-1 to 13443-1-9)	P-element insertion in chromosomes II and III.	María Domínguez's Lab.
UAS-3xFLAG::psqBΔ3::Tag-RFP-T (9 transgenic lines: 13443-2-1 to 13443-2-9)	P-element insertion in chromosomes II and III.	María Domínguez's Lab.
UAS-3xFLAG::psqBΔ4::Tag-RFP-T (8 transgenic lines: 9832-1-1 to 9832-1-8)	P-element insertion in chromosomes II and III.	María Domínguez's Lab.
UAS-3xFLAG::psqBΔ5::Tag-RFP-T (5 transgenic lines: 9832-2-1 to 9832-2-5)	P-element insertion in chromosomes II and III.	María Domínguez's Lab.

UNIVERSITAS Miguel Hernández

Appendix II. List of positive clones obtained in the yeast two-hybrid

The full list of the 154 positive clones identified in the Y2H is presented in the following tables, as well as a table with a description of the PBS categories used to assess the interaction reliability of the interactions detected. Clones identifying the same gene are clustered together.

Global PBS (for Interactions represented in the Screen)		Nb	%
A	Very high confidence in the interaction	5	23.8%
B	High confidence in the interaction	3	14.3%
C	Good confidence in the interaction	1	4.8%
D	Moderate confidence in the interaction This category is the most difficult to interpret because it mixes two classes of interactions : - False-positive interactions - Interactions hardly detectable by the Y2H technique (low representation of the mRNA in the library, prey folding, prey toxicity in yeast)	11	52.4%
E	Interactions involving highly connected prey domains, warning of non-specific interaction. The threshold for high connectivity is 10 for screens with Human, Mouse, Drosophila and Arabidopsis and 6 for all other organisms. They can be classified in different categories: - Prey proteins that are known to be highly connected due to their biological function - Proteins with a prey interacting domain that contains a known protein interaction motif or a biochemically promiscuous motif	1	4.8%
F	Experimentally proven technical artifacts	0	0.0%
Non Appliable			
N/A	The PBS is a score that is automatically computed through algorithms and cannot be attributed for the following reasons : - All the fragments of the same reference CDS are antisens - The 5p sequence is missing - All the fragments of the same reference CDS are either all OOF1 or all OOF2 - All the fragments of the same reference CDS lie in the 5' or 3' UTR		

Clone Name	Type Seq	Gene Name (Best Match)	Start..Stop (nt)	Frame	Sens	%Id 5p	%Id 3p	PBS
pB29_A-18	5p/3p	Drosophila melanogaster - Arc70	366..2837	IF		99.0	95.0	B
pB29_A-115	5p	Drosophila melanogaster - Arc70	366	IF		93.6		B
pB29_A-11	5p	Drosophila melanogaster - Arc70	387	IF		96.8		B
pB29_A-82	5p/3p	Drosophila melanogaster - CG10042-PA [Drosophilamelanogaster]	2236..1668	??	N	87.5	89.9	N/A
pB29_A-113	5p/3p	Drosophila melanogaster - CG17090	2256..3068	IF		97.2	91.2	D
pB29_A-61	5p	Drosophila melanogaster - CG2926-PA [Drosophilamelanogaster]	4944..5451	IF		86.9		C
pB29_A-91	5p/3p	Drosophila melanogaster - CG2926-PA [Drosophilamelanogaster]	5103..6121	IF		91.7	79.7	C
pB29_A-87	5p	Drosophila melanogaster - CG7185	2141	??	N	84.6		N/A
pB29_A-121	5p	Drosophila melanogaster - CG7222-PA	764	??	N	68.9		N/A
pB29_A-106	5p	Drosophila melanogaster - Chro	1485	IF		84.4		D
pB29_A-130	5p/3p	Drosophila melanogaster - EDTP	507..-275	??	N	99.4	98.7	N/A
pB29_A-28	5p/3p	Drosophila melanogaster - Eb1	898..405	??	N	94.8	89.6	N/A
pB29_A-15	5p/3p	Drosophila melanogaster - Eno	136..1183	OOF1		93.4	86.0	N/A
pB29_A-152	5p	Drosophila melanogaster - Fer1HCH	-79	IF		85.2		D
pB29_A-2	3p	Drosophila melanogaster - MICAL	..5826	??			98.7	A
pB29_A-19	5p/3p	Drosophila melanogaster - MICAL	4326..5061	IF		96.0	89.8	A
pB29_A-85	5p	Drosophila melanogaster - MICAL	4341	IF		89.8		A
pB29_A-37	5p/3p	Drosophila melanogaster - MICAL	4341..5503	IF		99.3	93.2	A
pB29_A-124	5p/3p	Drosophila melanogaster - MICAL	4341..5503	IF		96.0	94.6	A
pB29_A-105	5p	Drosophila melanogaster - MICAL	4365	IF		94.3		A
pB29_A-146	5p	Drosophila melanogaster - MICAL	4797	IF		96.8		A
pB29_A-78	5p/3p	Drosophila melanogaster - MICAL	4797..5529	IF		87.8	93.9	A
pB29_A-58	5p/3p	Drosophila melanogaster - MICAL	4797..5529	IF		91.3	91.5	A
pB29_A-89	5p	Drosophila melanogaster - MICAL	4797	IF		86.2		A
pB29_A-142	3p	Drosophila melanogaster - PR2	..3987	??			95.0	N/A
pB29_A-24	5p/3p	Drosophila melanogaster - Rbp2	1898..1548	??	N	86.3	88.2	N/A

Clone Name	Type Seq	Gene Name (Best Match)	Start..Stop (nt)	Frame	Sens	%Id 5p	%Id 3p	PBS
pB29_A-186	5p	Drosophila melanogaster - RpS3A	755	??	N	77.9		N/A
pB29_A-86	5p	Drosophila melanogaster - Su(var)2-10	1086	IF		89.7		B
pB29_A-75	5p/3p	Drosophila melanogaster - Su(var)2-10	1086..1805	IF		87.4	77.0	B
pB29_A-170	5p	Drosophila melanogaster - Su(var)2-10	1086	IF		98.3		B
pB29_A-109	5p	Drosophila melanogaster - Su(var)2-10	1224	IF		96.5		B
pB29_A-41	5p/3p	Drosophila melanogaster - Taf8	-41..836	OOF2		97.2	84.5	N/A
pB29_A-21	5p/3p	Drosophila melanogaster - Xbp1	484..1034 ✕	OOF1		89.1	93.6	N/A
pB29_A-161	5p	Drosophila melanogaster - abs	738	IF		96.4		B
pB29_A-126	5p/3p	Drosophila melanogaster - abs	738..1874 ✕	IF		90.2	79.9	B
pB29_A-165	5p/3p	Drosophila melanogaster - abs	738..1874 ✕	IF		93.5	94.0	B
pB29_A-20	5p/3p	Drosophila melanogaster - abs	739..1874 ✕	OOF1		92.7	100.0	B
pB29_A-25	5p/3p	Drosophila melanogaster - bip2	957..3046	IF		87.4	98.0	A
pB29_A-168	5p	Drosophila melanogaster - bip2	1515	IF		99.7		A
pB29_A-73	5p	Drosophila melanogaster - bip2	1515	IF		99.7		A
pB29_A-127	5p/3p	Drosophila melanogaster - bip2	1782..3053	IF		90.9	89.3	A
pB29_A-163	5p	Drosophila melanogaster - bip2	1818	IF		77.7		A
pB29_A-62	5p/3p	Drosophila melanogaster - bip2	1920..3198	IF		98.3	86.7	A
pB29_A-98	5p/3p	Drosophila melanogaster - bip2	1920..3198	IF		91.8	75.5	A
pB29_A-80	5p	Drosophila melanogaster - bip2	1920	IF		96.6		A
pB29_A-131	5p	Drosophila melanogaster - bip2	1920	IF		96.5		A
pB29_A-38	5p/3p	Drosophila melanogaster - bip2	1923..2634	IF		93.0	86.1	A
pB29_A-185	5p	Drosophila melanogaster - bip2	1941	IF		87.4		A
pB29_A-77	5p/3p	Drosophila melanogaster - bip2	1941..3192	IF		93.1	82.8	A
pB29_A-64	5p	Drosophila melanogaster - bip2	2262	IF		90.3		A
pB29_A-30	5p/3p	Drosophila melanogaster - bip2	2262..3945	IF		97.2	86.5	A
pB29_A-39	5p/3p	Drosophila melanogaster - bip2	2421..4209	IF		98.6	76.8	A
pB29_A-40	5p/3p	Drosophila melanogaster - bip2	2424..3252	IF		87.4	77.7	A
pB29_A-53	5p/3p	Drosophila melanogaster - bip2	2424..3252	IF		91.2	85.2	A
pB29_A-16	5p/3p	Drosophila melanogaster - bip2	2424..3252	IF		83.2	84.7	A
pB29_A-35	5p/3p	Drosophila melanogaster - bip2	2424..3252	IF		95.9	91.6	A
pB29_A-84	5p	Drosophila melanogaster - bip2	2448	IF		99.5		A
pB29_A-60	5p	Drosophila melanogaster - bip2	2520	IF		95.4		A
pB29_A-76	5p	Drosophila melanogaster - bip2	2520	IF		81.6		A
pB29_A-166	5p/3p	Drosophila melanogaster - bol	636..-136	??	N	89.1	82.3	N/A
pB29_A-34	5p/3p	Drosophila melanogaster - crm	267..683	IF		97.6	97.2	D
pB29_A-160	5p	Drosophila melanogaster - dCG6923-PA [Drosophilamelanogaster]	735	IF		90.5		D
pB29_A-176	5p	Drosophila melanogaster - ed	3172	??	N	86.0		N/A
pB29_A-79	5p	Drosophila melanogaster - hdc	2301	??	N	71.9		N/A
pB29_A-104	5p/3p	Drosophila melanogaster - lolal	-95..561 ✻ ✕	OOF2		89.9	82.6	N/A
pB29_A-42	5p/3p	Drosophila melanogaster - lolal	-95..486 ✻ ✕	OOF2		90.5	79.4	N/A
pB29_A-172	5p/3p	Drosophila melanogaster - lolal	-80..506 ✻ ✕	OOF2		97.3	90.6	N/A
pB29_A-66	5p/3p	Drosophila melanogaster - lolal	-77..865 ✻ ✕	OOF2		96.5	72.5	N/A
pB29_A-139	5p	Drosophila melanogaster - lolal	-71..571 ✻ ✕	OOF2		91.9		N/A
pB29_A-81	5p/3p	Drosophila melanogaster - lolal	-71..573 ✻ ✕	OOF2		90.4	88.9	N/A
pB29_A-156	5p	Drosophila melanogaster - lolal	-71	OOF2		95.5		N/A
pB29_A-43	5p/3p	Drosophila melanogaster - lolal	-71..573 ✻ ✕	OOF2		95.4	85.1	N/A
pB29_A-112	5p/3p	Drosophila melanogaster - lolal	-71..1016 ✻ ✕	OOF2		98.0	99.6	N/A

