PART I

Studies on the *Drosophila* selector gene *apterous* and compartment boundary formation

PART II

Development of synthetic intracellular protein binders

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Dimitri Christopher Bieli

aus Selzach, Solothurn

Basel, 2016

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel

edoc.unibas.ch

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. Dr. Markus Affolter

Prof. Dr. Renato Paro

Basel, den 10. November 2015

Prof. Dr. Jörg Schibler

Dekan

SUMMARY

This PhD Thesis is divided into two different, separate parts.

In the first part, I focus on the transcriptional regulation of the *Drosophila* selector gene *apterous* (*ap*).

During animal development, selector gene activity is known to be important for the subdivision of cell populations into distinct functional units, called compartments. *ap* is essential for the subdivision into a dorsal and ventral compartment of the wing imaginal disc. This compartmentalization is a prerequisite for proper wing development. While the function of *ap* as a dorsal selector gene has been studied extensively, the regulation of its expression during wing development is poorly understood. In the presented studies, the transcriptional regulation of *ap* was analyzed by classical means and extended by novel approaches, which allowed direct manipulation of the endogenous locus. By combining all those approaches, we identified three separable *cis*-regulatory elements that work in synergy to regulate the expression of *ap* during wing imaginal disc development and gained insight into the general patterning of the wing disc and the *de novo* formation of a compartment boundary.

In the second part, I focus on the development and application of a novel class of protein binders, called nanobodies.

Protein-protein interactions are key to almost all biological processes. So far, protein functions *in vivo* have been mostly studied by genetic manipulations. However, to describe and understand protein functions in their respective native environment, it is very important and necessary to manipulate proteins directly *in vivo.* Towards this end, the discovery and development of a new class of protein binders (nanobodies) was essential. Nanobodies are protein binders based on single-domain antibody scaffolds. Conveniently, randomized nanobody libraries have been engineered that hypothetically allow the isolation of nanobodies against any protein of interest. As the field of protein binders is still very young, we wanted to explore the possibility to generate novel specific nanobodies. Using phage display, I did isolate new specific nanobodies. Importantly, we demonstrated that these new nanobodies work intracellularly in cell culture and *in vivo*.

ACKNOWLEDGMENT

First of all, I want to thank my family for the great support during all the years. This thesis is dedicated to my grand-parents, especially my grandma, who always wanted to see the first-ever "Herr Dokter" in the family. Unfortunately, she passed away during the writing of this thesis. This one is for you.

I particularly want to thank Lara and all my other friends, especially my band mates. I'm truly happy to have you in my life.

I sincerely thank Prof. Markus Affolter for giving me the opportunity to spend all these years in his great lab. His never-ending enthusiasm to solve all kinds of biological questions is encouraging. He gave me all the freedom I needed and always supported me when I wanted to start something new.

At this point, I truly want to thank Dr. Martin Müller, who was some sort of an antipole to Markus. He kicked my ass when I needed it, which was sometimes hard, but always motivated me to become better. Without Martin, a lot of things you see in this thesis would not have been possible. As a hardcore, old-school geneticist, he helped me in countless experiments.

I also want to thank Aurélien Olichon, who taught me all the tricks in phage display. I was extraordinarily happy that he helped me to organize my stays in Toulouse, with all the paper work it involved. His hospitality was outstanding. Thanks also goes to all the youngsters from Toulouse, especially Nico, Guillaume, Adrien and Remi. They gave me the feeling of being home from the first second I arrived. Thanks for the parties, Ricard, haircut, food and all the other cool stuff besides working.

I'm also very grateful to Carlos Estella, with whom I could do this very cool collaboration about the regulation of *apterous*. From the very first moment, he was open and we could share all our results. I extremely enjoyed all the discussions we had during his two stays in Basel and via e-mail. Besides that, he also has a great taste in music.

I also want to thank Alexander Harms for all the beer and punk rock. He is a true friend.

A super special thanks goes out to Shinya Matsuda, just for being great.

Special thanks also goes to Mario Metzler, who helped me a lot with computer-related issues and proved to be an excellent driver during our road-trip in the US.

I would also like to say thanks to Vahap Aydogan for helping me with fish experiments.

PART I

TABLE OF CONTENTS

INTRODUCTION

In the development of multicellular organisms, the fusion of a male and female haploid gametocytes marks the beginning of a new life cycle. This process, called fertilization, results in a diploid unicellular zygote. Multicellular individuals arise from this single cell, which has to divide and differentiate to build a functional organism with its specialized tissues and specific organs. In order to form a proper functional body, cells within an organism and tissues must coordinate their functions and duties. During development, cell proliferation and growth must be tightly regulated. Additionally, the different tissues and organs must be precisely patterned in such a way that every cell in the organism knows where it is located with respect to other cells and what function it has to fulfill at this particular location. This spatial and functional integrity of the respective cells and consequently tissues has to be coordinated and maintained for lifetime to ensure survival.

How the shape and size of an animal is regulated during development is a key question in biology. The use of model organisms has proven to be very fruitful in the pursuit of answering this question. One of the first genetically well-established multicellular eukaryotic model organism is the common fruit fly *Drosophila melanogaster*. Introduced by Thomas Hunt Morgan in 1910, it exhibits an excellent genetic toolbox, which has been refined over the last hundred years to understand the genetic and molecular basis of development. Additionally, *Drosophila melanogaster* has a –at least for eukaryotic organisms- short generation time of about 10 days, shows a high fecundity, and is very easy to handle and relatively cheap to maintain in the laboratory.

The life cycle of *Drosophila* can be subdivided into four distinct parts; involving embryogenesis of the fertilized egg, three larval stages, called instars, a pupal stage, and the adult imago (see Figure 1). As a holometabolous insect, *Drosophila* undergoes a complete metamorphosis in which almost the whole larval body is reorganized and remodeled during the pupal stage. At the end of metamorphosis an adult fly hatches from the pupal case, which then becomes fertile and ready to mate to produce another generation of flies.
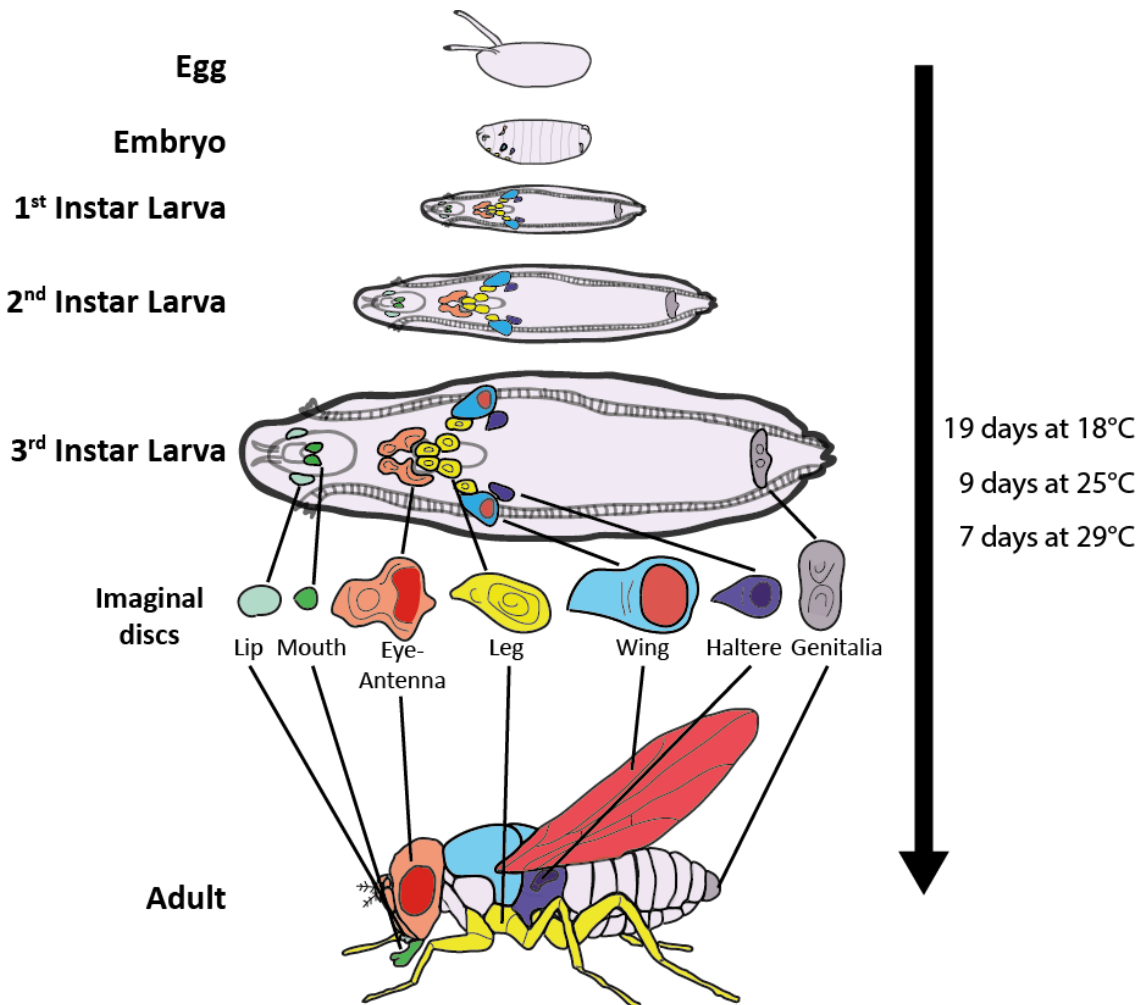
Figure 1. The development of *Drosophila melanogaster* and the role of imaginal discs. After the fertilization of the egg, and 24 h of embryogenesis (at 25°C), a 1st instar larva hatches. The larva lives in the food, grows and molds twice. After 5 days, the 3rd instar larva leaves the food and pupates. During this phase, called metamorphosis, most of the larval body is degenerated and the adult body is built from imaginal disc and imaginal histoblasts. These imaginal cells do not contribute to the larval body and are not essential for survival of the larvae. The adult fly hatches after 4 days from the pupal case. Male and female flies mate, and upon copulation, the female again lays fertilized eggs on the food, restarting a new life cycle starts of a new generation.

During embryogenesis, special 'imaginal' cells are set aside. These imaginal cells are the primordia of the adult fly body, which is built up during metamorphosis. The cells which will form the head, thorax, genitalia and the appendages, such as legs or wings, are organized as monolayered epithelial sacs, called imaginal discs (COHEN 1993) (Figure 1). These discs are comprising two different cell layers; the disc proper, which will form the principal adult structures, and the peripodial membrane, which forms the integumentary cuticle of the body wall. The abdomen and the internal organs, such as the salivary glands, gut, trachea, and brain form from more loosely connected histoblast nests (CURTISS and HEILIG 1995).

During the larval stages, the imaginal discs grow without contributing functionally to the larval body, and are dispensable for survival. However, to form a proper adult body during

11

metamorphosis, imaginal discs must be tightly organized and patterned. Due to the growth and patterning events occurring in the imaginal discs, they represent an excellent model to study tissue organization and organogenesis. Moreover, they can be manipulated without directly affecting the survival of the *Drosophila* larvae.

By the end of third larval stage, the imaginal discs are patterned to such an extent that a fate map can be extrapolated (HELD 2005). In the case of the imaginal wing disc, the tissue is also responsible for the formation of the notum, scutellum, wing hinge, and pleura (Figure 2). Thus, the wing disc is sometimes more correctly called dorsal metathoracic disc, since it is responsible for the formation of all dorsal metathoracic structures. Furthermore, it is possible to predict which region will be posterior, anterior, dorsal or ventral in the adult. The notum and scutellum will form from the most dorsal regions of the disc. The actual wing primordium is located in the center of the disc, which is termed the wing pouch. From regions that encircle the pouch, the wing hinge will develop. During metamorphosis, the wing blade evaginates from the wing pouch, and the monolayered, two-dimensional wing disc forms during the pupal stage the adult wing as well as the other dorsal metathoracic structures. As a result, the wing blade consists of dorsal and ventral cells of the wing pouch, which touch each other at their basal side forming a two-cell-layered sheet.
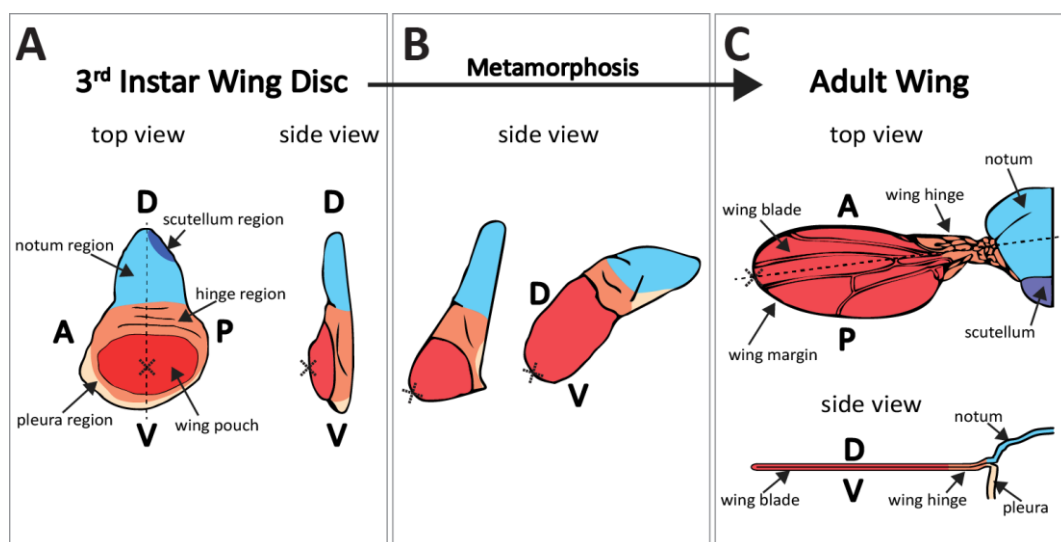
Figure 2. Fate map of *Drosophila* wing imaginal disc. (A) 3rd instar wing imaginal disc from top and side views. The regions of the future notum is colored in light blue, the future scutellum is in dark blue, the hinge region is in orange and the pleura region in light yellow. The wing pouch, the region of the future wing, is colored in red. The dotted cross in the middle of the wing pouch will form the most distal tip of the adult wing. D stands for dorsal, V for ventral, A means anterior and P posterior. **(B)** During metamorphosis, the wing pouch evaginates, resulting in the merge of the dorsal and ventral cells. **(C)** Top and side view of an adult wing. The wing is consisting of a two-cell-layered sheet of dorsal and ventral cells. The wing hinge (orange) is connecting the wing blade to the body wall, consisting of notum, scutellum and pleura. The pleura is connected to the ventral body wall, which originates from the leg disc (not shown). Note: discs and adult structures are not to scale.

Compartment and compartment boundaries

The concept of compartmentalization

Cells in a developing multicellular organism do not only divide, grow and differentiate, but they also sort out into and within different tissues. This sorting out subdivides the body and the different tissues into so-called compartments. This subdivision, or compartmentalization, of the tissue results in the juxtaposition of different compartments, each consisting of cells with unique properties. In order to keep the integrity of the tissue, cells from one compartment do not intermingle with cells from another compartment (DAHMANN *et al.* 2011). This leads to a straight interface between two adjacent compartments, which is termed compartment boundary (DAHMANN and BASLER 1999). Having a cell segregation mechanism ensures the integrity of the tissue during growth and morphogenesis (Figure 3).

Additionally, short-range signaling events between the compartments specify cells close to the boundary region. Cells in this region, also called organizers, play an important role in the patterning of the surrounding tissue by secreting long-range signaling molecules (morphogens). The concept of morphogen gradients, also known as the French flag model, was introduced as a hypothetical model by Wolpert in 1969 to explain the patterning of an organ or a tissue (WOLPERT 1969). In this model, an extracellular substance disperses from a localized, stable source, creating a concentration gradient that provides a series of thresholds, which elicit distinct cellular responses at different distances from the source. The morphogen gradient provides positional information to the cells in a tissue, leading to an even further subdivision of the tissue. This refinement can then result in further compartmentalization of the tissue. As a result, new compartments form, which possibly interact with each other at their boundaries to induce secondary organizers (MEINHARDT 1983). Thus, the formation of compartments and morphogen gradients can be seen as recurring themes and integral part of pattern formation in the development of multicellular organisms (IRVINE and RAUSKOLB 2001). Therefore, the formation and maintenance of the compartments and compartment boundaries is fundamental to development.

The observation of compartment boundaries established a new model in which the subdivision of cells into distinct populations could be viewed as a cause, rather than a consequence, of their development (IRVINE and RAUSKOLB 2001).

A

Proliferation   Morphogenesis

Intermingling   Segregation   Intermingling   Segregation

B

Short-range signaling   Short-range signaling

Wiggly organizer   Stable, straight organizer

C

Organizer

Morphogen
gradient

Target gene response
Mutual respression

Formation of new
compartments

14
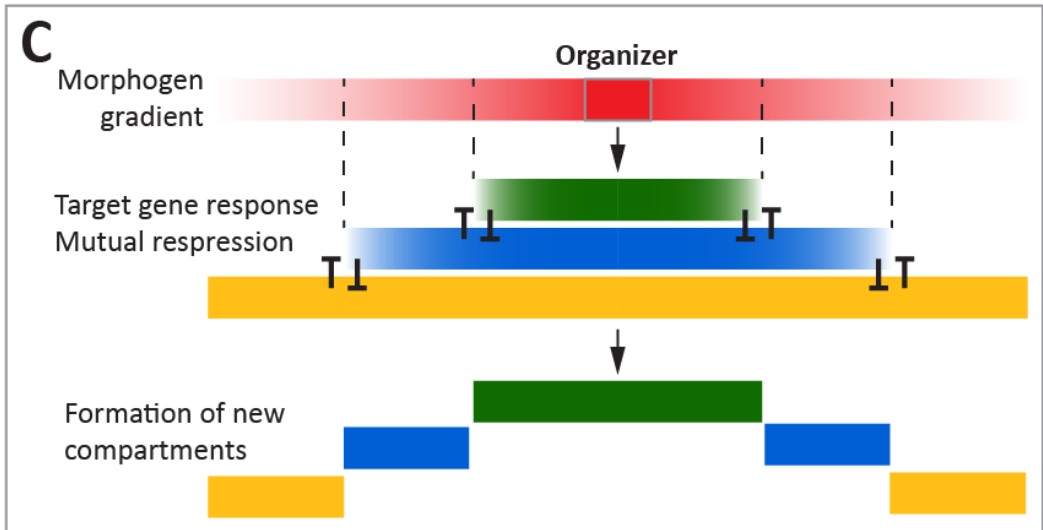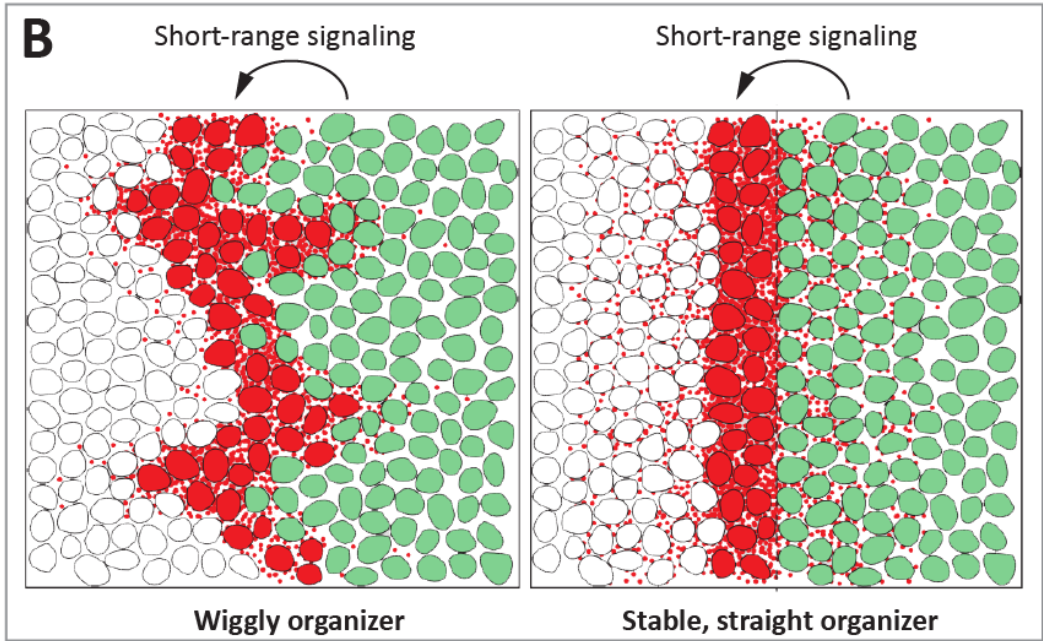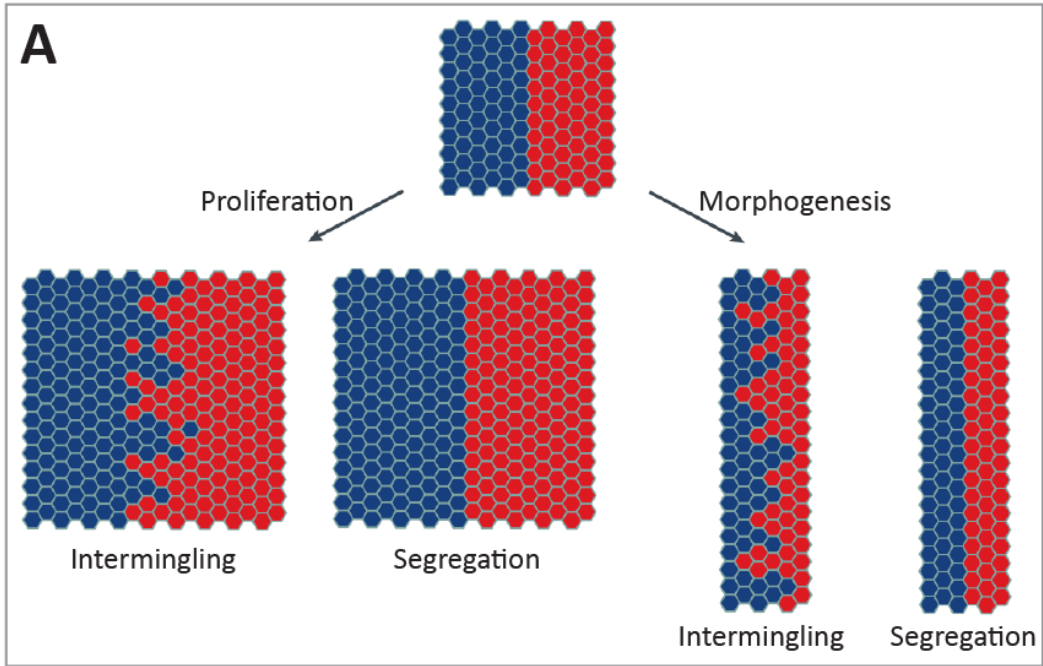
Figure 3. Concept of compartments, compartments boundaries and morphogens in development. (A) Two groups of cells (compartments, blue and red) form a straight, sharp compartment boundary. Proliferation and morphogenesis can result in the intermingling of the two compartments, if the two cell populations do not possess a cell segregation system. Modified from DAHMANN *et al.* 2011. **(B)** Short-range signaling across the compartment boundaries define cells at the boundary (red) to produce and secrete long-range signaling molecules (morphogens), which will pattern the tissue on a dose-dependent manner. If there is no straight boundary (left), the resulting organizer will be wiggly, and the subsequent patterning of the tissue is imprecise, leading to developmental malformations. However, a straight and stable compartment boundary (right) ensures the correct positioning of the organizer, and subsequently the correct pattern formation of the tissue. Modified from DAHMANN and BASLER 1999. **(C)** Morphogen gradient can induce different target genes at different threshold. If the target genes mutually repress each other, new compartments can form.

Features of compartments and selector genes

The concept of compartmentalization is largely based on observations made in *Drosophila*. Initially, compartment boundaries were discovered by clonal marker analysis in the fruit fly as groups of cells that give rise to a certain portion of the adult cuticle (GARCIA-BELLIDO *et al.* 1973). By investigating genetic mosaics, it became apparent that most of the time clones contributed to inconstant and irregularly shaped regions of the adult cuticle. In some cases, however, a smooth, straight boundary was observed that did not show any morphologically visible landmarks. Marked cells were not able to cross this boundary, but remained restricted to either side of the compartment boundary. Thus, compartments were defined as groups of cells that do not intermingle with each other, forming a straight boundary at the interface.

With further investigations it became apparent that compartments and their boundaries are defined by the activity of a special group of transcription factors, the so-called selector genes (GARCIA-BELLIDO 1975). The expression of a selector gene gets induced and maintained in cells that will contribute and form a particular compartment. The classical definition of a selector gene was based on the analysis of the genes *Ultrabithorax* (*Ubx*) and *engrailed* (*en*) (GARCIA-BELLIDO 1975). They fulfill the following properties: First, their functional domain is limited by the compartment boundary; second, they act in combination with other selector genes; third, their function is cell-autonomously; fourth, they provide cells with unique adhesion properties; and fifth, they induce signaling events at the compartment boundaries (DAHMANN and BASLER 1999). As already mentioned, these signaling events across the compartment boundaries were demonstrated to induce signaling centers, termed organizers.

The *Drosophila* wing disc served as a paradigm to study the theory of compartmentalization. In this case, the tissue is subdivided into anterior-posterior (A/P) and dorsal-ventral (D/V) compartments (Figure 4). At the compartment boundaries, organizers are defined that secrete

morphogens encoded by *decapentaplegic* (*dpp*) and *wingless* (*wg*) (DIAZ-BENJUMEA and COHEN 1993, 1995; ZECCA *et al.* 1995; NELLEN *et al.* 1996; LECUIT *et al.* 1996; NEUMANN and COHEN 1997) (see below for more detailed information). In addition to the compartments in the wing, the embryonic ectoderm, abdomen, legs, head as well as the proboscis were shown to be compartmentalized (STEINER 1976; LAWRENCE *et al.* 1978; MORATA and LAWRENCE 1978; STRUHL 1981).
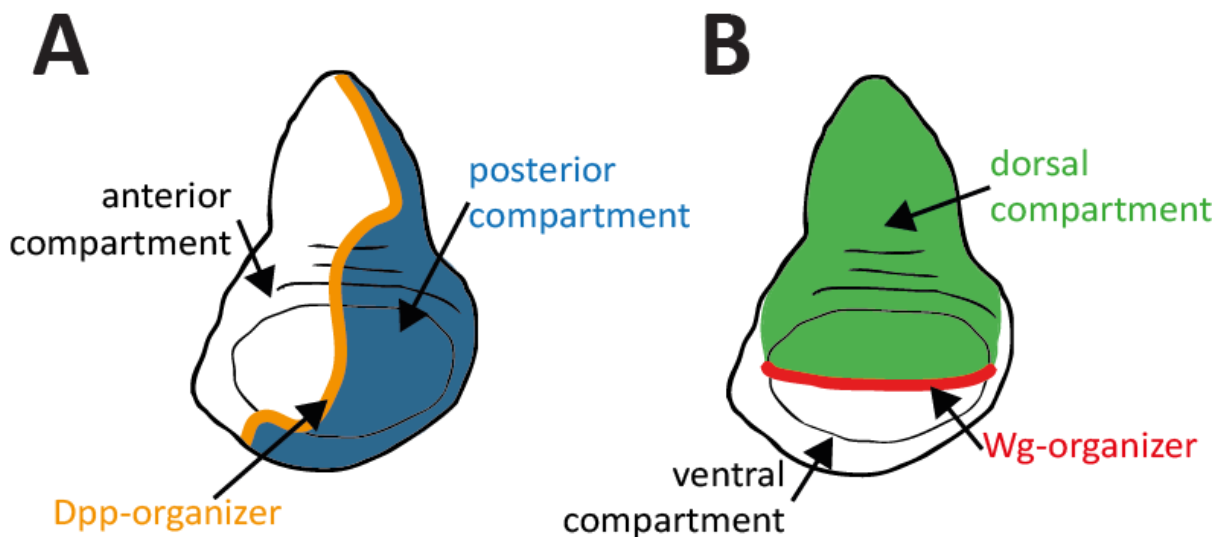
Figure 4 Compartmentalization of the 3rd instar *Drosophila* wing disc. (A) From the beginning of wing disc development**,** the tissue is subdivided into a posterior and anterior compartment. As a selector gene, *engrailed* is expressed in the entire posterior compartment. Cells from the anterior compartment close to the posterior cells are induced by short-range Hedgehog signaling to produce the Dpp morphogen. **(B)** During development, the disc gets also subdivided into a dorsal and ventral compartment (see below for more detailed information). The selector gene *apterous* is expressed in the dorsal compartment. Cells from each side of the D/V boundary are instructed to produce Wg, which will act as an organizer. Together with Dpp, Wg induces patterning and growth of the wing pouch.

In summary, compartments are functional units of cell populations that are demarcated by the activity of selector genes. To keep compartment integrity, selector genes regulate genes important for differentiation, genes that control the cell interactions at the compartment boundary and genes necessary for providing exclusive adhesion properties to the cells (see below). Moreover, compartments define new landmarks in a tissue, which subsequently can be used for patterning.

Compartments in vertebrates

After the discovery of compartments in *Drosophila*, researchers have also found compartments in vertebrate species. The first evidence for compartments was found in the subdivision of the developing chick hindbrain into morphologically distinguishable units, called rhombomeres (FRASER *et al.* 1990). Successively, rhomobomeres were also observed in mice (JIMENEZ-GURI *et al.* 2010). Moreover, the mid-hindbrain boundary displays a nice example for a compartment boundary in vertebrates (ZERVAS *et al.* 2004; LANGENBERG and BRAND 2005). Additionally to the rhombomeres, the developing mouse brain shows compartment boundaries in more anterior parts, like the zona limitans intrathalamica and in the telencephalon (INOUE *et al.* 2001; ZELTSER *et al.* 2001). Compartments have also been described outside of the developing brain, for example in limb buds of mice and chicken (ALTABEF *et al.* 1997; ARQUES *et al.* 2007; PEARSE *et al.* 2007; QIU *et al.* 2009). Cell-linage tracing experiments with retroviruses have shown that the growing chick gut is partially compartmentalized (SMITH and TABIN 2000). Furthermore, studies on the developing mouse presomitic mesoderm and the somites have revealed that each somite is subdivided into a rostral and a caudal compartment (TAM *et al.* 2000). Interestingly, it has been shown that the transcription factor MESP2 has features of a selector gene in the classical sense. Its expression domain correlates with the size of the rostral compartment and loss of MESP2 function results in the caudalization of the somites (SAGA *et al.* 1997; TAKAHASHI *et al.* 2000).

In addition, cells with organizing activity have also been observed along compartment boundaries in vertebrates. For example, specialized cells at the mid-hindbrain boundary express and secrete Wnt-1 and fibroblast growth factor 8 (FGF8) ligands, which are implicated in the pattern formation of this tissue (SIMEONE 2000; JOYNER *et al.* 2000). Another example is the apical ectodermal ridge (AER). The position of the AER corresponds to the boundary of the dorsal and ventral lineages in the developing limb bud. FGFs secreted from the AER are responsible of the development and outgrowth of the limb bud (MARTIN 1998).

Interestingly, it has been demonstrated that the general gene network involved in the D/V compartmentalization in the *Drosophila* wing and in the compartment formation in the vertebrate hindbrain is conserved to a large extend (BUCETA *et al.* 2007).

Altogether, the general concept of compartmentalization, with organizers as well as selector genes, seems to be conserved from invertebrates to higher mammals.

Maintenance of compartment integrity

In order to keep tissue integrity and a stable, precise organizer source, it is extremely important that compartments and their boundaries are maintained.

In addition to the lineage restriction, cells from each compartment are predicted to possess differential adhesion properties. Special proteins on the cell surface interacting with surface proteins on other cells have been shown to mediate cell-cell adhesion. The most prominent players in mediating differential cell adhesion are the 'cadherins' (Nose *et al.* 1988; Steinberg 2007). In this model, cells that express one type of cadherin segregate with cells that express the same type of cadherin (homotypic interaction). However, two cell populations expressing different cadherins (heterotypic interaction) causes a difference in affinity at the interface of the two populations. To minimize the tension between the two cell populations, the interface area is kept at its minimum, resulting in the formation of a straight boundary (Batlle and Wilkinson 2012). This is nicely illustrated by the differential expression of R-cadherin and cadherin-6 in the mouse telencephalon. While R-cadherin is expressed in one compartment, cadherin-6 is expressed in the other. At the interface of the two compartments a border forms, the cortico-striatal compartment boundary. Misexpression of the two surface proteins resulted in the intermingling cells from each compartment, ultimately leading to the disruption of the boundary (Inoue *et al.* 2001).

Another mechanism that drives cell sorting is based on differential mechanical tension at the boundary. It has been shown that mechanical tension is generated by contractile elements at the cell cortex. Therefore, actomyosin-based filaments become locally enriched at the compartment boundary. In *Drosophila*, this phenomenon has been observed at the parasegment boundaries of the embryonic epidermis (Monier *et al.* 2010). Additionally, actomyosin filaments are seen at the A/P as well as D/V boundaries in the wing disc (Major and Irvine 2005; Landsberg *et al.* 2009). The contraction of these filaments is thought to help cell sorting, thus facilitating the formation of clear and straight compartment boundaries (Brodland 2002).

Tissue growth is mostly accompanied with increased cell proliferation. However, cell divisions and the resulting cell rearrangements can perturb straightness of compartment boundaries (see Figure 3A). At the D/V boundary of the *Drosophila* wing disc, a strong reduction of cell proliferation as has been reported (O'Brochta and Bryant 1985). In vertebrates, a reduction of cell proliferation has also been observed at the rhombomere boundaries in the developing

chick hindbrain (GUTHRIE *et al.* 1991). These zones of non-(low)-proliferating cells have been proposed to contribute to the sharpness and maintenance of the compartment boundaries.

A further mechanism that is associated with cell segregation into different compartments is the bi-directional signaling between Ephrin receptors and their respective ephrin ligands. For example, in the segregation of rhombomeres in vertebrates, Ephrin signaling has been demonstrated to restrict intermingling by repulsion of two different cell populations (XU *et al.* 1999; MELLITZER *et al.* 1999). Therefore, adjacent rhombomeres express Ephrin receptors and their ligands in a reciprocal manner. Also during somitogenesis , the reciprocal patterns of Ephrin receptors and ligands has been reported to actively contribute to cell segregation (BARRIOS *et al.* 2003). Thus, the active repulsion mediated by Ephrin signaling is thought to contribute to the integrity of compartment boundaries. However, so far this mechanism of active repulsion for boundary formation has not been reported in *Drosophila*.

During development, tissue and cellular rearrangements, such as convergence extension, play an important role. The forces generated by these rearrangements represent a challenge for the formation as well as the maintenance of compartment boundaries (see Figure 3A). To strengthen cell segregation, boundaries between rhombomeres and somites have been shown to rely on the deposition of extracellular matrix (HEYMAN *et al.* 1993; KOSHIDA *et al.* 2005). Thereby, a fibronectin-based extracellular matrix forms a physical barrier, which helps to minimize mixing of different cell populations.

Altogether, the differential cell adhesion, mechanical tension, reduced cell proliferation, active repulsion and the extracellular matrix are important for cell segregation at compartment boundaries. These mechanisms ensure tissue integrity, which is absolutely essential for the correct positioning and maintenance of the organizer throughout growth and development.

The selector gene *apterous*

As previously illustrated, the Drosophila wing disc is divided into A/P and D/V compartments. The generation of the A/P and D/V compartments is instructed by the activity of the selector genes *en* and *apterous* (*ap*), respectively (LAWRENCE and STRUHL 1982; DIAZ-BENJUMEA and COHEN 1993; TABATA *et al.* 1995). Investigations of *ap* function during the development of the wing imaginal disc of *Drosophila* have demonstrated that it is a selector gene by the classical

definition (see previous chapter). First, its expression and functional domain is limited to the dorsal compartment (Cohen *et al.* 1992). Second, together with other selector genes, such as *en* and *Ubx*, it has an instructive role in patterning and growth of the wing and haltere discs (Guss *et al.* 2001). Third, mitotic recombination experiments confirmed the cell-autonomous action of Ap (Diaz-Benjumea and Cohen 1993). Fourth, the segregation of dorsal and ventral cells is partially mediated by the action of the transmembrane proteins Tartan (Trn) and Capricious (Caps), whose expression pattern in the wing disc largely coincide with early *ap* expression (Milán *et al.* 2001). Additionally, the PS1 and PS2 integrins are expressed in dorsal-specific and ventral-specific patterns, respectively, also controlled by Ap activity (Blair *et al.* 1994). Together with Trn and Caps, they prevent the intermingling of cells at the dorsal/ventral (D/V) compartment boundary. Fifth, Ap activity induces a signaling cascade that positions a Wg-organizer at the D/V boundary (see below).

The mature Ap protein is a transcription factor composed of two LIM-domains and one homeodomain (Cohen *et al.* 1992) (Figure 5A). LIM domains are important to mediate protein-protein interaction (Feuerstein *et al.* 1994), whereas the homeodomain is conferring sequence-specific DNA binding (Gehring *et al.* 1994). The activity of Ap depends on complex formation with the LIM-domain binding protein Chip (Fernández-Fúnez *et al.* 1998; O'Keefe *et al.* 1998; Van Meyel *et al.* 1999; Milán and Cohen 1999; Rincón-Limas *et al.* 2000; Pueyo and Couso 2004).
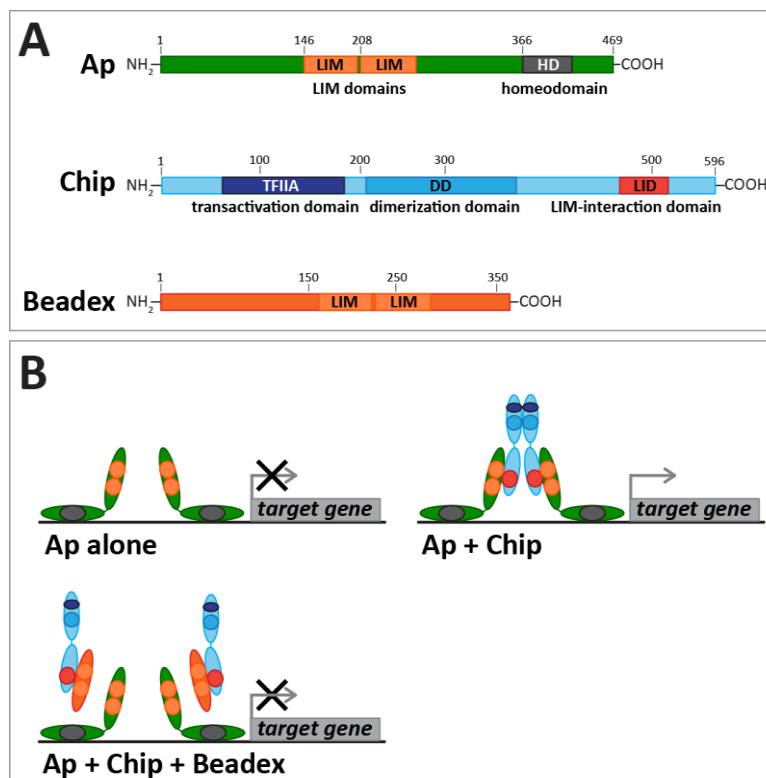
Figure 5 Structure and function of Ap protein complexes. (A) The mature Ap protein has two LIM domains and a C-terminal homeodomain (HD). Chip has transactivation, a dimerization (DD), and LIM-interaction domains (LID). The dLMO protein Beadex contains only LIM domains resembling the ones from Ap. **(B)** on its own, Ap can probably bind DNA, but is unable to activate transcription. The protein Chip interacts with the LIM domains of Ap with its LIM-interaction domain (LID). A tetrameric Ap-Chip complex is formed via the dimerization domains (DD) of Chip. This complex is activating transcription via the recruitment of the general transcription factor IIA (TFIIA). The LIM domains of Beadex are competing with Ap to bind to Chip. High levels of Beadex are interfering with ap-Chip complex formation, rendering Ap to be transcriptionally inactive.

On the other hand, the Drosophila LIM-only protein Beadex (Bx) acts as a repressor of the Ap-Chip complex activity (MILÁN *et al.* 1998; WEIHE *et al.* 2001; BRONSTEIN *et al.* 2010a). As expected, the interaction between Ap, Chip and Bx and the resulting Ap activity are very dynamic and depend on the relative amount of each protein (Figure 5B).

Interestingly, evo-devo experiments have found an evolutionary conserved role for *apterous* in development. The human orthologue of *ap, hLhx2*, is able to rescue the *ap* mutant phenotypes as effectively as the endogenous fly protein (RINCÓN-LIMAS *et al.* 1999). Furthermore, the investigation of expression pattern of *mLhx2*, the mouse orthologue of *ap*, revealed a remarkable similarity in the respective expression patterns (RINCÓN-LIMAS *et al.* 1999). In chicks, the *ap* orthologue *C-Lmx1* specifies dorsal cell fate of the developing limb bud, and misexpression of *C-Lmx1* in ventral mesoderm of the bud can transform ventral cell fate to dorsal (VOGEL *et al.* 1995). These results suggest that the function as well as the expression pattern of *ap* is evolutionary conserved across phyla.

The role of Ap in the development of the *Drosophila* wing disc

In contrast to the A/P wing division, which is established during embryonic development, Apterous activity subdivides the wing disc into D/V compartments while it is growing (DIAZ-BENJUMEA and COHEN 1993). *ap* expression is initiated during early second instar and subsequent Ap activity in the dorsal compartment of the wing discs of second and early third instar larvae induces the expression of *Serrate* (*Ser*) (BACHMANN and KNUST 1998) (Figure 6). The transmembrane protein and Notch ligand Serrate acts then as a short-range signal, which signals via Notch receptors to adjacent ventral cells at the D/V compartment boundary, subsequently inducing *wingless* (*wg*) and *Delta* (*Dl*) expression in those cells (KIM *et al.* 1995). Dl is another Notch ligand, via which the ventral cells signal back to dorsal cells, thereby inducing *wg* and maintaining *Ser* expression in dorsal cells. The resulting feedback mechanism is now thought to be independent of Ap function, and maintains *wg* expression during further development (RULIFSON and BLAIR 1995). Wg, a ligand of the Wnt family, regulates the expression domains of its target genes, *Distalless* (*Dll*) and *vestigial* (*vg*), which pattern the wing disc along the dorso-ventral axis and are responsible for wing identity (NEUMANN and COHEN 1997). Although its essential role in wing development is undisputed, the mode of action of Wg as a classical morphogen is currently questioned (ALEXANDRE *et al.* 2014).

Ap also induces expression of *fringe* (*fng*), whose protein product makes the dorsal cells refractory to their own Ser signal, but at the same time more sensitive to Dl signals from the ventral cells (PANIN *et al.* 1997). Furthermore, Ap induces the expression of *Bx*, which negatively regulates the activity of Ap (see Figure 5B), thus the Notch signaling is thought to become uncoupled from the initial Ap activity, resulting in symmetrical Notch feedback signaling (MILÁN and COHEN 2000).
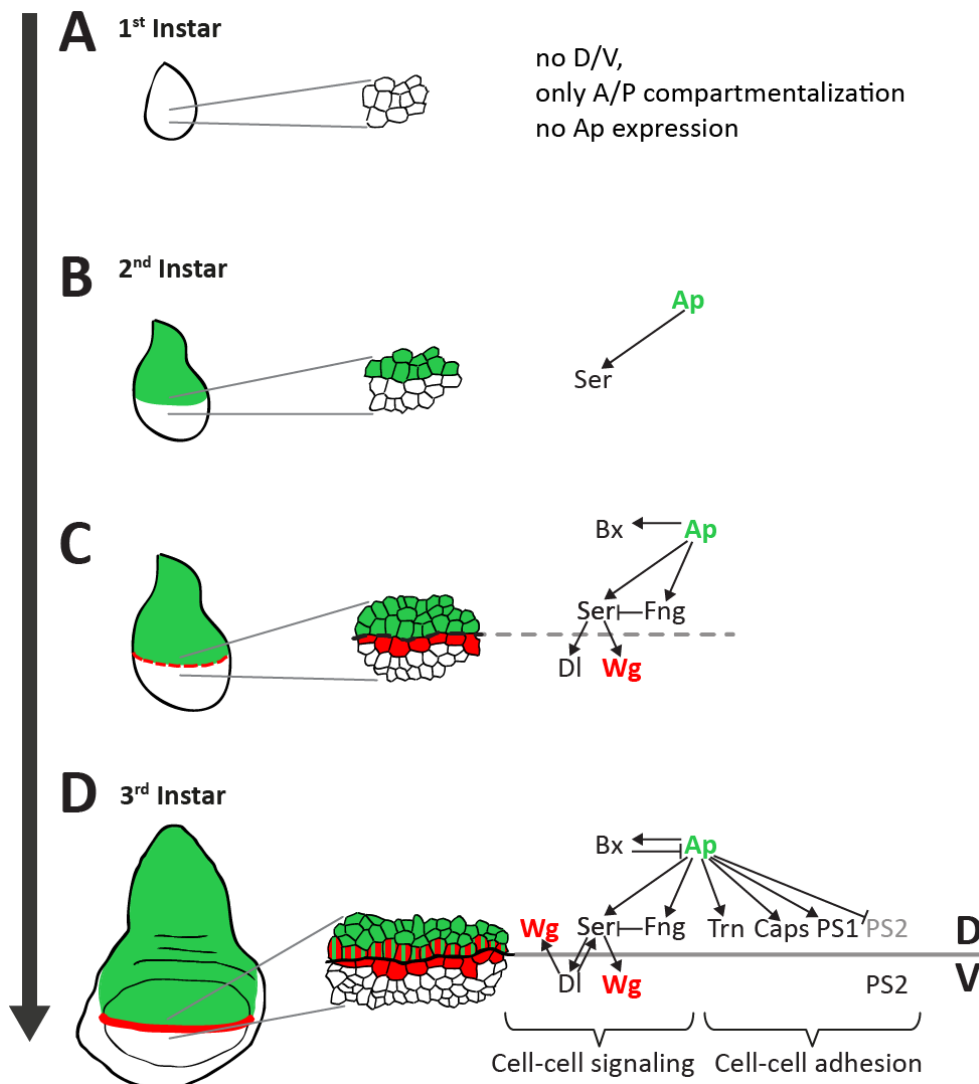
Figure 6 Apterous function and formation of D/V compartment boundary in the wing disc. (A) After embryogenesis and during the 1st instar larval stage, the wing disc is only subdivided into an anterior and posterior compartment (not shown here, see Figure 4). **(B)** At early-mid 2nd instar stage, Ap starts do get expressed in the dorsal portion of the developing wing disc (green). Subsequently, Ap induces Serrate (Ser) expression. As a Notch ligand, Ser signals to adjacent non-Ap-expressing cells to induce Delta (Dl) and Wingless (Wg, red). At the same time, Ap-positive cells are refractory to the induced Ser signal, because of the action of Fringe (Fng). Ap also induces Beadex (Bx), which is a negative regulator of Ap activity. (D) Ventral cells signal back to dorsal (Ap-positive) cells via Dl, where they maintain the expression of Ser and also induce Wg. This way, Wg is expressed on the dorsal as well as on the ventral side of the compartment boundary. To maintain tissue integrity, Ap also induces specific cell adhesion molecules (Tartan (Trn), Capricious (Caps), and PS1 integrin), On the other hand, Ap represses the expression of PS2 inegrin in dorsal cells, Thus PS2 is only expressed ventrally. The exact timing of induction these adhesion molecules has not been analyzed. Hence, only their respective expression patterns in late 3rd instar wing discs are displayed.

Taken together, the signaling cascade initiated by Ap ensures that the position of the organizer is stable in cells close to the D/V boundary, which is a prerequisite for normal wing development.

In contrast to the signaling events across the A/P boundary, which induces the Dpp organizer exclusively in the anterior compartment, short-range signaling at the D/V boundary is bidirectional. Hence, cells with organizing activity are found on either side of the compartment boundary. A possible explanation for this is that, unlike the anterior and posterior compartments, the ventral and dorsal part of the wing must have exactly the same size and shape to form a functional wing. Therefore, the organizer (Wg source) is positioned exactly symmetrically along the D/V boundary (DAHMANN and BASLER 1999).

Phenotypes of *apterous* mutants

Mutations in *ap* result in a variety of phenotypes. The most striking morphological defect in strong mutant (amorphic) alleles is the complete lack of wing (wing blade and hinge) and haltere structures (BUTTERWORTH and KING 1965). Based in this obvious and striking phenotype, the gene got its name *apterous* from the Greek words a- (without/not) -pteros (having wings). In the *Drosophila* wing, different *ap* alleles can lead to various phenotypes, regarding their severity and penetrance (Figure 7). The phenotypes can range from complete loss of all wing structures, to tiny wing stumps, and blistering of the wing, to only minor notching of the wing blade (GOHL *et al.* 2008). In general, mutations in the *ap* gene have been reported to be recessive, at least for Ap function in wing development. However, one dominant allele (*ap^Xasta^*) was isolated, causing a severe, fully penetrant notching of the distal wing blade (SEREBROVSKY and DUBININ 1930).
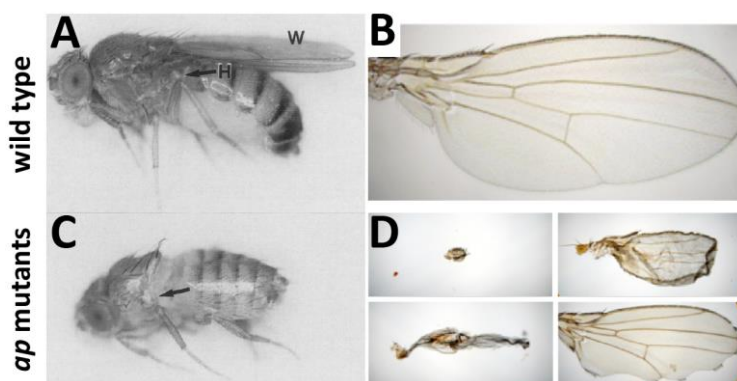
Figure 7 Phenotypes of *apterous* mutants. (A) Wild type fly, with normal wing (W) and haltere (H) structures. **(B)** Overview of a wild type *Drosophila* wing. **(C)** Amorphic *ap* mutants show loss of all wing and haltere structures (arrow)**. (D)** Variable wing defects observed in hypomorphic *ap* alleles. Phenotypes range from tiny wing stumps (top left), to notching of the wing blade (bottom right). Note: not to scale with wing depicted in **(B).** Image credit: **(A and C)** COHEN *et al.* 1992, **(B and D):** GOHL *et al.* 2008**.**

ap null mutant wing discs are characterized by the complete absence of the wing pouch (Zecca and Struhl 2007a). Interestingly, *ap* is expressed throughout the whole dorsal compartment of the wing disc (see below), including the future notum and scutellum. However, in wing discs of *ap* mutants these regions are largely unaffected and the notum and scutellum structures – apart from minor defects in bristle patterns- seem to develop normally in adult flies.

In addition to the wing phenotypes, *ap* mutants show other morphological as well as physiological defects. *ap*-null mutants survive embryogenesis, the larval stages and the pupal stadium and can eclose successfully; however, they die within one to two days after eclosion (Cohen *et al.* 1992). Mutant embryos also display neuronal fasciculation defects (Lundgren *et al.* 1995) and are characterized by the absence of distinct sets of muscles (Bourgouin *et al.* 1992). Another physiological defect is the decreased production of juvenile hormone (JH) (Ringo *et al.* 1991; Altaratz *et al.* 1991). The lack of JH was used to explain that *ap* mutant flies fail to undergo fat body histolysis and the observed female sterility, caused by non-vitellogenic eggs. Since *ap* expression is not detected in the tissue where JH hormone is produced, it is suggested that *ap* has an indirect effect, causing defects in nervous system structure and function and ultimately JH production (Cohen *et al.* 1992; Lundgren *et al.* 1995).

Endogenous *apterous* expression pattern in larval structures

By using a P-element *LacZ* enhancer trap that inserted into the *ap* promoter (*ap^{rk568}*), the endogenous expression pattern of *ap* in the imaginal discs and larval brain could be assayed (Cohen *et al.* 1992). In the wing and haltere discs of third instar larvae*, ap* was shown to be expressed throughout the regions that give rise to the dorsal structures, i.e. the dorsal compartment. The highest expression is observed in the region of the future dorsal hinge region, lower expression is detected in the regions of the disc that form the dorsal thoracic body wall, the notum and scutellum (Figure 8A). In all leg discs, *ap* is expressed in a ring in the region of the presumptive fourth tarsal segment. Additionally, *ap* expression is detected in a central spot in the antennal part of the eye-antennal disc (Figure 8B). In the larval brain, *ap* expression is seen in a complex pattern in the brain and in repeated clusters in the ventral nerve cord (VNC; Figure 8C).
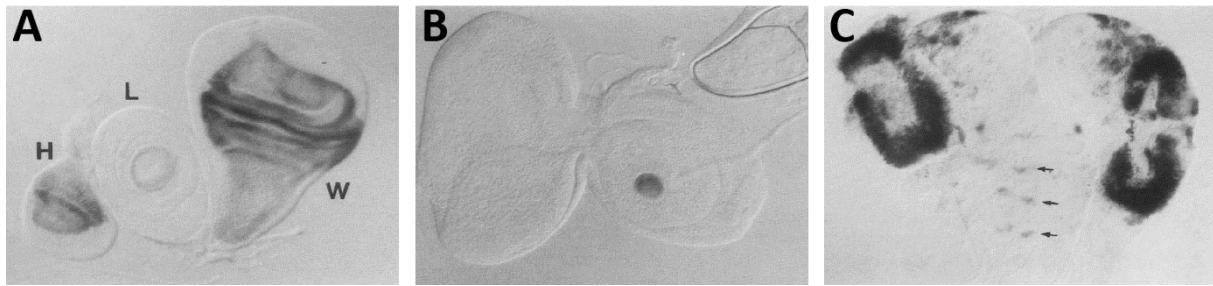
Figure 8 Endogenous *ap* expression pattern in imaginal discs and larval brain. β-Galactosidase activity patterns represent endogenous *ap* expression in the *LacZ*-enhancer trap line *ap^rk568*. **(A)** In the haltere (H) and wing disc (W), *ap* is expressed in the dorsal compartment. A ring of expression is observed at the region of the fourth tarsal segment. **(B)** In the eye-antennal disc, *ap* is expressed in a spot in the middle of the future antenna. **(C)** Complex *ap* expression pattern in the larval brain and clusters in the VNC. Modified from Cohen *et al.* 1992**.**

Regulation of *apterous* expression in the wing imaginal disc

It has been demonstrated that the activation of the epidermal growth factor receptor (EGFR) is necessary and sufficient to activate the expression of *ap* in the dorsal compartment of the wing disc (Zecca and Struhl 2002a). For the activation of the pathway, the EGF ligand Vein, which is expressed in the dorsal part of early second instar wing discs (Klein 2001), is required (Zecca and Struhl 2002b). However, by using a temperature-sensitive mutant of EGFR, it was shown that EGFR signaling is only required during early wing disc development to induce *ap*, later inactivation of the pathway had no effect on *ap* expression (Wang *et al.* 2000). These results were confirmed by ectopic activation of the pathway, where ectopic *ap* activation was restricted to early stages of wing development (Zecca and Struhl 2002a). Furthermore, early ventral *wg* expression correlates with the restricted *ap* expression in the dorsal part of the developing wing disc (Williams *et al.* 1994; Klein 2001).

At the molecular level, a 7.7 kilobase (kb) enhancer fragment was isolated, which has been reported to drive reporter gene expression in the wing disc similar to the endogenous *ap* expression pattern (Lundgren *et al.* 1995). This fragment, termed apC, resides about 7 kb upstream of the transcription start site of *ap* (see Figure 9). In addition to the wing disc expression, apC drives reporter gene expression in the *ap* neurons of the VNC. *Df(2R)ap^DG*, a 27 kb deletion, which includes the apC region but leaves the promoter intact, is homozygous viable and adult flies lack all wing structures (Gohl *et al.* 2008). These observations indicate that essential *ap* wing enhancers are uncovered by *ap^DG* and that the apC interval possibly

contains relevant *cis*-regulatory elements involved in the transcriptional regulation of *ap* expression.

A
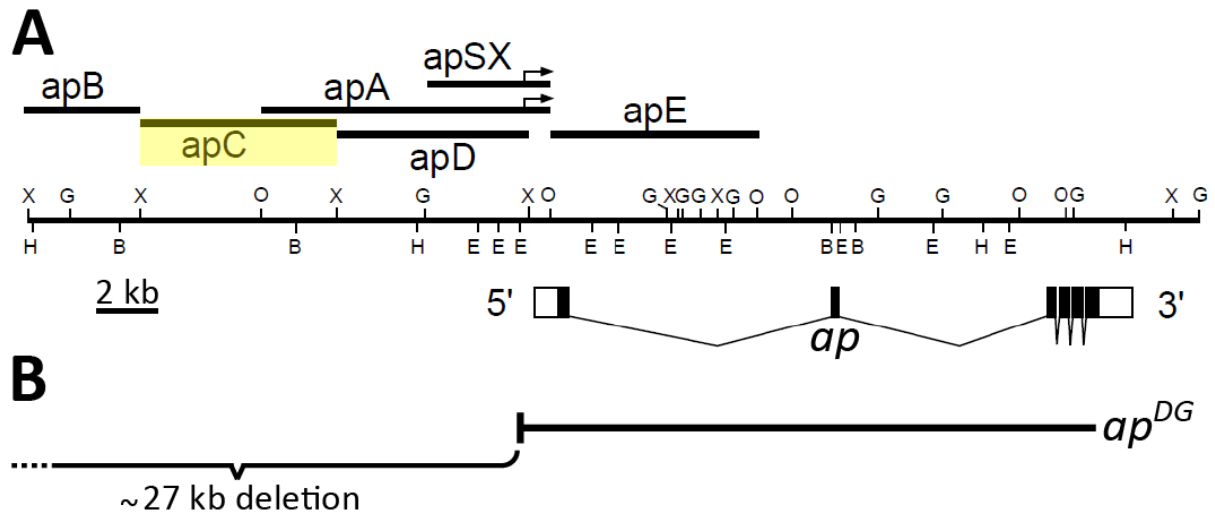
Figure 9 Molecular map of the *apterous* locus. (A) Various fragments already tested for their ability to drive reporter gene expression in an *ap*-specific pattern. Only apC (highlighted in yellow) was reported, but not shown, to recapitulate endogenous *ap* expression pattern in the wing imaginal disc. Map modified from Lundgren *et al.* 1995. **(B)** *ap^{DG}* is a deletion in the *ap* locus that removes approximately 27 kb of the upstream region, but leaves the open-reading frame and promoter of *ap* intact. Homozygous flies of this genotype were shown to lack all wing and haltere structures, similar to other complete *ap* loss-of-function alleles.

The concept of cellular memory and PRE/TREs

During development, many decisions, such as cell fate commitment, are made in response to transient positional signals. To form a functional body with all the various cell types, cells in an organism must ''remember'' where they are and what their respective function is. This process is often referred to as ''cellular memory''. Most importantly, the genes which confer the respective cellular identity, such as the selector genes, must 'know' whether they are expressed or repressed. This particularly applies to tissue which proliferate extensively during development. So daughter cells must maintain the gene expression profile inherited from their mother cells.

In Eukaryotes, the DNA is wrapped around histones forming a nucleosome. In a broader scale, nucleosomes build the chromatin. For transcriptional active gene loci, the chromatin structure is looser and more accessible to (specific) transcription factors and the general transcription machinery. On the contrary, compaction of the chromatin structure is generally associated

with repression/silencing of gene transcription. Thus, the regulation of chromatin structures at specific loci plays a big role in the concept of cellular memory.

Proteins of the Polycomb group (PcG) are conserved from flies to mammals and have been shown to silence gene expression via the modification of the local chromatin structure (SCHUETTENGRUBER *et al.* 2007). Broadly, the PcG proteins are associated into higher order protein complexes. The two most studied PcG protein complexes are the Polycomb repressive complexes 1 and 2 (PRC1 and PRC2; BEISEL and PARO 2011; KASSIS and BROWN 2013; GROSSNIKLAUS and PARO 2014). In *Drosophila*, PRC2 includes the proteins Enhancer of zeste (E(z)), Extra sex combs (Esc), Suppressor of zeste 12 (Su(z)12) (CZERMIN *et al.* 2002; MÜLLER *et al.* 2002). E(z) displays the catalytic component of PRC2 and trimethylates histone H3 lysine 27 creating the characteristic H3K27me3 mark of Polycomb-repressed genes (CAO *et al.* 2002). This chromatin mark is specifically recognized by the protein Polycomb (Pc) (CAO and ZHANG 2004). Pc together with other proteins of the Polycomb family, such as Sexcombs on midlegs (Scm), Polyhomeotic (Ph), Posteroir sex combs (Psc) and dRing, forms the PRC1 (PETERSON *et al.* 1997; SAURIN *et al.* 2001). Subsequently, activity of PRC1 leads to a further compaction of the chromatin and results in silencing of gene expression.

On the other hand, activity of proteins from the trithorax group (trxG) have been shown to maintain the active state of gene expression (KINGSTON and TAMKUN 2014). Prominent members of this group are encoded by the genes *trithorax* (*trx*), *absent, small or homeotic 1* (*ash1*), *trithorax-like* (*trl, GAF*), and *brahma* (*brm*). In *Drosophila,* these genes were initially identified as mutations that resemble loss-of-function of homeotic genes or can act as suppressors of Pc mutants (KENNISON and TAMKUN 1988; KENNISON 1995). Brm has been shown to be part of a conserved multiprotein complex that actively remodels the chromatin structure (TAMKUN *et al.* 1992; KRUGER *et al.* 1995). This ATP-dependent chromatin remodeling is very complex and not yet completely understood. However, it has been proposed that it can lead to disposal, sliding or exchange of histones, which would give transcription factor more space to interact with *cis*-regulatory DNA elements (KINGSTON and TAMKUN 2014). Trx and Ash1 were shown to specifically trimethylate H3K4 (BYRD and SHEARN 2003; DOU *et al.* 2005; WYSOCKA *et al.* 2005). This methylation mark has been mostly associated with highly active promoters (KIM *et al.* 2005).

Thus, PcG and TrxG proteins have opposing roles in either repressing or maintaining active transcription via alteration of the chromatin structure after having integrated signals from *cis*-regulatory elements (Figure 10).
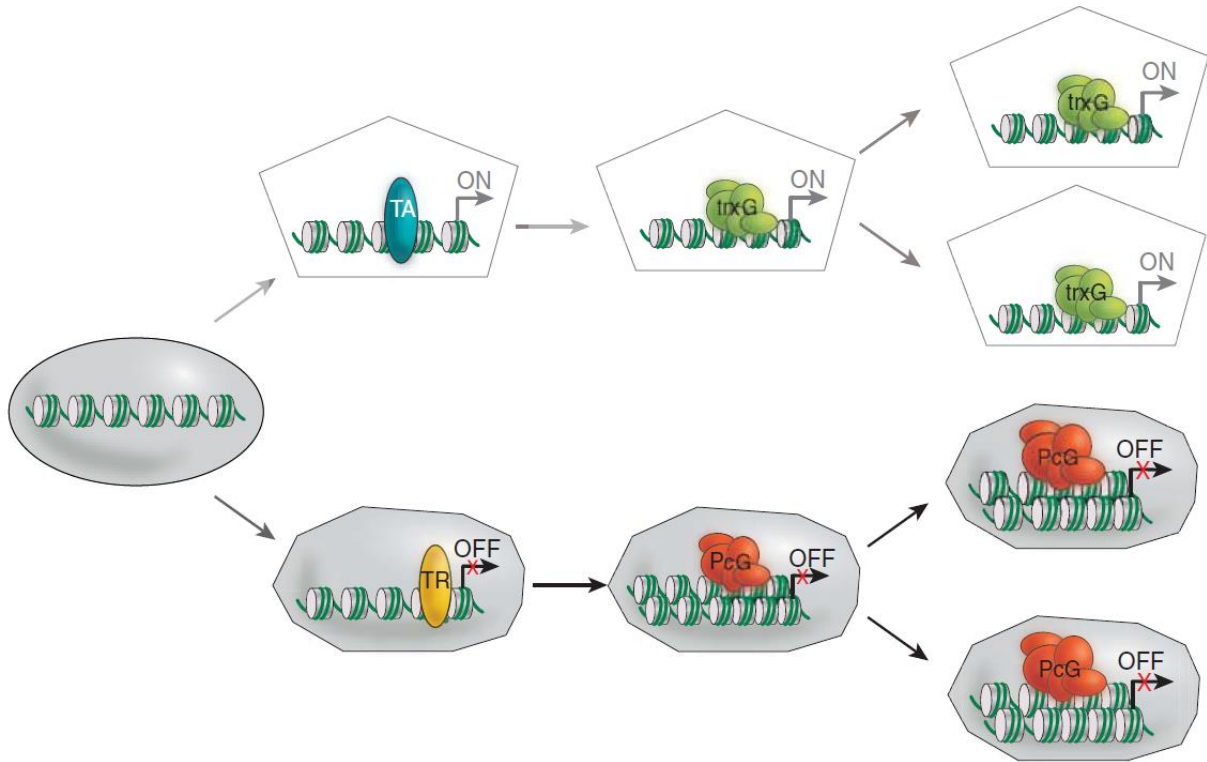
Figure 10 The role of trxG and PcG proteins in the establishment of the cellular memory. After transient activation, the activities of trxG proteins maintain gene expression in daughter cells. PcG proteins silence gene expression by compaction of chromatin structure. From GROSSNIKLAUS and PARO 2014.

But how are the PcG and trxG proteins brought to the right location in the genome? In *Drosophila*, researchers have shown that these complexes interact with specific factors that are bound to a special class of *cis*-regulatory elements, the so-called Polycomb/Trithorax response elements (PRE/TREs; BUSTURIA *et al.* 1989; MAEDA and KARCH 2006; RINGROSE and PARO 2007; KASSIS and BROWN 2013). The presence of PRE/TREs has been proposed by the cytological analysis of PcG/trxG proteins binding to *Drosophila* polytene chromosomes (ZINK and PARO 1989; DECAMILLIS *et al.* 1992; CHINWALLA *et al.* 1995). Detailed analyses of several different PREs have identified a number of different factors and respective DNA-binding sites important for proper PRE/TRE function (STRUTT *et al.* 1997; HAGSTROM *et al.* 1997; MAURANGE and PARO 2002; DEVIDO *et al.* 2008; CUNNINGHAM *et al.* 2009; reviewed in RINGROSE and PARO 2007; BROWN and KASSIS 2013). These factors include Pleiohomeotic and Pleiohomeotic-like (Pho/Phol), which bind to a GCCAT DNA motif and have been shown to interact with PcG proteins to induce silencing at the PRE (BROWN *et al.* 1998). Another protein implied in PRE-mediated silencing is

28

Dsp1, which binds to the GAAAA consensus site (DÉJARDIN *et al.* 2005). Two other DNA-binding proteins that have been shown to interact with PcG complexes are Sp1 and Grainy head (Grh), which specifically bind RRGGYGY and TGTTTTT consensus sites, respectively (BROWN *et al.* 2005; BLASTYÁK *et al.* 2006). The proteins GAF (aka Trl) and Pipsqueak (Psq) bind to the same DNA sequence (GAGAG; STRUTT *et al.* 1997; HODGSON *et al.* 2001). The Zeste protein binds to the YGAGYG DNA consensus site (BERGMAN *et al.* 2004). Zeste, GAF and Psq have all been demonstrated to be involved in gene activation as well as silencing (HAGSTROM *et al.* 1997; DÉJARDIN and CAVALLI 2004). Altogether, it seems that a functional PRE/TRE consists of a number of specific DNA motifs, which are cooperatively bound by proteins that recruit other PcG or trxG effectors (RINGROSE and PARO 2007; BROWN and KASSIS 2013).

The functionality of putative PREs can be tested by various means including the pairing-sensitive mini-*white* silencing assay (KASSIS *et al.* 1991; BROWN and KASSIS 2013). mini-*white* is a modified, small version of the *white* gene which lacks all of its tissue-specific enhancers. It is frequently used as a marker for transgenesis, because it results in an easily scoreable eye color phenotype. Transformed flies containing one copy of a mini-*white* harboring transgene can, depending on their genomic location, show a variety of eye colors ranging from weak yellow (low mini-*white* activity), to orange, to brown, to red (high mini-*white* activity). mini-*white* activity is dosage-dependent, meaning that the eye pigmentation is increased in homozygous flies as compared to their heterozygous siblings. It is this dosage dependence which is affected by the presence of a PRE cloned right next to the mini-*white* reporter gene: homozygous flies containing a {PRE, mini-*white*}-transgene can have less mini-*white* activity compared to their heterozygous siblings. In the most extreme causes, mini-*white* activity is completely suppressed. Furthermore, this PRE-mediated silencing effect is position-dependent. Kassis (1994) has also shown that {PRE, mini-*white*}-transgene insertions in the vicinity of an unrelated, endogenous PRE are likely to show the pairing-sensitive silencing effect (KASSIS 1994). So far, all the PREs tested with this assay displayed this pairing-sensitive mini-white silencing (KASSIS and BROWN 2013).

In this regard, it is important to mention that various ChIP studies have proposed the presence of a PRE directly upstream of the *ap* transcript isoforms *ap-RA* and *ap-RC* (SCHWARTZ *et al.* 2006; TOLHUIS *et al.* 2006; OKTABA *et al.* 2008). In silico analysis of the corresponding DNA interval revealed that it harbors clustered consensus sites for Dsp1, Zeste, GAF, Pho and Sp1 (Dimi Bieli and Daryl Gohl, unpublished, not shown). Additionally, a 400 bp fragment of this region was

shown to mediate pairing-sensitive mini-*white* silencing (Daryl Gohl and Martin Müller, unpublished). Further genetic studies at the endogenous *ap* locus demonstrated that the *ap* PRE dominantly silences the mini-*white* activity of nearby {mini-*white, yellow*}-transgenes (Daryl Gohl and Martin Müller, unpublished). Thus, there is molecular and genetic evidence that that *ap* possesses a PRE/TRE. However, its role during wing disc development remains elusive.

Aim of the project

Due to its key role in wing disc development and compartment formation, the role and function of Apterous has been studied extensively. However, comparatively little is known about the factors and mechanisms that regulate the expression of *ap* during the development of the wing.

Critical unanswered questions in wing development are: How is the expression of *ap* restricted to and maintained in the dorsal compartment? How is a sharp, stable compartment boundary of *ap*-positive and *ap*-negative cells generated *de novo* during the growth phase of the imaginal disc?

Therefore, investigations of the *cis*-regulatory elements of *ap* are crucial to provide more insights into the initiation and maintenance of a compartment boundary.

So far, *cis*-regulatory elements were mainly investigated using reporter-based assays. For this purpose, putative regulatory DNA fragments were tested for their ability to drive reporter gene expression when present on a transgene inserted randomly in the genome (SIMON *et al.* 1985; HIROMI and GEHRING 1987). However, with this method, enhancer fragments are tested in a genomic environment that may differ considerably from their endogenous locus. Moreover, this approach allows no prediction whether the investigated elements are required, sufficient, permissive or even dispensable for the regulation of gene expression at their original location.

We aimed to define and characterize the *cis*-regulatory elements of *ap* at their endogenous location in the genome and compare the results to the classical reporter-based assay.

To do so, we used several complementary genetic approaches. First, a classical enhancer *LacZ* reporter study was performed, using the fragment apC as a starting point. Second, deletions with defined breakpoints in the *ap* genomic locus were generated and their effects on wing development were analyzed. Third, we characterized several classical and newly generated *ap* alleles at the molecular level. Finally, a ΦC31-integrase-dependent *in situ* rescue system was engineered, which can be used to investigate the activity of *cis*-regulatory elements at the endogenous *ap* locus.

PUBLICATION I

Up to now, very little published information was available about how the selector gene *apterous* is regulated during wing disc development. We decided to improve our knowledge by dissecting the regulatory landscape of *ap*. In this first publication, we started with enhancer bashing of the previously published 8 kb apC (wing) enhancer fragment from LUNDGREN *et al.* 1995 and shortened it to 874 bp. Using large deletions at the endogenous *ap* locus uncovering the minimal enhancer region, we demonstrated that this enhancer is essential for *ap* expression and wing development. Moreover, molecular analyses of the classical mutants *ap^blot* and *ap^Xasta* showed that in both, the activity of this minimal enhancer is affected. Furthermore, we reported the generation and validation of an *in situ* rescue system and showed that the defined minimal wing enhancer was not sufficient to restore proper *ap* expression and rescue wing development. This first paper basically set the stage for PUBLICATION II, in which a more complete view of *ap* regulation in the wing disc is presented.

The *Drosophila melanogaster* Mutants *ap^blot* and *ap^Xasta* Affect an Essential *apterous* Wing Enhancer

Dimitri Bieli,* Oguz Kanca,* Daryl Gohl,[†,1] Alexandru Denes,* Paul Schedl,[†] Markus Affolter,*
and Martin Müller*[,2]
*Biozentrum, University of Basel, 4056 Basel, Switzerland, and [†]Department of Molecular Biology, Princeton University, New Jersey 08540

ABSTRACT The selector gene *apterous* (*ap*) plays a key role during the development of the *Drosophila melanogaster* wing because it governs the establishment of the dorsal-ventral (D-V) compartment boundary. The D-V compartment boundary is known to serve as an important signaling center that is essential for the growth of the wing. The role of Ap and its downstream effectors have been studied extensively. However, very little is known about the transcriptional regulation of *ap* during wing disc development. In this study, we present a first characterization of an essential wing-specific *ap* enhancer. First, we defined an 874-bp fragment about 10 kb upstream of the *ap* transcription start that faithfully recapitulates the expression pattern of *ap* in the wing imaginal disc. Analysis of deletions in the *ap* locus covering this element demonstrated that it is essential for proper regulation of *ap* and formation of the wing. Moreover, we showed that the mutations *ap^blot* and *ap^Xasta* directly affect the integrity of this enhancer, leading to characteristic wing phenotypes. Furthermore, we engineered an *in situ* rescue system at the endogenous *ap* gene locus, allowing us to investigate the role of enhancer fragments in their native environment. Using this system, we were able to demonstrate that the essential wing enhancer alone is not sufficient for normal wing development. The *in situ* rescue system will allow us to characterize the *ap* regulatory sequences in great detail at the endogenous locus.

The body wall and appendages of the adult fly are generated by specialized clusters of primordial cells in *Drosophila* larvae called imaginal discs. The patterning of cells in imaginal discs is initiated by establishing cell lineage boundaries, called compartments (Garcia-Bellido *et al.* 1973; Dahmann and Basler 1999). In the case of the wing imaginal disc, the tissue is subdivided into four different compartments, anterior (A) and posterior (P) as well as dorsal (D) and ventral (V). The A−P compartment is established during the process of segmentation in the embryo. The subdivision into dorsal and ventral compartments takes place later in development during the larval

stages when the wing tissue is growing extensively (Wieschaus and Gehring 1976; Lawrence and Morata 1977; Cohen *et al.* 1992; Williams *et al.* 1993; Diaz-Benjumea and Cohen 1993). Short-range signaling events between the A−P or D−V compartments specify cells close to the compartment boundaries. These cells, also called organizer, play an important role in patterning the surrounding tissue by secreting long-range signaling molecules, also referred to as morphogens (Struhl and Basler 1993; Diaz-Benjumea and Cohen 1995; Neumann and Cohen 1997; Affolter and Basler 2007).

Compartment specificity is conferred by the cell-autonomous activity of a special class of transcription factors, called selector genes. Selector genes regulate genes important for proper differentiation and genes that control cell−cell interactions at the compartment boundary. *apterous* (*ap*), which is expressed in the dorsal compartment of the wing disc, has been shown to act as a selector gene subdividing the wing disc into a D and a V portion (Cohen *et al.* 1992; Diaz-Benjumea and Cohen 1993; Williams *et al.* 1994; Blair *et al.* 1994). Different *ap* alleles can lead to a wide range of wing phenotypes (Stevens and Bryant 1985). The most striking morphological defect in strong *ap* alleles is the complete lack of wing and haltere structures (Butterworth and King 1965). Because *ap* is not essential for the progression through larval and pupal stages, the investigation of adult *ap* mutant wing phenotypes is possible.

The target genes of Ap and their downstream functions in the patterning of the wing disc are relatively well understood. The activity of Ap initiates a bidirectional Notch signaling cascade at the D−V compartment boundary, which subsequently induces the expression of *wingless* (*wg*) in a stripe along the compartment boundary (Diaz-Benjumea and Cohen 1993; Williams *et al.* 1994; Irvine and Wieschaus 1994; Rulifson and Blair 1995; Kim *et al.* 1995; Couso *et al.* 1995). Wg, a ligand of the Wnt family, is responsible for the growth of the wing pouch and patterning along the D−V-axis, although its mode of action as a classical morphogen currently is questioned (Neumann and Cohen 1997; Alexandre *et al.* 2014).

Despite the rather detailed knowledge about the functions of Ap in wing disc development, our knowledge of the mechanisms regulating *ap* expression is still limited. It has been shown that activation of the epidermal growth factor receptor by its ligand Vein is necessary and sufficient to activate the expression of *ap* in the dorsal compartment of the wing disc (Zecca and Struhl 2002a,b). Moreover, early ventral *wg* expression has been shown to restrict the expression of *ap* to the dorsal portion of the developing wing disc (Williams *et al.* 1994).

To identify the wing disc-specific *cis*-regulatory elements of *ap*, we used several genetic approaches. First, a classical *LacZ* enhancer reporter study was performed. Second, deletions with defined breakpoints in the *ap* genomic locus were generated. Third, we have characterized two classical *ap* alleles, *ap^blot^* and *ap^Xasta^* (*ap^Xa^*), at the molecular level and have associated their respective molecular alterations to the minimal wing enhancer. Finally, we engineered a *ΦC31*-integrase-dependent *in situ* rescue system, which enabled us to dissect the role of these *cis*-regulatory elements in their native environment.

Using these assays, we have defined an essential, but not sufficient, minimal 874-bp *ap* wing enhancer fragment that drives reporter gene expression in the dorsal compartment of the wing imaginal disc.