Clone Name	Type Seq	Gene Name (Best Match)	Start..Stop (nt)	Frame	Sens	%Id 5p	%Id 3p	PBS
pB29_A-4	5p/3p	Drosophila melanogaster - lolal	-65	OOF2		90.3	74.3	N/A
pB29_A-96	5p	Drosophila melanogaster - lolal	-56	OOF2		89.8		N/A
pB29_A-188	5p/3p	Drosophila melanogaster - mam	1989..1061	??	N	94.4	96.9	N/A
pB29_A-46	5p/3p	Drosophila melanogaster - ovo	2331..1603	??	N	83.1	80.0	N/A
pB29_A-8	5p	Drosophila melanogaster - psq	-23	OOF2		96.1		A
pB29_A-90	5p	Drosophila melanogaster - psq	-15	OOF1		82.5		A
pB29_A-6	5p/3p	Drosophila melanogaster - rdx	-21..373	OOF1		92.2	96.5	N/A
pB29_A-33	5p/3p	Drosophila melanogaster - ry	3461..3071	??	N	96.9	80.9	N/A
pB29_A-181	5p	Drosophila melanogaster - sca	2431	??	N	79.7		N/A
pB29_A-29	3p	Drosophila melanogaster - sta	..252	??	N		91.7	N/A
pB29_A-175	5p	Drosophila melanogaster - tou	1854	IF		94.0		E
pB29_A-71	5p/3p	Drosophila melanogaster - tou	1857..2961	IF		85.8	98.3	E
pB29_A-17	5p/3p	Drosophila melanogaster - tou	1857..2961	IF		95.3	95.5	E
pB29_A-97	5p/3p	Drosophila melanogaster - tou	2052..3135	IF		93.6	88.9	E
pB29_A-114	5p	Drosophila melanogaster - tou	2052	IF		95.5		E
pB29_A-65	5p	Drosophila melanogaster - tou	2130	IF		91.9		E
pB29_A-70	5p	Drosophila melanogaster - tou	2136	IF		80.4		E
pB29_A-14	5p	Drosophila melanogaster - tou	2136	IF		80.6		E
pB29_A-184	5p	Drosophila melanogaster - tou	2136	IF		77.8		E
pB29_A-119	5p/3p	Drosophila melanogaster - tou	2136..2699	IF		99.1	78.7	E
pB29_A-110	5p	Drosophila melanogaster - tou	2136..2701	IF		96.4		E
pB29_A-47	5p/3p	Drosophila melanogaster - tou	2136..2699	IF		93.2	90.8	E
pB29_A-88	5p	Drosophila melanogaster - tou	2136	IF		80.9		E
pB29_A-144	5p	Drosophila melanogaster - tou	2136	IF		99.2		E
pB29_A-49	5p	Drosophila melanogaster - tou	2136	IF		99.2		E
pB29_A-55	5p/3p	Drosophila melanogaster - tou	2157..3132	IF		85.3	95.3	E
pB29_A-99	5p/3p	Drosophila melanogaster - tou	2157..3132	IF		84.1	87.6	E
pB29_A-32	5p	Drosophila melanogaster - tou	2157	IF		90.6		E
pB29_A-48	5p	Drosophila melanogaster - tou	2157	IF		99.8		E
pB29_A-103	5p	Drosophila melanogaster - tou	2157	IF		100.0		E
pB29_A-182	5p/3p	Drosophila melanogaster - tou	2184..2985	IF		88.8	83.7	E
pB29_A-167	5p	Drosophila melanogaster - tou	2184	IF		87.4		E
pB29_A-155	5p	Drosophila melanogaster - tou	2205..2722	IF		94.8		E
pB29_A-143	5p	Drosophila melanogaster - tou	2205	IF		93.1		E
pB29_A-83	5p/3p	Drosophila melanogaster - tou	2205..2720	IF		96.1	99.8	E
pB29_A-187	5p/3p	Drosophila melanogaster - tou	2205..2720	IF		94.6	85.7	E
pB29_A-27	5p/3p	Drosophila melanogaster - tou	2205..2720	IF		92.6	95.7	E
pB29_A-164	5p	Drosophila melanogaster - tou	2205	IF		84.2		E
pB29_A-74	5p	Drosophila melanogaster - tou	2205	IF		90.8		E
pB29_A-56	5p	Drosophila melanogaster - tou	2205	IF		88.3		E
pB29_A-94	5p/3p	Drosophila melanogaster - CG1472	1677..2250	IF		88.6	94.2	D
pB29_A-177	5p/3p	Drosophila melanogaster - CG14438	7048..5754	??	N	96.8	97.7	N/A
pB29_A-95	3p	Drosophila melanogaster - CG1244	..3410	??			96.4	A
pB29_A-134	5p	Drosophila melanogaster - CG1244	630	IF		95.2		A
pB29_A-92	5p/3p	Drosophila melanogaster - CG1244	630..1439	IF		94.1	93.5	A
pB29_A-3	5p	Drosophila melanogaster - CG1244	2346	IF		97.0		A
pB29_A-133	5p	Drosophila melanogaster - CG1244	2472	IF		93.0		A
pB29_A-189	5p/3p	Drosophila melanogaster - CG1244	2502..3859	IF		96.1	83.2	A
pB29_A-137	5p	Drosophila melanogaster - CG1244	2502	IF		95.3		A
pB29_A-93	5p/3p	Drosophila melanogaster - CG1244	2547..3247	IF		91.5	88.8	A
pB29_A-140	5p	Drosophila melanogaster - CG1244	2547	IF		91.5		A

Clone Name	Type Seq	Gene Name (Best Match)	Start..Stop (nt)	Frame	Sens	%Id 5p	%Id 3p	PBS
pB29_A-45	5p/3p	Drosophila melanogaster - CG1244	2547..3247	IF		91.5	98.1	A
pB29_A-101	5p/3p	Drosophila melanogaster - CG1244	2577..3550 ✕	IF		87.5	94.5	A
pB29_A-51	5p	Drosophila melanogaster - CG1244	2577	IF		92.8		A
pB29_A-5	5p/3p	Drosophila melanogaster - CG1244	2577..3550 ✕	IF		94.7	99.8	A
pB29_A-138	5p	Drosophila melanogaster - CG1244	2601	IF		94.0		A
pB29_A-107	5p	Drosophila melanogaster - CG1244	2601	IF		96.0		A
pB29_A-129	5p/3p	Drosophila melanogaster - CG1244	2724..3686 ✕	IF		97.5	94.4	A
pB29_A-13	5p	Drosophila melanogaster - CG32046	3667	??	N	93.3		N/A
pB29_A-67	5p	Drosophila melanogaster - CG3252	1938 ↷	??	N	97.4		N/A
pB29_A-162	5p	Drosophila melanogaster - CG14991	271	??	N	97.7		N/A
pB29_A-154	5p/3p	Drosophila melanogaster - CG5708	93..800 ✕	IF		88.7	88.1	D
pB29_A-31	3p	Drosophila melanogaster - CG10712	..1915	??			97.1	A
pB29_A-111	5p	Drosophila melanogaster - CG10712	1266	IF		97.0		A
pB29_A-183	5p	Drosophila melanogaster - CG10712	1266	IF		83.7		A
pB29_A-141	5p	Drosophila melanogaster - CG10712	1266	IF		93.3		A
pB29_A-44	5p	Drosophila melanogaster - CG10712	1266	IF		93.8		A
pB29_A-52	5p/3p	Drosophila melanogaster - CG10712	1284..1929	IF		90.6	88.9	A
pB29_A-120	5p	Drosophila melanogaster - CG10712	1404	IF		87.5		A
pB29_A-69	5p	Drosophila melanogaster - CG10712	1428	IF		85.5		A
pB29_A-12	5p/3p	Drosophila melanogaster - CG10712	1428..2102	IF		99.0	89.6	A
pB29_A-63	5p	Drosophila melanogaster - CG12505	1625..1946 ↷ ✕	OOF2		97.5		N/A
pB29_A-153	3p	Drosophila melanogaster - CG8332	..-15 ↶	??	N		88.7	N/A
pB29_A-158	5p	Drosophila melanogaster - CG17295	-18..220	OOF1		99.2		N/A
pB29_A-148	5p	Drosophila melanogaster - CG12207	614	??	N	92.4		N/A
pB29_A-128	5p	Drosophila melanogaster - CG17612	-30	OOF1		95.8		N/A
pB29_A-151	5p	Drosophila melanogaster - CG4019	75..374	IF		98.0		D
pB29_A-26	5p/3p	Drosophila melanogaster - GenMatch GID: 8573373	-1..943 ✻ ✕	IF		95.8	99.2	D
pB29_A-132	5p	Drosophila melanogaster - GenMatch GID: 113194865	-1	IF		100.0		D
pB29_A-100	5p/3p	Drosophila melanogaster - GenMatch GID: 210062745	-1..342 ✕	IF		96.9	90.1	D
pB29_A-50	3p	Drosophila melanogaster - GenMatch GID: 113194556	..413	??			100.0	N/A

REFERENCES

UNIVERSITAS Miguel Hernández

UNIVERSITAS
Miguel
Hernández

Abdel-Hafiz H., Takimoto G. S., Tung L., Horwitz K. B. (2002). The inhibitory function in human progesterone receptor N termini binds SUMO-1 protein to regulate autoinhibition and transrepression. *J Biol Chem* 277: 33950-33956.

Adams J. (2003). The proteasome: structure, function, and role in the cell. *Cancer Treat Rev* 29 Suppl 1: 3-9.

Adams J. (2004). The proteasome: a suitable antineoplastic target. *Nat Rev Cancer* 4: 349-360.

Ahmad K. F., Engel C. K., Prive G. G. (1998). Crystal structure of the BTB domain from PLZF. *Proc Natl Acad Sci U S A* 95: 12123-12128.

Ahmad K. F., Melnick A., Lax S., Bouchard D., Liu J., Kiang C. L., Mayer S., Takahashi S., Licht J. D., Prive G. G. (2003). Mechanism of SMRT corepressor recruitment by the BCL6 BTB domain. *Mol Cell* 12: 1551-1564.

Albagli-Curiel O. (2003). Ambivalent role of BCL6 in cell survival and transformation. *Oncogene* 22: 507-516.

Albagli O., Dhordain P., Deweindt C., Lecocq G., Leprince D. (1995). The BTB/POZ domain: a new protein-protein interaction motif common to DNA- and actin-binding proteins. *Cell Growth Differ* 6: 1193-1198.

Alkalay I., Yaron A., Hatzubai A., Orian A., Ciechanover A., Ben-Neriah Y. (1995). Stimulation-dependent I kappa B alpha phosphorylation marks the NF-kappa B inhibitor for degradation via the ubiquitin-proteasome pathway. *Proc Natl Acad Sci U S A* 92: 10599-10603.

Alkuraya F. S., Saadi I., Lund J. J., Turbe-Doan A., Morton C. C., Maas R. L. (2006). SUMO1 haploinsufficiency leads to cleft lip and palate. *Science* 313: 1751.

Altschul S. F., Madden T. L., Schaffer A. A., Zhang J., Zhang Z., Miller W., Lipman D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-3402.

Andrews E. A., Palecek J., Sergeant J., Taylor E., Lehmann A. R., Watts F. Z. (2005). Nse2, a component of the Smc5-6 complex, is a SUMO ligase required for the response to DNA damage. *Mol Cell Biol* 25: 185-196.

Apionishev S., Malhotra D., Raghavachari S., Tanda S., Rasooly R. S. (2001). The Drosophila UBC9 homologue lesswright mediates the disjunction of homologues in meiosis I. *Genes Cells* 6: 215-224.

Aravind L., Koonin E. V. (2000). SAP - a putative DNA-binding motif involved in chromosomal organization. *Trends Biochem Sci* 25: 112-114.

Aravind L., Landsman D. (1998). AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucleic Acids Res* 26: 4413-4421.

Armstrong A. A., Mohideen F., Lima C. D. (2012). Recognition of SUMO-modified PCNA requires tandem receptor motifs in Srs2. *Nature* 483: 59-63.

Artavanis-Tsakonas S., Rand M. D., Lake R. J. (1999). Notch signaling: cell fate control and signal integration in development. *Science* 284: 770-776.

Aster J. C. (2005). Deregulated NOTCH signaling in acute T-cell lymphoblastic leukaemia/lymphoma: new insights, questions, and opportunities. *Int J Hematol* 82: 295-301.

Aster J. C., Pear W. S., Blacklow S. C. (2008). Notch signaling in leukaemia. *Annu Rev Pathol* 3: 587-613.

Aza-Blanc P., Ramirez-Weber F. A., Laget M. P., Schwartz C., Kornberg T. B. (1997). Proteolysis that is inhibited by hedgehog targets Cubitus interruptus protein to the nucleus and converts it to a repressor. *Cell* 89: 1043-1053.

Azuma Y., Tan S. H., Cavenagh M. M., Ainsztein A. M., Saitoh H., Dasso M. (2001). Expression and regulation of the mammalian SUMO-1 E1 enzyme. *FASEB J* 15: 1825-1827.

Baba D., Maita N., Jee J. G., Uchimura Y., Saitoh H., Sugasawa K., Hanaoka F., Tochio H., Hiroaki H., Shirakawa M. (2005). Crystal structure of thymine DNA glycosylase conjugated to SUMO-1. *Nature* 435: 979-982.

Bachant J., Alcasabas A., Blat Y., Kleckner N., Elledge S. J. (2002). The SUMO-1 isopeptidase Smt4 is linked to centromeric cohesion through SUMO-1 modification of DNA topoisomerase II. *Mol Cell* 9: 1169-1182.

Bailey D., O'Hare P. (2004). Characterization of the localization and proteolytic activity of the SUMO-specific protease, SENP1. *J Biol Chem* 279: 692-703.

Baker S. P., Grant P. A. (2005). The proteasome: not just degrading anymore. *Cell* 123: 361-363.

Bardwell V. J., Treisman R. (1994). The POZ domain: a conserved protein-protein interaction motif. *Genes Dev* 8: 1664-1677.

Bartel P. L., Chien C.-T., Sternglanz R., Fields S. (1993) Using the two- hybrid system to detect protein-protein interactions. In: Cellular Interactions in Development: A Practical Approach, pp 153–179. Oxford Oxford University Press.

Barton M. K., Schedl T. B., Kimble J. (1987). Gain-of-function mutations of fem-3, a sex-determination gene in Caenorhabditis elegans. *Genetics* 115: 107-119.

Belfiore M., Mathies L. D., Pugnale P., Moulder G., Barstead R., Kimble J., Puoti A. (2002). The MEP-1 zinc-finger protein acts with MOG DEAH box proteins to control gene expression via the fem-3 3' untranslated region in Caenorhabditis elegans. *RNA* 8: 725-739.

Beliakoff J., Sun Z. (2006). Zimp7 and Zimp10, two novel PIAS-like proteins, function as androgen receptor coregulators. *Nucl Recept Signal* 4: e017.

Benson M. D., Li Q. J., Kieckhafer K., Dudek D., Whorton M. R., Sunahara R. K., Iniguez-Lluhi J. A., Martens J. R. (2007). SUMO modification regulates inactivation of the voltage-gated potassium channel Kv1.5. *Proc Natl Acad Sci U S A* 104: 1805-1810.

Bereczki O., Ujfaludi Z., Pardi N., Nagy Z., Tora L., Boros I. M., Balint E. (2008). TATA binding protein associated factor 3 (TAF3) interacts with p53 and inhibits its function. *BMC Mol Biol* 9: 57.

Bergink S., Jentsch S. (2009). Principles of ubiquitin and SUMO modifications in DNA repair. *Nature* 458: 461-467.

Bernier-Villamor V., Sampson D. A., Matunis M. J., Lima C. D. (2002). Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. *Cell* 108: 345-356.

Bernstein B. E., Mikkelsen T. S., Xie X., Kamal M., Huebert D. J., Cuff J., Fry B., Meissner A., Wernig M., Plath K., Jaenisch R., Wagschal A., Feil R., Schreiber S. L., Lander E. S. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125: 315-326.

Best J. L., Ganiatsas S., Agarwal S., Changou A., Salomoni P., Shirihai O., Meluh P. B., Pandolfi P. P., Zon L. I. (2002). SUMO-1 protease-1 regulates gene transcription through PML. *Mol Cell* 10: 843-855.

Bettermann K., Benesch M., Weis S., Haybaeck J. (2012). SUMOylation in carcinogenesis. *Cancer Lett* 316: 113-125.

Bier E. (2005). Drosophila, the golden bug, emerges as a tool for human genetics. *Nat Rev Genet* 6: 9-23.

Bies J., Markus J., Wolff L. (2002). Covalent attachment of the SUMO-1 protein to the negative regulatory domain of the c-Myb transcription factor modifies its stability and transactivation capacity. *J Biol Chem* 277: 8999-9009.

Birck C., Poch O., Romier C., Ruff M., Mengus G., Lavigne A. C., Davidson I., Moras D. (1998). Human TAF(II)28 and TAF(II)18 interact through a histone fold encoded by atypical evolutionary conserved motifs also found in the SPT3 family. *Cell* 94: 239-249.

Bischof O., Schwamborn K., Martin N., Werner A., Sustmann C., Grosschedl R., Dejean A. (2006). The E3 SUMO ligase PIASy is a regulator of cellular senescence and apoptosis. *Mol Cell* 22: 783-794.

Bossis G., Melchior F. (2006). SUMO: regulating the regulator. *Cell Div* 1: 13.

Brand A. H., Perrimon N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118: 401-415.

Brandl A., Wagner T., Uhlig K. M., Knauer S. K., Stauber R. H., Melchior F., Schneider G., Heinzel T., Kramer O. H. (2012). Dynamically regulated sumoylation of HDAC2 controls p53 deacetylation and restricts apoptosis following genotoxic stress. *J Mol Cell Biol* 4: 284-293.

Bray S. J. (2006). Notch signalling: a simple pathway becomes complex. *Nat Rev Mol Cell Biol* 7: 678-689.

Brou C., Logeat F., Gupta N., Bessia C., LeBail O., Doedens J. R., Cumano A., Roux P., Black R. A., Israel A. (2000). A novel proteolytic cleavage involved in Notch signaling: the role of the disintegrin-metalloprotease TACE. *Mol Cell* 5: 207-216.

Buschmann T., Fuchs S. Y., Lee C. G., Pan Z. Q., Ronai Z. (2000). SUMO-1 modification of Mdm2 prevents its self-ubiquitination and increases Mdm2 ability to ubiquitinate p53. *Cell* 101: 753-762.

Carter S., Bischof O., Dejean A., Vousden K. H. (2007). C-terminal modifications regulate MDM2 dissociation and nuclear export of p53. *Nat Cell Biol* 9: 428-435.

Cavodeassi F., Diez Del Corral R., Campuzano S., Dominguez M. (1999). Compartments and organising boundaries in the Drosophila eye: the role of the homeodomain Iroquois proteins. *Development* 126: 4933-4942.

Ciechanover A., Iwai K. (2004). The ubiquitin system: from basic mechanisms to the patient bed. *IUBMB Life* 56: 193-201.

Colland F., Jacq X., Trouplin V., Mougin C., Groizeleau C., Hamburger A., Meil A., Wojcik J., Legrain P., Gauthier J. M. (2004). Functional proteomics mapping of a human signaling pathway. *Genome Res* 14: 1324-1332.

Croston G. E., Kerrigan L. A., Lira L. M., Marshak D. R., Kadonaga J. T. (1991). Sequence-specific antirepression of histone H1-mediated inhibition of basal RNA polymerase II transcription. *Science* 251: 643-649.

Chen B. S., Hampsey M. (2002). Transcription activation: unveiling the essential nature of TFIID. *Curr Biol* 12: R620-622.

Chen C. H., von Kessler D. P., Park W., Wang B., Ma Y., Beachy P. A. (1999). Nuclear trafficking of Cubitus interruptus in the transcriptional regulation of Hedgehog target gene expression. *Cell* 98: 305-316.

Chen W., Cooper T. K., Zahnow C. A., Overholtzer M., Zhao Z., Ladanyi M., Karp J. E., Gokgoz N., Wunder J. S., Andrulis I. L., Levine A. J., Mankowski J. L., Baylin S. B. (2004). Epigenetic and genetic loss of Hic1 function accentuates the role of p53 in tumorigenesis. *Cancer Cell* 6: 387-398.

Chen Z., Hagler J., Palombella V. J., Melandri F., Scherer D., Ballard D., Maniatis T. (1995). Signal-induced site-specific phosphorylation targets I kappa B alpha to the ubiquitin-proteasome pathway. *Genes Dev* 9: 1586-1597.

Cheng J., Kang X., Zhang S., Yeh E. T. (2007). SUMO-specific protease 1 is essential for stabilization of HIF1alpha during hypoxia. *Cell* 131: 584-595.

Cho K. O., Choi K. W. (1998). Fringe is essential for mirror symmetry and morphogenesis in the Drosophila eye. *Nature* 396: 272-276.

Chopra V. S., Srinivasan A., Kumar R. P., Mishra K., Basquin D., Docquier M., Seum C., Pauli D., Mishra R. K. (2008). Transcriptional activation by GAGA factor is through its direct interaction with dmTAF3. *Dev Biol* 317: 660-670.

Chun T. H., Itoh H., Subramanian L., Iniguez-Lluhi J. A., Nakao K. (2003). Modification of GATA-2 transcriptional activity in endothelial cells by the SUMO E3 ligase PIASy. *Circ Res* 92: 1201-1208.

Chung C. D., Liao J., Liu B., Rao X., Jay P., Berta P., Shuai K. (1997). Specific inhibition of Stat3 signal transduction by PIAS3. *Science* 278: 1803-1805.

Chung T. L., Hsiao H. H., Yeh Y. Y., Shia H. L., Chen Y. L., Liang P. H., Wang A. H., Khoo K. H., Shoei-Lung Li S. (2004). In vitro modification of human centromere protein CENP-C fragments by small ubiquitin-like modifier (SUMO) protein: definitive identification of the modification sites by tandem mass spectrometry analysis of the isopeptides. *J Biol Chem* 279: 39653-39662.

Da Ros V. G., Gutierrez-Perez I., Ferres-Marco D., Dominguez M. (2013). Dampening the signals transduced through hedgehog via microRNA miR-7 facilitates notch-induced tumourigenesis. *PLoS Biol* 11: e1001554.

Dadke S., Cotteret S., Yip S. C., Jaffer Z. M., Haj F., Ivanov A., Rauscher F., 3rd, Shuai K., Ng T., Neel B. G., Chernoff J. (2007). Regulation of protein tyrosine phosphatase 1B by sumoylation. *Nat Cell Biol* 9: 80-85.

Dawlaty M. M., Malureanu L., Jeganathan K. B., Kao E., Sustmann C., Tahk S., Shuai K., Grosschedl R., van Deursen J. M. (2008). Resolution of sister centromeres requires RanBP2-mediated SUMOylation of topoisomerase IIalpha. *Cell* 133: 103-115.

De Strooper B., Annaert W., Cupers P., Saftig P., Craessaerts K., Mumm J. S., Schroeter E. H., Schrijvers V., Wolfe M. S., Ray W. J., Goate A., Kopan R. (1999). A presenilin-1-dependent gamma-secretase-like protease mediates release of Notch intracellular domain. *Nature* 398: 518-522.

Deato M. D., Marr M. T., Sottero T., Inouye C., Hu P., Tjian R. (2008). MyoD targets TAF3/TRF3 to activate myogenin transcription. *Mol Cell* 32: 96-105.

Desterro J. M., Rodriguez M. S., Hay R. T. (1998). SUMO-1 modification of IkappaBalpha inhibits NF-kappaB activation. *Mol Cell* 2: 233-239.

Desterro J. M., Rodriguez M. S., Kemp G. D., Hay R. T. (1999). Identification of the enzyme required for activation of the small ubiquitin-like protein SUMO-1. *J Biol Chem* 274: 10618-10624.

Di Bacco A., Ouyang J., Lee H. Y., Catic A., Ploegh H., Gill G. (2006). The SUMO-specific protease SENP5 is required for cell division. *Mol Cell Biol* 26: 4489-4498.

Diaz-Benjumea F. J., Cohen S. M. (1995). Serrate signals through Notch to establish a Wingless-dependent organizer at the dorsal/ventral compartment boundary of the Drosophila wing. *Development* 121: 4215-4225.

Dobreva G., Dambacher J., Grosschedl R. (2003). SUMO modification of a novel MAR-binding protein, SATB2, modulates immunoglobulin mu gene expression. *Genes Dev* 17: 3048-3061.

Dominguez M. (2014). Oncogenic programmes and Notch activity: an 'organized crime'? *Semin Cell Dev Biol* 28: 78-85.

Dominguez M., Casares F. (2005). Organ specification-growth control connection: new in-sights from the *Drosophila* eye-antennal disc. *Dev Dyn* 232: 673-684.

Dominguez M., de Celis J. F. (1998). A dorsal/ventral boundary established by Notch controls growth and polarity in the Drosophila eye. *Nature* 396: 276-278.

Dominguez M., Ferres-Marco D., Gutierrez-Avino F. J., Speicher S. A., Beneyto M. (2004). Growth and specification of the eye are controlled independently by Eyegone and Eyeless in Drosophila melanogaster. *Nat Genet* 36: 31-39.

Donaghue C., Bates H., Cotterill S. (2001). Identification and characterisation of the Drosophila homologue of the yeast Uba2 gene. *Biochim Biophys Acta* 1518: 210-214.

Dorn R., Krauss V., Reuter G., Saumweber H. (1993). The enhancer of position-effect variegation of Drosophila, E(var)3-93D, codes for a chromatin protein containing a

conserved domain common to several transcriptional regulators. *Proc Natl Acad Sci U S A* 90: 11376-11380.

Duval D., Duval G., Kedinger C., Poch O., Boeuf H. (2003). The 'PINIT' motif, of a newly identified conserved domain of the PIAS protein family, is essential for nuclear retention of PIAS3L. *FEBS Lett* 554: 111-118.

Dye B. T., Schulman B. A. (2007). Structural mechanisms underlying posttranslational modification by ubiquitin-like proteins. *Annu Rev Biophys Biomol Struct* 36: 131-150.

Ellisen L. W., Bird J., West D. C., Soreng A. L., Reynolds T. C., Smith S. D., Sklar J. (1991). TAN-1, the human homolog of the Drosophila notch gene, is broken by chromosomal translocations in T lymphoblastic neoplasms. *Cell* 66: 649-661.

Epps J. L., Tanda S. (1998). The Drosophila semushi mutation blocks nuclear import of bicoid during embryogenesis. *Curr Biol* 8: 1277-1280.

Escalante R., Wessels D., Soll D. R., Loomis W. F. (1997). Chemotaxis to cAMP and slug migration in Dictyostelium both depend on migA, a BTB protein. *Mol Biol Cell* 8: 1763-1775.

Evdokimov E., Sharma P., Lockett S. J., Lualdi M., Kuehn M. R. (2008). Loss of SUMO1 in mice affects RanGAP1 localization and formation of PML nuclear bodies, but is not lethal as it can be compensated by SUMO2 or SUMO3. *J Cell Sci* 121: 4106-4113.