MATERIAL AND METHODS

Fly stocks and methods

Flies were grown on standard cornmeal agar at 25°, unless otherwise stated. *ap^e01573^* (*PBac{RB}e01573*), *ap^f08090^* (*PBac{WH}f08090*), *ap^f00451^* (*PBac{WH}f00451*), and *ap^f00878^* (*PBac{WH}f00878*) were purchased from the Exelixis stock collection at Harvard Medical School. *Df(2R) nap1* (BL#1006), *ap^blot^* (BL#4190), *w**; *T(2;3)ap^Xa^*, *ap^Xa^/CyO*; *TM3, Sb^1^* (BL#2475), *P{hsFLP}12, y^1 w** (BL#1929), *Df(2R)BSC696* (BL# 26548), *w**; *P{10XUAS-IVS-mCD8::GFP}attP40* (BL#32186), *Df(3R)Exel6176* (BL#7655), *TM3, ry^RK^ Sb^1^ Ser^1^ P{Δ2-3}99B* (BL#1808), *Bx-Gal4* (*w^1118^ P{GawB}Bx^MS1096^*, BL#8860), *y^1 w**; *Mi{y[+mDint2]=MIC}MI00964* (BL#34133), *y^1 w**; *Mi{y[+mDint2]=MIC}MI02330/SM6a* (BL#33205), *y^1 w^67c23^*, *P{EPgy2}EY03046* (BL#15619), *ptc-Gal4* (*P{GawB}ptc^559.1^*; BL#2017) were all obtained from the Bloomington Stock Center. *fng-Gal4* (*y* w**; *P{w^+mW.hs^= GawB}NP5399 / TM6, P{w=UAS-lacZ.UW23-1}UW23-1*, DGCR#104990) was obtained from Kyoto Drosophila Research Center. *Dad4-Gal4* was established in our laboratory as described in the sections to follow. *actin-Gal4* (*y w^1118^*; *P{actin5c::Gal4, w }/CyO*) and *GMR-Gal4* (*w^1118^*; *P{GMR::Gal4, w }/CyO*) were obtained from Steven Henikoff (Ahmad and Henikoff 2001). *salE-Gal4* was obtained from the Basler lab via Fisun Hamarotoglu (Mosimann *et al.* 2006). *dpp-Gal4* is described in Staehling-Hampton *et al.* (1994). *UAS-ap* was obtained from Marco Milán (Milán and Cohen 1999). *y w M{vas-int.Dm}zh-2A*, a stock producing *ΦC31*-integrase under the control of the *vasa* promoter, and insertion platform *M{3xP3-RFP.attP}zh-86Fb* were obtained from Johannes Bischof (Bischof *et al.* 2007). *ap^41F^/T(2;3)ap^Xa^* was obtained from John B. Thomas.

According to our genetic and molecular analysis, *ap^41F^* should not be listed as an allele of *ap*. First, and contrary to a previous report (Bourgouin *et al.* 1992), hemizygous *ap^41F^* flies have normal wings and halters. Second, although molecular analysis confirmed the presence of a *P*-element insertion just proximal to *vulcan* on the *ap^41F^* chromosome, polymerase chain reaction (PCR), and sequencing failed to provide evidence for a ~200-bp deletion within 1.5 kb of the longest *ap* cDNA (D. Bieli and M. Müller, unpublished data). The *GFP* knock-in allele *ap::GFP* is described in Caussinus *et al.* (2011) (BL#38423). *ap^MM^* has been described in Gohl *et al.* (2008). It contains an insertion ~400 bp upstream of the longest *ap* cDNA. *Dad4-GFP* (*P{Dad4::EGFPnuc, w^+^}*) was obtained from Jorgos Pyrowolakis (Vuilleumier *et al.* 2010). Nuclear enhanced green fluorescence protein (GFP) is expressed under the control of the Dad4 enhancer (Weiss *et al.* 2010). *dad^P1883Δ32^ / TM3, Sb* was obtained from Tetsuya Tabata. This deletion covers at least 24 kb downstream of the *Dad^P1883^* insertion, including the complete *Dad* open-reading frame (ORF) and three neighboring genes (*CG3983, CG5184,* and *CG3962*; T . Tabata, personal communication; Tsuneizumi *et al.* 1997; Henderson *et al.* 1999). A recombinant between *ap^Xa^* and *P{y[+t7.7] w[+mC]=10XUAS-IVS-mCD8::GFP}attP40*, inserted on 2R at 25C6 (Pfeiffer *et al.* 2010), was obtained by meiotic recombination and selection for the dominant *Xasta* and *mini-white* markers. The generation of deficiencies *ap^DG1^*, *ap^DG3^*, *ap^DG8^*, and *ap^DG11^* is described below in section *ΦC31-integrase–mediated transgenesis and generation of deletions*.

Adult wings were dissected and mounted in Hoyer's. Then, wing preparations were baked at 58° for a few hours. Preparations were allowed to harden at room temperature and flattened by applying a 40-g metal cylinder on the cover slip. Pictures were taken with a Nikon Microphot-FXA microscope with a Sony NEX-5RK digital camera. The notums of adult flies were photographed with a Leica M125 binocular equipped with a Leica DFC420C camera.

Introduction of *ΦC31*-integrase targets into the *ap* locus by gene conversion at the site of *ap^MM^*

A method known as direct gene conversion has previously been developed to engineer a desired DNA fragment into the genomic site of a *P*-element insertion (Gloor *et al.* 1991; Sipos *et al.* 2007). Upon exposure of a given *P*-element insertion to *P*-element transposase, the transposon is excised and a double strand break is created. It is normally repaired by the cellular machinery using the homologous chromosome as a template. However, the repair process may also use an exogenous plasmid containing the desired DNA fragment flanked by homology arms derived from either side of the *P*-element insertion site. Such a gene conversion template plasmid containing homology arms flanking the site of *ap^MM^* insertion, along with *hsp70-GFP* bracketed by a pair of inverted attB sites was constructed and named pLAPGPRA (see Figure 5A). The construction of this plasmid was a multi-step procedure. Details can be obtained upon request. In brief, left (899 bp long) and right (1981 bp long) *ap*-homology arms were amplified by PCR. To minimize sequence polymorphism which could decrease the efficiency of gene conversion, *ap^MM^* genomic DNA (gDNA; isolated as described in Ashburner 1989) was used as the template for PCR. As primers we used apLA-R, apLA-FNotI, apRA-F, and apRA-R (for primer sequences, see Supporting Information, Table S1). Inserts in the proper orientation for subsequent cloning were identified using diagnostic digests and sequencing.

pLAPGPRA was injected (650 ng/μL) along with pTurbo (250 ng/μlL) into embryos derived from a cross of *y w; ap^DG3^{w^+^}/+; TM3, Sb Δ2-3/+* males with *y w; ap^MM^{y^+^}* ; + virgins. Surviving injectees were transferred to fresh vials and carefully tended at 18°. Among the hatching

adults, males and virgins representing the two desired genotypes ($ap^{DG3}\{w^+\}/ap^{MM}\{y^+\}$ and $ap^{DG3}\{w^+\}/ap^{MM}\{y^+\}$; TM3, Sb Δ2-3/+) were selected for further work. ap^{DG3}/ap^{MM} flies have normal wings and halteres. The ap^{DG3} chromosome was included because it lacks the DNA corresponding to the homology arms of pLAPGPRA and hence cannot serve as a template for double strand gap repair. A total of 72 fertile crosses involving virgins (in pairs) or single males mated with y w; al b c sp/SM6a flies could be set up. Originally, it was intended to screen the larval progeny of these crosses for GFP expression. Unfortunately, this elegant approach failed in practice. Therefore, the progeny was screened for y^- w^- males. This phenotype indicates loss of the y marker and therefore most likely also of ap^{MM} and was, in the absence of the positive GFP selection, the only selectable marker to identify putative conversion candidates. A total of 105 y^- w^- males were selected from 32 (out of 72) crosses yielding such males. Balanced lines of potential gene conversion events were established and screened for GFP fluorescence in larval wing discs. Five candidate gene conversion lines with weak GFP expression in wing imaginal discs in an ap-like pattern were obtained from two independent dysgenic crosses (isolation numbers: c1.4a, c1.4b, c1.4d, and c1.4e; c1.13a).

To confirm that the five GFP-positive candidate gene conversion lines had the attP-flanked GFP construct integrated in the ap locus, gDNA was isolated and analyzed by PCR. PCR products were obtained for all five candidates between a primer (SV40out55) within the SV40 trailer sequence (just downstream of GFP) and a primer (apLOF2) in the ap gene outside of the left homology arm. Sequencing of all five lines confirmed the integrity of the ap promoter region and the presence of the ap proximal attP site (data not shown).

On the distal side, PCRs using primers in the hsp70 promoter (just upstream GFP) and several primers outside of the right homology arm initially failed to produce products (data not shown). Later, by use of one of the lines obtained by RMCE (see below), the integrity of the junction between the template plasmid and the right homology arm could be verified by PCR and sequencing using the Mcp-dir-y and ap-dir-3 as primers. Mcp-dir-y primes toward the end of the mini-yellow gene present on our Recombination-Mediated Cassette Exchange (RMCE) insertion cassette. We also tested whether the junction between the right homology arm and the flanking ap sequence is intact by PCR using a primer near the end of the right homology arm (apRAendF), and a primer in the flanking ap sequence (apROR2). A product of the expected size was observed, indicating that the junction is intact.

$ap^{c1.4a}$, $ap^{c1.4b}$, $ap^{c1.4d}$, $ap^{c1.4e}$, and $ap^{c1.13a}$ homozygotes all have wild-type wings, indicating that the function of the ap wing enhancer and promoter were not disrupted by the gene conversion event. Gene conversion events $ap^{c1.4a}$, $ap^{c1.4b}$, $ap^{c1.4d}$, $ap^{c1.4e}$ also have a rough eye phenotype when homozygous, but not over ap^{DG3}. The rough eye phenotype can be separated from the ap locus by meiotic recombination. Finally, only $ap^{c1.4b}$ was chosen for further work. One of its applications is the targeted insertion of exogenous DNA into the ap locus by RMCE (Bateman et al. 2006).

ΦC31-integrase−mediated transgenesis and generation of deletions

Constructs for ΦC31-integrase−mediated transgenesis were generated based on plasmid piB-LLFY(BI) [details about the construction of piB-LLFY(BI) can be acquired upon request]. As required for RMCE, it contains two inverted attB sites. Separating them are the following three genetic components: (1) two LoxP sites in direct orientation with a multiple cloning site in between them; (2) the LoxP cassette is followed by a single FRT site; (3) the mini-yellow transformation marker completes piB-LLFY(BI). mini-yellow refers to a yellow reporter gene

lacking all of its characterized tissue specific enhancers. It consists of the yellow cDNA fused to ~330 bp of 5′ genomic DNA, including the yellow promoter and extending up to a KpnI restriction site. The mini-yellow fragment was isolated from plasmid C4yellow (referred to as Dint in Geyer and Corces 1987; Gohl et al. 2008). In the context of the ap gene and in a y background, mini-yellow activity always manifests itself in phenotypically yellow+ wings. Depending on the transgene and orientation of insert, thoracic bristles may also acquire yellow+ pigmentation (D. Gohl and M. Müller, unpublished data).

Constructs were introduced into the ap locus by RMCE into two docking sites, Mi{y[+mDint2]=MIC}MI02330 (Venken et al. 2011) and $ap^{c1.4b}$. DNA was injected at a concentration of 300 ng/µL in 1× phosphate-buffered saline (PBS) into early embryos of the genotype y w M{vas-int.Dm}zh-2A; MI02330/CyO or y w M{vas-int.Dm}zh-2A; $ap^{c1.4b}$/CyO. The relevant transgenic lines obtained in this way are $ap^{DD35.34}$ and $ap^{D5f.1}$, respectively. Their position and the orientation of the FRT are depicted in Figure 1C together with four other FRT containing transposon insertions. Five of the six stocks are homozygous and hemizygous viable. Their wings and halteres are of wild-type appearance. This is not the case for ap^{f08090}. The lethality of this chromosome cannot be reverted by excision of the PBac{WH}, indicating that it is associated with a second site lethal. Rare homozygous revertant escapers as well as frequent hemizygous revertants have normal wings. Therefore, the PBac{WH} insert is responsible for the strong phenotype in hemizygous ap^{f08090} flies. However, this phenotype is not dependent on the gypsy insulator present in ap^{f08090} because the wing phenotype is not suppressed in a su(Hw)− background (M. Müller, unpublished data).

We have noted that in the Drosophila literature, two divergent definitions for FRT orientation are in use! In this study, FRT orientation is indicated according to Thibault et al. (2004).

In Drosophila, the production of deletions by Flipase-catalyzed recombination between two FRT sites either in cis or in trans has enabled the community to obtain a huge collection of tailor-made deficiencies (Golic and Golic 1996; Ryder et al. 2007). We have previously applied this technology to generate a ~27-kb deletion named $Df(2R)ap^{DG}$ between two FRT sites in ap^{MM} and ap^{e01573} (Gohl et al. 2008). Note that in this study, $Df(2R)ap^{DG}$ is referred to as ap^{DG1}. Applying analogous genetic crossing schemes, we have generated three further deletions:

$Df(2R)ap^{DG3}$: An ~44-kb deletion between two FRT sites located in ap^{f08090} and ap^{e01573}. It is referred to as ap^{DG3}. In this deficiency, a large part of the ap transcription unit is lost together with ~27 kb of intergenic DNA separating ap from l(2)09851. Although a considerable part of the ap ORF located proximal to the break in ap^{DG3} remains in place, genetic observations are consistent with it being a true null allele with respect to ap function in wing and haltere tissue. Flanking the new FRT junction are three genetic elements: gypsy insulator and mini-white (of ap^{f08090}) and a splice acceptor (of ap^{e01573}). Over several kilobases, the region of the new fusion is identical to PBac{RB}e01573 and, hence, no adequate ap^{DG3}-specific PCR primers could be designed. Thus, four PCR primer pairs distributed evenly over the ~44-kb interval missing on the ap^{DG3} chromosome were tested on w^1 and on $ap^{DG3}/Df(2R)$ nap1 flies (Df(2R)nap1 being a cytologically visible deletion also uncovering ap). The absence of the corresponding DNA in the latter could unambiguously be demonstrated (data not shown).

$Df(2R)ap^{DG8}\{w^+\}$: An ~20-kb deletion between two FRT sites located in ap^{f00878} and $ap^{D5f.1}$. It is referred to as ap^{DG8}. It corresponds to a rather clean deletion of the complete ap transcription unit. Its phenotypes are indistinguishable from those observed for ap^{DG3}. Flanking the new FRT junction are two genetic elements: a UAS-inducible promoter (of ap^{f00878}) and mini-white (of $ap^{D5f.1}$). It was verified by
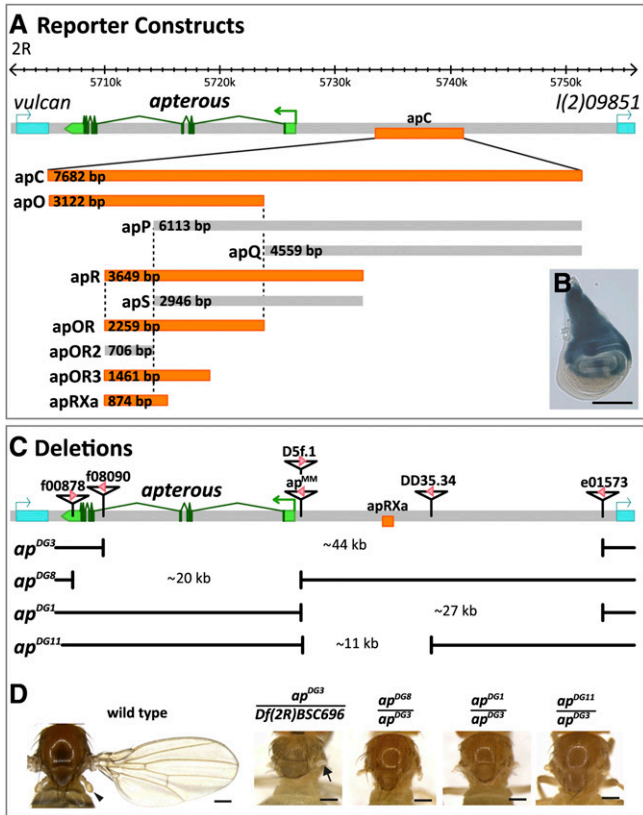
Figure 1 LacZ reporter assay and deletion analysis at the *apterous* locus. (A) Diagrammatic representation of the *ap* locus. As drawn at the top of the panel, it extends over roughly 50 kb. Its transcribed part is shown in green. *ap* is flanked by two genes indicated in blue: *vulcan* on the proximal and *l(2)09851* on its distal side. Arrows above the genomic interval specify the direction of transcription of the three genes. Fragment apC, indicated in orange, has been reported to drive reporter expression in the dorsal compartment of the pouch, the hinge and the notum of the wing imaginal disc, where *ap* is normally expressed. Below, the relative positions and dimensions of nine fragments tested with our LacZ reporter assay are depicted. Fragments colored in orange (apO, apR, apOR, apOR3, and apRXa) elicit the same expression pattern as apC. Fragments depicted in gray (apP, apQ, apS, apOR2) do not drive reporter gene expression in the wing disc. (B) X-Gal staining in the wing disc of an *apC-LacZ* transgenic fly. Scale bar: 100 μm. (C) Deletions generated at the endogenous *ap* locus with FRT-containing inserts. At the top of the panel, triangles along the *ap* locus indicate the position of six different inserts. Pink arrowheads within them mark the orientation of the FRT sites according to the definition of Thibault *et al.* (2004). The location of the apRXa fragment is shown in orange. *ap*^DG3^ deletes approximately 44 kb between inserts *ap*^f08090^ to *ap*^e01573^, thereby removing most of *ap* ORF and upstream sequences. *ap*^DG8^ is a 20-kb deficiency that deletes the complete *ap* ORF from *ap*^f00878^ to *ap*^D5f.1^. *ap*^DG1^ removes the complete intergenic spacer between *ap*^MM^ to *ap*^e01573^. *ap*^DG11^ deletes an 11-kb fragment from *ap*^MM^ to *ap*^DD35.34^. Note that *ap*^D5f.1^ and *ap*^MM^ have exactly the same insertion site. (D) Notum pictures of a wild-type fly and *trans*-heterozygous *ap* mutants. In the wild type, the wing and the haltere (arrowhead) are well formed and clearly visible. *Df(2R) BSC696* is a large deletion at the base of 2R, deleting approximately 360 kb, including the whole *ap* locus. When *Df(2R)BSC696* is crossed to *ap*^DG3^ all wing and haltere structures are lost. Only small stumps of amorphic tissue remain at the actual attachment site of the wing (see arrow).Very similar phenotypes are observed in *ap*^DG8^/*ap*^DG3^, *ap*^DG1^/*ap*^DG3^ and *ap*^DG11^/*ap*^DG3^ flies. Scale bar: 25 μm.

PCR and sequencing. Aprec-LA-AscI-F and WARIout#1 primers were used for PCR. For sequencing, we used primers apEnhDel-Seq-PBrev and WARIout#2.

Df(2R)ap^DG11^, al: An ~11-kb deletion between 2 FRT sites in *al ap^MM^ sp* and *ap^DD35.34^*. It is referred to as *ap^DG11^*. Because *ap^DG1^*, homozygous *ap^DG11^* flies have no wings. Both deficiencies share the same proximal break point. Previous transvection studies have suggested that the *ap* promoter immediately proximal to *ap^DG1^* (and hence also of *ap^DG11^*) remains intact (Gohl *et al.* 2008). Flanking the new FRT junction are two genetic elements: a LoxP site (of *ap^MM^*) and *mini-yellow* (of *ap^DD35.34^*). The new junction was verified by sequencing. For PCR amplification of the region, the apMM-200for and yellow5′out primer pair was used. Part of the fragment was sequenced with yellow5′out and Inverseappromfor.

Generation of a *ΦC31*-integrase based *in situ* rescue system at *ap*

Construction of pBSattBattPLoxFRTy: Two complementary oligos (attBPfor and attBPrev) containing attB and attP sites in tandem were purchased from Sigma-Aldrich. These oligos were annealed and cloned between the *Xho*I and *Kpn*I sites of pBSIIKS. The new plasmid's name is pBSIIKSattBattP. A *Xho*I-*Cla*I fragment containing LoxP, FRT, and mini-*yellow* was isolated from piB-LLFY(BI) and subcloned into pBSIIKSattBattP, thereby generating the pBSattBattPLoxFRTy vector used for *ΦC31* integrase mediated transgenesis (see Figure 5B). The attB and attP sites on this vector are separated by only 6 bp. It was assumed that therefore the two elements are too close for efficient intramolecular recombination. The fact the two desired insertions (one in each attP site present in *ap^c1.4b^*, see Figure 5B) could be isolated seems to support this assumption.

Generation of ap^attPΔEnh^, a platform for insertion of ap enhancer fragments: pBSattBattPLoxFRTy DNA was injected into *y w M{vas-int.Dm}zh-2A; ap^c1.4b^/CyO* embryos. Yellow⁺ marked flies could be isolated and mated. The desired insert orientation could be identified by PCR using the apdown-forN and aptransch_yw_rev primers. A stock with isolation number 6.1 was selected for correct orientation of insert pBSattBattPLoxFRTy. It is referred to as *ap^attBPFRTy2^* (see Figure 5C; the other insert orientation was also isolated and called *ap^attBPFRTy1^*). We wished to further modify this stock by introducing the same ~27-kb deletion as in *ap^DG1^*. Therefore, *y w; ap^attBPFRTy2^* males were mated with *y w hsFlp; PBac{RB}e01573* virgins. Progeny was heat-shocked at three subsequent days during its larval development for 1 hr in a 37° water bath. Hatchlings were individually crossed to *y w M{vas-int.Dm}zh-2A ; Sp Pin/CyO* flies. A total of 7 of 80 single crosses produced phenotypically yellow⁻ and white⁻ flies, indicating the loss of all DNA between the FRT sites of *ap^attBPFRTy2^* and *PBac{RB}e01573*, including *mini-yellow* and *mini-white*. The newly established deletion was named *ap^attPΔEnh^* (see Figure 4D) and kept as a *y w M{vas-int.Dm}zh-2A ; ap^attPΔEnh^/CyO* stock. The deletion was confirmed by PCR and sequencing with the primer pair apEnhDel_seq_dwnst_for and apEnhDel_seq_PB_rev.

The *ap^attPΔEnh^* chromosome contains a single functional attP site ready for *ΦC31*-integrase catalyzed insertion of pEnh-Reentry derived plasmids (see Figure 5D). Insertion events can be further modified by suitably placed LoxP and FRT sites, allowing for the deletion of the *yellow* marker or the enhancer fragment-*yellow*⁺ marker cassette, respectively (see Figure 5E).

Generation of pEnh-Reentry constructs: yellow⁺ coding sequence and body cuticle enhancer were subcloned into pBSIIKS as a BglII fragment from C4yellow, thereby generating plasmid pBSIIKS-yellow. Please note that the *yellow* wing enhancer is not part of the BglII fragment! attB and FRT LoxP fragments were cloned by first annealing and phosphorylating oligos attBtop and attBbottom as well as FRTLoxPtop and FRTLoxPbottom followed by three fragment ligation with pBSIIKS-yellow vector cut with SacI and XbaI. The resulting plasmid was called pEnh-Reentry and served as the backbone for all constructs described below.

The 27-kb full-length enhancer was recombineered in pEnh-Reentry from BACR45O18 (purchased from the Berkeley *Drosophila* Genome Project). The left homology arm was amplified with PCR with primers containing NotI and XhoI sites (primer pair: apenhrecLA_Not_for and apenhrecLA_XhoI_rev). The right homology arm was amplified with primers containing XhoI and BglII sites (primer pair: apenhrecRA_XhoI_for and apenhrecRA_BglII_rev). Homology arms were cloned in pEnh-Reentry cut with NotI/BglII as 3 fragment ligation. Recombineering was performed according to Thomason *et al.* 2007. In brief, the pEnh-Reentry-homologyarms vector was linearized with XhoI and transformed into bacterial strain DY380 (purchased from NCI at Frederick) pretransformed with BAC45O18 (purchased from BDGP), and pre-induced at 42° for 15 min. Recombinants were selected on ampicillin and screened by PCR. The correct recombineering product's name is pEnh-Reentry-Full-length.

Dad enhancer fragments and apRXa were amplified from apˣᵃ gDNA. First, fragments apRXaDadInt2, DadInt52, and Dad4 were cloned into a pBluescript II KS(+) vector, where the XbaI site was mutated previously into a AvrII site. For apRXaDadInt52, primers apR_AvrII_for and dadint52_XmaI_SpeI_rev were used. For DadInt52, primers dadint52_XmaI_SpeI_rev andXa_brkpnt_AvrII_for were used. To clone Dad4, we used the primer pair dad4_AvrII_for and dad4_XmaI_SpeI_rev. These fragments were combined via the respective SpeI or AvrII sites to produce apRXaDadInt52Dad4 and DadInt52Dad4 fusion fragments. These were subcloned from pBluescript II KS(+) via AvrII and XmaI sites into pEnh-Re-entry cut with AvrII and AgeI. apR, apRXa, apP, and apY were amplified from pEnh-Reentry-Full-length plasmid and cloned into pEnh-Re-entry via NotI, AvrII or AgeI sites. To clone apR, primers apR_AvrII_for and apR_XmaI_SpeI_rev were used. For apRXa, primer pair apR_AvrII_for and apRXa_AgeI_rev was used. To amplify apP, primers apP_NotI_for and apP_AvrII_rev were used. apY was amplified using primer pair apY_NotI_for and apY_AgeI_rev.

All pEnh-Reentry derived constructs were brought into the *ap* locus by ΦC31-integrase mediated recombination (see Figure 4, D and E). DNAs were injected at a concentration of 300 ng/μL in 1×PBS into *y w M{vas-int.Dm}zh-2A ; ap^{attPΔEnh}/CyO* embryos. Transgenic flies were selected with the help of the yellow⁺ marker and balanced stocks were generated according to standard genetic procedure.

Generation of LacZ-reporter lines

ap regulatory DNA were amplified via PCR from *y¹ w^{67c23}* gDNA with primers containing restriction enzyme sites as overhangs, and subsequently cloned into plasmid pAttBLaZ (Weiss *et al.* 2010) using the respective enzymes. apC was amplified with the primer pair apC_AscI_for and apC_BglII_rev. The apC fragment was defined by Lundgren *et al.* 1995. The apO fragment was cloned with the primers apC_AscI_for and apO_BglII_rev. For apP, primers apC_BglII_rev and apP_AscI_for were used. To clone apQ, primer pair apC_ BglII_rev and apQ_AscI_for was used. For apR, the primers apR_AscI_for and apR_BglII_rev were used. apS was cloned with the primers apS_AscI_for and apS_BglII_rev. apOR was amplified with apR_AscI_for and

apO_BglII_rev. For apOR2, primers apR_AscI_for and apOR2_XbaI_rev were used. apOR3 was amplified with apR_AscI_for and apOR3_XbaI_rev. apRXa was cloned with apR_AscI_for and apRXa_XbaI_rev.

All the reporter transgenes were generated with the ΦC31-based integration system using the landing platform M{3xP3-RFP.attP}zh-86Fb (Bischof *et al.* 2007).

Molecular characterization of ap^{blot}

Complementation crosses with ap^{blot} over a set of overlapping *ap* deletions mapped the mutation to am ~11-kb interval upstream of ap^{MM}. Therefore, a set of PCR primer pairs was designed to screen for a lesion in that region of ap^{blot} gDNA. *y¹ w^{67c23}* gDNA served as positive control. With one primer pair, a discontinuity could be identified on the ap^{blot} chromosome. It could be best reconciled with the presence of a larger insertion of DNA of unknown origin. Inverse PCR (iPCR) was subsequently used to obtain sequence information about the ends of the putative insertion. Toward that end, ap^{blot} gDNA was digested with BsaWI and ligated with T4 Ligase under conditions as previously described (Ochman *et al.* 1988). Primer pairs used for iPCR on the proximal side of the insertion were iPCR_for and iPCR_rev. Primer pairs used for iPCR on the distal side of the insertion were K_for and L_rev. Following this strategy, sequence information could be obtained for both ends of the inserted DNA. Sequence comparison identified them as LTRs of the *blood* retrotransposon (Bingham and Chapman 1986). To verify the insertion, primers out of *blood* 3′ and 5′ LTR (blood3prime and blood5prime, respectively) were used with primers binding in adjacent *ap* regions (iPCR_for and L_rev, respectively). Sequencing was performed by Microsynth AG, Switzerland.

Molecular characterization of apˣᵃ

The dominant *Xasta* allele was originally induced by X-ray mutagenesis in a stock already containing two large inversions on 2R and 3R (Serebrovsky and Dubinin 1930; Waddington 1940; Lewis 1951; Hetherington *et al.* 1968). The new rearrangement was classified as a reciprocal translocation with breakpoints 41F9-41F11;89E8-89F1. Allelism with *ap* was inferred from noncomplementation with known *ap* alleles (Butterworth and King 1965; Stevens and Bryant 1985). Complementation crosses with a set of small overlapping *ap* deletions failed to narrow down the location of apˣᵃ. Hence, the whole *ap* locus was screened by overlapping primer pairs. PCR products obtained from amplification of apˣᵃ/+ and *y¹ w^{67c23}* gDNA were compared. The analysis of these reactions identified a difference close to the insertion break point found in ap^{blot}. Again, this region was probed by iPCR. apˣᵃ/+ gDNA was cut with NlaIII and religated under diluted conditions. For iPCR, the primer pair iPCR_Xa_rev and 19_for was used. Sequencing of the iPCR product revealed that the reciprocal translocation had fused DNA originating from *dad* locus on 3R to *ap*-specific sequences. The fusion was confirmed by PCR and sequencing with 19_for and a primer in the *dad* region (primer dadint52out). The breakpoint associated with *Xasta* in *ap* was found to be identical in the two stocks ap^{41F}/T(2;3)apˣᵃ and *w**; *T(2;3)apˣᵃ, apˣᵃ/CyO; TM3, Sb¹*.

Generation of Dad4-Gal4 fly line

The minimal *hsp70* promoter was amplified from the pUAST vector with the primer pair hsp70_XbaI_for and hsp70_BamHI_rev, then cloned into pBluescript II KS(+) via the XbaI and BamHI sites. The Dad4 fragment was amplified from gDNA with the primers dad_NotI_for and dad_NheI_rev, followed by the insertion of the fragment into the NotI and XbaI digested pBS-hsp70 plasmid. Gal4 was amplified from a pCaSpeR4-Gal4 plasmid, obtained from the lab of Konrad Basler, with the primer pair Gal4_BglII_for and Gal4_HindIII_rev. The Gal4

fragment was subsequently cloned into the *Bam*HI and *Hin*dIII digested pBS-Dad4-hsp70 plasmid. We amplified the SV40-PA terminator sequence from the pUAST vector, using the primer pair SV40_*Hin*dIII_for and SV40_*Bam*HI_*Apa*I. The SV40_PA was subsequently inserted into the pBS-Dad4-hsp70-Gal4 plasmid using the *Hin*dIII and *Apa*I restriction sites. Finally, the Dad4-hsp70-Gal4-SV40_PA sequence was subcloned into the pCaSpeR4 vector, using the *Not*I and *Bam*HI restriction sites. Transgenic flies were selected in a y^1 w^{67c23} background with the help of the *mini-white* marker. The Dad4-Gal4 insert used in this study is linked to the X chromosome.

X-Gal staining of imaginal discs

Third instar larvae were cut in half, and the anterior part was inverted and subsequently fixed in 1% glutaraldehyde (Fluka) in PBS for 15 min on ice. After fixation, the fixative was removed and the larvae were washed twice with PBST (0.1% Tween 20 in PBS). The tissue was then stained as previously described (Ashburner 1989). Afterward, the imaginal discs were dissected and mounted in 80% glycerol. Discs were analyzed under the Zeiss Axiophot microscope and photographed with a Sony NEX-5RK digital camera.

In situ hybridization

A 1.5-kb fragment from the 3′ end of the *ap* cDNA was amplified from the cDNA clone HL02012 (purchased from DGRC) with primers insitu_*SacI*_for and insitu_*KpnI*_rev. The fragment was cloned between *Sac*I and *Kpn*I sites of pBluescript II KS(+) vector. Then, the resulting plasmid was linearized with *Acc*65I and digoxigenin-(DIG)-labeled RNA was produced from T7 promoter according to the manufacturer's protocol (Roche, Switzerland). *In situ* hybridizations were performed as described in Tautz and Pfeifle (1989). Wing imaginal discs were dissected and mounted in 80% glycerol and photographed under a Nikon Microphot-FXA microscope with a Sony NEX-5RK digital camera.

Immunostaining

The anterior part of third instar larvae was inverted and fixed with 4% paraformaldehyde in PBS for 25 min at room temperature. Standard protocols were used to perform immunostaining. As primary antibodies, rabbit α-GFP (1:1000; Abcam) and mouse α-Wg (1:120, DSHB, University of Iowa) were used. α-rabbit AlexaFlour488 and α-mouse AlexaFlour568 (Molecular Probes) were used at a 1:750 dilution. Samples were mounted in Vectashield (Vector Laboratories, Inc.). Confocal imaging was performed using a Leica SP5 microscope with a vertical step size of 1 μm. Image processing was done with the ImageJ software.

RESULTS

Defining a short wing-specific enhancer element in apC

At *apterous*, four different transcripts starting from three different promoters have been annotated (see www.flybase.org). In this study, the transcription start site for transcripts *ap-RA* and *ap-RC* will be referred to as *ap* TSS.

An ~8-kb DNA fragment named apC located several kilobases upstream of the *ap* TSS had been shown to drive reporter gene expression in an *ap*-specific pattern in the wing disc (Lundgren *et al.* 1995). We used a *LacZ* reporter assay to analyze the *cis*-regulatory elements in apC in more detail. apC was first sub-divided into four overlapping fragments, apO, apR, apP and apQ (Figure 1A). Of these, only the two promoter proximal fragments, apO and apR were found to drive reporter gene expression in the wing disc. To further pinpoint the wing disc enhancers, we generated five subfragments that span the DNA sequences covered by apO and apR. As shown in Figure 1A, this

analysis defines a minimal 874-bp fragment, apRXa. Because apP and apOR2 together cover the minimal apRXa element, but neither showed any expression in the wing disc, key *ap* wing enhancer elements are likely to be on both sides of the breakpoint that divides these two fragments.

To determine whether the wing enhancer element identified in the *LacZ* reporter assay is necessary for the proper regulation of the endogenous *ap* gene, we generated several deletions with defined breakpoints (Figure 1C; for details see the section *Materials and Methods*). The largest of these, ap^{DG3}, removes almost the entire *ap* locus, from the 4th intron to a site located about 500 bp upstream of the flanking distal gene, *l(2)09851*. Previous observations suggested that *l(2)09851* activity is not affected by the proximity of ap^{DG3}'s distal break (Gohl *et al.* 2008). As a homozygote or when *in trans* to a large deficiency, *Df(2R)BSC696*, that includes the entire ap locus, ap^{DG3} flies displayed a complete loss of all wing and haltere structures (Figure 1D). This result suggests that ap^{DG3} represents an amorphic allele of *ap*, at least with respect to Ap function during wing development. Furthermore, we generated a deletion, called ap^{DG8}, which removes the whole ORF from the end of the 3′UTR to ~400 bp upstream of the *ap* TSS. Trans-heterozygous ap^{DG8}/ap^{DG3} flies again showed a typical *ap*-null mutant phenotype. Finally, two deficiencies affecting only the 5′ regulatory region were generated, namely ap^{DG1} and ap^{DG11}. They share the same proximal break located ~400 bp upstream of the *ap* TSS. Previous transvection studies suggested that the activity of the *ap* promoter is not affected by this breakpoint (Gohl *et al.* 2008). ap^{DG1} extends ~27 kb distally to the same position as ap^{DG3}. ap^{DG11} removes only ~11 kb of upstream DNA, including the whole apC fragment. ap^{DG1}/ap^{DG3} as well as ap^{DG11}/ap^{DG3} flies lacked all wing and haltere structures (Figure 1D). In these two deletions, the minimal wing enhancer element defined by apRXa is removed, suggesting that elements within apRXa are indeed necessary for the regulation of *ap* in the endogenous locus (see Figure 1C).

Although *ap* is expressed in the presumptive notum of the developing wing disc, the phenotypic appearance of the adult dorsal thorax is only mildly affected in flies lacking any *ap* activity (*e.g.*, ap^{DG8}/ap^{DG3} or homozygous ap^{DG8}). Apart from a few missing macro- and microchaetae in the vicinity of the wing appendage, it appears largely normal (Figure 1D and data not shown). The reduced size of the dorsal thorax and the aberrant bristle pattern in $ap^{DG3}/Df(2R)BSC696$ flies can probably be attributed to other genetic loci deleted in *Df(2R)BSC696*.

Apart from the dominant ap^{Xa} allele, lesions in the *ap* gene have been reported as recessive in genetic character. Careful inspection of wings obtained from flies heterozygous for any of the 4 deletions presented in Figure 1C corroborated this fact. However, ~2% of them had small margin defects, indicating a mild dominance of strong *ap* loss-of-function alleles (data not shown).

Mutations in the apR region result in wing phenotypes

In the course of investigating the *cis*-regulatory region of *ap*, we identified two classical *ap* alleles, ap^{blot} and ap^{Xa}, that map to the apR region. ap^{blot} was isolated as a spontaneous, hypomorphic mutation that causes notching mostly of the posterior wing margin in homozygous mutant flies, while the anterior wing margin remains largely unaffected (Figure 2A; Butterworth and King 1965; Whittle 1979). To narrow down the genomic site affected by the mutation, intragenic complementation crosses with the aforementioned deletions were analyzed. They showed that ap^{DG11} was the smallest deletion that failed to complement ap^{blot}. This observation suggested that ap^{blot} maps to the ~11-kb interval defined by ap^{DG11}. Consequently, this region was screened with a set of overlapping PCR primer pairs. One primer pair did not yield a PCR product and thus identified the site of the putative

lesion on the *ap^blot* chromosome. Using iPCR, we identified the insertion of a retrotransposable element of the *blood* family in the apRXa sequence (Figure 2B, see the section *Materials and Methods* for details). This event caused the typical 4-bp duplication at the insertion site characteristic for *blood* family transposons (Figure 2C; Bingham and Chapman 1986; Wilanowski *et al.* 1995).

Phenotypes caused by *blood* insertions at other loci are sometimes temperature-sensitive (Bingham and Chapman 1986). To test this possibility, we raised homozygous *ap^blot* flies at different temperatures and scored their wing phenotypes (Table 1). At 18°, only 28% of the wings displayed minor defects. In most of these, the posterior cross vein failed to connect with the 4th wing vein (Figure 2A). At greater temperatures, more severe wing phenotypes were detected with a higher penetrance. At 25° and 29°, 52% and 70%, respectively, of the wings showed extensive notching within the posterior compartment and reduced wing size (Figure 2A).

The dominant *ap^Xa* allele was generated by X-ray mutagenesis and is associated with a reciprocal translocation between chromosome arms 2R and 3R. The breakpoints were mapped to 41F and 89EF, respectively (Serebrovsky and Dubinin 1930; Waddington 1940; Lewis 1951; Hetherington *et al.* 1968). When heterozygous, *ap^Xa* flies show the characteristic dominant mitten-shaped wing phenotype, in which the distal tip of the wing is missing leading to a deep notching of the wing blade. In hemizygous *ap^Xa* flies, only long wing stumps with little or no wing margin and unstructured vein patterns are formed (Figure 2D). The break on 2R has long been known to affect the *ap* locus (Butterworth and King 1965; Stevens and Bryant 1985). However, our attempt to map *ap^Xa* by intragenic complementation was not successful, suggesting that the lesion in *apterous* prevents this type of genetic analysis (see also Figure 3D). Thus, we screened the entire *ap* locus with overlapping PCR primer pairs. We identified a discontinuity in the apR region and determined the molecular nature of the breakpoint (Figures 2, E and F; for details see the section *Materials and Methods*). It localized right at the edge of the apRXa fragment, 142 bp distal to the insertion site of the *blood* transposon in *ap^blot*. Only the proximal 874 bp of apR remain associated with the *ap* transcription unit (see Figure 2E). The DNA on the other side of the breakpoint is from the

daughters against dpp (*dad*) locus located at 89E on 3R. As predicted from the cytological mapping of the rearranged *ap^Xa* chromosomes, the *dad* locus is inverted compared to its wild-type orientation on 3R (for a comprehensive drawing of the *ap^Xa* polytene chromosomes, see Hetherington *et al.* 1968). We were not able to determine the breakpoint at the reciprocal site of the translocation. Nevertheless, based on its reciprocal nature, it is conceivable that the *dad* locus is split within its 4th intron and hence destroyed. Because *dad* is expressed in the imaginal wing disc, it is formally possible that the Xasta phenotype is due to the loss of Dad activity. This possibility was addressed by crossing *ap^Xa* with 2 known *dad* deletions, *Df(3R)Exel6176* and *dad^P1883Δ32*. The wings of *trans*-heterozygous animals displayed the characteristic mitten phenotype seen in *ap^Xa* heterozygous flies, suggesting that an amorphic *dad* background does not modify the Xasta phenotype. Hence Dad function is not relevant for the production of the Xasta phenotype (data not shown). This is not unexpected, since *dad* mutants show no visible phenotype in the adult wing (Ogiso *et al.* 2011).

The proximity of *dad* enhancers to the *ap* transcription unit in the *ap^Xa* chromosome suggests a plausible explanation for the Xasta wing phenotype. Two *cis*-regulatory elements, Dad4 and DadInt52, are located in the *dad* introns (Figure 2E) and are known to drive reporter gene expression in the wing disc in a stripe along the A−P compartment boundary in response to Dpp signaling (Weiss *et al.* 2010). Because *dad* territory encompasses not only the dorsal but also the ventral compartment of the presumptive wing pouch, a likely scenario is that the *ap* promoter responds to these two *dad* enhancers, leading to ectopic Ap expression in the ventral compartment of the pouch.

Ectopic expression of *ap* in *ap^Xa* leads to the ectopic expression of Ap target genes

To further characterize the effect of the *ap^blot* and *ap^Xa* mutations on wing development, we examined *ap* mRNA and Wingless protein (Wg) expression in 3rd instar larval wing discs (Figure 3, A and B). In wild-type discs, *ap* mRNA is restricted to the dorsal compartment of the wing pouch, the hinge and the notum (Figure 3A). In the pouch, Ap activity is required to direct the expression of Wg in a stripe along
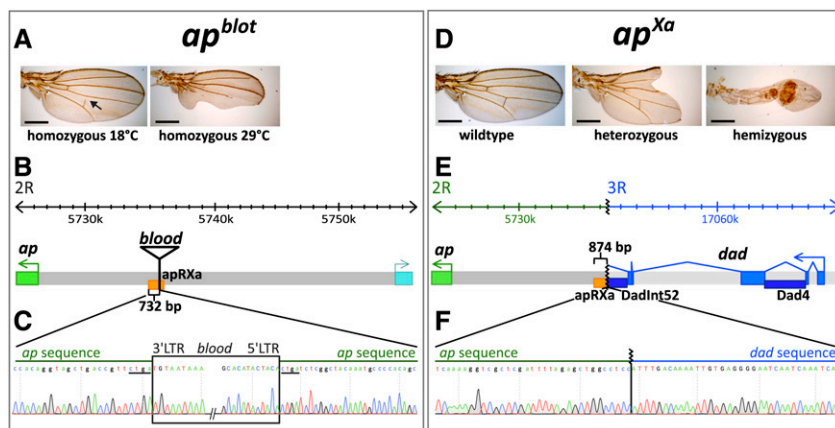
Figure 2 The mutations *ap^blot* and *ap^Xa* affect the *ap* wing enhancer region. (A) Temperature-sensitive wing phenotypes obtained for the homozygous *ap^blot* allele. At 18°, less than 30% of the wings are affected and most of them only show a disruption of the posterior crossvein (arrow). At 29°, ~70% of the wings have a phenotype. In many of them, the posterior compartment is severely affected. (B) At the top of the panel, the coordinates of the *apterous* locus are indicated. The insertion site of *blood*, a retrotransposable element, within the apRXa wing enhancer is depicted. (C) Sequence data close to the insertion site of the *blood* element in *ap^blot*. The insertion causes a four bp duplication (CTGA, underlined). Exact coordinates of the 4 bp duplication: 2R:5735176.0.5735179 (Flybase Release FB2014_06). (D) Preparations of wild type and *ap^Xa* mutant wings. All *ap^Xa*/+ flies show a dominant phenotype: the distal part of the wing blade is lost and the characteristic mitten phenotype is formed. In hemizygous condition, the wing tissue of *ap^Xa*/*ap^DG3* flies forms a short tube-like structure. Margin bristles are absent except for sometimes a few at the tip. All scale bars are 50 μm. (E) Molecular characteristics of the *ap^Xa* mutation. A reciprocal translocation involving the right arms of the second and third chromosome causes a breakpoint just upstream of the apRXa wing enhancer (indicated in orange) and juxtaposes the *daughters against dpp* (*dad*) locus (indicated in blue) next to the *ap* gene. The dark blue rectangles represent the well-studied *cis*-regulatory elements Dadint52 and Dad4 which are active in the wing disc (Weiss *et al.* 2010). (F) Chromatograph of the *ap^Xa* sequence across the rearrangement break point. The coordinates of the breakpoints are: 2R:5375319 and 3R:17065902 (Flybase Release FB2014_06).

Temperature	Total Wings Scored	Normal Wings	Wings with Phenotypes
18°	294	72%	28%
25°	284	48%	52%
29°	242	30%	70%

the D−V compartment boundary (Figure 3B). This Wg stripe is essential for the proper formation of the wing margin (Couso *et al.* 1994).

The temperature sensitivity of *ap^blot* was faithfully recapitulated by the expression patterns in 3rd instar wing discs. Although *ap* mRNA levels were reduced at 18° as well as at 29°, an obvious deviation of the *ap* mRNA pattern was only observed at 29° in the posterior compartment of the pouch. This change correlated with a size reduction of the posterior compartment and the appearance of additional tissue folding in this region (arrow in Figure 3A). Consistent with the sharp boundary of the *ap* mRNA expression pattern at 18°, the Wg stripe along the D-V compartment boundary remained unchanged (Figure 3B). In contrast, at 29°, the fuzzy appearance of the *ap* mRNA pattern in the posterior compartment correlated with the disruption of the Wg stripe. In summary, these results are consistent with the adult wing phenotypes and provide an explanation for the abnormalities in the posterior wing margin as well as for the reduced size of the posterior compartment in *ap^blot* flies raised at elevated temperature.

In *ap^Xa* heterozygotes, a strong ectopic misexpression of the *ap* transcript was detected in the ventral compartment of the wing disc,

with the highest signal along the medial part of the disc (Figure 3A). As a consequence, the Wg stripe was disrupted in the medial region of the wing pouch (Figure 3B). Remarkably, the disruption of the Wg stripe correlated well with the expression domain of the *Dad4-GFP* reporter construct (Figure 3C). Wherever GFP was detected, the expression of Wg was either very low or absent. Wing discs of hemizygous *ap^Xa/ap^DG3* flies showed strong *ap* expression in the entire pouch region. The characteristic Wg stripe in the wing pouch was lost, leaving behind only a small dot of Wg expression in the middle of the pouch. Moreover, the dimension of the wing pouch was reduced to about half the size of a wild-type pouch.

In *Drosophila*, the somatic pairing of the two homologous chromosomes can lead to a special situation of gene regulation called transvection (Lewis 1954; Sipos *et al.* 1998; Morris *et al.* 1999; Coulthard *et al.* 2005). In this case the regulatory elements of a gene can regulate the expression of its homolog *in trans*. Transvection has been described for many gene loci (for reviews see Wu and Morris 1999; Duncan 2002) including the *ap* locus (Gohl *et al.* 2008). Therefore, we decided to test the transvection ability of *ap^Xa* by crossing it with ap::GFP. In this combination only the gene *in trans* is labeled with GFP, allowing for the independent detection of the gene product from this chromosome. Trans-heterozygous *ap::GFP/ap^Xa* flies displayed no ectopic expression of Ap-GFP in the ventral wing pouch (Figure 3D). This result demonstrates that the misexpression of *ap* is limited to the chromosome affected by the rearrangement.

As a selector gene, *ap* is known to regulate multiple downstream genes (Bronstein *et al.* 2010). We wished to know whether the ectopic
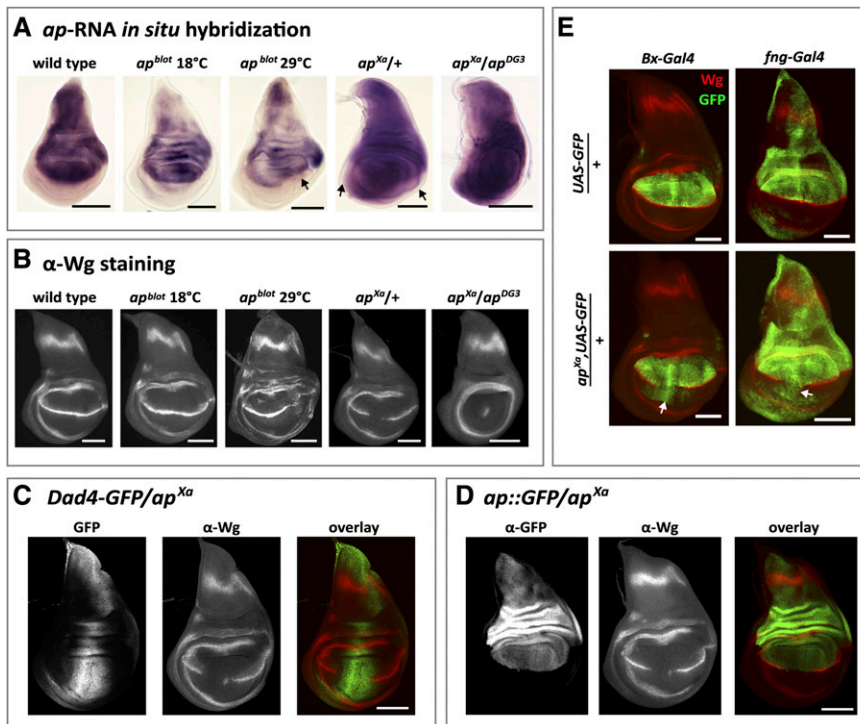
Figure 3 Wing disc phenotypes in *ap^blot* and *ap^Xa*. All discs are shown anterior to the left and dorsal side up. (A) *in situ* hybridization against *ap* mRNA in late 3rd instar larval wing discs. In wild type, the dorsal compartment of the wing pouch is filled and outlined by the *ap* transcript. *ap^blot* discs show reduced *ap* mRNA levels. At 18°, the *ap* expression pattern remains very similar as that in wild type. At 29°, expression of *ap* in the posterior compartment is disturbed and the tissue is deformed (see arrow). In heterozygous *ap^Xa* discs, ectopic *ap* expression is seen in the ventral part of the wing disc, with the strongest signal in median regions. The black arrows point to the edges of the disc where *ap* transcript is absent. In hemizygous *ap^Xa/ap^DG3* larvae, a similar pattern is observed. Note the change in shape of the wing disc. (B) α-Wg antibody staining of 3rd instar wing discs. In wild type, a characteristic thin stripe of Wg traverses the wing pouch along the D-V compartment boundary. In *ap^blot*, Wg expression is normal at 18°C. At 29°C, the Wg stripe is much weaker and less well defined in posterior regions of the wing pouch. In *ap^Xa*/+ discs, the Wg stripe is interrupted in the median pouch region. In hemizygous *ap^Xa* discs, the Wg stripe is lost and only a dot of Wg expression in the middle of the pouch is visible. In addition, the size of the pouch is reduced. (C) GFP expression driven by the Dad4 enhancer is detected in the central part of an *ap^Xa*/+ wing disc. Note that absence of Wg stripe correlates well with higher GFP levels. Therefore, stripe formation is more affected in the anterior than in the posterior compartment. (D) α-GFP and α-Wg antibody staining of an *ap::GFP/ap^Xa* wing disc. GFP expression is restricted to the dorsal compartment of the wing pouch. In particular, Ap-GFP fusion protein does not spread ventrally where the Wg stripe is interrupted. This indicates that *dad* enhancers on the *ap^Xasta* chromosome are unable to activate *ap::GFP* located on the homologous chromosome. (E) Expression of *Beadex*- and *fringe-Gal4* enhancer trap lines in wild type and *ap^Xa*/+ discs. Note that ectopic expression (white arrows) of these two validated Ap targets in the ventral compartment is only detected where the Wg stripe is interrupted. All scale bars are 100 μm.

Ap expression observed in heterozygous ap^{Xa} flies was sufficient to induce its targets also in the ventral compartment of the pouch. Toward that end, we analyzed Gal4-enhancer trap lines of two validated Ap target genes, *Beadex* (*Bx*) and *fringe* (*fng*) (Irvine and Wieschaus 1994; Milán *et al.* 2004). Their activity was monitored with the help of a UAS-CD8-GFP transgene. Under wild-type conditions, *Beadex > GFP* expression was detected exclusively in the dorsal wing pouch, whereas *fng > GFP* was observed predominantly in the whole dorsal compartment with weak ventral GFP outside the wing pouch (Figure 3E). When analyzed in a heterozygous ap^{Xa} background, the expression of both reporters extended to the medial-ventral region of the wing discs (Figure 3E). In summary, the data presented in Figure 3 strongly suggest that in ap^{Xa}, *ap* is ectopically expressed due to the juxtaposition of *dad* wing enhancer elements and the *ap* promoter. As a consequence, Ap target genes are up-regulated in the ventral compartment of the wing pouch. These molecular events correlate well with the disruption of the Wg stripe in median parts of the wing disc and, finally, the altered adult wing morpholgy.

Requirements for ectopic wing margin induction

A puzzling observation is that Ap expression does not induce Wg in the ventral part of the pouch, ultimately leading to extra margin formation in adult wings. One explanation for this finding is that compartment and compartment boundaries must be defined by a clear on-off state of selector gene activity. However, the *dad* gene and its enhancers are regulated by a Dpp concentration gradient (high Dpp in medial parts, low Dpp in lateral parts of the wing disc; for a review see Affolter and Basler 2007). As a consequence, in ap^{Xa}, Ap is expressed in the ventral part of the wing pouch in response to the *dad* enhancers in a gradient-like manner. Thus, no clear selector gene on-off state between neighboring cells would be generated. In this case, the initiation of the signaling cascade that usually induces the *wg* gene at the compartment boundary fails to be activated. To explore this possibility in more detail, we used the Gal4/UAS-system (Brand and Perrimon 1993). Preliminary test crosses indicated that upon Gal4 activation, a UAS-ap transgene leads to lethality or pleiotropic phenotypes with all Gal4 drivers tested except *dpp-Gal4*. For this reason, we tested an insertion into *ap*, EY03046, which contains a UAS-driven promoter located several 100 bp upstream of the *ap* TSS (Figure 4F). In contrast to *Gal4 > UAS-ap* combinations, *Gal4 > EY03046* flies were viable and obvious phenotypes were restricted to the dorsal thoracic appendages. One possible explanation for the difference is that the activation of EY03046 by Gal4 is reduced or eliminated by the *ap* PRE (Schwartz *et al.* 2006; Tolhuis *et al.* 2006; Oktaba *et al.* 2008; D. Bieli *et al.*, unpublished data) in most tissues outside of the wing disc. To activate EY03046 expression in the wing pouch, we used the following Gal4 drivers: *actin-Gal4*, *dad4-Gal4*, *salE-Gal4*, *dpp-Gal4*, and *ptc-Gal4*. Their expression domains are depicted in Figure 4, A−E. For our purposes, they can be grouped into three classes: (1) *actin > GFP* is found in all cells of the pouch; (2) *Dad4 > GFP* and *salE > GFP* expression domains are rather broad with a rather ill-defined edge and centered on the A-P axis; and (3) *dpp > GFP* and *ptc > GFP* form a narrow stripe along the A-P axis.

To analyze the effects of ectopic Ap expression, we examined Wg stripe formation along the D-V compartment boundary (see Figure 4, A′−E′) and adult wing morphology (Figure 4, A′′−E′′). Ubiquitous Ap expression in the pouch using *actin::Gal4* prevents Wg activation. As a consequence, margin formation in the tiny adult wings was abolished. As expected from the data in Figure 3, Dad4 and *spalt*-mediated Ap expressions led to Xasta phenocopies. Wg stripe formation was abolished in the center of the pouch. Occasionally, small

ectopic Wg stripes extended into the ventral compartment in *Dad4 > EY03046* wing discs. Nevertheless, both Gal4 drivers elicited similar moderate Xasta-like phenotypes in adult wings. Finally, although the expression patterns mediated by the *dpp-Gal4* and *ptc-Gal4* drivers were remarkably similar, their phenotypic consequences were dramatically different. *dpp > EY03046* caused the appearance of a faint ectopic Wg patch on the A-P axis in the ventral compartment. A tuft of ectopic bristles was observed on the ventral side at the intersection of the A-P axis and the wing margin in less than 10% of the adult wings (black arrowhead in Figure 4D′′). *ptc > EY03046*, on the other hand, interrupted the Wg stripe in the center of the pouch and frequently induced a well-defined Wg stripe which traversed the whole ventral compartment. A Wg stripe of variable length also is formed in the anterior compartment. Adult wings of this genotype often formed three-dimensional, balloon-like structures with an oval-shaped posterior margin extending from the proximal edge of the wing appendage to its distal end and back to proximal. In addition, anterior and posterior margins were not continuous at the tip of the wing.

These observations corroborate our expectations. First, the presence of Ap in the ventral compartment at sufficiently high levels impedes the activation of the signaling cascade that induces Wg expression along the D−V compartment boundary. Second, an ectopic compartment boundary can only be formed between cells with sharp on-off levels of Ap. This prerequisite is only satisfied by the *ptc-Gal4* driver. In the wing disc, Ptc is expressed in a straight line immediately abutting the posterior compartment where it serves as a receptor for the Hedgehog ligand (Capdevila *et al.* 1994; Alexandre *et al.* 1996). Its anterior limit of expression is more graded and less well defined and ectopic anterior margin in the adult wing can only rarely be observed. The question remains why *dpp > EY03046* is only marginally active in this experiment. It is possible that ectopic Ap levels remain below a certain threshold because the levels of Gal4 do not suffice. Alternatively, the onset of Gal4 activity in this driver line might be delayed. However, very similar observations were made with a UAS-ap transgene in place of EY03046 and also with a different *dpp-Gal4* driver line. In addition, Klein *et al.* have reported a similar phenotype for ectopic Ser expression by *dpp-Gal4* (Klein *et al.* 1998).

The *in situ* rescue system

To extend our analysis of the *cis*-regulatory elements directing *ap* expression, we decided to characterize and manipulate possible regulatory sequences directly at the endogenous locus. For this purpose, we engineered an *in situ* rescue system. The establishment of this system was a multistep procedure and is described in detail in the section *Materials and Methods*. A diagrammatic summary is presented in Figure 5. In brief, we deleted the 27-kb intergenic spacer between the *ap* and *l(2) 09851* loci and replaced it with an attP site located 400 bp upstream of the *ap* TSS (Figure 5, A−D). This *ap* allele is referred to as $ap^{attP\Delta Enh}$ (Figure 5D). The deleted region is identical to that of ap^{DG1}. Therefore, homo- or hemi-zygous $ap^{attP\Delta Enh}$ flies have no wings (data not shown). The attP site of $ap^{attP\Delta Enh}$ serves as docking site for $\Phi C31$-mediated integration of any desired DNA located on a plasmid containing an attB site and the *yellow* selection marker (Figure 5, D and E).

As proof of principle, two control plasmids were first introduced into $ap^{attP\Delta Enh}$: (1) the empty pEnh-Reentry vector gave rise to a fly line called ap^{empty}; (2) pEnh-Reentry-full-length contained the complete 27-kb intergenic spacer and the corresponding transgenic line was called $ap^{full-length}$ (Figure 6A). The "wing-forming" activity of these two controls as well as all subsequent transgenic lines analyzed in this study was determined in hemizygous condition. Therefore, balanced ap^{empty} and $ap^{full-length}$ males were crossed with ap^{DG3}/SM6 virgins and
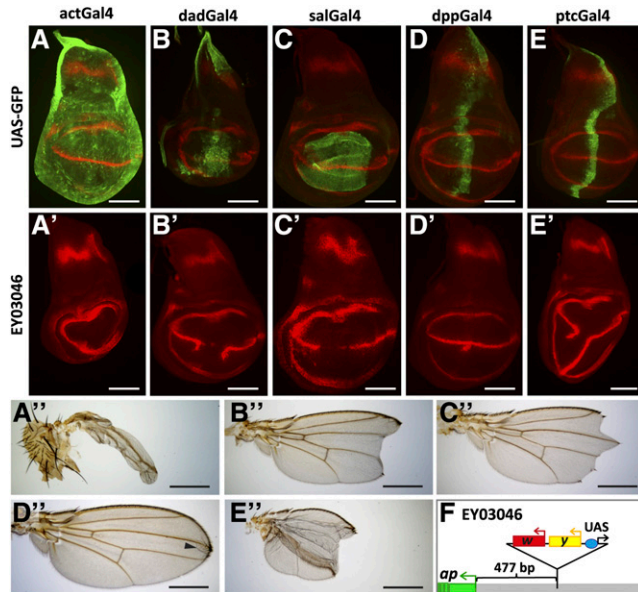
Figure 4 Margin formation in adult wings depends on well-defined On-Off Apterous expression levels during larval development. All discs are shown anterior to the left and dorsal side up. (A−E) 3rd instar imaginal wing discs showing UAS-GFP patterns (in green) elicited by the five Gal4 drivers indicated at the top of the panel. α-Wg antibody staining (in red) outlines the pouch and the position of the D−V compartment boundary. (A′−E′) α-Wg antibody staining. The effect of ectopic Ap production as a consequence of Gal4 > EY03046 on D-V boundary formation is shown. (A″−E″) Adult wings as obtained after ectopic Ap expression in (A′) *actin* > *EY03046*, (B′) *Dad4* > *EY03046*, (C′) *salE* > *EY03046*, (D′) *dpp* > *EY03046*, and (E′) *ptc* > *EY03046* animals. In (D′), the arrowhead points to a small lesion near the tip of the wing. Scale bars in (A−F) and (A′−F′) are 100 μm. Scale bars in (A″−F″) are 50 μm. (F) Insertion site of *P{EPgy2}EY03046* relative to the *ap* TSS is shown. The triangle depicts the structure of the transgene. The red box corresponds to the *mini-white* marker, the yellow box to the *yellow* marker and the blue oval to an array of UAS sites. Arrows specify the transcriptional direction of *mini-white*, *yellow*, and the UAS-driven promoter. *P{EPgy2}* transgenes are intended for regulated expression of genes proximate to the site of the insertion: genes in direct orientation with respect to the UAS-controlled promoter can be conditionally expressed via transgene-derived Gal4 activity (Bellen *et al.* 2004). Note that at *apterous*, the UAS-driven promoter is at a considerable distance from and in opposite orientation to the *apterous* promoter (shown in green). We propose that in *Gal4 > EY03046* flies, Gal4 activates *ap* transcription in much the same way as the eye-specific *GMR-Gal4{w}* driver boosts *mini white* expression in *GMR-Gal4{w}/EY03046* flies. These have red eyes while the eye pigmentation in *EY03046/+* flies is faint yellow (M. Müller, unpublished data). Drawing not to scale.

the wings of *trans*-heterozygous progeny were carefully inspected. As expected, ap^{empty}/ap^{DG3} flies generated no detectable wing material. In contrast, the reconstituted *ap* locus produced wild-type wings in $ap^{full-length}/ap^{DG3}$ flies. Taken together, these observations demonstrate the feasibility of our *in situ* rescue system and suggest that the backbone of the pEnh-Reentry plasmid does not cause any disturbances.

DadInt52 and Dad4 enhancers contribute significantly to the Xasta phenotype

Our model for the Xasta wing phenotype posits that the wing specific *dad* enhancers Dad4 and DadInt52 are responsible for ectopic Ap expression in the ventral pouch compartment. We wished to test this

hypothesis with the *in situ* rescue system. Two fly lines were established: $ap^{apRXaDad52.4}$ and $ap^{Dad52.4}$ (Figure 6B). The former combined the three identified wing specific enhancers apRXa, DadInt52 and Dad4. The latter contained only the two *dad* regulatory elements. When the 2 transgenics were initially isolated, it was immediately apparent that both phenocopied the dominant *Xasta* allele. However, a semi quantitative analysis also showed that the severity of their phenotypes was weaker than observed for $ap^{Xa}/+$ wings (Table 2). Although roughly 50% of the $ap^{Dad52.4}/+$ wings were as strongly affected as those of $ap^{Xa}/+$ flies, hardly any such wings appeared in $ap^{apRXaDad52.4}/+$ flies. These observations indicate that apart from DadInt52 and Dad4, other factors contribute to the production of a full blown Xasta wing phenotype.

The wings of $ap^{apRXaDad52.4}$ and $ap^{Dad52.4}$ were also analyzed in hemizygous condition. The phenotypes were comparable to the one seen in ap^{Xa}/ap^{DG3} flies: only tube-like wing stumps were formed which lacked wing margin completely except for the occasional occurrence of a few margin hairs at the very tip. It is conceivable that the latter arise due to the Wg spot seen in the center of the pouch of ap^{Xa}/ap^{DG3} wing discs (see Figure 3B). We have never seen homozygous ap^{Xa} flies but did inspect adult wings of $ap^{Xa}/ap^{Dad52.4}$ animals. They appeared as even smaller versions of those observed in hemizygous ap^{Xa} flies (data not shown).

The apRXA enhancer is required but not sufficient for wing formation

In Figure 1 of this paper, we have presented evidence that the ~8 kb apC fragment harbors an 874-bp wing specific enhancer that is essential for wing formation. However, the experimental approaches we used are not adequate to test whether the enhancer is also sufficient for the formation of a wild-type wing. Therefore, four overlapping fragments covering the whole apC were introduced into the *ap* locus and the corresponding transgenic lines were obtained: ap^{apP}, ap^{apY}, ap^{apR}, and ap^{apRXa}. Their wing enhancer activity was tested in a hemizygous genetic background (Figure 6C). ap^{apP}/ap^{DG3} flies, which contained the apP fragment that did not yield any LacZ reporter activity (see Figure 1A), also did not develop any wing or haltere tissue and phenotypically resembled *ap* null alleles. When apY, a fragment which is shifted by 2 kb toward the *ap* TSS, was tested in ap^{apY}/ap^{DG3} flies, wing development was partially restored. However, most of the margin, the alula and the hinge region were poorly formed. Similar phenotypes as for ap^{apY} were observed in ap^{apR}/ap^{DG3} and ap^{apRXa}/ap^{DG3} flies. Note that these three apC derivatives were sufficient to drive *ap*-specific *LacZ* expression in our reporter assay (see Figure 1A). "Homozygotes" obtained by pairwise combinations of ap^{apY}, ap^{apR} or ap^{apRXa} were also studied. Such wings looked improved compared to the phenotypes observed in hemizygotes, because the margins, particularly along the anterior but also along the posterior edges of the wing, were formed to a large degree (data not shown). Somewhat unexpectedly, heterozygous ap^{apY}, ap^{apR} and ap^{apRXa} flies showed a weak dominant wing phenotype, associated with a small notch in the tip region in 10–20% of the cases. This phenotype was not observed in $ap^{apY}/+$ or $ap^{full-length}/ap^{DG3}$ flies (data not shown).

These results demonstrate that the 874 bp apRXa wing enhancer element is required but not sufficient in the endogenous context to correctly regulate *ap* expression. Our observations imply the existence of further unidentified wing enhancer elements elsewhere in the *ap* region.

DISCUSSION

In the past, *cis*-regulatory elements were mainly investigated using reporter-based assay systems, in which putative regulatory DNA
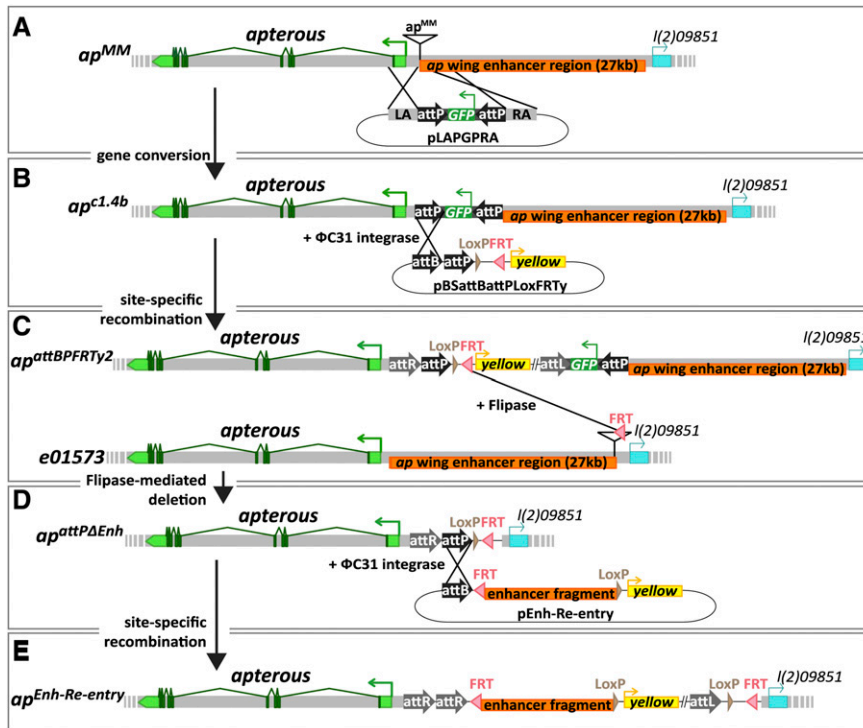
Figure 5 Generation of the *in situ* rescue system at the endogenous *ap* locus. (A−B) Direct gene conversion at *apterous*. P-element insertion *ap^MM* located ~400 bp upstream of the *ap* TSS was previously isolated. By mobilization of *ap^MM* and concomitant injection of plasmid pLAPGPRA, fly line *ap^c1.4b* could be isolated. It contains two inverted attP sites flanking a GFP reporter. (B−C) *ΦC31*-integrase mediated site-specific recombination. By injection of plasmid pBSattBattPLoxFRTy, new attP, LoxP, and FRT sites were introduced into the *ap* locus. Note that pBSattBattPLoxFRTy can insert in two different attP sites leading to oppositely oriented insertions. *ap^attBPFRTy2* is the appropriate one for our purpose. (C−D) Flipase-mediated deletion. Trans-heterozygous *ap^attBPFRTy2*/*ap^e01573* animals were repeatedly treated with Flipase during larval stages. Among the progeny of these flies, *ap^attPΔEnh* could be isolated. It lacks the 27kb intergenic spacer but retains a strategically positioned attP site. (D−E) *ap^attPΔEnh* serves as a platform to reinsert enhancer fragments. These are cloned into pEnh-Reentry. This plasmid is injected into young embryos and integrates into the *ap* locus by *ΦC31*-integrase mediated recombination. Transgenics of the type *ap^Enh-Reentry* can be isolated thanks to the *yellow* marker. If desired, *yellow* can be removed by Cre-treatment. In addition, the complete insert can be excised by Flipase treatment.

fragments were tested for their ability to drive reporter gene expression when present on a transgene inserted randomly in the genome (Simon *et al.* 1985; Hiromi and Gehring 1987). Although this method proved to be a highly useful and valuable approach, it has some shortcomings. Enhancer fragments are tested in a genomic environment that may differ considerably from their native position. Additionally, the results of such studies yield little or no information about whether the investigated elements are sufficient, permissive or even dispensable for the regulation of gene expression at their original location. Recently, some improvements were achieved by using bacterial artificial chromosomes to investigate *cis*-regulatory elements in a broader genomic context (Dunipace *et al.* 2013).

To circumvent the problem of positional effects, we performed our classical reporter assay at a single *ΦC31*-system docking site located on 3R. Our laboratory has successfully used this insertion site for the analysis of wing specific enhancer elements (Weiss *et al.* 2010). Furthermore, we investigated the relevance of the reporter data with two powerful genetic approaches. We used methods from the *Drosophila* genetic tool kit and generated useful materials for the *in situ* dissection of regulatory elements directly at the *ap* locus. First, a set of small overlapping deletions within the *ap* region was isolated with the help of different transposable elements carrying FRT sites. Second, the *in situ* rescue system was established. The novel fly strain generated greatly facilitates the introduction of any DNA fragment by means of integrase-mediated recombination into the *apterous* locus. It has and will serve us as a tool to dissect important *ap* regulatory sequences in great detail.

In this study, the combined application of reporter assay, deletion analysis and *in situ* rescue system has allowed us to firmly establish the 874 bp apRXa fragment as an essential wing-specific regulatory element

for *apterous* transcription. We show that apRXa is sufficient to drive reporter gene expression within the dorsal compartment of the wing pouch. Flies hemizygous for an 11-kb deletion encompassing the apRXa element develop no wing structures. This observation proves that this larger DNA interval including apRXa is essential for *ap* function. Finally, when tested in the context of the endogenous *ap* locus, we document that apRXa is required but not sufficient to form wild-type wings.

The importance of the apRXa enhancer element is further highlighted through the molecular characterization of 2 classical *ap* alleles, *ap^blot* and *ap^Xa*. *ap^blot* contains an insertion of a retrotransposon from the *blood* family. This insertion is located within the apRXa enhancer. We have not attempted to prove the presence of the full length 7.4-kb *blood* element in *ap^blot*, but we have completely sequenced both LTRs. So far, all *blood* elements detected in the *Drosophila* genome are full-length insertions. None of them was found to be truncated (Kaminker *et al.* 2002). Hence, it appears likely that *ap^blot* also contains an intact, full length *blood* element and that it is accountable for the mutagenic effect. For example, it is possible that the insertion destroys an important transcription factor binding site within the apRXa wing enhancer. Alternatively, the inserted DNA might separate important transcription factor binding sites.

The other *ap* allele we investigated is *ap^Xa*. In this mutant, a reciprocal translocation event between the right arms of the second and third chromosomes caused a breakpoint immediately upstream of the apRXa wing enhancer. This rearrangement juxtaposes the *dad* locus next to *apterous*. Our experimental evidence strongly indicates that in this mutant, *ap* transcription falls under the control of *dad* wing specific enhancers Dad4 and DadInt52. As a consequence, *ap* and its target genes are ectopically expressed in the medial section of the ventral part of the wing disc, conferring ventral cells with a dorsal cell fate identity.
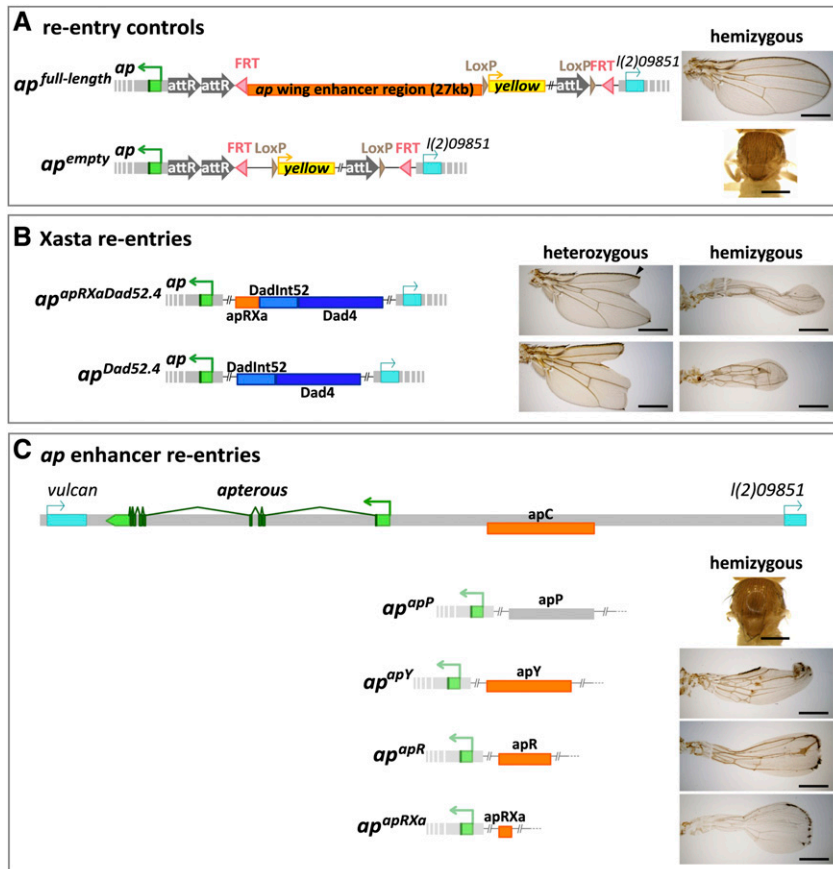
Figure 6 Testing *dad* and *ap* enhancers in the endogenous *apterous* locus. (A) Positive control: the whole 27-kb *ap* wing enhancer region was re-inserted and *ap^full-length* flies obtained. In *ap^full-length*/*ap^DG3* animals, perfectly wild-type wings are formed. Negative control: the empty pEnh-Reentry plasmid gave rise to *ap^empty* flies. No wing tissue is formed in *ap^empty*/*ap^DG3* adults. (B) *Xasta* phenocopies are obtained with *ap^apRXaDad52.4* and *ap^Dad52.4* alleles. *ap^apRXaDad52.4* contains three enhancer elements: apRXa, DadInt52, and Dad4. Heterozygous flies only produce rather weak phenocopies. The junction between wing vein L2 and the margin (see arrow head) is present in almost 100% of the wings. *ap^Dad52.4* contains only the 2 *dad* enhancers. Faithful phenocopies of *ap^Xa*/+ wings where the junction between vein L2 and the margin is missing are often observed. The wing phenotypes of hemizygous *ap^apRXaDad52.4*, *ap^Dad52.4* and *ap^Xa* are similar: tube-like wing stumps of variable length are formed. Margin bristles are absent except for sometimes a few at the tip of the wing. (C) Testing the wing enhancer activity of four apC derivatives. At the top of the panel, the *ap* locus is depicted. Below, the positions of fragments apP, apY, apR and apRXa are shown relative to apC. The respective wing phenotypes in hemizygous condition are shown to the right of the corresponding fragments. Flies transgenic for the gray apP fragment behave like a true null allele: no wings are formed. Fragments drawn in orange have partial rescue activity: inflated wings are formed, where most of the margin and the alula are missing. The hinge is poorly formed. Note that in B and C, for space reasons, parts of the reentry plasmid have been omitted. All scale bars are 50 μm.

This, in turn, likely interferes with signaling at the D-V compartment boundary and causes the disruption of the Wg stripe in the center of the wing pouch. *dad*-controlled *ap* expression also provides an explanation why the anterior compartment is more strongly affected than the posterior one in adult wings. As evidenced by asymmetric Dad4-GFP expression along the A-P compartment boundary, a wider domain with higher levels of GFP is produced in the anterior compartment (see Figure 3C). We propose that a similar asymmetrical distribution of Ap causes differential Wg stripe expression in the two compartments.