Everett R. D., Earnshaw W. C., Findlay J., Lomonte P. (1999). Specific destruction of kinetochore protein CENP-C and disruption of cell division by herpes simplex virus immediate-early protein Vmw110. *EMBO J* 18: 1526-1538.

Farkas G., Gausz J., Galloni M., Reuter G., Gyurkovics H., Karch F. (1994). The Trithorax-like gene encodes the Drosophila GAGA factor. *Nature* 371: 806-808.

Ferguson E. L., Horvitz H. R. (1985). Identification and characterization of 22 genes that affect the vulval cell lineages of the nematode Caenorhabditis elegans. *Genetics* 110: 17-72.

Ferres-Marco D. (2010) Identificación de genes y mecanismos causativos de cáncer y metástasis en *Drosophila melanogaster*. PhD Thesis.

Ferres-Marco D., Gutierrez-Garcia I., Vallejo D. M., Bolivar J., Gutierrez-Avino F. J., Dominguez M. (2006). Epigenetic silencers and Notch collaborate to promote malignant tumours by Rb silencing. *Nature* 439: 430-436.

Fishbain S., Prakash S., Herrig A., Elsasser S., Matouschek A. (2011). Rad23 escapes degradation because it lacks a proteasome initiation region. *Nat Commun* 2: 192.

Flotho A., Melchior F. (2013). Sumoylation: a regulatory protein modification in health and disease. *Annu Rev Biochem* 82: 357-385.

Formstecher E., Aresta S., Collura V., Hamburger A., Meil A., Trehin A., Reverdy C., Betin V., Maire S., Brun C., Jacq B., Arpin M., Bellaiche Y., Bellusci S., Benaroch P., Bornens M., Chanet R., Chavrier P., Delattre O., Doye V., Fehon R., Faye G., Galli T., Girault J. A., Goud B., de Gunzburg J., Johannes L., Junier M. P., Mirouse V., Mukherjee A., Papadopoulo D., Perez F., Plessis A., Rosse C., Saule S., Stoppa-Lyonnet D., Vincent A., White M., Legrain P., Wojcik J., Camonis J., Daviet L. (2005). Protein interaction mapping: a Drosophila case study. *Genome Res* 15: 376-384.

Fortini M. E. (2009). Notch signaling: the core pathway and its posttranslational regulation. *Dev Cell* 16: 633-647.

Fraser A. G., Kamath R. S., Zipperlen P., Martinez-Campos M., Sohrmann M., Ahringer J. (2000). Functional genomic analysis of C. elegans chromosome I by systematic RNA interference. *Nature* 408: 325-330.

Fromont-Racine M., Rain J. C., Legrain P. (1997). Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat Genet* 16: 277-282.

Fryer C. J., Lamar E., Turbachova I., Kintner C., Jones K. A. (2002). Mastermind mediates chromatin-specific transcription and turnover of the Notch enhancer complex. *Genes Dev* 16: 1397-1411.

Furukawa M., He Y. J., Borchers C., Xiong Y. (2003). Targeting of protein ubiquitination by BTB-Cullin 3-Roc1 ubiquitin ligases. *Nat Cell Biol* 5: 1001-1007.

Gangloff Y. G., Pointud J. C., Thuault S., Carre L., Romier C., Muratoglu S., Brand M., Tora L., Couderc J. L., Davidson I. (2001a). The TFIID components human TAF(II)140 and Drosophila BIP2 (TAF(II)155) are novel metazoan homologues of yeast TAF(II)47 containing a histone fold and a PHD finger. *Mol Cell Biol* 21: 5109-5121.

Gangloff Y. G., Sanders S. L., Romier C., Kirschner D., Weil P. A., Tora L., Davidson I. (2001b). Histone folds mediate selective heterodimerization of yeast TAF(II)25 with TFIID components yTAF(II)47 and yTAF(II)65 and with SAGA component ySPT7. *Mol Cell Biol* 21: 1841-1853.

Gangloff Y. G., Werten S., Romier C., Carre L., Poch O., Moras D., Davidson I. (2000). The human TFIID components TAF(II)135 and TAF(II)20 and the yeast SAGA components ADA1 and TAF(II)68 heterodimerize to form histone-like pairs. *Mol Cell Biol* 20: 340-351.

Geiss-Friedlander R., Melchior F. (2007). Concepts in sumoylation: a decade on. *Nat Rev Mol Cell Biol* 8: 947-956.

Geyer R., Wee S., Anderson S., Yates J., Wolf D. A. (2003). BTB/POZ domain proteins are putative substrate adaptors for cullin 3 ubiquitin ligases. *Mol Cell* 12: 783-790.

Ghisletti S., Huang W., Ogawa S., Pascual G., Lin M. E., Willson T. M., Rosenfeld M. G., Glass C. K. (2007). Parallel SUMOylation-dependent pathways mediate gene- and signal-specific transrepression by LXRs and PPARgamma. *Mol Cell* 25: 57-70.

Gill G. (2003). Post-translational modification by the small ubiquitin-related modifier SUMO has big effects on transcription factor activity. *Curr Opin Genet Dev* 13: 108-113.

Girdwood D., Bumpass D., Vaughan O. A., Thain A., Anderson L. A., Snowden A. W., Garcia-Wilson E., Perkins N. D., Hay R. T. (2003). P300 transcriptional repression is mediated by SUMO modification. *Mol Cell* 11: 1043-1054.

Godt D., Couderc J. L., Cramton S. E., Laski F. A. (1993). Pattern formation in the limbs of Drosophila: bric a brac is expressed in both a gradient and a wave-like pattern and is required for specification and proper segmentation of the tarsus. *Development* 119: 799-812.

Gong L., Millas S., Maul G. G., Yeh E. T. (2000). Differential regulation of sentrinized proteins by a novel sentrin-specific protease. *J Biol Chem* 275: 3355-3359.

Gong L., Yeh E. T. (2006). Characterization of a family of nucleolar SUMO-specific proteases with preference for SUMO-2 or SUMO-3. *J Biol Chem* 281: 15869-15877.

Goodson M. L., Hong Y., Rogers R., Matunis M. J., Park-Sarge O. K., Sarge K. D. (2001). Sumo-1 modification regulates the DNA binding activity of heat shock transcription factor 2, a promyelocytic leukaemia nuclear body associated transcription factor. *J Biol Chem* 276: 18513-18518.

Gordon W. R., Arnett K. L., Blacklow S. C. (2008). The molecular logic of Notch signaling--a structural and biochemical perspective. *J Cell Sci* 121: 3109-3119.

Gostissa M., Hengstermann A., Fogal V., Sandy P., Schwarz S. E., Scheffner M., Del Sal G. (1999). Activation of p53 by conjugation to the ubiquitin-like protein SUMO-1. *EMBO J* 18: 6462-6471.

Grabher C., von Boehmer H., Look A. T. (2006). Notch 1 activation in the molecular pathogenesis of T-cell acute lymphoblastic leukaemia. *Nat Rev Cancer* 6: 347-359.

Greenwald I. S., Sternberg P. W., Horvitz H. R. (1983). The lin-12 locus specifies cell fates in Caenorhabditis elegans. *Cell* 34: 435-444.

Grillo M., Furriols M., Casanova J., Luschnig S. (2011). Control of germline torso expression by the BTB/POZ domain protein pipsqueak is required for embryonic terminal patterning in Drosophila. *Genetics* 187: 513-521.

Groll M., Bajorek M., Kohler A., Moroder L., Rubin D. M., Huber R., Glickman M. H., Finley D. (2000). A gated channel into the proteasome core particle. *Nat Struct Biol* 7: 1062-1067.

Groll M., Koguchi Y., Huber R., Kohno J. (2001). Crystal structure of the 20 S proteasome:TMC-95A complex: a non-covalent proteasome inhibitor. *J Mol Biol* 311: 543-548.

Guo D., Li M., Zhang Y., Yang P., Eckenrode S., Hopkins D., Zheng W., Purohit S., Podolsky R. H., Muir A., Wang J., Dong Z., Brusko T., Atkinson M., Pozzilli P., Zeidler A., Raffel L. J., Jacob C. O., Park Y., Serrano-Rios M., Larrad M. T., Zhang Z., Garchon H. J., Bach J. F., Rotter J. I., She J. X., Wang C. Y. (2004). A functional variant of SUMO4, a new I kappa B alpha modifier, is associated with type 1 diabetes. *Nat Genet* 36: 837-841.

Gutierrez-Aviño F. J., Ferres-Marco D., Dominguez M. (2009). The position and function of the Notch-mediated eye growth organizer: the roles of JAK/STAT and four-jointed. *EMBO Rep* 10: 1051-1058.

Gutiérrez L., Zurita M., Kennison J. A., Vázquez M. (2003). The Drosophila trithorax group gene tonalli (tna) interacts genetically with the Brahma remodeling complex and encodes an SP-RING finger protein. *Development* 130: 343-354.

Haltiwanger R. S. (2002). Regulation of signal transduction pathways in development by glycosylation. *Curr Opin Struct Biol* 12: 593-598.

Hanahan D., Weinberg R. A. (2011). Hallmarks of cancer: the next generation. *Cell* 144: 646-674.

Hang J., Dasso M. (2002). Association of the human SUMO-1 protease SENP2 with the nuclear pore. *J Biol Chem* 277: 19961-19966.

Hannich J. T., Lewis A., Kroetz M. B., Li S. J., Heide H., Emili A., Hochstrasser M. (2005). Defining the SUMO-modified proteome by multiple approaches in Saccharomyces cerevisiae. *J Biol Chem* 280: 4102-4110.

Hardeland U., Steinacher R., Jiricny J., Schar P. (2002). Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. *EMBO J* 21: 1456-1464.

Harder Z., Zunino R., McBride H. (2004). Sumo1 conjugates mitochondrial substrates and participates in mitochondrial fission. *Curr Biol* 14: 340-345.

Hari K. L., Cook K. R., Karpen G. H. (2001). The Drosophila Su(var)2-10 locus regulates chromosome structure and function and encodes a member of the PIAS protein family. *Genes Dev* 15: 1334-1348.

Hayashi T., Seki M., Maeda D., Wang W., Kawabe Y., Seki T., Saitoh H., Fukagawa T., Yagi H., Enomoto T. (2002). Ubc9 is essential for viability of higher eukaryotic cells. *Exp Cell Res* 280: 212-221.

He L. Z., Guidez F., Tribioli C., Peruzzi D., Ruthardt M., Zelent A., Pandolfi P. P. (1998). Distinct interactions of PML-RARalpha and PLZF-RARalpha with co-repressors determine differential responses to RA in APL. *Nat Genet* 18: 126-135.

Heberlein U., Borod E. R., Chanut F. A. (1998). Dorsoventral patterning in the Drosophila retina by wingless. *Development* 125: 567-577.

Hecker C. M., Rabiller M., Haglund K., Bayer P., Dikic I. (2006). Specification of SUMO1- and SUMO2-interacting motifs. *J Biol Chem* 281: 16117-16127.

Hershko A., Ciechanover A. (1998). The ubiquitin system. *Annu Rev Biochem* 67: 425-479.

Hicke L. (1997). Ubiquitin-dependent internalization and down-regulation of plasma membrane proteins. *FASEB J* 11: 1215-1226.

Hietakangas V., Anckar J., Blomster H. A., Fujimoto M., Palvimo J. J., Nakai A., Sistonen L. (2006). PDSM, a motif for phosphorylation-dependent SUMO modification. *Proc Natl Acad Sci U S A* 103: 45-50.

Hochstrasser M. (2001). SP-RING for SUMO: new functions bloom for a ubiquitin-like protein. *Cell* 107: 5-8.

Hodgkin J., Doniach T., Shen M. (1985). The sex determination pathway in the nematode Caenorhabditis elegans: variations on a theme. *Cold Spring Harb Symp Quant Biol* 50: 585-593.

Hoege C., Pfander B., Moldovan G. L., Pyrowolakis G., Jentsch S. (2002). RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* 419: 135-141.

Holmstrom S., Van Antwerp M. E., Iniguez-Lluhi J. A. (2003). Direct and distinguishable inhibitory roles for SUMO isoforms in the control of transcriptional synergy. *Proc Natl Acad Sci U S A* 100: 15758-15763.

Hombria J. C., Brown S. (2002). The fertile field of Drosophila Jak/STAT signalling. *Curr Biol* 12: R569-575.

Hong S. H., David G., Wong C. W., Dejean A., Privalsky M. L. (1997). SMRT corepressor interacts with PLZF and with the PML-retinoic acid receptor alpha (RARalpha) and PLZF-RARalpha oncoproteins associated with acute promyelocytic leukaemia. *Proc Natl Acad Sci U S A* 94: 9028-9033.