Our observations also suggest that *dad*-mediated transcriptional activation of *ap* is not the sole cause for the explanation of the *Xasta* phenotype. The dominant phenotypes of *ap^apRXaDad52.4*/+ and *ap^Dad52.4*/+ flies are clearly less pronounced than that documented for *ap^Xa*/+. Why should this be the case? It is known that the *apterous* locus is a target of the repressive Polycomb Group (PcG) system. *Scm^-/-* clones reaching into the ventral compartment elicit ectopic Ap expression (Oktaba *et al.* 2008). In addition, it is well documented that the silencing activity of isolated Polycomb Response Elements is pairing dependent (Kassis *et al.* 1991; Fauvarque and Dura 1993; Chan *et al.* 1994; Gindhart and Kaufman 1995; Muller *et al.* 1999; reviewed in Kassis 2002). It is therefore conceivable that in *ap^Xa*/+ flies, the chromosomal rearrangement prevents efficient homologous chromosome pairing and thus reduced PcG-mediated silencing. This hypothesis is supported by the fact that the *mini-white* markers of transgenes inserted in the *ap* locus are partially derepressed in a *Xasta* heterozy-

gous background (M. Müller, unpublished data). An alternative explanation could be that as yet-uncharacterized wing-specific enhancers are present in the genomic *dad* locus which lack in *ap^apRXaDad52.4*/+ and *ap^Dad52.4*/+ flies. It thus might be that the stronger phenotype observed in *ap^Xa*/+ flies is a consequence of stronger Ap misexpression due to the combined effect of more than two enhancers.

Our findings imply the existence of other, as yet unidentified wing-specific regulatory elements within the realms of the *apterous* locus. A hint about the possible location of such sequences has previously been obtained through the genetic analysis of insertion *PBac{WH}f00451*. This transposon is located about 3 kb distal to apRXa. The PBac{WH} element contains an array of Su(Hw) binding sites at its 3' end (Thibault *et al.* 2004). It is well established that on transgenes, a cluster of Su(Hw) binding sites acts as enhancer blocker. It interferes with enhancer-promoter interaction when placed in between two such regulatory

Table 2 Penetrance of the dominant *ap^Xa* wing phenotype

Genotype	Number of Wings Scored[a]	L2 Junction Present	L2 Junction Absent
ap^Xa/+	262	6.5%	93.5%
ap^apRXaDad52.4/+	160	98.1%	1.9%
ap^Dad52.4/+	546	58.6%	41.4%

[a] Wings were scored for the presence or absence of the junction between wing vein L2 and the wing margin.

elements (Holdridge and Dorsett 1991; Hagstrom *et al.* 1996; Scott *et al.* 1999). Furthermore, the mutagenic effect of an array of Su (Hw) binding sites located on the *gypsy* mobile genetic element can often be attributed to a similar mechanism (Geyer *et al.* 1986; Peifer and Bender 1988; Dorsett 1993; Hogga *et al.* 2001). Homozygous as well as hemizygous *ap^{f00451}* flies have been reported to cause a rather strong wing phenotype. Importantly, the phenotype is completely suppressed in a *su(Hw)^-* background (Gohl *et al.* 2008). This observation suggests the presence of other *ap* wing enhancer elements distally to *PBac{WH}f00451*. We are currently exploring this possibility with analogous experimental approaches as outlined above, in particular through the use of the *in situ* rescue system established and described in this study (D. Bieli *et al.*, unpublished results).

LITERATURE CITED

Affolter, M., and K. Basler, 2007   The decapentaplegic morphogen gradient: from pattern formation to growth regulation. Nat. Rev. Genet. 8: 663–674.

Ahmad, K., and S. Henikoff, 2001   Modulation of a transcription factor counteracts heterochromatic gene silencing in *Drosophila*. Cell 104: 839–847.

Alexandre, C., A. Jacinto, and P. W. Ingham, 1996   Transcriptional activation of hedgehog target genes in *Drosophila* is mediated directly by the cubitus interruptus protein, a member of the GLI family of zinc finger DNA-binding proteins. Genes Dev. 10: 2003–2013.

Alexandre, C., A. Baena-Lopez, and J.-P. Vincent, 2014   Patterning and growth control by membrane-tethered Wingless. Nature 505: 180–185.

Ashburner, M., 1989   Drosophila: A Laboratory Handbook and Manual (two volumes). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Bateman, J. R., A. M. Lee, and C. T. Wu, 2006   Site-specific transformation of *Drosophila* via phiC31 integrase-mediated cassette exchange. Genetics 173: 769–777.

Bellen, H. J., R. W. Levis, G. Liao, Y. He, J. W. Carlson *et al.*, 2004   The BDGP gene disruption project: single transposon insertions associated with 40% of *Drosophila* genes. Genetics 167: 761–781.

Bingham, P. M., and C. H. Chapman, 1986   Evidence that white-blood is a novel type of temperature-sensitive mutation resulting from temperature-dependent effects of a transposon insertion on formation of white transcripts. EMBO J. 5: 3343–3351.

Bischof, J., R. K. Maeda, M. Hediger, F. Karch, and K. Basler, 2007   An optimized transgenesis system for *Drosophila* using germ-line-specific phiC31 integrases. Proc. Natl. Acad. Sci. USA 104: 3312–3317.

Blair, S. S., D. L. Brower, J. B. Thomas, and M. Zavortink, 1994   The role of apterous in the control of dorsoventral compartmentalization and PS integrin gene expression in the developing wing of *Drosophila*. Development 120: 1805–1815.

Bourgouin, C., S. E. Lundgren, and J. B. Thomas, 1992   apterous is a drosophila LIM domain gene required for the development of a subset of embryonic muscles. Neuron 9: 549–561.

Bronstein, R., L. Levkovitz, N. Yosef, M. Yanku, E. Ruppin *et al.*, 2010   Transcriptional regulation by CHIP/LDB complexes. PLoS Genet. 6: e1001063.

Butterworth, F. M., and R. C. King, 1965   The developmental genetics of apterous mutants of *Drosophila melanogaster*. Genetics 52: 1153–1174.

Capdevila, J., M. P. Estrada, E. Sánchez-Herrero, and I. Guerrero, 1994   The *Drosophila* segment polarity gene patched interacts with decapentaplegic in wing development. EMBO J. 13: 71–82.

Caussinus, E., O. Kanca, and M. Affolter, 2011   Fluorescent fusion protein knockout mediated by anti-GFP nanobody. Nat. Struct. Mol. Biol. 19: 117–121.

Chan, C. S., L. Rastelli, and V. Pirrotta, 1994   A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression. EMBO J. 13: 2553–2564.

Cohen, B., M. E. McGuffin, C. Pfeifle, D. Segal, and S. M. Cohen, 1992   apterous, a gene required for imaginal disc development in *Drosophila* encodes a member of the LIM family of developmental regulatory proteins. Genes Dev. 6: 715–729.

Coulthard, A. B., N. Nolan, J. B. Bell, and A. J. Hilliker, 2005   Transvection at the vestigial locus of *Drosophila melanogaster*. Genetics 170: 1711–1721.

Couso, J. P., S. A. Bishop, and A. Martinez Arias, 1994   The wingless signalling pathway and the patterning of the wing margin in *Drosophila*. Development 120: 621–636.

Couso, J. P., E. Knust, and A. Martinez Arias, 1995   Serrate and wingless cooperate to induce vestigial gene expression and wing formation in *Drosophila*. Curr. Biol. 5: 1437–1448.

Dahmann, C., and K. Basler, 1999   Compartment boundaries: at the edge of development. Trends Genet. 15: 320–326.

Diaz-Benjumea, F. J., and S. M. Cohen, 1993   Interaction between dorsal and ventral cells in the imaginal disc directs wing development in *Drosophila*. Cell 75: 741–752.

Diaz-Benjumea, F. J., and S. M. Cohen, 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.

Dorsett, D., 1993   Distance-independent inactivation of an enhancer by the suppressor of Hairy-wing DNA-binding protein of *Drosophila*. Genetics 134: 1135–1144.

Duncan, I. W., 2002   Transvection effects in Drosophila. Annu. Rev. Genet. 36: 521–556.

Dunipace, L., A. Saunders, H. Ashe, and A. Stathopoulos, 2013   Autoregulatory feedback controls sequential action of cis-regulatory modules at the brinker locus. Dev. Cell 26: 536–543.

Fauvarque, M. O., and J. M. Dura, 1993   polyhomeotic regulatory sequences induce developmental regulator-dependent variegation and targeted P-element insertions in *Drosophila*. Genes Dev. 7: 1508–1520.

Garcia-Bellido, A., P. Ripoll, and G. Morata, 1973   Developmental compartmentalisation of the wing disk of *Drosophila*. Nat. New Biol. 245: 251–253.

Geyer, P. K., and V. G. Corces, 1987   Separate regulatory elements are responsible for the complex pattern of tissue-specific and developmental transcription of the yellow locus in *Drosophila melanogaster*. Genes Dev. 1: 996–1004.

Geyer, P. K., C. Spana, and V. G. Corces, 1986   On the molecular mechanism of gypsy-induced mutations at the yellow locus of *Drosophila melanogaster*. EMBO J. 5: 2657–2662.

Gindhart, J. G., and T. C. Kaufman, 1995   Identification of Polycomb and trithorax group responsive elements in the regulatory region of the *Drosophila* homeotic gene Sex combs reduced. Genetics 139: 797–814.

Gloor, G. B., N. A. Nassif, D. M. Johnson-Schlitz, C. R. Preston, and W. R. Engels, 1991 Targeted gene replacement in *Drosophila* via P element-induced gap repair. Science 253: 1110–1117.

Gohl, D., M. Müller, V. Pirrotta, M. Affolter, and P. Schedl, 2008 Enhancer blocking and transvection at the *Drosophila apterous* locus. Genetics 178: 127–143.

Golic, K. G., and M. M. Golic, 1996 Engineering the *Drosophila* genome: chromosome rearrangements by design. Genetics 144: 1693–1711.

Hagstrom, K., M. Muller, and P. Schedl, 1996 Fab-7 functions as a chromatin domain boundary to ensure proper segment specification by the *Drosophila* bithorax complex. Genes Dev. 10: 3202–3215.

Henderson, K. D., D. D. Isaac, and D. J. Andrew, 1999 Cell fate specification in the *Drosophila* salivary gland: the integration of homeotic gene function with the DPP signaling cascade. Dev. Biol. 205: 10–21.

Hetherington, C. M., W. J. Whittington, M. A. Hossain, and W. E. Peat, 1968 The genetics of the Xasta mutant of *Drosophila melanogaster*. Genet. Res. 12: 285–294.

Hiromi, Y., and W. J. Gehring, 1987 Regulation and function of the *Drosophila* segmentation gene fushi tarazu. Cell 50: 963–974.

Hogga, I., J. Mihaly, S. Barges, and F. Karch, 2001 Replacement of Fab-7 by the gypsy or scs insulator disrupts long-distance regulatory interactions in the Abd-B gene of the bithorax complex. Mol. Cell 8: 1145–1151.

Holdridge, C., and D. Dorsett, 1991 Repression of hsp70 heat shock gene transcription by the suppressor of hairy-wing protein of *Drosophila melanogaster*. Mol. Cell. Biol. 11: 1894–1900.

Irvine, K. D., and E. Wieschaus, 1994 fringe, a Boundary-specific signaling molecule, mediates interactions between dorsal and ventral cells during *Drosophila* wing development. Cell 79: 595–606.

Kaminker, J. S., C. M. Bergman, B. Kronmiller, J. Carlson, R. Svirskas et al., 2002 The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. Genome Biol. 3: RESEARCH0084.

Kassis, J. A., 2002 Pairing-sensitive silencing, polycomb group response elements, and transposon homing in *Drosophila*. Adv. Genet. 46: 421–438.

Kassis, J. A., E. P. VanSickle, and S. M. Sensabaugh, 1991 A fragment of engrailed regulatory DNA can mediate transvection of the white gene in *Drosophila*. Genetics 128: 751–761.

Kim, J., K. D. Irvine, and S. B. Carroll, 1995 Cell recognition, signal induction, and symmetrical gene activation at the dorsal-ventral boundary of the developing *Drosophila* wing. Cell 82: 795–802.

Klein, T., J. P. Couso, and A. Martinez Arias, 1998 Wing development and specification of dorsal cell fates in the absence of apterous in *Drosophila*. Curr. Biol. 8: 417–420.

Lawrence, P. A., and G. Morata, 1977 The early development of mesothoracic compartments in *Drosophila*. An analysis of cell lineage and fate mapping and an assessment of methods. Dev. Biol. 56: 40–51.

Lewis, E. B., 1951 Additions and corrections to the cytology of rearrangements. Drosoph. Inf. Serv. 25: 108–109.

Lewis, E. B., 1954 The theory and application of a new method of detecting chromosomal rearrangements in *Drosophila melanogaster*. Am. Nat. 88: 225–239.

Lundgren, S. E., C. A Callahan, S Thor, and J. B. Thomas, 1995 Control of neuronal pathway selection by the *Drosophila* LIM homeodomain gene apterous. Development 121: 1769–1773.

Milán, M., and S. M. Cohen, 1999 Regulation of LIM homeodomain activity in vivo: a tetramer of dLDB and apterous confers activity and capacity for regulation by dLMO. Mol. Cell 4: 267–273.

Milán, M., T. T. Pham, and S. M. Cohen, 2004 Osa modulates the expression of Apterous target genes in the *Drosophila* wing. Mech. Dev. 121: 491–497.

Morris, J. R., J. Chen, S. T. Filandrinos, R. C. Dunn, R. Fisk et al., 1999 An analysis of transvection at the yellow locus of *Drosophila melanogaster*. Genetics 151: 633–651.

Mosimann, C., G. Hausmann, and K. Basler, 2006 Parafibromin/Hyrax activates Wnt/Wg target gene transcription by direct association with beta-catenin/Armadillo. Cell 125: 327–341.

Muller, M., K. Hagstrom, H. Gyurkovics, V. Pirrotta, and P. Schedl, 1999 The mcp element from the *Drosophila melanogaster* bithorax complex mediates long-distance regulatory interactions. Genetics 153: 1333–1356.

Neumann, C. J., and S. M. Cohen, 1997 Long-range action of Wingless organizes the dorsal-ventral axis of the *Drosophila* wing. Development 124: 871–880.

Ochman, H., A. S. Gerber, and D. L. Hartl, 1988 Genetic applications of an inverse polymerase chain reaction. Genetics 120: 621–623.

Ogiso, Y., K. Tsuneizumi, N. Masuda, M. Sato, and T. Tabata, 2011 Robustness of the Dpp morphogen activity gradient depends on negative feedback regulation by the inhibitory Smad, Dad. Dev. Growth Differ. 53: 668–678.

Oktaba, K., L. Gutiérrez, J. Gagneur, C. Girardot, A. K. Sengupta et al., 2008 Dynamic regulation by polycomb group protein complexes controls pattern formation and the cell cycle in *Drosophila*. Dev. Cell 15: 877–889.

Peifer, M., and W. Bender, 1988 Sequences of the gypsy transposon of *Drosophila* necessary for its effects on adjacent genes. Proc. Natl. Acad. Sci. USA 85: 9650–9654.

Pfeiffer, B. D., T.-T. B. Ngo, K. L. Hibbard, C. Murphy, A. Jenett et al., 2010 Refinement of tools for targeted gene expression in *Drosophila*. Genetics 186: 735–755.

Rulifson, E. J., and S. S. Blair, 1995 Notch regulates wingless expression and is not required for reception of the paracrine wingless signal during wing margin neurogenesis in *Drosophila*. Development 121: 2813–2824.

Ryder, E., M. Ashburner, R. Bautista-Llacer, J. Drummond, J. Webster et al., 2007 The DrosDel deletion collection: a *Drosophila* genomewide chromosomal deficiency resource. Genetics 177: 615–629.

Schwartz, Y. B., T. G. Kahn, D. A. Nix, X. Y. Li, R Bourgon et al., 2006 Genome-wide analysis of Polycomb targets in *Drosophila melanogaster*. Nat. Genet. 38: 700–705.

Scott, K. C., A. D. Taubman, and P. K. Geyer, 1999 Enhancer blocking by the *Drosophila* gypsy insulator depends upon insulator anatomy and enhancer strength. Genetics 153: 787–798.

Serebrovsky, A. S., and N. P. Dubinin, 1930 X-ray experiments with *Drosophila*. J. Hered. 21: 259–265.

Simon, J. A., C. A. Sutton, R. B. Lobell, R. L. Glaser, and J. T. Lis, 1985 Determinants of heat shock-induced chromosome puffing. Cell 40: 805–817.

Sipos, L., J. Mihály, F. Karch, P. Schedl, J. Gausz et al., 1998 Transvection in the *Drosophila* Abd-B domain: extensive upstream sequences are involved in anchoring distant cis-regulatory regions to the promoter. Genetics 149: 1031–1050.

Sipos, L., G. Kozma, E. Molnár, and W. Bender, 2007 In situ dissection of a Polycomb response element in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA 104: 12416–12421.

Staehling-Hampton, K., P. D. Jackson, M. J. Clark, A. H. Brand, and F. M. Hoffmann, 1994 Specificity of bone morphogenetic protein-related factors: cell fate and gene expression changes in *Drosophila* embryos induced by decapentaplegic but not 60A. Cell Growth Differ. 5: 585–593.

Stevens, M. E., and P. J. Bryant, 1985 Apparent genetic complexity generated by developmental thresholds: the apterous locus in *Drosophila melanogaster*. Genetics 110: 281–297.

Struhl, G., and K. Basler, 1993 Organizing activity of wingless protein in *Drosophila*. Cell 72: 527–540.

Tautz, D., and C. Pfeifle, 1989 A non-radioactive in situ hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene hunchback. Chromosoma 98: 81–85.

Thibault, S. T., M. A. Singer, W. Y. Miyazaki, B. Milash, N. A. Dompe et al., 2004 A complementary transposon tool kit for *Drosophila melanogaster* using P and piggyBac. Nat. Genet. 36: 283–287.

Thomason, L., D. L. Court, M. Bubunenko, N. Costantino, H. Wilson, S. Datta, and A. Oppenheim, 2007 Recombineering: Genetic engineering in bacteria using homologous recombination. Curr. Protoc. Mol. Biol. Chapter 1: Unit 1.16.

Tolhuis, B., E. de Wit, I. Muijrers, H. Teunissen, W. Talhout *et al.*, 2006   Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in *Drosophila melanogaster*. Nat. Genet. 38: 694–699.

Tsuneizumi, K., T. Nakayama, J. L. Christian, and T. Tabata, 1997   Daughters against dpp modulates dpp organizing activity in *Drosophila* wing development. Nature 389: 627–631.

Venken, K. J. T., K. L. Schulze, N. A. Haelterman, H. Pan, Y. He *et al.*, 2011   MiMIC: a highly versatile transposon insertion resource for engineering *Drosophila melanogaster* genes. Nat. Methods 8: 737–743.

Vuilleumier, R., A. Springhorn, L. Patterson, S. Koidl, M. Hammerschmidt *et al.*, 2010   Control of Dpp morphogen signalling by a secreted feedback regulator. Nat. Cell Biol. 12: 611–617.

Waddington, C. H., 1940   The genetic control of wing development in *Drosophila*. J. Genet. 41: 75–139.

Weiss, A., E. Charbonnier, E. Ellertsdóttir, A. Tsirigos, C. Wolf *et al.*, 2010   A conserved activation element in BMP signaling during *Drosophila* development. Nat. Struct. Mol. Biol. 17: 69–76.

Whittle, J. R., 1979   Replacement of posterior by anterior structures in the *Drosophila* wing caused by the mutation apterous-blot. J. Embryol. Exp. Morphol. 53: 291–303.

Wieschaus, E., and W. Gehring, 1976   Clonal analysis of primordial disc cells in the early embryo of *Drosophila melanogaster*. Dev. Biol. 50: 249–263.

Wilanowski, T. M., J. B. Gibson, and J. E. Symonds, 1995   Retrotransposon insertion induces an isozyme of sn-glycerol-3-phosphate dehydrogenase in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA 92: 12065–12069.

Williams, J. A., S. W. Paddock, and S. B. Carroll, 1993   Pattern formation in a secondary field: a hierarchy of regulatory genes subdivides the developing *Drosophila* wing disc into discrete subregions. Development 117: 571–584.

Williams, J. A., S. W. Paddock, K. Vorwerk, and S. B. Carroll, 1994   Organization of wing formation and induction of a wing-patterning gene at the dorsal/ventral compartment boundary. Nature 368: 299–305.

Wu, C. T., and J. R. Morris, 1999   Transvection and other homology effects. Curr. Opin. Genet. Dev. 9: 237–246.

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

Zecca, M., and G. Struhl, 2002b   Control of growth and patterning of the *Drosophila* wing imaginal disc by EGFR-mediated signaling. Development 129: 1369–1376.

Communicating editor: H. D. Lipshitz

PUBLICATION II

Publication II is based on a collaboration with the Laboratory of Carlos Estella (Centro De Biología Molecular Severo Ochoa, University of Madrid, Spain) and Matthew Slattery (Department of Biomedical Sciences, University of Minnesota Medical School, USA). By using a plethora of different genetic and biochemical approaches, we could show that the interplay between distinct *cis*-regulatory elements is responsible for the proper regulation of *ap* and compartment boundary formation.

Establishment of a Developmental Compartment Requires Interactions between Three Synergistic *Cis*-regulatory Modules

Dimitri Bieli[1], Oguz Kanca[1], David Requena[2], Fisun Hamaratoglu[3], Daryl Gohl[4¤],
Paul Schedl[4], Markus Affolter[1], Matthew Slattery[5], Martin Müller[1], Carlos Estella[2]*

1 Biozentrum, University of Basel, Basel, Switzerland, 2 Departamento de Biología Molecular and Centro de Biología Molecular Severo Ochoa, Universidad Autónoma de Madrid (UAM), Madrid, Spain, 3 Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland, 4 Department of Molecular Biology, Princeton University, Princeton, New Jersey, United States of America, 5 Department of Biomedical Sciences, University of Minnesota Medical School. Duluth, Minnesota, United States of America

¤ Current address: University of Minnesota Genomics Center, Minneapolis, Minnesota, United States of America

* cestella@cbm.csic.es

Abstract

The subdivision of cell populations in compartments is a key event during animal development. In *Drosophila*, the gene *apterous (ap)* divides the wing imaginal disc in dorsal *vs* ventral cell lineages and is required for wing formation. *ap* function as a dorsal selector gene has been extensively studied. However, the regulation of its expression during wing development is poorly understood. In this study, we analyzed *ap* transcriptional regulation at the endogenous locus and identified three *cis*-regulatory modules (CRMs) essential for wing development. Only when the three CRMs are combined, robust *ap* expression is obtained. In addition, we genetically and molecularly analyzed the trans-factors that regulate these CRMs. Our results propose a three-step mechanism for the cell lineage compartment expression of *ap* that includes initial activation, positive autoregulation and Trithorax-mediated maintenance through separable CRMs.

Author Summary

The separation of cell populations into distinct functional units is essential for both vertebrate and invertebrate animal development. A classical paradigm for this phenomenon is the establishment of developmental compartments during *Drosophila* wing development. These compartments depend on the restricted expression of two selector genes, *engrailed* in the posterior compartment and *apterous* (*ap*) in the dorsal compartment. Yet, despite the central role these genes and their restricted expression patterns play in *Drosophila* development, we still do not understand how these patterns are established or maintained. Here, by dissecting the regulatory sequences required for *ap* expression, we solve this problem for this critical selector gene. We used a combination of experimental approaches to identify and functionally characterize the *cis*-regulatory modules (CRMs) that regulate

ap expression during *Drosophila* wing development. For these analyses we implement a novel technique allowing us to study the function of these CRMs *in vivo*, at the native *ap* locus. We found three *ap* CRMs crucial for wing development: the Early (apE) and the D/V (apDV) enhancers and the *ap* PRE (apP). Only when all three regulatory elements are combined is a uniform and complete *ap* expression domain generated. In summary, our results indicate that *ap* is regulated in time and space by a three-step mechanism that generates a lineage compartment by integrating input from separate CRMs for the initiation, refinement and maintenance of its expression.

Introduction

Animal development requires the segregation of cell populations using both lineage and non-lineage boundaries. These cell boundaries act as signaling centers that organize the growth and patterning of specific tissues (reviewed in [1]). A paradigmatic example is the subdivision of the *Drosophila* wing disc into anterior-posterior (A/P) and dorsal-ventral (D/V) compartments. At the compartment boundaries, ligands encoded by *decapentaplegic* (*dpp*) and *wingless* (*wg*) are secreted and activate signaling pathways that orchestrate wing development [2–8]. The generation of the A/P and D/V compartments is directed by specific transcription factors, the selector genes *engrailed* (*en*) and *apterous* (*ap*), respectively, that define the identity of the cells using a binary code (on or off) [9–15]. Once the compartmental fates have been assigned, the cells in which *en* and *ap* are expressed as well as their descendants maintain that "determined" state. Unlike the A/P wing division, which is established during embryonic development, the D/V boundary is defined in the wing disc during the second larval stage by the expression of *ap* [16]. *ap* encodes a LIM-type homeodomain transcription factor and its activity depends on the formation of a complex with the LIM-domain binding protein Chip [17–19]. Since *ap* function is crucial to initiate the signaling center at the D/V boundary [20,21], *ap* null mutants completely lack the wing [16].

Due to its key role in wing disc development, *ap* function has been studied extensively. However, the transcriptional regulation of *ap* is poorly understood. How a sharp border of *ap*-expressing and non-expressing cells is generated *de novo* during the growth phase of the imaginal disc, and how the expression of *ap* is maintained and restricted to the dorsal compartment are critical unanswered aspects of wing development.

The spatial and temporal regulation of gene expression is mediated by the binding of transcription factors to discrete DNA sequences named *cis*-regulatory modules (CRMs). CRMs can be located up to hundreds of kilobases away from their target promoters. Synergistic interactions between CRMs may be required to faithfully regulate gene transcription (reviewed in [22]). Several CRMs have been identified controlling *ap* expression in different tissues, such as in muscle progenitors and in the embryonic nervous system [23,24]. A wing disc specific enhancer, named apC, has been reported to drive expression in the dorsal wing disc [24]. However, it has been demonstrated that this element is not sufficient for proper *ap* regulation in the wing [25]. *ap* expression is initially activated in future dorsal cells by the *Drosophila* Epidermal Growth Factor Receptor (EGFR) pathway through the secreted neuregulin-like signaling protein Vein (Vn) [26,27]. However, it is still unknown how *ap* expression is regulated after this initial EGFR-mediated activation. This is particularly critical in a highly proliferating tissue such as the wing imaginal disc.

The maintenance of selector gene expression domains through multiple rounds of cell divisions partially depends on the activity of the Polycomb and Trithorax group gene products

(PcG and TrxG). These proteins either repress (PcG) or activate (TrxG) the expression of their target genes through *cis*-regulatory sequences called Polycomb Response Elements (PREs) (reviewed in [28,29]). It has been suggested that *ap* expression is repressed by PcG protein complexes in ventral wing disc cells [30].

In this study, we have analyzed the regulation of *ap* at the endogenous locus and identified three *ap* CRMs crucial for wing development: the Early (apE) and the D/V (apDV) enhancers and the *ap* PRE (apP). Importantly, we analyzed these CRMs in the endogenous locus using a novel *in situ* rescue system. We find that only when the three regulatory elements are combined, a uniform and complete *ap* expression domain is observed. Our results indicate that *ap* is regulated by a three-step mechanism that generates a lineage compartment through the integration of input from separate CRMs for the initiation, refinement and maintenance of its expression.

Results

Genetic characterization of the *apterous* promoter region

ap is expressed in multiple tissues during embryonic and larval stages. Four different transcripts starting from three different promoters have been annotated which give rise to three unique polypeptides (FlyBase). We have generated a series of deletions to identify which *ap* non-coding sequences are required for *ap* expression in the wing imaginal disc (Fig 1A–1C; see Materials and Methods for information about each allele). Unless otherwise stated, hemizygous phenotypes resulting from these deletions were analyzed over *ap^{DG3}*, a large deletion removing the bulk of the *ap* locus [25], and were classified as amorphs or hypomorphs depending on their severity. *ap* amorphs were defined by the absence of wing tissue in discs and adults.

The shortest deletion in our collection with an *ap* null phenotype is *ap^{t11b}* (Fig 1C and S1D Fig). It specifically deletes the transcription start site (TSS) of transcripts *ap-RA* and *ap-RC*. Our *in silico* analysis indicates that this TSS is not controlled by a TATA-box promoter, but rather contains an Initiator (Inr) and a Downstream Promoter Element (DPE) (for review see [31]). In addition, *ap^{t11b}* removes a PRE located around 100 bp upstream of the *ap-RA/ap-RC* TSS. This PRE was defined by several chromatin immunoprecipitation studies with various anti-PcG antibodies [30,32–34]. The putative PRE core as defined by Oktaba *et al* [30] is indicated in Fig 1C. Two other small deletions with the same distal break point as *ap^{t11b}* were isolated: *ap^{c2.73c}* and *ap^{c1.2b}*. They leave the TSS, Inr and DPE intact, but remove the PRE core. In hemizygous flies, small wing stumps are often formed (Fig 1G and S1E Fig). In addition the wing pouches of 3rd instar wing discs are larger than in amorphic mutants (compare Fig 1G" with 1F"). Small amounts of Ap can only be detected in the presumptive hinge and notum (arrows in Fig 1G' and S1E Fig). The Wg stripe along the compartment boundary is absent (Fig 1G" and S1E Fig). Hence, *ap^{c2.73c}* and *ap^{c1.2b}* behave as strong hypomorphic alleles. A dramatic improvement of the adult wing phenotype is observed for deletion *ap^{c1.60c}* which is a mere 113 bp shorter than *ap^{c1.2b}* (Fig 1C and S1F Fig). Note that it keeps the TSS, Inr, DPE as well as the PRE core in place. A weak phenotype becomes apparent in hemizygous condition: similar to other weak *ap* loss-of-function alleles, most wing margins have notches. Unexpectedly, this phenotype is brought about by partial ectopic *ap* expression in the ventral pouch compartment which correlates with gaps in the Wg stripe along the compartment boundary (S1F Fig).

In summary, these observations provide strong genetic evidence for an important contribution of the *ap* PRE to wing development. In addition, a region defined by *ap^{c1.60c}* appears to act as an auxiliary module, which helps to confine the established Ap pattern to the dorsal compartment.
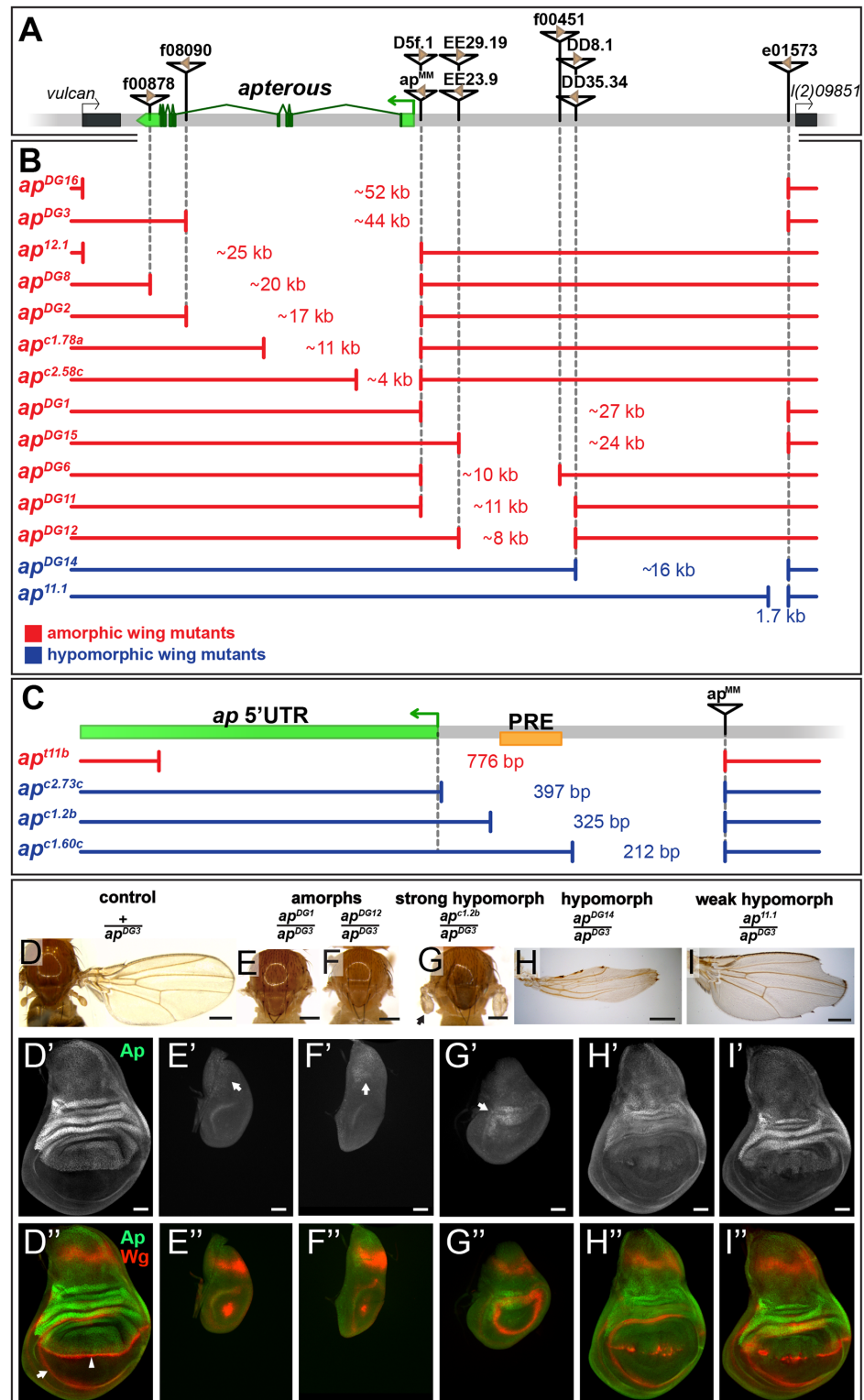
Fig 1. Deletion analysis at the *ap* locus. (A) Overview of the *ap* gene locus. *ap* transcript *ap-RA* is indicated in green and the arrow at the 5' end demarcates its TSS. The flanking genes (indicated by black boxes) are *vulcan* on the proximal and *l(2)09851* on the distal side. Relevant transposable elements used for the generation of deletions by *flp*-mediated recombination are displayed as black triangles with FRT sites within them as brown triangles. FRT orientation is indicated as defined by [82]. **(B and C)** Deletions at the *ap* locus.

The number in between the break points indicates the approximate length of the deletion. Phenotypically, the deletions can be divided into amorphic (in red) or hypomorphic (in blue) wing alleles when hemizygous over ap^{DG3}. Please note the different scales of the maps depicted in B and C. **(B)** Deletions that affect the coding sequence all lead to a no wing phenotype (ap^{DG16}, ap^{DG3}, $ap^{12.1}$, ap^{DG8}, ap^{DG2}, $ap^{c1.78a}$, and $ap^{c2.58c}$). Deletions in the upstream noncoding region between ap^{MM} and $l(2)09851$ either lead to amorphic (ap^{DG1}, ap^{DG15}, ap^{DG6}, ap^{DG11}, and ap^{DG12}) or hypomorphic wing phenotypes (ap^{DG14} and $ap^{11.1}$). **(C)** Blow up of the *ap* promoter region specific for transcripts *ap-RA* and *ap-RC*. The PRE core is depicted by a yellow box. ap^{t11b}, a deletion which removes the TSS as well as the PRE core, results in a no wing phenotype. The two deletions $ap^{c2.73c}$ and $ap^{c1.2b}$ leave the TSS intact but both remove the PRE core and both yield strong hypomorphic wing phenotypes. The weak hypomorphic allele $ap^{c1.60c}$ leaves TSS and PRE core untouched. **(D-I)** Wings and 3rd instar wing discs of representative *ap* wing mutants stained for Wg (red) and Ap (green). **(D)** Wing and notum of a hemizygous +/ap^{DG3} fly. Almost 100% of the wings look normal [25] **(D')** Ap staining in the wing disc demarcates the dorsal compartment. **(D")** Wg staining: the inner ring outlines the wing pouch (white arrow) and the stripe traversing it corresponds to the D/V compartment boundary (white arrowhead). **(E and F)** All wing tissue is lost in amophic wing mutants (ap^{DG1}/ap^{DG3} and ap^{DG12}/ap^{DG3}). **(E' and F')** Only weak Ap staining is detectable in the notum (white arrow). **(E" and F")** The wing pouch is completely lost and the inner Wg ring is reduced to a dot. **(G)** In strong hypomorphic conditions ($ap^{c1.2b}$/ap^{DG3}), only small wing and haltere stumps form (black arrow). **(G')** Low Ap protein levels are detected (white arrow) mainly in the hinge region. **(G")** The size of the wing pouch is drastically reduced and no Wg stripe along the D/V boundary is visible. **(H)** Hypomorphic mutants (ap^{DG14}/ap^{DG3}) developed considerably more wing tissue with no or little wing margin or hinge. **(H')** Compared to control discs, weaker Ap staining is observed in the pouch region. **(H")** The size of the wing pouch is comparable to wild type while the D/V Wg stripe is disrupted. **(I)** In weak hypomorphic mutants ($ap^{11.1}$/ap^{DG3}) notching of the wing blade is prominent. **(I')** Compared to control discs, *ap* expression is mostly compromised in the pouch region. **(I")** Pouch size is similar to wild type and the Wg D/V stripe is locally disrupted. All scale bars are 50 µm.

doi:10.1371/journal.pgen.1005376.g001

Deletions affecting the intergenic spacer separating *ap* and *l(2)09851* identify two regions important for *ap* function

As a next step, we generated alleles that retain an intact PRE/promoter region, but remove upstream non-coding regions of *ap*. In the mutant ap^{DG1}, 27 kb of the upstream non-coding region are deleted (Fig 1B) [35]. Hemizygous flies of this genotype can be considered as amorphic mutants, since no wing tissue was formed despite weak residual *ap* expression in the notum (Fig 1E'). Removing proximal upstream regions (ap^{DG6}, ap^{DG11}, and ap^{DG12}) also resulted in amorphic phenotypes (Fig 1B and 1F). These deletions remove the previously identified *ap* wing enhancer apC [24,25]. The distal part of the interval defined by ap^{DG1} was deleted in ap^{DG14}. Hemizygous flies form wing stumps of almost normal length but wing margin formation is severely impaired (Fig 1H). In the corresponding wing discs, *ap* expression in the wing pouch is reduced and the Wg stripe along the D/V boundary is critically perturbed (Fig 1H' and 1H"). The large size of the ap^{DG14} deficiency precludes a precise localization of an enhancer element within the ~16 kb interval. However, a few small deletions extending only proximal to e01573 allowed us to narrow down its approximate distal end: one of them, $ap^{11.1}$, deletes 1654 bp (Fig 1B; see also Materials and Methods). It can be maintained as a homozygous stock and wings look wild-type. Its weak hypomorphic nature is revealed in hemizygous $ap^{11.1}$ flies: all wings have notches along the margin (Fig 1I). Their origin can be traced to gaps in the Wg stripe along the D/V compartment boundary due to reduced Ap levels in the pouch (Fig 1I' and 1I").

Two conserved regions harbor essential wing enhancer elements

To further characterize the intergenic spacer between *ap* and *l(2)09851*, we engineered and validated a system which allowed us to investigate the role of given DNA stretches at the *ap* locus [25]. Briefly, we deleted the 27 kb (ap^{DG1}) upstream region of *ap*, and replaced it by an attP site juxtaposing the promoter/PRE ($ap^{attPΔEnh}$; Fig 2B'). In this amorphic situation, we were able to bring back sub-fragments of the previously deleted regulatory regions by ΦC31-integrase
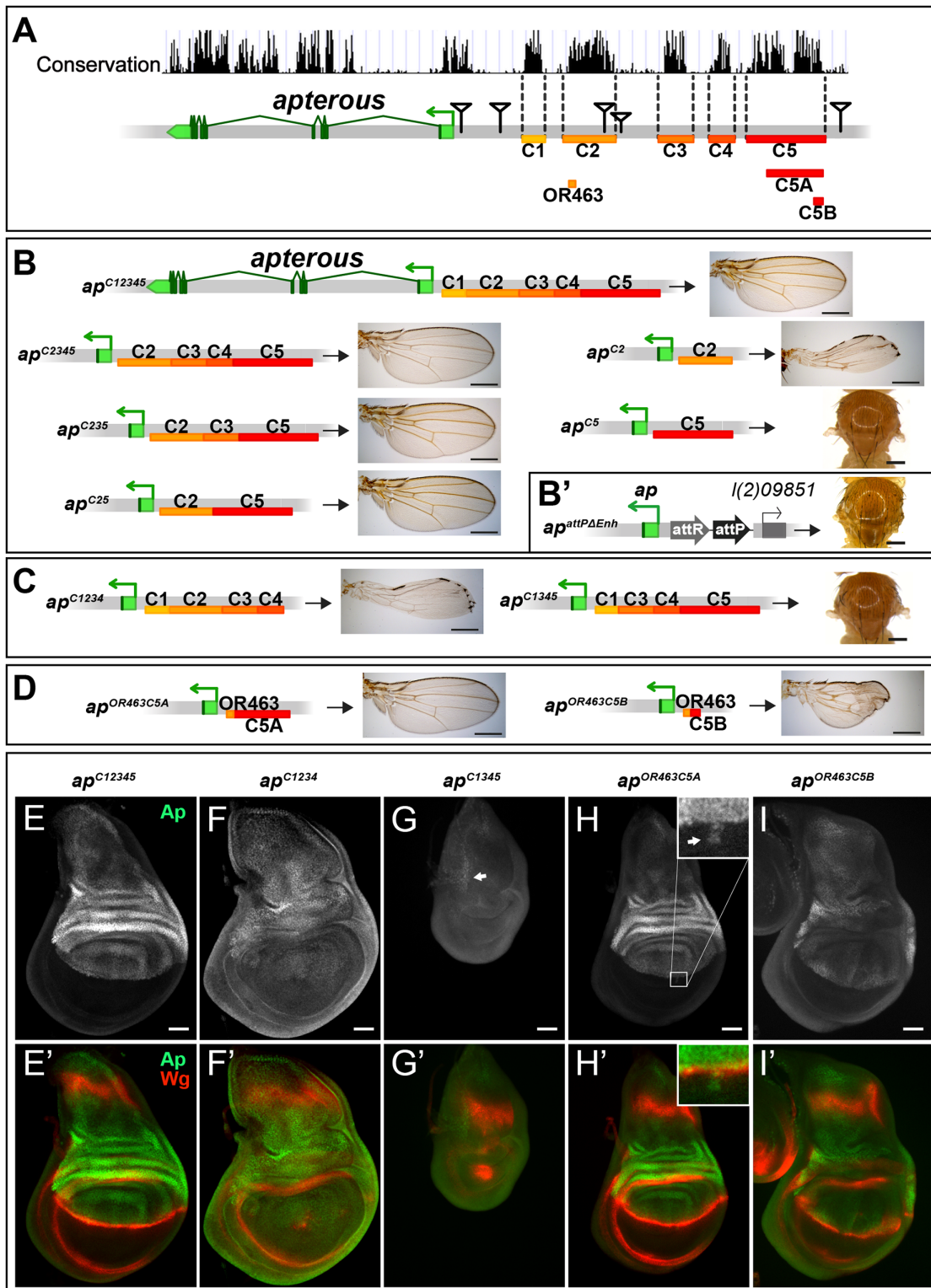
Fig 2. Analysis of the *ap* wing enhancer region. (A) Conservation of the *ap* locus (data from UCSC genome browser) and subdivision of the 27 kb intergenic region between *ap* and *l(2)09851* into 5 conserved blocks (C1–C5) is shown. OR463, C5A and C5B are subfragments of C2 and C5, respectively. Black triangles mark the locations of the transposon used for the generation of the deletions. **(B)** Six different constructs consisting of variable combinations of conserved blocks and the corresponding hemizygous wing phenotypes are depicted. When all 5 conserved regions are present (*ap^C12345^*), a normal sized and patterned wing develops. Gradual removal of C1 (*ap^C2345^*), C4 (*ap^C235^*), and C3 (*ap^C25^*) has no effect on wing morphology. Removing C5 from *ap^C25^*

results in hypomorphic wings (ap^{C2}). C5 alone (ap^{C5}) is an amophic allele, as no wings are formed. **(B')** $ap^{attPΔEnh}$: the docking site of the *in situ* rescue system for the evaluation of DNA fragments originating from the 27 kb intergenic spacer is shown. An attP site located ~400 bp upstream of the *ap* TSS juxtaposes the promoter/PRE region. As in ap^{DG1}, the 27 kb intergenic region is deleted. **(C)** Removing C2 or C5 in the context of ap^{C12345} (ap^{C1234} and ap^{C1345}) leads to the same phenotype as each element alone (ap^{C2} or ap^{C5}, respectively). **(D)** Enhancer bashing of C2 and C5 regions. OR463 and C5A in combination are the shortest fragments that still result in a normal wing ($ap^{OR463C5A}$). C5B, a sub-fragment of C5A, in combination with OR463 does not fully rescue wing formation ($ap^{OR463C5B}$). Wing size is reduced, but all margin structures are formed. **(E-I)** Third instar wing discs of different genotypes stained for Ap (green) and Wg (red). **(E-E')** ap^{C12345}: Ap and Wg pattern is indistinguishable from wild type. **(F-F')** ap^{C1234}: a significant reduction of Ap levels in the wing pouch is observed. The Wg stripe along the D/V border is almost completely lost. **(G-G')** ap^{C1345}: scattered cells with little Ap protein are detectable in the notum (see arrow). Wing pouch is reduced to a small dot of *wg* expression. **(H-H')** $ap^{OR463C5A}$: Ap and Wg patterns are similar to wild type. Ap protein can sometimes be detected in some cells of the ventral part of the disc (arrow in inset). **(I-I')** $ap^{OR463C5B}$: although protein levels are reduced, the Ap pattern is close to wild type. Nevertheless, the appearance of the Wg stripe along the D/V border is not as smooth as in wild type. All scale bars are 50 μm.

mediated insertion and test their ability to rescue wing development. Again, all the newly generated alleles were tested in hemizygous condition. According to sequence conservation and histone H3 lysine 4 trimethylation (H3K4me3) patterns, which have been reported to correlate with active promoters and enhancers [36], we have divided the upstream non-coding region of *ap* into 5 blocks (C1–5, Fig 2A). Combining all 5 conserved blocks and reintroducing them into $ap^{attPΔEnh}$ (ap^{C12345}), fully rescued wing formation (Fig 2B) as well as the Ap and Wg pattern in wing imaginal discs (compare Figs 2E with 1D). Deleting the conserved blocks that showed no H3K4me3 mark (C1 and C4), had no consequence on wing phenotype (Fig 2B). Next, we deleted conserved regions with a methylation mark. Deleting C3 had no influence on wing formation (ap^{C25}, Fig 2B). In contrast, upon removal of C5, wing development was critically disturbed (ap^{C2}, Fig 2B). Long wing stumps with defective wing margin and hinge were formed that resembled the hypomorphic ap^{DG14} mutant, which completely lacks the conserved C5 block (compare Figs 2B and 1H). In flies containing only C5, no wing tissue was formed (ap^{C5}).

We then tested whether the C2 and C5 fragments were also necessary when all the other conserved elements were present. Removing C5 only (ap^{C1234}) had the same effect as maintaining C2 alone, since long wing stumps with little margin and hinge were formed (compare Fig 2C with 2B). Wing discs of this genotype showed drastically reduced *ap* expression in the pouch region and in most cases lost the Wg stripe along the D/V boundary (Fig 2F). As expected, when only C2 was removed (ap^{C1345}, see Fig 2C), no wing tissue was formed, Ap protein was only weakly detected in the notum and the Wg pattern was equivalent to amophic *ap* wing mutants (Fig 2G).

We also investigated a possible role of the positions of C2 and C5 relative to each other in ap^{C52} and ap^{C15342} flies. Both alleles yield wild-type wings in hemizygous flies, indicating that their order on the chromosome is not important (S2A Fig).

Next, we aimed at defining the minimal CRMs which were able to direct wing development. We have recently found that shorter sub-fragments of C2 retain its wing disc specific activity [25]. The combination of a 463 bp fragment of C2 (OR463) and 3.8 kb of C5 (C5A) in $ap^{OR463C5A}$ fully rescued wing development (Fig 2D and 2H). Replacing C5A by C5B, a 600 bp subfragment of C5A, indicated that it lacks certain regulatory input (Fig 2D). The expression of *ap* in $ap^{OR463C5B}$ wing discs was restricted to the dorsal compartment, but reduced compared to ap^{C12345} (compare Fig 2I with 2E). Nevertheless, apart from small disruptions at the D/V boundary, *wg* expression appeared almost normal (Fig 2I').

Finally, to investigate whether additional wing-specific CRMs reside within the intronic sequences, we replaced the coding sequences with an *ap* cDNA lacking most intronic sequences ($ap^{cDNAint2.3}$). This allele produces normal wings (S2B and S2C Fig). Thus, we conclude that no essential wing CRMs are present in the intronic regions of *ap*. In agreement with this notion, fragments taken from intronic sequences (see below and S2D Fig) failed to drive reporter gene expression in the wing disc. Note that the cDNA used for the construction of $ap^{cDNAint2.3}$ corresponds to the *ap-RA/ap-RC* transcripts.

Combining the results from the two complementary *in vivo* approaches (deletion analysis and the *in situ* rescue system), we have defined three distinct regions which are absolutely required for the correct *ap* expression in the wing disc: a region next to the *ap* TSS which contains a PRE and two enhancers with distinct regulatory input located in homology blocks C2 and C5.

Identification of *ap cis*-regulatory modules active in the wing imaginal disc

In parallel to the *ap* deletion and *in situ* rescue strategies, we performed an unbiased search for *ap* CRMs active in the wing imaginal disc. Using the Fly Light database [37] and self-made constructs (see Materials and Methods), we screened the *ap* genomic region for DNA fragments that activate the *Gal4* gene in an *ap*-like expression pattern (Fig 3A). We found that 4 of the 17 lines tested partially recapitulated *ap*-like expression pattern in third instar wing imaginal disc (S3 Fig). Interestingly, lines 1 and 2 were active in a similar pattern in the wing pouch and hinge but were not active in the notum, while lines 7 and 8 showed identical expression pattern in the notum and hinge with low levels in the dorsal wing pouch (S3 Fig). Subsequently, we cloned the overlapping sequences between lines 1–2 and 7–8 in reporter constructs and compared their activity with *ap* expression as well as with each other during wing imaginal disc development (Fig 3B–3F; see Materials and Methods). apE (Early), the first element to be activated in early to mid-second instar imaginal discs, drove expression in all *ap*-expressing cells (Fig 3B and 3D). The other element, named apDV (Dorso-Ventral), was activated a few hours later in dorsal cells close to the D/V boundary (Fig 3C and 3E). As the wing imaginal disc developed, the activity of apE became mainly restricted to the notum and hinge with low expression remaining in the wing pouch (Fig 3B). In contrast, apDV was always restricted to dorsal wing pouch cells close to the D/V boundary, with some cells expressing the reporter in the dorsal wing hinge (Fig 3C).

In line with our previous results, apE and apDV are located within the C2 and C5 regulatory fragments identified with the *in situ* rescue system, respectively, and overlap with OR463 and C5B (Fig 3A). Moreover, in a reporter gene construct, C2 and C5 reproduced the same expression pattern described for apE and apDV, respectively (S3A Fig).

It should be noted that none of the single *ap*-CRMs identified, apDV or apE, nor the combination of them, apDV-*lacZ*+apE-*GFP*, was able to completely reproduce the endogenous *ap* expression pattern, suggesting that additional elements are necessary (Fig 3F and see below).

The EGFR pathway transiently regulates the apE element

The initial *ap* expression in the wing disc is activated by the EGFR signalling pathway at early stages of wing development (from early to mid-second instar), while its later expression is EGFR-independent [26,27]. Since the apE element was active in the entire *ap* expression domain in early wing discs, we tested whether this CRM is regulated by the EGFR pathway. Clones of cells expressing a dominant-negative form of the pathway effector Raf (*Raf*^{DN}) generated early in larval development (24–48 hrs after egg laying, AEL) were unable to activate apE (Fig 4A), while no effect was observed in clones generated later (72–96h AEL, Fig 4B). The same temporal EGFR-dependency of apE was found when the pathway was reduced in the entire wing disc using a temperature-sensitive EGFR allele (S4 Fig).

Consistent with the low levels of apE activity in dorsal wing pouch cells, misexpression of a constitutive active version of the EGFR receptor (*EGFR*$^{λtop4.2}$) by *dpp-Gal4* activated apE-*lacZ* expression in cells of the ventral pleura, while wing pouch cells were resilient to activate it (Fig 4C). To rule out a potential auto-regulatory input on the apE element, gain of function clones of the Ap activity repressor dLMO were made [38]. However, dLMO expression had no effect
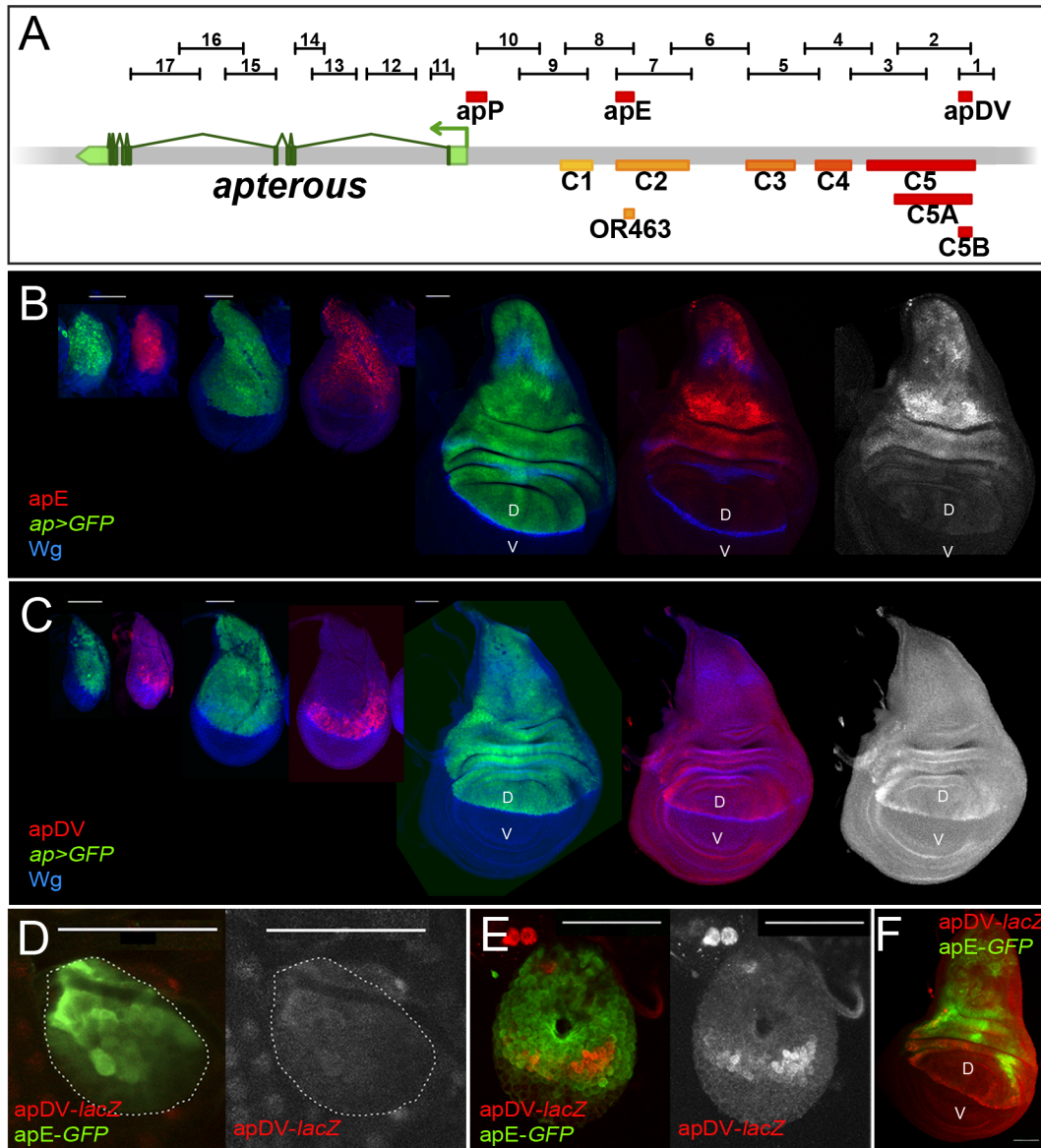
Fig 3. Activity patterns of apE and apDV enhancers. (A) Schematic representation of the *ap* genomic region is shown as a grey bar. *ap* transcript *ap-RA* is shown in green. In the upper part of the panel, horizontal bars represent the DNA elements for which Gal4 drivers were generated by the Janelia Farm consortium except for line 6 (see Materials and Methods) (http://flweb.janelia.org/cgi-bin/flew.cgi). Red bars represent regulatory elements apP, apE and apDV. At the bottom of the panel, 8 fragments tested with the *ap in situ* rescue system are indicated. **(B-C)** Pairs of wing imaginal discs isolated from second, early and late third instar larvae are shown (from left to right). They were stained for GFP (*ap-Gal4>*UAS-*GFP*, green) and Wg (blue) or apE-*lacZ* (red) in (B) or apDV-*lacZ* (red) in (C). Note that *ap>GFP* represents the complete *ap* pattern to which those of apE-*lacZ* and apDV-*lacZ* are compared. **(B)** In early discs, apE is active in all the cells that express *ap*. Later, its activity is restricted to a subset of *ap*-expressing cells mainly in the notum and hinge region. Expression in the wing pouch is very low. **(C)** apDV is active in dorsal-distal cells in early discs. Later, its activity is restricted to dorsal wing pouch cells close to the D/V boundary. **(D-F)** Early second **(D)**, mid-second **(E)** and third instar imaginal discs **(F)** stained for apDV-*lacZ* (red) and apE-*GFP* (green) are shown. **(D)** apE is activated earlier than apDV in proximal wing disc cells. **(E)** apDV is activated in dorsal-distal cells that already have apE activity. **(F)** In third instar imaginal discs, apE and apDV occupy complementary territories. apDV is restricted to dorsal wing cells close to the D/V border. apE remains mainly active in the hinge and notum. All scale bars are 50 μm. D, dorsal and V, ventral.

doi:10.1371/journal.pgen.1005376.g003

on apE activity (Fig 4D). Taken together, these results suggest that apE is activated by the EGFR pathway and that other factors regulate its expression afterwards.

To understand how the EGFR pathway regulates apE activity, we searched for putative binding sites of the ETS transcription factor Pnt [39]. Two highly conserved and one less
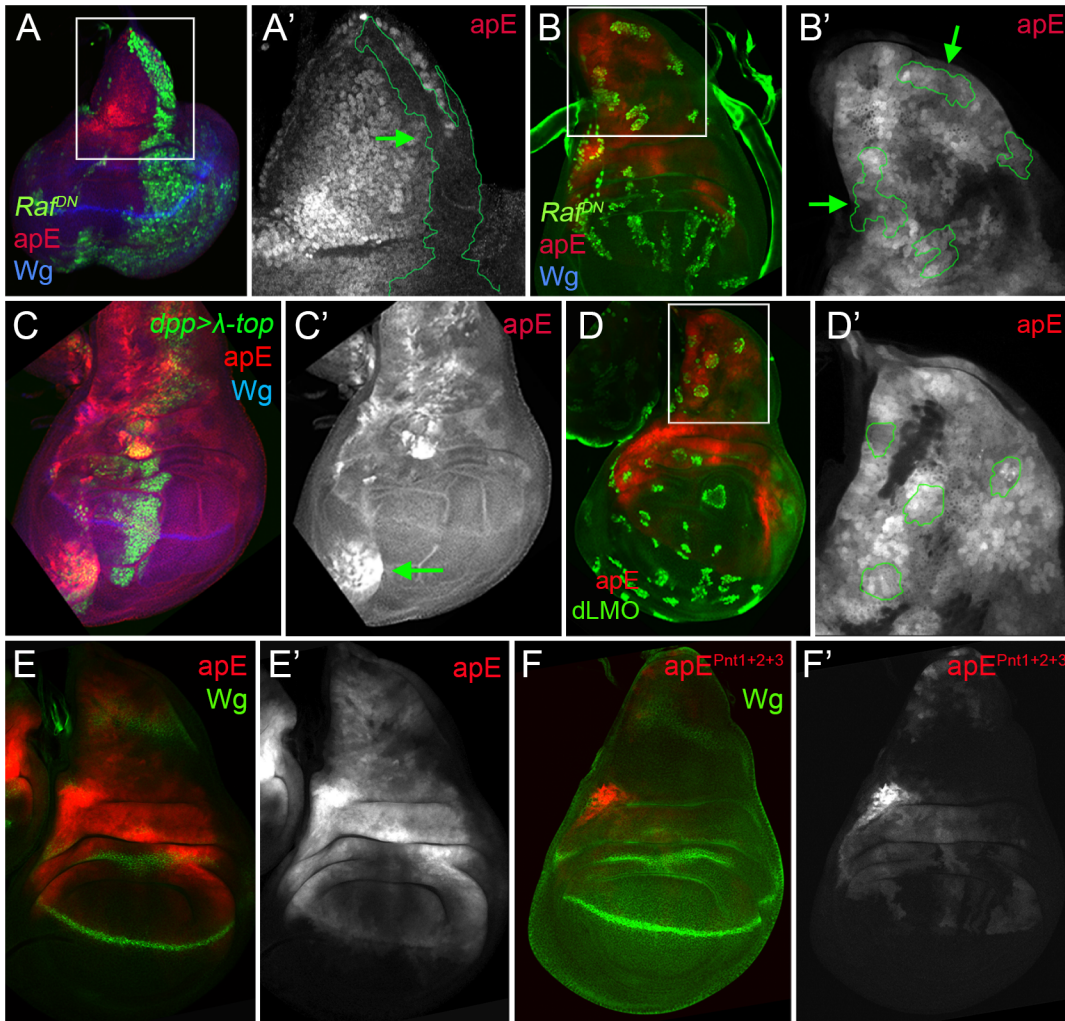
Fig 4. apE is regulated by the EGFR pathway. (A-B) Third instar wing imaginal disc with clones expressing a dominant negative version of Raf (*Raf^DN*) induced at different time points of larval development are marked by GFP (green). Discs were also stained for Wg (blue) and apE-*lacZ* activity (red). **(A)** Early induced *Raf^DN* clones (24–48hrs after egg laying, AEL) are unable to activate apE. **(Aʹ)** Close-up of the disc in **(A)** with the clone outlined in green (green arrow). Note that apE is not activated within the clone. **(B)** Late induced *Raf^DN* clones (72–96hrs AEL) have no effect on apE activity. **(Bʹ)** Close-up of the disc in **(B)** with clones outlined in green (green arrows). **(C)** *dpp-Gal4*; UAS-*GFP*, UAS-*EGFR^Atop4.2* (*λ-top*, green) wing disc stained for apE-*lacZ* (red) and Wg (blue). Ectopic activation of the EGFR pathway induces apE in ventral pleural cells. **(Cʹ)** Single channel for apE. Green arrow points to ectopic apE activity. **(D)** Gain of function clones of the Ap activity repressor dLMO (green) has no effect on apE activity (red). **(Dʹ)** Close-up of the disc in **(D)** with clones outlined in green. **(E-F)** Wing imaginal discs stained for Wg (green) and apE (**E**, red) and apE^pnt1+2+3 (**F**, red) activity. Note that apE activity is strongly reduced after mutating the three identified Pnt binding sites (apE^pnt1+2+3). All constructs have been inserted in the same genomic location. Images were obtained keeping the confocal settings constant.

doi:10.1371/journal.pgen.1005376.g004

conserved sites were identified. When all these sites were mutated simultaneously, the activity of apE was strongly reduced (compare Fig 4E with 4F). Altogether, these results suggest that apE is initially activated by the EGFR pathway and that this activation requires Pnt function.

Ap and Vg/Sd regulate the apDV CRM

While the apE element was activated in all *ap*-expressing cells in early second instar wing discs, the apDV element was induced only later and restricted to a subset of apE-positive cells (the wing pouch cells). Therefore, we tested whether the Ap protein itself is needed in an autoregulatory fashion for the restricted activity of the apDV element in the dorsal compartment.

dLMO expressing clones cell-autonomously repressed the apDV element, while forced expression of *ap* in the ventral compartment cells ectopically activated it (Fig 5A and 5B). This suggests that Ap restricts the dorsal activity of apDV. Although *ap* is expressed in all dorsal wing disc cells, apDV is only active in dorsal wing cells close to the D/V boundary, which suggests
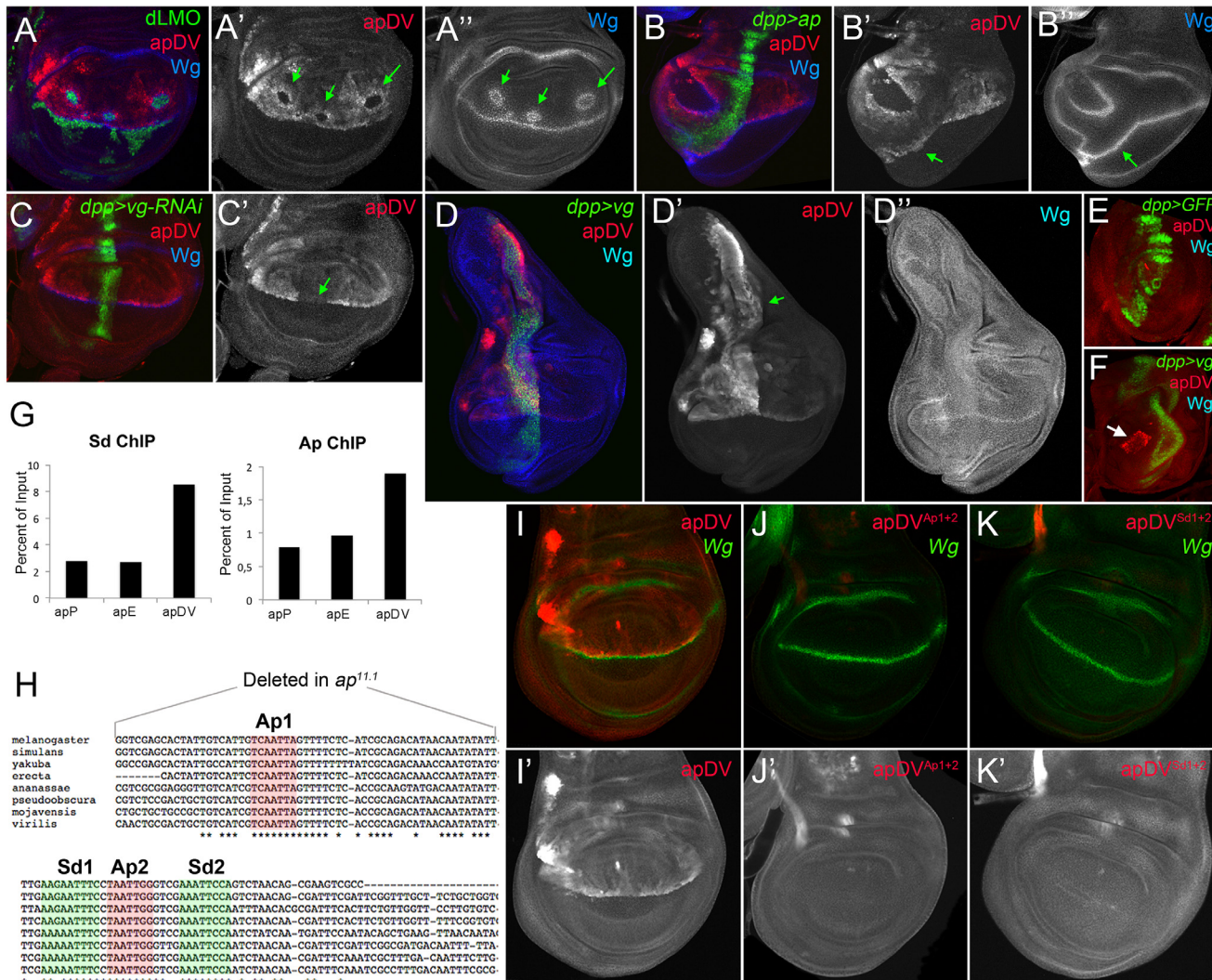
Fig 5. apDV is regulated by Ap and Sd/Vg. (A) In dLMO expressing clones (green), apDV-*lacZ* activity (red) is repressed. Wg (blue) is non-autonomously activated in cells surrounding the clones. Single channels are displayed for apDV-*lacZ* **(A')** and Wg **(A")**. Green arrows point to dLMO expressing clones. **(B)** *dpp*-Gal4; UAS-*GFP*, UAS-*ap* (green): upon ectopic Ap induction by *dpp*-Gal4, apDV-*lacZ* (red) and Wg (blue) are induced in ventral compartment cells. Single channels are displayed for apDV-*lacZ* **(B')** and Wg **(B")**. Green arrow points to ectopic apDV and *wg* expression. **(C)** *dpp*-Gal4; UAS-*GFP*, UAS-*vg*-RNAi: RNAi-induced knockdown of Vg in *dpp* domain (green). Wg is in blue. Note that apDV-*lacZ* expression (red) is strongly downregulated in the central part of the pouch. **(C')** Single channel display of apDV-*lacZ*. Green arrow points to discontinuity in the apDV pattern. **(D)** *dpp*-Gal4; UAS-*GFP*, UAS-*vg*: ectopic *vg* expression induces apDV (red) along the *dpp* domain (green), but only in the dorsal compartment. Note that *wg* is not induced upon the ectopic expression of *vg*. Single channels are displayed for apDV-*lacZ* **(D'**; green arrow points to ectopic apDV-*lacZ*) and Wg **(D")**. **(E-F)** *dpp*-Gal4; UAS-*GFP* **(E)** and *dpp*-Gal4; UAS-*GFP*, UAS-*vg* leg discs **(F)**: *dpp*-Gal4; UAS-*GFP* (green), apDV-*lacZ* (red) and Wg (blue) patterns are shown. Ectopic *vg* expression induces apDV activity in the distal domain of the leg disc (white arrow in **F**). **(G)** ChIP experiments with anti-Sd and anti-Ap antibodies. Quantifications of apP, apE and apDV DNA in immunoprecipitates demonstrate that Sd and Ap are preferentially bound to the apDV regulatory region. Representative enrichment values are shown for a single experiment that was conducted in triplicate. **(H)** DNA sequences of various *Drosophilae* species surrounding the identified Ap (red shade) and Sd (green shade) binding sites are shown. Note that Ap1 site is deleted in *ap^{11.1}* flies. **(I-K)** Wing imaginal discs stained for Wg (green) and apDV (**I**, red), apDV^{Ap1+2} (**J**, red) and apDV^{Sd1+2} (**K**, red) activity. Mutation of the Ap sites **(J)** or Sd sites **(K)** results in loss of apDV activity. **(I'-K')** Single channel pictures are depicted for each apE wild type and mutant condition. All constructs have been inserted in the same genomic location and images were obtained keeping the confocal settings constant.

additional input into this element. Therefore, we tested whether apDV activity is controlled by *wg* or *vestigial* (*vg*) [40], two key genes required for wing development. Downregulation or ectopic activation of the Wg pathway did not significantly affect apDV-*lacZ* expression (S4 Fig). However, knockdown of *vg* in the *dpp* domain eliminated apDV-*lacZ* expression (Fig 5C). Additionally, ectopic expression of *vg* strongly activated apDV in the dorsal compartment (Fig 5D). Remarkably, while apDV is not activated in the leg disc, forced expression of *vg* in this disc induced its activity in the distal domain of the leg, where a ring of endogenous *ap* expression has been described (Fig 5E and 5F)[16,41].

As a next step, we tested whether Ap and Scalloped proteins (Sd), the transcriptional companion of Vg, directly bind to the apDV CRM. Using Ap and Sd chromatin immunoprecipitation (ChIP), we found that Ap and Sd were significantly enriched at the apDV regulatory region in comparison to apE or apP (Fig 5G). Moreover, we identified two conserved consensus-binding sites for Sd as well as for Ap in the apDV region. Mutation of these sites completely eliminated apDV activity (Fig 5H–5K). Intriguingly, loss of one of these Ap binding sites likely contributes to the wing defects seen in the *ap*[11.1] mutant described previously (see Fig 1I).

Taken together, these results suggest that Ap and Vg/Sd directly regulate apDV in the wing pouch, with an Ap autoregulatory input restricting its activity to the dorsal compartment.

Synergistic effect of apE and apDV with the *ap* promoter directs *ap* expression in the wing disc

We have identified two *ap* CRMs (apE and apDV) that, when combined in a reporter assay, partially recapitulated *ap* expression in the wing disc (see Fig 3F), suggesting that other CRMs are needed for full expression. Since PRE-containing sequences are necessary for correct *ap* expression and proper wing development (Fig 1C), we tested if a region around the *ap* TSS including the PRE, named apPRE (apP), had an impact on the activity of the distal *ap* CRMs (Fig 3A). On its own, the apP drove weak expression in the wing disc in a pattern not related to the characteristic *ap* expression (Fig 6A). When placed together in a reporter construct with either apDV or apE (resulting in apDV+P-*lacZ* or apE+P-*lacZ*), the activity of the resulting reporter gene construct was the sum of both elements and did not reproduce faithfully the *ap* expression pattern (Fig 6B and 6C). Interestingly, when the three CRMs were placed together, the expression of the apDV+E+P-*lacZ* in third instar wing discs was more accurate than the expression of the previous CRMs combinations or the apDV+E-*lacZ* and more precisely reproduced the expression pattern of *ap* (compare Fig 6D with 6E).

Therefore, we tested whether these *ap* CRM combinations placed next to *ap* cDNA were sufficient to rescue wing development in an *ap* mutant background. As expected, apE+P-*apcDNA* was partially able to rescue wing growth, but completely lacks the D/V margin, whereas an apDV+P-*apcDNA* transgene, that lacks the apE enhancer, did not rescue wing formation (S5 Fig). Interestingly, the apDV+E+P-*apcDNA* transgene rescued wing formation in *ap* mutants. Although the rescue was not fully wild type, a clear wing margin was observed in wing discs and adult wings (S5 Fig).

In summary, we have identified three *ap* CRMs that, only when combined, can accurately reproduce the endogenous *ap* expression pattern in the wing imaginal disc.

Trx maintains robust *ap* expression via the apP element

It has been proposed that PcG proteins repress *ap* expression in ventral wing disc cells and that sequences around the *ap* TSS could function as a PRE [30]. However, the role of TrxG proteins in the control of *ap* expression in wing imaginal discs has not been tested previously. Therefore, we generated *trx*[E2] mutant clones and studied *ap* expression with a *lacZ*-enhancer trap inserted
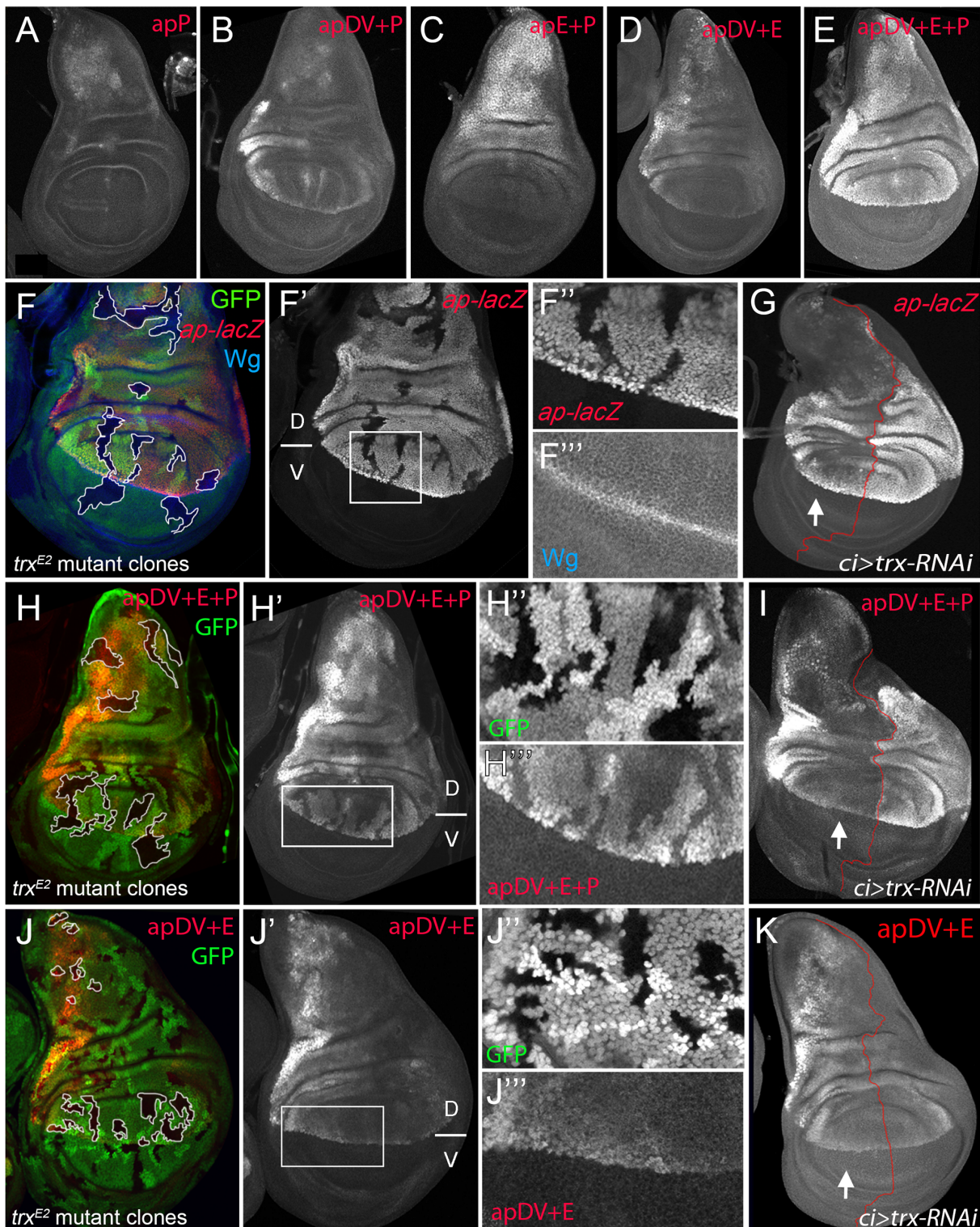
Fig 6. apP mediates *ap* expression maintenance and depends on Trx input (A-E) Third instar wing imaginal discs stained with α-βGal antibody to visualize *lacZ* activity. (A) apP activity is weak and not related to the endogenous *ap* expression pattern. **(B)** apDV+P activity is the sum of both elements. **(C)** The combination of apE+P leads to stronger and more homogeneous *lacZ* expression in the notum and hinge. Note that expression levels remain low in the dorsal wing pouch. **(D)** apDV+E activity is the sum of apDV and apE and does not reproduce the complete *ap* expression pattern. **(E)** Only the combination of apDV+E+P reproduces the endogenous *ap* pattern. All constructs were inserted in the same genomic location. **(F, H and J)** *trx*[E2] mutant

clones were generated 48–72hrs AEL and are marked by the absence of GFP (in each disc, several of them are outlined in white). Discs were stained for Wg (blue) and *ap-lacZ* (red, **F**), apDV+E+P-*lacZ* (red, **H**) and apDV+E-*lacZ* (red, **J**). (**F',H',I'**) single channel image of *lacZ* staining. (**G, I, and K**) *ci-Gal4*; UAS-*trx*-RNAi: RNAi-induced knockdown of Trx-activity in the anterior wing disc compartment. Imaginal disc were stained for β-Gal protein. White arrow points to anterior wing compartment. (**G**) *ap-lacZ*: enhancer trap *ap^rK568^*, (**I**) apDV+E+P-*lacZ* and (**K**) apDV+E-*lacZ*. (**F, F'**) *trx^E2^* mutant clones show downregulation of *ap* expression. (**F"** and **F'"**) Close-up of *ap-lacZ* and *wg* expression shown in **F'**. (**G**) Knockdown of Trx in the anterior compartment downregulates *ap-lacZ* expression. Note that the reduction of *ap-lacZ* is stronger in the notum and the wing pouch close to the hinge. (**H and H'**) *trx^E2^* mutant clones show downregulation of apDV+E+P-*lacZ* expression. (**H"** and **H'"**) Close-up of GFP and apDV+E+P activity in **H'**. (**I**) Knockdown of Trx in the anterior compartment (arrow) downregulates apDV+E+P-*lacZ* expression. As *ap-LacZ* in (**G**), apDV+E+P activity is reduced in a spatial dependent manner. (**J and J'**) *trx^E2^* mutant clones show no effect on apDV+E-*lacZ* expression. (**J"** and **J'"**) Close-up of GFP and apDV+E activity in **J'**. (**K**) Reducing Trx in the anterior compartment has no effect on apDV+E-*lacZ* expression. D, dorsal and V, ventral.

immediately 5' to the *ap* TSS (*ap^rK568^*). We found that cells devoid of *trx* function show reduced *ap-lacZ* expression ([Fig 6F](#)). To analyze this result in more detail, we reduced *trx* mRNA levels in the anterior wing disc compartment (*ci-Gal4>trx-RNAi*) and compared the levels of *ap-lacZ* expression with the posterior control compartment ([Fig 6G](#)). Consistent with *trx* mutant clones, *ap-lacZ* expression was strongly reduced in the anterior compartment, although the reduction was more prominent in the notum and in the dorsal wing pouch close to the hinge.

To genetically confirm that the *ap* PRE (apP) functions as a Trithorax response element (TRE), we eliminated or downregulated Trx activity and analyzed the expression of the apDV +E+P-*lacZ* reporter construct. In *trx^E2^* mutant clones, apDV+E+P-*lacZ* levels were strongly reduced, as it was the case for *ap-lacZ* ([Fig 6H](#)). In contrast, the same construct without the *ap* promoter, apDV+E-*lacZ*, was not altered in these *trx* mutant clones ([Fig 6J](#)). Accordingly, reducing the levels of Trx in the anterior compartment cells (*ci>trx-RNAi*) did not affect expression of apDV+E-*lacZ* (compare [Fig 6K with 6I](#)). Interestingly, the expression pattern of apDV+E+P-*lacZ* was strongly reduced upon Trx downregulation and resembled the pattern of wild type apDV+E-*lacZ* (the same construct without the apP, compare [Fig 6G and 6I with 6K](#)).

Altogether, our results suggest that the *ap* promoter region behaves as a PRE/TRE providing the information required to maintain *ap* expression.

Direct and continuous contact of the apDV and apE CRMs together with the apP element for *ap* maintenance

Classical transvection experiments usually deal with chromosomes harboring genes lacking either a functional promoter region or a functional enhancer. For combinations of members of the two groups, intragenic complementation can be observed, i.e. the corresponding phenotype is much less severe than seen in allelic combinations involving only one or the other group [42,43]; reviewed in [44]. We have previously reported that transvection is at work at *ap* [35]. For example, *ap^DG12^/ap^DG1^* flies have no wings because both alleles delete wing enhancer apE. The same phenotype is observed in *ap^t11b^/ap^DG8^* flies because both alleles remove the promoter region as well as the 5' end of *ap*. In contrast, the wing phenotype of *ap^t11b^/ap^DG1^* flies is much improved ([S7B Fig](#)). Models for transvection posit that the apE and apDV enhancers on chromosome *ap^t11b^* can activate the transcription machinery of the functional *ap* gene on chromosome *ap^DG1^*. However, the *ap^t11b^/ap^DG1^* wings are consistently less well formed than those obtained from *ap^DG3^/+* flies (see [S7 Fig](#)). These observations suggest that the apP region on the one hand and the two enhancers on the other interact more efficiently if they are located *in cis*.

In our study, we have shown that the two *ap* wing enhancers are clearly separable units: (1) they lie ~10 kb apart and (2) the activity of apE is essential for auto-regulatory activation of apDV. From these premises, one would not *a priori* expect that the two enhancers must be *in cis* for full function. However, several allelic combinations containing only one or the other enhancer element (apE or apDV) generated discs and adult wings with defects at the D/V boundary: similar results were obtained for genotypes *ap^C1234^/ap^C1345^*, *ap^DG14^/ap^DG12^*, *ap^C2^/*

ap^{DG12} or when a $su(Hw)$ insulator element was inserted between apDV and apE in ap^{f00451}/ap^{DG3} animals (Fig 7A and S7 Fig). Our transvection studies suggest that all three CRMs need to be in *cis* to fully rescue wing development.

To better understand how the synergy between the three regulatory elements is achieved, we used chromosome conformation capture (3C) [45], which allowed us to test *in vivo* whether there is direct physical contact between the apE and apDV CRMs with the apP. Indeed, as seen in Fig 7B, we found that in whole third instar larvae, apP preferentially contacted the apE and apDV elements, and did so more frequently when compared to sequences outside the *ap* genomic locus. This suggests that the distal apE and apDV regions are in close physical proximity to apP *in vivo*.

Next, we tested whether apDV and apE CRMs are required either continuously or only transiently to direct *ap* expression during wing disc development. To distinguish between these two possibilities, we generated an apDV+E+P-*lacZ* construct, in which the apDV+E is flanked by FRT sequences (FRT-apDV+E-FRT+P-*lacZ*, Fig 7C). This allowed us to remove the apDV+E cassette at different time points of wing development using Flp-mediated recombination [46,47] (see Materials and Methods). Deletion of apDV+E early in development in the posterior compartment completely abolished reporter expression compared to anterior control cells (compare Fig 7D and 7E). Deletion of the apDV+E at later stages also strongly decreased reporter gene expression (Fig 7F).

In summary, these experiments suggest that there is a direct contact between the apE and apDV with the *ap* promoter and that these three elements need to be in *cis* throughout wing disc development to confer optimal *ap* expression.

Discussion

The selector gene *ap* encodes for a transcription factor that confers dorsal identity in the wing imaginal disc. A precise border of *ap*-expressing and non-expressing cells is absolutely necessary for wing growth and pattern formation. Although the role of *ap* as a dorsal selector gene has been extensively studied, how its specific spatial expression pattern is brought about during wing development has remained unclear. In this work, we have used complementary strategies to identify and molecularly characterize the endogenous CRMs that regulate *ap* expression during wing development.

ap cis-regulatory logic for Dorso-Ventral identity in the wing imaginal disc

Our genetic and *cis*-regulatory analysis provides information about the logic of *ap* expression during wing development. We propose that *ap* expression is controlled by at least three CRMs that act in combination (Fig 7G). The first element, apE is the earliest to be activated in proximal wing disc cells via the EGFR pathway; its expression subsequently weakens in the wing pouch. Deletion of this early enhancer (e.g. ap^{DG12} or ap^{C1345}) completely abolishes wing formation. The asymmetry of *ap* expression to the proximal domain of the wing disc is probably due to the localized activation of the EGFR pathway by its ligand Vn and a distal repression by Wg signaling [26,48–50]. We have genetically and molecularly confirmed the initial activation of the apE by the EGFR pathway; however, other inputs are required for the continuous activation of this CRM in later wing discs.

A few hours after apE activation, a second CRM, apDV, is activated in a subset of apE positive cells. In contrast to apE, apDV is restricted to the dorsal-distal domain of the wing pouch by direct positive inputs from Ap and Vg/Sd (Fig 7G). The direct Ap autoregulatory input defines the time window when the apDV element is activated; apDV can only be active after the induction of Ap by the early enhancer (apE). It has been shown that Ap induces *vg*
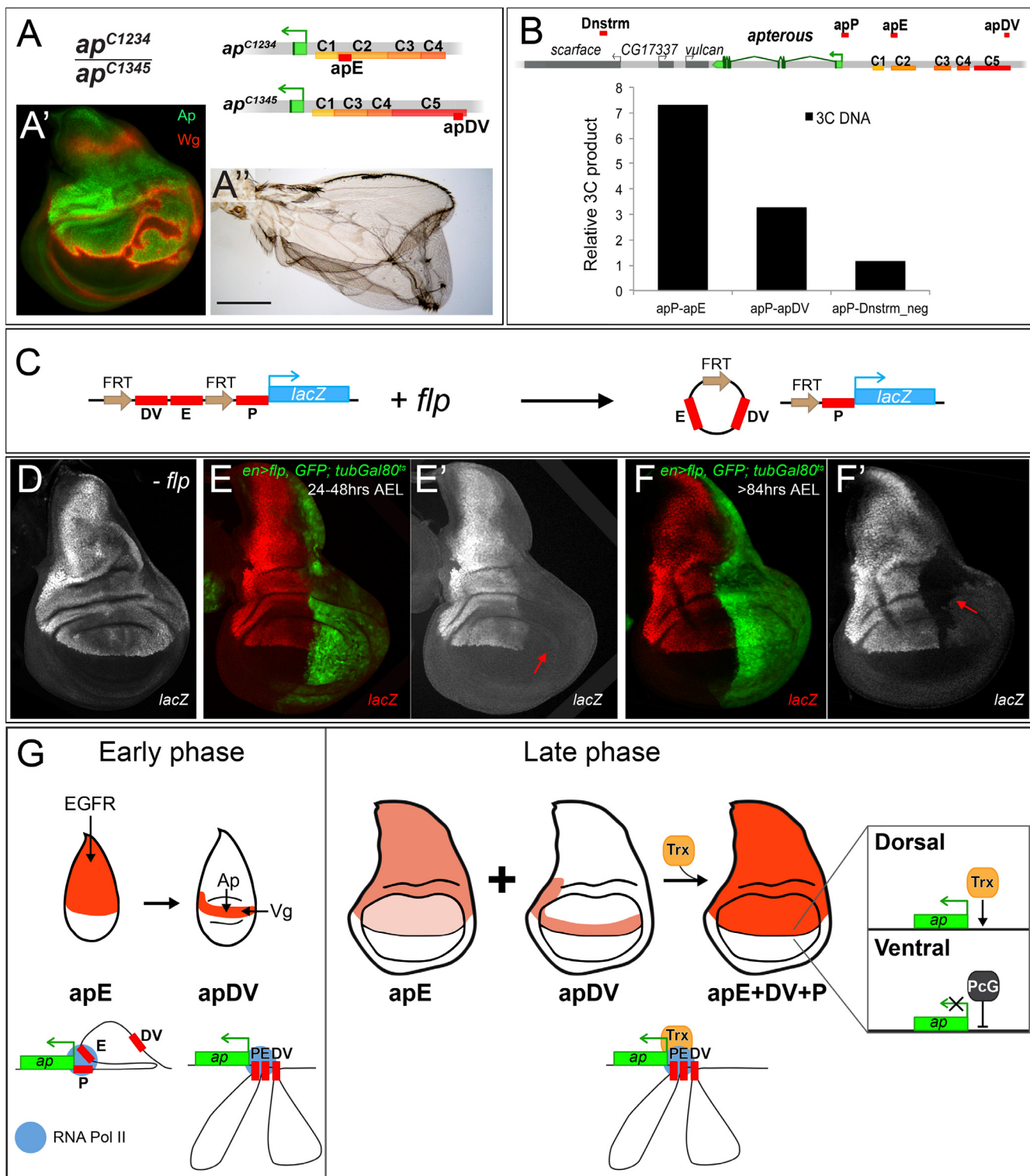
Fig 7. Evidence for genetic and physical interaction between apDV, apE and apP. (A) At the top of the panel, the genetic constitution of ap^{C1234}/ap^{C1345} flies is shown. Note that apE and apDV are present in *trans* and that apP (not indicated) is present on both chromosomes. (A') Ap (green) in the wing disc is uneven leading to derepression of *wg* in cells with no Ap (red). (A") Wings of ap^{C1234}/ap^{C1345} flies frequently show wing patterning defects and outgrowths. (B) At the top of the panel, a schematic representation of the *ap* genomic locus and several flanking genes is shown. C1–C5 indicates the conserved homology regions. Red bars above the chromosome represent the different regions tested for direct interaction with apP using the 3C technique. A region

downstream of *ap*, named Dnstrm, was used as a negative control. The diagram in the lower part of the panel summarizes the 3C data. In whole third instar larvae, apE and apDV elements are more frequently close to apP than a control DNA element located downstream (Dnstrm) the *ap* genomic locus. **(C)** Diagram of the FRT-apDV+E-FRT+P-*lacZ* reporter gene is depicted. Upon *flp* induction, the apDV+E cassette is deleted. The *lacZ* reporter remains under the control of apP only. **(D-F)** Expression of the FRT-apDV+E-FRT+P-*lacZ* reporter gene in third instar wing imaginal discs in the absence of *flp* **(D)** or after *flp*-induction at different times of larval development **(E-F)**. Controlled *flp*-induction in the posterior pouch compartment (red arrow) was achieved and monitored in an *en-Gal4*; UAS-*flp*; UAS-*GFP*, tubGal80*ts* background. Wing discs were stained for *lacZ* (red) and GFP (green). **(D)** *lacZ* pattern without *flp*-induction resembles wild type *ap* expression. **(E)** *flp*-induction 24–48hrs AEL: deletion of the apDV+E cassette results in loss of *lacZ* expression in the posterior compartment. **(F)** *flp*-induction >84hrs AEL: Loss of *lacZ* expression in the posterior compartment after late *flp*-induction demonstrates the continuing requirement of apE and apDV. **(G)** *ap cis*-regulatory model for the establishment of Dorso-Ventral identity in the wing imaginal disc. During the early phase of *ap* activation, the EGFR pathway triggers *ap* expression via the apE CRM that directly interacts with apP. A few hours later, apDV is activated in dorsal cells close to the D/V boundary. It is activated by Vg/Sd in the future wing cells but its activity is restricted dorsally by Ap itself. apDV is also recruited to apP. In the late phase, apP maintains *ap* expression through Trx input in dorsal cells. Persistent *ap* transcription is required for the generation of a dorsal lineage compartment. It is dependent on the permanent presence of the apE and apDV enhancers and their continuing interaction with apP. PcG proteins repress *ap* in ventral compartment cells.

expression by triggering Notch signaling at the D/V boundary [20,21,48,51,52]. Thus, the (direct) input of Vg/Sd on apDV can be regarded as an indirect positive autoregulation, which delimits the spatial domain where apDV can be activated. Consequently, the interface of Ap and Vg expression defines the region of apDV activity via positive autoregulation.

The third *ap* CRM is the *ap* PRE/TRE region (apP), that, when deleted, leads to a strong hypomorphic wing phenotype (*ap*$^{c1.2b}$). The apP requires Trx input and maintains *ap* expression when placed in *cis* with the apDV and apE CRMs (Fig 7G). Only the combination of the three CRMs faithfully reproduces *ap* expression in the wing disc. Moreover, our regulatory *in locus* deletion and *in situ* rescue analysis provide strong functional relevance for these CRMs.

Ultimately, this cascade of *ap* CRMs provides a mechanism to initiate, refine and maintain *ap* expression during wing imaginal disc development, in which the later CRMs depend on the activity of the early ones (Fig 7G). A similar mechanism has been described for *Distal-less* (*Dll*) regulation in the leg primordia where separate CRMs trigger and maintain *Dll* expression in part by an autoregulatory mechanism [46,53].

It has been proposed that positive autoregulation may help to maintain the epigenetic memory of differentiation [54]. In the case of *ap*, we demonstrate that autoregulation works in conjunction with a PRE/TRE system; this might make the system very robust and refractory to perturbations.

The role of *ap* promoter in maintenance

ChIP experiments have shown that many developmentally important genes are associated with a promoter proximal PRE as found at *ap* [30]. The role of such a PRE has been studied at the *engrailed* (*en*) locus. It has been demonstrated that in imaginal discs, the promoter as well as the promoter proximal PRE are important for the long-range action of *en* enhancers [55,56]. These authors propose that this PRE brings chromatin together, allowing both positive and negative regulatory interactions between distantly located DNA fragments.

Our results indicate that sequences around the transcription start of *ap* (apP) may serve a similar function. First, this element, when placed in *cis* with the *ap* CRMs (apE and apDV), maintains the *ap* expression pattern and keeps reporter gene expression off in cells where low or no activity of apDV and apE has been observed. Second, in the absence of *trx*, the expression of *ap* and apDV+E+P-*lacZ* is strongly reduced. All these data suggest that sequences within the apP integrate Trx input, thereby maintaining *ap* expression in a highly proliferative tissue such as the wing disc. Interestingly, *trx* mutant clones were not round and did not show ectopic *wg* activation (Fig 6F), which is a hallmark of *ap* loss-of-function clones. This suggests that in *trx* mutant clones enough Ap protein is still present to maintain *wg* expression off. However, we

found derepression of the ventral-specific integrin *αPS2* in *trx* mutant clones in the wing pouch as previously described for *ap* mutant clones [14] (S6 Fig).

It has been suggested that TrxG proteins could act passively antagonizing PcG silencing, rather than playing an active role as co-activators of gene transcription [57,58]. For example, *Ubx* expression in the leg and haltere does not require Trx in the absence of Polycomb repression [59]. We tested these possibilities and generated *trx* mutant clones that were also mutant for the PcG member *Sex combs on midlegs* (*Scm*). Dorsally-located *Scm⁻ trx⁻* double mutant clones still downregulate *ap-lacZ* expression while ventral-induced ones are unable to derepress *ap-lacZ* as we observed for *Scm⁻* single mutant clones (S6 Fig). Therefore, our results, in addition with previous findings by Oktaba *et al* [30], suggest that TrxG maintains *ap* expression in dorsal cells, while *ap* expression is repressed in the ventral compartment by PcG proteins. Moreover, it has been shown that the sequences around the *ap* transcription start, including the PRE, are occupied by PcG complexes PRC1 and PRC2, as well as Trx [30,60].

Direct and continuous interactions between the apE and apDV with the apP

Enhancers-promoter interactions initiate transcription but their dynamics during development have remained poorly understood. Our chromosome conformation capture (3C) experiment provides evidence for the direct interaction between the *ap* CRMs apE and apDV with the maintenance element encoded by the apP. Beyond this, we also find that these elements cooperate continuously during wing development. Our flip-out experiments, in which we removed the apDV and apE CRMs at different time points, suggest that these elements need to be present continuously to ensure correct *ap* expression. Additionally, flies carrying apE only on one chromosome and apDV only on the homologue were unable to fully rescue wing development suggesting that these CRMs need to be in *cis*. It is conceivable that *in cis* configuration of the three *ap* CRMs facilitates and stabilizes enhancer-promoter looping. It could also help to rapidly establish relevant chromatin contacts after each cell division. These results are in accordance with previous observations, in which constant interactions between *ap* enhancers and promoter during embryogenesis have been described [61]. Our results extend these observations to the wing disc, a highly proliferative tissue, where the expression of the trans-factors that regulate the activity of the apE and apDV is very dynamic. This raises the question on how this contact is re-assembled over many cell generations. It is possible that some epigenetic modifications are laid down in the activated apE and apDV CRMs, which are then inherited during cell divisions to ensure contact with apP. Studies of the chromatin status of these elements will be required to fully understand this process.

Developmental transcriptional regulation during tissue growth

A key question in developmental biology is how transcriptional regulation is coupled to tissue growth to precisely regulate gene expression in a spatio-temporal manner. For example, during *Drosophila* leg development, initial activation of the ventral appendage gene *Dll* by high levels of Wg and Dpp initiates a cascade of cross-regulation between Dll and Dachshund (Dac) and positive feedback loops that patterns the proximo-distal axis [46,62]. Other mechanisms to expand gene expression patterns depend on memory modules such as PREs, as it is the case for the Hox genes or other developmental genes like *hh* [63–65]. To direct wing formation, expression of *ap* in the highly proliferative tissue of the wing disc must be precisely induced to generate and maintain the D/V border. Our in-depth analyses at the *ap* locus provide a functional and molecular explanation of how expression of this dorsal selector gene is initiated, refined at the D/V border, and maintained during wing disc development. We propose that this three-

step mechanism may be common for developmental patterning genes to make the developmental program robust to perturbations.

Materials and Methods

Stocks used in this study

Flies were grown on standard cornmeal agar. *ap-lacZ* (*P{PZ}ap^{rK568}*), *ap-Gal4*, *ap^{UGO35}*, *trx^{E2}*, *Scm^{D1}*, *trx^{E2} Scm^{D1}* (gift from Jürg Müller)[59], *EGFR^{tsa}*, UAS-*EGFR$^{λtoP4.2}$*, UAS-*Raf^{Dn}* UAS-*arm^{S10}*, UAS-*TCF^{DN}*, UAS-*vg*, UAS-*dLMO*, UAS-*ap*, *dpp*-Gal4; UAS-*GFP*, *ci*-Gal4; UAS-*GFP*, *tubGal80^{ts}*. *act5C>stop>lacZ*; UAS-*flp*, *P{hsFLP}12*, y^1 w*, *TM3*, ry^{RK} Sb^1 Ser^1 *P{Δ2–3}99B*, *P{EPgy2}l(2)09851^{EY06365}*, *al^1 b^1 c^1 sp^1*, y^1 w^{67c23}; *noc^{Sco}* / CyO, *P{Crew}DH1*, y^1 w*; *Mi{y[+-mDint2] = MIC}MI00964*, y^1 w*; *Mi{y[+mDint2] = MIC}MI02330/SM6a* as well as all the Janelia Farm Gal4 drivers were obtained from the Bloomington Drosophila Stock Center except as indicated. These are described in the Fly light data base (http://flweb.janelia.org/cgi-bin/flew.cgi): 1-GMR_39E04, 2-GMR_42A06, 3-GMR_42D11, 4-GMR_41B09, 5- GMR_41E03, 7-GMR_42B11, 8-GMR_41D11, 9-GMR_41D03, 10-GMR_40H04, 11-GMR_39B07, 12-GMR_40A08, 13-GMR_39G10, 14-GMR_ 39C09, 15-GMR_40A07, 16-GMR_41A02, 17-GMR_41C10. For the lineage analyses of the Janelia lines we used the *act5C>stop>lacZ*; UAS-*flp* [47]. UAS-*vg*-RNAi, UAS-*sd*-RNAi and UAS-*trx*-RNAi are available at the Vienna Drosophila Resource Center (VDRC). RNAi knock-down experiments were performed in a UAS-*Dcr–2* background. *en-Gal4; UAS-flp, UAS-GFP* was a gift from Laura Johnston. *PBac{RB}e01573*, *ap^{f08090}* (*PBac{WH}f08090*),*ap^{f00878}* (*PBac{WH}f08090*) and *ap^{f00451}* (*PBac{WH}f00451*) were purchased from the Exelixis stock collection at Harvard Medical School. *y w M {vas-int.Dm}zh-2A*, a stock producing *ΦC31*-integrase under the control of the vasa promoter, and docking site *M{3xP3-RFP.attP}zh-86Fb* were obtained from Johannes Bischof [66]. The *GFP* knock-in allele *ap^{GFP}* is described in Caussinus *et al*, 2012 [67]. *ap^{MM}* and *ap^{MM-Mcp}* have been described previously [35]. They contain a P-element insertion ~400 bp upstream of the *ap* TSS. The *ΦC31*-integrase platforms *ap^{attPΔCDS}* and *ap^{attPΔEnh}* used for the *in situ* rescue system are described in detail in Caussinus *et al*, 2012 and Bieli *et al*, 2015, respectively [25,67]. The generation of all deficiencies shown in Fig 1A and 1B is described below.

Adult wings were dissected and mounted in Hoyer's and baked at 58°Celsius for a few hours. Pictures were taken with a Nikon Microphot-FXA microscope with a Sony NEX-5RK digital camera.

The notums of adult flies were photographed with a Leica M125 binocular equipped with a Leica DFC420C camera.

Generation of deletions

Df(2R)ap^{DG1} is described in Gohl *et al*, 2008, where it is called *ap^{DG}* [35].

Df(2R)ap^{DG3}, *Df(2R)ap^{DG8}* and *Df(2R)ap^{DG11}* are described in Bieli *et al*, 2015 [25].

Df(2R)ap$^{12.1}$, al b was obtained in an attempt to isolate male-recombination events to the right of P-element insertion *ap^{MM-Mcp}*. Molecular characterization identified the proximal breakpoint in a P-element insertion hot spot at the 5' end of the *vulcan* gene (Genome release R6 FB2015_01: 2R:5702133). It also verified the integrity of ap^{MM-Mcp} at its original insertion site. This deletion is referred to as *Df(2R)ap$^{12.1-Mcp}$, al b*. In order to delete the *Mcp* element located between 2 loxP sites on *ap^{MM-Mcp}*, *Df(2R)ap$^{12.1-Mcp}$, al b* was treated with *Cre* recombinase [68] and *Df(2R)ap$^{12.1}$, al b* was obtained. Homozygous flies of this genotype make it to the pharate adult stage. Dissected individuals have neither wings nor halteres. In this study, *Df(2R)ap$^{12.1}$, al b* is referred to as *ap$^{12.1}$*. Note that it is associated with a FRT site left within *ap^{MM}*.

The following 6 deletions were created by *flp*-mediated recombination [69] between 2 FRT sites located *in trans* to each other in 2 different transposons (below, their names are indicated in parenthesis; their positions within the *ap* locus is depicted in Fig 1A):

Df(2R)ap^{DG16}, al b (*ap^{12.1}*; e01573). Referred to in the text as *ap^{DG16}*. This chromosome is deficient for *vulcan* and *ap*. Homozygous *ap^{DG16}* flies are pharate adult lethal. Dissected individuals have neither wings nor halteres.

Df(2R)ap^{DG2} (f08090; ap^{MM-Mcp}). Referred to in the text as *ap^{DG2}*. Note that *Mcp* is lost upon flp-mediated recombination and that this deletion is associated with an array of Su(Hw) binding sites originating from f08090.

Df(2R)ap^{DG15} (*ap^{EE23.9}*; e01573). Referred to in the text as *ap^{DG15}*. *ap^{EE23.9}* (as well as *ap^{EE29.19}*) is a ΦC31-integrase mediated insertion of a plasmid containing *mini-white*, *FRT* and *mini-yellow* in docking site *MI00964* [70].

Df(2R)ap^{DG6}, al (ap^{D5f.1}; f00451). Referred to in the text as *ap^{DG6}*. *ap^{D5f.1}* is a ΦC31-integrase mediated insertion of a plasmid containing *mini-white*, *FRT* and *mini-yellow* in docking site *ap^{c1.4b}* [25]. Note that this deletion is associated with an array of Su(Hw) binding sites originating from f00451. *ap^{DG6}* flies have neither wings nor halteres. These phenotypes are not modified in a *su(Hw)^-* background.

Df(2R)ap^{DG12} (*ap^{EE29.19}*; *ap^{DD8.1}*). Referred to in the text as *ap^{DG12}*. *ap^{DD8.1}* (as well as *ap^{DD35.34}*) is a ΦC31-integrase mediated insertion of a plasmid containing *mini-white*, *FRT* and *mini-yellow* in docking site *MI02330* [70].

Df(2R)ap^{DG14} (*ap^{DD35.34}*; e01573). Referred to in the text as *ap^{DG14}*.

Df(2R)ap^{c2.73c}: this short deletion was obtained by direct gene conversion [71,72]. A detailed account on our experimental approach is given in Bieli *et al*, 2015 [25]. *ap^{c2.73c}* was obtained according to the exact same procedure as *ap^{c1.4b}*, except that the left homology arm on the gene conversion template plasmid was only 502 bp long, leading to a 397 bp deletion just proximal to *ap^{MM}*. Our gene conversion approach also introduced a cassette consisting of a GFP reporter driven by a minimal *hsp70* promoter flanked by two inverted attP sites for Recombination Mediated Cassette Exchange (RMCE) [73].

The following five deletions were obtained by imprecise excision of insert *ap^{MM}* during the generation of gene conversion events *ap^{c1.4b}* and *ap^{c2.73c}*. In all five cases, the deletion extends only to the left of *ap^{MM}*.

Df(2R)ap^{c1.78a}: 12 bp are left between the break points, 8 of them can be identified as belonging to the end of the P-element 3' foot. Referred to in the text as *ap^{c1.78a}*.

Df(2R)ap^{c2.58c}: the most 3' ~1.6 kb of *ap^{MM}* are left at the break point, including the wing enhancer of the *yellow* gene. Referred to in the text as *ap^{c2.58c}*.

Df(2R)ap^{t11b}: the terminal 17 bp of the P-element 3'-foot are left between the breakpoints. Referred to in the text as *ap^{t11b}*.

Df(2R)ap^{c1.2b}: the intact *ap^{MM}* insert is left at the break point. Referred to in the text as *ap^{c1.2b}*.

Df(2R)ap^{c1.60c}, sp: the intact *ap^{MM}* insert is left at the break point. This small deletion can be maintained as a homozygous stock and most wings look wild-type. Referred to in the text as *ap^{c1.60c}*.

Finally, *Df(2R)ap^{11.1}, c sp* was obtained by transposase treatment of *EY06365*, a P-element inserted in the 5' end of *l(2)09851*, the gene immediately distal to *ap* (see Fig 1A). In an attempt to isolate deletions extending proximal to *EY06365*, dysgenic males of the genotype *y w*; *al ap^{DG3}{w^+}* + + + / + + *EY06365{y^+ w^+} c sp*; TM3, Sb, Δ2–3 / + were crossed with *y w*; *al b c sp* / SM6, *al sp* females. Progeny was screened for candidates with no eye colour, *y^+* body colour and carrying the c and sp markers. 2 of the candidate chromosomes (isolation numbers 11.1 and 34.1) gave rise to notched wings *in trans* to *ap^{DG3}*, a phenotype reminiscent of weak *ap*

alleles. SM6 balanced stocks were established. Homozygous flies readily hatch and show no or only very weak wing phenotypes. Molecular characterization of *EY06365* and the 2 candidates detected in all three a ~400 bp LTR of the *springer* retrotransposon at position 2R:5751931 (Genome release R6 FB2015_01). *EY06365/ap^{DG3}* flies have normal wings indicating that the LTR doesn't have phenotypic consequences. Furthermore, remarkably similar rearrangements could be detected in candidates 11.1 and 34.1: *EY06365* has relocated into exactly the same site in the hybrid piggyBac present on *ap^{DG3}* (obtained by *flp*-mediated recombination between *FRT*s in *f08090* and *e01573*) in between *mini-white* and *FRT*. On the proximal side of the relo-cated *EY* element and next to the 3' P-element foot, 11.1 contains ~100 bp of DNA originating from the 5' end of *CR44953*, while 34.1 contains ~200 bp of DNA originating from the *rosy* locus. These insertions of heterologous DNA normally found on chromosome arm 3R are abutted by a ~1.7 kb deletion that extends to the left into the *apterous* region, 11.1 removing 8 bp more than 34.1. The two rearrangements are referred to as *ap^{11.1}* and *ap^{34.1}*. Apart from these, two other very similar rearrangements associated with smaller deficiencies were isolated. Their names are *ap^{72.2}* and *ap^{62.3}*. Their distal break point is the same as for *ap^{11.1}* and *ap^{34.1}* but they are smaller: 657 bp and 480 bp are missing, respectively. In both cases, hemizygous flies have normal wings, implying that the different position of their proximal deletion break is responsible for the wing phenotype observed for *ap^{11.1}* and *ap^{34.1}*. These observations map the distal end of the *ap* regulatory domain to a 1 kb interval between the proximal ends of deficien-cies *ap^{11.1}* and *ap^{72.2}*.

Generation of α-Ap antibody

DNA corresponding to amino acids 312 to 469 of *ap* cDNA clone HL02012 (DGRC, Indiana University) was amplified by PCR and cloned into pET22b(+) bacterial expression vector (Novagen) via NcoI and NotI sites. This fragment contains the Ap homeodomain (apHD), which is shared by all different Ap isoforms. The pelB leader sequence of pET22b vector was subsequently removed via mutagenesis PCR [74], resulting in the final expression plasmid pETapHD. BL21(DE3) bacteria (NEB) were transformed with pETapHD, grown to OD_{600nm} 0.6. T7 polymerase was induced with 0.1 mM IPTG. The protein was produced overnight at 18°C. Bacterial cells were lysed using a French press, then the lysate was loaded on a HisTrap HP column (GE Healthcare Life Sciences). apHD was purified with an ÄKTA HPLC machine. 3 mg of pure apHD were sent to Perbio Sciences Switzerland, where two rabbits were immu-nized. After 80 days, the serum of one positive rabbit was used to perform affinity purification of polyclonal antibody pool (final concentration: 0.67 mg/ml). For imaginal disc staining, the antibody is used at a dilution of 1:1000–2000.

Cloning of *in situ* rescue constructs

First, fragments C1 (size: 1.6 kb), C2 (3.6 kb), C3 (2.5 kb), C4 (1.6 kb), C5 (5.3 kb), C5A (3.8 kb), C5B (600 bp) and OR463 (463 bp) were amplified by PCR from clone BACR45O18 (Berkeley Drosophila Genome Project). The PCR primers had AvrII or XmaI sites overhangs, respectively (see S1 Table for Primer sequences). PCR-fragments were cut with AvrII and XmaI and subcloned into pBS KSII(+) vector, in which the XbaI site had previously been mutated into an AvrII site. Primers containing the XmaI site additionally had a SpeI site. AvrII and SpeI produce compatible sticky ends, which –when ligated- cannot be cut again by any of these enzymes. To combine the different fragments in the desired order, the following strategy was used: one fragment was cut out with AvrII and XmaI, and cloned into another pBSKSII subclone, that had a different fragment, via SpeI and XmaI sites. In the new subclone two dif-ferent fragments were combined, which could be cut out again via AvrII and XmaI sites and

cloned into another SpeI/XmaI cut plasmid. Subsequently, the combined fragments were cut out with AvrII/XmaI and cloned into AvrII/AgeI cut pEnh-Reentry plasmids, resulting in the final pEnh-Reentry constructs. Detailed description of the pEnh-Reentry plasmid can be found elsewhere [25]. Transgenic flies were obtained by injecting these plasmids (300ng/µl final concentration) into *y w M{vas-int.Dm}zh-2A; ap^{attPΔEnh}/CyO* embryos and stocks were established according to standard genetic practice [75].

Cloning of *ap* coding sequence *in situ* rescue constructs

ap cDNA was amplified from clone HL02012, the *ap* promoter region was PCRed from BAC clone BACR45O18 (Berkeley Drosophila Genome Project). The two fragments where combined by fusion PCR, and subcloned into pCR-XL-TOPO (Invitrogen). The *ap* promoter-cDNA fusion fragment was cloned into pCDS-Reentry vector [67] via NotI and AscI sites, to produce plasmid pCDS-Reentry-apcDNA. The pCDS-Reentry-apcDNAint2.3 construct, which contains the intron 2 and 3 of *ap* at the correct position, was synthesized by Genewiz, Inc. Transgenic flies were obtained by injecting these plasmids (300ng/µl final concentration) into *y w M{vas-int.Dm}zh-2A; ap^{attPΔCDS}/CyO* embryos and stocks were established according to standard genetic practice.

Generation of *lacZ* reporter and rescue transgenic lines

To generate C1–C5 and int2.3 reporter constructs, DNA from *ap* locus was amplified by PCR from *y^1 w^{67c23}* genomic DNA with primers containing restriction enzyme sites as overhangs, and subsequently cloned into plasmid pAttBLaZ [76] sing the respective enzymes (See S1 Table for primers and restriction enzymes). apE, apDV and apP were cloned into two reporter genes vectors, *attB-hs43-nuc-lacZ* [62] and *attB-pHPdesteGFP* [77]. The putative Pnt, Ap and Sd binding sites were identified on the basis of a bioinformatics analysis combining data from the JASPAR CORE Insecta database (http://jaspar.genereg.net/) and the Target Explorer tool [78].

Mutagenesis of the putative Pnt, Sd and Ap binding sites was performed using the Quik-Change Site-Directed Mutagenesis Kit (Stratagene). See S1 Table for sequence of all primers used in this study. All the reporter constructs were inserted and analysed at the same landing attP site. The reporter FRT-apDV+E-FRT-P-*lacZ* was generated cloning PCR FRT sequences flanking the *apDV* and *apE* elements with the *apP* following the last FRT. To delete the *apDV* and *apE* casette at different time points of development we drove *flp* in the posterior compartment by crossing FRT-apDV+E-FRT+P-*lacZ* containing flies to *en-Gal4*, UAS-*flp*, UAS-*GFP*; *tubGal80^{ts}*. Larvae were kept at 17°C to keep Gal4 off. At the desired time of development, the fly vials were shifted to 29°C for *flp* induction.

ap rescue experiments were done replacing the *lacZ* reporter gene of the *attB-hs43-nuc-lacZ* with the *ap* cDNA using EcoRI and KpnI in the different *ap* CRMs combinations. All *ap* rescue transgenes were inserted in the same attP site (86Fb).

trx and *Scm* mutant clonal analysis

Loss-of-function clones were generated by heat shocking the larvae for 1 hour at 37°C. The following genotypes were used:

y w hs FLP122; FRT 82B ubiGFP/ FRT 82B trx^{E2}
y w hs FLP122; FRT 82B ubiGFP/ FRT 82B Scm^{D1}
y w hs FLP122; FRT 82B ubiGFP/ FRT 82B trx^{E2} Scm^{D1}

Immunostaining

Imaginal discs were prepared and stained using standard procedures. The primary antibodies used were: rabbit and mouse anti-β-Gal (1:1000, Cappel and Promega), mouse anti-Wg (1: 50, Developmental Studies Hybridoma Bank), rat-αPS2 (1: 5, gift from Martín Bermudo) and rabbit anti-Ap (1:1000, this study)

Chromatin immunoprecipitation experiments

Third instar larvae were dissected and wing imaginal discs were collected in PBS on ice. Discs were fixed with 1.8% formaldehyde. Chromatin preparation and immunoprecipitation were performed as described [79]. For Ap ChIPs, 1.5 µg anti-Ap (dN–20, Santa Cruz Biotechnologies) was used for each immunoprecipitation, and specificity was tested by parallel "mock" immunoprecipitations carried out with normal goat IgG (Santa Cruz Biotechnologies). ChIP enrichment values were normalized relative to "mock" enrichment values to control for any signal that could be attributed to highly accessible chromatin [80]. Three real-time PCR amplicons surrounding the apP (chr2R, 1614425–1614545; coordinates based on dm3 build of *Drosophila* genome), apE (chr2R, 1622079–1622182), or apDV (chr2R, 1639774–1639867) elements were used to quantify immunoprecipitated chromatin. For Sd ChIP, maximum enrichment signals from Sd ChIP-chip data [79] for the corresponding apP, apE, and apDV regions were normalized to the same "mock" enrichment values used in the Ap ChIP experiments. Importantly, the Sd peak at apDV was called as statistically significant in the previously published genome-wide ChIP data [79].

Chromosome conformation capture (3C)

Chromosome conformation capture (3C) was performed as described in Webber *et al*, 2013 [81] with slight modification. Approximately 200 early third instar larvae were homogenized at room temperature in a crosslinking solution (1.8% formaldehyde, 50 mM HEPES, 1 mM EDTA, 0.5 mM EGTA, 100 mM NaCl). Total crosslinking time was limited to 20 minutes and followed by a 5-minute quench with glycine (0.125 M Glycine, 1xPBS, 0.01% Triton). Crude, fixed homogenate was then washed twice with PBS with 1% Triton, washed twice with a HEPES buffer (10 mM HEPES pH 7.6, 10 mM EDTA, 0.5 mM EGTA, 0.25% Triton), then Dounce homogenized in Buffer A (15 mM HEPES at pH 7.6, 10 mM KCl, 5 mM MgCl2, 0.1 mM EDTA, 0.5 mM EGTA, 350 mM sucrose, 1 mM DTT). After a brief centrifugation (400g for 1 minute) to remove cuticle and large debris, homogenate was centrifuged for 15 min at 10,000 rpm. Nuclei were resuspended in 100 µl of 1.2X DpnII Buffer with BSA (New England BioLabs), and then passed through a 27G syringe needle 10 times. 1.5 µl of 20% SDS was added to the nuclei-containing solution, which was then incubated for 30 min at 37°C, followed by 10 minutes at 65°C, addition of 10 µl 20% Triton X–100, and then incubation for 1 hour at 37°C. 100 units of DpnII were then added to the nuclei-containing solution, followed by overnight incubation at 37°C. The digestion reaction was stopped by adding 16 µl 10% SDS and incubating at 65°C for 10 minutes. From this point on, 3C was carried out as described in [81]. Ligation products were analyzed by qPCR (primer sequences available upon request). The amount of 3C amplicon product was normalized relative to an amplicon in the *ap* promoter that does not span a DpnII site and gives a measure of the total DNA in the reaction.

Supporting Information

S1 Fig. Wing discs and wing/notum preps of additional *ap* alleles. All 3rd instar wing discs were stained for Ap (green) and Wg (red). **(A-A")** ap^{MM}/ap^{DG3}: Ap and Wg patterns are

indistinguishable from wild type. Wings look normal. This indicates that ap^{MM} does not hamper ap function. **(B-D)** No Ap protein is detectable in hemizygous amorphic wing mutants ap^{DG16}, ap^{DG8}, and ap^{t11b} (over ap^{DG3}). **(B'-D')** Inner Wg ring is reduced to a dot, and wing pouch is lost. **(B"-D")** No wing tissue is formed in adult flies. **(E)** $ap^{c2.73c}/ap^{DG3}$: Ap is weakly detected in the dorsal part of the wing disc (white arrow). **(E')** Wing pouch is larger than in amophic mutants, but no D/V sub-division is observed. **(E")** Wing stumps or small tube-like structures are often formed in adults. **(F-F"')** $ap^{c1.60c}/ap^{DG3}$: in the weak hypomorphic mutant $ap^{c1.60c}$, ap is ectopically expressed in the ventral compartment correlating with the disruption of the Wg stripe at the D/V boundary (white arrow in F"'). All adult wings show notches along the wing margin (F""). All scale bars are 50 μm.
(TIF)

S2 Fig. *in situ* rescue system for *ap* coding sequences. (A) Relative order of C2 and C5 relative to apP has no influence on wing development. Hemizygous ap^{C15342} and ap^{C52} over ap^{DG3} flies develop normal wings. **(B)** Construction of $ap^{attP\Delta CDS}$: this ap allele harbors an attP docking site for the "coding sequence *in situ* rescue system". Initially, attP, FRT and LoxP sites were introduced at the ap^{MM} insertion site by direct gene conversion and ΦC31-mediated recombination. This intermediate allele is referred to as $ap^{attBPFRTy1}$ (for details see [25]). In a second step, the complete ap coding sequence was deleted by *flp*-mediated recombination between the two FRT sites in $ap^{attBPFRTy1}$ and ap^{f00878} and $ap^{attP\Delta CDS}$ was obtained. This deletion corresponds exactly to that in ap^{DG8} which leads to loss of all wing and haltere structures (see Fig 1B and S1C Fig). Its attP site allows the integration of ap coding sequences into the endogenous ap locus with the help of a plasmids like pCDS-Re-entry. The offspring can be screened for transgenics thanks to the *yellow* selection marker. **(C)** At the top of the panel, allele ap^{GFP} is shown. It contains the entire ap coding sequences with all introns specific for transcript ap-RA. The Ap protein is tagged with GFP at its C-terminal end (see [67] for a more detailed description). In ap^{GFP} hemizygous flies, ap function is fully complemented. The cDNA used for the construction of ap^{cDNA} and $ap^{cDNAint2.3}$ is also specific for transcript ap-RA. Introducing an intronless cDNA is not sufficient to re-establish wild type appearing wings (ap^{cDNA}). However, it has been proposed that intron-containing genes are often transcribed more efficiently than non-intronic genes, independently of putative enhancers in intronic sequences [83]. Thus, we engineered a cDNA/gDNA hybrid containing the two short introns 2 and 3 of ap. The corresponding allele $ap^{cDNAint2.3}$ was obtained. Hemizygous $ap^{cDNAint2.3}$ / ap^{DG3} flies fully rescue wing formation. **(D)** A 2.3 kb fragment containing intron 2 and 3 does not drive any detectable reporter gene expression in wing imaginal discs. Scale bars are 50 μm.
(TIF)

S3 Fig. Wing-disc specific expression patterns obtained with a collection of *ap* Gal4-driver lines and reporter constructs. (A) A schematic representation of the ap genomic region is depicted by a gray bar in the center of the panel. In green, the ap-RA transcript is indicated along with the five conserved regions C1–C5. apP, apE and apDV correspond to the regulatory elements characterized in this study. At the top of the panel, the location of two previously reported apE containing fragments apC [24] and apRXa [25] is indicated. The horizontal bars below represent the 17 DNA elements available as Gal4 drivers (Janelia Farm database) or *lacZ*-reporter constructs. At the bottom of the panel, the wing disc specific enhancer activity of conserved regions C1 to C5 in a *lacZ* reporter assay is shown. **(B)** 4 out of 17 DNA fragments tested show activity in the dorsal wing imaginal disc. All Janelia *Gal4* lines were crossed with a stock containing UAS-*GFP* (green) and act5C>stop>lacZ; UAS-*flp* to lineage-trace all the cells that at one point have activated *Gal4*. Wing discs were stained for GFP (green), *lacZ* (red), Wg (blue) and Cut (red) for line 2. Note that lines 1 and 2 are active in a similar pattern in the wing

pouch and hinge but are not active in the notum. Line 2 is more broadly expressed than line 1, with few cells showing activity in the ventral compartment (see arrow, Cut is in red). The other two lines active in dorsal wing disc cells are 7 and 8. They showed similar activity patterns in the notum and hinge regions with low levels in the dorsal wing pouch. Note that cells that have activated these DNA elements almost mark the entire dorsal compartment (*lacZ* in red). Also note that some cells labeled with *lacZ* of line 7 appear in the ventral compartment.
(TIF)

S4 Fig. *EGFR^ts* and *wg* experiments. (A-D) To reduce EGFR activity, a temperature-sensitive allele, *EGFR^ts* was used. Larvae of the genotype *EGFR^ts*; apE-*lacZ* were maintained at 17°C and shifted to 29°C to reduce EGFR activity for a 24hr period at different time points of larval development (time interval at 29°C is indicated below each imaginal disc picture). Then larvae were returned to 17°C until dissection at around 120hrs AEL. Imaginal discs stained for apE-*lacZ* (red) and Wg (green). (**A**) Control wing imaginal disc of a larva maintained at 17°C until dissection. (**B-D**) Wing imaginal disc shifted to 29°C at mid-third (**B**), early-third (**C**) and early-second (**D**) instar for a 24hr period. Note that apE is still active after *EGFR* removal at mid-third or early-third imaginal disc stage (**B** and **C**). Only removal of *EGFR* function at early-second instar completely abolishes apE activity (**D**). The resulting wing imaginal disc is strongly reduced in size and *wg* expression is lost. (**E**) *dpp-Gal4*; UAS-*TCF^DN*, UAS-*GFP* wing imaginal disc stained for apDV-*lacZ* (red) and GFP (green). Note that apDV activity is reduced (arrow in **E'**) although not eliminated after knockdown of the Wg pathway. Single channel is displayed for apDV-*lacZ* (**E'**). (**F**) Ectopic activation of the Wg pathway in *dpp-Gal4*; UAS-*arm^S10*, UAS-*GFP* wing imaginal disc does not ectopically activate apDV, with the exception of some scattered cells in the notum (arrow in **F'**). Single channel is displayed for apDV-*lacZ* (**F'**). Scale bars are 50 μm.
(TIF)

S5 Fig. apDV+E+P placed next to *ap*-cDNA is sufficient to rescue wing development.
(A-D) Wing imaginal discs of different genotypes stained for *ap-lacZ* (green), Wg (red) and αPs2 (white, in separate channels). For each genotype, the corresponding adult wing pheno-type is shown at the bottom of each panel along with details of the wing margin. *ap-lacZ* stands for *ap^rk568*. This is a *lacZ* enhancer trap line which behaves as a very strong *ap* allele. (**A**) *ap-lacZ*/+ wing imaginal discs show normal *ap-lacZ* and Wg pattern. αPS2 is restricted to ventral cells. Adult wings look normal. Dorsal and ventral patterning of the anterior wing margin is as in wild type. (**B**) *ap-lacZ*/*ap^UGO35* flies are amorphic and wing imaginal discs have no wing pouch. Adult flies do not develop any wings. (**C**) *ap-lacZ*/*ap^UGO35*; apDV+E-*apcDNA* homozy-gous flies: wing imaginal discs lack the D/V Wg stripe and αPS2 is observed in the entire pouch. Wing development is partially complemented but wing margin fails to form. (**D**) *ap-lacZ*/*ap^UGO35*; apDV+E+P-*apcDNA* homozygous flies: in wing imaginal discs, a normal Wg D/V stripe is present and with the exception of some dorsal cells (arrow), αPS2 is restricted to ventral cells. Although wing rescue is not perfect, a clear D/V margin is observed.
(TIF)

S6 Fig. Role of Trx and Scm in the regulation of *ap* expression. (**A**) Wild type wing imaginal disc stained for Wg (blue) and αPS2 (red). Note that αPS2 positive cells are confined to the ven-tral compartment. (**B**) *trx^E2* mutant clones generated 48–72hrs AEL: clones are marked by the absence of GFP. Discs were stained for Wg (blue) and αPS2 (red). (**B'**) single-channel picture of (**B**): αPS2 is derepressed in dorsal *trx^E2* clones (green arrow). (**C**) *ap-lacZ* (*ap^rK568*) expression in *Scm^D1* clones generated 48–72hrs AEL: clones are marked by the absence of GFP (several out-lined in white). Discs were stained for Wg (blue) and *ap-lacZ* (red). (**C'**) *ap-lacZ* expression is

derepressed in ventral clones close to the D/V boundary. **(C"-C"')** Close-up of **(C)**. Note *ap-lacZ* derepression in *Scm^{D1}* mutant cells close to the D/V **(C")**. *wg* expression does not follow *ap-lacZ* derepression **(C"')**. **(D)** *ap-LacZ* expression in *Scm^{D1} trx^{E2}* double mutant clones generated 48–72hrs AEL: clones are marked by the absence of GFP (several outlined in white). Discs were stained for Wg (blue) and for *ap-lacZ* (red). **(D')** *ap-lacZ* expression is downregulated in dorsal cells but no derepression is observed in ventral cells. **(D"-D"')** Close-up of (D'). Note *ap-lacZ* downregulation in the dorsal compartment in *Scm^{D1} trx^{E2}* mutant cells **(D")**. *wg* expression is not altered in *Scm^{D1} trx^{E2}* mutant cells. In particular, *wg* is not ectopically expressed along the edge of clones with reduced *ap-lacZ* activity **(D"')**. D, dorsal and V, ventral.
(TIF)

S7 Fig. apP, apE and apDV cooperate best when *in cis*. (A) In hemizygous +/*ap^{DG3}* flies, *ap* and *wg* expression patterns in wing discs are normal **(A', A")**. Apart from rare, mild margin defects, most wings are indistinguishable from wild type **(A"')**. Note that the 3 *ap* CRMs are all *in cis*. **(B)** *ap^{DG1}*/*ap^{t11b}*: apP is on one chromosome and apE and apDV are on the other. *ap^{DG1}* and *ap^{t11b}* alleles are amorphic when tested in hemizygous condition. *In trans* to each other, wing development is much improved. Typically, <20% of the wings appear normal. Among the rest, wings displaying an enlarged posterior compartment are frequent **(B"')**. Wing margin is rather well formed. Consistent with the adult phenotype, and although *ap* expression appears fairly normal, the posterior compartment is often overgrown in imaginal wing discs and the Wg stripe along the D/V border is wavy. **(B' and B")**. **(C)** *ap^{DG14}*/*ap^{DG12}*: formally, this genotype is equivalent to *ap^{C1345}*/*ap^{C1234}* shown in [Fig 7A](). apE and apDV are present *in trans* to each other. **(C' and C")** Expression of *ap* is affected in the dorsal compartment, leading to *wg* misexpression. **(C"')** All adult wings have similar phenotypes, including large, unstructured outgrowths. **(D)** *ap^{f00451}*/*ap^{DG3}*: on *ap^{f00451}*, apE and apDV enhancers are separated by a cluster of Su(Hw) binding sites. Many studies have shown that such clusters interfere with enhancer-promoter communication. **(D'-D"')** The phenotypes observed in all *ap^{f00451}*/ *ap^{DG3}* discs and adult wings suggest that apDV is not completely excluded from *ap* regulation. Their appearances are similar to those observed for *ap^{DG14}*/*ap^{DG12}* animals. From the similarities of the phenotypes, it may be inferred that *in trans* configuration of apE and apDV is equivalent to partially blocking apDV from interaction with apP. Scale bars are 50 μm.
(TIF)

S1 Table. Primer sequences used in this study. Primers used for the cloning of the different CRMs (the respective restriction enzymes used for cloning are indicated in the primer names). Mutagenesis of the Pnt, Sd and Ap putative binding sites (in bold) was performed using the QuikChange Site-Directed Mutagenesis Kit (Stratagene).
(TIF)

Acknowledgments

Author Contributions

Conceived and designed the experiments: DB MM MA CE. Performed the experiments: DB MM FH CE MS DR. Analyzed the data: DB MM MS CE. Contributed reagents/materials/analysis tools: DB MM OK DG PS CE. Wrote the paper: DB MM MA CE.

References

1. Irvine KD, Rauskolb C (2001) Boundaries in development: formation and function. Annu Rev Cell Dev Biol 17: 189–214. PMID: 11687488

2. Garcia-Bellido A, Ripoll P, Morata G (1976) Developmental compartmentalization in the dorsal mesothoracic disc of Drosophila. Dev Biol 48: 132–147. PMID: 1245256

3. Diaz-Benjumea FJ, Cohen SM (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. PMID: 8575321

4. Basler K, Struhl G (1994) Compartment boundaries and the control of Drosophila limb pattern by hedgehog protein. Nature 368: 208–214. PMID: 8145818

5. Nellen D, Burke R, Struhl G, Basler K (1996) Direct and long-range action of a DPP morphogen gradient. Cell 85: 357–368. PMID: 8616891

6. Zecca M, Basler K, Struhl G (1996) Direct and long-range action of a wingless morphogen gradient. Cell 87: 833–844. PMID: 8945511

7. Lecuit T, Brook WJ, Ng M, Calleja M, Sun H, et al. (1996) Two distinct mechanisms for long-range patterning by Decapentaplegic in the Drosophila wing. Nature 381: 387–393. PMID: 8632795

8. Neumann CJ, Cohen SM (1997) Long-range action of Wingless organizes the dorsal-ventral axis of the Drosophila wing. Development 124: 871–880. PMID: 9043068

9. Morata G, Lawrence PA (1975) Control of compartment development by the engrailed gene in Drosophila. Nature 255: 614–617. PMID: 1134551

10. Lawrence PA, Struhl G (1982) Further studies of the engrailed phenotype in Drosophila. EMBO J 1: 827–833. PMID: 6152896

11. Tabata T, Schwartz C, Gustavson E, Ali Z, Kornberg TB (1995) Creating a Drosophila wing de novo, the role of engrailed, and the compartment border hypothesis. Development 121: 3359–3369. PMID: 7588069

12. Zecca M, Basler K, Struhl G (1995) Sequential organizing activities of engrailed, hedgehog and decapentaplegic in the Drosophila wing. Development 121: 2265–2278. PMID: 7671794

13. Lawrence PA, Morata G (1976) Compartments in the wing of Drosophila: a study of the engrailed gene. Dev Biol 50: 321–337. PMID: 1278589

14. Blair SS, Brower DL, Thomas JB, Zavortink M (1994) The role of apterous in the control of dorsoventral compartmentalization and PS integrin gene expression in the developing wing of Drosophila. Development 120: 1805–1815. PMID: 7924988

15. Diaz-Benjumea FJ, Cohen SM (1993) Interaction between dorsal and ventral cells in the imaginal disc directs wing development in Drosophila. Cell 75: 741–752. PMID: 8242746

16. Cohen B, McGuffin ME, Pfeifle C, Segal D, Cohen SM (1992) apterous, a gene required for imaginal disc development in Drosophila encodes a member of the LIM family of developmental regulatory proteins. Genes Dev 6: 715–729. PMID: 1349545

17. Fernandez-Funez P, Lu CH, Rincon-Limas DE, Garcia-Bellido A, Botas J (1998) The relative expression amounts of apterous and its co-factor dLdb/Chip are critical for dorso-ventral compartmentalization in the Drosophila wing. EMBO J 17: 6846–6853. PMID: 9843490

18. Milan M, Cohen SM (1999) Regulation of LIM homeodomain activity in vivo: a tetramer of dLDB and apterous confers activity and capacity for regulation by dLMO. Mol Cell 4: 267–273. PMID: 10488342

19. van Meyel DJ, O'Keefe DD, Jurata LW, Thor S, Gill GN, et al. (1999) Chip and apterous physically interact to form a functional complex during Drosophila development. Mol Cell 4: 259–265. PMID: 10488341

20. Kim J, Sebring A, Esch JJ, Kraus ME, Vorwerk K, et al. (1996) Integration of positional signals and regulation of wing formation and identity by Drosophila vestigial gene. Nature 382: 133–138. PMID: 8700202

21. Rulifson EJ, Blair SS (1995) Notch regulates wingless expression and is not required for reception of the paracrine wingless signal during wing margin neurogenesis in Drosophila. Development 121: 2813–2824. PMID: 7555709

22. Spitz F, Furlong EE (2012) Transcription factors: from enhancer binding to developmental control. Nat Rev Genet 13: 613–626. doi: 10.1038/nrg3207 PMID: 22868264

23. Capovilla M, Kambris Z, Botas J (2001) Direct regulation of the muscle-identity gene apterous by a Hox protein in the somatic mesoderm. Development 128: 1221–1230. PMID: 11262224

24. Lundgren SE, Callahan CA, Thor S, Thomas JB (1995) Control of neuronal pathway selection by the Drosophila LIM homeodomain gene apterous. Development 121: 1769–1773. PMID: 7600992

25. Bieli D, Kanca O, Gohl D, Denes A, Schedl P, et al. (2015) The Drosophila melanogaster Mutants apblot and apXasta Affect an Essential apterous Wing Enhancer. G3 5:1129–1143. doi: 10.1534/g3.115.017707 PMID: 25840432

26. Wang SH, Simcox A, Campbell G (2000) Dual role for Drosophila epidermal growth factor receptor signaling in early wing disc development. Genes Dev 14: 2271–2276. PMID: 10995384

27. Zecca M, Struhl G (2002) Control of growth and patterning of the Drosophila wing imaginal disc by EGFR-mediated signaling. Development 129: 1369–1376. PMID: 11880346

28. Kassis JA, Brown JL (2013) Polycomb group response elements in Drosophila and vertebrates. Adv Genet 81: 83–118. doi: 10.1016/B978-0-12-407677-8.00003-8 PMID: 23419717

29. Kennison JA (1995) The Polycomb and trithorax group proteins of Drosophila: trans-regulators of homeotic gene function. Annu Rev Genet 29: 289–303. PMID: 8825476

30. Oktaba K, Gutierrez L, Gagneur J, Girardot C, Sengupta AK, et al. (2008) Dynamic regulation by polycomb group protein complexes controls pattern formation and the cell cycle in Drosophila. Dev Cell 15: 877–889. doi: 10.1016/j.devcel.2008.10.005 PMID: 18993116

31. Butler JE, Kadonaga JT (2002) The RNA polymerase II core promoter: a key component in the regulation of gene expression. Genes Dev 16: 2583–2592. PMID: 12381658

32. Negre N, Hennetin J, Sun LV, Lavrov S, Bellis M, et al. (2006) Chromosomal distribution of PcG proteins during Drosophila development. PLoS Biol 4: e170. PMID: 16613483

33. Tolhuis B, de Wit E, Muijrers I, Teunissen H, Talhout W, et al. (2006) Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in Drosophila melanogaster. Nat Genet 38: 694–699. PMID: 16628213

34. Schwartz YB, Kahn TG, Nix DA, Li XY, Bourgon R, et al. (2006) Genome-wide analysis of Polycomb targets in Drosophila melanogaster. Nat Genet 38: 700–705. PMID: 16732288

35. Gohl D, Muller M, Pirrotta V, Affolter M, Schedl P (2008) Enhancer blocking and transvection at the Drosophila apterous locus. Genetics 178: 127–143. doi: 10.1534/genetics.107.077768 PMID: 18202363

36. Pekowska A, Benoukraf T, Zacarias-Cabeza J, Belhocine M, Koch F, et al. (2011) H3K4 tri-methylation provides an epigenetic signature of active enhancers. EMBO J 30: 4198–4210. doi: 10.1038/emboj.2011.295 PMID: 21847099

37. Jory A, Estella C, Giorgianni MW, Slattery M, Laverty TR, et al. (2012) A survey of 6,300 genomic fragments for cis-regulatory activity in the imaginal discs of Drosophila melanogaster. Cell Rep 2: 1014–1024. doi: 10.1016/j.celrep.2012.09.010 PMID: 23063361

38. Milan M, Diaz-Benjumea FJ, Cohen SM (1998) Beadex encodes an LMO protein that regulates Apterous LIM-homeodomain activity in Drosophila wing development: a model for LMO oncogene function. Genes Dev 12: 2912–2920. PMID: 9744867

39. Rebay I (2002) Keeping the receptor tyrosine kinase signaling pathway in check: lessons from Drosophila. Dev Biol 251: 1–17. PMID: 12413894

40. Halder G, Polaczyk P, Kraus ME, Hudson A, Kim J, et al. (1998) The Vestigial and Scalloped proteins act together to directly regulate wing-specific gene expression in Drosophila. Genes Dev 12: 3900–3909. PMID: 9869643

41. Pueyo JI, Galindo MI, Bishop SA, Couso JP (2000) Proximal-distal leg development in Drosophila requires the apterous gene and the Lim1 homologue dlim1. Development 127: 5391–5402. PMID: 11076760

42. Morris JR, Petrov DA, Lee AM, Wu CT (2004) Enhancer choice in cis and in trans in Drosophila melanogaster: role of the promoter. Genetics 167: 1739–1747. PMID: 15342512

43. Morris JR, Chen J, Filandrinos ST, Dunn RC, Fisk R, et al. (1999) An analysis of transvection at the yellow locus of Drosophila melanogaster. Genetics 151: 633–651. PMID: 9927457

44. Duncan IW (2002) Transvection effects in Drosophila. Annu Rev Genet 36: 521–556. PMID: 12429702

45. Dekker J, Rippe K, Dekker M, Kleckner N (2002) Capturing chromosome conformation. Science 295: 1306–1311. PMID: 11847345

46. Estella C, McKay DJ, Mann RS (2008) Molecular integration of wingless, decapentaplegic, and autoregulatory inputs into Distalless during Drosophila leg development. Dev Cell 14: 86–96. doi: 10.1016/j.devcel.2007.11.002 PMID: 18194655

47. Struhl G, Basler K (1993) Organizing activity of wingless protein in Drosophila. Cell 72: 527–540. PMID: 8440019

48. Williams JA, Paddock SW, Carroll SB (1993) Pattern formation in a secondary field: a hierarchy of regulatory genes subdivides the developing Drosophila wing disc into discrete subregions. Development 117: 571–584. PMID: 8330528

49. Paul L, Wang SH, Manivannan SN, Bonanno L, Lewis S, et al. (2013) Dpp-induced Egfr signaling triggers postembryonic wing development in Drosophila. Proc Natl Acad Sci U S A 110: 5058–5063. doi: 10.1073/pnas.1217538110 PMID: 23479629

50. Simcox AA, Grumbling G, Schnepp B, Bennington-Mathias C, Hersperger E, et al. (1996) Molecular, phenotypic, and expression analysis of vein, a gene required for growth of the Drosophila wing disc. Dev Biol 177: 475–489. PMID: 8806825

51. de Celis JF, Bray S (1997) Feed-back mechanisms affecting Notch activation at the dorsoventral boundary in the Drosophila wing. Development 124: 3241–3251. PMID: 9310319

52. Klein T, Arias AM (1998) Different spatial and temporal interactions between Notch, wingless, and vestigial specify proximal and distal pattern elements of the wing in Drosophila. Dev Biol 194: 196–212. PMID: 9501029

53. McKay DJ, Estella C, Mann RS (2009) The origins of the Drosophila leg revealed by the cis-regulatory architecture of the Distalless gene. Development 136: 61–71. doi: 10.1242/dev.029975 PMID: 19036798

54. Crews ST, Pearson JC (2009) Transcriptional autoregulation in development. Curr Biol 19: R241–246. doi: 10.1016/j.cub.2009.01.015 PMID: 19321138

55. DeVido SK, Kwon D, Brown JL, Kassis JA (2008) The role of Polycomb-group response elements in regulation of engrailed transcription in Drosophila. Development 135: 669–676. doi: 10.1242/dev.014779 PMID: 18199580

56. Kwon D, Mucci D, Langlais KK, Americo JL, DeVido SK, et al. (2009) Enhancer-promoter communication at the Drosophila engrailed locus. Development 136: 3067–3075. doi: 10.1242/dev.036426 PMID: 19675130

57. Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, et al. (2002) MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol Cell 10: 1107–1117. PMID: 12453418

58. Poux S, Horard B, Sigrist CJ, Pirrotta V (2002) The Drosophila trithorax protein is a coactivator required to prevent re-establishment of polycomb silencing. Development 129: 2483–2493. PMID: 11973279

59. Klymenko T, Muller J (2004) The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. EMBO Rep 5: 373–377. PMID: 15031712

60. Schuettengruber B, Ganapathi M, Leblanc B, Portoso M, Jaschek R, et al. (2009) Functional anatomy of polycomb and trithorax chromatin landscapes in Drosophila embryos. PLoS Biol 7: e13. doi: 10.1371/journal.pbio.1000013 PMID: 19143474

61. Ghavi-Helm Y, Klein FA, Pakozdi T, Ciglar L, Noordermeer D, et al. (2014) Enhancer loops appear stable during development and are associated with paused polymerase. Nature 512: 96–100. doi: 10.1038/nature13417 PMID: 25043061

62. Giorgianni MW, Mann RS (2011) Establishment of medial fates along the proximodistal axis of the Drosophila leg through direct activation of dachshund by Distalless. Dev Cell 20: 455–468. doi: 10.1016/j.devcel.2011.03.017 PMID: 21497759

63. Perez L, Barrio L, Cano D, Fiuza UM, Muzzopappa M, et al. (2011) Enhancer-PRE communication contributes to the expansion of gene expression domains in proliferating primordia. Development 138: 3125–3134. doi: 10.1242/dev.065599 PMID: 21715425

64. Bejarano F, Milan M (2009) Genetic and epigenetic mechanisms regulating hedgehog expression in the Drosophila wing. Dev Biol 327: 508–515. doi: 10.1016/j.ydbio.2009.01.006 PMID: 19210960

65. Steffen PA, Ringrose L (2014) What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory. Nat Rev Mol Cell Biol 15: 340–356. doi: 10.1038/nrm3789 PMID: 24755934

66. Bischof J, Maeda RK, Hediger M, Karch F, Basler K (2007) An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. Proc Natl Acad Sci U S A 104: 3312–3317. PMID: 17360644

67. Caussinus E, Kanca O, Affolter M (2012) Fluorescent fusion protein knockout mediated by anti-GFP nanobody. Nat Struct Mol Biol 19: 117–121.

68. Siegal ML, Hartl DL (1996) Transgene Coplacement and high efficiency site-specific recombination with the Cre/loxP system in Drosophila. Genetics 144: 715–726. PMID: 8889532

69. Golic KG, Golic MM (1996) Engineering the Drosophila genome: chromosome rearrangements by design. Genetics 144: 1693–1711. PMID: 8978056

70. Venken KJ, Schulze KL, Haelterman NA, Pan H, He Y, et al. (2011) MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes. Nat Methods 8: 737–743. PMID: 21985007

71. Sipos L, Kozma G, Molnar E, Bender W (2007) In situ dissection of a Polycomb response element in Drosophila melanogaster. Proc Natl Acad Sci U S A 104: 12416–12421. PMID: 17640916

72. Gloor GB, Nassif NA, Johnson-Schlitz DM, Preston CR, Engels WR (1991) Targeted gene replacement in Drosophila via P element-induced gap repair. Science 253: 1110–1117. PMID: 1653452

73. Bateman JR, Lee AM, Wu CT (2006) Site-specific transformation of Drosophila via phiC31 integrase-mediated cassette exchange. Genetics 173: 769–777. PMID: 16547094

74. Makarova O, Kamberov E, Margolis B (2000) Generation of deletion and point mutations with one primer in a single cloning step. Biotechniques 29: 970–972. PMID: 11084856

75. Greenspan R (1997) Fly pushing: The theory and practice of Drosophila genetics.

76. Weiss A, Charbonnier E, Ellertsdottir E, Tsirigos A, Wolf C, et al. (2010) A conserved activation element in BMP signaling during Drosophila development. Nat Struct Mol Biol 17: 69–76. doi: 10.1038/nsmb. 1715 PMID: 20010841

77. Boy AL, Zhai Z, Habring-Muller A, Kussler-Schneider Y, Kaspar P, et al. (2010) Vectors for efficient and high-throughput construction of fluorescent drosophila reporters using the PhiC31 site-specific integration system. Genesis 48: 452–456. doi: 10.1002/dvg.20637 PMID: 20506180

78. Sosinsky A, Bonin CP, Mann RS, Honig B (2003) Target Explorer: An automated tool for the identification of new target genes for a specified set of transcription factors. Nucleic Acids Res 31: 3589–3592. PMID: 12824372

79. Slattery M, Voutev R, Ma L, Negre N, White KP, et al. (2013) Divergent transcriptional regulatory logic at the intersection of tissue growth and developmental patterning. PLoS Genet 9: e1003753. doi: 10. 1371/journal.pgen.1003753 PMID: 24039600

80. Teytelman L, Thurtle DM, Rine J, van Oudenaarden A (2013) Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. Proc Natl Acad Sci U S A 110: 18602–18607. doi: 10.1073/pnas.1316064110 PMID: 24173036

81. Webber JL, Zhang J, Mitchell-Dick A, Rebay I (2013) 3D chromatin interactions organize Yan chromatin occupancy and repression at the even-skipped locus. Genes Dev 27: 2293–2298.

82. Thibault ST, Singer MA, Miyazaki WY, Milash B, Dompe NA, et al. (2004) A complementary transposon tool kit for Drosophila melanogaster using P and piggyBac. Nat Genet 36: 283–287. PMID: 14981521

83. Moabbi AM, Agarwal N, El Kaderi B, Ansari A (2012) Role for gene looping in intron-mediated enhancement of transcription. Proc Natl Acad Sci U S A 109: 8505–8510. doi: 10.1073/pnas.1112400109 PMID: 22586116

SUPPLEMENTARY FIGURES

Figure S1

Bieli et al

$\dfrac{ap^{MM}}{ap^{DG3}}$

$\dfrac{ap^{DG16}}{ap^{DG3}}$

$\dfrac{ap^{DG8}}{ap^{DG3}}$

$\dfrac{ap^{t11b}}{ap^{DG3}}$

$\dfrac{ap^{c2.73c}}{ap^{DG3}}$

$\dfrac{ap^{c1.60c}}{ap^{DG3}}$

Figure S2   Bieli et al

A

ap^{C15342} C1 C5 C3 C4 C2 →

ap^{C52} C5 C2 →

B

f00878   *apterous* ap^{MM}
PRE

gene conversion
+ deletion of downstream region

$ap^{attPΔCDS}$ FRTLoxP attP

LoxP coding sequences
yellow attB
pCDS-Re-entry

ΦC31 integrase
+
re-entry construct with attB site and coding sequences

C

ap^{GFP} attL // *yellow* *GFP* attR →
PRE

ap^{cDNA} attL // *ap cDNA* attR →
PRE

$ap^{cDNAint2.3}$ attL // attR →
PRE

D

apint2.3 (+) *lacZ*

apint2.3 (-) *lacZ*

80

Figure S4   Bieli et al

A Wg / apE — Control

B Mid-third instar to 29°C

C Early-third instar to 29°C

D Early-second to 29°C

E dpp>TCF^DN / apDV **E'** apDV

F dpp>arm^S10 / apDV **F'** apDV

Figure S5   Bieli et al

A *ap-lacZ/+*

B *ap^UGO35/ap-lacZ*

C *ap^UGO35/ap-lacZ* apE+P-*apcDNA*

D *ap^UGO35/ap-lacZ* apDV+E+P-*apcDNA*

ap-lacZ / Wg

D / V

Wg / αPS2

Dorsal / Ventral

No DV

No DV

Wing

Figure S6   Bieli et al

A αPS2 Wg — wild type
B αPS2 GFP Wg — trx^E2 clones
B' αPS2
C GFP ap-lacZ Wg — Scm^D1 mutant clones
C' ap-lacZ
C'' ap-lacZ
C''' Wg
D GFP ap-lacZ Wg — trx^E2 Scm^D1 mutant clones
D' ap-lacZ
D'' ap-lacZ
D''' Wg

Figure S7   Bieli et al

A apP apE apDV ap^DG3
A' Ap   A'' Ap Wg   A'''
B ap^DG1 apP ap^t11b apE apDV
B'   B''   B'''
C ap^DG14 apP apE ap^DG12 apP apDV
C'   C''   C'''
D ap^f00451 apP Su(Hw) apE apDV ap^DG3
D'   D''   D'''

83

UNPUPLISHED RESULTS AND DISCUSSION

Dissection of the *apterous* regulatory landscape

Generation and validation of a new polyclonal *apterous* antibody

Soon after initiation of this project, we realized that it would be absolutely necessary to have a functional anti-Ap antibody. Aliquots of a polyclonal antibody generated 25 years ago by the Thomas-Lab (San Diego) were no longer available. Moreover, commercially available antibodies against Ap (Santa Cruz Biotechnology) were not functioning with a number of tissue staining protocols (data not shown). Therefore, we tried to raise antibodies against a number of specifically selected Ap peptides (Peptide Specialty Laboratories GmbH). However, these attempts to isolate functional antibodies failed (data not shown). Thus, we decided to go for a more conventional approach; to produce recombinant Ap protein and inject it into rabbits in order to induce an immune response.

First, we tried to express full length Ap protein in *E. coli*. Although recombinant protein was detected in induced bacterial lysates via Western Blots (data not shown), the resulting protein amounts would not have been sufficient to purify large amounts needed for antibody production. Thus, we divided the protein into two parts, an N-terminal part including the LIM domains and a C-terminal part which encodes for the homeodomain (HD). While the part with the LIM domains was not expressed, the apHD was produced in large amounts in the cytoplasm of expression *E. coli* (data not shown). Importantly, the apHD peptide is shared by all Ap protein isoforms. As a next step, two polyclonal rabbit antibodies were raised against the apHD, and one showed specific signal in the wing imaginal disc. Based on the Ap wing disc expression pattern, signal-to-noise ratio was increased by antibody affinity purification, choice of best secondary antibody (AlexaFluor 568, Abcam), dilution factor (1:1000-1:2000) and usage of Pierce Immunostain Enhancer (Thermo Fisher Scientific). Finally, the new polyclonal serum was validated by analyzing Ap expression pattern in all embryonic and larval tissues where Ap expression had been reported.

During embryogenesis, *ap* shows a very dynamic and complex expression pattern. Earliest expression is observed during germ band extension in the mandibular lobe (m) and later in cells forming the anus (a). By the end of germ band extension, the expression shows segmental repetition in dorsal- and ventral-lateral groups of cells (COHEN *et al.* 1992). By the end of germ

84

band retraction, those cells have invaginated under the ectoderm, where they are partially associated with cells of the peripheral nervous system (Figure 11A). Some of these lateral cells are also precursor cells for a distinct set of muscles (asterisk in Figure 11A). Expression in the brain is also observed, starting when the germ band has fully extended and persisting throughout embryogenesis into larval stages. During germ band retraction, *ap* is expressed in the ventral nerve cord (VNC) and brain hemispheres (Figure 11B). There, it is detected in three cells, one dorsal and two ventral, in each hemisegment and in an additional cluster of four lateral cells in each thoracic hemisegment (Lundgren *et al.* 1995). Additionally, *ap* gets expressed in the filzkörper at the posterior spiracles.
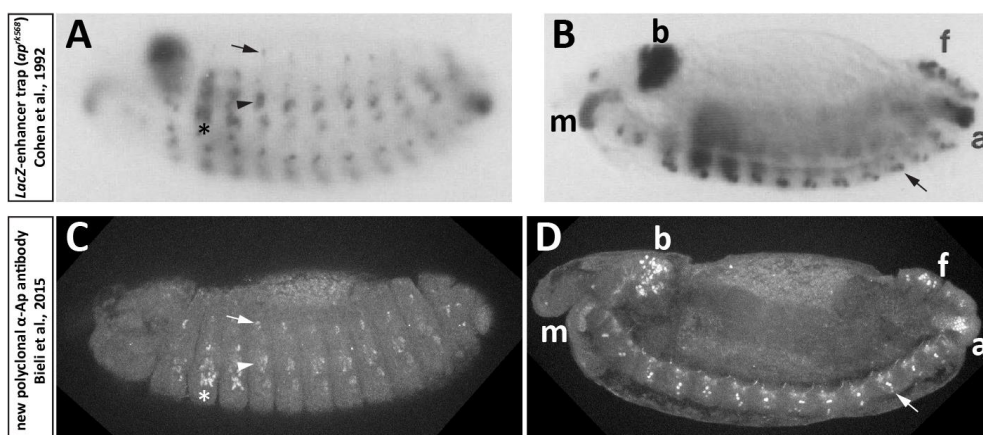
Figure 11 Comparison of reported *ap* expression in the embryo to expression pattern detected with the newly generated anti-Ap antibody. (A and B) Complex endogenous *ap* expression pattern by *LacZ*-enhancer trap *ap^{rk568}* as reported in Cohen et al. 1992, modified. **(C and D)** Embryonic expression patterns as detected with the new anti-Ap serum generated for this study. **(A)** *ap* is expressed in repeated lateral clusters beneath the epidermis (arrow and arrowhead) and in muscle precursor cells (asterisk). **(B)** *ap* expression is detected in the mandibular region (m), brain (b), filzkörper (f), anus (a) and the ventral nerve cord (arrow). **(C)** Lateral cross-section of an embryo stained with the newly developed anti-Ap antibody. Expression is seen in the lateral clusters (arrow and arrowhead) and the presumptive muscle precursors (asterisk). **(D)** Medial cross-section of the same embryo seen in C. Specific signal is observed in the brain (b), filzkörper (f), anus (a), mouth (m) and the VNC (arrow).