Hong Y., Rogers R., Matunis M. J., Mayhew C. N., Goodson M. L., Park-Sarge O. K., Sarge K. D. (2001). Regulation of heat shock transcription factor 1 by stress-induced SUMO-1 modification. *J Biol Chem* 276: 40263-40267.

Hoppe T., Rape M., Jentsch S. (2001). Membrane-bound transcription factors: regulated release by RIP or RUP. *Curr Opin Cell Biol* 13: 344-348.

Horowitz H., Berg C. A. (1996). The Drosophila pipsqueak gene encodes a nuclear BTB-domain-containing protein required early in oogenesis. *Development* 122: 1859-1871.

Huang D. H., Chang Y. L., Yang C. C., Pan I. C., King B. (2002). pipsqueak encodes a factor essential for sequence-specific targeting of a polycomb group protein complex. *Mol Cell Biol* 22: 6261-6271.

Huang H. W., Tsoi S. C., Sun Y. H., Li S. S. (1998). Identification and characterization of the SMT3 cDNA and gene encoding ubiquitin-like protein from Drosophila melanogaster. *Biochem Mol Biol Int* 46: 775-785.

Huang T. T., Wuerzberger-Davis S. M., Wu Z. H., Miyamoto S. (2003). Sequential modification of NEMO/IKKgamma by SUMO-1 and ubiquitin mediates NF-kappaB activation by genotoxic stress. *Cell* 115: 565-576.

Hurlbut G. D., Kankel M. W., Lake R. J., Artavanis-Tsakonas S. (2007). Crossing paths with Notch in the hyper-network. *Curr Opin Cell Biol* 19: 166-175.

Husnjak K., Dikic I. (2012). Ubiquitin-binding proteins: decoders of ubiquitin-mediated cellular functions. *Annu Rev Biochem* 81: 291-322.

Irvine K. D. (1999). Fringe, Notch, and making developmental boundaries. *Curr Opin Genet Dev* 9: 434-441.

Irvine K. D., Rauskolb C. (2001). Boundaries in development: formation and function. *Annu Rev Cell Dev Biol* 17: 189-214.

Ivanov A. V., Peng H., Yurchenko V., Yap K. L., Negorev D. G., Schultz D. C., Psulkowski E., Fredericks W. J., White D. E., Maul G. G., Sadofsky M. J., Zhou M. M., Rauscher F. J., 3rd (2007). PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. *Mol Cell* 28: 823-837.

Jackson P. K. (2001). A new RING for SUMO: wrestling transcriptional responses into nuclear bodies with PIAS family E3 SUMO ligases. *Genes Dev* 15: 3053-3058.

Jasper H., Benes V., Atzberger A., Sauer S., Ansorge W., Bohmann D. (2002). A genomic switch at the transition from cell proliferation to terminal differentiation in the Drosophila eye. *Dev Cell* 3: 511-521.

Jentsch S., Psakhye I. (2013). Control of nuclear activities by substrate-selective and protein-group SUMOylation. *Annu Rev Genet* 47: 167-186.

Jentsch S., Pyrowolakis G. (2000). Ubiquitin and its kin: how close are the family ties? *Trends Cell Biol* 10: 335-342.

Jia J., Zhang L., Zhang Q., Tong C., Wang B., Hou F., Amanai K., Jiang J. (2005). Phosphorylation by double-time/CKIepsilon and CKIalpha targets cubitus interruptus for Slimb/beta-TRCP-mediated proteolytic processing. *Dev Cell* 9: 819-830.

Jiang J., Struhl G. (1998). Regulation of the Hedgehog and Wingless signalling pathways by the F-box/WD40-repeat protein Slimb. *Nature* 391: 493-496.

Johnson E. S. (2004). Protein modification by SUMO. *Annu Rev Biochem* 73: 355-382.

Johnson E. S., Gupta A. A. (2001). An E3-like factor that promotes SUMO conjugation to the yeast septins. *Cell* 106: 735-744.

Johnson E. S., Schwienhorst I., Dohmen R. J., Blobel G. (1997). The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer. *EMBO J* 16: 5509-5519.

Jones D., Crowe E., Stevens T. A., Candido E. P. (2002). Functional and phylogenetic analysis of the ubiquitylation system in Caenorhabditis elegans: ubiquitin-conjugating enzymes, ubiquitin-activating enzymes, and ubiquitin-like proteins. *Genome Biol* 3: RESEARCH0002.

Joseph J., Liu S. T., Jablonski S. A., Yen T. J., Dasso M. (2004). The RanGAP1-RanBP2 complex is essential for microtubule-kinetochore interactions in vivo. *Curr Biol* 14: 611-617.

Joseph J., Tan S. H., Karpova T. S., McNally J. G., Dasso M. (2002). SUMO-1 targets RanGAP1 to kinetochores and mitotic spindles. *J Cell Biol* 156: 595-602.

Kang S. I., Chang W. J., Cho S. G., Kim I. Y. (2003). Modification of promyelocytic leukaemia zinc finger protein (PLZF) by SUMO-1 conjugation regulates its transcriptional repressor activity. *J Biol Chem* 278: 51479-51483.

Kang S. I., Choi H. W., Kim I. Y. (2008). Redox-mediated modification of PLZF by SUMO-1 and ubiquitin. *Biochem Biophys Res Commun* 369: 1209-1214.

Karanu F. N., Murdoch B., Gallacher L., Wu D. M., Koremoto M., Sakano S., Bhatia M. (2000). The notch ligand jagged-1 represents a novel growth factor of human hematopoietic stem cells. *J Exp Med* 192: 1365-1372.

Kasturi P., Zanetti S., Passannante M., Saudan Z., Muller F., Puoti A. The C. elegans sex determination protein MOG-3 functions in meiosis and binds to the CSL co-repressor CIR-1. *Dev Biol* 344: 593-602.

Kerrigan L. A., Croston G. E., Lira L. M., Kadonaga J. T. (1991). Sequence-specific transcriptional antirepression of the Drosophila Kruppel gene by the GAGA factor. *J Biol Chem* 266: 574-582.

Kerscher O. (2007). SUMO junction-what's your function? New insights through SUMO-interacting motifs. *EMBO Rep* 8: 550-555.

Kidd S., Kelley M. R., Young M. W. (1986). Sequence of the notch locus of Drosophila melanogaster: relationship of the encoded protein to mammalian clotting and growth factors. *Mol Cell Biol* 6: 3094-3108.

Kim J., Cantwell C. A., Johnson P. F., Pfarr C. M., Williams S. C. (2002a). Transcriptional activity of CCAAT/enhancer-binding proteins is controlled by a conserved inhibitory domain that is a target for sumoylation. *J Biol Chem* 277: 38037-38044.

Kim J. H., Choi H. J., Kim B., Kim M. H., Lee J. M., Kim I. S., Lee M. H., Choi S. J., Kim K. I., Kim S. I., Chung C. H., Baek S. H. (2006). Roles of sumoylation of a reptin chromatin-remodelling complex in cancer metastasis. *Nat Cell Biol* 8: 631-639.

Kim K. I., Baek S. H. (2006). SUMOylation code in cancer development and metastasis. *Mol Cells* 22: 247-253.

Kim K. I., Baek S. H. (2009). Small ubiquitin-like modifiers in cellular malignancy and metastasis. *Int Rev Cell Mol Biol* 273: 265-311.

Kim K. I., Baek S. H., Jeon Y. J., Nishimori S., Suzuki T., Uchida S., Shimbara N., Saitoh H., Tanaka K., Chung C. H. (2000). A new SUMO-1-specific protease, SUSP1, that is highly expressed in reproductive organs. *J Biol Chem* 275: 14102-14106.

Kim S. W., Fang X., Ji H., Paulson A. F., Daniel J. M., Ciesiolka M., van Roy F., McCrea P. D. (2002b). Isolation and characterization of XKaiso, a transcriptional repressor that associates with the catenin Xp120(ctn) in Xenopus laevis. *J Biol Chem* 277: 8202-8208.

Koonin E. V., Senkevich T. G., Chernos V. I. (1992). A family of DNA virus genes that consists of fused portions of unrelated cellular genes. *Trends Biochem Sci* 17: 213-214.

Kopan R., Ilagan M. X. (2009). The canonical Notch signaling pathway: unfolding the activation mechanism. *Cell* 137: 216-233.

Kotaja N., Karvonen U., Janne O. A., Palvimo J. J. (2002). PIAS proteins modulate transcription factors by functioning as SUMO-1 ligases. *Mol Cell Biol* 22: 5222-5234.

Kovall R. A. (2008). More complicated than it looks: assembly of Notch pathway transcription complexes. *Oncogene* 27: 5099-5109.

Kraut D. A., Matouschek A. (2011). Proteasomal degradation from internal sites favors partial proteolysis via remote domain stabilization. *ACS Chem Biol* 6: 1087-1095.

Krumova P., Meulmeester E., Garrido M., Tirard M., Hsiao H. H., Bossis G., Urlaub H., Zweckstetter M., Kugler S., Melchior F., Bahr M., Weishaupt J. H. (2011). Sumoylation inhibits alpha-synuclein aggregation and toxicity. *J Cell Biol* 194: 49-60.

Kunert N., Wagner E., Murawska M., Klinker H., Kremmer E., Brehm A. (2009). dMec: a novel Mi-2 chromatin remodelling complex involved in transcriptional repression. *EMBO J* 28: 533-544.

Lai E. C. (2004). Notch signaling: control of cell communication and cell fate. *Development* 131: 965-973.

Lawrence P. A. (1992). The Making of a Fly: The Genetics of Animal Design. *Blackwell Science Publications, Oxford*.

Le Borgne R., Bardin A., Schweisguth F. (2005). The roles of receptor and ligand endocytosis in regulating Notch signaling. *Development* 132: 1751-1762.

Le H. D., Donaldson K. M., Cook K. R., Karpen G. H. (2004). A high proportion of genes involved in position effect variegation also affect chromosome inheritance. *Chromosoma* 112: 269-276.

Lee C., Prakash S., Matouschek A. (2002). Concurrent translocation of multiple polypeptide chains through the proteasomal degradation channel. *J Biol Chem* 277: 34760-34765.

Lehembre F., Badenhorst P., Muller S., Travers A., Schweisguth F., Dejean A. (2000). Covalent modification of the transcriptional repressor tramtrack by the ubiquitin-related protein Smt3 in Drosophila flies. *Mol Cell Biol* 20: 1072-1082.

Lehmann M., Siegmund T., Lintermann K. G., Korge G. (1998). The pipsqueak protein of Drosophila melanogaster binds to GAGA sequences through a novel DNA-binding domain. *J Biol Chem* 273: 28504-28509.

Leight E. R., Glossip D., Kornfeld K. (2005). Sumoylation of LIN-1 promotes transcriptional repression and inhibition of vulval cell fates. *Development* 132: 1047-1056.

Leong K. G., Niessen K., Kulic I., Raouf A., Eaves C., Pollet I., Karsan A. (2007). Jagged1-mediated Notch activation induces epithelial-to-mesenchymal transition through Slug-induced repression of E-cadherin. *J Exp Med* 204: 2935-2948.

Li S. J., Hochstrasser M. (1999). A new protease required for cell-cycle progression in yeast. *Nature* 398: 246-251.

Li S. J., Hochstrasser M. (2000). The yeast ULP2 (SMT4) gene encodes a novel protease specific for the ubiquitin-like Smt3 protein. *Mol Cell Biol* 20: 2367-2377.

Li S. J., Hochstrasser M. (2003). The Ulp1 SUMO isopeptidase: distinct domains required for viability, nuclear envelope localization, and substrate specificity. *J Cell Biol* 160: 1069-1081.

Li X., Peng H., Schultz D. C., Lopez-Guisa J. M., Rauscher F. J., 3rd, Marmorstein R. (1999). Structure-function studies of the BTB/POZ transcriptional repression domain from the promyelocytic leukaemia zinc finger oncoprotein. *Cancer Res* 59: 5275-5282.

Lieber T., Kidd S., Young M. W. (2002). kuzbanian-mediated cleavage of Drosophila Notch. *Genes Dev* 16: 209-221.

Liefke R., Oswald F., Alvarado C., Ferres-Marco D., Mittler G., Rodriguez P., Dominguez M., Borggrefe T. (2010). Histone demethylase KDM5A is an integral part of the core Notch-RBP-J repressor complex. *Genes Dev* 24: 590-601.

Lin R. J., Nagy L., Inoue S., Shao W., Miller W. H., Jr., Evans R. M. (1998). Role of the histone deacetylase complex in acute promyelocytic leukaemia. *Nature* 391: 811-814.

Linder B., Newman R., Jones L. K., Debernardi S., Young B. D., Freemont P., Verrijzer C. P., Saha V. (2000). Biochemical analyses of the AF10 protein: the extended LAP/PHD-finger mediates oligomerisation. *J Mol Biol* 299: 369-378.