With the newly developed antibody, Ap protein was detected in all the described structures with high specificity (Figure 11C and D).

Besides *ap* expression in the dorsal compartment of the wing imaginal disc, *LacZ* expression by *ap^{rk568}* was detected in the haltere disc, leg disc, eye-antennal disc and brain (Figure 12; Cohen *et al.* 1992).

Using the new antibody, we could reveal Ap expression in the dorsal compartment of the wing and haltere, the ring-like structure of the fourth tarsal segment in the leg disc and the central dot in the antennal disc (Figure 12D-G). Even though the brain expression pattern was not

analyzed in detail, the staining was comparable to the published expression pattern, with rings in the brain hemispheres and clusters in the VNC (compare 12H to C).

However, additionally to the dot in the antennal disc, strong staining was observed in the developing ommatidia in the eye disc (see arrow in Figure 12G). This pattern has never been reported for *ap* and was still present in homozygous amorphic *ap^DG8* mutants that remove all the protein-coding sequences (data not shown). Therefore, it is very unlikely that the staining in the ommatidia is *ap*-specific. Thus, the sequence of the apHD was blasted (flybase.org) and all the proteins/genes from BLAST-results that showed similarity to the apHD were checked for their available expression patterns. With an overall identity score of 36.6%, the protein LIM3 was the lowest on the BLAST list. However, LIM3 shares a lot of structural similarities to Ap, as it also has two LIM domains and one HD (THOR *et al.* 1999). Additionally, it has been shown to be expressed in ommatidia of the developing eye-antennal disc (ROIGNANT *et al.* 2010). LIM3 expression in the larval brain, wing disc, and leg disc has not been reported. Thus, it seems that the newly developed polyclonal anti-Ap antibody shows some cross-reactivity with the LIM3 protein.

Overall, the anti-Ap antibody seems to be very specific in all of the tested tissues with minor Ap-unspecific binding in the eye-disc. As LIM3 is also expressed in the embryonic VNC, carefully re-examination of expression patterns might be necessary.
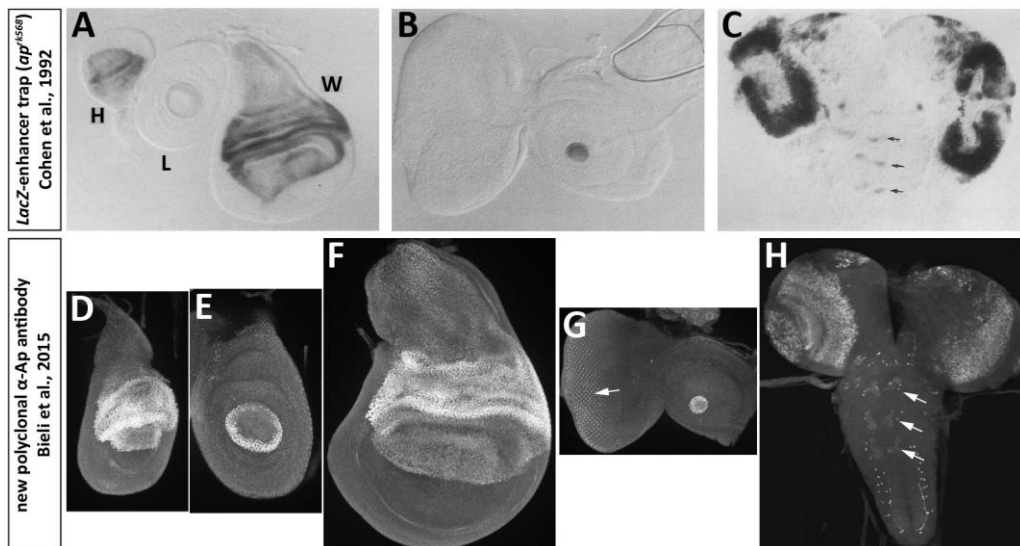
Figure 12 Comparison of reported *ap* expression in larval structures to expression pattern detected with the newly generated anti-Ap antibody. (A-B) *LacZ* expression of *ap^rk568* in haltere (H), leg (L) wing (W), and eye-antennal discs according to COHEN et al. 1992, modified. **(C)** Complex *ap* expression pattern in the brain lobes and VNC. **(D-H)** Staining patterns obtained with newly developed anti-Ap antibody. **(D)** Ap is expressed in the dorsal compartment of the haltere disc. **(E)** Ring-like expression of Ap in the leg disc. **(F)** Ap expression pattern in the dorsal compartment of the wing imaginal disc. **(G)** Eye-antennal disc with specific dot in the antennal part and unspecific staining in the ommatidia (arrow). **(H)** Circular Ap expression is seen in the brain lobes, and in repeated clusters in the VNC (arrows). Additionally, single cells on both sites of the along the entire VNC are stained.

Extended expression analysis of selected *apterous* reporter constructs

As nicely shown in the previous section, *ap* is expressed in many other tissues apart from the wing disc. So far, in this study, expression patterns of enhancer fragments have been investigated primarily by focusing on the wing imaginal discs. In order to define further *cis*-regulatory elements relevant for the complex and dynamic expression pattern of *ap* during fly development, we have analyzed reporter gene constructs in additional larval tissues and in the embryo.
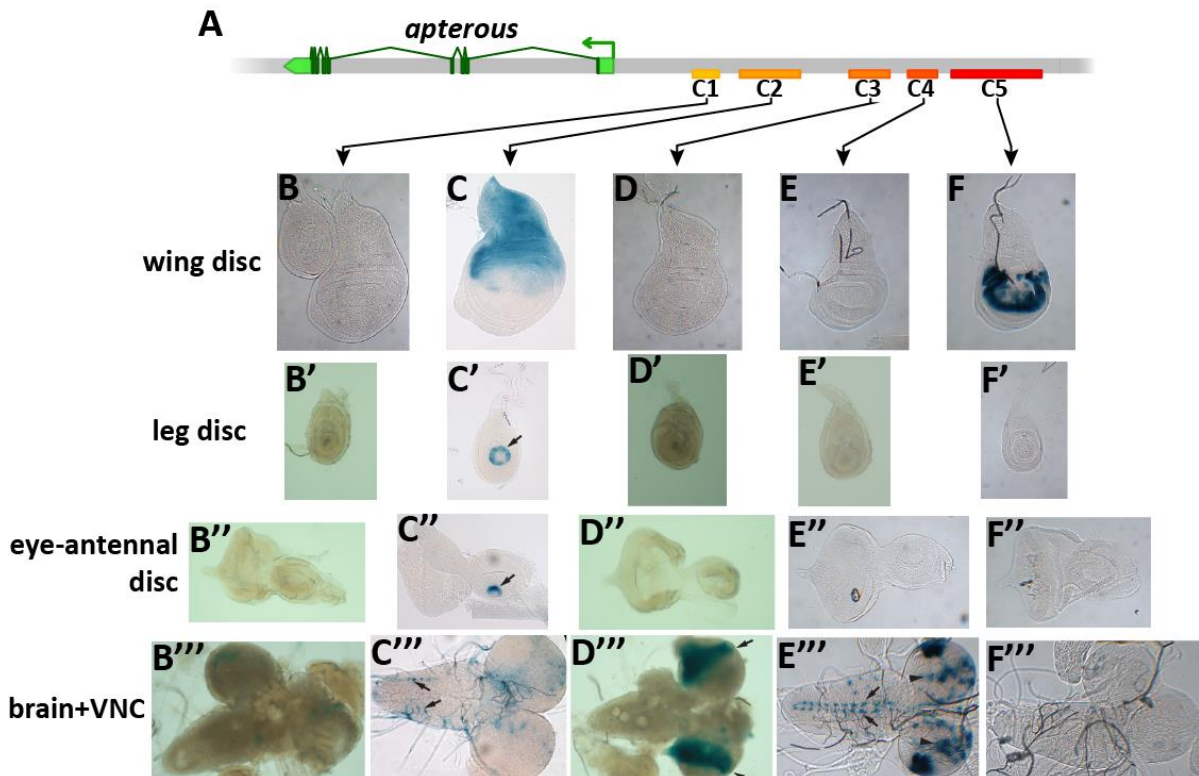
Figure 13 Summary of expression patterns of conserved fragments in imaginal discs and the larval brain. (A) Overview of the *ap* locus and the conserved regions (C1-C5) tested for their respective enhancer activity. **(B-B''')** C1-*LacZ* is not expressed in any of the tested larval structures. **(C-C''')** C2 drives expression throughout the dorsal compartment, in a ring in the leg disc, in a dot in the antennal disc and in the VNC (arrows). **(D-D''')** C3-*LacZ* driving reporter gene expression only in a ring in each brain lobe (arrows). **(E-E''')** C4-*LacZ* activity is seen in the brain lobes (arrowheads) and in repeating clusters in the VNC (arrows). **(F-F''')** C5-*LacZ*, which contains the apDV element, is only active in the dorsal compartment along the D/V compartment boundary.

First, we examined the expression patterns of the five conserved fragments upstream of *ap* in other imaginal discs (Figure 13). As already described comprehensively in the previous chapters, C2 and C5 contain the apE and apDV wing enhancers, respectively, and show activity in the wing discs (Figure 13C and F). While C5-*LacZ* was only active in the wing disc (Figure 13F-F'''), C2-*LacZ* activity was detected in the leg disc and antennal disc in a pattern resembling endogenous *ap* expression (Figure 13C' and C''). C2 also drove weak reporter gene expression in the VNC (Figure 13C'''). The VNC expression of C2 can be separated from the imaginal disc

expression. Interestingly, the smallest version of the wing disc enhancer (apE-1) still contains the leg and antennal disc enhancer (data not shown, Master Thesis Dimitri Bieli, 2011). The significance of this observation is discussed below. The reporter C1-*LacZ* showed no activity in all the examined larval tissues (Figure 13B). C3-*LacZ* was only active in the larval brain lobes (Figure 13E'''). C4 drove detectable β-Gal expression in the VNC and the brain lobes, but no activity was observed in the imaginal discs (Figure 13E-E''').

Next, we also investigated the expression pattern generated by the *ap* promoter (approm-*LacZ*; see Figure 14A). It is well known that reporter constructs containing a Polycomb Response Element (PRE) are very sensitive to position effects they encounter at the respective insertion sites (Kassis 2002; Müller and Kassis 2006). Since the landing site (at 86F) we generally used to insert our other reporter constructs was demonstrated to exhibit minimal position effect (Bischof *et al.* 2007; Weiss *et al.* 2010), we wanted to investigate the expression pattern of approm-*LacZ* at different genomic locations. To do so, approm-*LacZ* was introduced to a P-element vector (see Material and Methods), and independent transgenic insertion events were isolated.
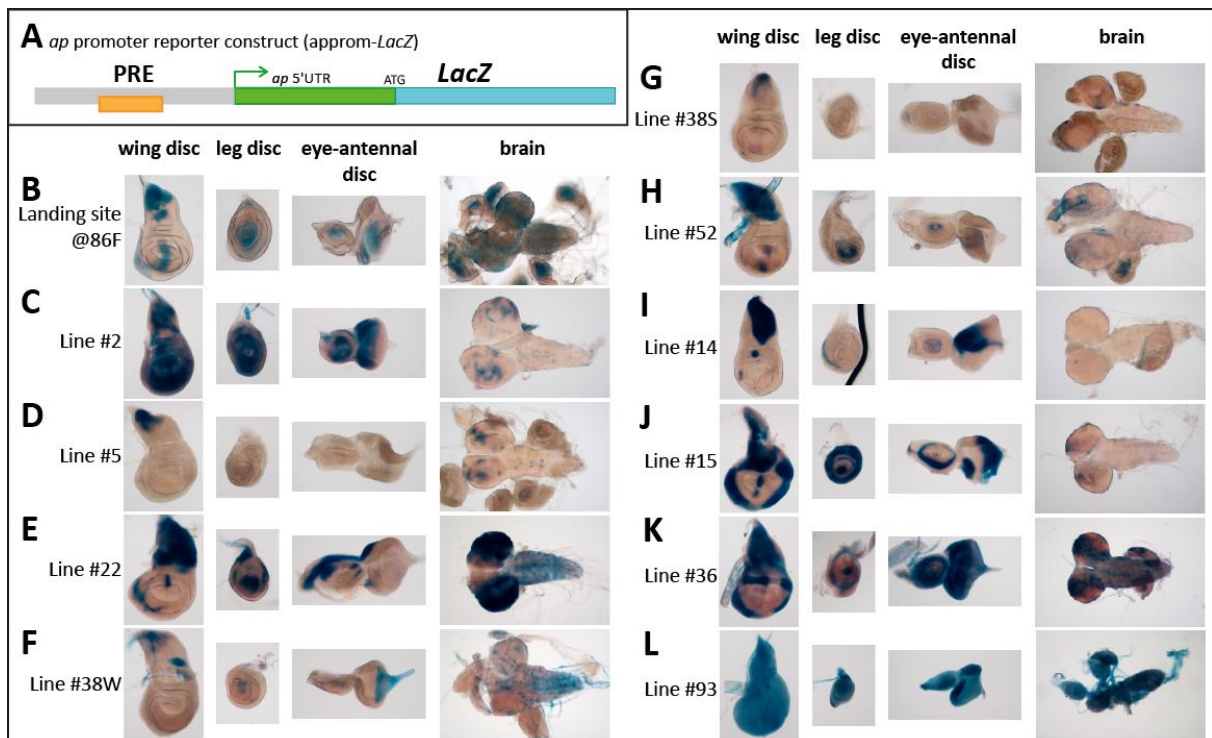
Figure 14 Summary of expression patterns of approm-LacZ in imaginal discs and the larval brain. (A) Schematic representation of the approm-*LacZ* reporter construct. It spans from the 400 upstream region of TSS of *ap-RA/RC*, covering the core PRE, over the 5'UTR of *ap-RA/RC* to the ATG, where it is fused to the ORF of the LacZ gene. **(B)** Expression pattern generated by approm-*LacZ* inserted in the attB-landing site at 86F. **(C-L)** Expression patterns generated by the random P-element insertions of approm-LacZ.

Figure 14 gives an overview of all the expression patterns generated in the wing, leg, eye-antennal discs and larval brain by the various reporter lines. As expected, no reproducible expression was observed when the different reporter lines were compared to each other, and certainly no pattern resembling the endogenous *ap* expression was generated (Figure 14B-L). The only common expression generated by all the investigated reporter lines was detected in the most dorsal tip of the future notum.

In addition to the larval tissues, we also studied the embryonic expression patterns of the aforementioned reporter gene constructs.

No activity was observable in approm-*LacZ*, C3-*LacZ*, C4-*LacZ* and C5-*LacZ* embryos (data not shown). Fragment C1 was weakly active in the most dorsal-lateral cells and in some lateral cells of the peripheral nervous system (PNS; Figure 15A'). The C2-LacZ reporter also displayed activity in the lateral PNS, but in more cells towards the ventral side than C1 (Figure 15B'). Additionally, the C2 construct drove reporter gene expression in the mandibular lobe and anus (Figure 15C'). None of the tested reporters showed expression in the embryonic VNC, the brain, the filzkörper or in the lateral muscle precursors (Figure 15B and C).
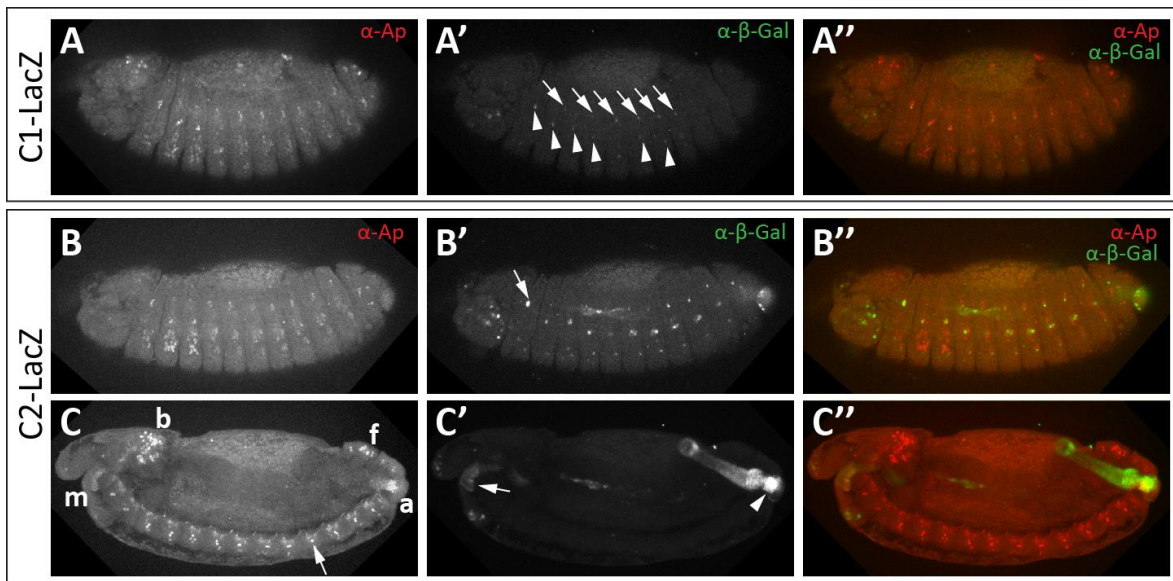
Figure 15 Expression analysis in embryonic structures. (A and B) Embryonic expression of *ap* in the lateral cells. **(C)** Ap expression in the brain, VNC, filzkörper anus and mouth. **(A')** C1-*LacZ* is active in most dorsal cells of the PNS (arrows and arrowheads). **(B')** β-Gal reporter is detected laterally (dorsal and ventral) in PNS cells. **(C')** C2-*LacZ* drives also reporter gene expression in the mandibular lobes (arrow) and anus (arrowhead). **(A'''-C''')** Merge of the Ap signal (red) and β-Gal reporters (green).

However, we (Master Thesis Dimitri Bieli, 2011) and others (DE TAFFIN *et al.* 2015) have demonstrated that the embryonic VNC enhancer resides just distal to the apE wing enhancer.

A *cis*-regulatory element important for the expression in the embryonic muscles had been isolated from the fourth intron (CAPOVILLA *et al.* 2001). Right next to the muscle enhancer, a fragment was found to drive reporter gene expression the embryonic brain lobes and in the filzkörper (BERGMAN *et al.* 2002).

Overview of the *cis*-regulatory architecture of the *apterous* locus

As described above, the LIM-HD transcription factor Ap is expressed in complex patterns and important for the development of various tissues throughout development (COHEN *et al.* 1992). Based on the transcriptional analysis performed by of us (PUBLICATION I and II, Master Thesis Dimitri Bieli) and by others (LUNDGREN *et al.* 1995; CAPOVILLA *et al.* 2001; BERGMAN *et al.* 2002; DE TAFFIN *et al.* 2015), it is now possible to draw a more comprehensive picture of the *cis*-regulatory landscape at the *ap* locus. Figure 16 summarizes all the known enhancer elements and shows their relative positions in the locus.
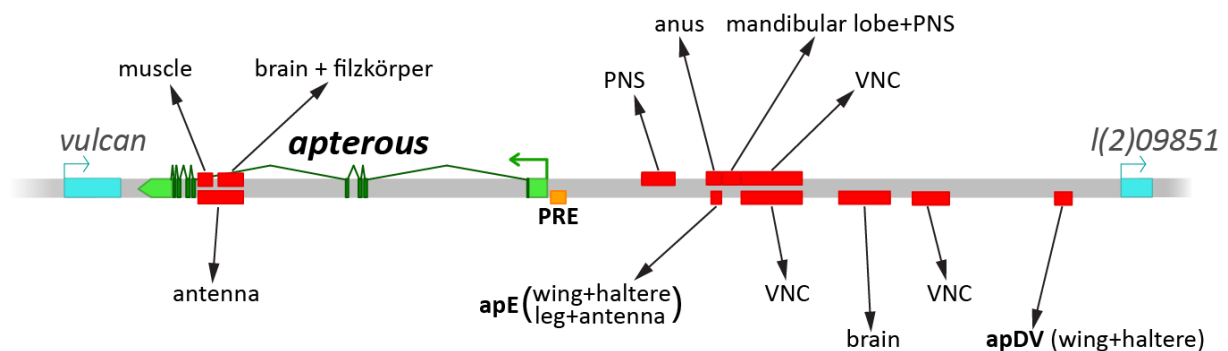
Figure 16 Overview of the *cis*-regulatory landscape of *apterous*. Red bars depict the different enhancers. On top: enhancers that are active in embryonic tissues. Bottom: enhancers active in the larva.

Ap expression is seen in in cells of the embryonic muscles, brain, filzkörper, peripheral nervous system (PNS), anus, mandibular lobe, ventral nerve cord (VNC) as well as in larval tissues including the wing, haltere, leg, and antennal discs as well as the brain and VNC. For all these tissues, one or more *cis*-regulatory elements were isolated.

We have only characterized the elements important for wing development in greater detail; it is very likely that the other elements can also be shortened to reach an even higher resolution of the regulatory landscape of *ap*. Enhancer-reporter constructs available from the FlyLight database and stock center (Janelia Farm, Virginia) and careful re-examination of all available deletions and mutants will certainly help to analyze the other *cis*-regulatory elements of *ap* in the respective tissues.

Interestingly, the *cis*-regulatory elements show a remarkable flexibility in their relative positions. At least with respect to wing development, we demonstrated that the position of the two enhancers can be swapped (i.e. apE and apDV) without affecting overall wing development. Whether this observation is an exception or whether this modularity also applies to the other *cis*-regulatory elements remains to be investigated.

Also the role of the Polycomb Response Element (PRE) in the transcriptional regulation of *ap* has been mainly analyzed in wing disc development. However, deletion of the PRE also resulted in loss or strong reduction of Ap expression in the leg disc, larval brain and VNC (data not shown). However, Ap expression in the center of the antennal disc seemed to be rather unaffected. As shown in Figure 16, there are two *cis*-regulatory elements which showed expression in the antennal disc; one in the 4th intron (Carlos Estella, personal communication) and one close to the apE wing enhancer in the intergenic region between *ap* and *l(2)09851*. The transcription start site (TSS) of the *ap-RB* isoform is located directly downstream of the antennal enhancer in the 4th intron. In contrast to the characterized *ap* PRE, no binding of Polycomb group factors has been reported in proximity of this particular TSS. Based on these observations, it is be possible that distinct transcriptional units, which consist of different TSS, isoforms and *cis*-regulatory elements are present at the *ap* locus. To test this hypothesis, it is absolutely necessary to perform a more detailed analysis of the various mutants (e.g. ap^{cDNA}, $ap^{cDNAint2.3}$, or $ap^{c1.2b}$) and reporter gene constructs, also in embryonic tissues. Since the anti-Ap antibody we isolated should recognize all the reported isoforms, it will be necessary to perform transcript-specific *in situ* hybridization or qPCR.

In recent years, it became apparent that genomes are organized into so-called 'topologically-associated domains' or TADs (SEXTON *et al.* 2012; DIXON *et al.* 2012; CIABRELLI and CAVALLI 2014). Genes residing in these higher-order chromatin structures seem to show a certain degree of co-regulation (HURST *et al.* 2004) and have common active or repressive chromatin marks (ERNST and KELLIS 2010; KHARCHENKO *et al.* 2011). In general, the average TAD size ranges from 20-200 kb in *Drosophila* (SPELLMAN and RUBIN 2002) to 1 Mb in mammals (CARON *et al.* 2001). The TADs are thought to be delimited by boundary elements to which CTCF and other so-called insulator proteins, like BEAF and CP190, are bound (SEXTON *et al.* 2012; DIXON *et al.* 2012; HOU *et al.* 2012).

ChIP-Chip data on chromatin states available on modencode.org show repressive marks that cover the entire *vulcan* and *ap* locus to the 5' end of *l(2)09851*. Usually, *Drosophila* transgenic

lines containing a transposable element marked by mini-*white* display reddish eye color. However, eye color of transposon inserted into the *vulcan* or *ap* locus was demonstrated to be actively repressed (Martin Müller and Daryl Gohl, unpublished). This 'trans-repression domain' happens to coincide with the repressive histone marks found in these loci. Moreover, putative BEAF-binding sites were found distal to the apDV enhancer. Based on these observations, it could be that the *ap* and *vulcan* loci, which are spanning a region of about 50 kb, are forming a TAD. In this regard, it is interesting to mention that the deletions isolated between apDV and *l(2)09851* remove various BEAF sites, without having a observable effect on wing development (see below, Figure 22). However, it cannot be excluded that neighboring genes are mis-regulated in these small deletions without having obvious phenotypic consequences. Alternatively, the 11 kb P-element present at the distal end of these deletions ight also be sufficient to shield the flanking genes. As there will certainly be more precise high-resolution data on whole genome interactions available in the future, the question with which other genes *ap* forms a TAD can likely be answered more precisely. Moreover, it will also be interesting be see which role the *ap* PRE plays in the formation of these higher-order chromosome structures.

Role of *ap* in the evolution of insect wings

The appearance of insect wings is a nice example for a morphological novelty in evolution; however, the origin of insect wings remains unclear. At the moment, there are two prominent hypotheses to explain the evolution of insect wings: The paranotal hypothesis and the exite hypothesis. In the paranotal hypothesis, the wing is thought to have evolved from lateral outgrowths of the notum, which might have helped ancient insects to glide (Rᴀsɴɪᴛsʏɴ 1981). This theory is largely based on fossil records of wing-like structures on the notum of the prothoracic segment of Paleozoic insects (Wᴏᴏᴛᴛᴏɴ and Kᴜᴋᴀʟᴏᴠᴀ́-Pᴇᴄᴋ 2000). The other theory, the exite hypothesis, proposes that wings originated from ancestral outer leg branches (exites), which appeared to have fused into the body wall (Sɴᴏᴅɢʀᴀss 1935). The derivation of the wings from exites would also provide a nice explanation for the development and attachment of the flight muscles, which would have simply been provided by the ancestral leg segment. This hypothesis is mainly supported by Evo-Devo data like the shared expression of wing-relevant genes, such as *ap* and *Dll*, in *Drosophila* wings and crustacean exites (Aᴠᴇʀᴏꜰ and Cᴏʜᴇɴ 1997).

Averof and Cohen also proposed that the evolution of the insect wings involved genetic and developmental changes. More specifically, it was suggested that ''*apterous expression and its restriction to the dorsal surface of the wing* [is used] *to initiate a DV patterning system not used in the ancestral limb or present-day insect legs*'' (AVEROF and COHEN 1997).

In this regard, it is very interesting to note that the *cis*-regulatory element needed for initiation of *ap* expression, apE, also shows activity in the leg disc. Importantly, in our experiments, the wing expression of apE was always accompanied by expression in the leg and the antennal discs and could never be separated in any of the numerous sub-fragments from this region (see above). Thus, it could very well be that the regulatory network important for *ap* expression in the dorsal region of the metathoracic tissue was adopted from the leg or exite structures. However, *ap* is dispensable for the development of the notum (see below). The main purpose of the early, dorsal *ap* expression is the initiation of the *ap*-Notch-Wg-*vg* autoregulatory system or –in other words- a ''*DV patterning system*''. Remarkably, the *cis*-regulatory element involved in the maintenance of this D/V patterning system, apDV, shows activity exclusively in the wing disc (not active in ''*present-day insect legs*''). Hence, the apDV enhancer could represent a novelty in the evolution of the insect wing.

Moreover, ectopic *vg* expression in the leg disc of *Drosophila* activates the apDV enhancer, but only in the cells that express *ap* (see PUBLICATION II, Figure 5F). This illustrates that the leg cells are basically 'ready' to start an autoregulatory system controlled by *ap* and the wing fate gene *vg*, which is able to coordinate the outgrowth (patterning and growth) of wing tissue (see below).

Altogether, these observations mainly support the idea that the wing originates from ancestral lateral leg branches (exites).

Further analyses of the apE and apDV enhancers

Activating and repressive input is integrated by the apE element

In PUBLICATION II, we have isolated and characterized the apE element in some detail. apE is essential for the initiation of *ap* expression in early 2nd instar wing discs. It is activated by EGFR signaling, but later becomes independent of it. Deleting consensus sites of PntP2 (REBAY 2002)

in apE strongly reduced but did not completely abolish its activity (see PUBLICATION II, Figure 4F). These observations suggest that more input is integrated in apE.

To identify other molecular inputs that control apE activity, a deletion analysis was performed. First, we shortened apE from about 900 bp to 463 bp (apE-1, aka OR463) (see Figure 17A). Afterwards, fragment apE-1 was analyzed with a newly developed program (MotEvo), which uses all annotated transcription factor binding site weight matrices and checks the conservation of the putative binding site (ARNOLD *et al.* 2012). Within apE-1, the putative binding sites are clustered into four distinct regions apE (Figure 17B, 1-4). According to the prevailing transcription factor (TF) type, we defined binding sites for ETS domain and zinc finger TFs (cluster 1 and 4) and homeodomain TFs (cluster 2, 3 and 4). It is well known that homeodomain/selector TFs and TFs from signaling pathways synergize in functional *cis-*regulatory elements (AFFOLTER and MANN 2001). PntP2 binds to DNA via an ETS domain; interestingly, none of the previously identified PntP2 sites falls in one of the defined clusters. Additionally, together with Ken Cadigan (University of Michigan), we found conserved repressive TCF (Wg-signaling input) and corresponding helper sites (CADIGAN 2012; ZHANG *et al.* 2014), which are directly adjacent to the conserved homeodomain clusters 2 and 3.

Based on these *in silico* analyses, the following hypothesis can be proposed; (I) cluster 1 and 4 could represent additional PntP2 binding sites, which positively integrate EGF signaling and activate transcription in the dorsal portion of the developing wing disc; (II) clusters 2 and 3 may represent repressive inputs which suppress *ap* transcription in the ventral part, where Wg is expressed in early discs, via the TCF and helper sites. Thus, according to these predictions, the positive and negative input could be separated on apE.

In order to test this hypothesis, discrete 30-45bp long deletions of the highly conserved clusters were made and the fragments were tested with a *LacZ*-based reporter assay (Figure 17C). Deleting either ETS cluster 1 or 4 had no effect on reporter gene expression in the wing disc. However, and in contrast to our hypothesis deleting both clusters simultaneously led to robust de-repression of apE-1 activity in the ventral compartment. Deleting the TCF and homeodomain sites (apE-1 m2, m3 and m23) resulted in a strong reduction or complete loss of reporter gene expression, respectively.

Strikingly, these results are entirely contradictory to our hypothesis! Nevertheless, they reveal the existence of activating as well as repressive inputs into apE. Hence, further analyses are

required to understand the regulatory logic controlling the expression of this essential, early-acting *ap* wing enhancer.
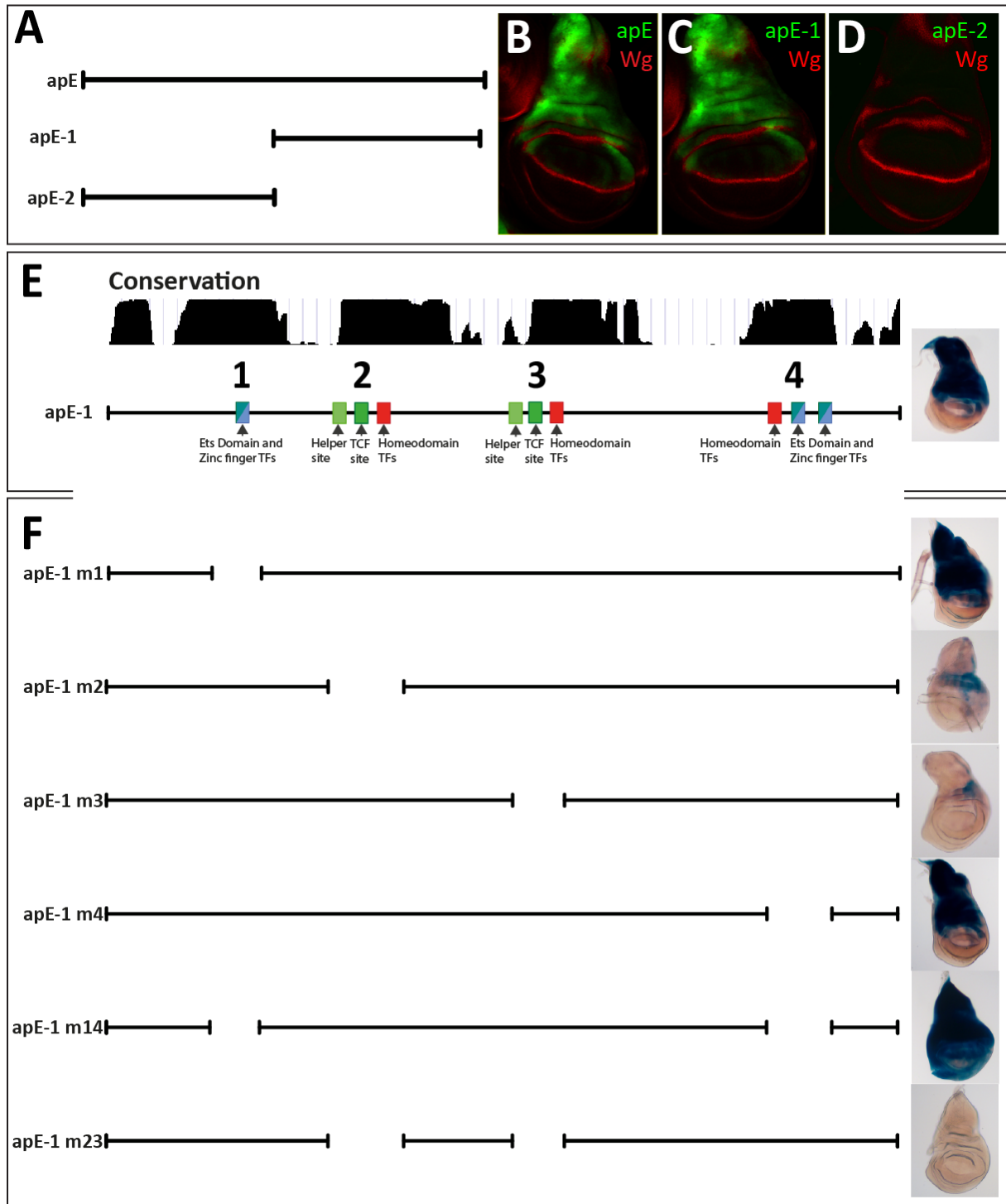
Figure 17 Deletion analysis of the apE element. (A) apE is spilt into apE-1 and apE-2. (B and C) apE and apE-1 show indistinguishable expression pattern. **(D)** apE-2 has no activity in wing imaginal discs. **(E)** Conservation of the apE-1 element and subdivision into four transcription factor (TF) binding site clusters. **(F)** Deletion of individual and combinatorial clusters. apE-1 m1 shows same expression pattern as apE-1. apE-1 m2 and apE-1 m3 have only residual activity in posterior, dorsal hinge region. apE-1 m4 is similar to apE-1 m1. Combination of m1 and m4 (apE-1 m14) results in strong reporter activity in the entire ventral compartment. When m2 and m3 (apE-1 m23) are deleted, no activity is observed anymore. Pictures in B, C, and D kindly provided by Carlos Estella, University of Madrid.

Auto-regulation of *apterous* via apDV

To test whether apE or apDV are regulated in an auto-regulatory mode, we previously generated clones overexpressing dLMO (see PUBLICATION II, Figures 4 and 5). dLMO/Beadex is a well-characterized antagonist of Ap activity (Milán *et al.* 1998; Bejarano *et al.* 2008). Thus, overexpression of dLMO is thought to have the same effect as removing Ap. Clones that overexpress dLMO in the dorsal compartment have been shown to round up and induce Wg expression at their edge (Milan and Cohen 2003), a feature shared with *ap* mutant clones (Diaz-Benjumea and Cohen 1993).

To evaluate an auto-regulatory role more directly, we furthermore analyzed apE-*LacZ* and apDV-*LacZ* reporter constructs in *ap* null mutant backgrounds (Figure 18A-D). While the apE-*LacZ* was still expressed in such *ap* mutant discs, apDV showed no activity. Usually, apDV-*LacZ* shows expression in the dorsal pouch and hinge region in wild type discs (Figure 18B). Since amorphic wing discs lack all pouch and hinge regions, it is reasonable to argue that the cells in which apDV is usually active are not formed, explaining lack of reporter gene detection in these discs.

To rule out this possibility, we made *ap* mutant MARCM clones (see MATERIAL AND METHODS) and analyzed their effect on apDV-*LacZ* (Figure 18E).

MARCM (Mosaic Analysis with a Repressible Cell Marker) is a technique to positively label mutant clones, which facilitates clonal analyses (Lee and Luo 2001). Usually, *ap* mutant clones get expelled from the dorsal compartment, when induced prior to D/V boundary formation. However, if clones are induced after D/V boundary formation, *ap* null-mutant clones in the dorsal compartment do not get expelled, round up and induce Wg expression at their borders (Hamaratoglu and Müller, unpublished observation; see below and MATERIAL AND METHODS).

Additionally, *ap* mutant clones displayed no detectable expression from the apDV-*LacZ* reporter (Figure 18E').

Thus, we confirmed that *ap* expression is directly auto-regulated via the apDV *cis*-regulatory element.
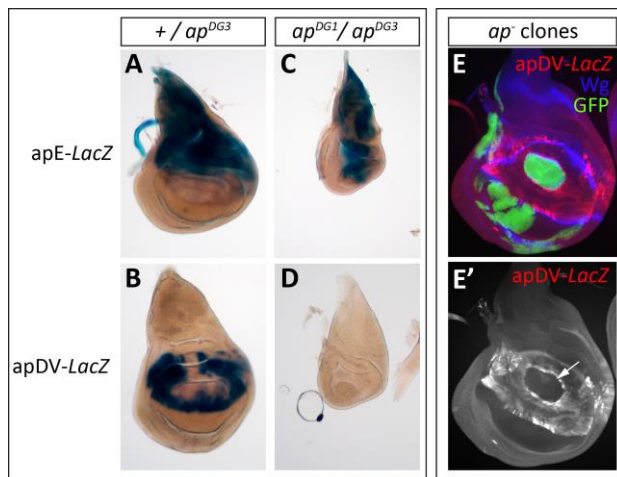
Figure 18 Effect of *ap* null mutants on apE and apDV reporter gene expression. (A and B) Wild type expression patterns of apE and apDV. **(C and D)** Expression patterns of apE and apDV in amorphic *ap* mutants. apE is still expressed, apDV activity is not detectable. **(E)** MACRM of *ap* null mutant (*ap^DG1*, green) and expression of apDV-*LacZ* (red), Wg staining in blue. Clone in the dorsal compartment is round and surrounded with Wg expression. **(E')** Single channel of apDV-*LacZ*. No β-Gal protein is detected in the dorsal ap mutant clone (arrow).

Continuous requirement of apE and apDV during wing disc development?

Previously, we have used an apE+apDV-containing flip-out cassette to demonstrate that these two enhancers must be together in *cis* to the *ap* promoter and PRE throughout wing development (see PUBLICATION II, Figure 7). However, with this experimental setup, it was not possible to tell whether both enhancers were indeed needed throughout wing development. More specifically, we could not test whether the early enhancer apE is also required for proper *ap* expression and wing disc development in late wing discs. To validate the requirement of each enhancer during wing disc development, we used the MARCM technique. The necessary stocks with deletions that remove either both enhancers (*ap^DG1*), only apE (*ap^DG12*) or only apDV (*ap^DG14*) were generated (see MATERIAL AND METHODS). Mutant clones were induced at two different time points; before *ap* gets activated and D/V boundary formation (24-48h after egg laying (AEL)) and after D/V boundary formation has been initiated (48-72 h AEL).

Early clones that lack both enhancers (*ap^DG1*-clones) were eliminated from the dorsal pouch and expelled to the ventral compartment (Figure 19A). Ap and Wg expression patterns in discs with early clones were indistinguishable from wild type wing discs (Figure 19A'-A''). When *ap^DG1*-mutant clones were induced later, they remained in the dorsal compartment, rounded up, were lacking Ap protein and induced Wg at their edges (Figure 19B-B'').

Removing only apE early in development had the same effect as removing both enhancers, as mutant clones were ejected from the dorsal wing pouch (compare Figure 19C to A). In sharp contrast, later *ap^DG12*-clones in the dorsal wing pouch did not round up, Ap levels were not altered and no additional Wg was induced (Figure 19D-D''). Interestingly, clones located at

the border of hinge and notum regions lacked Ap expression, but did not induce Wg (see arrow in Figure 19D').

When the apDV element was removed prior to D/V boundary formation, mutant clones were able to remain in the dorsal compartment. Importantly, they did not round up (Figure 19E). Moreover, these clones showed decreased Ap protein levels (Figure 19E'). Clones that were touching or crossing the D/V boundary resulted in reduced Wg expression at the boundary (arrows in Figure 19E''). Late apDV clones had only minor effects on Ap expression, resulting in a slightly fuzzy D/V boundary (Figure 19F-F'). Also Wg expression, apart from some disruptions in the D/V stripe, seemed to be largely normal (Figure 19F'').
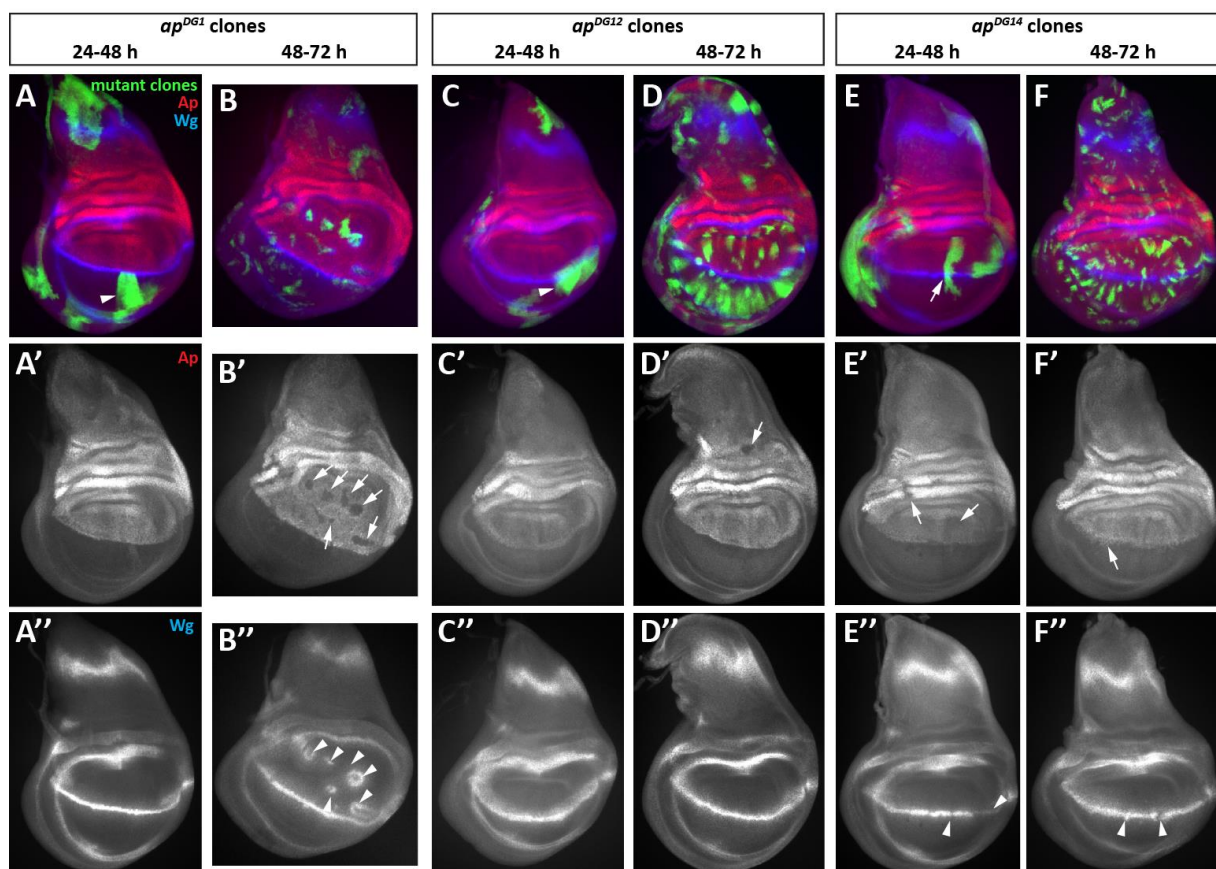
Figure 19 MACRM clones of mutants removing apE and/or apDV *cis*-regulatory elements. (A) ap^{DG1}-clones, deleting apE and apDV, are pushed out of the dorsal wing pouch, when induced 24-48 h AEL (arrowhead). **(A'-A'')** Ap and Wg expression is normal. **(B)** Dorsal ap^{DG1}-clones induced 48-72 h AEL are round, lack Ap protein (arrows in B') and induce Wg surrounding the clones (arrowheads in B''). **(C)** Early ap^{DG12}-clones, removing apE only, are expelled from the dorsal compartment (arrowhead). **(C' and C'')** Ap and Wg expression is indistinguishable to wild type. **(D-D'')** Late apE-missing clones only show reduced Ap expression in the notum (arrow), clones in the wing pouch have no effect on Ap and Wg expression. **(E)** Early ap^{DG14}-clones, that remove only the apDV element, stay in the dorsal compartment and can freely cross the D/V boundary (arrow). **(E')** Ap expression is slightly reduced in clones (arrows). **(E'')** Wg-stripe is disrupted (arrowheads), in places where ap^{DG14}-clones cross the D/V boundary. **(F)** apDV-clones induced 48-72 h AEL stay in the dorsal compartment. **(F')** Ap expression is weakly fuzzy at the D/V compartment boundary (arrow). **(F'')** D/V stripe of Wg shows tiny disruptions (arrowheads).

The fact that early *ap*[DG14]-clones were not expelled can be explained, because these mutant clones still have the early apE enhancer. This enhancer induced *ap* expression, thus cells in *ap*[DG14]-clones are *ap*-positive and probably express specific adhesion molecules which allows them to stay in the dorsal compartment. Only later, and due to the lack of the second enhancer (apDV), they cannot maintain *ap* expression and show perturbations of Wg expression at the compartment boundary. This way, the apE enhancer can also be regarded to confer dorsal identity important for the correct, early expression of various cell adhesion molecules.

The most important result of these MARCM clone analyses is that late apE-missing clones had no detectable effect in the wing pouch. This means that apE must be present in early wing development, but after *ap* induction, it is no longer used for the progression of wing pouch development. The pouch expression of *ap* seems to be exclusively depended on the apDV *cis*-regulatory element. However, apE is still needed for proper *ap* expression in the notum. Based on these results, the model proposed in PUPLICATION II Figure 7 can be refined and expanded (Figure 20).
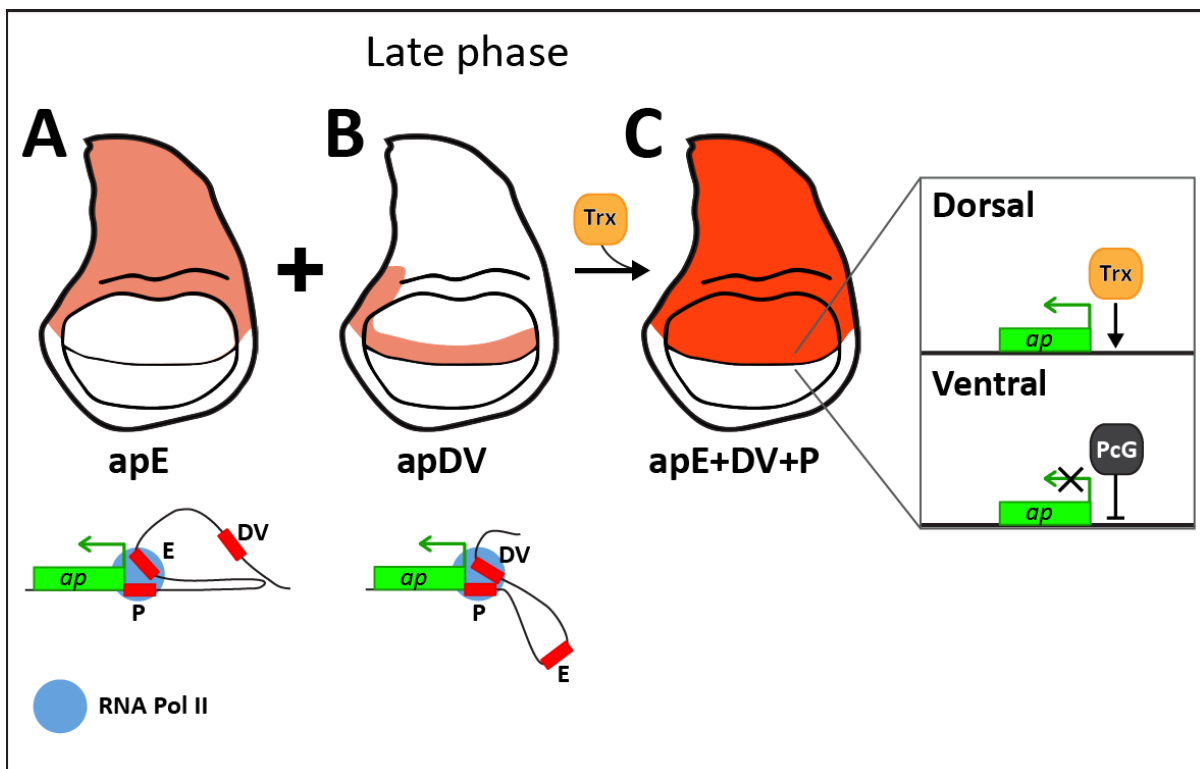
Figure 20 Refined model of *ap* regulation in late wing imaginal discs. (A) Late notum expression of *ap* depends on apE exclusively. **(B)** Late wing pouch and hinge *ap* expression is mainly depended on the apDV element. **(C)** Together with the promoter element (apP), which integrates Trx and PcG input, apE and apDV ensure proper and stable *ap* in the entire dorsal compartment of the wing disc.

Transcriptional autoregulation in development

A key feature of animal development is the exact regulation of gene expression. This can either be a quick change in spatial or temporal expression levels or the maintenance of constant expression and activity levels. Our analysis of the various *ap* mutants revealed that it is crucial to maintain Ap levels during wing disc development, as mis-expression or loss of expression of *ap* lead to severe phenotypes. Thus, *ap* as a selector gene must exhibit stable on-off states. A common mechanism to control gene expression in development is autoregulation (CREWS and PEARSON 2009). Mechanisms of autoregulation are either positive or negative and they can either be direct or indirect.

A nice example of direct positive autoregulation is the *fushi tarazu* (*ftz*) gene in *Drosophila*. Ftz is a homeodomain transcription factor important for embryonic pattern formation. Once Ftz expression is initiated, the Ftz protein binds to a *cis*-regulatory element in the *ftz* gene locus itself. This boosts *ftz* expression, which confers robustness and stability to the *ftz* expression domain (HIROMI and GEHRING 1987; PICK *et al.* 1990; SCHIER and GEHRING 1992). A similar autoregulation has also been observed for the *Distalless* gene (ESTELLA *et al.* 2008). We have demonstrated that during wing disc development *ap* is also directly auto-activated via the apDV *cis*-regulatory element.

This direct positive autoregulation can be further reinforced in a so-called "feed-forward positive autoregulatory" loop. In such cases, a transcription factor maintains its own expression and induces another second transcription factor, which also positively regulates itself as well as the initial transcription factor triggering its expression. The *Myocyte enhancing factor 2/twist* feed-forward loop in the embryonic muscles of Drosophila is a prominent example of this kind of autoregulation (LEE *et al.* 1997; CRIPPS *et al.* 1998, 2004; SANDMANN *et al.* 2007). For *ap*, no evidence has yet been obtained arguing for such a feed-forward autoregulatory circuit.

Furthermore, positive autoregulation can be rather indirect. This is nicely demonstrated in the case of *Ubx* regulation in the visceral mesoderm (THÜRINGER *et al.* 1993). Ubx expression in parasegment 7 (ps7) activates *decapentaplegic* (*dpp*), the BMP2/4 homologue in *Drosophila* (PADGETT *et al.*). Dpp-signaling then induces *wg* in cells of ps8, which then secrete the Wg ligand. In turn, Wg together with Dpp maintain *Ubx* expression in ps7. For the maintenance for *ap* expression in the wing disc, we described a similar situation: Notch signaling initiated by Ap induces *wg* expression along the D/V compartment boundary. Subsequently, Wg

100

signaling activates *vg*, which induces *ap* expression via the apDV element. This feedback loop can be considered as indirect positive autoregulation. Since both, Vg and Ap, are needed to maintain *ap* expression, the apDV *cis*-regulatory element thus integrates direct and indirect positive autoregulation.

In contrast to the given examples, autoregulation can also be negative. *Ubx* has been shown to contain *cis*-regulatory elements, which are directly negatively regulated (IRVINE *et al.* 1993). This negative autoregulation can have different outcomes, ranging from complete repression to fine tuning of expression levels of the autoregulated target gene. In the wing disc, Ap induces *Beadex* (*Bx*) expression. Bx is known to interfere with Ap activity by disruption of the Ap-Chip transcription factor complexes, which also interferes with the direct autoregulation via apDV (MILÁN *et al.* 1998; WEIHE *et al.* 2001; BRONSTEIN *et al.* 2010b). However, *ap* expression is not completely repressed upon *Bx* expression in the wing disc. In this case, negative autoregulation is indirect and can be regarded to stabilize Ap activity levels.

Altogether, the apDV wing enhancer can also be considered as the autoregulatory element. Figure 21 summarizes all the autoregulatory mechanisms observed in the control of *ap* expression during wing disc development.

Generally, transcriptional autoregulation of *trans*-acting factors in important processes have been demonstrated to be conserved from bacteriophages to humans (DODD *et al.* 2005; ODOM *et al.* 2006). Furthermore, it has been proposed that all developmentally important transcription factors show autoregulatory loops that stabilize, fine-tune and/or repress their own transcription. In this regard, autoregulation can be seen as a kind of epigenetic memory of the cell (CREWS and PEARSON 2009; PTASHNE 2013). Given the importance of *ap* in compartmentalization and boundary formation during wing disc development, it is certainly no surprise that such autoregulatory mechanisms are also found to be at work in the regulation of *ap*.
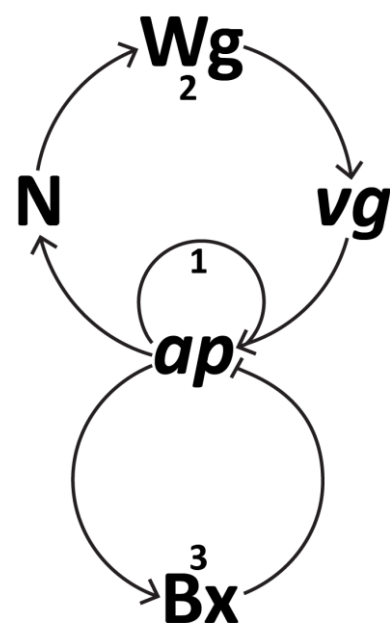
Figure 21 Autoregulation of apterous. (1) Direct positive autoregulation. **(2)** Indirect positive autoregulation. **(3)** Indirect negative autoregulation.

Why do all the *cis*-regulatory elements required for wing development have to be on the same chromosome?

Using genetic complementation experiments, we have demonstrated that only when the three *cis*-regulatory elements apP, apE and apDV are on the same chromosome (*in cis*), *ap* expression is normal and a wild type wing develops (see PUBLICATION II, Figure S8). However, in the situation where apP (PRE) and apE are on one, and apP and apDV on the other chromosome (*in trans*), *ap* expression is perturbed resulting in severe malformations in the adult wing (e.g. ap^{DG14}/ap^{DG12}).

One possible explanation for these observations is that the *in cis* configuration of the three *cis*-regulatory elements facilitates and stabilizes enhancer-promoter looping. Furthermore, it is possible that the *in cis* arrangement helps to rapidly re-establish chromatin contacts after each round of cell division. This is in accordance with previous observations, in which constant interactions between *ap* enhancers and promoter have been described during embryogenesis (GHAVI-HELM *et al.* 2014).

In another scenario trying to explain this phenomena, the PRE at the promoter plays a much more active role. It is possible that transcriptional repression represents the ground state of genes containing a PRE (MÜLLER and KASSIS 2006). Thus, prior to induction, *ap* expression is actively repressed by the PRE present on both chromosomes. This statement is corroborated by the fact that ChIP-data from whole wing discs show repressive histone marks, which span the entire *ap* locus (modencode.org). When *ap* is induced by EGFR-signaling via the apE element, the repressive input of the Polycomb group proteins present at the PRE may be replaced by the chromatin-activating input of the Trithorax (Trx) group proteins. In PUBLICATION II, we have demonstrated that Trx is actively involved in the transcriptional regulation of *ap*. The activity of Trx group proteins might then open up the repressed chromatin (also seen by respective histone mark data available on modencode.org), allowing the binding of additional transcription factors to *cis*-regulatory elements, most notably to the (autoregulatory) apDV enhancer. In the case where the *cis*-regulatory elements are *in trans*, apE on one chromosome activates Trx proteins primarily on 'its' PRE *in cis*. The PRE *in trans* possibly sees no activating signal, as it does not contain the apE element, thus the repressive Polycomb group proteins on this PRE prevail and keep the chromosome 'off'. This obstructs DNA-protein interactions on the apDV, which usually maintains *ap* expression in the progression of wing disc development.

This second model might be very difficult to test by biochemical means such as Chromatin Immunoprecipitation (ChIP), because it is -at the moment- not possible to distinguish between differing chromatin states of the homologous chromosomes. However, the hypothesis can be investigated by genetic experiments. For this, *ap* controlled only by the apDV element must be uncoupled from PRE regulation. The easiest way to achieve this is via transgenics, which contain an apDV-*ap*cDNA fusion outside the endogenous *ap* locus. This transgene could then be tested in hemizygous *ap* alleles which lack the apDV enhancer (e.g. *ap*^{DG12}, *ap*^{C2}). If the hypothesis holds true, wing development should be normal under these circumstances. We are currently working on experiments along these lines.

How does *apterous* instruct growth?

While the genetic regulation of pattern formation in the wing disc is fairly well understood and described, the understanding of proliferation/growth control is more rudimentary and under debate. Many models and factors are currently discussed with respect to their contribution to growth regulation, and many results are surprising and challenging current models (HARIHARAN 2015). For example, while the requirement of Dpp for growth in the wing disc is undisputed, its mode of action remains unclear (AFFOLTER and BASLER 2007; WARTLICK *et al.* 2011; RESTREPO *et al.* 2014). The same is true for Wg, whose role as a classical morphogen is seriously questioned (ALEXANDRE *et al.* 2014). Additionally, more global pathways, like the mTOR, Hippo and Insulin signaling pathways, have been implicated in the regulation of growth of the wing disc tissue (HALDER and JOHNSON 2011; PARISI *et al.* 2011).

Since *ap* null mutants lack all wing blade/pouch and hinge tissue in imaginal discs and in the adult, the question of how *ap* is involved in the growth control of the wing imaginal disc can and must be raised.

One first obvious observation made, is that the notum in *ap* mutants forms rather normally in terms of size and patterning, even though *ap* expression is seen throughout the dorsal compartment, including the most dorsal region, from which the entire notum forms. Thus, not all cells that usually express *ap* are affected in *ap* null mutants. This observations also most likely excludes a direct instructive involvement of Ap in wing disc growth. Moreover, it divides the 'wing disc' into at least two different regions; an *ap*-independent body wall/notum and an *ap*-dependent appendage/wing region. The formation and growth of the notum structures

has been demonstrated to rely on the EGFR pathway and its ligand Vein throughout wing disc development (SIMCOX *et al.* 1996; WANG *et al.* 2000; ZECCA and STRUHL 2002a).

But how is growth regulated in the *ap*-dependent wing region? As already mentioned several times, only the early *ap* expression relies on EGFR signaling. Thus, later *ap* expression and wing development is EGFR-independent and is maintained by autoregulation (see above). After its induction, Ap starts a Notch signaling cascade across the D/V boundary (see INTRODUCTION). Notch signaling in the early wing disc has been shown to induce growth and ensure cell survival (RAFEL and MILÁN 2008). Moreover, this Notch-signaling cascade across the D/V boundary directly induces *vg*, via the vg-Boundary-Enhancer (vgBE). *vg* is considered to be the 'master control gene' for wing development, since it can ectopically induce wing fate (KIM *et al.* 1996). However, growth of this transdetermined tissue is only induced when Dpp and Wg are both present at the same time (MAVES and SCHUBIGER 1998). In terms of growth control, Vg has been implied to be rather permissive, as *vg* mutants completely lack this tissue. But, Zecca and Struhl have proposed a model in which Vg is more instructive for the growth of the wing pouch (ZECCA and STRUHL 2007a; b). Furthermore, positive autoregulation maintains *vg* expression via the vg-Quadrant-Enhancer (vgQE) and was shown to recruit more cells into the wing pouch primordium (ZECCA and STRUHL 2007a). Interestingly, *ap*, *Ser* and *Su(H)*(Notch signaling), and *vg* mutants all show very similar malformed wing discs, where the inner Wg ring, usually surrounding the pouch, is shrunk to a dot (KLEIN and ARIAS 1998; ZECCA and STRUHL 2007a), underlining the epistasis of these factors in this aspect of wing development. Simultaneously, the Notch signaling cascade at the D/V boundary induces Wg, which in turn also activates and maintains *vg* expression (NEUMANN and COHEN 1997).

But which role does Dpp play in this model? In early second instar, Dpp secreted from the peripodial membrane (PE) of the wing disc has been suggested to activate *vein* (*vn*) in the most dorsal cells (PAUL *et al.* 2013). Subsequently, Vn triggers EGFR signaling which is needed to activate *ap* expression (see above). *Dpp* expression at the A/P compartment boundary in the disc proper (DP) only starts at mid-second instar (PAUL *et al.* 2013). From there, the Dpp morphogen forms a concentration gradient, which is crucial for the patterning and growth of the wing disc (AFFOLTER and BASLER 2007). Via the vgQE element, Dpp signaling cooperates with the activation and maintenance of *vg* expression (KIM *et al.* 1996).
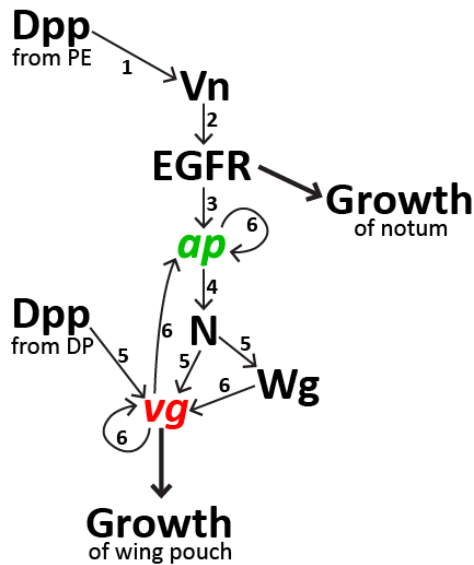
Figure 22 summarizes all the described regulatory interactions that are essential for *Drosophila* wing patterning and growth. Altogether, *vg* is suggested to play a central role in growth control of the wing pouch and wing fate commitment. The crucial role of *ap* in this cascade is the correct activation of the Notch signaling, which subsequently results in *vg* expression.

Figure 22 Summary of the gene network involved in the development and growth of the wing imaginal disc. Numbers indicate the relative time in development.

Isolation and characterization of additional deletions in the apDV region

The first imprecise excision screen we performed in *ap* was done with a P-element insertion 400 bp upstream of the *ap-RA/RC* TSS (apMM; see Material and Methods of PUBLICATION II). Deletions affecting only the upstream region of *ap* either compromised none or both of the defined wing enhancers (apE and apDV; data not shown). In order to isolate deletions that only affect the apDV element, the most distal region of the *ap* locus was scanned for available and suitable transposon insertions. One possible candidate would have been the PiggyBac insertion e01573, whose FRT site was used to generate some deletions, such as *ap^{DG3}* or *ap^{DG14}*. Although PiggyBac insertions were also shown to cause imprecise excisions, the resulting deletions were rather small (KIM *et al.* 2012), which did not fit our purposes.

Thus, we decided to use a P-element insertion in the flanking gene *l(2)09851* (EY06365, Figure 23A). This gene is largely uncharacterized, but mutations affecting *l(2)09851* were shown to be lethal (SPRADLING *et al.* 1999). To isolate excisions that leave the ORF of *l(2)09851* intact but remove distal *ap* DNA, we applied transposase to EY06365 *in trans* to *ap^{DG3}*. As in *ap^{DG3}* the *l(2)09851* gene is unaffected, it may serve as a template for the DNA repair after an imprecise excision event. Importantly, *ap^{DG3}* is missing almost the entire *ap* locus, thus we reasoned that imprecise excision events would only cause deletions in this region (Figure 23B).

Indeed, we isolated four different deletions that showed wild type *l(2)09851* sequences and had the same distal breakpoint as *ap^{DG3}* (Figure 23C). Interestingly, all deletions show the same signature at the *ap^{DG3}* breakpoint; they contain the entire 3' PiggyBac end of e01573 to the

FRT site; the P-element from EY06365 is inserted in the opposite direction of the original insert adjacent to the intact FRT site of e01573; and they contain a 400 bp insert of the retrotransposable springer long terminal repeat (LTR) element proximal to apDV (Figure 23C). The springer LTR was shown to be already present in the original EY06365 stock (data not shown).

Two of the deletions removed half of the apDV element ($ap^{11.1}$ and $ap^{34.1}$), which resulted in notching of the wing margin in the adult fly wing (Figure 23D). Deletions not compromising the integrity of apDV exhibited wild type wings ($ap^{72.2}$ and $ap^{62.3}$).

These results again confirm the location of apDV and its importance in maintaining the D/V compartment boundary in the developing wing.
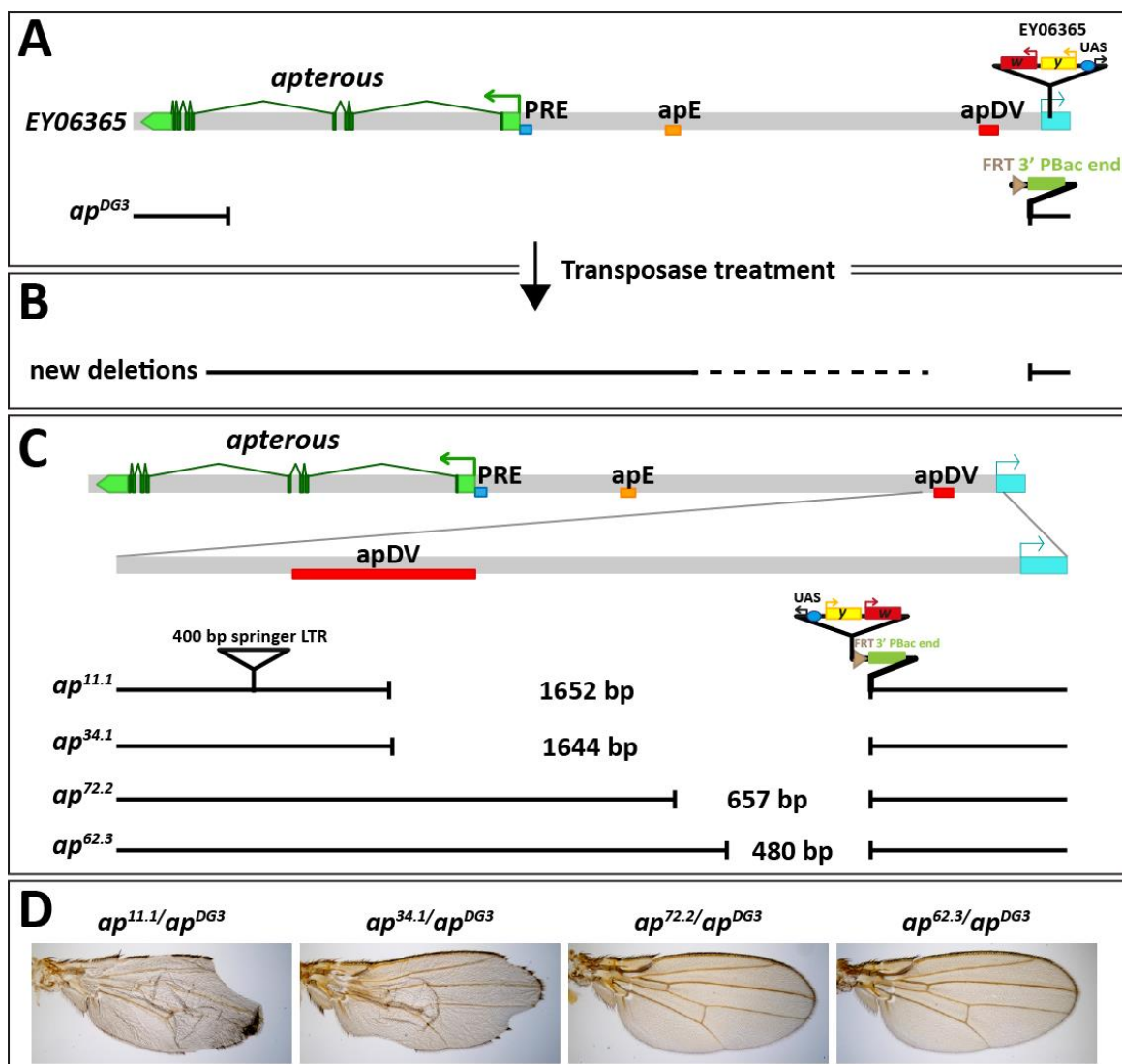
Figure 23 Setup to generate deletions affecting the apDV element. (A) EY06365 containing a P-element in *l(2)09851* is crossed to ap^{DG3}. Offspring were treated with transposase to cause imprecise excisions of EY06365, resulting in new deletions **(B). (C)** Deletions isolated had P-element insertions proximal to 3' PiggyBac ends and long terminal repeat insertion (LTR). Note: all the isolated deletions showed these features, but it's only depicted for $ap^{11.1}$. **(D)** Wing phenotypes caused by deletions. $ap^{11.1}$ and $ap^{34.1}$ display severe notching and blistering of the wing blade, while $ap^{72.2}$ and $ap^{62.3}$ do not cause any observable phenotypes.

Outlook

During the course of my PhD thesis, we have deciphered the *cis*-regulatory logic involved in the regulation of *ap* expression during wing disc development. Our final model proposes a three-step mechanism that generates a lineage compartment through the integration of input from separate *cis*-regulatory elements important for the initiation, refinement and maintenance of *ap* expression.

Although we have uncovered a number of important inputs into these regulatory elements, the regulation of the apE enhancer is still not completely understood. A first attempt to investigate apE in more detail was done with a candidate-based approach and has revealed repressive and activating inputs. These preliminary results show that there is still a lot to be learned about how transcription is actively repressed and how activating and repressive inputs are integrated to control gene expression. For example, it is still not clear which transcription factors, besides PntP2, are binding to this element and are regulating the initiation of *ap* expression in wing disc development. To find new and possibly unknown factors, apE regulation should be studied in an unbiased manner, preferably at the endogenous locus. For this, random mutation could be introduced into the sequence of the apE-1 element and, together with the C5 fragment, inserted into the *ap* locus with the *in situ* rescue system. If the random mutation affects an important transcription factor binding site, then the wing might show an obvious phenotype easy to recognize on adult flies. To increase the transcription factor site resolution, transgenic animals, with or without phenotypes, can then be sequenced. This approach does not directly reveal the respective trans-acting factors binding to this element. Therefore, one should also consider to use new, unbiased biochemical approaches to better study the regulation of apE. A promising method for this purpose is the 'proteomics of isolated chromatin segments (PICh)' approach (DÉJARDIN and KINGSTON 2009). In this approach, defined DNA stretches, together with the chromatin and the bound transcription factors, are purified with the help of a complementing, DNA-derivate locked nucleic acid (LNA). The purified complexes can then be analyzed by mass-spectrometry. Potential results obtained when applying this method to the *ap* regulatory elements would probably give more insights into compartment boundary formation, basic principles of gene activation and repression, and also in the evolution of insect wings.

Moreover, it will be interesting to study the exact role of the *ap* PRE in processes such as transvection, chromosome pairing and cellular memory formation. Again, the developed and

validated *in situ* rescue system will be extremely helpful in this regard, since tailor-made deletions and insertions can be generated.

To better understand the growth of the wing disc, it is absolutely essential to (re-)validate the relationship between Dpp, Ap, Wg, Vg and other known players. Interesting questions are; how does *ap* expression in various *Dpp* mutants look like? Or how does *ap* influence *Dpp* expression? Additionally, dissection of the *cis*-regulatory network of the factors involved in wing disc development will be necessary. While for *vg* and now *ap* detailed analyses of their *cis*-regulatory elements were done, such data are missing for *dpp* and *wg*.

MATERIAL AND METHODS

Preparation of electro-competent *E. coli* bacteria

Solutions:
- LB without salt (For 1l: 10g tryptone, 5g yeast extract add ca. 980mL water, autoclave)
- 10% Glycerol, sterile

Protocol:

A 50 ml culture (LB without salt) was inoculated with desired bacteria strain (e.g. TOP10, Invitrogen) and incubated overnight at 37°C. As TOP10 has a Streptomycin-resistance cassette, Streptomycin (50 µg/ml) was added to the overnight culture. The next morning, the culture 1:100 or 1:50 was diluted in 1l of LB without salt (NO antibiotics) in 2x 3L Erlenmeyer flasks with baffles and grown until OD_{600}= 0.6-0.8. Subsequently, the cultures were immediately cooled in an ice slurry, and never warmed up from then on. It's important to keep the bacteria cold, thus all the following steps are performed on ice. The 1l culture was distributed to 4 collection bottles for the SLA-3000 rotor and centrifuged for 10 min at 3300 rpm at 4°C. After centrifugation, it was normal that the supernatant was still a bit turbid. The supernatant was discarded and the each pellet re-suspended pellet in 5 ml cool 10% glycerol. Then the suspension was distributed to 2ml Eppendorf tubes and centrifuged at 8000 rpm for 5 min in a cooled benchtop centrifuge. Afterwards, the supernatant was carefully pipetted off and the pellets were re-suspended in 1 ml 10% glycerol again (now the culture was 100x concentrated). The suspension was distributed as 50 µl aliquots and flash-frozen in liquid nitrogen. The electro-competent bacteria can be stored in a -80°C freezer.

Isolation of genomic DNA from adult flies

Material:
- Mortar and pestle
- Dounce homogenizer with A pestle

Solutions:
- Homogenization buffer (for 20 ml: 200 µl 1M Tris-HCl (pH 7.5), 240 µl 5M NaCl, 400 µl 500 mM EDTA (pH 8.0), 83.3 µl 300 mM tetra-HCl with spermine, 25 µl 1 M tri-HCl with spermidine)

- Chloroform/*iso*-amyl alcohol mixture (480 µl chloroform plus 20 µl *iso*-amyl alcohol)

Procedure:

About 200 adult flies were collected and left in a clean bottle for two hours to minimize ingested yeast. Afterwards, the flies were ground with a pestle in liquid nitrogen in a precooled mortar. The fly powder was transferred to a chilled Dounce homogenizer containing 5 ml homogenizing buffer and treated with a few strokes of the A pestle to free the nuclei from the carcasses. To remove debris, this mixture was centrifuged at 1000 rpm for 1 min. The supernatant was transferred to a fresh tube (15 ml falcon tube) and centrifuged at 4000 rpm for 8 min at 4°C. The pellet was resupended in 0.5 ml homogenization buffer and transferred to a 1.5 ml Eppendorf tube and re-suspended. Subsequently, proteinase K (to a final concentration of 100 µg/ml) and 50 µl 10 % SDS were added to the suspension. The solution was gently mixed by inverting the tube several times and incubated at 37°C for one hour. Then, the lysate was extracted twice with 0.5 ml phenol and once with 0.5 ml chloroform/*iso*-amyl alcohol. The supernatant was transferred to a round-bottom tube (Falcon) and NaCl was added to a final concentration of 200 mM. Then 2 volumes of 100 % ethanol were added and the mixture was mixed gently by swirling. The DNA appeared at the interface as a clump. When the phases were completely mixed the clump of DNA was pulled out with a glass hook, rinsed in 70% ethanol and suspended in ddH2O. The ethanol precipitation was repeated once and the pellet was air-dried to remove all the ethanol. Finally, the dry pellet was dissolved in distilled H_2O and the DNA concentration was determined with the NanoDrop spectrophotometer (Thermo Fischer Scientific). The DNA was stored at -20°C.

Preparation of DNA from single flies

Material:
- Heating block

Solutions:
- Squishing buffer (for 1ml: 10 µl 1M Tris-HCl (pH 7.5), 5 µl 500 mM EDTA (pH 8.0), 5 µl 5 M NaCl, 200 µg proteinase K, 980 µl H_2O)

Procedure:

To isolate DNA from individual flies, one fly was placed in a 1.5 ml Eppendorf tube and mashed for 5 - 10 sec with a pipette tip. Subsequently, the carcass was re-suspended in 50 µl squishing buffer. The sample was incubated at 37°C for 30 min. The sample was heated to 95°C for 2

min to inactivate the proteinase K. Afterwards, the DNA sample could be stored at 4°C for at least one month. 1 µl of the sample was used as a template for PCR.

Polymerase Chain Reaction (PCR)

Standard protocol:
PCR reaction mix:
10 µl 5x Phusion® HF Reaction Buffer (Finnzymes)/ 5µl 10x ThermoPol Reaction Buffer (NEB)
1 µl 10mM dNTP mix
1 µl 20µM forward primer
1 µl 20µM reverse primer
x µl template (gDNA), usually 50 ng
0.5 µl Phusion® Polymerase (Finnzymes)/ Taq DNA Polymerase (NEB)
H_2O to 50µl

Standard PCR program:
98°C 120sec
98°C 20sec
56°C 20sec   35x
72°C time alters with product length
72°C 10min
4°C ∞

The PCR reactions were carried out in a T3 Thermocycler (Biometra). For Phusion® Polymerase the elongation time was 30 sec per 1000 bp fragment size. For Taq DNA Polymerase the elongation time was 1 min per 1000 bp.

If unspecific, spurious PCR products appeared on the agarose gel the annealing temperature was increased by one or two °C. Alternatively, the following 'Touchdown' PCR program was used (DON et al. 1991):

98°C 3 min
98°C 20 sec
66°C 20 sec 10x/-1°C each cycle
72°C alters with product length
98°C 20 sec
56°C 20 sec 35x
72°C alters with product length
72°C 10 min
4°C ∞

The PCR products were purified directly or from Agoarose gels using the NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel) according to the manufacturer's protocol.

To clone a putative enhancer, fragments were amplified from genomic DNA with PCR. Primers were designed in the way that about 20 nucleotides (nt) were complementary to the template, the GC content was between 40 % and 60 %, and when necessary restrictions site were introduced at the 5' end. To increase cleavage efficiency of these restriction sites, the primers were extended by four additional nt, usually ATTA (SAMBROOK and RUSSELL 2006a). The oligo-nucleotides were synthesised by Sigma-Aldrich Corp. (United Kingdom).

Restriction digestion

The PCR products and the respective plasmid were usually digested at 37°C for 2 h in a volume of 30 µl with the appropriate restriction enzymes and corresponding buffers (NEB or Roche). If a sequential restriction digestion was necessary, the fragment was cut with the first enzyme for two hours and purified with the NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel). The clean fragment was then digested with the second enzyme. If the DNA was only cut with one enzyme, dephosphorylation of the vector backbone was necessary and achieved by using Shrimp Alkaline Phosphatase (Promega). Afterwards the digested DNA was separated by agarose gel electrophoresis (SAMBROOK and RUSSELL 2006b), purified with the NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel), and the DNA concentration was determined by the NanoDrop spectrophotometer (Thermo Fischer Scientific).

Ligation

The digested inserts and the appropriate plasmid were ligated in a volume of 15 µl in the presence of T4 ligase and the respective buffer (NEB). The insert/plasmid ratio was calculated according to the equation mentioned below, generally 50 ng of plasmid was used. The ligation was incubated at 18°C overnight or at room temperature (RT) for two hours (SAMBROOK and RUSSELL 2006c).

$$insert\ mass\ in\ ng = 6 \times \left[\frac{insert\ lenght\ in\ bp}{plasmid\ length\ in\ bp}\right] \times plasmid\ mass\ in\ ng$$

Transformation

Material:

- Electro-competent E.coli
- Gene Pulser (Biorad)

Solution:

- Luria-Bertani (LB) medium (10g Bacto-tryptone; 5g yeast extract; 10g NaCl in 1l H_2O, adjust to pH 7.5, autoclave)
- LB Agar plates (add 15 g Agar to 1l of LB, autoclave, add appropriate antibiotics, pour hot liquid LB Agar into plates (20ml in 10cm diameter plates) and let the Agar harden by cooling down to RT)

Procedure:

Electro-competent *E. coli* bacteria were used for transformation. To increase transformation and avoid arcing, the salt concentration was reduced by dialysing the ligations against H_2O on a 0.025 µm membrane filter paper (Millipore®) for 15 min. In the meantime, frozen electro-competent *E. coli* (50 µl in Eppendorf tubes) from the -80°C freezer were thawed on ice and 1-3 µl of the DNA ligation was added. Then the bacteria were transferred to a 1 mm Gene Pulser cuvette (Biorad). The electro-transformation was performed with a Gen Pulser (Biorad) at 1.8 mV, 200 Ω and 25 µF. After transformation, the bacteria were propagated in 1 ml LB medium without antibiotics on a shaker at 37°C for one hour. Subsequently, the bacteria were spread on LB Agar plates supplemented with appropriate antibiotics and grown over night in a 37°C incubator.

Minipreps

Single colonies from the LB plate were picked and grown in 2-3 ml LB medium with respective antibiotics at 37°C overnight on a shaker. Plasmid DNA was isolated with the NucleoSpin® Plasmid from Macherey-Nagel following the instruction manual for the manufacturer. Positive candidates were selected with appropriate restriction digests followed by agarose gel electrophoresis (Sᴀᴍʙʀᴏᴏᴋ and Rᴜssᴇʟʟ 2006d).

Sequencing

The identity of positive miniprep candidates and PCR products were sequenced. Therefore 0.8 µg of the plasmid were mixed with the appropriate primer in a final volume of 15 µl.

Sequencing was performed by Microsnyth AG (Balgach, Switzerland). Sequences were analyzed with the ApE© or SnapGene DNA software.

Midiprep

If large amounts of plasmid or highly purified plasmids were needed for downstream applications, such as injection into *Drosophila or fish* embryos, a midi prep was done.

For this, a positive clone was picked and grown in 100 ml LB medium with the respective antibiotics at 37°C on shaker overnight. Injection grade plasmid DNA was isolated using a NucleoBond® Xtra Midi EF (Macherey-Nagel) from. After the isolation, the DNA concentration was determined with the NanoDrop spectrophotometer (Thermo Fischer Scientific) and the plasmids were subsequently stored at -20°C.

Generation of transgenic flies

P-element transgenesis

Stable integration of foreign DNA into the germline of an organism is tremendously important to carry out modern genetic research. In *Drosophila*, the conversion of the P-element transposon into a transgenic vector was the starting point into a new era of genetic manipulation (SPRADLING and RUBIN 1982). Transposons are mobile genetic elements which are capable to integrate into the genome of their hosts. The functional *Drosophila* P-element consists of terminal repeats (the 5' and 3' P-element ends) and between these ends it contains a gene encoding for a transposase protein. The transposase binds to the terminal ends and catalyzes the transposition of the element (RIO 1990). For transgenesis, a P-element vector was engineered that contains the P-element ends, a selectable marker and a multiple cloning site for the integration of any desired DNA (RUBIN and SPRADLING 1983). However, these plasmids are unable to integrate into the genome on their own, since they lack the transposase. Thus, transposase is supplied by the 'helper plasmid'. It lacks P-element ends and is thus unable to integrate into the genome. A mixture of modified P-element and helper plasmid is injected into the posterior end of freshly laid *Drosophila* embryos, where the germ cells will soon form. P-elements integrate rather randomly into the genome. Transformed animals are selected by means of the selectable marker present on the modified P-element.

The marker gene is expressed in adult flies were it causes an easily recognizable phenotypes (PIRROTTA 1988). Frequently used marker genes are *white* or *yellow*.

Additionally, stably integrated P-elements can be excised again by treatment with transposase. Sometimes these excision events can cause deletions close to the insertion site. These so-called imprecise excision events can be used to generate deletions of a given DNA segment in the genome (VOELKER *et al.* 1984).

For this thesis, a new P-element was cloned containing the *ap* promoter (approm) fused to the *LacZ* reporter gene and a GFP marker cassette (plasmid name: pPapprom5'UTRLacZ-3xP3GFP). Starting vector was the pBSmin-mame2 (obtained from M. Metzler), which contains a multiple cloning site (MCS) between the minimal P-element ends. The approm-LacZ was cut from the pAttBlacZ approm 5'UTR plasmid and cloned via SmaI and NotI sites into the pBSmin-mame2. Afterwards, the 3xP3-GFP, which drives GFP expression the *Drosophila* eyes and larval brain, was inserted via the EcoRI site (see APPENDIX).

The attB/attP system

To investigate the expression pattern of putative enhancers, germ-line transformation of P-elements is a very powerful and widely used tool. The integration of such P-elements into the genome is random. This property may be advantageous for many applications, such as insertional mutagenesis or when combined with the Gal4/UAS system (BRAND and PERRIMON 1993). But for these random P-element insertions, position effects acting on reporter gene expression can never be excluded and is considered as a major drawback when used for comparative enhancer studies, as performed in this study. To overcome this problem and achieve targeted integration, a new method (the φC31 integrase system) was used (BISCHOF *et al.* 2007).

This integration system is based on the site-specific integrase of the bacteriophage φC31 (THORPE and SMITH 1998), which has been adopted to use in *D. melanogaster* (GROTH *et al.* 2004). The integrase catalyzes the recombination between the bacterial attachment site (attB) and the non-identical phage attachment site (attP), resulting in attL and attR sites flanking the recombination event (Figure MM1). Since these 'new' sites differ from the original sites, the integrase is unable to recognize these sequences, which makes an excision event impossible, resulting in a stable recombination product.
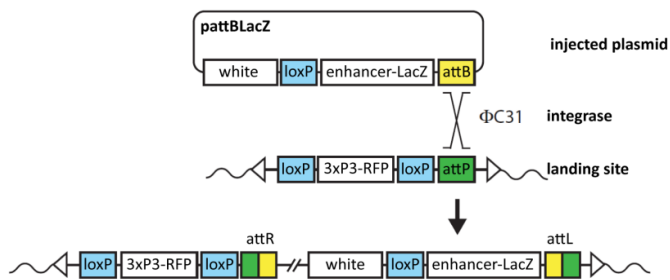
Figure MM1 Integration at attachment sites mediated by φC31 integrase. A plasmid containing the enhancer of interest fused to the LacZ reporter, a *mini-white* gene and an attB site. the attB site can recombine with the attP site in the genome when φC31 integrase is abundant and catalyzes the reaction. The marker genes, e.g. *white*, can be removed applying Cre recombinase. Figure modified from Bischof et al., 2007.

Successful production of transgenics with the φC31 system requires two components. First, a stable source of φC31 integrase must be available. This was achieved by combining the integrase gene with the regulatory elements of the *vasa* gene (BISCHOF *et al.* 2007). This way, the integrase is produced specifically in the germ line. Second, an attP landing site is required somewhere in the genome. To generate transgenic flies, a plasmid containing an attB site and a selectable marker gene, e.g *white*, is injected into embryos which contain the integrase source and an attP landing site.

For this study, flies homozygous for the attP landing site on the third chromosome (zh-86Fb) and the integrase on the fourth chromosome were used. In the Affolter-Lab, this strain is referred to as J5. This landing site was originally selected for the generation of transgenic animals for three reasons. (I) the landing site was reported to be in an intergenic region. (II) It experiences no position effect (tested with empty vectors) and (III) it allows efficient integration. It is noteworthy to mention that in the most recent release of the annotated fly genome (Release R5.34), this insertion site turns out to be located in the second intron of the gene *Chloride channel-a* (*ClC-a*).

Additionally for this thesis, flies were engineered to harbor an attP in the promoter of *ap* (see BIELI *et al.* 2015; PUBLICATION I and II).

Injection

Appropriate fly lines were set up in cages (about 100 flies per cage) covered with yeasted grape juice agar plates. After about 5 days, the flies began to the lay sufficient amounts of eggs for injection.

Early embryos were collected in 20-30 min intervals. Eggs were dechorionized in 3.5% sodium hypochloride solution for 2 min. After the dechorionation, the eggs were transferred into a

container equipped with a fine-mashed net and extensively washed with water. The embryos were transferred to a block of LB Agar (without antibiotics) and aligned with a coarse needle under a binocular. Next, the aligned embryos were picked up with a cover slip with a streak of ''embryo glue'' on it. ''Embryo glue'' had been obtained by dissolving the glue of an adhesive tape with heptane. The embryos were then dried with a hair dryer without heating for 4.5 min and covered with Voltalef PCTFE oil (Atofina), to avoid further desiccation. Now the embryos were ready to be injected with plasmid solution (0.3 µg/µl plasmid in 5 mM KCl, 100µM Na$_2$HPO$_4$ and 100µM NaH$_2$PO$_4$) at their posterior end.

Injected embryos immersed in oil were kept at 18°C for two days. Surviving first instar larvae were transferred to a tube containing fly food and bred at 25°C. Adult injectees were crossed with y⁻w⁻-flies and the offspring was screened for positive transformants (e.g. GFP⁺, white⁺, or yellow⁺). Independent stocks were obtained by crossing single transformants with conventional second or third chromosome balancers (LINDSLEY and ZIMM 1992).

DIG-labeling of anti-sense *ap* RNA

A 1.5 kb fragment from the 3' end of the *ap* cDNA was amplified from the cDNA clone HL02012 (purchased from DGRC) with respective primers (insitu probe_SacI_for and insitu probe_long_KpnI_rev). The fragment was cloned between SacI and KpnI sites of pBluescript II KS(+) vector (name of the plasmid: pBSIIKS ap 3'UTR and CDS long). Then, the resulting plasmid was linearized with Acc65I and digoxigenin-(DIG)-labeled RNA was produced from T7 promoter according to the manufacturer's protocol (Roche, Switzerland). To produce sense probe for the negative control, the plasmid was cut with SacI and RNA was produced from the SP6 promoter.

Expression analysis in fly

RNA-*in situ* hybridization in imaginal discs

Material:
- Anti-Digoxigenin-AP, Fab fragments (Roche; Cat. No. 11093274910)

Solutions:
- 10x phosphate buffered saline (PBS) (for 1l: 2g KH$_2$PO$_4$, 80g NaCl, 2 g KCl, 6.1 g Na$_2$HPO$_4$, 1.25 ml 10N NaCl and fill with H$_2$O to 1l, autoclave)
- PBT (for 1l: 890ml H2O, 100 ml 10x PBS and 10 ml Tween-20)

- Fix solution: 4% paraformaldehyde (PFA) in PBT
- PBTN (PBT plus 2% normal donkey serum)
- RNA Hybridization solution (50% formamide, 5x SSC, 0.1% Tween-20, 100µg/ml heparin, 100µg/ml fish sperm DNA (Roche, Switzerland)
- Probe solution: 0.5ng/µl/kb of probe in 50µl of hybridization solution, heated to 80°C for 3 min and cooled on ice for 5 min.
- AP Buffer (100mM NaCl, 50mM $MgCl_2$, 100mM Tris pH 9.5, 0,2% Tween-20)
- Staining solution (20µl NBT/BCIP stock solution (Roche, Switzerland), fill up with AP buffer to 1ml)

Procedure:

Larvae were dissected in cold 1x PBS by cutting into half and inverting the anterior part. Subsequently, dissected larvae were incubated in 600 µl Fix solution for 25 min on a slow rocking shaker at RT. The fixed tissue was washed 4 times for 5 min with PBT. Washing buffer was removed and replaced by 1:1 PBT:RNA hybridization solution, in which the samples were incubated for 10 min. Afterwards, the solution was changed to 100% hybridization solution and the samples were rocked at 55°C. After on hour, hybridization solution was renewed and incubated for another hour before probe solution was applied. The sample with the probe solution was incubated at 55°C for 20-24 h. The day after, the samples were rinsed once quickly and twice for 30 min with pre-warmed hybridization solution. After cooling the samples to RT, they were washed in 1:1 PBT:Hybridization solution for 10min. Next, the carcasses were washed once quickly and three times for 10 min with PBT. Then, the samples were incubated in antibody/PBTN solution (1:250, Anti-Digoxigenin-AP) for 2 h. Afterwards, it was washed three times for 5 min in AP buffer. To stain the tissue, 500 µl staining solution was added to the samples. The color reaction was followed under a binocular and stopped at the right time point by rinsing quickly and washing three times for 10 min with PBT. To remove unspecific staining the stained samples were washed six times for 10 min with 10 min. The tissue was mounted in 80% Glycerol and analysed under a microscope. Pictures were taken with a Nikon Microphot-FXA microscope with a Sony NEX-5RK digital camera.

XGal staining of imaginal discs

Solutions:
- Buffer B (for 10 ml: 500µl 200 mM NaP_i (pH 7.2), 300µl 5 M NaCl, 10µl 1 M $MgCl_2$ plus 9.5 ml H_2O)
- XGal staining solution (960µl Buffer B, 10µl 333 mM $K_4[FeII(CN)_6]$, 10µl 333 mM $K_3[FeIII(CN)_6]$, 5µl 10% Triton X-100 (Fluka), 16µl 5% XGal (AppliChem) in dimethylformamide)

Procedure:

Third instar larvae were torn apart, inverted and stored temporarily in PBS on ice. After enough samples had been collected, the carcasses were fixed in 1 % glutaraldehyde (Fluka) in PBS for 15 min on ice. After fixation, the fixative was removed and the larvae were washed twice with PBST (0.1% Tween-20 in PBS) and 500 µl of the XGal staining solution was added. The larvae were incubated at 37°C until the staining was clearly visible (usually after 1 hour). Afterwards the carcasses were rinsed twice with PBST and dissected in PBS. The imaginal disc were transferred onto a microscope slide in 80% Glycerol and analyzed under the Zeiss Axiophot microscope.

Imaginal disc antibody staining

Solutions:
- PBS fixative (for 1ml: 650 µl H2O, 250 µl PFA (16% stock solution (Electron Microscopy Sciences)), 100 µl 10xPBS)
- PBST (1xPBS with 0.3% Triton X-100)
- PBTN (PBST plus 2% normal donkey serum)

Procedure:

The anterior part of larvae was inverted and stored in cold 1xPBS until enough samples were collected. Afterwards, the tissue was fixed in PBS fixative for 25 min at RT on a rotating wheel. Then rinsed once and washed 4 times for 15 min in PBST to stop fixing and permeabilize the tissue. To counteract unspecific antibody binding later, the samples were blocked for 20 min at RT or 2h at 4°C in PBTN. Then, the samples were incubated with the primary antibodies (dilutions for most used antibodies; 1:120 for α-Wg, 1:1000 for α-Ap, or 1:1000 α- β-Gal) in PBTN at 4°C over night. The next day, the primary antibody solution was removed carefully (can be used again to achieve staining with less background signal!) and the samples were rinsed three times quickly with PBST. The larvae were then washed extensively six times for 20 min with PBST. When unbound primary antibodies were washed off, the secondary antibodies were applied (1:1000 of respective abcam®-antibodies in PBTN) and incubated on a rotator for 1-2 h at RT. Again, the samples were washed like after the removal of the primary antibodies (three times rinsing, 6 times 20 min with PBST). After the extensive washing, the PBST was completely removed and the sample were stored in VECTASHIELD H-1000 (vector Laboratories).

The imaginal discs were mounted on glass slides and analyzed with a confocal fluorescence microscope (Leica SP-5 MP).

Collection and fixation of *Drosophila* embryos

Solutions:
- Embryo fix 1x PBS with 4% paraformaldehyde (Fluka)
- N-heptane
- Methanol

Procedure:
Fly cages with grape juice plates and yeast were set up and 0-24 h old embryos were collected. The embryos were dechorionized with 3.5% sodium hypochloride solution for 3 min and subsequently rinsed with water. For fixation, the embryos were transferred into a 1.5 ml tube with 500 µl n-heptane and 500 µl embryo fix and incubated for 30 min on a fast turning rotor. To remove the vitelline membrane, the lower aqueous phase was removed and replaced with 500 µl methanol. The embryos were quickly vortexed or shaken vigorously until they sank to the bottom of the tube. The embryos were then washed 3 times with 1 ml methanol for 20 min on a slow turning rotor and finally stored in methanol at -20°C or directly used for antibody staining.

Whole mount antibody staining of *Drosophila* embryos

The fixed embryos were rehydrated in 1 ml PBST/methanol 1:1 for 2 min and washed twice with PBST for 10 min on a rotor. After incubation in PBTN for 20 min, the embryos were incubated with primary antibodies over night at 4°C on a rotor.

The next day, the embryos were treated as described in the imaginal disc antibody staining protocol.

The stained embryos were mounted in 80% Glycerol, transferred to a microscope slide and analyzed with confocal microscopy.

Generation of allele-specific *apterous* MARCM clones

In general, MARCM clones for alleles on 2R are induced by recombining them with either FRT G13 (at 42B) or FRT 42D located at the base of this chromosome arm. However, the *apterous* gene maps to 41F and therefore just proximal to these FRT sites, making them useless for the production of *ap* MARCM clones. Hence, *ap*-specific fly stocks had to be established for this purpose. Many of the deletions in the *ap* locus produced for this study were obtained by flp-induced recombination. Therefore, such deletions keep a single FRT right at the break point. In combination with the more proximal of the two FRT-containing transgenes used to generate the deletion, this FRT can be utilized to induce MARCM clones. This strategy allowed us to establish fly stocks for the production of allele-specific MARCM clones.

Below, the stocks for the production of ap^{DG1}, ap^{DG12} and ap^{DG14} MARCM clones in the wing disc are described. It is important to note that ap^{DG1} clones are indistinguishable from those obtained with ap^{DG8}, a deletion removing the complete *apterous* structural gene (Hamaratoglu and Müller, unpublished).

For <u>*Df(2R)ap^{DG1}* clones, the following genotypes were crossed:</u>

Virgins: y w P{hs-flp} ; ap^{DG1} / CyO, GFP ; P{tub-Gal4} P{UAS-GFP} / TM6B, Hu

Males: y w ; ap^{MM} P{tub-Gal80} ; +

Since the FRT of ap^{MM} was used to construct ap^{DG1}, these two chromosomes contain a single FRT at exactly the same site.

Genotype of interest:

y w P{hs-flp} / y w (or Y); ap^{DG1} / ap^{MM} P{tub-Gal80} ; P{tub-Gal4} P{UAS-GFP} / +

3rd instar larvae of this genotype can be identified by selecting against CyO,GFP. Among the remaining larvae, animals of interest show a distinct pattern of GFP-positive spots in the anterior part of the animal where imaginal tissue has been developing. The same selection procedure applies for the clones described below.

For the demonstration of autoregulation of the apDV enhancer, the following males were crossed with the same virgins as described above:

Males: y w ; ap^{MM} P{tub-Gal80} ; M{apDV-LacZ }zh-86Fb / TM3, Sb

Genotype of interest:

y w P{hs-flp} / y w (or Y); ap^{DG1} / ap^{MM} P{tub-Gal80} ; P{tub-Gal4} P{UAS-GFP} / M{apDV-LacZ }zh-86Fb

Note that in larvae, it is not possible to select for the presence of M{apDV-LacZ }zh-86Fb. Only immune-detection of β-Gal in fixed imaginal discs will reveal its presence or absence.

For $Df(2R)ap^{DG12}$ clones, the following genotypes were crossed:

Virgins: y w P{hs-flp}; ap^{DG12} / CyO, GFP ; P{tub-Gal4} P{UAS-GFP} / TM6B, Hu

Males: y w ; $ap^{EE29.19}$ P{tub-Gal80} ; +

Since the FRT of $ap^{EE29.19}$ was used to construct ap^{DG12}, these two chromosomes contain a single FRT at exactly the same site.

Genotype of interest:

y w P{hs-flp} / y w (or Y); ap^{DG12} / $ap^{EE29.19}$ P{tub-Gal80} ; P{tub-Gal4} P{UAS-GFP} / +

For $Df(2R)ap^{DG14}$ clones, the following genotypes were crossed:

Virgins: y w P{hs-flp}; ap^{DG14} / CyO, GFP ; P{tub-Gal4} P{UAS-GFP} / TM6B, Hu

Males: y w ; $ap^{DD35.34}$ P{tub-Gal80} ; +

Since the FRT of $ap^{DD35.34}$ was used to construct ap^{DG14}, these two chromosomes contain a single FRT at exactly the same site.

Genotype of interest:

y w P{hs-flp} / y w (or Y); ap^{DG14} / $ap^{DD35.34}$ P{tub-Gal80} ; P{tub-Gal4} P{UAS-GFP} / +

REFERENCES

AFFOLTER M., MANN R., 2001 Development. Legs, eyes, or wings--selectors and signals make the difference. Science **292**: 1080–1.

AFFOLTER M., BASLER K., 2007 The Decapentaplegic morphogen gradient: from pattern formation to growth regulation. Nat. Rev. Genet. **8**: 663–74.

ALEXANDRE C., BAENA-LOPEZ A., VINCENT J.-P., 2014 Patterning and growth control by membrane-tethered Wingless. Nature **505**: 180–5.

ALTABEF M., CLARKE J. D., TICKLE C., 1997 Dorso-ventral ectodermal compartments and origin of apical ectodermal ridge in developing chick limb. Development **124**: 4547–56.

ALTARATZ M., APPLEBAUM S. W., RICHARD D. S., GILBERT L. I., SEGAL D., 1991 Regulation of juvenile hormone synthesis in wild-type and apterous mutant Drosophila. Mol. Cell. Endocrinol. **81**: 205–16.

ARNOLD P., ERB I., PACHKOV M., MOLINA N., NIMWEGEN E. VAN, 2012 MotEvo: integrated Bayesian probabilistic methods for inferring regulatory sites and motifs on multiple alignments of DNA sequences. Bioinformatics **28**: 487–94.

ARQUES C. G., DOOHAN R., SHARPE J., TORRES M., 2007 Cell tracing reveals a dorsoventral lineage restriction plane in the mouse limb bud mesenchyme. Development **134**: 3713–22.

AVEROF M., COHEN S. M., 1997 Evolutionary origin of insect wings from ancestral gills. Nature **385**: 627–630.

BACHMANN A., KNUST E., 1998 Positive and negative control of Serrate expression during early development of the Drosophila wing. Mech. Dev. **76**: 67–78.

BARRIOS A., POOLE R. J., DURBIN L., BRENNAN C., HOLDER N., WILSON S. W., 2003 Eph/Ephrin signaling regulates the mesenchymal-to-epithelial transition of the paraxial mesoderm during somite morphogenesis. Curr. Biol. **13**: 1571–82.

BATLLE E., WILKINSON D. G., 2012 Molecular mechanisms of cell segregation and boundary formation in development and tumorigenesis. Cold Spring Harb. Perspect. Biol. **4**: a008227.

BEISEL C., PARO R., 2011 Silencing chromatin: comparing modes and mechanisms. Nat. Rev. Genet. **12**: 123–35.

BEJARANO F., LUQUE C. M., HERRANZ H., SORROSAL G., RAFEL N., PHAM T. T., MILÁN M., 2008 A gain-of-function suppressor screen for genes involved in dorsal-ventral boundary formation in the Drosophila wing. Genetics **178**: 307–23.

BERGMAN C. M., PFEIFFER B. D., RINCÓN-LIMAS D. E., HOSKINS R. A, GNIRKE A., MUNGALL C. J., WANG A. M., KRONMILLER B., PACLEB J., PARK S., STAPLETON M., WAN K., GEORGE R. A, JONG P. J. DE, BOTAS

J., RUBIN G. M., CELNIKER S. E., 2002 Assessing the impact of comparative genomic sequence data on the functional annotation of the Drosophila genome. Genome Biol. **3**: RESEARCH0086.

BERGMAN C. M., CARLSON J. W., CELNIKER S. E., 2004 Drosophila DNase I footprint database: a systematic genome annotation of transcription factor binding sites in the fruitfly, Drosophila melanogaster. Bioinformatics **21**: 1747–1749.

BIELI D., KANCA O., GOHL D., DENES A., SCHEDL P., AFFOLTER M., MÜLLER M., 2015 The Drosophila melanogaster Mutants apblot and apXasta Affect an Essential apterous Wing Enhancer. G3 (Bethesda). **5**: 1129–1143.

BISCHOF J., MAEDA R. K., HEDIGER M., KARCH F., BASLER K., 2007 An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. Proc. Natl. Acad. Sci. U. S. A. **104**: 3312–7.

BLAIR S. S., BROWER D. L., THOMAS J. B., ZAVORTINK M., 1994 The role of apterous in the control of dorsoventral compartmentalization and PS integrin gene expression in the developing wing of Drosophila. Development **120**: 1805–15.

BLASTYÁK A., MISHRA R. K., KARCH F., GYURKOVICS H., 2006 Efficient and specific targeting of Polycomb group proteins requires cooperative interaction between Grainyhead and Pleiohomeotic. Mol. Cell. Biol. **26**: 1434–44.

BOURGOUIN C., LUNDGREN S. E., THOMAS J. B., 1992 apterous is a drosophila LIM domain gene required for the development of a subset of embryonic muscles. Neuron **9**: 549–561.

BRAND A. H., PERRIMON N., 1993 Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development **118**: 401–15.

BRODLAND G. W., 2002 The Differential Interfacial Tension Hypothesis (DITH): a comprehensive theory for the self-rearrangement of embryonic cells and tissues. J. Biomech. Eng. **124**: 188–97.

BRONSTEIN R., LEVKOVITZ L., YOSEF N., YANKU M., RUPPIN E., SHARAN R., WESTPHAL H., OLIVER B., SEGAL D., 2010a Transcriptional regulation by CHIP/LDB complexes. PLoS Genet. **6**: e1001063.

BRONSTEIN R., LEVKOVITZ L., YOSEF N., YANKU M., RUPPIN E., SHARAN R., WESTPHAL H., OLIVER B., SEGAL D., 2010b Transcriptional regulation by CHIP/LDB complexes. PLoS Genet. **6**: e1001063.

BROWN J. L., MUCCI D., WHITELEY M., DIRKSEN M. L., KASSIS J. A., 1998 The Drosophila Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1. Mol. Cell **1**: 1057–64.

BROWN J. L., GRAU D. J., DEVIDO S. K., KASSIS J. A., 2005 An Sp1/KLF binding site is important for the activity of a Polycomb group response element from the Drosophila engrailed gene. Nucleic Acids Res. **33**: 5181–5189.

Brown J. L., Kassis J. A., 2013 Architectural and functional diversity of polycomb group response elements in Drosophila. Genetics **195**: 407–19.

Buceta J., Herranz H., Canela-Xandri O., Reigada R., Sagués F., Milán M., 2007 Robustness and stability of the gene regulatory network involved in DV boundary formation in the Drosophila wing. PLoS One **2**: e602.

Busturia A., Casanova J., Sánchez-Herrero E., González R., Morata G., 1989 Genetic structure of the abd-A gene of Drosophila. Development **107**: 575–83.

Butterworth F. M., King R. C., 1965 The developmental genetics of apterous mutants of Drosophila melanogaster. Genetics **52**: 1153–74.

Byrd K. N., Shearn A., 2003 ASH1, a Drosophila trithorax group protein, is required for methylation of lysine 4 residues on histone H3. Proc. Natl. Acad. Sci. U. S. A. **100**: 11535–40.

Cadigan K. M., 2012 TCFs and Wnt/β-catenin signaling: more than one way to throw the switch. Curr. Top. Dev. Biol. **98**: 1–34.

Cao R., Wang L., Wang H., Xia L., Erdjument-Bromage H., Tempst P., Jones R. S., Zhang Y., 2002 Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science **298**: 1039–43.

Cao R., Zhang Y., 2004 The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. Curr. Opin. Genet. Dev. **14**: 155–64.

Capovilla M., Kambris Z., Botas J., 2001 Direct regulation of the muscle-identity gene apterous by a Hox protein in the somatic mesoderm. Development **128**: 1221–30.

Caron H., Schaik B. van, Mee M. van der, Baas F., Riggins G., Sluis P. van, Hermus M. C., Asperen R. van, Boon K., Voûte P. A., Heisterkamp S., Kampen A. van, Versteeg R., 2001 The human transcriptome map: clustering of highly expressed genes in chromosomal domains. Science **291**: 1289–92.

Chinwalla V., Jane E. P., Harte P. J., 1995 The Drosophila trithorax protein binds to specific chromosomal sites and is co-localized with Polycomb at many sites. EMBO J. **14**: 2056–65.

Ciabrelli F., Cavalli G., 2014 Chromatin-Driven Behavior of Topologically Associating Domains. J. Mol. Biol. **427**: 608–625.

Cohen B., McGuffin M. E., Pfeifle C., Segal D., Cohen S. M., 1992 apterous, a gene required for imaginal disc development in Drosophila encodes a member of the LIM family of developmental regulatory proteins. Genes Dev. **6**: 715–729.

Cohen S. M., 1993 Imaginal Disc Development. In: Bate M, Martinez Arias A (Eds.), *The Development of Drosophila melanogaster*, Cold Spring Harbor Laboratory Press, New

York, pp. 747–842.

CREWS S. T., PEARSON J. C., 2009 Transcriptional autoregulation in development. Curr. Biol. **19**: R241–6.

CRIPPS R. M., BLACK B. L., ZHAO B., LIEN C. L., SCHULZ R. A., OLSON E. N., 1998 The myogenic regulatory gene Mef2 is a direct target for transcriptional activation by Twist during Drosophila myogenesis. Genes Dev. **12**: 422–34.

CRIPPS R. M., LOVATO T. L., OLSON E. N., 2004 Positive autoregulation of the Myocyte enhancer factor-2 myogenic control gene during somatic muscle development in Drosophila. Dev. Biol. **267**: 536–47.

CUNNINGHAM M. D., BROWN J. L., KASSIS J. A., 2009 Characterization of the Polycomb Group Response Elements of the Drosophila melanogaster invected Locus. Mol. Cell. Biol. **30**: 820–828.

CURTISS J., HEILIG J. S., 1995 Establishment of Drosophila imaginal precursor cells is controlled by the Arrowhead gene. Development **121**: 3819–28.

CZERMIN B., MELFI R., MCCABE D., SEITZ V., IMHOF A., PIRROTTA V., 2002 Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. Cell **111**: 185–96.

DAHMANN C., BASLER K., 1999 Compartment boundaries: at the edge of development. Trends Genet. **15**: 320–6.

DAHMANN C., OATES A. C., BRAND M., 2011 Boundary formation and maintenance in tissue development. Nat. Rev. Genet. **12**: 43–55.

DECAMILLIS M., CHENG N. S., PIERRE D., BROCK H. W., 1992 The polyhomeotic gene of Drosophila encodes a chromatin protein that shares polytene chromosome-binding sites with Polycomb. Genes Dev. **6**: 223–32.

DÉJARDIN J., CAVALLI G., 2004 Chromatin inheritance upon Zeste-mediated Brahma recruitment at a minimal cellular memory module. EMBO J. **23**: 857–68.

DÉJARDIN J., RAPPAILLES A., CUVIER O., GRIMAUD C., DECOVILLE M., LOCKER D., CAVALLI G., 2005 Recruitment of Drosophila Polycomb group proteins to chromatin by DSP1. Nature **434**: 533–8.

DÉJARDIN J., KINGSTON R. E., 2009 Purification of proteins associated with specific genomic Loci. Cell **136**: 175–86.

DEVIDO S. K., KWON D., BROWN J. L., KASSIS J. A., 2008 The role of Polycomb-group response elements in regulation of engrailed transcription in Drosophila. Development **135**: 669–676.

Diaz-Benjumea F. J., Cohen S. M., 1993 Interaction between dorsal and ventral cells in the imaginal disc directs wing development in Drosophila. Cell **75**: 741–52.

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**: 4215–25.

Dixon J. R., Selvaraj S., Yue F., Kim A., Li Y., Shen Y., Hu M., Liu J. S., Ren B., 2012 Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature **485**: 376–80.

Dodd I. B., Shearwin K. E., Egan J. B., 2005 Revisited gene regulation in bacteriophage lambda. Curr. Opin. Genet. Dev. **15**: 145–52.

Don R. H., Cox P. T., Wainwright B. J., Baker K., Mattick J. S., 1991 "Touchdown" PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res. **19**: 4008.

Dou Y., Milne T. A., Tackett A. J., Smith E. R., Fukuda A., Wysocka J., Allis C. D., Chait B. T., Hess J. L., Roeder R. G., 2005 Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. Cell **121**: 873–85.

Ernst J., Kellis M., 2010 Discovery and characterization of chromatin states for systematic annotation of the human genome. Nat. Biotechnol. **28**: 817–25.

Estella C., McKay D. J., Mann R. S., 2008 Molecular integration of wingless, decapentaplegic, and autoregulatory inputs into Distalless during Drosophila leg development. Dev. Cell **14**: 86–96.

Fernández-Fúnez P., Lu C. H., Rincón-Limas D. E., García-Bellido a, Botas J., 1998 The relative expression amounts of apterous and its co-factor dLdb/Chip are critical for dorso-ventral compartmentalization in the Drosophila wing. EMBO J. **17**: 6846–6853.

Feuerstein R., Wang X., Song D., Cooke N. E., Liebhaber S. A., 1994 The LIM/double zinc-finger motif functions as a protein dimerization domain. Proc. Natl. Acad. Sci. U. S. A. **91**: 10655–9.

Fraser S., Keynes R., Lumsden A., 1990 Segmentation in the chick embryo hindbrain is defined by cell lineage restrictions. Nature **344**: 431–5.

Garcia-Bellido A., Ripoll P., Morata G., 1973 Developmental compartmentalisation of the wing disk of Drosophila. Nat. New Biol. **245**: 251–3.

Garcia-Bellido A., 1975 Genetic control of wing disc development in Drosophila. Ciba Found. Symp.: 161–82.

Gehring W. J., Qian Y. Q., Billeter M., Furukubo-Tokunaga K., Schier A. F., Resendez-Perez D., Affolter M., Otting G., Wüthrich K., 1994 Homeodomain-DNA recognition. Cell **78**: 211–23.

127

GHAVI-HELM Y., KLEIN F. A., PAKOZDI T., CIGLAR L., NOORDERMEER D., HUBER W., FURLONG E. E. M., 2014 Enhancer loops appear stable during development and are associated with paused polymerase. Nature **512**: 96–100.

GOHL D., MÜLLER M., PIRROTTA V., AFFOLTER M., SCHEDL P., 2008 Enhancer blocking and transvection at the Drosophila apterous locus. Genetics **178**: 127–43.

GROSSNIKLAUS U., PARO R., 2014 Transcriptional silencing by polycomb-group proteins. Cold Spring Harb. Perspect. Biol. **6**: a019331.

GROTH A. C., FISH M., NUSSE R., CALOS M. P., 2004 Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. Genetics **166**: 1775–82.

GUSS K. A., NELSON C. E., HUDSON A., KRAUS M. E., CARROLL S. B., 2001 Control of a genetic regulatory network by a selector gene. Science **292**: 1164–7.

GUTHRIE S., BUTCHER M., LUMSDEN A., 1991 Patterns of cell division and interkinetic nuclear migration in the chick embryo hindbrain. J. Neurobiol. **22**: 742–754.

HAGSTROM K., MULLER M., SCHEDL P., 1997 A Polycomb and GAGA dependent silencer adjoins the Fab-7 boundary in the Drosophila bithorax complex. Genetics **146**: 1365–80.

HALDER G., JOHNSON R. L., 2011 Hippo signaling: growth control and beyond. Development **138**: 9–22.

HARIHARAN I. K., 2015 Organ Size Control: Lessons from Drosophila. Dev. Cell **34**: 255–265.

HELD L. I. J., 2005 *Imaginal Discs - The Genetic and Cellular Logic of Pattern Formation*. Cambridge Universitiy Press.

HEYMAN I., KENT A., LUMSDEN A., 1993 Cellular morphology and extracellular space at rhombomere boundaries in the chick embryo hindbrain. Dev. Dyn. **198**: 241–53.

HIROMI Y., GEHRING W. J., 1987 Regulation and function of the Drosophila segmentation gene fushi tarazu. Cell **50**: 963–74.

HODGSON J. W., ARGIROPOULOS B., BROCK H. W., 2001 Site-specific recognition of a 70-base-pair element containing d(GA)(n) repeats mediates bithoraxoid polycomb group response element-dependent silencing. Mol. Cell. Biol. **21**: 4528–43.

HOU C., LI L., QIN Z. S., CORCES V. G., 2012 Gene density, transcription, and insulators contribute to the partition of the Drosophila genome into physical domains. Mol. Cell **48**: 471–84.

HURST L. D., PÁL C., LERCHER M. J., 2004 The evolutionary dynamics of eukaryotic gene order. Nat. Rev. Genet. **5**: 299–310.

INOUE T., TANAKA T., TAKEICHI M., CHISAKA O., NAKAMURA S., OSUMI N., 2001 Role of cadherins in maintaining the compartment boundary between the cortex and striatum during development. Development **128**: 561–9.

IRVINE K. D., BOTAS J., JHA S., MANN R. S., HOGNESS D. S., 1993 Negative autoregulation by Ultrabithorax controls the level and pattern of its expression. Development **117**: 387–99.

IRVINE K. D., RAUSKOLB C., 2001 Boundaries in development: formation and function. Annu. Rev. Cell Dev. Biol. **17**: 189–214.

JIMENEZ-GURI E., UDINA F., COLAS J.-F., SHARPE J., PADRÓN-BARTHE L., TORRES M., PUJADES C., 2010 Clonal analysis in mice underlines the importance of rhombomeric boundaries in cell movement restriction during hindbrain segmentation. PLoS One **5**: e10112.

JOYNER A. L., LIU A., MILLET S., 2000 Otx2, Gbx2 and Fgf8 interact to position and maintain a mid-hindbrain organizer. Curr. Opin. Cell Biol. **12**: 736–41.

KASSIS J. A., VANSICKLE E. P., SENSABAUGH S. M., 1991 A fragment of engrailed regulatory DNA can mediate transvection of the white gene in Drosophila. Genetics **128**: 751–61.

KASSIS J. A., 1994 Unusual properties of regulatory DNA from the Drosophila engrailed gene: three "pairing-sensitive" sites within a 1.6-kb region. Genetics **136**: 1025–38.

KASSIS J. A., 2002 Pairing-sensitive silencing, polycomb group response elements, and transposon homing in Drosophila. Adv. Genet. **46**: 421–38.

KASSIS J. A., BROWN J. L., 2013 Polycomb group response elements in Drosophila and vertebrates. Adv. Genet. **81**: 83–118.

KENNISON J. A., TAMKUN J. W., 1988 Dosage-dependent modifiers of polycomb and antennapedia mutations in Drosophila. Proc. Natl. Acad. Sci. U. S. A. **85**: 8136–40.

KENNISON J. A., 1995 The Polycomb and trithorax group proteins of Drosophila: trans-regulators of homeotic gene function. Annu. Rev. Genet. **29**: 289–303.

KHARCHENKO P. V, ALEKSEYENKO A. A., SCHWARTZ Y. B., MINODA A., RIDDLE N. C., ERNST J., SABO P. J., LARSCHAN E., GORCHAKOV A. A., GU T., LINDER-BASSO D., PLACHETKA A., SHANOWER G., TOLSTORUKOV M. Y., LUQUETTE L. J., XI R., JUNG Y. L., PARK R. W., BISHOP E. P., CANFIELD T. K., SANDSTROM R., THURMAN R. E., MACALPINE D. M., STAMATOYANNOPOULOS J. A., KELLIS M., ELGIN S. C. R., KURODA M. I., PIRROTTA V., KARPEN G. H., PARK P. J., 2011 Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. Nature **471**: 480–5.

KIM J., IRVINE K. D., CARROLL S. B., 1995 Cell recognition, signal induction, and symmetrical gene activation at the dorsal-ventral boundary of the developing Drosophila wing. Cell **82**: 795–802.

KIM J., SEBRING a, ESCH J. J., KRAUS M. E., VORWERK K., MAGEE J., CARROLL S. B., 1996 Integration of positional signals and regulation of wing formation and identity by Drosophila vestigial gene. Nature **382**: 133–138.

KIM T. H., BARRERA L. O., ZHENG M., QU C., SINGER M. A., RICHMOND T. A., WU Y., GREEN R. D., REN B., 2005 A high-resolution map of active promoters in the human genome. Nature **436**:

876–80.

Kim H., Kim K., Kim J., Kim S.-H., Yim J., 2012 Mutagenesis by imprecise excision of the piggyBac transposon in Drosophila melanogaster. Biochem. Biophys. Res. Commun. **417**: 335–9.

Kingston R. E., Tamkun J. W., 2014 Transcriptional regulation by trithorax-group proteins. Cold Spring Harb. Perspect. Biol. **6**: a019349.

Klein T., Arias A. M., 1998 Different spatial and temporal interactions between Notch, wingless, and vestigial specify proximal and distal pattern elements of the wing in Drosophila. Dev. Biol. **194**: 196–212.

Klein T., 2001 Wing disc development in the fly: The early stages. Curr. Opin. Genet. Dev. **11**: 470–475.

Koshida S., Kishimoto Y., Ustumi H., Shimizu T., Furutani-Seiki M., Kondoh H., Takada S., 2005 Integrinalpha5-dependent fibronectin accumulation for maintenance of somite boundaries in zebrafish embryos. Dev. Cell **8**: 587–98.

Kruger W., Peterson C. L., Sil A., Coburn C., Arents G., Moudrianakis E. N., Herskowitz I., 1995 Amino acid substitutions in the structured domains of histones H3 and H4 partially relieve the requirement of the yeast SWI/SNF complex for transcription. Genes Dev. **9**: 2770–9.

Landsberg K. P., Farhadifar R., Ranft J., Umetsu D., Widmann T. J., Bittig T., Said A., Jülicher F., Dahmann C., 2009 Increased cell bond tension governs cell sorting at the Drosophila anteroposterior compartment boundary. Curr. Biol. **19**: 1950–5.

Langenberg T., Brand M., 2005 Lineage restriction maintains a stable organizer cell population at the zebrafish midbrain-hindbrain boundary. Development **132**: 3209–16.

Lawrence P. A., Green S. M., Johnston P., 1978 Compartmentalization and growth of the Drosophila abdomen. J. Embryol. Exp. Morphol. **43**: 233–45.

Lawrence P. A., Struhl G., 1982 Further studies of the engrailed phenotype in Drosophila. EMBO J. **1**: 827–33.

Lecuit T., Brook W. J., Ng M., Calleja M., Sun H., Cohen S. M., 1996 Two distinct mechanisms for long-range patterning by Decapentaplegic in the Drosophila wing. Nature **381**: 387–93.

Lee Y. M., Park T., Schulz R. A., Kim Y., 1997 Twist-mediated activation of the NK-4 homeobox gene in the visceral mesoderm of Drosophila requires two distinct clusters of E-box regulatory elements. J. Biol. Chem. **272**: 17531–41.

Lee T., Luo L., 2001 Mosaic analysis with a repressible cell marker (MARCM) for Drosophila neural development. Trends Neurosci. **24**: 251–4.

Lindsley D. L., Zimm G. G., 1992 *The Genome of Drosophila Melanogaster*. Elsevier.

LUNDGREN S. E., CALLAHAN C. a, THOR S., THOMAS J. B., 1995 Control of neuronal pathway selection by the Drosophila LIM homeodomain gene apterous. Development **121**: 1769–73.

MAEDA R. K., KARCH F., 2006 The ABC of the BX-C: the bithorax complex explained. Development **133**: 1413–22.

MAJOR R. J., IRVINE K. D., 2005 Influence of Notch on dorsoventral compartmentalization and actin organization in the Drosophila wing. Development **132**: 3823–33.

MARTIN G. R., 1998 The roles of FGFs in the early development of vertebrate limbs. Genes Dev. **12**: 1571–86.

MAURANGE C., PARO R., 2002 A cellular memory module conveys epigenetic inheritance of hedgehog expression during Drosophila wing imaginal disc development. Genes Dev. **16**: 2672–83.

MAVES L., SCHUBIGER G., 1998 A molecular basis for transdetermination in Drosophila imaginal discs: interactions between wingless and decapentaplegic signaling. Development **125**: 115–24.

MEINHARDT H., 1983 Cell determination boundaries as organizing regions for secondary embryonic fields. Dev. Biol. **96**: 375–85.

MELLITZER G., XU Q., WILKINSON D. G., 1999 Eph receptors and ephrins restrict cell intermingling and communication. Nature **400**: 77–81.

MEYEL D. J. VAN, O'KEEFE D. D., JURATA L. W., THOR S., GILL G. N., THOMAS J. B., 1999 Chip and Apterous physically interact to form a functional complex during Drosophila development. Mol. Cell **4**: 259–265.

MILAN M., COHEN S. M., 2003 A re-evaluation of the contributions of Apterous and Notch to the dorsoventral lineage restriction boundary in the Drosophila wing. Development **130**: 553–562.

MILÁN M., DIAZ-BENJUMEA F. J., COHEN S. M., 1998 Beadex encodes an LMO protein that regulates Apterous LIM-homeodomain activity in Drosophila wing development: a model for LMO oncogene function. Genes Dev. **12**: 2912–20.

MILÁN M., COHEN S. M., 1999 Regulation of LIM homeodomain activity in vivo: a tetramer of dLDB and apterous confers activity and capacity for regulation by dLMO. Mol. Cell **4**: 267–73.

MILÁN M., COHEN S. M., 2000 Temporal regulation of apterous activity during development of the Drosophila wing. Development **127**: 3069–78.

MILÁN M., WEIHE U., PÉREZ L., COHEN S. M., 2001 The LRR proteins capricious and Tartan mediate cell interactions during DV boundary formation in the Drosophila wing. Cell **106**: 785–94.

MONIER B., PÉLISSIER-MONIER A., BRAND A. H., SANSON B., 2010 An actomyosin-based barrier inhibits cell mixing at compartmental boundaries in Drosophila embryos. Nat. Cell Biol. **12**: 60–5; sup pp 1–9.

MORATA G., LAWRENCE P. A., 1978 Anterior and posterior compartments in the head of Drosophila. Nature **274**: 473–4.

MÜLLER J., HART C. M., FRANCIS N. J., VARGAS M. L., SENGUPTA A., WILD B., MILLER E. L., O'CONNOR M. B., KINGSTON R. E., SIMON J. A., 2002 Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell **111**: 197–208.

MÜLLER J., KASSIS J. a, 2006 Polycomb response elements and targeting of Polycomb group proteins in Drosophila. Curr. Opin. Genet. Dev. **16**: 476–84.

NELLEN D., BURKE R., STRUHL G., BASLER K., 1996 Direct and long-range action of a DPP morphogen gradient. Cell **85**: 357–68.

NEUMANN C. J., COHEN S. M., 1997 Long-range action of Wingless organizes the dorsal-ventral axis of the Drosophila wing. Development **124**: 871–80.

NOSE A., NAGAFUCHI A., TAKEICHI M., 1988 Expressed recombinant cadherins mediate cell sorting in model systems. Cell **54**: 993–1001.

O'BROCHTA D. A., BRYANT P. J., 1985 A zone of non-proliferating cells at a lineage restriction boundary in Drosophila. Nature **313**: 138–41.

O'KEEFE D. D., THOR S., THOMAS J. B., 1998 Function and specificity of LIM domains in Drosophila nervous system and wing development. Development **125**: 3915–23.

ODOM D. T., DOWELL R. D., JACOBSEN E. S., NEKLUDOVA L., ROLFE P. A., DANFORD T. W., GIFFORD D. K., FRAENKEL E., BELL G. I., YOUNG R. A., 2006 Core transcriptional regulatory circuitry in human hepatocytes. Mol. Syst. Biol. **2**: 2006.0017.

OKTABA K., GUTIÉRREZ L., GAGNEUR J., GIRARDOT C., SENGUPTA A. K., FURLONG E. E. M., MÜLLER J., 2008 Dynamic regulation by polycomb group protein complexes controls pattern formation and the cell cycle in Drosophila. Dev. Cell **15**: 877–89.

PADGETT R. W., ST JOHNSTON R. D., GELBART W. M., A transcript from a Drosophila pattern gene predicts a protein homologous to the transforming growth factor-beta family. Nature **325**: 81–4.

PANIN V. M., PAPAYANNOPOULOS V., WILSON R., IRVINE K. D., 1997 Fringe modulates Notch-ligand interactions. Nature **387**: 908–12.

PARISI F., RICCARDO S., DANIEL M., SAQCENA M., KUNDU N., PESSION A., GRIFONI D., STOCKER H., TABAK E., BELLOSTA P., 2011 Drosophila insulin and target of rapamycin (TOR) pathways regulate GSK3 beta activity to control Myc stability and determine Myc expression in vivo. BMC Biol. **9**: 65.

Paul L., Wang S., Manivannan S. N., Bonanno L., Lewis S., Austin C. L., Simcox A., 2013 Dpp-induced Egfr signaling triggers postembryonic wing development in Drosophila. Proc. Natl. Acad. Sci. U. S. A. **110**: 5058–63.

Pearse R. V, Scherz P. J., Campbell J. K., Tabin C. J., 2007 A cellular lineage analysis of the chick limb bud. Dev. Biol. **310**: 388–400.

Peterson A. J., Kyba M., Bornemann D., Morgan K., Brock H. W., Simon J., 1997 A domain shared by the Polycomb group proteins Scm and ph mediates heterotypic and homotypic interactions. Mol. Cell. Biol. **17**: 6683–92.

Pick L., Schier A., Affolter M., Schmidt-Glenewinkel T., Gehring W. J., 1990 Analysis of the ftz upstream element: germ layer-specific enhancers are independently autoregulated. Genes Dev. **4**: 1224–39.

Pirrotta V., 1988 Vectors for P-mediated transformation in Drosophila. Biotechnology **10**: 437–56.

Ptashne M., 2013 Epigenetics: core misconcept. Proc. Natl. Acad. Sci. U. S. A. **110**: 7101–3.

Pueyo J. I., Couso J. P., 2004 Chip-mediated partnerships of the homeodomain proteins Bar and Aristaless with the LIM-HOM proteins Apterous and Lim1 regulate distal leg development. Development **131**: 3107–3120.

Qiu Q., Chen H., Johnson R. L., 2009 Lmx1b-expressing cells in the mouse limb bud define a dorsal mesenchymal lineage compartment. Genesis **47**: 224–33.

Rafel N., Milán M., 2008 Notch signalling coordinates tissue growth and wing fate specification in Drosophila. Development **135**: 3995–4001.

Rasnitsyn A. P., 1981 A modified paranotal theory of insect wing origin. J. Morphol. **168**: 331–338.

Rebay I., 2002 Keeping the receptor tyrosine kinase signaling pathway in check: lessons from Drosophila. Dev. Biol. **251**: 1–17.

Restrepo S., Zartman J. J., Basler K., 2014 Coordination of patterning and growth by the morphogen DPP. Curr. Biol. **24**: R245–55.

Rincón-Limas D. E., Lu C. H., Canal I., Calleja M., Rodríguez-Esteban C., Izpisúa-Belmonte J. C., Botas J., 1999 Conservation of the expression and function of apterous orthologs in Drosophila and mammals. Proc. Natl. Acad. Sci. U. S. A. **96**: 2165–70.

Rincón-Limas D. E., Lu C. H., Canal I., Botas J., 2000 The level of DLDB/CHIP controls the activity of the LIM homeodomain protein apterous: evidence for a functional tetramer complex in vivo. EMBO J. **19**: 2602–2614.

Ringo J., Werczberger R., Altaratz M., Segal D., 1991 Female sexual receptivity is defective in

juvenile hormone-deficient mutants of the apterous gene of Drosophila melanogaster. Behav. Genet. **21**: 453–69.

RINGROSE L., PARO R., 2007 Polycomb/Trithorax response elements and epigenetic memory of cell identity. Development **134**: 223–232.

RIO D. C., 1990 Molecular mechanisms regulating Drosophila P element transposition. Annu. Rev. Genet. **24**: 543–78.

ROIGNANT J.-Y., LEGENT K., JANODY F., TREISMAN J. E., 2010 The transcriptional co-factor Chip acts with LIM-homeodomain proteins to set the boundary of the eye field in Drosophila. Development **137**: 273–81.

RUBIN G. M., SPRADLING A. C., 1983 Vectors for P element-mediated gene transfer in Drosophila. Nucleic Acids Res. **11**: D1–4.

RULIFSON E. J., BLAIR S. S., 1995 Notch regulates wingless expression and is not required for reception of the paracrine wingless signal during wing margin neurogenesis in Drosophila. Development **121**: 2813–24.

SAGA Y., HATA N., KOSEKI H., TAKETO M. M., 1997 Mesp2: a novel mouse gene expressed in the presegmented mesoderm and essential for segmentation initiation. Genes Dev. **11**: 1827–39.

SAMBROOK J., RUSSELL D. W., 2006a Cloning PCR products by addition of restriction sites to the termini of amplified DNA. CSH Protoc. **2006**.

SAMBROOK J., RUSSELL D. W., 2006b Agarose gel electrophoresis. CSH Protoc. **2006**.

SAMBROOK J., RUSSELL D. W., 2006c Directional cloning into plasmid vectors. CSH Protoc. **2006**.

SAMBROOK J., RUSSELL D. W., 2006d Detection of DNA in agarose gels. CSH Protoc. **2006**.

SANDMANN T., GIRARDOT C., BREHME M., TONGPRASIT W., STOLC V., FURLONG E. E. M., 2007 A core transcriptional network for early mesoderm development in Drosophila melanogaster. Genes Dev. **21**: 436–49.

SAURIN A. J., SHAO Z., ERDJUMENT-BROMAGE H., TEMPST P., KINGSTON R. E., 2001 A Drosophila Polycomb group complex includes Zeste and dTAFII proteins. Nature **412**: 655–60.

SCHIER A. F., GEHRING W. J., 1992 Direct homeodomain-DNA interaction in the autoregulation of the fushi tarazu gene. Nature **356**: 804–7.

SCHUETTENGRUBER B., CHOURROUT D., VERVOORT M., LEBLANC B., CAVALLI G., 2007 Genome regulation by polycomb and trithorax proteins. Cell **128**: 735–45.

SCHWARTZ Y. B., KAHN T. G., NIX D. a, LI X.-Y., BOURGON R., BIGGIN M., PIRROTTA V., 2006 Genome-wide analysis of Polycomb targets in Drosophila melanogaster. Nat. Genet. **38**: 700–5.

SEREBROVSKY A. S., DUBININ N. P., 1930 X-ray experiments with Drosophila. J. Hered. **21**: 259–265.

SEXTON T., YAFFE E., KENIGSBERG E., BANTIGNIES F., LEBLANC B., HOICHMAN M., PARRINELLO H., TANAY A., CAVALLI G., 2012 Three-dimensional folding and functional organization principles of the Drosophila genome. Cell **148**: 458–72.

SIMCOX A. A., GRUMBLING G., SCHNEPP B., BENNINGTON-MATHIAS C., HERSPERGER E., SHEARN A., 1996 Molecular, phenotypic, and expression analysis of vein, a gene required for growth of the Drosophila wing disc. Dev. Biol. **177**: 475–89.

SIMEONE A., 2000 Positioning the isthmic organizer where Otx2 and Gbx2meet. Trends Genet. **16**: 237–40.

SIMON J. A., SUTTON C. A., LOBELL R. B., GLASER R. L., LIS J. T., 1985 Determinants of heat shock-induced chromosome puffing. Cell **40**: 805–17.

SMITH D. M., TABIN C. J., 2000 Clonally related cells are restricted to organ boundaries early in the development of the chicken gut to form compartment boundaries. Dev. Biol. **227**: 422–31.

SNODGRASS R., 1935 *The Thorax. Principles of Insect Morphology*. Cornell Univeristy Press, New York.

SPELLMAN P., RUBIN G., 2002 Evidence for large domains of similarly expressed genes in the Drosophila genome. J. Biol. **1**: 5.

SPRADLING A. C., RUBIN G. M., 1982 Transposition of cloned P elements into Drosophila germ line chromosomes. Science **218**: 341–7.

SPRADLING A. C., STERN D., BEATON A., RHEM E. J., LAVERTY T., MOZDEN N., MISRA S., RUBIN G. M., 1999 The Berkeley Drosophila Genome Project gene disruption project: Single P-element insertions mutating 25% of vital Drosophila genes. Genetics **153**: 135–77.

STEINBERG M. S., 2007 Differential adhesion in morphogenesis: a modern view. Curr. Opin. Genet. Dev. **17**: 281–6.

STEINER E., 1976 Establishment of compartments in the developing leg imaginal discs of Drosophila melanogaster. Wilhelm Roux's Arch. Dev. Biol. **180**: 9–30.

STRUHL G., 1981 Anterior and posterior compartments in the proboscis of Drosophila. Dev. Biol. **84**: 372–85.

STRUTT H., CAVALLI G., PARO R., 1997 Co-localization of Polycomb protein and GAGA factor on regulatory elements responsible for the maintenance of homeotic gene expression. EMBO J. **16**: 3621–32.

TABATA T., SCHWARTZ C., GUSTAVSON E., ALI Z., KORNBERG T. B., 1995 Creating a Drosophila wing de

novo, the role of engrailed, and the compartment border hypothesis. Development **121**: 3359–69.

Taffin M. de, Carrier Y., Dubois L., Bataillé L., Painset A., Gras S. Le, Jost B., Crozatier M., Vincent A., 2015 Genome-Wide Mapping of Collier In Vivo Binding Sites Highlights Its Hierarchical Position in Different Transcription Regulatory Networks. PLoS One **10**: e0133387.

Takahashi Y., Koizumi K., Takagi A., Kitajima S., Inoue T., Koseki H., Saga Y., 2000 Mesp2 initiates somite segmentation through the Notch signalling pathway. Nat. Genet. **25**: 390–6.

Tam P. P., Goldman D., Camus A., Schoenwolf G. C., 2000 Early events of somitogenesis in higher vertebrates: allocation of precursor cells during gastrulation and the organization of a meristic pattern in the paraxial mesoderm. Curr. Top. Dev. Biol. **47**: 1–32.

Tamkun J. W., Deuring R., Scott M. P., Kissinger M., Pattatucci A. M., Kaufman T. C., Kennison J. A., 1992 brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. Cell **68**: 561–72.

Thor S., Andersson S. G., Tomlinson A., Thomas J. B., 1999 A LIM-homeodomain combinatorial code for motor-neuron pathway selection. Nature **397**: 76–80.

Thorpe H. M., Smith M. C., 1998 In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. Proc. Natl. Acad. Sci. U. S. A. **95**: 5505–10.

Thüringer F., Cohen S. M., Bienz M., 1993 Dissection of an indirect autoregulatory response of a homeotic Drosophila gene. EMBO J. **12**: 2419–30.

Tolhuis B., Wit E. de, Muijrers I., Teunissen H., Talhout W., Steensel B. van, Lohuizen M. van, 2006 Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in Drosophila melanogaster. Nat. Genet. **38**: 694–9.

Voelker R. A., Greenleaf A. L., Gyurkovics H., Wisely G. B., Huang S. M., Searles L. L., 1984 Frequent Imprecise Excision among Reversions of a P Element-Caused Lethal Mutation in Drosophila. Genetics **107**: 279–94.

Vogel A., Rodríguez C., Warnken W., Izpisúa-Belmonte J. C., 1995 Dorsal cell fate specified by chick Lmx1 during vertebrate limb development. Nature.

Wang S.-H., Simcox A., Campbell G., 2000 Dual role for Drosophila epidermal growth factor receptor signaling in early wing disc development. Genes Dev. **14**: 2271–2276.

Wartlick O., Mumcu P., Jülicher F., Gonzalez-Gaitan M., 2011 Understanding morphogenetic growth control -- lessons from flies. Nat. Rev. Mol. Cell Biol. **12**: 594–604.

Weihe U., Milán M., Cohen S. M., 2001 Regulation of Apterous activity in Drosophila wing development. Development **128**: 4615–22.

WEISS A., CHARBONNIER E., ELLERTSDÓTTIR E., TSIRIGOS A., WOLF C., SCHUH R., PYROWOLAKIS G., AFFOLTER M., 2010 A conserved activation element in BMP signaling during Drosophila development. Nat. Struct. Mol. Biol. **17**: 69–76.

WILLIAMS J. A., PADDOCK S. W., VORWERK K., CARROLL S. B., 1994 Organization of wing formation and induction of a wing-patterning gene at the dorsal/ventral compartment boundary. Nature **368**: 299–305.

WOLPERT L., 1969 Positional information and the spatial pattern of cellular differentiation. J. Theor. Biol. **25**: 1–47.

WOOTTON R. J., KUKALOVÁ-PECK J., 2000 Flight adaptations in Palaeozoic Palaeoptera (Insecta). Biol. Rev. Camb. Philos. Soc. **75**: 129–67.

WYSOCKA J., SWIGUT T., MILNE T. A., DOU Y., ZHANG X., BURLINGAME A. L., ROEDER R. G., BRIVANLOU A. H., ALLIS C. D., 2005 WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell **121**: 859–72.

XU Q., MELLITZER G., ROBINSON V., WILKINSON D. G., 1999 In vivo cell sorting in complementary segmental domains mediated by Eph receptors and ephrins. Nature **399**: 267–71.

ZECCA M., BASLER K., STRUHL G., 1995 Sequential organizing activities of engrailed, hedgehog and decapentaplegic in the Drosophila wing. Development **121**: 2265–78.

ZECCA M., STRUHL G., 2002a Subdivision of the Drosophila wing imaginal disc by EGFR-mediated signaling. Development **129**: 1357–68.

ZECCA M., STRUHL G., 2002b Control of growth and patterning of the Drosophila wing imaginal disc by EGFR-mediated signaling. Development **129**: 1369–76.

ZECCA M., STRUHL G., 2007a Recruitment of cells into the Drosophila wing primordium by a feed-forward circuit of vestigial autoregulation. Development **134**: 3001–3010.

ZECCA M., STRUHL G., 2007b Control of Drosophila wing growth by the vestigial quadrant enhancer. Development **134**: 3011–3020.

ZELTSER L. M., LARSEN C. W., LUMSDEN A., 2001 A new developmental compartment in the forebrain regulated by Lunatic fringe. Nat. Neurosci. **4**: 683–4.

ZERVAS M., MILLET S., AHN S., JOYNER A. L., 2004 Cell behaviors and genetic lineages of the mesencephalon and rhombomere 1. Neuron **43**: 345–57.

ZHANG C. U., BLAUWKAMP T. A., BURBY P. E., CADIGAN K. M., 2014 Wnt-mediated repression via bipartite DNA recognition by TCF in the Drosophila hematopoietic system. PLoS Genet. **10**: e1004509.

ZINK B., PARO R., 1989 In vivo binding pattern of a trans-regulator of homoeotic genes in Drosophila melanogaster. Nature **337**: 468–71.

PART II

TABLE OF CONTENTS

INTRODUCTION

Why do we need new tools in basic research and developmental biology?

Protein-protein interactions are fundamental to almost all biological processes. To a large extend, protein functions *in vivo* have been studied by genetic manipulations. With the rise of TALEN and CRISPR-Cas9 technologies (for a review see CARROLL 2014), the possibilities of artificial genome engineering have reached new spheres, especially in non-model organisms. Other well-established methods to investigate cellular processes are RNA interference or morpholinos. However, these applications affect gene function upstream of the protein level (either at the DNA or RNA level). This might be a disadvantage when studying proteins and their respective localization, modifications, trafficking, or interaction partners *in situ*. Therefore, it is considered desirable to investigate and manipulate proteins directly, especially in the case of non-dividing cells, in which the use of classical genetics is also limited.

To better describe and understand protein functions in their respective native environment, it is therefore highly important and necessary to manipulate proteins in *in vivo* studies. Classical means to study protein functions, such as the use of small inhibitory molecules, are prone to deleterious off-target effects, which might complicate and falsify the analyses and conclusions. Antibodies have been successfully used to investigate and inhibit proteins in the extracellular environment. Additionally, they are well-established tools to characterize protein distribution, interactions and functions, mostly in fixed tissues or in permeabilized cells. But due to their complex structure, their large size and their stabilization via disulphide bridges, antibodies function rather poorly in living cells.

In the past two decades, new approaches has been taken to identify highly specific protein binders. Towards this end, many small protein scaffolds were adapted and engineered, the best known being the so-called single-domain antibodies (sdAb), also known as nanobodies. The following chapters deal with the discovery, features, selection and applications of these sdAbs.

From conventional to single domain antibodies

Upon infection with foreign harmful agents, such as bacteria or viruses, the body starts to produce antibodies against the invaders. With the help of different antibodies, the immune system then fights infections within the body. The antibody variety is subdivided into different classes, referred to as antibody isotypes. In mammals, five main common isotypes have been described; IgA, IgD, IgE, IgG, and IgM. The prefix Ig stands for immunoglobulin. The most abundant antibodies in our body belong to the IgG isotype (also called conventional antibodies). They are mostly found in the blood and in extracellular fluids (ABBAS and LICHTMAN 2005).

Conventional antibodies (IgG) are composed of two identical large heavy (H-) and two identical small light (L-) chains. Each of the chains is composed in a large part of a constant (C) region and a smaller variable (V) region. These four polypeptide chains are connected by disulphide bonds and build a tetrameric, Y-shaped protein-complex (Figure 1A). Since the C domains prevail, the overall structure of different antibodies is very similar. The CH (constant region of heavy chain) domains play important roles in immunity since they recruit immune cells or activate the complement system. At the tips of each antibody, which are composed of the V domains of the H and the L chains (VH and VL, respectively), a tremendous diversity at the sequence as well as at the structural level is observed. This hyper-variable region, also called antigen binding site or paratope, confers the specificity to the epitope of the antigen.

Since their discovery, antibodies have been applied as valuable tools in basic research, diagnostics and as therapeutics. However, the large size of antibodies and their complex structural organization with various inter- and intramolecular disulphide bridges have proven to be a major obstacles in synthetic protein engineering approaches. Although derivatives of the IgGs (e.g. Fab fragments or scFvs; see Figure 1B) are considerably smaller, they are still composed of two different polypeptide chains that need to be linked via disulphide bonds and/or artificial linkers. Nevertheless, the intracellular expression and use of such antibodies is compromised due to misfolding, insolubility and aggregation.

Some twenty years ago, another set of antibodies has been identified, in addition to conventional antibodies, in *Camelidae* species and cartilaginous fish (HAMERS-CASTERMAN *et al.* 1993; GREENBERG *et al.* 1995). These antibodies are lacking the L-chain as well as the CH1 domain and are called heavy-chain antibodies (HCAb, Figure 1C). The antigen binding site in these antibodies is thus only composed of the variable domain of the heavy chain (VHH,

nanobody©), and they are often also referred to as single-domain antibodies (sdAb). In HCAb, the VHH is considered as the structural and functional equivalent of the Fab fragment (antigen-binding fragment) of conventional antibodies, with the main difference that it consists only of a single polypeptide chain. Hence, it has only approximately half of the size (about 15 kD) of the Fab fragment. This monomeric structure of the VHH makes protein engineering as well as recombinant production considerably easier.
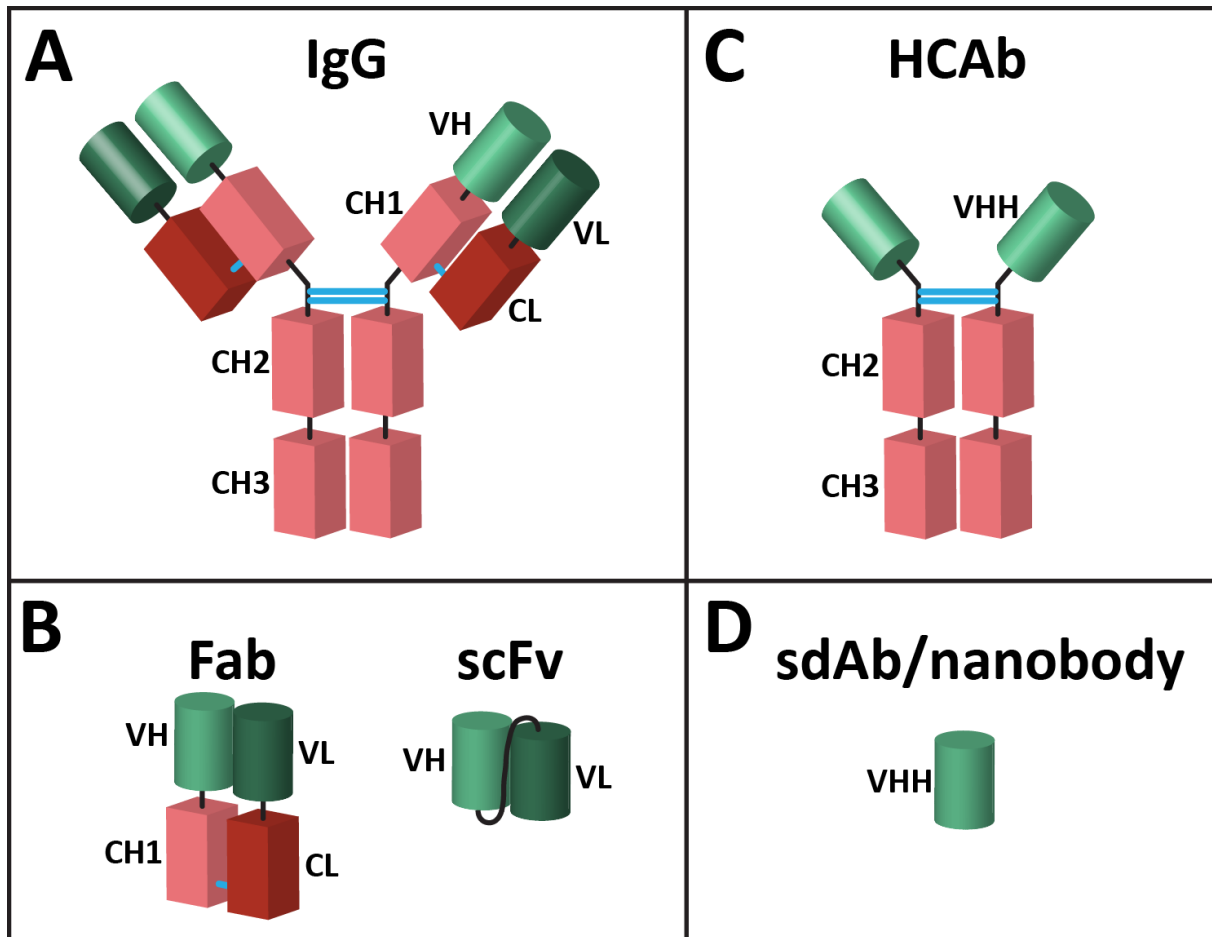
Figure 1 Schematic illustrations of the structures of conventional IgG antibodies and heavy chain antibodies. **(A)** Conventional antibody with constant regions of the heavy chains (CH1-3, light red) and light chains (CL, dark red). The antigen-specificity is conferred by the variable domains of the heavy (VH, light green) and light chains (VL, dark green). Intermolecular disulphide bonds are depicted as blue bars. Approximate size: 150 kD. **(B)** Derivatives of IgGs. *Fragment antigen binding* (Fab) is generated by digestion of IgGs with Papain. It consists of constant and variable domains of the heavy and light chains (approx. size: 50kD). The single chain variable fragment (scFV) is a synthetic fusion protein of the variable heavy and light chains (approx. size: 25 kD). **(C)** Heavy chain antibody (HCAb) as observed in *Camelidae* species. CH1 domain is missing, CH2 is directly fused to the variable domain of the heavy chain antibody (VHH). Approximate size: 100 kD. **(D)** single domain antibody (sdAb), also known as nanobody, consists of the variable domain of the heavy chain antibodies (VHH). Approximate size: 15 kD.

The structure of VHHs

All VHHs share a common conserved framework (FR) structure, which consists of 9 β-strands (see Figure 2). These β-strands fold into 4- and 5-stranded β-sheets, altogether folding into a so-called β-sandwich structure. In addition to the FR, VHHs contain three hypervariable regions, called complement-determining regions (CDRs). The CDRs are located within loops that connect different β-strands and form a continuous surface which is conferring the recognition of the epitope (MUYLDERMANS et al. 1994; MUYLDERMANS 2013). The overall structure resembles very much the conventional VH; however, VHH lack hydrophobic residues present in FR2 in VH, which are important for interaction with VL. In VHHs, these residues are replaced by more hydrophilic amino acids, prohibiting interaction with VL (MUYLDERMANS et al. 1994; VU et al. 1997). As it turns out, the lack of the hydrophobic residues in the FR beneficially affects the solubility of VHHs in recombinant expression approaches.

Following the rationale that the overall structure of the VH, VL and VHH, respectively, is similar, it is apparent that the paratope of conventional antibodies is consisting of six CDR loops (three from VL and three from VH). In the VHHs, the paratope consists of three CDR loops only, but the length of these loops is considerably longer and confers a similar variability and specificity as that observed in conventional antibodies. As the loops are longer, they need to be stabilized, and this is mainly brought about by an intramolecular disulfide bond between CDR1 and 3. This intramolecular disulfide bridge has been linked to a stronger paratope-epitope interaction (GOVAERT et al. 2012). In conventional antibodies, the six CDR loops have been shown to form either a cavity, a groove, or a flat surface, recognizing small molecules, linear peptides, or proteins (SUNDBERG and MARIUZZA 2002). In contrast, VHHs have been shown to form a convex paratope surface, which has been implied in favoring interactions with cavities in proteins, often found at active sites in enzymes (DE GENST et al. 2006). Other paratope architectures have also been observed, including protruding loops and flat surfaces (DESMYTER et al. 2001). Besides, it has also been reported that the paratopes of VHHs can form cavities (SPINELLI et al. 2001). As a result, VHHs are mostly recognizing three-dimensional structures of antigens, which make them optimally suited for applications in living cells. However, VHHs against small polypeptides, such as myc-tag (ChromoTek GmbH) or C-tag (Life Technologies), have recently become available. This demonstrates the possibility of VHHs to interact also with small (poly-) peptides, a feature that has mainly been associated with conventional antibodies. Nanobodies can also bind to small molecules, as exemplified by the isolation of VHHs that can recognize

caffeine (LADENSON *et al.* 2006). These nanobodies were shown to be functional even at high temperatures (around 70°C), demonstrating that VHHs are also very thermostable. Furthermore, VHHs were shown to retain functionality after refolding from denaturing conditions, such as high temperature, pressure, or high guanidinium chloride or urea concentrations (DUMOULIN *et al.* 2002). Overall, the small size, high solubility, specificity, adaptability and stability make the VHH a perfect tool for basic research as well as diagnostic and therapeutic applications.
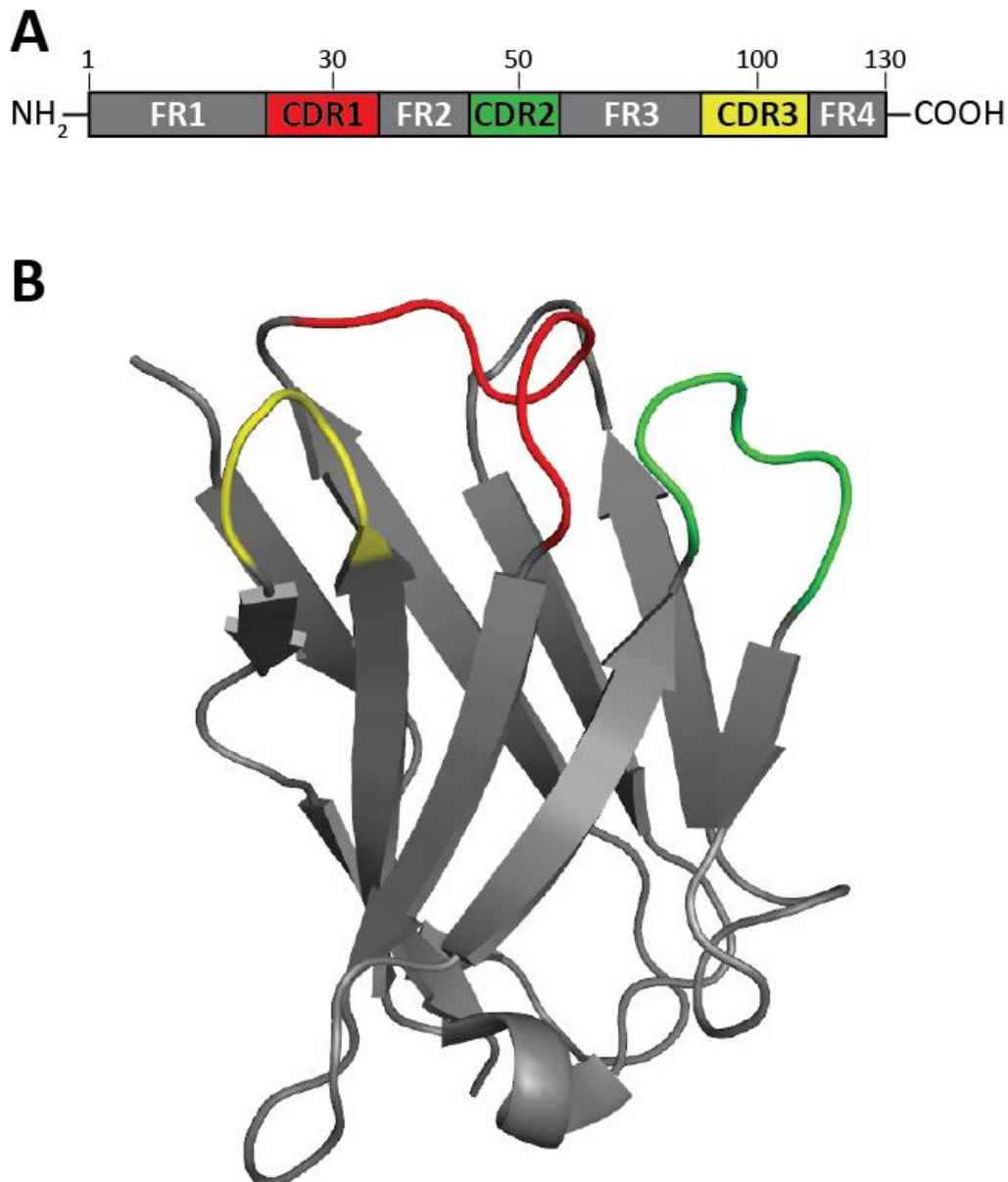
Figure 2 Structure of a VHH. (A) Schematic drawing of a basic VHH structure. In grey are the conserved framework regions (FR1-4), which are interspersed by three complement-determining regions (CDR1-3). On top, the approximate positions of amino acids. **(B)** 3D structure of a GFP-binding VHH (GBP, cABGFP4; from KUBALA *et al.* 2010). The CDRs are in loops (red, green, yellow) connecting the β-strands of the FR (grey). Figure created with the help of the PyMOL program.