Liu B., Liao J., Rao X., Kushner S. A., Chung C. D., Chang D. D., Shuai K. (1998). Inhibition of Stat1-mediated gene activation by PIAS1. *Proc Natl Acad Sci U S A* 95: 10626-10631.

Liu B., Mink S., Wong K. A., Stein N., Getman C., Dempsey P. W., Wu H., Shuai K. (2004). PIAS1 selectively inhibits interferon-inducible genes and is important in innate immunity. *Nat Immunol* 5: 891-898.

Liu B., Yang R., Wong K. A., Getman C., Stein N., Teitell M. A., Cheng G., Wu H., Shuai K. (2005). Negative regulation of NF-kappaB signaling by PIAS1. *Mol Cell Biol* 25: 1113-1123.

Liu B., Yang Y., Chernishof V., Loo R. R., Jang H., Tahk S., Yang R., Mink S., Shultz D., Bellone C. J., Loo J. A., Shuai K. (2007a). Proinflammatory stimuli induce IKKalpha-mediated phosphorylation of PIAS1 to restrict inflammation and immunity. *Cell* 129: 903-914.

Liu C. W., Corboy M. J., DeMartino G. N., Thomas P. J. (2003). Endoproteolytic activity of the proteasome. *Science* 299: 408-411.

Liu X., Liu Z., Jang S. W., Ma Z., Shinmura K., Kang S., Dong S., Chen J., Fukasawa K., Ye K. (2007b). Sumoylation of nucleophosmin/B23 regulates its subcellular localization, mediating cell proliferation and survival. *Proc Natl Acad Sci U S A* 104: 9679-9684.

Liu Z., Scannell D. R., Eisen M. B., Tjian R. (2011). Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. *Cell* 146: 720-731.

Long X., Griffith L. C. (2000). Identification and characterization of a SUMO-1 conjugation system that modifies neuronal calcium/calmodulin-dependent protein kinase II in Drosophila melanogaster. *J Biol Chem* 275: 40765-40776.

Mabb A. M., Wuerzberger-Davis S. M., Miyamoto S. (2006). PIASy mediates NEMO sumoylation and NF-kappaB activation in response to genotoxic stress. *Nat Cell Biol* 8: 986-993.

Madhavan M. M., Madhavan K. (1980). Morphogenesis of the epidermis of adult abdomen of Drosophila. *J Embryol Exp Morphol* 60: 1-31.

Maeda T., Hobbs R. M., Merghoub T., Guernah I., Zelent A., Cordon-Cardo C., Teruya-Feldstein J., Pandolfi P. P. (2005). Role of the proto-oncogene Pokemon in cellular transformation and ARF repression. *Nature* 433: 278-285.

Mahajan R., Delphin C., Guan T., Gerace L., Melchior F. (1997). A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell* 88: 97-107.

Martin S., Nishimune A., Mellor J. R., Henley J. M. (2007). SUMOylation regulates kainate-receptor-mediated synaptic transmission. *Nature* 447: 321-325.

Matsubayashi H., Sese S., Lee J. S., Shirakawa T., Iwatsubo T., Tomita T., Yanagawa S. (2004). Biochemical characterization of the Drosophila wingless signaling pathway based on RNA interference. *Mol Cell Biol* 24: 2012-2024.

Matunis M. J., Coutavas E., Blobel G. (1996). A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *J Cell Biol* 135: 1457-1470.

Matunis M. J., Wu J., Blobel G. (1998). SUMO-1 modification and its role in targeting the Ran GTPase-activating protein, RanGAP1, to the nuclear pore complex. *J Cell Biol* 140: 499-509.

Maurel-Zaffran C., Treisman J. E. (2000). pannier acts upstream of wingless to direct dorsal eye disc development in Drosophila. *Development* 127: 1007-1016.

Melchior F., Schergaut M., Pichler A. (2003). SUMO: ligases, isopeptidases and nuclear pores. *Trends Biochem Sci* 28: 612-618.

Melnick A., Ahmad K. F., Arai S., Polinger A., Ball H., Borden K. L., Carlile G. W., Prive G. G., Licht J. D. (2000a). In-depth mutational analysis of the promyelocytic leukaemia zinc finger BTB/POZ domain reveals motifs and residues required for biological and transcriptional functions. *Mol Cell Biol* 20: 6550-6567.

Melnick A., Carlile G., Ahmad K. F., Kiang C. L., Corcoran C., Bardwell V., Prive G. G., Licht J. D. (2002). Critical residues within the BTB domain of PLZF and Bcl-6 modulate interaction with corepressors. *Mol Cell Biol* 22: 1804-1818.

Melnick A. M., Westendorf J. J., Polinger A., Carlile G. W., Arai S., Ball H. J., Lutterbach B., Hiebert S. W., Licht J. D. (2000b). The ETO protein disrupted in t(8;21)-associated acute myeloid leukaemia is a corepressor for the promyelocytic leukaemia zinc finger protein. *Mol Cell Biol* 20: 2075-2086.

Meluh P. B., Koshland D. (1995). Evidence that the MIF2 gene of Saccharomyces cerevisiae encodes a centromere protein with homology to the mammalian centromere protein CENP-C. *Mol Biol Cell* 6: 793-807.

Merzlyak E. M., Goedhart J., Shcherbo D., Bulina M. E., Shcheglov A. S., Fradkov A. F., Gaintzeva A., Lukyanov K. A., Lukyanov S., Gadella T. W., Chudakov D. M. (2007). Bright monomeric red fluorescent protein with an extended fluorescence lifetime. *Nat Methods* 4: 555-557.

Miele L., Miao H., Nickoloff B. J. (2006). NOTCH signaling as a novel cancer therapeutic target. *Curr Cancer Drug Targets* 6: 313-323.

Mikkelsen T. S., Ku M., Jaffe D. B., Issac B., Lieberman E., Giannoukos G., Alvarez P., Brockman W., Kim T. K., Koche R. P., Lee W., Mendenhall E., O'Donovan A., Presser A., Russ C., Xie X., Meissner A., Wernig M., Jaenisch R., Nusbaum C., Lander E. S., Bernstein B. E. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448: 553-560.

Minty A., Dumont X., Kaghad M., Caput D. (2000). Covalent modification of p73alpha by SUMO-1. Two-hybrid screening with p73 identifies novel SUMO-1-interacting proteins and a SUMO-1 interaction motif. *J Biol Chem* 275: 36316-36323.

Miura K., Jin J. B., Hasegawa P. M. (2007). Sumoylation, a post-translational regulatory process in plants. *Curr Opin Plant Biol* 10: 495-502.

Mohr S. E., Boswell R. E. (1999). Zimp encodes a homologue of mouse Miz1 and PIAS3 and is an essential gene in Drosophila melanogaster. *Gene* 229: 109-116.

Moohr O. L. (1919). *Genetics*: 252.

Mossessova E., Lima C. D. (2000). Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol Cell* 5: 865-876.

Mukhopadhyay D., Ayaydin F., Kolli N., Tan S. H., Anan T., Kametaka A., Azuma Y., Wilkinson K. D., Dasso M. (2006). SUSP1 antagonizes formation of highly SUMO2/3-conjugated species. *J Cell Biol* 174: 939-949.

Muller S., Berger M., Lehembre F., Seeler J. S., Haupt Y., Dejean A. (2000). c-Jun and p53 activity is modulated by SUMO-1 modification. *J Biol Chem* 275: 13321-13329.

Muller S., Ledl A., Schmidt D. (2004). SUMO: a regulator of gene expression and genome integrity. *Oncogene* 23: 1998-2008.

Mumm J. S., Kopan R. (2000). Notch signaling: from the outside in. *Dev Biol* 228: 151-165.

Nacerddine K., Lehembre F., Bhaumik M., Artus J., Cohen-Tannoudji M., Babinet C., Pandolfi P. P., Dejean A. (2005). The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. *Dev Cell* 9: 769-779.

Nakatani Y., Bagby S., Ikura M. (1996). The histone folds in transcription factor TFIID. *J Biol Chem* 271: 6575-6578.

Nishida T., Tanaka H., Yasuda H. (2000). A novel mammalian Smt3-specific isopeptidase 1 (SMT3IP1) localized in the nucleolus at interphase. *Eur J Biochem* 267: 6423-6427.

Nussbaum A. K., Dick T. P., Keilholz W., Schirle M., Stevanovic S., Dietz K., Heinemeyer W., Groll M., Wolf D. H., Huber R., Rammensee H. G., Schild H. (1998). Cleavage motifs of the yeast 20S proteasome beta subunits deduced from digests of enolase 1. *Proc Natl Acad Sci U S A* 95: 12504-12509.

Ohsako S., Takamatsu Y. (1999). Identification and characterization of a Drosophila homologue of the yeast UBC9 and hus5 genes. *J Biochem* 125: 230-235.

Okuma T., Honda R., Ichikawa G., Tsumagari N., Yasuda H. (1999). In vitro SUMO-1 modification requires two enzymatic steps, E1 and E2. *Biochem Biophys Res Commun* 254: 693-698.

Orian A., Gonen H., Bercovich B., Fajerman I., Eytan E., Israel A., Mercurio F., Iwai K., Schwartz A. L., Ciechanover A. (2000). SCF(beta)(-TrCP) ubiquitin ligase-mediated processing of NF-kappaB p105 requires phosphorylation of its C-terminus by IkappaB kinase. *EMBO J* 19: 2580-2591.

Palombella V. J., Rando O. J., Goldberg A. L., Maniatis T. (1994). The ubiquitin-proteasome pathway is required for processing the NF-kappa B1 precursor protein and the activation of NF-kappa B. *Cell* 78: 773-785.

Palomero T., Sulis M. L., Cortina M., Real P. J., Barnes K., Ciofani M., Caparros E., Buteau J., Brown K., Perkins S. L., Bhagat G., Agarwal A. M., Basso G., Castillo M., Nagase S., Cordon-Cardo C., Parsons R., Zuniga-Pflucker J. C., Dominguez M., Ferrando A. A. (2007). Mutational loss of PTEN induces resistance to NOTCH1 inhibition in T-cell leukaemia. *Nat Med* 13: 1203-1210.

Pamer E., Cresswell P. (1998). Mechanisms of MHC class I--restricted antigen processing. *Annu Rev Immunol* 16: 323-358.

Panse V. G., Kressler D., Pauli A., Petfalski E., Gnadig M., Tollervey D., Hurt E. (2006). Formation and nuclear export of preribosomes are functionally linked to the small-ubiquitin-related modifier pathway. *Traffic* 7: 1311-1321.

Panse V. G., Kuster B., Gerstberger T., Hurt E. (2003). Unconventional tethering of Ulp1 to the transport channel of the nuclear pore complex by karyopherins. *Nat Cell Biol* 5: 21-27.

Papayannopoulos V., Tomlinson A., Panin V. M., Rauskolb C., Irvine K. D. (1998). Dorsal-ventral signaling in the Drosophila eye. *Science* 281: 2031-2034.

Parker J. L., Ulrich H. D. (2012). A SUMO-interacting motif activates budding yeast ubiquitin ligase Rad18 towards SUMO-modified PCNA. *Nucleic Acids Res* 40: 11380-11388.

Pascual G., Fong A. L., Ogawa S., Gamliel A., Li A. C., Perissi V., Rose D. W., Willson T. M., Rosenfeld M. G., Glass C. K. (2005). A SUMOylation-dependent pathway mediates transrepression of inflammatory response genes by PPAR-gamma. *Nature* 437: 759-763.

Perez-Torrado R., Yamada D., Defossez P. A. (2006). Born to bind: the BTB protein-protein interaction domain. *Bioessays* 28: 1194-1202.

Pickart C. M. (2001). Mechanisms underlying ubiquitination. *Annu Rev Biochem* 70: 503-533.

Pichler A., Gast A., Seeler J. S., Dejean A., Melchior F. (2002). The nucleoporin RanBP2 has SUMO1 E3 ligase activity. *Cell* 108: 109-120.

Pichler A., Knipscheer P., Saitoh H., Sixma T. K., Melchior F. (2004). The RanBP2 SUMO E3 ligase is neither HECT- nor RING-type. *Nat Struct Mol Biol* 11: 984-991.

Pintard L., Willems A., Peter M. (2004). Cullin-based ubiquitin ligases: Cul3-BTB complexes join the family. *EMBO J* 23: 1681-1687.

Pintard L., Willis J. H., Willems A., Johnson J. L., Srayko M., Kurz T., Glaser S., Mains P. E., Tyers M., Bowerman B., Peter M. (2003). The BTB protein MEL-26 is a substrate-specific adaptor of the CUL-3 ubiquitin-ligase. *Nature* 425: 311-316.

Pinte S., Stankovic-Valentin N., Deltour S., Rood B. R., Guerardel C., Leprince D. (2004). The tumor suppressor gene HIC1 (hypermethylated in cancer 1) is a sequence-specific transcriptional repressor: definition of its consensus binding sequence and analysis of its DNA binding and repressive properties. *J Biol Chem* 279: 38313-38324.