Selection of high affinity binders

Usually, the first step in the selection of specific sdAbs is the immunization of *Camelidae* with a protein of interest, which allows a first somatic maturation and selection of binders within the body of the animal. After the immunization, antibody libraries can be cloned from isolated lymphocytes. This approach for the natural selection of VHHs has been proven to yield VHHs with high affinity and specificity. However, since the immunization as well as the preparation of the library has to be repeated for each antigen anew, this method is rather time consuming and costly. Recently, naïve (from non-immunized animals) or synthetic libraries have been generated and have allowed the isolation of VHHs with high affinity, which were suitable for the standard immune-applications (OLICHON and DE MARCO 2012) (Moutel et al, unpublished).

Having a functional library (immunized, naïve, or synthetic) is only the start of a labor-intensive and tenacious selection process. Several techniques have been successfully applied to identify specific binders *in vitro* as well as *in vivo*. The techniques must feature several key aspects, such as the possibility to screen large libraries and –most importantly- the (physical) linking of the VHHs with their respective encoding DNA, thus the direct connection of phenotype (protein) to genotype (DNA). Furthermore, by cycling the selection process (called panning), enrichment of potential binders with high affinity is possible.

The most common *in vitro* selection method is the so-called phage display, where the VHHs are displayed on the surface of bacteriophages, which contain the gene encoding the VHH on a phagemid within the phage particle (see below for more detailed information). Other methods used are bacteria or yeast displays, which are very similar to phage displays (BODER and WITTRUP 1997; SALEMA *et al.* 2013), with the difference of using bacteria or yeast, respectively, as carriers, instead of phages. Ribosome and mRNA-display have also been used successfully in the screening process for functional protein binders (MATTHEAKIS *et al.* 1994; ROBERTS and SZOSTAK 1997). In ribosome and mRNA-display, PCR-based amplification is needed between different rounds of selection, offering the possibility of introducing casual mutations, which additionally increases the diversity of the starting library (for a review on different selection methods see HOOGENBOOM 2005). However, with none of these techniques, it is possible to predict whether the selected binders are correctly folded and functional in the desired environment of the final application (e.g. intracellular compartments).

A conventional high-throughput test to validate protein interaction is the enzyme-linked immunosorbent assay, better known as ELISA. In this assay, the antigen is coated on a surface

and the probe (in this case different VHHs) is either labelled directly or indirectly with an enzyme, whose activity can be measured by various means (VOLLER *et al.* 1978). However, even though a larger number of potential binders can be tested in a short time period, again each binder has to be evaluated for functionality with an appropriate assay.

The major limitations of that these screening/selection methods is the fact that they work exclusively *in vitro*. A first progress towards *in vivo* selection of large antibody/binder libraries has been achieved with the introduction of two-hybrid systems, in which protein-protein interactions can be assayed in intracellular environments of yeast or bacteria. In the yeast two-hybrid (Y2H) assay, the bait (antigen) is fused to the Gal4 DNA-binding domain, and the prey (VHH/binder) to the Gal4 trans-activator domain (FIELDS and SONG 1989). If bait and prey interact with each other in the cell, they are able to activate selection markers or reporter genes, which are under the control of a UAS promoter. This activation of gene transcription can then be monitored to select interacting proteins. The bacteria two-hybrid selection assay works with the same basic principle as the yeast 2H system (HU *et al.* 2000). Recently, these method has been optimized to identify antigen-specific sdAbs (PELLIS *et al.* 2012). Although these two hybrid systems are clearly an improvement towards intrabody selection, there is still no guarantee to isolate binders working in higher eukaryotic systems, such as in cultured cells or in *in vivo* settings.

An elegant and visual way of testing interaction of selected potential binders with their respective antigen in cell culture is a simple co-localization assay. In order to test for co-localization, the antigen and the putative binder are fused with different, spectroscopically separable fluorescent proteins and are co-expressed in cells. Interaction of binder and antigen can be visualized by the merge of the two fluorescent signals. The versatility and applicability of this approach has been nicely demonstrated in the so-termed fluorescent-two-hybrid assay (ZOLGHADR *et al.* 2012). It is also possible to screen for interaction in a degradation assay. Therefore, the potential binders are fused with an F-Box, which ubiquitinates proteins and thereby leads to their degradation by the proteasome (CAUSSINUS *et al.* 2011). Hence, interaction of VHH and antigen can be monitored by the degradation of the antigen (Moutel et al., unpublished). To find the optimal binder for a desired application, other customized functional assays might be useful and/or necessary.

Functionalization of nanobodies/protein binders

After the isolation and characterization of a good binder, the potential applications of such nanobodies are virtually unlimited and open up room for new and innovative ideas. It is feasible to think of linking any protein modification domain to the respective protein binder. Towards this end, a chimeric protein is engineered, which consists of the functional binder/VHH fused to a functional protein domain, such as a modifying enzyme or a localization signal. Below, some examples of the functionalization and application of nanobodies are discussed.

First, it is possible to convert VHHs into monoclonal antibodies by engineering a chimeric heavy chain antibody consisting of a selected, specific VHH fused to a constant antibody domains (Fc) (ZHANG *et al.* 2009; AGRAWAL *et al.* 2012). These 'monoclonals' can be further used to perform classical immunostainings, immunoprecipitations or even as drugs in therapeutics. Furthermore, the Fc fragment is readily interchanged between different species, allowing more flexibility in the final use. Additionally, since the genetic information is available for the selected VHHs, they provide a renewable source of antibodies, which is an important advantage over conventional polyclonal antibodies.

Nanobodies can also be functionalized via their fusion with fluorescent proteins. Such so-called 'chromobodies' have been used to target and trace the distribution and dynamics of recombinant and endogenous proteins in live cells (ROTHBAUER *et al.* 2006; TRAENKLE *et al.* 2015). Chromobodies were also successfully used to perform live imaging of endogenous proteins in zebrafish (PANZA *et al.* 2015). With the development of super-resolution microscopes, it becomes apparent that classical staining methods with primary and secondary antibodies leads to a so-called linkage error. This means that the detected signal is significantly off from the actual signal. Attaching the fluorophore directly to the nanobody thus also increases the resolution of protein distribution detection (RIES *et al.* 2012). Therefore, chomobodies can be used to visualize proteins *in vivo* with a minimal linkage error.

In many cases, protein localization and function are closely linked to each other. Protein function was shown to be altered upon mislocalization induced by chromobody expression in plant cells (SCHORNACK *et al.* 2009). This study also demonstrated that the use of nanobodies is not limited to the animal or bacterial kingdom, but can also be applied in plant research. Fascinatingly, plants have been used to express chimeric VHH antibodies protective against pathogens, such as enterotoxigenic *E. coli* or rotavirus (VIRDI *et al.* 2013; TOKUHARA *et al.* 2013).

Due to their high stability, nanobodies can even be administered orally via plant seeds allowing passive immunization.

To probe protein function and protein-protein interactions in more detail, it is crucial to generate high-resolution structural information. The most widely used technique to solve protein structures is X-ray crystallography, which –as the name implies- relies on the formation of protein crystals and their analysis using X-ray irradiation. Introduction of nanobodies as crystallization chaperones has been demonstrated to facilitate crystal formation of 'difficult' proteins, as they help to stabilize membrane proteins, large protein complexes or proteins with highly flexible, transient conformations (LORIS *et al.* 2003; KOROTKOV *et al.* 2009; RASMUSSEN *et al.* 2011a; b; DOMANSKA *et al.* 2011; BARANOVA *et al.* 2012; KRUSE *et al.* 2013; ABSKHARON *et al.* 2014; CHAIKUAD *et al.* 2014; HELMA *et al.* 2015).

Although having a protein structure is a major step towards understanding the function of a given protein, to confirm and elaborate on protein function in living organisms, it is absolutely necessary to perform *in vivo* studies. Based on a nanobody which has previously been helpful to generate the crystal structure of a G-protein coupled receptor (GPCR) (RASMUSSEN *et al.* 2011a), a chromobody or 'biosensor' has been engineered (IRANNEJAD *et al.* 2013). This way, it was possible to follow conformational changes in the GPCR with high spatiotemporal resolution in living cells. Thus, nanobodies that detect specific conformations can be used as biosensors to monitor conformational changes in real-time *in vivo*. In addition, structure-specific nanobodies have been employed which specifically interfere with protein function (DELANOTE *et al.* 2010; DE CLERCQ *et al.* 2013). Moreover, it has been shown that conformational changes can be induced in proteins using nanobodies within living cells (KIRCHHOFER *et al.* 2010). Using GFP as an antigen, it was possible to alter the fine structure of the fluorophore site upon the expression of specific nanobodies in order to either enhance or minimize GFP fluorescence. Thus, nanobodies are powerful tools to detect and induce conformational changes of proteins in living cells.

Since posttranslational modifications also likely alter protein conformation, it is reasonable to argue that nanobodies can be used to sense posttranslational modifications. However, no sensor of posttranslational modifications based on a nanobody-scaffold has been reported in the literature yet. However, DARPins, synthetic binders based on a different scaffold (for details see below), have been isolated and demonstrated to discriminately bind to the

unphosphorylated and the doubly phosphorylated form of mitogen-activated protein kinase ERK2 (extracellular signal-regulated kinase 2) (Kummer *et al.* 2012).

Nanobodies can be used to specifically alter the stability of a target protein. This was nicely demonstrated by the DeGradFP system (Caussinus *et al.* 2011). In this method, a GFP-specific nanobody is fused with an F-Box domain. The binding of F-box proteins usually leads to the ubiquitination of the target proteins, which are then degraded by the proteasome. The functionalized nanobody-F-Box fusion protein was able to induce the degradation of GFP-fusion proteins in Drosophila and mammalian cell culture. More recently, it has been shown that GFP-fusion proteins can also be specifically phosphorylated *in vivo* by using a kinase domain fused to the anti-GFP nanobody (Caussinus et al., unpublished). These examples illustrate the possibility to modulate proteins at the posttranslational level and study their impact on development or homeostasis of organs and organisms.

In order to study protein function and interactions, it is further possible to use nanobodies that functionally inhibit protein function by blocking protein-protein interactions. Nanobodies can either inhibit protein function by binding to the active site of an enzyme, as exemplified by anti-lysozyme binders (De Genst *et al.* 2006), or can block posttranslational modifications of a substrate, as seen by the delayed phosphorylation of L-plastin (Delanote *et al.* 2010; De Clercq *et al.* 2013). Different nanobodies against the human fic-domain protein HypE, which AMPylates target proteins, were shown to not only inhibit but also overactivate target AMPylation (Truttmann *et al.* 2015). By using a nanobody against gelsolin, the gelsolin-G-action interaction was blocked and the effect on actin polymerization was assayed (Van den Abbeele *et al.* 2010). Upon cytosolic expression of nanobodies directed against influenza nucleoprotein, replication of virus particles was disrupted (Ashour *et al.* 2015). Thus, nanobodies can also be used to alter enzyme function and/or protein-protein interactions. So far, protein function and visualization in living cells or organisms was studied using transient expression of labelled protein from virus-based vectors or via the UAS/Gal4 system (Brand and Perrimon 1993). These expression systems usually lead to a massive overexpression of the protein, which might falsify the results. Importantly, nanobodies do not affect endogenous gene expression of the proteins studied, which is an enormous improvement to study protein functions by means of overexpression of labelled proteins.

Similar to conventional antibodies, nanobodies can be used to trap proteins in all kind of applications, such as co-immunoprecipitation of native protein complexes or affinity

purification (ROTHBAUER *et al.* 2008). Based on a VHH scaffold, many so-called Nano-Traps® (ChromoTek GmbH) have been developed, with binders against different commonly used tags, such as GFP, RFP, GST or myc.

Additionally, these proteins traps can be used *in vivo* to study developmental questions. For example, by tagging Dpp, the *Drosophila* homologue of BMP-2/4, with GFP, the role of this protein was assessed by new means with an anti-GFP nanobody. Dpp forms an extracellular concentration gradient, this gradient formation was successfully blocked when combined with the anti-GFP nanobody (Harmansa et al, unpublished).

Reporter genes, such as GFP, fused to tissue-specific enhancers are used extensively used in cell and developmental biology to label different cell populations. However, while these reporters allow the observation of a developing organ or tissue, the direct manipulation of the observed tissue is not possible by this simple labelling technique. By using two nanobodies that recognize different epitopes on the GFP surface, GFP can be used as a scaffold to assemble functionally significant biological activities. In an elegant study, Tang et al. fused one GFP-binding nanobody to the DNA binding domain of Gal4, while the other GFP recognizing nanobody was fused to a transactivation domain. Each individual nanobody fusion was insufficient to induce UAS-dependent transcription. However, when the two functionalized nanobodies were used in conjunction with GFP, UAS-dependent gene activation and manipulation in the GFP-positive tissue was achieved (TANG *et al.* 2013). Moreover, the modular design of this system allows the application to already existing transgenic animal lines. Most protein-protein interaction studies are based on biochemical *in vitro* assays, thereby neglecting the complex cellular environment they are usually embedded in. A newly developed, nanobody-based method to visualize protein-protein interactions is the fluorescent-three-hybrid (F3H) system (HERCE *et al.* 2013). In this system, a GFP binder is anchored to a specific subcellular compartment. The proteins of interest are then labelled either with GFP or RFP. The GFP-labelled protein will be recruited by the nanobody, and if the RFP-labelled protein interacts with the GFP labelled protein, a co-localization of the green and red fluorescent signal is observed. Additionally, drugs and their effect on protein complex formation can be tested. Thus, the F3H system allows dynamic *in vivo* monitoring of protein-protein interactions.

All in all, nanobodies, sdAbs, intrabodies, VHHs (or whatever you like to call them) are convenient, versatile reagents to study protein properties, structures, functions, and complexes within their native *in vivo* environment.

Additional protein binder scaffolds

Besides the VHH antibody scaffold described above, various other scaffolds based on non-immunoglobulin proteins have recently been developed.

One of these scaffolds is based on the fibronectin type III domain (FN3) and resembles the VHH at the structural level, as it also has a β-sandwich fold and CDR-like loops (KOIDE *et al.* 1998). This antibody-like proteins, also known as 'monobodies', were shown to be engineerable and have proven to be functional in living cells (GROSS *et al.* 2013). Additionally, they have the advantage of not being dependent on intramolecular disulphide bridges for proper folding, and, with a molecular mass of approximately 10 kD, they are slightly smaller than nanobodies (WOJCIK *et al.* 2010).

With a size of around 6-7 kD, the 'affibodies' are even smaller binding proteins. This class of protein binders is based on the immunoglobulin binding staphylococcal protein A and consist of three α-helices with a repetitive structure and no disulphide bonds. Randomization of 13 surface residues has led to selection of high affinity binders to various different antigens (NORD *et al.* 1997).

Reshaping the ligand binding pocket of lipocalin by site-directed random mutagenesis resulted in the isolation of highly specific binders against small molecules as well as polypeptide chains exposing from proteins. These so-called 'anticalins' fold into an eight-stranded beta-barrel structure with a molecular mass of 20 kD, which does not require any posttranslational modifications for proper folding (KORNDÖRFER *et al.* 2003; SKERRA 2008).

Other scaffolds that are used to generate protein binders with desired specificities are based on repeat proteins. In nature, repetitive protein domains are found in many proteins and are often involved in protein-protein interactions (BOERSMA and PLÜCKTHUN 2011). Remarkably, jawless fish are using leucine-rich repeats as surrogates for conventional antibodies in their adaptive immune system (PANCER and COOPER 2006). The protein repeat scaffold with the most progress in engineering and applications is based on ankyrin repeat proteins. The designed ankyrin repeat proteins (DARPins) are composed of structural repeats (one β-turn and two

anti-parallel α-helices). In general, DARPins contain four to six of these structural units, which results in an approximate protein size from 14 to 22 kD. A randomized library of defined residues was generated based on structural analysis of the repeats (Binz *et al.* 2003). The modular fashion in protein design gives the DARPins a highly adjustable, flexible structure allowing the isolation of a variety of specific binders (Plückthun 2015).

In general, all protein binder scaffolds should lead to products that are robust, soluble, monomeric, rather small, and easily expressible at large scales. Moreover, the availability or construction of large randomized libraries to perform directed evolution screens is a must. Overall, all the different protein binder scaffolds should be considered as complementary rather than competitive to find the best possible binder for each specific application.

The phage display technology

Phage display is a powerful method to select peptides or proteins (such as antibody fragments) that are binding to specific proteins (SCOTT and SMITH 1990; DE KRUIF et al. 1995; VAUGHAN et al. 1996). Therefore, large randomized libraries of proteins binders are constructed. The binding proteins or peptides to be selected are then displayed on the surface of phage particles. On the other hand, the target protein is immobilized to beads or specialized surfaces in vitro. The phages displaying polypeptides with the best binding ability to their target protein are then selected in a process called biopanning. The following chapters deal with the features and functionality of the phage display technology.

The components for phage display

Bacteriophages

Phages, more precisely bacteriophages, are viruses that infect bacteria. Upon infection they release their genome, in the form of either DNA or RNA, into their hosts. The phage genome is relatively simple and encodes mainly proteins. Some of these proteins are required for the replication of the phage genome itself, while others encapsulate the genome to form a new phage particle or help in the release of the particle from the host (RUSSELL et al. 2004). After the replication, the infectious particles leave their host to infect new bacteria.

For phage display, mainly filamentous phages are used, with M13 bacteriophage being the most widespread (see Figure 3). The M13 bacteriophage contains a single-stranded DNA genome with the length of approximately 6.4 kb encoding only eleven genes (gI-gXI) (VAN WEZENBEEK et al. 1980). The coat of M13 has a cylindrical structure and is composed of five different coat proteins (pIII, pVI-IX), encoded by the genes gIII and gVI-IX. pVIII is the most abundant coat protein and builds the filamentous capsule in a modular fashion, which allows the adjustment of the length of the phage capsule to the size of the genome. Five pIII proteins form the tip of the phage particle and are essential for the infectivity of the virion (RUSSELL et al. 2004). pIII interacts with F pili from E. coli, leading to a contraction of the pili. This contraction brings the phage in close proximity to TolA, a structural anchor complex in the cell membranes of E. coli. Attachment to these co-receptors are a shared feature of viruses, resulting in the release of the phage/virus genome into the bacteria (SLONCZEWSKI and FOSTER

2010). Consequently, M13 phages can only infect bacteria that have pili (F+ strains), such as TG1 and SS320 (Lucigen). In contrast to other phages, the filamentous bacteriophage M13 has a non-lytic lifecycle, thus bacteria are not lysed but constantly release virions. However, infection of *E. coli* leads to reduced growth and infected colonies can be identified as turbid plaques.
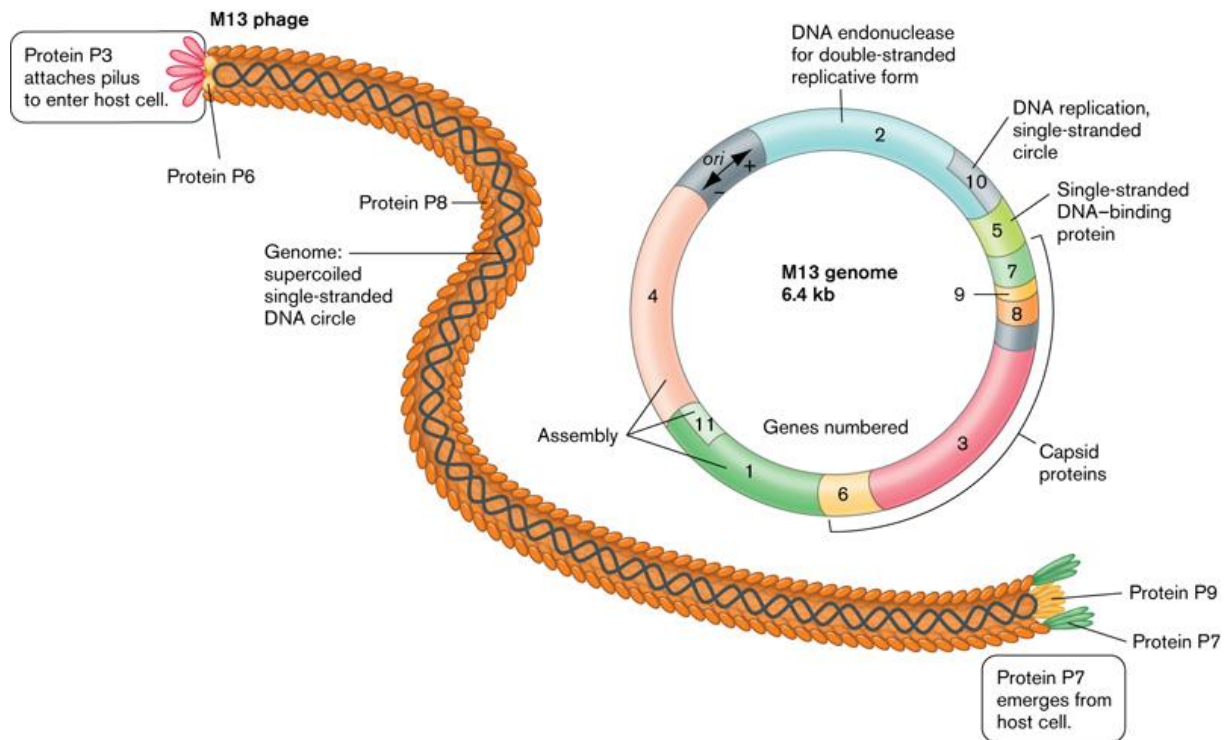
Figure 3 Structure and genome of M13 bacteriophage. On the left side: Drawing of a M13 bacteriophage with the capsule (P8, pVIII, respectively) in orange. pIII/P3 (red) is important for infectivity of the phage. On the right side: M13 genome. Picture adapted from SLONCZEWSKI and FOSTER 2010.
.

In principle, all five phage coat proteins can be and have been used to display polypeptides on the virion surface. However, pIII and pVIII are the most frequently used surface proteins for phage display. It has been demonstrated that large recombinant proteins can be inserted into pIII, which were displayed and accessible on the surface of the phage, without significantly affecting its infectivity (SMITH 1985). pVIII is mainly used to display short polypeptides, since addition of more than 8 amino acids prevent packaging and assembly of the phage. For improvement, pVIII has been specifically engineered allowing to display longer polypeptides (SIDHU *et al.* 2000). However, it is important to note that when displaying presumptive binders as a pIII-fusion, they are less abundant on the surface compared to pVIII-fusions. This has proven beneficial in selecting high affinity antibodies. Thus, mainly pIII-fusions are used in antibody phage display (RUSSELL *et al.* 2004).

Importantly, phage particles are very stable as they can withstand high temperatures, extreme pH, proteases, DNases as well as non-aqueous environment. Thus, due to their ability to display foreign proteins on their surface and the simple, very stable but flexible structure, M13 phages are the tools of choice in phage display.

Phagemids

For an efficient phage display, the development of so-called phagemids was tremendously important and helpful. Phagemids are vectors derived from filamentous phages. In addition to the replication origin of phages, they contain an origin of replication of a bacterial plasmid and an antibiotic resistance cassette for selection. However, phagemids usually do not contain all phage genes, thus they cannot form functional virions on their own. Interestingly, most of the cloning plasmids in use today are based on a phagemid scaffold (e.g. pBluescript or pUC).

For phage display, phagemids have been engineered to contain single engineered coat proteins. Signal peptides (for export, release) and a multiple cloning site (MCS) are fused to the N-terminus of the coat protein (QI *et al.* 2012). Gene libraries (e.g. antibody libraries) are inserted into the MCS, and subsequently recombinant proteins/polypeptides are displayed on the surface of phage particles as fusions to coat proteins (BARBAS *et al.* 1991; SIDHU 2001).

For our purposes, we used the phagemid pHEN2 (Griffiths 1. Library; GRIFFITHS *et al.* 1994) in which a synthetic VHH/nanobody library (called 'NaLi-H1') was inserted (Moutel et al., unpublished; see Figure 4). The pHEN2 phagemid has the following structure; a P*lac* promoter which is followed by an ORF with pelB leader sequence for export, then a MCS (in our case containing VHHs), which is followed by 6xHis and three myc tags and the coat protein gIII. Between the tags and *gIII*, a TAG amber codon is present. In suppressor strains *sup*E (such as TG1), TAG is read as a glutamic acid codon instead of being a stop codon. Therefore, the ORF is translated into pelB-VHH-His-myc-pIII. In a non-suppressor strain, the TAG is a stop codon and ORF is translated into pelB-VHH-His-myc and secreted into the periplasm after pelB cleavage. In other cases, this has proven to facilitate subsequent production and analysis of selected protein binders from a given library (HOOGENBOOM *et al.* 1991).
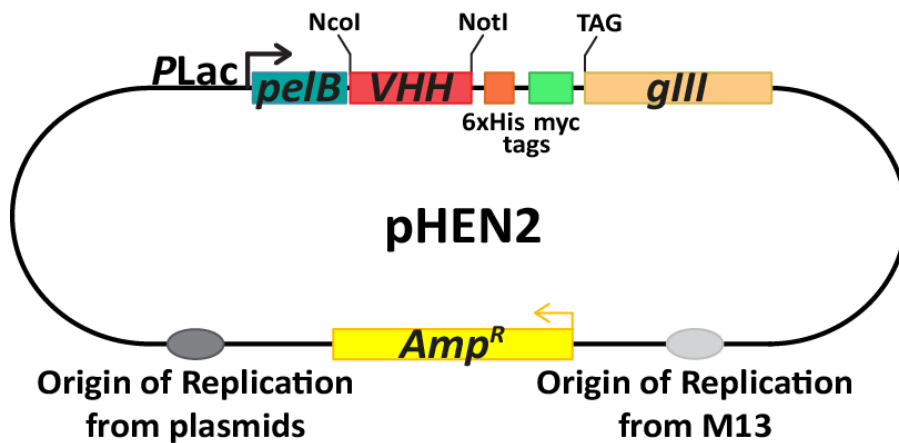
Figure 4 Illustration of the pHEN2 phagemid. It contains two origins of replication; one for bacteria (dark grey) and one for phages (light grey). To select infected bacteria, the phagemid contains an Ampicillin-resistance cassette (*Amp^R*). Synthetic, library *VHHs* (red) are cloned between a periplasm export signal peptide (*pelB*, marine blue) and the gene for the surface coat protein pIII (*gIII*, green) via NcoI and NotI sites. The VHH is followed by a 6xHis-Tag and myc-Tags. The *gIII* is preceded by amber codon TAG. Approximate size: 5000 bp.

Helper phages

As mentioned above, phagemids alone are unable to produce mature phages. To do so, they rely on so-called 'helper' phages. Helper phages provide the additional genes *in trans* required to produce an infectious virion. When infected, bacteria contain a phagemid and the helper phage genome that together now start to produce chimeric phage particles. The infected bacteria produce various phage particles composed of helper coat and helper genome, helper coat and phagemid, helper coat plus pIII-fusion protein from phagemid and helper genome. But most importantly, they produce phages with helper coat plus pIII-fusion protein from the phagemid which encapsulate the phagemid. These are the important ones for the phage display, since they link genotype to phenotype (see Figure 5 for overview of different phages produced in phage display).

Helper phages have been optimized in various ways to make phage display even more efficient. For example, by introducing a proteolytic sensitive site in pIII, M13 phages became non-infective after trypsin treatment (named KM13; Kristensen and Winter 1998). This way, proteolytic-sensitive pIII (coming from helper phage genome) are eliminated, and phages with 'wild type' pIII fused with the displayed peptide(from the phagemid) are selected during the selection process (Chasteen *et al.* 2006). Another improvement is the development of hyperphages, which have a wild type pIII but lack the gIII. This way the pIII comes exclusively from the phagemid, which increases the library peptides displayed on the phage surface from one to five binding domains (Rondot *et al.* 2001). Additionally, helper phage genomes contain

antibiotic-resistance cassettes to facilitate their selection and increase their production in bacteria.

Thus, to start a promising phage display screen, it is absolutely necessary to have; (1) a good and large library that is fused to a surface protein of the phage encoded on a phagemid, (2) helper phages that provide the rest of the phage genome needed to construct functional phage particles and (3) bacterial host cells producing phages.
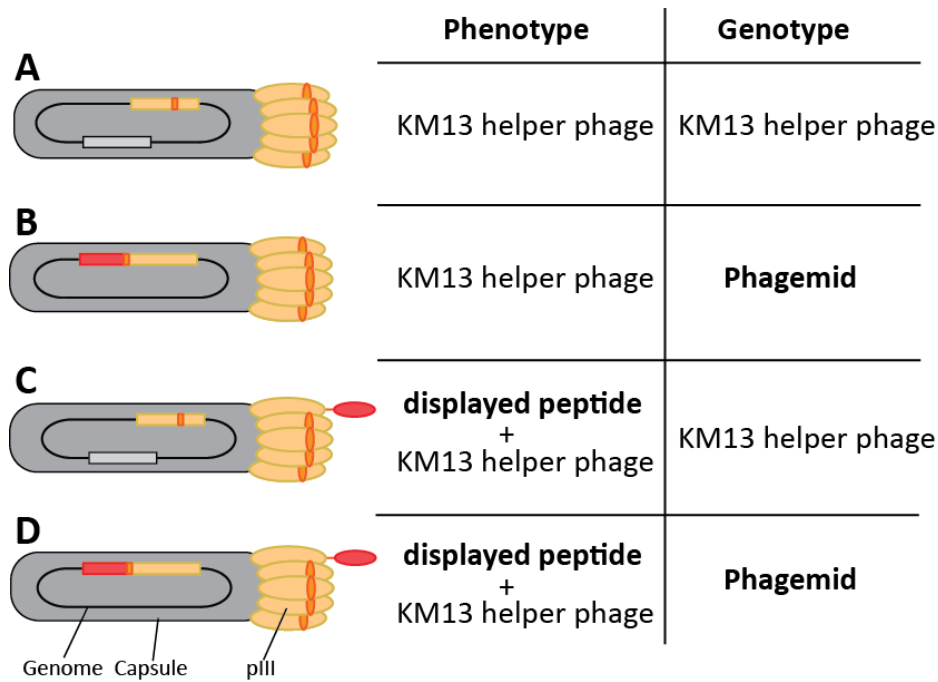
Figure 5 Different phages produced after helper phage infection of phagemid-containing bacteria. The capsule is depicted in grey, the genome as an oval and the pIII is in light brown, with a trypsin-sensitive site in orange. The displayed polypeptide is shown in red. **(A)** Typical helper phage, in phenotype and genotype. Is non-infectious after trypsin treatment. Helper phage genome contains a Kanamycin-resistance cassette (not shown). **(B)** Helper phage phenotype containing a phagemid. Can also be rendered non-infectious by trypsin treatment. **(C)** Has at least one pIII displaying a polypeptide from the library, but contains KM13 helper phage genome. Treatment with trypsin doesn't affect infectivity, since one functional pIII is sufficient. However, infected bacteria can be destroyed by Ampicillin. **(D)** Important phages for the phage display, since they combine phenotype (displayed protein, red) with genotype (phagemid). Keep infectivity after trypsin digestion, and can be positively selected by Ampicillin treatment.

The principle of phage display and biopanning

Figure 6 gives a schematic overview of a typical phage display screen. A comprehensive protocol is provided in the MATERIAL AND METHODS section. The first two steps in a phage display are performed in parallel; (1a) production of recombinant protein displaying phage library from a pre-existing phagemid-library with the help of bacteria and helper phages and (1b) immobilization of the target protein on either (magnetic) beads or on a dish. In a second step, phages and target protein are mixed and incubated to allow interaction (step 2; affinity binding, biopanning). Phages that do not bind to the target protein are removed by extensive washing (step 3). Afterwards, phages are eluted from the protein of interest (step 4), usually by increase of the pH. In the washing and elution steps, a wide variety of conditions can be implemented suiting the environment of the final application of the binder. For example, to isolate nanobodies against the cause of dandruff, panning was performed in the presence of shampoo (DOLK *et al.* 2005). In the case of 'intrabody' selection, panning is best done in isotonic conditions, such as one time phosphate buffered saline (1xPBS). When KM13 were used as helper phages, the eluted fraction is treated with trypsin to eliminate all non-displaying phages (step 5). Afterwards, bacteria with F(+) pili (in this case TG1) are infected with the eluted phages. Bacteria that received a phagemid are selected with antibiotics, resulting in a bacterial phagemid sub-library. This sub-library can then be reinfected with helper phages, and another round of panning is started. To enrich for high affinity, good quality binders, this procedure is usually repeated three to four times. At each round of panning, the protocol can be modified towards harsher binding conditions, such as additional washing steps or competitive binding to other proteins.
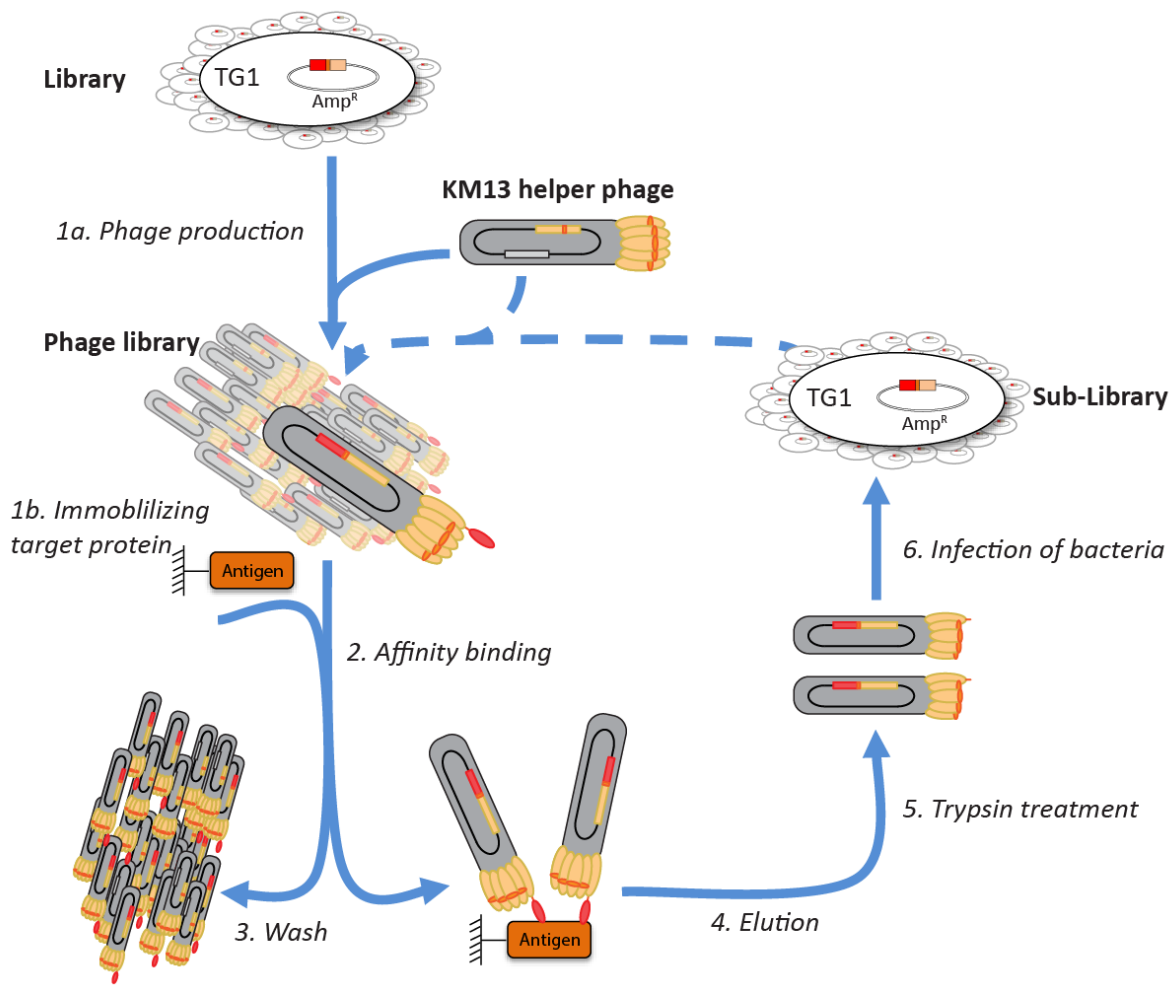
Figure 6 Workflow a typical phage display.

Aim of the project

So far, many nanobody-based applications are depended on the GFP-binding protein (also known as VHH4 or GBP; Rᴏᴛʜʙᴀᴜᴇʀ *et al.* 2006). Since many different fluorescent proteins are used in molecular, cellular and developmental biology, we tried to isolate intrabodies against other fluorescent proteins, which are not derivatives of GFP. Additionally, we wanted to isolate binders against an endogenous *Drosophila* protein to explore the possibility of modulating endogenous, untagged proteins directly. For this purpose, we chose Apterous (Ap), because at that moment there was no conventional antibody against Ap available. Additionally, phenotypes of *ap* mutants are easily visible (see PART I).

To isolate high-affinity nanobodies, we collaborated with Aurélien Olichon from the University of Toulouse/Cancer Research Center of Toulouse, who, has engineered a new library specialized for the isolation of intrabodies (Moutel et al., unpublished). Additionally, Aurélien Olichon is an expert in phage display. Thus, the plan was to perform phage display against additional fluorescent proteins and against Ap in Toulouse, and to screen the sub-libraries and validate and characterize the potential intrabodies with various assays in Basel. The ultimate goal of this project was to validate their functionality of such new nanobodies in model organisms used in the lab, such as the fruit fly *Drosophila melanogaster* and zebrafish *Danio rerio*.

As the field of nanobodies in developmental and cell biology is rather small and the topic rather new, it has not yet been established which is the best practice to isolate binders against a particular protein of interest. Thus, in parallel to the phage display done in collaboration with Aurélien Olichon, we tested the possibility to isolate specific nanobodies with the help of a commercial provider (Hybrigenics).

The results of these two approaches should provide insight in the optimal workflow in the isolation of new recombinant binders.

RESULTS AND DISCUSSION

As this project is still ungoing, only very preliminary data can be provided here. The results are directly discussed in the respective sections.

Preparations for phage display

Find appropriate antigens: Choice of suitable fluorescent proteins

The prominent green fluorescent protein (GFP) from the Pacific jellyfish *Aequorea victoria* was the first fluorescent protein discovered and characterized (Sʜɪᴍᴏᴍᴜʀᴀ *et al.* 1962; Mᴏʀɪsᴇ *et al.* 1974). GFP folds into a characteristic structure, the so-called β-barrel structure, which consists of eleven β-strands (Oʀᴍö *et al.* 1996). The chromophore, which is responsible for the fluorescence of the protein, is covalently bound in the center of the β-barrel (Yᴀɴɢ *et al.* 1996). Researchers quickly realized the great potential of fluorescent proteins, especially in applications in biology as a live marker for cells (Cʜᴀʟғɪᴇ *et al.* 1994; Tsɪᴇɴ 1998). Over time, more and more, structurally very similar fluorescent proteins have been discovered in other marine species, such as corals or anemones (reviewed in Aʟɪᴇᴠᴀ *et al.* 2008; Dᴀʏ and Dᴀᴠɪᴅsᴏɴ 2009). Despite the large distribution of fluorescent proteins in marine animals, their native function(s) remain(s) unclear, speculations reach from light protection to regulation of symbiosis with algal endosymbionts (Sᴀʟɪʜ *et al.* 2000; Fɪᴇʟᴅ *et al.* 2006).

Until today, many different fluorescent proteins have been engineered that work in distinct light spectra. However, when dealing with a fluorescent protein exhibiting a different spectra, it does not directly mean that the fluorescent protein is from another species. This is nicely demonstrated by the various derivatives of GFP that together cover a large spectra (Hᴇɪᴍ *et al.* 1995; Hᴀsᴇʟᴏғғ 1999; Pᴀᴋʜᴏᴍᴏᴠ and Mᴀʀᴛʏɴᴏᴠ 2008). As the chromophore defines the excitation and emission spectra, the overall β-barrel structure of these GFP derivatives remains largely unaffected. This is important to keep in mind when generating protein binders to 'different' fluorescent proteins. Due to their size and structure, VHHs can only bind to the surface of the β-barrel structure. Therefore, nanobodies against GFP, such as VHHGFP4 (Rᴏᴛʜʙᴀᴜᴇʀ *et al.* 2006), also recognize GFP derivatives, such as the yellow fluorescent protein

(YFP; Caussinus *et al.* 2011). Thus, it is extremely important to choose suitable fluorescent proteins as novel potential antigens.

We chose the antigen based on different criteria; they must exhibit distinguishable β-barrel surfaces, they should be used extensively and in many different applications, they should preferably be monomeric, exhibit an acceptable brightness, and they must be easily produced as recombinant protein in bacteria.

Based on these prerequisites, we decided to try to isolate single domain antibodies against mCherry (Shaner *et al.* 2004), mKate2 (Shcherbo *et al.* 2009), and mTFP1 (Ai *et al.* 2006). Figure 7 shows a phylogenetic tree of the chosen fluorescent proteins and some of their derivatives and their relationship to GFP.
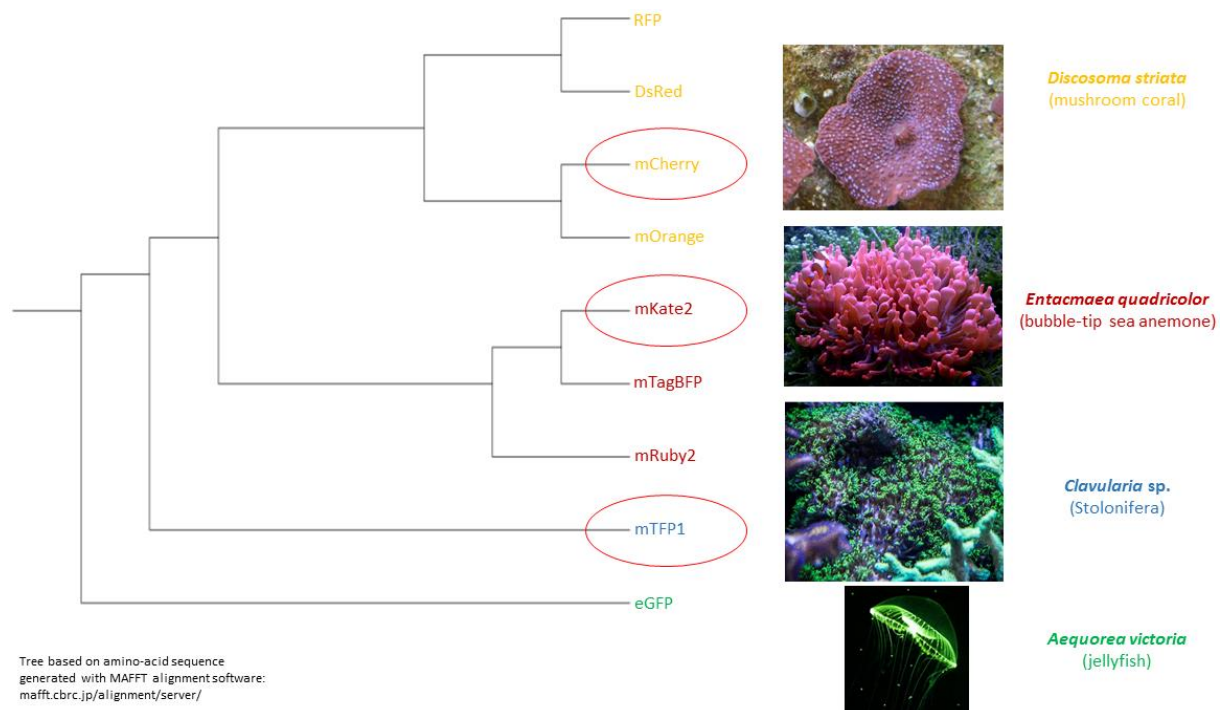
Figure 7 Phylogenetic tree of selected fluorescent proteins derived from different species. Red encircled proteins were chosen to be used in phage display to find new nanobodies. Image credits: *Discosoma striata* from www.ultimatereef.net; *Entacmaea quadricolor* from www.korallenfarm.de; *Clavularia* sp. from home aquarium of Dimi Bieli; *Aequorea victoria* from mcdb-webarchive.mcdb.ucsb.edu.

mCherry was generated using directed evolution starting from DsRed fluorescent protein, which had been isolated from the mushroom coral *Discosoma striata* (Shaner *et al.* 2004). mCherry, a red fluorescent protein, is used in many application in cell and developmental biology as a fusion protein or reporter.

mKate2 is derived from the bubble-tip sea anemone *Entacmaea quadricolor* fluorescent protein eqFP578 (SHCHERBO *et al.* 2007). Its emission maximum is at 633 nm wavelength, and it is thus considered as a far-red fluorescent protein. In our lab, mKate2 is frequently used in transgenic zebrafish and flies (LENARD *et al.* 2015).

mTFP1 was engineered using directed evolution starting from the colonial soft coral *Clavularia* tetrameric fluorescent protein cFP484 (AI *et al.* 2006). It has been reported to be one of the brightest and most photostable fluorescent proteins characterized so far. However, it is not used very often yet in cell biology.

In general, the prefix 'm-' stands for monomeric, so all of the chosen fluorescent proteins had been engineered to work in monomeric conformations. This is an important consideration when these proteins are used as fluorescent fusion partners to study the function of a protein of interest. For a comprehensive table of fluorescent proteins and their features see http://nic.ucsf.edu/dokuwiki/doku.php?id=fluorescent_proteins.

Expression of antigens in *E. coli*

As a next step, we tested whether the selected fluorescent proteins and the endogenous *Drosophila* protein Apterous (Ap) can be produced in bacteria and purified with the appropriate fusion tags (His-, Streptavidin-, and Chitin binding tags) needed for the subsequent phage display. For Ap, we decided to express only the homeodomain (apHD), due to expression problems encountered with full length Ap protein (see PART I). For expression, purification and loading-on-beads protocols see MATERIAL AND METHODS.

As shown in Figure 8, we successfully expressed and purified all the selected fluorescent proteins. Moreover, we made sure that all the fluorescent proteins and the apHD could be loaded to the specific beads and stayed bound to the beads for significant rounds of washing (min. 30 rounds), as it has been performed during phage display (Figure 8B and C, and data not shown).

In summary, all the tests were positive and we could proceed with the phage display.
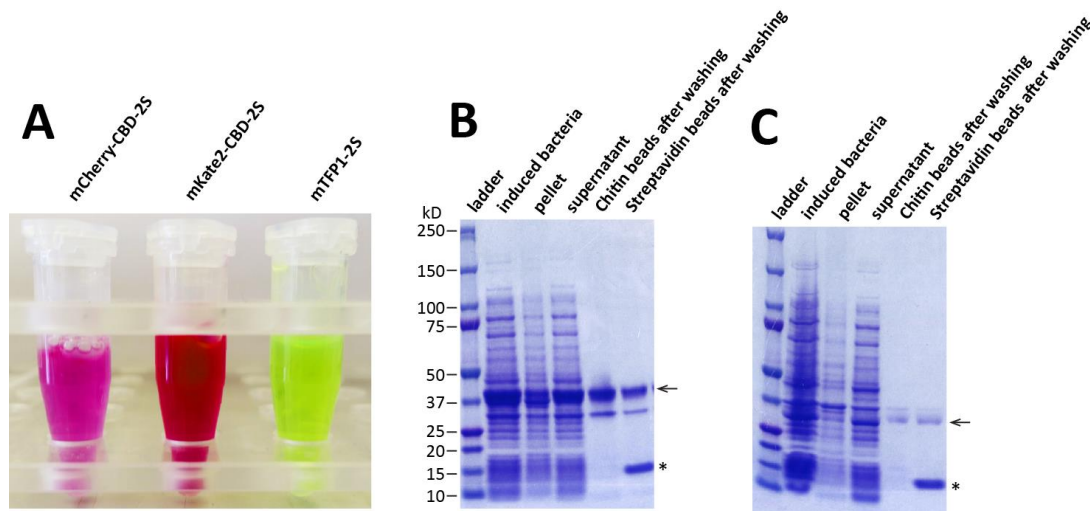
Figure 8 Expression of fluorescent proteins and loading on beads. (A) Picture of fluorescent proteins in solution. On the left: mCherry with Chitin binding domain and two Streptags (mCherry-CBD-2S). In the middle: mKate2-CBD-2S is darker red than purple mCherry protein. On the right: mTFP1 shows greenish fluorescence. **(B)** SDS-PAGE stained with Coomassie. lane 1: ladder (Precision Protein Plus Dual Color (Biorad)); lane2: whole bacterial extract after induction with ITPG; lane 3: bacterial pellet after lysis and centrifugation; lane 4: supernatant of bacterial lysate; After loading mKate2 to the Chitin or Strptavidin beads, respectively, beads were washed 30 with Phage display Washing Buffer (see MATERIAL AND METHODS). Afterwards, beads were cooked in loading buffer for 2 min on loaded in lane 5 and 6. mKate2-CBD-2S has a molecular weight of 36.9 kD (arrow). Monomeric streptavidin from the Streptavidin beads has a size of approx. 15 kD (asterisk). **(C)** Testing beads bindings of apHD. Same loading order as in B. apHD-CBD-2S has a molecular weight of 28.9 kD (arrow).

Phage display

To isolate potential nanobodies against the proteins mCherry, mKate2 and apHD, we performed 4 rounds of phage panning in the Laboratory of Prof. Gilles Favre, University of Toulouse, in collaboration with Aurélien Olichon. For a complete protocol see MATERIAL AND METHODS.

To immobilize the antigens during phage display, we used magnetic Chitin (NEB) as well as magnetic Streptactin (IBA) beads. For each round of phage display, magnetic beads were alternated to avoid enrichment of nanobodies that bind to the beads. To eliminate binders that might bind to both mCherry and mKate2 proteins, we performed a counter-selection during the second, third and fourth round of panning. Therefore, the phage sub-library from the previous mCherry panning round was pre-incubated with mKate2 before the next round of panning with mCherry protein. Conversely, the mKate2 binders were incubated with potential mCherry binders. For apHD, we did not perform a counter-selection, due to lack of a suitable counterparts to the apHD.

During the course of the phage display, there is no real control to check for the potential enrichment or quality of the binders. Thus, we collected and stored the sub-libraries from each round of panning to have the possibility to return to an appropriate step and repeat the panning again.

A possibility to test the enrichment of potential binders by the end of 3 or 4 rounds of phage display is the sequencing of individual clones. The analysis of the clone sequences can give an insight about enrichment of clones and diversity of the sub-libraries. If several clones show the same sequence, this might be a hint that enrichment worked. Thus, we decided to sequence 30 random clones from each sub-library.

Of the 30 clones tested from the mKate2 binder sub-library, we found 4 independent clones with identical sequences (see Figure 9). For the mCherry and apHD, we identified 30 individual clones.

These results suggest that enrichment worked for one of the targets (mKate2), however, there is still a large diversity seen in all the different sub-libraries.

Figure 9 DNA sequence alignment of various potential mKate2 binders. DNA sequence of the first two CDR loops is shown. 4 clones (blue circles) have identical sequences.

Additionally, to compare and test the production of specific VHH binders via an external provider, we decided to send recombinant mTFP1 to Hybrigenics (Paris, France). As a standard protocol, they first performed 3 rounds of phage display.

Screening sub-libraries for specific intrabodies

As our main goal was to find new nanobodies that also work in the intracellular environment ('intrabodies'), we decided to directly screen each 4th round sub-library with a simple co-localization assay in cell culture using HeLa cells (SCHERER *et al.* 1953; ROTHBAUER *et al.* 2006). In order to visualize co-localization, the antigen (bait) and the putative binder (prey) were fused

to different, spectroscopically separable fluorescent proteins and were co-expressed in cells. Interaction of binder and antigen was visualized by the merge of the two fluorescent signals. For our purposes, we fused the VHHs from the sub-library with a fluorescent marker protein and cloned it into a mammalian expression vector (see MATERIAL AND METHODS). Potential mCherry, mKate2 and Apterous binders were fused with GFP, whereas mTFP1 binders were fused with mCherry. In other words, we produced 'chromobodies' of the potential binders.

Since all of our target proteins, with the exception of Ap, are fluorescent proteins, it was important to use an appropriate, spectroscopically separable marker protein for the respective bait and prey pairs. The fluorescent target proteins mCherry, mKate2 and mTFP1 were fused to H2B, which translocated the proteins to the cell nucleus (BUSTIN *et al.* 1975). To screen for potential Ap binders, we fused the Ap homeodomain (apHD) with mKate2, to visualize its distribution inside the cell. Interestingly, the heterologous fusion protein containing the apHD translocated to the nucleus, probably due to the nuclear location signal (NLS) present in the homeodomain (data not shown).

In a first round, we tested 48 individual clones from each sub-library (mKate2, mCherry and Ap) using this localization assay. Labelled VHHs (chromobodies) were co-transfected with their respective potential target protein, which was enriched in the nucleus. A potential interaction was scored by using a special script from the program CellProfiler that measures co-localization (Alessandra Vigano, personal communication).

In a second round, potential candidates were re-examined by 'eye' with higher resolution and magnification using a confocal microscope. Additionally, we tested the distribution of each VHH without the bait (respective antigen). For mCherry, mKate2 and mTFP1 binders we also tested whether the potential binders can be re-localized to a different cellular compartment (i.e. the cell membrane, by fusing a CAAX localization signal to the bait).

Validation of α-mCherry intrabodies in cell culture

In the sub-library of mCherry binders, we found 2 very promising hits (clone #10 and #23), called VHHmCherry10 and VHHmCherry23. When these VHH-GFP fusions were expressed in the cells without a bait, the GFP signal was seen throughout the cytoplasm, with slightly stronger signal in the nucleus (Figure 10A and D). Upon co-expression with H2B-mCherry, the

α-mCherry intrabodies re-distributed to the nucleus, where a strong co-localization was observed at distinct dots (presumably nucleoli) within the nucleus (arrows in Figure 10B and E). When the VHHs were co-transfected with a construct that translocated mCherry to the plasma membrane (mCherry-CAAX), accumulation of the GFP signal from binders VHHmCherry10 and VHHmCherry23 was also seen at the plasma membrane (Figure 10C and F). However, VHHmCherry23 also showed a slight nuclear signal.
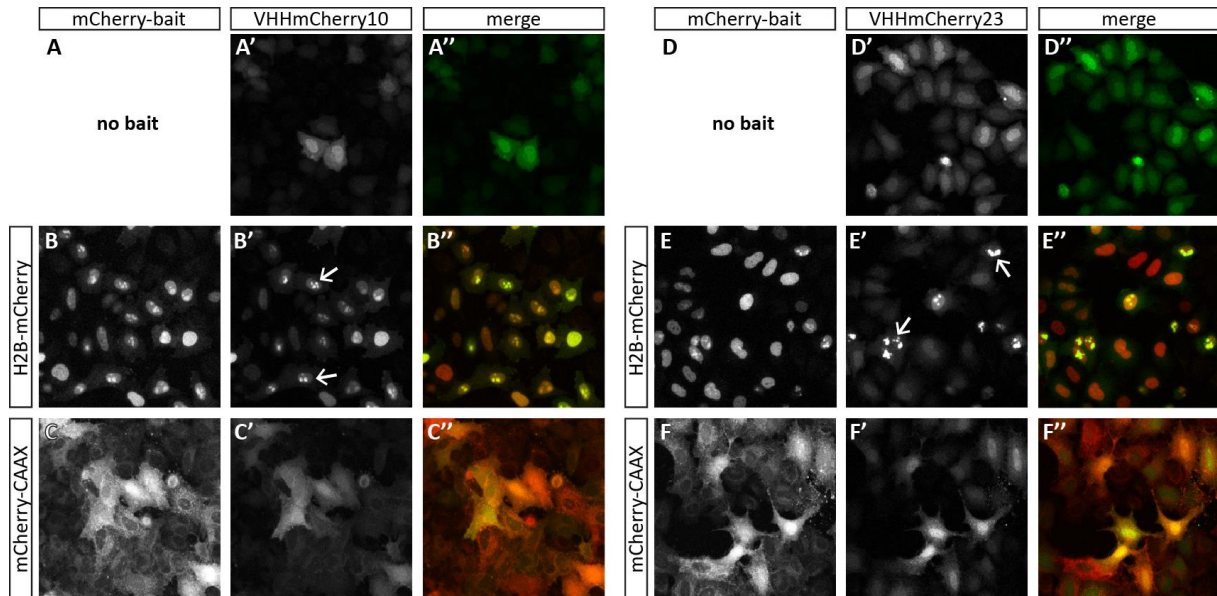
Figure 10 VHHmCherry10-GFP and VHHmCherry23-GFP expressed in HeLa cells. (A and D) When clones #10 and #23 are expressed alone without mCherry baits, the GFP signal is more or less evenly distributed in the cell, with a bit stronger signal in the nucleus. **(B and E)** Co-expression of VHHmCherry10-GFP and VHHmCherry23-GFP, respectively, with H2B-mCherry leads to strong accumulation of GFP in the nucleus and nucleoli (arrows). **(C and F)** Co-expression with mCherry-CAAX re-distributes GFP signal to the cell membrane. Pictures taken by Alessandra Vigano.

Validation of α-mKate2 intrabodies in cell culture

As for mCherry, we also found two promising mKate2 binders, VHHmKate2_2 and VHHmKate2_16. Interestingly, clone #16 is one of the 4 independent clones that shared the same DNA sequence (see above). GFP signal was detected in the cytoplasm and nucleus, when these chromobodies were expressed in cells without a bait (Figure 11A and D). Upon co-expression with H2B-mKate2, the GFP signal from α-mKate2 chromobodies was enriched in the nucleoli within the nucleus (arrows in Figure 11B and E). Co-expression with mKate2-CAAX

led to an accumulation of the GFP signal from binders #2 and #16 at the plasma membrane (Figure 11C and F). However, both mKate2 intrabodies still showed a nuclear fluorescent signal.
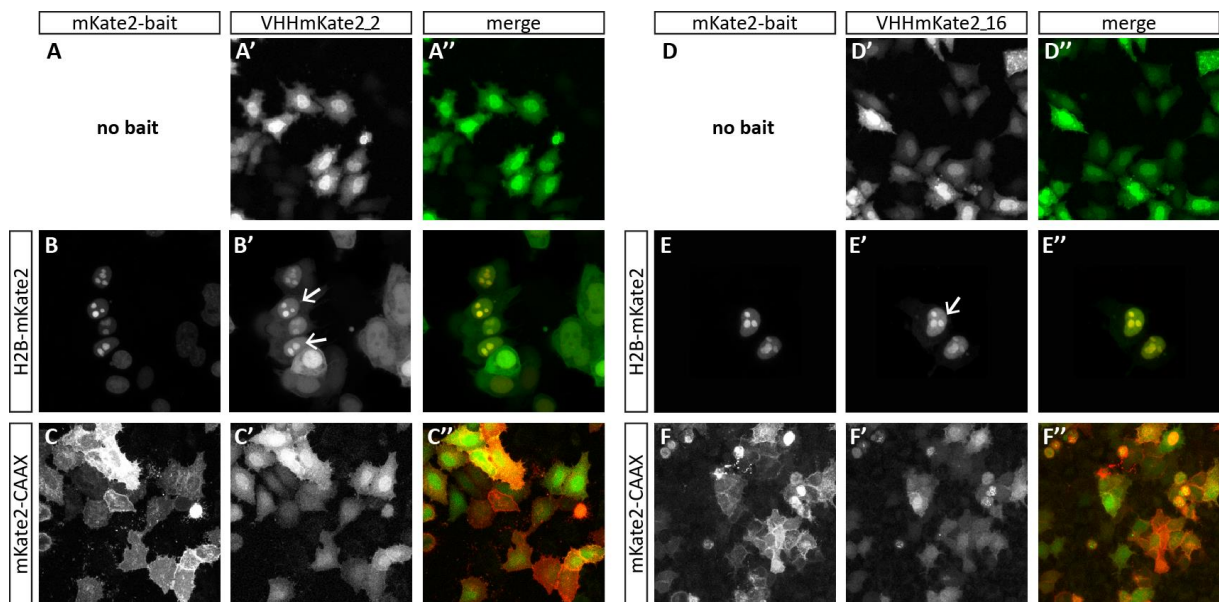
Figure 11 VHHmKate2_2-GFP and VHHmKate2_16-GFP expressed in HeLa cells. (A and D) When clones #2 and #26 are expressed alone without mKate2 baits, the GFP signal distributed in the cell, with a stronger signal in the nucleus. **(B and E)** Co-expression of VHHmKate2_2-GFP and VHHmKate2_2-GFP, respectively, with H2B-mKate2 leads to strong accumulation of GFP in the nucleus and nucleoli (arrows). **(C and F)** Co-expression with mCherry-CAAX re-distributes GFP signal to the cell membrane. Most pictures taken by Alessandra Vigano.

Taken together, these results demonstrate that we could successfully enrich for specific binders against mCherry and mKate2 using phage display and isolate functional intrabodies with a medium-throughput cell culture-based assay.

α-Apterous binders in cell culture

For binders against the homeodomain of Ap, the situation was a bit more difficult. Although we identified many potential binders in a first round (with co-expression of apHD-mKate2 bait), none of them could be validated in further assays. When the presumable hits were expressed without a bait, they still showed a very high nuclear localization with similar sub-nuclear distribution patterns (data not shown). The three 'most' promising clones are shown in Figure 12.

At this point, it is not clear why we encountered the problem with the nuclear localization of the Ap binders when no bait was present. Since we have not performed a counter-selection

for apHD, we cannot exclude that the enriched VHHs also interact with other homeodomain proteins, of which there are plenty in every cell type. Thus, even a low unspecific binding to other homeodomains might be enough in sum to re-locate the putative Apterous homeodomain binders to the nucleus. Trials to mislocalize the apHD to other cellular compartments, such as the cell membrane, have failed (data not shown).

All in all, the results from the cell culture assay for potential Ap binders were not satisfying. Still, we decided to test some of the binders *in vivo* in transgenic *Drosophila* lines (see below).
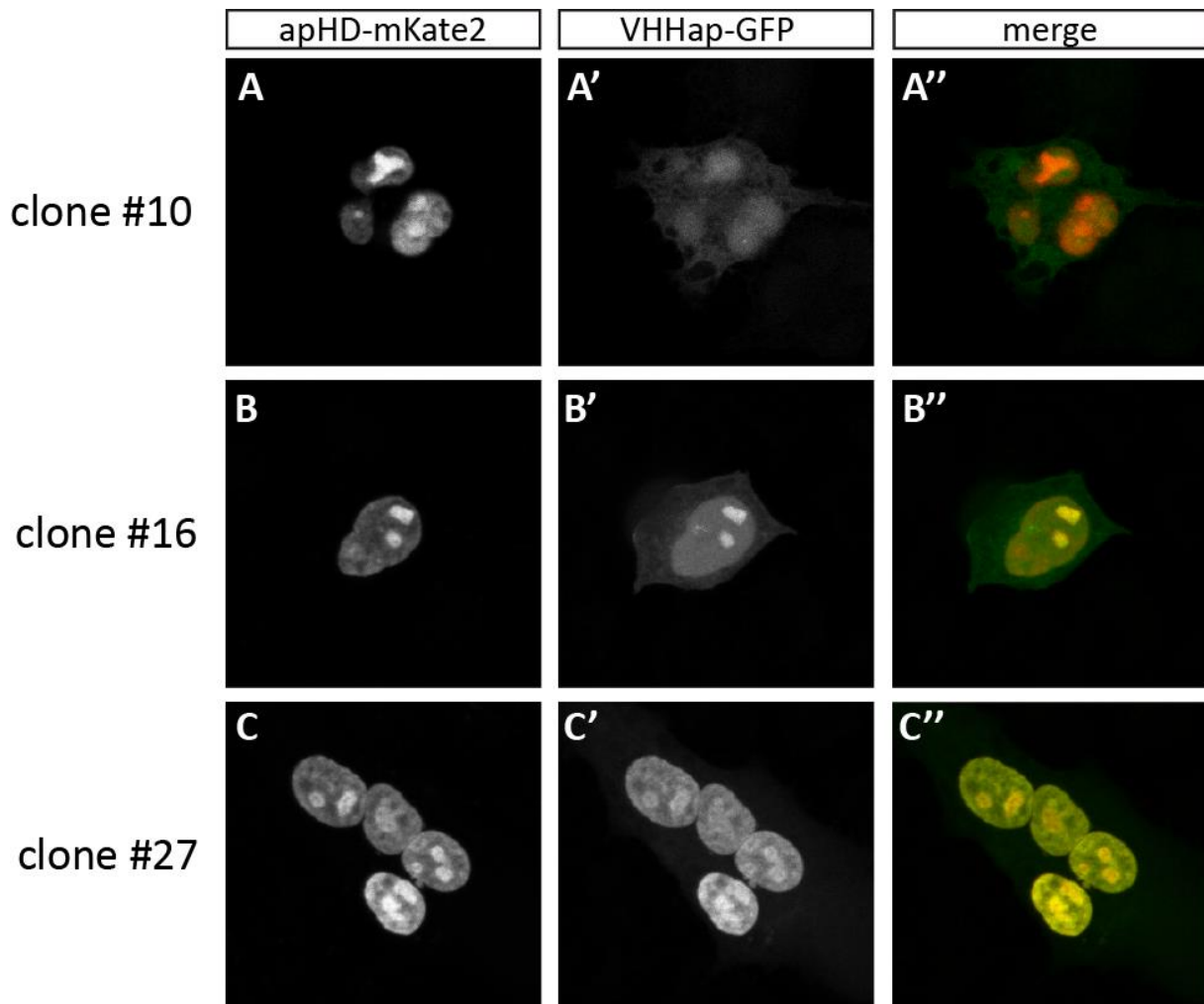
Figure 12 Three potential Apterous binders expressed in HeLa cells with apHD-mKate2 bait (A) clone#10 fused with GFP shows weak co-localization with mKate2 signal from apHD-mKate2. **(B and C)** clone#16 and #27 show stronger co-localization with apHD-mKate2.

Testing α-mTFP1 VHHs from Hybrigenics in cell culture

After 3 rounds of phage display using mTFP1 as a bait, Hybrigenics performed a high-throughput *in vitro* ELISA test to find the strongest binders to mTFP1. In total, they sent us 7

putative α-mTFP1 VHHs. However, when these VHHs were expressed in cell culture as chromobodies, the expression levels were very low, in some cases no expression was observed. Moreover, all VHHs tested showed strong aggregation in transfected cells. Upon co-expression with H2B-mTFP1 or mTFP1-CAAX, no re-localization was detected and the VHHs still aggregated (Alessandra Vigano, personal communication).

Thus, none the nanobodies provided by Hybrigenics was working as a functional intrabody. This result was very unsatisfying and might be due to the fact that the 'binders' were isolated with an *in vitro* test.

Therefore, Hybrigenics offered us to perform a yeast-two-hybrid (Y2H) screen with mTFP1, starting with the sub-library from the first round of phage display. After the Y2H screen, they sent us the two most highly enriched VHHs.

The two clones (VHHmTFP1_B69 and VHHmTFP1_B334) were fused with mCherry to generate chromobodies for expression and detection in cell culture. When these VHHs were expressed without a bait, the mCherry signal was seen throughout the cytoplasm and nucleus, without visible aggregations (Figure 13A and D). Co-transfection with H2B-mTFP1 re-localized the α-mTFP1 nanobodies to the nucleus (Figure 13B and E). When VHHmTFP1_B69 or VHHmTFP1_B334 were co-expressed with mTFP1-CAAX, accumulation of the mCherry signal was also observed at the plasma membrane (Figure 13C and F).
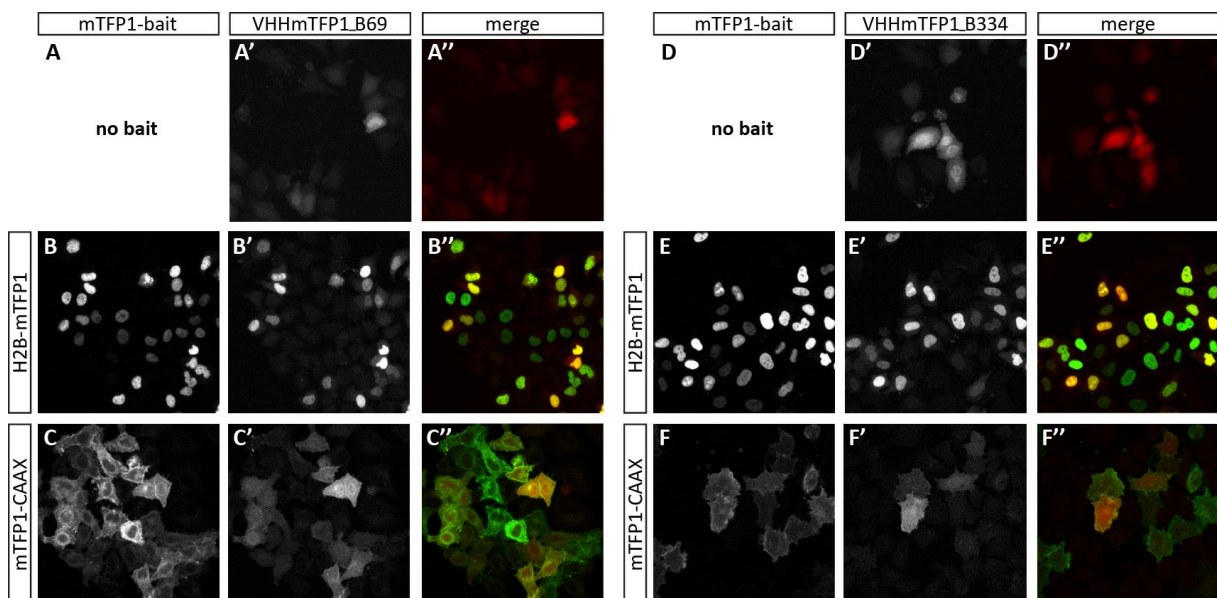
Figure 13 VHHmTFP1_B69-GFP and VHHmTFP1_B334-GFP expressed in HeLa cells. (A and D) When clones B69 and B334 are expressed alone without mTFP1 baits, the GFP signal distributed in the cell, with a stronger signal in the nucleus. **(B and E)** Co-expression of VHHmTFP1_B69-GFP and VHHmTFP1_B334-GFP, respectively, with H2B-mTFP1 leads to strong accumulation of GFP in the nucleus **(C and F)** Co-expression with mTFP1-CAAX re-distributes GFP signal to the cell membrane. Pictures taken by Alessandra Vigano.

These results show that the nanobodies provided after the Y2H screen functionally worked as specific intrabodies. Moreover, these results also demonstrate that functional intrabodies can be isolated by other means than phage display.

Testing intrabodies *in vivo*

After the successful selection and validation of potential new binders to fluorescent proteins, we also investigated whether the intrabodies work *in vivo*. As a first trial, we tested the VHHmCherry10 intrabody in live zebrafish. To this end, we injected plasmids coding for VHHmCherry10-GFP and the respective baits into early stage fertilized zebrafish embryos. After 24 hours of development, fish embryos transiently expressing intrabody and bait were analyzed under a confocal microscope. We analyzed expression in epithelial cells due to their relative large size.

When VHHmCherry10-GFP was injected without a bait, GFP signal was detected throughout the cytoplasm and nucleus (Figure 14A). Upon co-injection with H2B-mCherry, the GFP signal the mCherry10 intrabody was enriched in the nucleoli within the nucleus (Figure 14B). Co-injection with mCherry-CAAX led to an almost perfect co-localization of the GFP and mCherry signals at the plasma membrane (Figure 14C). This observation confirms the positive results from the cell culture assay and demonstrates that the VHHmCherry intrabodies can also be applied in live animals, as previously shown for other chromobodies (PANZA *et al.* 2015).
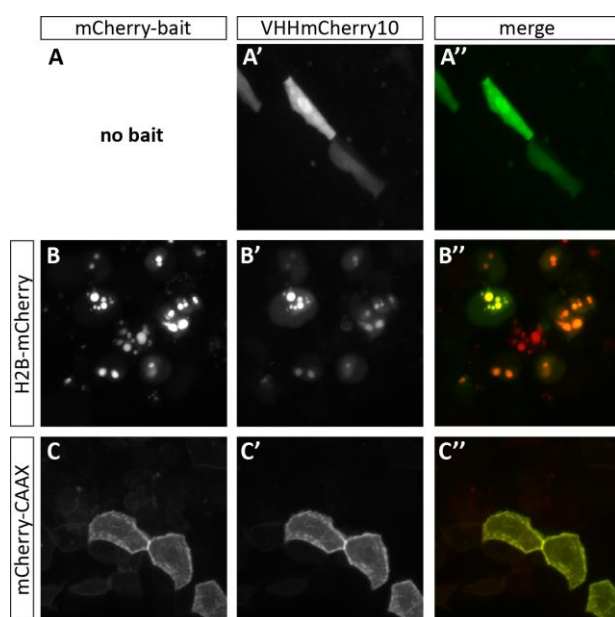
Figure 14 Transient expression of VHHmCherry10-GFP in live zebrafish. (A) When clones #10 is expressed alone without mCherry baits, the GFP signal is more or less evenly distributed in the cell, with slightly stronger signal in the nucleus. **(B)** Co-injection of VHHmCherry10-GFP with H2B-mCherry leads to strong accumulation of GFP in the nucleus and nucleoli **(C)** Co-expression with mCherry-CAAX completely re-distributes GFP signal to the cell membrane.

As described above, the cell culture experiments with α-Ap binders were not very encouraging. Nevertheless, we decided to test the 3 most promising binders (clone #10, #16 and #27) directly in *Drosophila*.

Therefore, we engineered chromobodies by cloning the VHH coding sequences in frame with GFP on a UAS-driven attB vector (see MATERIAL AND METHODS of PART I). Additionally, we tested interaction of putative binders with apHD using a degradation assay (CAUSSINUS *et al.* 2011). For this purpose, VHHs were fused with the fly F-Box protein Slmb (JIANG and STRUHL 1998), and cloned in a UAS-controlled attB plasmid. The constructs were injected into embryos containing an attP landing site on the third chromosome (BISCHOF *et al.* 2007). Transgenic lines were established and crossed with different Gal4 drivers available. As mis-regulation of *ap* leads to visible wing phenotypes (see PART I), we thought that lowering the protein level with a degradation assay might provoke similar wing phenotypes. Along this line, degradation of Ap-GFP with α-GFP nanobodies led to measurable reduction of the GFP signal and caused dramatic wing phenotypes (CAUSSINUS *et al.* 2011).

Ubiquitous expression of chromobodies with tubulin-Gal4, no effect on wing development was observed, and no re-localization of putative binders was seen in Ap-positive cells (e.g. cells in the dorsal compartment of the wing disc). Ap is only expressed at low levels (DAINES *et al.* 2011), thus it might be that Ap levels are too low to attract enough chromobodies to visualize the distribution of Ap protein in cells. However, also the ubiquitous expression of slmb-VHH fusion proteins had no effect on wing development (data not shown).

As *ap* is a recessive gene, even an assumed protein level reduction of 50% might have no effect. Thus, we decided to use an Gal4-enhancer trap in *ap* (apGal4; CALLEJA *et al.* 1996), which has shown to be a strong hypomorphic allele of *ap* (MILÁN and COHEN 1999). Hence, in this genetic background the Ap protein level is already strongly reduced compared to wild type levels.

Only flies expressing the Slmb-VHH fusion construct of clone #27 (apGAl4>UASslmbVHHap27) showed a weak wing phenotype (Figure 15A and D). The other two VHHs tested displayed normal wings (data not shown). Wings of apGAl4>UASslmbVHHap27-flies showed bending and additional veins in the wing blade. However, none of the many known *ap* alleles shows a similar wing phenotype.

Thus, based on the results obtained in these experiment, we conclude that we failed to isolate high affinity binders recognizing the Ap HD so far. Further screening might be necessary to find functional α-Ap nanobodies.
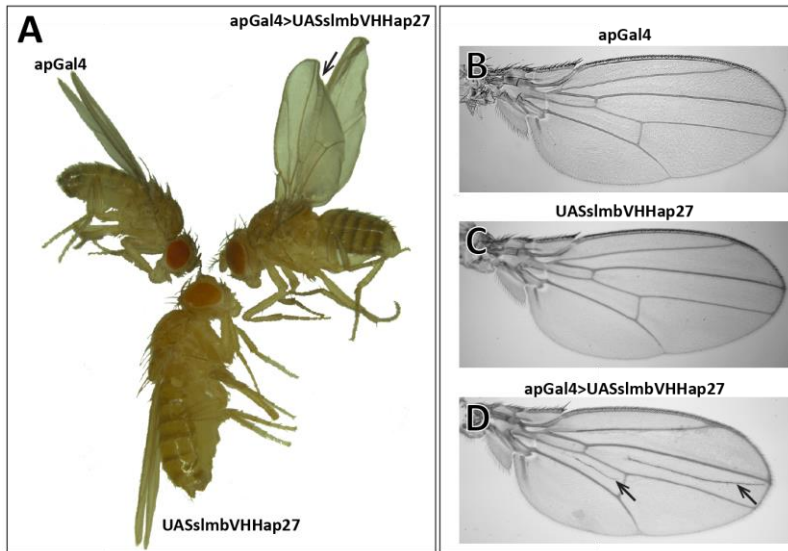
Figure 15 Effect of Slmb-VHHap27 on wing development. (A) Compared to wild type wings of apGal4 and UASslmbVHHap27 flies, wings of apGal4> UASslmbVHHap27 flies are bend downwards and have a nick at the tip of the wing (arrow). **(B and C)** apGal4 and UASslmbVHHap27 wings are wild type. **(D)** apGal4> UASslmbVHHap27 wings show additional veins between wing veins L5 and L4, and L4 and L3 (arrows).

Sequence alignment of positive intrabodies against fluorescent proteins

After the validation of the positive intrabodies against the different fluorescent proteins, we compared the respective amino acid sequences. Hence, we aligned the amino acid sequences using the CLUSTAL omega alignment program (SIEVERS *et al.* 2011).

The mTFP1 binders clone B69 and B334 are basically identical, except one amino acid substitution in the framework 1 at position 13 (F to S). They share the same sequences in the CDR loops (Figure 16). Thus, it is very likely that B69 and B334 recognize the same epitope on the surface of mTFP1. All the other VHHs showed different amino acid sequences in the CDR loops, which is a strong sign that they also interact with different epitopes. Interestingly, the sequence of the framework of the binders selected by us and Hybrigenics was identical.

```
                 CDR1              CDR2                       CDR3
VHHmCherry10 ...SGYTSFWDTMGW...VSAISWWHDEVAYYADSVK...YYCAVWSRK---------RKYHYWG
VHHmCherry23 ...SGGTYKGTNMGW...VSAISYTNSNAEYYADSVK...YYCAYYFDDL---TLSLPTPNMYWG
VHHmKate2_2  ...SGSTSASDAMGW...VSAISSHNGIELYYADSVK...YYCADIPWEYANNARTKHRPGAYWG
VHHmKate2_16 ...SGRTYFNTIMGW...VSAISSEPSPYTYYADSVK...YYCAPSYLTTGT------FIPMYWG
VHHmTFP1_B69 ...SGDTSYWDDMGW...VSAISWGWGLSTYYADSVK...YYCARGLLMGSIWRTYHESKAHYWG
VHHmTFP1_B334...SGDTSYWDDMGW...VSAISWGWGLSTYYADSVK...YYCARGLLMGSIWRTYHESKAHYWG
                 ** *      **********    **************            ***
```

Figure 16 Multiple sequence alignment of the different fluorescent protein binders. Only sequences of CDR loops are shown.

Conclusions and Outlook

Using phage display and medium-throughput selection in cell culture, we, in collaboration with Aurélien Olichon in Toulouse, have successfully isolated various intrabodies that show strong interaction with mCherry or mKate2 fluorescent proteins in cell culture and also *in vivo* in zebrafish embryos.

On the other hand, phage display and ELISA screening, performed by Hybrigenics, failed to isolate functional intrabodies, due to aggregation problems. However, we have also used the same service at Hybrigenics to produce binders against the Drosophila BMP2/4 homologue Dpp (Shinya Matsuda, personal communication, data not shown). For this project, we have obtained functional binders. However, these binders were only tested in the extracellular environment. Thus, it appears that selection via the *in vitro* ELISA assay is favoring the isolation of nanobodies that only fold properly outside the cell.

A second attempt to isolate mTFP1 binders was done by Hybrigenics using the yeast-2-hybrid system. From this screen, we obtained two almost identical VHHs that nicely worked as intrabodies. Though, it seems that the diversity from the Y2H screen is lower than compared to the phage display, which is possible due to the fact that the size of the library is limited by yeast transformation.

Every protein of interest shows different, unique features. Thus, before every attempt to isolate functional binders against proteins is initiated, several key aspects have to be considered. Important questions that have to be answered beforehand are the following: 1) Is the final application of the binder intra- or extracellularly; 2) Can the recombinant protein of interest be expressed as a soluble protein in heterologous systems; 3) Has the protein of interest a mono- or multimeric structure?

Is the final application of a binder in the extracellular space, it might be favorable to use conventional phage display and high-throughput ELISA selection. The same holds true if the antigen is a multimeric protein. For easily expressible monomeric proteins (e.g. fluorescent proteins), the Y2H screen provides a valuable alternative to the labor-intensive phage display technique.

When using an external provider, the costs are always an issue. In this regard, it is important to mention that the Y2H screen is considerably cheaper than the phage display.

As described in the previous section, we failed to isolate high-affinity binders to the endogenous *Drosophila* protein Apterous. As the preparations for the phage display were successful and the phage display for Ap was performed in parallel to the mCherry and mKate2 displays, it seems rather unlikely that something went wrong before and during the panning steps. Based on these presumptions, it is very probable that the sub-library of the 4th round of the Ap phage display also contains positive Ap binders. The major problem seems to have been the screening in the HeLa cells. A possible way to screen would be directly in *Drosophila* using a degradation assay. However, this approach would be very tedious and involve a lot of work. Initially, our goal was to find an α-Ap antibody to perform conventional immunostaining in flies and we started this project in parallel to the production of conventional antibodies. As by now, we have an excellent polyclonal α-Ap antibody (see PART I), it might not be very helpful to pursue the development of α-Ap nanobodies.

Thus, as a next step, we will focus on the characterization of the functional intrabodies against the fluorescent proteins mCherry, mKate2 and mTFP1. We will test whether the specific binders show cross-reactivity with other fluorescent proteins. Additionally, we will perform *in vivo* experiment for all the binders in zebrafish and flies. Together with the Biophysics Core Facility of the Biozentrum (University of Basel), we are planning to perform affinity measurements of each binder to its respective target. To understand the exact interaction of binder and antigen, we will try to produce high resolution protein structures in collaboration with the Group of Prof. Timm Maier of the Structural Biology Department (Biozentrum, University of Basel).

With these results, we will be able to extend the nanobody toolbox with new and well characterized binders, which will be beneficial for future (more sophisticated) experiments in cell and developmental biology.

MATERIAL AND METHODS

Plasmids

All plasmids used in this section were cloned using standard cloning protocols as described in MATERIAL AND METHODS of PART I. Plasmids are listed in the APPENDIX Plasmid list.

General protein expression protocol

Material:

- Bacterial expression plasmids (see Plasmid table)
- Rosetta BL21(DE3)pLysS (Novagen) expression bacteria

Solutions:

- Luria-Bertani (LB) medium (10g Bacto-tryptone; 5g yeast extract; 10g NaCl in 1l H_2O, adjust to pH 7.5, autoclave)
- Terrific Broth (TB) medium (for 1l: 12g Bacto-tryptone; 24g yeast extract, 5 ml Glycerol add H_2O to 900 ml, autoclave. Then add 100 ml of 0.17 M KH_2PO_4 (2.31 g), 0.17 M K_2HPO_4 (12.54 g) solution)
- 1M Isopropyl-β-D-thiogalactopyranosid (IPTG) stock solution
- Filter-sterilized 20% Glucose

The open reading frame (ORF) on the expression plasmids is controlled by a T7 promoter, expression bacteria (Rosetta BL21(DE3)pLysS) contain a T7 polymerase under the control of a *lac Operator* promoter. Lactose usually induces the *lac Operon* to produce enzymes that digest Lactose, high Glucose levels on the other hand shut down expression from *lac Operon.* IPTG is a molecular mimic of Allolactose, a metabolite of Lactose, which cannot be further degraded and thus acts as a 'gratuitous' inducer. By applying IPTG to expression bacteria, they also start to produce T7 polymerase, which on the other hand triggers expression of recombinant protein from the expression plasmid (Studier and Moffatt 1986).

Procedure:

First, expression bacteria were transformed with expression plasmids, plated on LB Agar plates containing and incubated at 37°C overnight. The next day, a 10 ml liquid culture (LB with 100 μg/ml Ampicillin, 50 μg/ml Kanamycin and 1% Glucose was inoculated with a single colony and grown at 37°C with a shaking incubator. After the culture had reached an OD_{600} of 2-3, the bacteria were pelleted by centrifugation at 1000 rpm for 15 min at RT. The pellet was then re-

suspended in 500 ml TB containing appropriate antibiotics, but no Glucose, and was adjusted to an OD_{600} of 0.1. The culture was grown at 37°C to reach an OD_{600} of 0.4-0.6. At this point, IPTG (final concentration: 1mM) was added to induce T7 polymerase and recombinant protein expression. The protein expression was performed at RT overnight at low shaking. Afterwards, the bacteria were harvested with an ultracentrifuge with 8000 rpm for 20 min at 4°C. The pellets, which contain bacteria with the produced recombinant proteins, can be stored at -80°C.

This expression protocol described above can be regarded as starting point for protein expression. The yield and quality of recombinant proteins produced in *E. coli* is highly variable and mainly depends on the features of the respective proteins. Thus, for each protein the optimal conditions have to be evaluated. Possible variables are the temperature at induction (from 20°C to 37°C), the length of production time (from 24h to 3 h), the concentration of IPTG (0.1mM to 2mM). But also the medium (LB, 2xYT, MagicMedia™ (Life Technologies), or TB) can be changed to reach optimal conditions. Additionally, different *E. coli* expression strains can be used, optimally suited to fit the characteristics of a given protein (see http://openwetware.org/wiki/E._coli_genotypes for a description of expression strains; and MAKRIDES 1996,BANEYX 1999 and ROSANO and CECCARELLI 2014 for reviews on recombinant protein expression in *E.coli*).

Lysis of bacterial cells

Material:

- Ultrasonicator (QSONICA)

Solutions:

- Lysis Buffer
- Lysozyme stock solution (500mg/ml) (Life Technologies)
- DNaseI stock solution (50mg/ml) (Sigma-Aldrich)
- Lysis Buffer (for 50 ml: 0.3 g Tris (final concentration 50 mM); 0.44 g NaCl (final conc.: 150 mM); 0.05 g MgCl2 (final conc.: 10 mM), plus one tablet of Roche cOmplete protease inhibitor mix®; add 100 µl Lysozyme stock solution and 100 µl DNaseI stock solution)

Procedure:

Bacterial pellets containing recombinant protein were re-suspended in cold 10 ml Lysis Buffer. To completely homogenize, the suspension was vortex vigorously for 1 min. Subsequently, the

mixture was incubated for 1 h on ice, this weakened the cell walls. To completely destroy cell walls and lyse the bacteria, the lysis solution was sonicated with 0.5 sec pulses at medium strength for 2-3 min in an Ultrasonicator. This step was performed on ice to counteract the warming of the cell lysate resulting from sonication. In addition to cell lysis, sonication will shear the bacterial DNA, which otherwise would form a viscous solution that may interfere with the subsequent steps of protein purification. After the sonication, the lysate was centrifuged at 19000 rpm at 4°C for 30 min in an Ultracentrifuge. This step separated the soluble (supernatant) from the insoluble fraction (pellet). For a comprehensive review on bacterial lysis by sonication see FELIU *et al.* 1998 and HO *et al.* 2006.

Usually, it is preferred when the recombinant protein is soluble (see protein expression protocol). If protein is misfolded and thus insoluble, the bacteria produce so-called inclusion bodies, which will be in the pellet fraction after centrifugation. In this case, the pellet can be solubilized in 8 M urea or 6 M guanidine HCl (SAMBROOK and RUSSELL 2006a).

For phage display, we aimed for correctly folded and soluble fluorescent proteins. This could easily be verified by the color of the cleared bacterial lysate.

Protein loading to magnetic beads for phage display

Material:

- Chitin magnetic beads (NEB)
- Streptavidin magnetic beads (IBA)

Solutions:

- Washing Buffer (for 50 ml: 0.3 g Tris (final concentration 50 mM); 0.44 g NaCl (final conc.: 150 mM); 0.05 g MgCl2 (final conc.: 10 mM), plus one tablet of Roche cOmplete protease inhibitor mix®)
- 1xPBS plus 0.05% Tween-20

Procedure:

Prior to protein loading, the magnetic beads (100 µl) were transferred to 1.5ml Eppendorf tubes and had to be washed extensively (5-6 times with 1.5 ml with Washing Buffer) to remove storage solution. After the centrifugation of the bacterial lysate, the supernatant was added to the washed magnetic beads and incubated on a spinning wheel for 45 min at 4°C. To remove unbound and other bacterial proteins, beads were washed with 5 times with 1.5 mL Washing

Buffer. Afterwards, magnetic beads with bound proteins were washed once with Washing Buffer mixed with 1xPBS plus 0.05% Tween-20 (ratio 1:1) and twice with 1.5 ml 1xPBS plus 0.05% Tween-20. This was done to adjust the buffer conditions used during phage display.

For the first round of phage display Chitin magnetic beads (NEB) were used. For the second round, antigens were bound to Streptavidin magnetic beads (IBA). In the third round of panning, Chitin magnetic beads were used again. Thus, for each round magnetic beads were alternated to eliminate phages specific for any beads.

Phage display protocol

During the process of phage display, many steps have to be performed in parallel. For example the production of the antigen and the preparation of the phage library. Thus, exact planning and preparation of all the steps is required.

Material:

- TG1 *E. coli* (Lucigen)
- Phagemid library in TG1 *E. coli*
- KM13 helper phages

Solutions:

- M9 media (For 1l: 6g Na_2HPO_4; 3g KH_2PO_4; NaCl: 0.5g; NH_4Cl: 1g; Agar: 15g, autoclave. Then add 10 mL of filter sterilized 100mM $MgSO_4$ (246.5mg); 20% Glucose (2g) ; 10mM $CaCl_2$ (14.7mg) and 100mM Thiamine-HCl (337.3 mg))
- 2xYT media (16g Bacto Tryptone; 10g Bacto-Yeast Extract; 5g NaCl in 1L H_2O)
- 500 cm^2 plates with 2xYT agar plus 100µl/mL Ampicillin and 0.5% Glucose
- LB agar plates without antibiotics
- Ampicillin stock solution (100 mg/ml)
- Kanamycin stock solution (50 mg/ml)
- Cold PEG6000 20%/NaCl 2.5 M
- Triethylamine (TEA) Buffer (28 µL 99% TEA (Sigma) in 2 ml H_2O, pH 10)
- 1.5 M Tris, pH 7.4
- 1 mg/ml Trypsin in PBS
- 20% Glucose in H_2O, filter-sterilized
- Washing buffer (1xPBS plus 150mM NaCl and 0.05% Tween-20)
- Washing Buffer without Tween-20
- Blocking Buffer 1 (Washing buffer plus 0.5% Casein)
- Blocking Buffer 2 (Washing buffer plus 6% BSA)
- Blocking Buffer 3 (Washing buffer plus 4% milk powder)

Production of helper phages

Plate KM13 helper phages

First, TG1 bacteria were streaked out on a M9 Agar plate. On this minimal medium plate F(+)-pili containing bacteria are more likely to be formed. Pili are necessary for the infection with bacteriophages (see Introduction). After overnight incubation at 37°C, one colony was picked and grown in 5 ml 2xYT media to $OD_{600}=1$. Then, sterile, melted M9 Agar (also called Top Agar, in this case) was prepared by warming to 47°C in a water bath. In the meantime, a serial dilution (e.g. 10^{-1} to 10^{-9}) of KM13 helper phage stock in 2xYT medium is prepared. Afterwards, 100 µl of grown TG1 bacteria were mixed with 100 µl of the each serial dilution by gently vortexing. Then, the infected bacteria were added to 3 ml of warm M9 Agar. This mixture was then directly poured onto a LB Agar plate and evenly distributed. This step was repeated for each serial dilution of the helper phage stock. Afterwards, the LB plates with the Top Agar (containing infected bacteria) were incubated overnight at 37°C. Infected colonies are recognized as so-called 'plaques'. These are areas of infected, slower growing clones that exhibit a turbid appearance (Sambrook and Russell 2006b).

Produce helper phages

The next day, TG1 bacteria (from the M9 plate) were inoculated in 1-2l 2xYT and incubated at 37°C to an $OD_{600}=0.4$. During the incubation of the TG1 bacteria, the Agar plates were checked and an isolated plaque was carefully picked by stamping out from the Agar with a specialized capillary. The stamped out plaque was mixed with 500 µL of 2xYT medium incubated for 1-2 h at RT. This time is needed for the phages to diffuse from the Agar into the medium. Then, the 2xYT containing phages was added to the prepared 2xYT medium and incubated at 37°C for 45 min without shaking, to infect the bacteria. Afterwards, Kanamycin was added to a final concentration of 75 µg/ml and the culture was incubated overnight at 37°C with moderate shaking. As KM13 helper phages contain a Kanamycin-resistance cassette, phage containing and producing cells can thus be positively selected (Sambrook and Russell 2006c).

Purification of helper phages

The next day, the bacteria were spun down for 15 min at 8000 rpm and 4°C in an ultracentrifuge. Now, the supernatant contained the produced and secreted helper phages, which were precipitated by mixing with ice cold PEG6000 20%/NaCl 2.5 M in a 5:1 ratio (supernatant: PEG6000). The supernatant/precipitation mix was then incubated for 60 min on

ice. Afterwards, the phages were centrifuged for 30 min at 4000 rpm and 4°C in an ultracentrifuge. Supernatant was removed and the phage pellets were dried for 20-30 min by inverting the centrifugation tube. To harvest, the phages were re-suspended in 50 ml cold 1xPBS and again re-precipitated with cold PEG6000 20%/NaCl 2.5 M on ice. After 30 min, the phages were centrifuged at 4000 rpm for 20 min, subsequently the supernatant was removed and the pellet was re-suspended in a smaller volume (2-4 ml) cold 1xPBS. This helper phage suspension was centrifuged several times at high speed in a microcentrifuge to remove residual bacteria. Afterwards, the KM13 helper phage can be stored at 4°C for about 2-3 weeks. By adding 7% DMSO, the phages can be stored at -20°C indefinitely.

Production of phages displaying library

Inoculation of (synthetic) phagemid library

1ml of glycerol stock of the phagemid library in TG1 bacteria (about 10^9 individual clones) was thawed and re-suspended in 600 mL 2xYT media. Subsequently, incubated for 20 min at 37°C on a shaker. When OD_{600} was at 0.07, 630 µl of Ampicillin stock solution and 30 ml of 20% Glucose was added. Then, incubated another 2h at 37°C on a shaker until the OD_{600} reaches 0.2, which corresponds to a multiplication of 100 the starting library.

Production of phages displaying library

For the first round of phage display, bacteria containing the library were infected with hyperphages (purchased from PROGEN) with multiplicity of infection (MOI) of 20:1 (phage-to-cells ratio), that means add $2x10^{12}$ phages/ml. The bacterial culture was quickly shaken to mix the phages and bacteria, but then incubated at 37°C for 45 min without shaking. Afterwards, the infected culture was incubated for 15 min at 37°C on a shaker. During this time, infected bacteria start to express the Kanamycin-resistance gene. Subsequently, bacteria were spun down for 30 min at 2000 rpm and 4°C in an ultracentrifuge. The supernatant, containing helper phages that did not infect bacteria, was discarded. Then, the pellet (containing library bacteria infected with helper phages) was re-suspended in 1.2l 2xYT with 100 µg/ml Ampicillin and 75 µg/ml Kanamycin. This culture was incubated overnight at 30°C, shaking moderately. The next day, bacteria were centrifuged for 15 min at 8000 rpm and 4°C in an ultracentrifuge. The supernatant then contained the phages displaying the library, which were precipitated by mixing with ice cold 240 ml PEG6000 20%/NaCl 2.5 M. The supernatant/precipitation mix was

then incubated for 60 min on ice. Afterwards, the phage precipitation was collected by centrifugation for 30 min at 4000 rpm and 4°C in an ultracentrifuge. Supernatant was removed and the phage pellets were dried for 20-30 min by inverting the centrifugation tube. The phage particles were re-suspended in 50 ml 1xPBS in total, foaming was tried to be avoided. The resuspension was vortexed at maximum for about 30 sec and centrifuged again at for 20 min 4000 rpm and 4°C to eliminate residual bacteria. The phage supernatant was precipitated again with 12.5 ml PEG6000 20%/NaCl 2.5 M for 30-60 min. Then, centrifuged for 30 min at 4000 rpm and 4°C, the supernatant was discarded and phage pellet re-suspended in 2ml Washing Buffer without Tween-20. To effectively remove all the residual bacteria, the phage solution was centrifuged at 4000 rpm for 3 min in a benchtop centrifuge and the supernatant was transferred to a new tube. This procedure was repeated as long as there was no more bacterial pellet visible. It is possible to store the phage library at 4°C overnight.

Biopanning

1ml of phage library (see previous chapter) was mixed with 1 ml of Washing Buffer 1 with 0.05% Tween-20 and 0.5% Casein to block unspecific high-affinity binders. In the next round of phage display, the blocking of the phages was performed with Blocking Buffer 2 (Washing buffer plus 6% BSA). The blocking mixture was incubated for at least an hour at 4°C. Afterwards, the antigen bound to magnetic beads was added to the blocked phage library and incubated for about 2 h at 4°C on a rotating wheel. Then, tube with phages and magnetic beads was inserted in a tube holder possessing a magnet on its rear panel. After all the beads were precipitated at the back of the tube, the supernatant (unbound phages) was removed and discarded. Magnetic pellet was re-suspended in 1mL Washing buffer and mix for 5 min at RT on a spinning wheel. Later, tube with washing buffer and magnetic beads (phages bound to the antigen) was re-inserted into the magnetic tube holder. After magnetic precipitation, the supernatant was removed and replaced by new Washing Buffer. These washing steps were repeated at least 20 times, to effectually remove all unbound and low-binding binders displayed on the phage surface. Approximately every fifth washing steps the magnetic beads were transferred to a new tube, thereby the phages binding to the test tube were removed. The last two washing steps were performed in Washing Buffer without Tween-20. To eliminate phages with helper phage phenotype and elute bound phages, 2 µl of Trypsin solution was

added to the magnetic beads in the Washing Buffer without Tween-20. After 30 min of incubation at RT, the supernatant (eluted positive phage particles) was transferred to a new tube. The remaining phages still bound to the beads/antigen were eluted with a pH-shock by adding 1 mL of TEA Buffer. To neutralize, this elution fraction was mixed with 750 µl of 1.5 M Tris pH 7.4. Afterwards, the two elution fractions are pooled.

Establishing a sub-library

Once the phages have been eluted from the antigen, they were used to infect 'empty' TG1 bacteria. Therefore, 9 ml of TG1 bacteria with an OD_{600}=0.3 were mixed with the combined elution fractions and incubated at RT for 1 h. In the meantime, to collect phages which could not be eluted by any of the two elution procedures, the beads were incubated in 500 µL of TG1 bacteria with an OD_{600}=0.3. Afterwards, the bacteria from the beads were mixed with the bacteria infected with the elution fractions. The infected bacteria were plated on six pre-warmed 500 cm^2 plates with 2xYT agar plus 100µg/ml Ampicillin and 0.5% Glucose and incubated overnight at 30°C. This way, bacteria that received a phagemid (Amp-resistance) were positively selected. The plating on Agar plates additionally ensures that fast-growing colonies were not outcompeting the slow-growing bacterial colonies. This way one antagonizes an overrepresentation of fast-growing clones, which do not necessarily correspond to the binders with the highest affinities.

The next day, the TG1 colonies were scraped from the Agar plates and re-suspended in approx. 5 mL LB per Agar plates. The scraped bacteria were pooled and mixed with Glycerol (final: 20% Glycerol) by vortexing for 1 min. This sub-library can now be stored in a -80°C freezer. To start a new round of phage display, a 100 µL aliquot of the sub-library glycerol stock was used, amplified and re-infected with KM13 helper phages, instead of hyperphages, as described in previous sections. Usually, after 3-4 rounds of phage display a sufficient enrichment for good binders is achieved (Aurélien Olichon, personal communication).

Phagemid purification from sub-library

50 µl of the sub-library glycerol stock was re-suspended in 100 ml LB with 100 µg/ml Ampicillin and incubated at 37°C on a shaker overnight. The next day, bacteria were harvested by centrifugation for 15 min with 4000 rpm at 4°C. Afterwards, the pHEN2 phagemids were

purified using a conventional plasmid midiprep kit (e.g. NucleoBond® Xtra Midi from Macherey-Nagel), according to the manufacturer's instructions manual.

Cloning VHHs into mammalian expression vector

From the pHEN2 phagemid, selected VHH-coding genes can be cloned out via NcoI and NotI restriction enzyme sites. Thus, the sub-library plasmid/phagemid pool was digested with NcoI/NotI, loaded and run on an Agarose gel and the about 400 bp-sized band was cut from the gel. The pool of VHH-coding genes was inserted into pAOINT2 plasmid (mammalian expression vector with ubiquitous CMV promoter N-terminal NcoI/NotI sites and C-terminal GFP fusion, obtained from Aurélien Olichon), transformed into electro-competent TOP10 *E.coli* and plated on LB Agar with 100 µg/ml Ampicillin. Individual single colonies were picked and grown in 2 ml LB plus 100 µg/ml Ampicillin at 37°C overnight. The following day, conventional plasmid mini prep (e.g. NucleoSpin® Plasmid from Macherey-Nagel) was performed according to the manufacturer's instructions manual.

Mammalian cell culture

Material:

- HeLa S3 cell line (obtained from Group of Prof. Nigg, Biozentrum Basel)
- Incubator 37°C, 5% CO_2
- Water bath

Solutions:

- Cell culture medium (500 mL Dulbecco's Modified Eagle's Medium (DMEM, Gibco®); 55 ml fetal bovine serum (FBS, Gibco®); 5.5 ml Penicillin-Streptomycin Solution (Gibco®))
- PBS, pH 7.4 (Gibco®)
- Trypsin-ETDA (0.25%) solution (Gibco®)

Procedure:

Frozen cells were thawed, re-suspended in fresh cell culture medium and plated in common tissue culture flasks. The cells attached to the surface of the flasks and were cultivated in a 37°C incubator with 5% CO_2. Every 2-3 days, the old cell culture medium was replaced by fresh, pre-warmed medium. When the cell culture reached 70-90% confluency, the cells were 'passaged' into new flasks. Therefore, the old tissue culture medium was aspirated and the

cells were washed with pre-warmed PBS to remove residual cell culture medium. Afterwards, the PBS was aspirated and Trypsin-ETDA (0.25%) solution was applied to the culture and incubated for 2-3 min at 37°C. During this step, called trypsinization, the cells detached from the surface d the tissue culture dish. Subsequently, trypsinization was blocked by addition of fresh cell culture medium. The cells, which were now in suspension, were diluted, distributed by new tissue culture flasks and incubated at 37°C 5% CO_2.

Transfection of mammalian cell culture

Material:

- 24-well glass bottom dish (MatTek Corporation)

Solutions:

- *Trans*IT ®-LT1 Transfection Reagent (Mirus)
- Opit-MEM®I Reduced Serum Medium (Gibco®)

Procedure:

HeLa cells were first equally distributed to a 24-well glass bottom dish with 0.5 mL cell culture medium and grown to a confluency of 50-70%.

For each well, 500 ng of plasmid DNA was dissolved in 50 µl Opit-MEM®I Reduced Serum Medium. For co-transfection, plasmids were mixed in a 1:1 ratio. 1.5 µl of *Trans*IT ®-LT1 Transfection Reagent was added to the DNA-Reduced Serum Medium mixture, quickly vortexed and incubated at RT for 30 min. Subsequently, the transfection suspension was applied dropwise to the cell culture. After 24-48 h incubation at 37°C, the cells were transfected and express the heterologous proteins from the plasmid(s).

For fixation, the cells were washed 2-3 times with PBS, and 4% paraformaldehyde (PFA) in PBS was added for 5 min at RT. Afterwards, the PFA was removed by 2-3 washing steps with PBS. The fixed cells can be stored at 4°C in PBS for several days.

The fixed, transfected cells were directly analyzed under a confocal fluorescence microscope, such as the Leica SP-5 MP.

REFERENCES

ABBAS A. K., LICHTMAN A. H., 2005 Cellular and Molecular Immunology. Saunders.

ABBEELE A. VAN DEN, CLERCQ S. DE, GANCK A. DE, CORTE V. DE, LOO B. VAN, SOROR S. H., SRINIVASAN V., STEYAERT J., VANDEKERCKHOVE J., GETTEMANS J., 2010 A llama-derived gelsolin single-domain antibody blocks gelsolin-G-actin interaction. Cell. Mol. Life Sci. **67**: 1519–35.

ABSKHARON R. N. N., GIACHIN G., WOHLKONIG A., SOROR S. H., PARDON E., LEGNAME G., STEYAERT J., 2014 Probing the N-terminal β-sheet conversion in the crystal structure of the human prion protein bound to a nanobody. J. Am. Chem. Soc. **136**: 937–44.

AGRAWAL V., SLIVAC I., PERRET S., BISSON L., ST-LAURENT G., MURAD Y., ZHANG J., DUROCHER Y., 2012 Stable expression of chimeric heavy chain antibodies in CHO cells. Methods Mol. Biol. **911**: 287–303.

AI H., HENDERSON J. N., REMINGTON S. J., CAMPBELL R. E., 2006 Directed evolution of a monomeric, bright and photostable version of Clavularia cyan fluorescent protein: structural characterization and applications in fluorescence imaging. Biochem. J. **400**: 531–40.

ALIEVA N. O., KONZEN K. a., FIELD S. F., MELESHKEVITCH E. a., HUNT M. E., BELTRAN-RAMIREZ V., MILLER D. J., WIEDENMANN J., SALIH A., MATZ M. V., 2008 Diversity and evolution of coral fluorescent proteins. PLoS One **3**.

ASHOUR J., SCHMIDT F. I., HANKE L., CRAGNOLINI J., CAVALLARI M., ALTENBURG A., BREWER R., INGRAM J., SHOEMAKER C., PLOEGH H. L., 2015 Intracellular expression of camelid single-domain antibodies specific for influenza virus nucleoprotein uncovers distinct features of its nuclear localization. J. Virol. **89**: 2792–800.

BANEYX F., 1999 Recombinant protein expression in Escherichia coli. Curr. Opin. Biotechnol. **10**: 411–21.

BARANOVA E., FRONZES R., GARCIA-PINO A., GERVEN N. VAN, PAPAPOSTOLOU D., PÉHAU-ARNAUDET G., PARDON E., STEYAERT J., HOWORKA S., REMAUT H., 2012 SbsB structure and lattice reconstruction unveil Ca2+ triggered S-layer assembly. Nature **487**: 119–22.

BARBAS C. F., KANG A. S., LERNER R. A., BENKOVIC S. J., 1991 Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. Proc. Natl. Acad. Sci. U. S. A. **88**: 7978–82.

BINZ H. K., STUMPP M. T., FORRER P., AMSTUTZ P., PLÜCKTHUN A., 2003 Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. J. Mol. Biol. **332**: 489–503.

BISCHOF J., MAEDA R. K., HEDIGER M., KARCH F., BASLER K., 2007 An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. Proc. Natl. Acad. Sci. U. S. A. **104**: 3312–7.

BODER E. T., WITTRUP K. D., 1997 Yeast surface display for screening combinatorial polypeptide

libraries. Nat. Biotechnol. **15**: 553–7.

BOERSMA Y. L., PLÜCKTHUN A., 2011 DARPins and other repeat protein scaffolds: advances in engineering and applications. Curr. Opin. Biotechnol. **22**: 849–57.

BRAND A. H., PERRIMON N., 1993 Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development **118**: 401–15.

BUSTIN M., SPERLING R., GOLDBLAT D., 1975 Proceedings: Use of antibodies against histones for the localization of histones in chromatin. Isr. J. Med. Sci. **11**: 1212.

CALLEJA M., MORENO E., PELAZ S., MORATA G., 1996 Visualization of gene expression in living adult Drosophila. Science **274**: 252–5.

CARROLL D., 2014 Genome engineering with targetable nucleases. Annu. Rev. Biochem. **83**: 409–39.

CAUSSINUS E., KANCA O., AFFOLTER M., 2011 Fluorescent fusion protein knockout mediated by anti-GFP nanobody. Nat. Struct. Mol. Biol. **19**: 117–121.

CHAIKUAD A., KEATES T., VINCKE C., KAUFHOLZ M., ZENN M., ZIMMERMANN B., GUTIÉRREZ C., ZHANG R.-G., HATZOS-SKINTGES C., JOACHIMIAK A., MUYLDERMANS S., HERBERG F. W., KNAPP S., MÜLLER S., 2014 Structure of cyclin G-associated kinase (GAK) trapped in different conformations using nanobodies. Biochem. J. **459**: 59–69.

CHALFIE M., TU Y., EUSKIRCHEN G., WARD W. W., PRASHER D. C., 1994 Green fluorescent protein as a marker for gene expression. Science **263**: 802–5.

CHASTEEN L., AYRISS J., PAVLIK P., BRADBURY A. R. M., 2006 Eliminating helper phage from phage display. Nucleic Acids Res. **34**: e145.

CLERCQ S. DE, ZWAENEPOEL O., MARTENS E., VANDEKERCKHOVE J., GUILLABERT A., GETTEMANS J., 2013 Nanobody-induced perturbation of LFA-1/L-plastin phosphorylation impairs MTOC docking, immune synapse formation and T cell activation. Cell. Mol. Life Sci. **70**: 909–22.

DAINES B., WANG H., WANG L., LI Y., HAN Y., EMMERT D., GELBART W., WANG X., LI W., GIBBS R., CHEN R., 2011 The Drosophila melanogaster transcriptome by paired-end RNA sequencing. Genome Res. **21**: 315–24.

DAY R. N., DAVIDSON M. W., 2009 The fluorescent protein palette: tools for cellular imaging. Chem. Soc. Rev. **38**: 2887–921.

DELANOTE V., VANLOO B., CATILLON M., FRIEDERICH E., VANDEKERCKHOVE J., GETTEMANS J., 2010 An alpaca single-domain antibody blocks filopodia formation by obstructing L-plastin-mediated F-actin bundling. FASEB J. **24**: 105–18.

DESMYTER A., DECANNIERE K., MUYLDERMANS S., WYNS L., 2001 Antigen specificity and high affinity binding provided by one single loop of a camel single-domain antibody. J. Biol. Chem. **276**:

26285–90.

DOLK E., VAART M. VAN DER, LUTJE HULSIK D., VRIEND G., HAARD H. DE, SPINELLI S., CAMBILLAU C., FRENKEN L., VERRIPS T., 2005 Isolation of llama antibody fragments for prevention of dandruff by phage display in shampoo. Appl. Environ. Microbiol. **71**: 442–50.

DOMANSKA K., VANDERHAEGEN S., SRINIVASAN V., PARDON E., DUPEUX F., MARQUEZ J. A., GIORGETTI S., STOPPINI M., WYNS L., BELLOTTI V., STEYAERT J., 2011 Atomic structure of a nanobody-trapped domain-swapped dimer of an amyloidogenic beta2-microglobulin variant. Proc. Natl. Acad. Sci. U. S. A. **108**: 1314–9.

DUMOULIN M., CONRATH K., MEIRHAEGHE A. VAN, MEERSMAN F., HEREMANS K., FRENKEN L. G. J., MUYLDERMANS S., WYNS L., MATAGNE A., 2002 Single-domain antibody fragments with high conformational stability. Protein Sci. **11**: 500–15.

FELIU J. X., CUBARSI R., VILLAVERDE A., 1998 Optimized release of recombinant proteins by ultrasonication of E. coli cells. Biotechnol. Bioeng. **58**: 536–40.

FIELD S. F., BULINA M. Y., KELMANSON I. V, BIELAWSKI J. P., MATZ M. V, 2006 Adaptive evolution of multicolored fluorescent proteins in reef-building corals. J. Mol. Evol. **62**: 332–9.

FIELDS S., SONG O., 1989 A novel genetic system to detect protein-protein interactions. Nature **340**: 245–246.

GENST E. DE, SILENCE K., DECANNIERE K., CONRATH K., LORIS R., KINNE J., MUYLDERMANS S., WYNS L., 2006 Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. Proc. Natl. Acad. Sci. U. S. A. **103**: 4586–91.

GOVAERT J., PELLIS M., DESCHACHT N., VINCKE C., CONRATH K., MUYLDERMANS S., SAERENS D., 2012 Dual beneficial effect of interloop disulfide bond for single domain antibody fragments. J. Biol. Chem. **287**: 1970–9.

GREENBERG A. S., AVILA D., HUGHES M., HUGHES A., MCKINNEY E. C., FLAJNIK M. F., 1995 A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. Nature **374**: 168–73.

GRIFFITHS A. D., WILLIAMS S. C., HARTLEY O., TOMLINSON I. M., WATERHOUSE P., CROSBY W. L., KONTERMANN R. E., JONES P. T., LOW N. M., ALLISON T. J., 1994 Isolation of high affinity human antibodies directly from large synthetic repertoires. EMBO J. **13**: 3245–60.

GROSS G. G., JUNGE J. A., MORA R. J., KWON H.-B., OLSON C. A., TAKAHASHI T. T., LIMAN E. R., ELLIS-DAVIES G. C. R., MCGEE A. W., SABATINI B. L., ROBERTS R. W., ARNOLD D. B., 2013 Recombinant probes for visualizing endogenous synaptic proteins in living neurons. Neuron **78**: 971–85.

HAMERS-CASTERMAN C., ATARHOUCH T., MUYLDERMANS S., ROBINSON G., HAMERS C., SONGA E. B., BENDAHMAN N., HAMERS R., 1993 Naturally occurring antibodies devoid of light chains.

Nature **363**: 446–8.

Haseloff J., 1999 GFP variants for multispectral imaging of living cells. Methods Cell Biol. **58**: 139–51.

Heim R., Cubitt A. B., Tsien R. Y., 1995 Improved green fluorescence. Nature **373**: 663–4.

Helma J., Cardoso M. C., Muyldermans S., Leonhardt H., 2015 Nanobodies and recombinant binders in cell biology. J. Cell Biol. **209**: 633–644.

Herce H. D., Deng W., Helma J., Leonhardt H., Cardoso M. C., 2013 Visualization and targeted disruption of protein interactions in living cells. Nat. Commun. **4**: 2660.

Ho C. W., Chew T. K., Ling T. C., Kamaruddin S., Tan W. S., Tey B. T., 2006 Efficient mechanical cell disruption of Escherichia coli by an ultrasonicator and recovery of intracellular hepatitis B core antigen. Process Biochem. **41**: 1829–1834.

Hoogenboom H. R., Griffiths A. D., Johnson K. S., Chiswell D. J., Hudson P., Winter G., 1991 Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. Nucleic Acids Res. **19**: 4133–7.

Hoogenboom H. R., 2005 Selecting and screening recombinant antibody libraries. Nat. Biotechnol. **23**: 1105–1116.

Hu J. C., Kornacker M. G., Hochschild A., 2000 Escherichia coli one- and two-hybrid systems for the analysis and identification of protein-protein interactions. Methods **20**: 80–94.

Irannejad R., Tomshine J. C., Tomshine J. R., Chevalier M., Mahoney J. P., Steyaert J., Rasmussen S. G. F., Sunahara R. K., El-Samad H., Huang B., Zastrow M. von, 2013 Conformational biosensors reveal GPCR signalling from endosomes. Nature **495**: 534–8.

Jiang J., Struhl G., 1998 Regulation of the Hedgehog and Wingless signalling pathways by the F-box/WD40-repeat protein Slimb. Nature **391**: 493–6.

Kirchhofer A., Helma J., Schmidthals K., Frauer C., Cui S., Karcher A., Pellis M., Muyldermans S., Casas-Delucchi C. S., Cardoso M. C., Leonhardt H., Hopfner K.-P., Rothbauer U., 2010 Modulation of protein properties in living cells using nanobodies. Nat. Struct. Mol. Biol. **17**: 133–8.

Koide A., Bailey C. W., Huang X., Koide S., 1998 The fibronectin type III domain as a scaffold for novel binding proteins. J. Mol. Biol. **284**: 1141–51.

Korndörfer I. P., Beste G., Skerra A., 2003 Crystallographic analysis of an "anticalin" with tailored specificity for fluorescein reveals high structural plasticity of the lipocalin loop region. Proteins **53**: 121–9.

Korotkov K. V, Pardon E., Steyaert J., Hol W. G. J., 2009 Crystal structure of the N-terminal domain of the secretin GspD from ETEC determined with the assistance of a nanobody.

Structure **17**: 255–65.

KRISTENSEN P., WINTER G., 1998 Proteolytic selection for protein folding using filamentous bacteriophages. Fold. Des. **3**: 321–8.

KRUIF J. DE, BOEL E., LOGTENBERG T., 1995 Selection and application of human single chain Fv antibody fragments from a semi-synthetic phage antibody display library with designed CDR3 regions. J. Mol. Biol. **248**: 97–105.

KRUSE A. C., RING A. M., MANGLIK A., HU J., HU K., EITEL K., HÜBNER H., PARDON E., VALANT C., SEXTON P. M., CHRISTOPOULOS A., FELDER C. C., GMEINER P., STEYAERT J., WEIS W. I., GARCIA K. C., WESS J., KOBILKA B. K., 2013 Activation and allosteric modulation of a muscarinic acetylcholine receptor. Nature **504**: 101–6.

KUBALA M. H., KOVTUN O., ALEXANDROV K., COLLINS B. M., 2010 Structural and thermodynamic analysis of the GFP:GFP-nanobody complex. Protein Sci. **19**: 2389–401.

KUMMER L., PARIZEK P., RUBE P., MILLGRAMM B., PRINZ A., MITTL P. R. E., KAUFHOLZ M., ZIMMERMANN B., HERBERG F. W., PLÜCKTHUN A., 2012 Structural and functional analysis of phosphorylation-specific binders of the kinase ERK from designed ankyrin repeat protein libraries. Proc. Natl. Acad. Sci. U. S. A. **109**: E2248–57.

LADENSON R. C., CRIMMINS D. L., LANDT Y., LADENSON J. H., 2006 Isolation and characterization of a thermally stable recombinant anti-caffeine heavy-chain antibody fragment. Anal. Chem. **78**: 4501–8.

LENARD A., DAETWYLER S., BETZ C., ELLERTSDOTTIR E., BELTING H.-G., HUISKEN J., AFFOLTER M., 2015 Endothelial Cell Self-fusion during Vascular Pruning. PLoS Biol. **13**: e1002126.

LORIS R., MARIANOVSKY I., LAH J., LAEREMANS T., ENGELBERG-KULKA H., GLASER G., MUYLDERMANS S., WYNS L., 2003 Crystal structure of the intrinsically flexible addiction antidote MazE. J. Biol. Chem. **278**: 28252–7.

MAKRIDES S. C., 1996 Strategies for achieving high-level expression of genes in Escherichia coli. Microbiol. Rev. **60**: 512–38.

MATTHEAKIS L. C., BHATT R. R., DOWER W. J., 1994 An in vitro polysome display system for identifying ligands from very large peptide libraries. Proc. Natl. Acad. Sci. **91**: 9022–9026.

MILÁN M., COHEN S. M., 1999 Notch Signaling Is Not Sufficient to Define the Affinity Boundary between Dorsal and Ventral Compartments. Mol. Cell **4**: 1073–1078.

MORISE H., SHIMOMURA O., JOHNSON F. H., WINANT J., 1974 Intermolecular energy transfer in the bioluminescent system of Aequorea. Biochemistry **13**: 2656–62.

MUYLDERMANS S., ATARHOUCH T., SALDANHA J., BARBOSA J. A., HAMERS R., 1994 Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. Protein Eng. **7**: 1129–35.

Muyldermans S., 2013 Nanobodies: Natural Single-Domain Antibodies. Annu. Rev. Biochem. **82**: 775–797.

Nord K., Gunneriusson E., Ringdahl J., Ståhl S., Uhlén M., Nygren P. a, 1997 Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain. Nat. Biotechnol. **15**: 772–777.

Olichon A., Marco A. de, 2012 Preparation of a naïve library of camelid single domain antibodies. Methods Mol. Biol. **911**: 65–78.

Ormö M., Cubitt A. B., Kallio K., Gross L. A., Tsien R. Y., Remington S. J., 1996 Crystal structure of the Aequorea victoria green fluorescent protein. Science **273**: 1392–5.

Pakhomov A. A., Martynov V. I., 2008 GFP family: structural insights into spectral tuning. Chem. Biol. **15**: 755–64.

Pancer Z., Cooper M. D., 2006 The evolution of adaptive immunity. Annu. Rev. Immunol. **24**: 497–518.

Panza P., Maier J., Schmees C., Rothbauer U., Sollner C., 2015 Live imaging of endogenous protein dynamics in zebrafish using chromobodies. Development **142**: 1879–1884.

Pellis M., Muyldermans S., Vincke C., 2012 Bacterial two hybrid: a versatile one-step intracellular selection method. Methods Mol. Biol. **911**: 135–50.

Plückthun A., 2015 Designed Ankyrin Repeat Proteins (DARPins): Binding Proteins for Research, Diagnostics, and Therapy. Annu. Rev. Pharmacol. Toxicol. **55**: 489–511.

Qi H., Lu H., Qiu H.-J., Petrenko V., Liu A., 2012 Phagemid vectors for phage display: properties, characteristics and construction. J. Mol. Biol. **417**: 129–43.

Rasmussen S. G. F., Choi H.-J., Fung J. J., Pardon E., Casarosa P., Chae P. S., Devree B. T., Rosenbaum D. M., Thian F. S., Kobilka T. S., Schnapp A., Konetzki I., Sunahara R. K., Gellman S. H., Pautsch A., Steyaert J., Weis W. I., Kobilka B. K., 2011a Structure of a nanobody-stabilized active state of the β(2) adrenoceptor. Nature **469**: 175–80.

Rasmussen S. G. F., DeVree B. T., Zou Y., Kruse A. C., Chung K. Y., Kobilka T. S., Thian F. S., Chae P. S., Pardon E., Calinski D., Mathiesen J. M., Shah S. T. A., Lyons J. A., Caffrey M., Gellman S. H., Steyaert J., Skiniotis G., Weis W. I., Sunahara R. K., Kobilka B. K., 2011b Crystal structure of the β2 adrenergic receptor-Gs protein complex. Nature **477**: 549–55.

Ries J., Kaplan C., Platonova E., Eghlidi H., Ewers H., 2012 A simple, versatile method for GFP-based super-resolution microscopy via nanobodies. Nat. Methods **9**: 582–4.

Roberts R. W., Szostak J. W., 1997 RNA-peptide fusions for the in vitro selection of peptides and proteins. Proc. Natl. Acad. Sci. **94**: 12297–12302.

Rondot S., Koch J., Breitling F., Dübel S., 2001 A helper phage to improve single-chain antibody

presentation in phage display. Nat. Biotechnol. **19**: 75–78.

Rosano G. L., Ceccarelli E. A., 2014 Recombinant protein expression in Escherichia coli: advances and challenges. Front. Microbiol. **5**: 172.

Rothbauer U., Zolghadr K., Tillib S., Nowak D., Schermelleh L., Gahl A., Backmann N., Conrath K., Muyldermans S., Cardoso M. C., Leonhardt H., 2006 Targeting and tracing antigens in live cells with fluorescent nanobodies. Nat. Methods **3**: 887–9.

Rothbauer U., Zolghadr K., Muyldermans S., Schepers A., Cardoso M. C., Leonhardt H., 2008 A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. Mol. Cell. Proteomics **7**: 282–9.

Russell M., Lowman H. B., Clackson T., 2004 Introduction to phage biology and phage display. In: Clackson T, Lowman HB (Eds.), *Phage Display - A Practical Approach*, Oxford University Press, pp. 1–26.

Salema V., Marín E., Martínez-Arteaga R., Ruano-Gallego D., Fraile S., Margolles Y., Teira X., Gutierrez C., Bodelón G., Fernández L. Á., 2013 Selection of single domain antibodies from immune libraries displayed on the surface of E. coli cells with two β-domains of opposite topologies. PLoS One **8**: e75126.

Salih A., Larkum A., Cox G., Kühl M., Hoegh-Guldberg O., 2000 Fluorescent pigments in corals are photoprotective. Nature **408**: 850–3.

Sambrook J., Russell D. W., 2006a Purification of expressed proteins from inclusion bodies. CSH Protoc. **2006**.

Sambrook J., Russell D. W., 2006b Plating bacteriophage m13. CSH Protoc. **2006**.

Sambrook J., Russell D. W., 2006c Growing bacteriophage m13 in liquid culture. CSH Protoc. **2006**.

Scherer W. F., Syverton J. T., Gey G. O., 1953 Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix. J. Exp. Med. **97**: 695–710.

Schornack S., Fuchs R., Huitema E., Rothbauer U., Lipka V., Kamoun S., 2009 Protein mislocalization in plant cells using a GFP-binding chromobody. Plant J. **60**: 744–754.

Scott J. K., Smith G. P., 1990 Searching for peptide ligands with an epitope library. Science **249**: 386–90.

Shaner N. C., Campbell R. E., Steinbach P. A., Giepmans B. N. G., Palmer A. E., Tsien R. Y., 2004 Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat. Biotechnol. **22**: 1567–72.

Shcherbo D., Merzlyak E. M., Chepurnykh T. V, Fradkov A. F., Ermakova G. V, Solovieva E. a,

Lukyanov K. a, Bogdanova E. a, Zaraisky A. G., Lukyanov S., Chudakov D. M., 2007 Bright far-red fluorescent protein for whole-body imaging. Nat. Methods **4**: 741–746.

Shcherbo D., Murphy C. S., Ermakova G. V, Solovieva E. a, Chepurnykh T. V, Shcheglov A. S., Verkhusha V. V, Pletnev V. Z., Hazelwood K. L., Roche P. M., Lukyanov S., Zaraisky A. G., Davidson M. W., Chudakov D. M., 2009 Far-red fluorescent tags for protein imaging in living tissues. Biochem. J. **418**: 567–574.

Shimomura O., Johnson F. H., Saiga Y., 1962 Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. J. Cell. Comp. Physiol. **59**: 223–39.

Sidhu S. S., Weiss G. A., Wells J. A., 2000 High copy display of large proteins on phage for functional selections. J. Mol. Biol. **296**: 487–95.

Sidhu S. S., 2001 Engineering M13 for phage display. Biomol. Eng. **18**: 57–63.

Sievers F., Wilm A., Dineen D., Gibson T. J., Karplus K., Li W., Lopez R., McWilliam H., Remmert M., Söding J., Thompson J. D., Higgins D. G., 2011 Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. **7**: 539.

Skerra A., 2008 Alternative binding proteins: anticalins - harnessing the structural plasticity of the lipocalin ligand pocket to engineer novel binding activities. FEBS J. **275**: 2677–83.

Slonczewski J. L., Foster J. W., 2010 *Microbiology- An Evolving Science*. W. W. Norton & Company.

Smith G. P., 1985 Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science **228**: 1315–7.

Spinelli S., Tegoni M., Frenken L., Vliet C. van, Cambillau C., 2001 Lateral recognition of a dye hapten by a llama VHH domain. J. Mol. Biol. **311**: 123–9.

Studier F. W., Moffatt B. A., 1986 Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189**: 113–30.

Sundberg E. J., Mariuzza R. A., 2002 *Protein Modules and Protein-Protein Interaction*. Elsevier.

Tang J. C. Y., Szikra T., Kozorovitskiy Y., Teixiera M., Sabatini B. L., Roska B., Cepko C. L., 2013 A nanobody-based system using fluorescent proteins as scaffolds for cell-specific gene manipulation. Cell **154**: 928–39.

Tokuhara D., Álvarez B., Mejima M., Hiroiwa T., Takahashi Y., Kurokawa S., Kuroda M., Oyama M., Kozuka-Hata H., Nochi T., Sagara H., Aladin F., Marcotte H., Frenken L. G. J., Iturriza-Gómara M., Kiyono H., Hammarström L., Yuki Y., 2013 Rice-based oral antibody fragment prophylaxis and therapy against rotavirus infection. J. Clin. Invest. **123**: 3829–38.

Traenkle B., Emele F., Anton R., Poetz O., Haeussler R. S., Maier J., Kaiser P. D., Scholz A. M.,

Nueske S., Buchfellner A., Romer T., Rothbauer U., 2015 Monitoring interactions and dynamics of endogenous beta-catenin with intracellular nanobodies in living cells. Mol. Cell. Proteomics **14**: 707–23.

Truttmann M. C., Wu Q., Stiegeler S., Duarte J. N., Ingram J., Ploegh H. L., 2015 HypE-specific nanobodies as tools to modulate HypE-mediated target AMPylation. J. Biol. Chem. **290**: 9087–100.

Tsien R. Y., 1998 The green fluorescent protein. Annu. Rev. Biochem. **67**: 509–44.

Vaughan T. J., Williams A. J., Pritchard K., Osbourn J. K., Pope A. R., Earnshaw J. C., McCafferty J., Hodits R. A., Wilton J., Johnson K. S., 1996 Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. Nat. Biotechnol. **14**: 309–14.

Virdi V., Coddens A., Buck S. De, Millet S., Goddeeris B. M., Cox E., Greve H. De, Depicker A., 2013 Orally fed seeds producing designer IgAs protect weaned piglets against enterotoxigenic Escherichia coli infection. Proc. Natl. Acad. Sci. U. S. A. **110**: 11809–14.

Voller A., Bartlett A., Bidwell D. E., 1978 Enzyme immunoassays with special reference to ELISA techniques. J. Clin. Pathol. **31**: 507–20.

Vu K. B., Ghahroudi M. A., Wyns L., Muyldermans S., 1997 Comparison of llama VH sequences from conventional and heavy chain antibodies. Mol. Immunol. **34**: 1121–31.

Wezenbeek P. M. van, Hulsebos T. J., Schoenmakers J. G., 1980 Nucleotide sequence of the filamentous bacteriophage M13 DNA genome: comparison with phage fd. Gene **11**: 129–48.

Wojcik J., Hantschel O., Grebien F., Kaupe I., Bennett K. L., Barkinge J., Jones R. B., Koide A., Superti-Furga G., Koide S., 2010 A potent and highly specific FN3 monobody inhibitor of the Abl SH2 domain. Nat. Struct. Mol. Biol. **17**: 519–27.

Yang F., Moss L. G., Phillips G. N., 1996 The molecular structure of green fluorescent protein. Nat. Biotechnol. **14**: 1246–51.

Zhang J., Liu X., Bell A., To R., Baral T. N., Azizi A., Li J., Cass B., Durocher Y., 2009 Transient expression and purification of chimeric heavy chain antibodies. Protein Expr. Purif. **65**: 77–82.

Zolghadr K., Rothbauer U., Leonhardt H., 2012 The fluorescent two-hybrid (F2H) assay for direct analysis of protein-protein interactions in living cells. Methods Mol. Biol. **812**: 275–82.

APPENDIX

Name	Size (bp)	Starting plasmid	Restriction enzymes used		Resistance cassette	Reference/details
pAttBLacZ	12665	made with pUASattB and pH-Pelican			Ampicillin	from Alex Weiss
pAttBLacZ apP	18712	pAttBLacZ	XbaI	AscI	Ampicillin	exact details see Master Thesis Dimi Bieli
pAttBLacZ apC	20322	pAttBLacZ	AscI	BglII	Ampicillin	exact details see Master Thesis Dimi Bieli
pAttBLacZ apO	15762	pAttBLacZ	AscI	BglII	Ampicillin	exact details see Master Thesis Dimi Bieli
pAttBLacZ apQ	17201	pAttBLacZ	AscI	BglII	Ampicillin	exact details see Master Thesis Dimi Bieli
pAttBLacZ apR	16289	pAttBLacZ	AscI	BglII	Ampicillin	exact details see Master Thesis Dimi Bieli
pAttBLacZ approm	12865	pAttBLacZ			Ampicillin	exact details see Master Thesis Dimi Bieli (apterous promoter combined with 5'UTR of hsp70)
pAttBLacZ approm apP	18925	pAttBLacZ approm	AscI	BglII	Ampicillin	exact details see Master Thesis Dimi Bieli
pAttBLacZ approm apC	20522	pAttBLacZ approm	AscI	BglII	Ampicillin	exact details see Master Thesis Dimi Bieli
pAttBLacZ approm apO	15962	pAttBLacZ approm	AscI	BglII	Ampicillin	exact details see Master Thesis Dimi Bieli
pAttBLacZ approm apQ	17414	pAttBLacZ approm	AscI	BglII	Ampicillin	exact details see Master Thesis Dimi Bieli
pAttBLacZ approm apR	16489	pAttBLacZ approm	AscI	BglII	Ampicillin	exact details see Master Thesis Dimi Bieli
pAttBLacZ approm dad4	?	pAttBLacZ approm	AscI	SphI	Ampicillin	exact details see Master Thesis Dimi Bieli, dadint2 in thesis
pAttBLacZ apCON1	14231	pAttBLacZ	XbaI	AscI	Ampicillin	PCRed from yw gDNA
pAttBLacZ apCON2	15135	pAttBLacZ	XbaI	AscI	Ampicillin	PCRed from yw gDNA
pAttBLacZ apCON3	14274	pAttBLacZ	EcoRI	NotI	Ampicillin	PCRed from yw gDNA
pAttBLacZ apCON4	17956	pAttBLacZ	NotI	AscI	Ampicillin	PCRed from yw gDNA
pAttBLacZ apOR	14899	pAttBLacZ	AscI	BglII	Ampicillin	PCRed from yw gDNA
pAttBLacZ apOR2	13354	pAttBLacZ	AscI	XbaI	Ampicillin	PCRed from yw gDNA
pAttBLacZ apOR3	14109	pAttBLacZ	AscI	XbaI	Ampicillin	PCRed from yw gDNA
pAttBLacZ apRXa	13521	pAttBLacZ	AscI	XbaI	Ampicillin	PCRed from pReentry whole enhancer full length
pAttBLacZ apRXadadint52	14548	pAttBLacZ	NotI	KpnI	Ampicillin	apRXadadint52 cut out of pBSIIKSmut plus apRXadadint52 with KpnI/NotI
pAttBLacZ apRXadadint52dad4	16654	pAttBLacZ	SphI	AscI	Ampicillin	PCRed from pBSIIKSmut plus apRXadadint52 plus dad4
pAttBLacZ apR plus CON4	21689	pAttBLacZ	NotI	KpnI	Ampicillin	apRCON4 cut out of pBSIIKSmut plus apR plus CON4 with KpnI/NotI
pAttBLacZ OR463	13120	pAttBLacZ	NotI	AscI	Ampicillin	PCRed from pReentry whole enhancer full length
pAttBLacZ OR463 m1	13092	pAttBLacZ	NotI	AscI	Ampicillin	cut out from pBS OR463 m1 with AscI/NotI
pAttBLacZ OR463 m2	13076	pAttBLacZ	NotI	AscI	Ampicillin	cut out from pBS OR463 m2 with AscI/NotI
pAttBLacZ OR463 m3	13090	pAttBLacZ	NotI	AscI	Ampicillin	cut out from pBS OR463 m3 with AscI/NotI
pAttBLacZ OR463 m4	13082	pAttBLacZ	NotI	AscI	Ampicillin	cut out from pBS OR463 m4 with AscI/NotI
pAttBLacZ OR463 m14	13054	pAttBLacZ	NotI	AscI	Ampicillin	cut out from pBS OR463 m14 with AscI/NotI
pAttBLacZ OR463 m23	13046	pAttBLacZ	NotI	AscI	Ampicillin	cut out from pBS OR463 m23 with AscI/NotI
pAttBLacZ int2.3 -	13649	pAttBLacZ	EcoRI		Ampicillin	PCRed from pReentry with apcDNA plus intron 2 and 3
pAttBLacZ int2.3 +	13649	pAttBLacZ	EcoRI		Ampicillin	PCRed from pReentry with apcDNA plus intron 2 and 3
pGH-approm5'UTR LacZ fusion	5005		NotI	AatI	Ampicillin	synthesized by ATG:biosynthetics GmbH
pAttBLacZ approm 5'UTR	13797	pAttBLacZ	NotI	AatI	Ampicillin	approm5'UTR LacZ fusion cut out of pGH-approm5'UTR LacZ fusion
pAttBLacZapprom 5'UTR plus apR	17421	pAttBLacZ approm 5'UTR	AscI	BglII	Ampicillin	PCRed from pReentry whole enhancer full length
pAttBLacZapprom 5'UTR apCON4	19088	pAttBLacZ approm 5'UTR	AscI	NotI	Ampicillin	PCRed from pReentry whole enhancer full length
pPapprom5'UTRLacZ	8992	pBSmin mame2	NotI	SmaI	Ampicillin	cut out from pAttBLacZ approm
pPapprom5'UTRLacZ-3xP3GFP	10310	pPapprom5'UTRLacZ	EcoRI		Ampicillin	EGFP cut from 3xp3-EGFP-SV40 in pGEM-T Easy (promega)
pBSIIKS	2961				Ampicillin	general plasmid
pBS apC		pBSIIKS			Ampicillin	exact details see Master Thesis Dimi Bieli
pBSKSmut XbaI to AvrII	2961	pBSIIKS			Ampicillin	made by mutagenic PCR
pBSIIKS ap 3'UTR short	3395	pBSIIKS	KpnI	SacI	Ampicillin	PCRed from pReentry with apcDNA
pBSIIKS ap 3'UTR and CDS long	4245	pBSIIKS	KpnI	SacI	Ampicillin	PCRed from pReentry with apcDNA
pBSIIKSmut plus apCON1	4538	pBSKSmut XbaI to AvrII	AvrII	XmaI	Ampicillin	PCRed from pReentry whole enhancer full length
pBSIIKSmut plus apR	6604	pBSKSmut XbaI to AvrII	AvrII	XmaI	Ampicillin	PCRed from pReentry whole enhancer full length
pBSIIKSmut plus apCON2	5442	pBSKSmut XbaI to AvrII	AvrII	XmaI	Ampicillin	PCRed from pReentry whole enhancer full length
pBSIIKSmut plus apCON3	4582	pBSKSmut XbaI to AvrII	AvrII	XmaI	Ampicillin	PCRed from pReentry whole enhancer full length
pBSIIKSmut plus apCON4	8254	pBSKSmut XbaI to AvrII	AvrII	XmaI	Ampicillin	PCRed from pReentry whole enhancer full length
pBSIIKSmut plus apRXadadint52	4768	pBSKSmut XbaI to AvrII	AvrII	XmaI	Ampicillin	PCRed from apXasta gDNA
pBSIIKSmut plus dad4	5162	pBSKSmut XbaI to AvrII	AvrII	XmaI	Ampicillin	PCRed from apXasta gDNA
pBSIIKSmut plus dadint52 leftover	3896	pBSKSmut XbaI to AvrII	AvrII	XmaI	Ampicillin	PCRed from apXasta gDNA
pBSIIKSmut plus apR plus CON2	9097	pBSIIKSmut plus apR	SpeI	XmaI	Ampicillin	CON2 cut out of pBSIIKSmut plus apCON2 with AvrII/XmaI
pBSIIKSmut plus apR plus CON4	11909	pBSIIKSmut plus apR	SpeI	XmaI	Ampicillin	CON4 cut out of pBSIIKSmut plus apCON4 with AvrII/XmaI
pBSIIKSmut plus apCON3 plus CON4	9887	pBSIIKSmut plus apCON3	SpeI	XmaI	Ampicillin	CON4 cut out of pBSIIKSmut plus apCON4 with AvrII/XmaI
pBSIIKSmut plus apR plus CON2 plus CON4	14402	pBSIIKSmut plus apR plus CON2	SpeI	XmaI	Ampicillin	CON4 cut out of pBSIIKSmut plus apCON4 with AvrII/XmaI
pBSIIKSmut plus apR plus CON2 plus CON3CON	16035	pBSIIKSmut plus apR plus CON2	SpeI	XmaI	Ampicillin	CON3CON4 cut out of pBSIIKSmut plus apCON3 plus apCON4 with AvrII/XmaI
pBSIIKSmut plus apCON1apRCON2CON3CON4	17624	pBSIIKSmut plus apCON1	SpeI	XmaI	Ampicillin	apRCON2CON3CON4 cut ot of pBSIIKSmut plus apR plus CON2 plus CON3CON4 with AvrII/XmaI
pBSIIKSmut plus apRXadadint52 plus dad4	6981	pBSIIKSmut plus apRXadadint52	SpeI	XmaI	Ampicillin	dad4 cut of of pBSIIKSmut plus dad4 with AvrII/XmaI
pBSIIKSmut plus dadint52 leftover plus dad4	6109	pBSIIKSmut plus dadint52 leftover	SpeI	XmaI	Ampicillin	dad4 cut of of pBSIIKSmut plus dad4 with AvrII/XmaI
pBSIIKSmut plus apRXa	3829	pBSKSmut XbaI to AvrII	AvrII	XmaI	Ampicillin	PCRed from pReentry whole enhancer full length
pBSIIKSmut plus apROR463	3418	pBSKSmut XbaI to AvrII	AvrII	XmaI	Ampicillin	PCRed from pReentry whole enhancer full length
pBSIIKSmut plus apCON4A	6737	pBSKSmut XbaI to AvrII	AvrII	XmaI	Ampicillin	PCRed from pReentry whole enhancer full length
pBSIIKSmut plus apCON4B	3555	pBSKSmut XbaI to AvrII	AvrII	XmaI	Ampicillin	PCRed from pReentry whole enhancer full length
pBSIIKSmut plus apRXaCON4A	7617	pBSIIKSmut plus apRXa	SpeI	XmaI	Ampicillin	CON4A cut out of pBSIIKSmut plus apCON4A with AvrII/XmaI
pBSIIKSmut plus apRXaCON4B	4435	pBSIIKSmut plus apRXa	SpeI	XmaI	Ampicillin	CON4B cut out of pBSIIKSmut plus apCON4B with AvrII/XmaI
pBSIIKSmut plus apROR463 CON4A	7206	pBSIIKSmut plus apROR463	SpeI	XmaI	Ampicillin	CON4A cut out of pBSIIKSmut plus apCON4A with AvrII/XmaI
pBSIIKSmut plus apROR463 CON4B	4024	pBSIIKSmut plus apROR463	SpeI	XmaI	Ampicillin	CON4B cut out of pBSIIKSmut plus apCON4B with AvrII/XmaI
pBSIIKSmut plus apCON1 apR	8193	pBSIIKSmut plus apCON1	SpeI	XmaI	Ampicillin	apR cut out of pBSIIKSmut plus apR with AvrII/XmaI
pBSIIKSmut plus apCON1 CON2	7031	pBSIIKSmut plus apCON1	SpeI	XmaI	Ampicillin	CON2 cut out of pBSIIKSmut plus apCON2 with AvrII/XmaI
pBSIIKSmut plus apCON1 CON4	9843	pBSIIKSmut plus apCON1	SpeI	XmaI	Ampicillin	CON4 cut out of pBSIIKSmut plus apCON4 with AvrII/XmaI
pBSIIKSmut plus apCON2 CON3	7075	pBSIIKSmut plus apCON2	SpeI	XmaI	Ampicillin	CON3 cut out of pBSIIKSmut plus apCON3 with AvrII/XmaI
pBSIIKSmut plus apCON1 CON4	13969	pBSIIKSmut plus apCON1 CON4	SpeI	XmaI	Ampicillin	CON23 cut out of pBSIIKSmut plus apCON2CON3 with AvrII/XmaI
pBSIIKSmut plus apCON1423apR	17624	pBSIIKSmut plus apCON1423	SpeI	XmaI	Ampicillin	PCRed from pReentry whole enhancer full length
pBSIIKSmut plus apCON4apR	11909	pBSIIKSmut plus apCON4	SpeI	XmaI	Ampicillin	PCRed from pReentry whole enhancer full length
pBSIIKSmut plus apCON123	8664	pBSIIKSmut plus apCON1	SpeI	XmaI	Ampicillin	CON23 PCRed from allCON
pBSIIKSmut plus apCON1234	13969	pBSIIKSmut plus apCON123	SpeI	XmaI	Ampicillin	CON4 cut out of pBSIIKSmut plus apCON4 with AvrII/XmaI
pBSIIKSmut plus apCON1apR23	12319	pBSIIKSmut plus apCON1apR2	SpeI	XmaI	Ampicillin	CON3 PCRed
pBS OR463	3420	pBSIIKS	NotI	BamHI	Ampicillin	PCRed from pReentry whole enhancer full length
pBS OR463 m1	3392	pBS OR463	NotI	BamHI	Ampicillin	made by mutagenic PCR
pBS OR463 m2	3376	pBS OR463	NotI	BamHI	Ampicillin	made by mutagenic PCR
pBS OR463 m3	3390	pBS OR463	NotI	BamHI	Ampicillin	made by mutagenic PCR
pBS OR463 m4	3382	pBS OR463	NotI	BamHI	Ampicillin	made by mutagenic PCR
pBS OR463 m14	3354	pBS OR463 m4	NotI	BamHI	Ampicillin	made by mutagenic PCR
pBS OR463 m23	3346	pBS OR463 m2	NotI	BamHI	Ampicillin	made by mutagenic PCR

Name	Size	Parent	RE1	RE2	Resistance	Description
pReentry for enhancer	7080	pBSIIKS			Ampicillin	aka pBSIIKS attB FRT LoxP yellow clone 3 made by Oguz Kanca
pReentry whole enhancer full length	33860	pReentry for enhancer			Ampicillin	made by Oguz Kanca
pReentry allCON	21755	pReentry for enhancer	AvrII	AgeI	Ampicillin	allCON cut out of pBSIIKSmut plus apCON1apRCON2CON3CON4 with AvrII/XmaI
pReentry apR	10735	pReentry for enhancer	AvrII	AgeI	Ampicillin	apR cut out of pBSIIKSmut plus apR with AvrII/XmaI
pReentry apRXa	7952	pReentry for enhancer	AvrII	AgeI	Ampicillin	PCRed from apXasta gDNA
pReentry apRCON2	13228	pReentry for enhancer	AvrII	AgeI	Ampicillin	apRCON2 cut out of pBSIIKSmut plus apR plus CON2 with AvrII/XmaI
pReentry apRCON2CON3CON4	20166	pReentry for enhancer	AvrII	AgeI	Ampicillin	apRCON2CON3CON4 cut out of pBSIIKSmut plus apR plus CON2 plus CON3CON4 with AvrII/XmaI
pReentry apRCON2CON4	18533	pReentry for enhancer	AvrII	AgeI	Ampicillin	apRCON2CON4 cut out of pBSIIKSmut plus apR plus CON2 plus CON4 with AvrII/XmaI
pReentry apRCON4	16040	pReentry for enhancer	AvrII	AgeI	Ampicillin	apRCON4 cut out of pBSIIKSmut plus apR plusCON4 with AvrII/XmaI
pReentry CON4	12385	pReentry for enhancer	AvrII	AgeI	Ampicillin	CON4 cut out of pBSIIKSmut plusCON4 with AvrII/XmaI
pReentry apRXadadint52	8899	pReentry for enhancer	AvrII	AgeI	Ampicillin	apRXadadint52 cut out of pBSIIKSmut plus apRXadadint52 with AvrII/XmaI
pReentry apRXadadint52dad4	11112	pReentry for enhancer	AvrII	AgeI	Ampicillin	apRXadadint52dad4 cut out of pBSIIKSmut plus apRXadadint52 plus dad4 with AvrII/XmaI
pReentry dadint52dad4	10240	pReentry for enhancer	AvrII	AgeI	Ampicillin	dadint52dad4 cut out of pBSIIKSmut plus dadint52 plus dad4 with AvrII/XmaI
pReentry apRXaCON4A	11748	pReentry for enhancer	AvrII	AgeI	Ampicillin	apRXaCON4A cut out of pBSIIKSmut plus apRXaCON4A with AvrII/XmaI
pReentry apRXaCON4B	8566	pReentry for enhancer	AvrII	AgeI	Ampicillin	apRXaCON4B cut out of pBSIIKSmut plus apRXaCON4B with AvrII/XmaI
pReentry apOR463 CON4A	11337	pReentry for enhancer	AvrII	AgeI	Ampicillin	apOR463CON4A cut out of pBSIIKSmut plus apOR463CON4A with AvrII/XmaI
pReentry apOR463 CON4B	8155	pReentry for enhancer	AvrII	AgeI	Ampicillin	apOR463CON4B cut out of pBSIIKSmut plus apOR463CON4B with AvrII/XmaI
pReentry CON1CON4CON2CON3apR	21755	pReentry for enhancer	AvrII	AgeI	Ampicillin	CON1CON4CON2CON3apR cut from pBSIIKSmut plus apCON1423apR with AvrII/XmaI
pReentryCON4apR	16040	pReentry for enhancer	AvrII	AgeI	Ampicillin	CON4apR cut from pBSIIKSmut plus apCON4apR with AvrII/XmaI
pReentry CON1CON2CON3	12795	pReentry for enhancer	AvrII	AgeI	Ampicillin	CON1CON2CON3 cut from pBSIIKSmut plus apCON123 with AvrII/XmaI
pReentry CON1CON2CON3CON4	18100	pReentry for enhancer	AvrII	AgeI	Ampicillin	CON1CON2CON3CON4 cut from pBSIIKSmut plus apCON1234 with AvrII/XmaI
pReentry CON1apRCON2CON3	16450	pReentry for enhancer	AvrII	AgeI	Ampicillin	CON1apRCON2CON3 cut from pBSIIKSmut plus apCON1apR23 with AvrII/XmaI
pReentry for CDS recombineering	7040	pBSIIKS			Ampicillin	made by Oguz Kanca
HL02012	?				Ampicillin	ap cDNA clone from DGRC long 3'UTR
SD05618	?				Chloramphenicol	ap cDNA clone from DGRC long 5'UTR
apCDS TOPO XL	8053	TOPO XL			Kanamycin	made a fusion/overlap of ap cDNA (from HL02012) and ap promoter 5'UTR (from yw gDNA).
pReentry with apcDNA	11551	pReentry for CDS recombineering	NotI	AscI	Ampicillin	CDS cut out from apCDS TOPO XL NotI/AscI
pReentry with apcDNA plus intron 2 and 3	12100	pReentry for CDS recombineering	NotI	AscI	Ampicillin	synthesized by Genewiz Inc.
pH-mCherryNLS	9991				Ampicillin	from Amanda Ochoa (made from pH-Stinger, Barolo et al., 2000)
pH-mKate2	9974	pH-mCherryNLS	AgeI	SpeI	Ampicillin	mKateNLS PCRed from pBSattPmKate2NLS (from Oguz)
pH-eGFP Kozak	10288	pH-mCherryNLS	AgeI	SpeI	Ampicillin	EGFP was PCRed from pH-Stinger, basically this plasmis is the same as pH-Stinger but with Drosophila Kozak seq)
pUASattB	8489				Ampicillin	from Amanda Ochoa
pattBmKate2	9131	pUASattB	XbaI	SphI	Ampicillin	hsp70-mKateNLS was cut out from pH-mKate2 with SpeI/SphI
pattBeGFPKozak	9445	pUASattB	XbaI	SphI	Ampicillin	hsp70-eGFPKozak was cut out from pH-eGFPKozak with SpeI/SphI
pattBmKate2 vgBE	9868	pattBmKate2	XbaI	AscI	Ampicillin	vgBE PCRed from yw gDNA (Jim A. Williams et al., 1994)
pattBmKate2 wg-c extended	11172	pattBmKate2	XbaI	AscI	Ampicillin	wg-c PCRed from yw gDNA also took the conserved peak next to the fragment (Neumann and Cohen 1996)
pattBmKate2 patched	9805	pattBmKate2	KpnI	AgeI	Ampicillin	patched promoter PCRed from yw gDNA and replaced hsp70 promoter in pattBmKate2 (C Alexandre et al., 1996)
pattBeGFPKozak apR	13164	pattBeGFPKozak	KpnI	NotI	Ampicillin	apR from pBSmut apR
pattBeGFPKozak CON4	14814	pattBeGFPKozak	KpnI	NotI	Ampicillin	apCON4 from pBSmut apCON4
pattBmKate2 apR	12850	pattBmKate2	KpnI	NotI	Ampicillin	apR from pBSmut apR
pattBmKate2 CON4	14500	pattBmKate2	KpnI	NotI	Ampicillin	apCON4 from pBSmut apCON4
pET-22b (+)	5493				Ampicillin	Novagen, has pelB leader sequence (exports protein into periplasm)
pET22b(+) plus apcDNA	6849	pET-22b (+)	NcoI	NotI	Ampicillin	ap cDNA was amplified from pReentry with apcDNA
pET22b(+) delpelB plus apcDNA	6783	pET22b(+) plus apcDNA			Ampicillin	made by mutatenic PCR
pET22b(+) plus apHD	5919	pET-22b (+)	NcoI	NotI	Ampicillin	
pET22b(+) plus apHD delpelB	5853	pET22b(+) plus apHD			Ampicillin	made by mutatenic PCR, used to purify apterous homeodomain and produce antibody
pET22b(+)delpelB plus ap delHD	6372	pET22b(+) delpelB plus apcDNA			Ampicillin	made by mutatenic PCR
pET22b(+) dpp	5787	pET-22b (+)	NcoI	XhoI	Ampicillin	
pAOT7-mKate2-CBD2S	5350	pAOT7-VHRhoC-CBD2S	NcoI	NotI	Kanamycin	
pAOT7-apHD-CBD2S	5134	pAOT7-VHRhoC-CBD2S	NcoI	NotI	Kanamycin	
pAOT7-mKate2-His	5041	pAOT7-mKate2-CBD2S			Kanamycin	
pAOT7-ap-CBD2S	6054	pAOT7-VHRhoC-CBD2S	NcoI	NotI	Kanamycin	useless because of frame shift
pAOT7 mKate2VHH clone2	4706	pAOT7-mKate2-His	NcoI	BamHI	Kanamycin	PCRed from pAOint mKateVHH-GFP clone2 (mammalian expression vector)
pAOT7 mKate2VHH clone12	4706	pAOT7-mKate2-His	NcoI	BamHI	Kanamycin	PCRed from pAOint mKateVHH-GFP clone12 (mammalian expression vector)
pAOT7 mKate2VHH clone16	4715	pAOT7-mKate2-His	NcoI	BamHI	Kanamycin	PCRed from pAOint mKateVHH-GFP clone16 (mammalian expression vector)
pAOT7 mKate2VHH clone40	4724	pAOT7-mKate2-His	NcoI	BamHI	Kanamycin	PCRed from pAOint mKateVHH-GFP clone40 (mammalian expression vector)
p2strep-mKate2	4764	p2strepRhoBQ63L	KpnI	BamHI	Kanamycin/Neomycin	mammalian expression vector, left in toulouse
pCBD-mKate2	4845	pCBD-Tev-RhoBQ63L	KpnI	BamHI	Kanamycin/Neomycin	mammalian expression vector, left in toulouse
pmKate2-N	4711				Kanamycin/Neomycin	Evrogen
pH2BmKate2	5085	pmKate2-N	KpnI	BamHI	Kanamycin/Neomycin	mammalian expression vector nuclear mKate2
papHD-mKate2	5197	pmKate2-N	KpnI	PstI	Kanamycin/Neomycin	mKate signal is exclusively seen in nucleus
papHD-H2B-mKate2	5572	pH2BmKate2	KpnI	PstI	Kanamycin/Neomycin	mKate signal is exclusively seen in nucleus
pFLAG-VKL-HA-H2B-mKate2	5154	pH2BmKate2	SalI	KpnI	Kanamycin/Neomycin	synthesized 2 complementary oligos (containing FLAG HA) which i anneal and then cut
pRIL					Chloramphenicol	from Aurelien Olichon, Toulouse

List of all the plasmids generated for this PhD thesis