Pointud J. C., Larsson J., Dastugue B., Couderc J. L. (2001). The BTB/POZ domain of the regulatory proteins Bric a brac 1 (BAB1) and Bric a brac 2 (BAB2) interacts with the novel Drosophila TAF(II) factor BIP2/dTAF(II)155. *Dev Biol* 237: 368-380.

Polo J. M., Dell'Oso T., Ranuncolo S. M., Cerchietti L., Beck D., Da Silva G. F., Prive G. G., Licht J. D., Melnick A. (2004). Specific peptide interference reveals BCL6 transcriptional and oncogenic mechanisms in B-cell lymphoma cells. *Nat Med* 10: 1329-1335.

Potter C. J., Turenchalk G. S., Xu T. (2000). Drosophila in cancer research. An expanding role. *Trends Genet* 16: 33-39.

Poukka H., Karvonen U., Janne O. A., Palvimo J. J. (2000). Covalent modification of the androgen receptor by small ubiquitin-like modifier 1 (SUMO-1). *Proc Natl Acad Sci U S A* 97: 14145-14150.

Poulson D. F. (1937). Chromosomal Deficiencies and the Embryonic Development of Drosophila Melanogaster. *Proc Natl Acad Sci U S A* 23: 133-137.

Prakash S., Tian L., Ratliff K. S., Lehotzky R. E., Matouschek A. (2004). An unstructured initiation site is required for efficient proteasome-mediated degradation. *Nat Struct Mol Biol* 11: 830-837.

Prince F., Katsuyama T., Oshima Y., Plaza S., Resendez-Perez D., Berry M., Kurata S., Gehring W. J. (2008). The YPWM motif links Antennapedia to the basal transcriptional machinery. *Development* 135: 1669-1679.

Proschel M., Zhang Z., Parsch J. (2006). Widespread adaptive evolution of Drosophila genes with sex-biased expression. *Genetics* 174: 893-900.

Radtke F., Raj K. (2003). The role of Notch in tumorigenesis: oncogene or tumour suppressor? *Nat Rev Cancer* 3: 756-767.

Raff J. W., Kellum R., Alberts B. (1994). The Drosophila GAGA transcription factor is associated with specific regions of heterochromatin throughout the cell cycle. *EMBO J* 13: 5977-5983.

Rain J. C., Selig L., De Reuse H., Battaglia V., Reverdy C., Simon S., Lenzen G., Petel F., Wojcik J., Schachter V., Chemama Y., Labigne A., Legrain P. (2001). The protein-protein interaction map of Helicobacter pylori. *Nature* 409: 211-215.

Rape M., Jentsch S. (2002). Taking a bite: proteasomal protein processing. *Nat Cell Biol* 4: E113-116.

Read M. A., Neish A. S., Luscinskas F. W., Palombella V. J., Maniatis T., Collins T. (1995). The proteasome pathway is required for cytokine-induced endothelial-leukocyte adhesion molecule expression. *Immunity* 2: 493-506.

Reddy B. A., Bajpe P. K., Bassett A., Moshkin Y. M., Kozhevnikova E., Bezstarosti K., Demmers J. A., Travers A. A., Verrijzer C. P. (2010). Drosophila transcription factor Tramtrack69 binds MEP1 to recruit the chromatin remodeler NuRD. *Mol Cell Biol* 30: 5234-5244.

Reuter G., Wolff I. (1981). Isolation of dominant suppressor mutations for position-effect variegation in Drosophila melanogaster. *Mol Gen Genet* 182: 516-519.

Ringrose L., Paro R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu Rev Genet* 38: 413-443.

Riquelme C., Barthel K. K., Liu X. (2006). SUMO-1 modification of MEF2A regulates its transcriptional activity. *J Cell Mol Med* 10: 132-144.

Rizzo P., Osipo C., Foreman K., Golde T., Osborne B., Miele L. (2008). Rational targeting of Notch signaling in cancer. *Oncogene* 27: 5124-5131.

Rock K. L., Goldberg A. L. (1999). Degradation of cell proteins and the generation of MHC class I-presented peptides. *Annu Rev Immunol* 17: 739-779.

Rodriguez M. S., Desterro J. M., Lain S., Midgley C. A., Lane D. P., Hay R. T. (1999). SUMO-1 modification activates the transcriptional response of p53. *EMBO J* 18: 6455-6461.

Rogers R. S., Horvath C. M., Matunis M. J. (2003). SUMO modification of STAT1 and its role in PIAS-mediated inhibition of gene activation. *J Biol Chem* 278: 30091-30097.

Rørth P. (1996). A modular misexpression screen in Drosophila detecting tissue-specific phenotypes. *Proc Natl Acad Sci U S A* 93: 12418-12422.

Ross S., Best J. L., Zon L. I., Gill G. (2002). SUMO-1 modification represses Sp3 transcriptional activation and modulates its subnuclear localization. *Mol Cell* 10: 831-842.

Rytinki M. M., Kaikkonen S., Pehkonen P., Jaaskelainen T., Palvimo J. J. (2009). PIAS proteins: pleiotropic interactors associated with SUMO. Cell Mol Life Sci 66: 3029-3041.

Sachdev S., Bruhn L., Sieber H., Pichler A., Melchior F., Grosschedl R. (2001). PIASy, a nuclear matrix-associated SUMO E3 ligase, represses LEF1 activity by sequestration into nuclear bodies. *Genes Dev* 15: 3088-3103.

Saitoh H., Hinchey J. (2000). Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. *J Biol Chem* 275: 6252-6258.

Saitoh H., Pu R., Cavenagh M., Dasso M. (1997). RanBP2 associates with Ubc9p and a modified form of RanGAP1. *Proc Natl Acad Sci U S A* 94: 3736-3741.

Santagata S., Demichelis F., Riva A., Varambally S., Hofer M. D., Kutok J. L., Kim R., Tang J., Montie J. E., Chinnaiyan A. M., Rubin M. A., Aster J. C. (2004). JAGGED1 expression is associated with prostate cancer metastasis and recurrence. *Cancer Res* 64: 6854-6857.

Sapetschnig A., Rischitor G., Braun H., Doll A., Schergaut M., Melchior F., Suske G. (2002). Transcription factor Sp3 is silenced through SUMO modification by PIAS1. *EMBO J* 21: 5206-5215.

Saracco S. A., Miller M. J., Kurepa J., Vierstra R. D. (2007). Genetic analysis of SUMOylation in Arabidopsis: conjugation of SUMO1 and SUMO2 to nuclear proteins is essential. *Plant Physiol* 145: 119-134.

Sato A., Tomlinson A. (2007). Dorsal-ventral midline signaling in the developing Drosophila eye. *Development* 134: 659-667.

Schmidt D., Muller S. (2002). Members of the PIAS family act as SUMO ligases for c-Jun and p53 and repress p53 activity. *Proc Natl Acad Sci U S A* 99: 2872-2877.

Schulz S., Chachami G., Kozaczkiewicz L., Winter U., Stankovic-Valentin N., Haas P., Hofmann K., Urlaub H., Ovaa H., Wittbrodt J., Meulmeester E., Melchior F. (2012). Ubiquitin-specific protease-like 1 (USPL1) is a SUMO isopeptidase with essential, non-catalytic functions. *EMBO Rep* 13: 930-938.

Schwendemann A., Lehmann M. (2002). Pipsqueak and GAGA factor act in concert as partners at homeotic and many other loci. *Proc Natl Acad Sci U S A* 99: 12883-12888.

Seeler J. S., Dejean A. (2003). Nuclear and unclear functions of SUMO. Nat Rev Mol Cell Biol 4: 690-699.

Seet B. T., Dikic I., Zhou M. M., Pawson T. (2006). Reading protein modifications with interaction domains. *Nat Rev Mol Cell Biol* 7: 473-483.

Sekiyama N., Ikegami T., Yamane T., Ikeguchi M., Uchimura Y., Baba D., Ariyoshi M., Tochio H., Saitoh H., Shirakawa M. (2008). Structure of the small ubiquitin-like modifier (SUMO)-interacting motif of MBD1-containing chromatin-associated factor 1 bound to SUMO-3. *J Biol Chem* 283: 35966-35975.

Shigenobu S., Kitadate Y., Noda C., Kobayashi S. (2006). Molecular characterization of embryonic gonads by gene expression profiling in Drosophila melanogaster. *Proc Natl Acad Sci U S A* 103: 13728-13733.

Shih H. P., Hales K. G., Pringle J. R., Peifer M. (2002). Identification of septin-interacting proteins and characterization of the Smt3/SUMO-conjugation system in Drosophila. *J Cell Sci* 115: 1259-1271.

Shin E. J., Shin H. M., Nam E., Kim W. S., Kim J. H., Oh B. H., Yun Y. (2012). DeSUMOylating isopeptidase: a second class of SUMO protease. *EMBO Rep* 13: 339-346.

Siegel V., Jongens T. A., Jan L. Y., Jan Y. N. (1993). pipsqueak, an early acting member of the posterior group of genes, affects vasa level and germ cell-somatic cell interaction in the developing egg chamber. *Development* 119: 1187-1202.

Siggs O. M., Beutler B. (2012). The BTB-ZF transcription factors. *Cell Cycle* 11: 3358-3369.

Smith M., Bhaskar V., Fernandez J., Courey A. J. (2004). Drosophila Ulp1, a nuclear pore-associated SUMO protease, prevents accumulation of cytoplasmic SUMO conjugates. *J Biol Chem* 279: 43805-43814.

Sobko A., Ma H., Firtel R. A. (2002). Regulated SUMOylation and ubiquitination of DdMEK1 is required for proper chemotaxis. *Dev Cell* 2: 745-756.

Song J., Durrin L. K., Wilkinson T. A., Krontiris T. G., Chen Y. (2004). Identification of a SUMO-binding motif that recognizes SUMO-modified proteins. *Proc Natl Acad Sci U S A* 101: 14373-14378.

Song J., Zhang Z., Hu W., Chen Y. (2005). Small ubiquitin-like modifier (SUMO) recognition of a SUMO binding motif: a reversal of the bound orientation. *J Biol Chem* 280: 40122-40129.

Spengler M. L., Kurapatwinski K., Black A. R., Azizkhan-Clifford J. (2002). SUMO-1 modification of human cytomegalovirus IE1/IE72. *J Virol* 76: 2990-2996.

Sriramachandran A. M., Dohmen R. J. (2014). SUMO-targeted ubiquitin ligases. *Biochim Biophys Acta* 1843: 75-85.

Stade K., Vogel F., Schwienhorst I., Meusser B., Volkwein C., Nentwig B., Dohmen R. J., Sommer T. (2002). A lack of SUMO conjugation affects cNLS-dependent nuclear protein import in yeast. *J Biol Chem* 277: 49554-49561.

Stankovic-Valentin N., Deltour S., Seeler J., Pinte S., Vergoten G., Guerardel C., Dejean A., Leprince D. (2007). An acetylation/deacetylation-SUMOylation switch through a phylogenetically conserved psiKXEP motif in the tumor suppressor HIC1 regulates transcriptional repression activity. *Mol Cell Biol* 27: 2661-2675.

Stead K., Aguilar C., Hartman T., Drexel M., Meluh P., Guacci V. (2003). Pds5p regulates the maintenance of sister chromatid cohesion and is sumoylated to promote the dissolution of cohesion. *J Cell Biol* 163: 729-741.

Steelman L. S., Abrams S. L., Whelan J., Bertrand F. E., Ludwig D. E., Basecke J., Libra M., Stivala F., Milella M., Tafuri A., Lunghi P., Bonati A., Martelli A. M., McCubrey J. A. (2008). Contributions of the Raf/MEK/ERK, PI3K/PTEN/Akt/mTOR and Jak/STAT pathways to leukaemia. *Leukaemia* 22: 686-707.

Stehmeier P., Muller S. (2009). Phospho-regulated SUMO interaction modules connect the SUMO system to CK2 signaling. *Mol Cell* 33: 400-409.

Steinacher R., Schar P. (2005). Functionality of human thymine DNA glycosylase requires SUMO-regulated changes in protein conformation. *Curr Biol* 15: 616-623.

Stelter P., Ulrich H. D. (2003). Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. *Nature* 425: 188-191.

Stielow B., Sapetschnig A., Kruger I., Kunert N., Brehm A., Boutros M., Suske G. (2008). Identification of SUMO-dependent chromatin-associated transcriptional repression components by a genome-wide RNAi screen. *Mol Cell* 29: 742-754.

Stoynova L., Solorzano R., Collins E. D. (2004). Generation of large deletion mutants from plasmid DNA. *Biotechniques* 36: 402-404, 406.

Struhl G., Basler K. (1993). Organizing activity of wingless protein in Drosophila. *Cell* 72: 527-540.

Strunnikov A. V., Aravind L., Koonin E. V. (2001). Saccharomyces cerevisiae SMT4 encodes an evolutionarily conserved protease with a role in chromosome condensation regulation. *Genetics* 158: 95-107.

Subramanian L., Benson M. D., Iniguez-Lluhi J. A. (2003). A synergy control motif within the attenuator domain of CCAAT/enhancer-binding protein alpha inhibits transcriptional synergy through its PIASy-enhanced modification by SUMO-1 or SUMO-3. *J Biol Chem* 278: 9134-9141.

Suliman B. A., Xu D., Williams B. R. (2012). The promyelocytic leukaemia zinc finger protein: two decades of molecular oncology. *Front Oncol* 2: 74.

Takahashi Y., Kahyo T., Toh E. A., Yasuda H., Kikuchi Y. (2001a). Yeast Ull1/Siz1 is a novel SUMO1/Smt3 ligase for septin components and functions as an adaptor between conjugating enzyme and substrates. *J Biol Chem* 276: 48973-48977.

Takahashi Y., Toh-e A., Kikuchi Y. (2001b). A novel factor required for the SUMO1/Smt3 conjugation of yeast septins. *Gene* 275: 223-231.

Talamillo A., Sanchez J., Barrio R. (2008). Functional analysis of the SUMOylation pathway in Drosophila. *Biochem Soc Trans* 36: 868-873.

Tan J. A., Hall S. H., Hamil K. G., Grossman G., Petrusz P., French F. S. (2002). Protein inhibitors of activated STAT resemble scaffold attachment factors and function as interacting nuclear receptor coregulators. *J Biol Chem* 277: 16993-17001.

Tanaka K., Nishide J., Okazaki K., Kato H., Niwa O., Nakagawa T., Matsuda H., Kawamukai M., Murakami Y. (1999). Characterization of a fission yeast SUMO-1 homologue, pmt3p, required for multiple nuclear events, including the control of telomere length and chromosome segregation. *Mol Cell Biol* 19: 8660-8672.

Tang Z., El Far O., Betz H., Scheschonka A. (2005). Pias1 interaction and sumoylation of metabotropic glutamate receptor 8. *J Biol Chem* 280: 38153-38159.

Tatham M. H., Jaffray E., Vaughan O. A., Desterro J. M., Botting C. H., Naismith J. H., Hay R. T. (2001). Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. *J Biol Chem* 276: 35368-35374.

Terradot L., Durnell N., Li M., Ory J., Labigne A., Legrain P., Colland F., Waksman G. (2004). Biochemical characterization of protein complexes from the Helicobacter pylori protein interaction map: strategies for complex formation and evidence for novel interactions within type IV secretion systems. *Mol Cell Proteomics* 3: 809-819.

Thirkettle H. J., Mills I. G., Whitaker H. C., Neal D. E. (2009). Nuclear LYRIC/AEG-1 interacts with PLZF and relieves PLZF-mediated repression. *Oncogene* 28: 3663-3670.

Tian L., Holmgren R. A., Matouschek A. (2005). A conserved processing mechanism regulates the activity of transcription factors Cubitus interruptus and NF-kappaB. *Nat Struct Mol Biol* 12: 1045-1053.

Tian S., Poukka H., Palvimo J. J., Janne O. A. (2002). Small ubiquitin-related modifier-1 (SUMO-1) modification of the glucocorticoid receptor. *Biochem J* 367: 907-911.

Toba G., Ohsako T., Miyata N., Ohtsuka T., Seong K. H., Aigaki T. (1999). The gene search system. A method for efficient detection and rapid molecular identification of genes in Drosophila melanogaster. *Genetics* 151: 725-737.

Tora L. (2002). A unified nomenclature for TATA box binding protein (TBP)-associated factors (TAFs) involved in RNA polymerase II transcription. *Genes Dev* 16: 673-675.

Treisman J. E. (2013). Retinal differentiation in Drosophila. *Wiley Interdiscip Rev Dev Biol* 2: 545-557.

Tsukiyama T., Becker P. B., Wu C. (1994). ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor. *Nature* 367: 525-532.

Ullmann R., Chien C. D., Avantaggiati M. L., Muller S. (2012). An acetylation switch regulates SUMO-dependent protein interaction networks. *Mol Cell* 46: 759-770.

Ungureanu D., Vanhatupa S., Kotaja N., Yang J., Aittomaki S., Janne O. A., Palvimo J. J., Silvennoinen O. (2003). PIAS proteins promote SUMO-1 conjugation to STAT1. *Blood* 102: 3311-3313.

Unhavaithaya Y., Shin T. H., Miliaras N., Lee J., Oyama T., Mello C. C. (2002). MEP-1 and a homolog of the NURD complex component Mi-2 act together to maintain germline-soma distinctions in C. elegans. *Cell* 111: 991-1002.

van der Veen A. G., Ploegh H. L. (2012). Ubiquitin-like proteins. *Annu Rev Biochem* 81: 323-357.

Verger A., Perdomo J., Crossley M. (2003). Modification with SUMO. A role in transcriptional regulation. *EMBO Rep* 4: 137-142.

Vermeulen M., Mulder K. W., Denissov S., Pijnappel W. W., van Schaik F. M., Varier R. A., Baltissen M. P., Stunnenberg H. G., Mann M., Timmers H. T. (2007). Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* 131: 58-69.

Voas M. G., Rebay I. (2004). Signal integration during development: insights from the Drosophila eye. *Dev Dyn* 229: 162-175.

Vogelstein B., Kinzler K. W. (2004). Cancer genes and the pathways they control. *Nat Med* 10: 789-799.

Vogt B., Hofmann K. (2012). Bioinformatical detection of recognition factors for ubiquitin and SUMO. *Methods Mol Biol* 832: 249-261.

Vojtek A. B., Hollenberg S. M. (1995). Ras-Raf interaction: two-hybrid analysis. *Methods Enzymol* 255: 331-342.

Wallberg A. E., Pedersen K., Lendahl U., Roeder R. G. (2002). p300 and PCAF act cooperatively to mediate transcriptional activation from chromatin templates by notch intracellular domains in vitro. *Mol Cell Biol* 22: 7812-7819.

Wang J., Chen L., Wen S., Zhu H., Yu W., Moskowitz I. P., Shaw G. M., Finnell R. H., Schwartz R. J. (2011). Defective sumoylation pathway directs congenital heart disease. *Birth Defects Res A Clin Mol Teratol* 91: 468-476.

Wang Q. (2009). BTB domain dimerization: Development of a protein-protein interaction assay. *PhD Thesis*.

Weber U., Siegel V., Mlodzik M. (1995). pipsqueak encodes a novel nuclear protein required downstream of seven-up for the development of photoreceptors R3 and R4. *EMBO J* 14: 6247-6257.

Weng A. P., Ferrando A. A., Lee W., Morris J. P. t., Silverman L. B., Sanchez-Irizarry C., Blacklow S. C., Look A. T., Aster J. C. (2004). Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukaemia. *Science* 306: 269-271.

Wharton K. A., Johansen K. M., Xu T., Artavanis-Tsakonas S. (1985a). Nucleotide sequence from the neurogenic locus notch implies a gene product that shares homology with proteins containing EGF-like repeats. *Cell* 43: 567-581.

Wharton K. A., Yedvobnick B., Finnerty V. G., Artavanis-Tsakonas S. (1985b). opa: a novel family of transcribed repeats shared by the Notch locus and other developmentally regulated loci in D. melanogaster. *Cell* 40: 55-62.

Wilkinson K. D. (1999). Ubiquitin-dependent signaling: the role of ubiquitination in the response of cells to their environment. *J Nutr* 129: 1933-1936.

Wilson M. D., Harreman M., Taschner M., Reid J., Walker J., Erdjument-Bromage H., Tempst P., Svejstrup J. Q. (2013). Proteasome-mediated processing of Def1, a critical step in the cellular response to transcription stress. *Cell* 154: 983-995.

Wojcik J., Boneca I. G., Legrain P. (2002). Prediction, assessment and validation of protein interaction maps in bacteria. *J Mol Biol* 323: 763-770.

Wong K. H., Todd R. B., Oakley B. R., Oakley C. E., Hynes M. J., Davis M. A. (2008). Sumoylation in Aspergillus nidulans: sumO inactivation, overexpression and live-cell imaging. *Fungal Genet Biol* 45: 728-737.

Wu J., Matunis M. J., Kraemer D., Blobel G., Coutavas E. (1995). Nup358, a cytoplasmically exposed nucleoporin with peptide repeats, Ran-GTP binding sites, zinc fingers, a cyclophilin A homologous domain, and a leucine-rich region. *J Biol Chem* 270: 14209-14213.

Wu J. Y., Rao Y. (1999). Fringe: defining borders by regulating the notch pathway. *Curr Opin Neurobiol* 9: 537-543.

Wu S. Y., Chiang C. M. (2009a). Crosstalk between sumoylation and acetylation regulates p53-dependent chromatin transcription and DNA binding. *EMBO J* 28: 1246-1259.

Wu S. Y., Chiang C. M. (2009b). p53 sumoylation: mechanistic insights from reconstitution studies. *Epigenetics* 4: 445-451.

Wuerzberger-Davis S. M., Nakamura Y., Seufzer B. J., Miyamoto S. (2007). NF-kappaB activation by combinations of NEMO SUMOylation and ATM activation stresses in the absence of DNA damage. *Oncogene* 26: 641-651.

Xie X., Kokubo T., Cohen S. L., Mirza U. A., Hoffmann A., Chait B. T., Roeder R. G., Nakatani Y., Burley S. K. (1996). Structural similarity between TAFs and the heterotetrameric core of the histone octamer. *Nature* 380: 316-322.

Xu L., Wei Y., Reboul J., Vaglio P., Shin T. H., Vidal M., Elledge S. J., Harper J. W. (2003). BTB proteins are substrate-specific adaptors in an SCF-like modular ubiquitin ligase containing CUL-3. *Nature* 425: 316-321.

Yang S. H., Jaffray E., Hay R. T., Sharrocks A. D. (2003). Dynamic interplay of the SUMO and ERK pathways in regulating Elk-1 transcriptional activity. *Mol Cell* 12: 63-74.

Yang S. H., Sharrocks A. D. (2004). SUMO promotes HDAC-mediated transcriptional repression. *Mol Cell* 13: 611-617.

Yokoyama N., Hayashi N., Seki T., Pante N., Ohba T., Nishii K., Kuma K., Hayashida T., Miyata T., Aebi U., et al. (1995). A giant nucleopore protein that binds Ran/TC4. *Nature* 376: 184-188.

Yuan H., Zhou J., Deng M., Liu X., Le Bras M., de The H., Chen S. J., Chen Z., Liu T. X., Zhu J. (2010). Small ubiquitin-related modifier paralogs are indispensable but functionally redundant during early development of zebrafish. *Cell Res* 20: 185-196.

Yunus A. A., Lima C. D. (2009). Structure of the Siz/PIAS SUMO E3 ligase Siz1 and determinants required for SUMO modification of PCNA. *Mol Cell* 35: 669-682.

Zhang F. P., Mikkonen L., Toppari J., Palvimo J. J., Thesleff I., Janne O. A. (2008). Sumo-1 function is dispensable in normal mouse development. *Mol Cell Biol* 28: 5381-5390.

Zhang H., Saitoh H., Matunis M. J. (2002). Enzymes of the SUMO modification pathway localize to filaments of the nuclear pore complex. *Mol Cell Biol* 22: 6498-6508.

Zhang Y., Wang Z., Ahmed F., Banerjee S., Li Y., Sarkar F. H. (2006). Down-regulation of Jagged-1 induces cell growth inhibition and S phase arrest in prostate cancer cells. *Int J Cancer* 119: 2071-2077.

Zhao J. (2007). Sumoylation regulates diverse biological processes. *Cell Mol Life Sci* 64: 3017-3033.

Zhao X., Blobel G. (2005). A SUMO ligase is part of a nuclear multiprotein complex that affects DNA repair and chromosomal organization. *Proc Natl Acad Sci U S A* 102: 4777-4782.

Zhong S., Muller S., Ronchetti S., Freemont P. S., Dejean A., Pandolfi P. P. (2000a). Role of SUMO-1-modified PML in nuclear body formation. *Blood* 95: 2748-2752.

Zhong S., Salomoni P., Pandolfi P. P. (2000b). The transcriptional role of PML and the nuclear body. *Nat Cell Biol* 2: E85-90.

Zunino R., Schauss A., Rippstein P., Andrade-Navarro M., McBride H. M. (2007). The SUMO protease SENP5 is required to maintain mitochondrial morphology and function. *J Cell Sci* 120: 1178-1188.

UNIVERSITAS Miguel Hernández

UNIVERSITAS
Miguel
Hernández

UNIVERSITAS
Miguel
Hernández