Evolution of divergent functions of the Hox protein Ultrabithorax in insects: comparison in transgenic *Drosophila melanogaster*

IISER PUNE

Thesis Submitted towards the Partial fulfillment of

BS-MS dual degree programme

By

Dhanashree Prakash Khanale

To

The Department of Biology

Indian Institute of Science Education and Research

Pune

Under the Supervision of

Prof. L. S. Shashidhra

Indian Institute of Science Education and Research

Pune

CERTIFICATE

This is to certify that this dissertation entitled **"Evolution of divergent functions of the Hox protein Ultrabithorax in insects: comparison in transgenic *Drosophila melanogaster*"** towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research Pune, represents original research carried out by **Dhanashree Prakash Khanale** at the Department of Biology, Indian Institute of Science Education and Research, Pune under the supervision of **Prof. L. S. Shashidhara** during the academic year 2013-2014.

Prof. L. S. Shasidhara

(Head of the Department of Biology)

Date: 2nd April, 2014

Place: Pune

DECLARATION

I hereby declare that the matter embodied in the report entitled **"Evolution of divergent functions of the Hox protein Ultrabithorax in insects: comparison in transgenic *Drosophila melanogaster*"** are the results of the investigations carried out by me at the Department of Biology, Indian Institute of Science Education and Research, Pune, under the supervision of **Prof. L S. Shashidhara** and the same has not been submitted elsewhere for any other degree.

Dhanashree Prakash Khanale

Date: 2nd April, 2014

Place: Pune

Abstract

The Hox gene, Ultrabithorax (Ubx) mediates haltere development which is the third thoracic segment structure in *Drosophila* by repressing some of the most important wing patterning genes like *wingless*, *cut* and *vestigial*. Diversity in the T3 segment dorsal morphology is evident in different insect groups. Although Ubx is expressed in the T3 segments in these insects, it has been shown that above mentioned targets of *Drosophila* Ubx are not differentially expressed between the forewing and hindwings of all insects. Here we chose three insects representing three distinct wing morphology groups viz. *Apis mellifera*, *Bombyx mori* and *Tribolium custaneum* and checked the ability of Ubx protein from these insects to repress the wing patterning genes in *Drosophila* and thus carry out wing to haltere transformation. We have generated transgenic *Drosophila* expressing Ubx from *Apis*, *Bombyx* and *Tribolium*. We have shown that Ubx from *Apis* and *Bombyx* is able to repress wing patterning genes when over-expressed in the wing imaginal discs of *Drosophila* and induce wing to haltere transformation. Ability of Ubx from other one insect group to specify haltere development is being investigated. Results suggest that diversity in the T3 segment morphology of these insects is not attributed to the differences in the sequences of Ubx genes and the differential regulation must be brought about downstream to Ubx, perhaps due to evolutionary changes in the enhancer sequences around the region bound by Ubx. To test this hypothesis, we have generated transgenic *Drosophila* expressing GFP under the control of enhancer of *wingless*, *cut*, and *vestigial* of *Apis.* These transgenic constructs are currently being tested for their expression patterns under different genetic backgrounds.

List of figures

Acknowledgements

This is indeed the right place to place a bouquet of my sincere and heartfelt gratitude to my project supervisor Prof. L. S. Shahidhara. I am indebted to him for his constant encouragement and invaluable guidance during the course of this project work. The atmosphere of academic freedom I enjoyed in his presence was really helpful to me. It is my pleasure to keep on records the deep respect and appreciation I have towards him.

I extend my sincere gratitude to Prof. K. N. Ganesh, for permitting me to take up my research project IISER, Pune. I am grateful for all kinds of facilities provided by IISER, Pune. I am also very grateful to Naveen Prasad, Savita Singh, Harsha, Jayprakash and Shital for all the help rendered during the course of my work.With much appreciation, I would like to mention the help and support that I have received from all my labmates. I sincerely acknowledge CDFD, Hyderabad for providing the *Bombyx* cDNA and NCBS fly facility, Bangalore for their prompt execution over transgenic requests. A big thank you goes to my friends Ankita Sharma, Upendra Singh, Tirthusree Das for helping me to get through tough times. My love to all my family members especially to my parents and my 'aaji' for constant guidance, moral support and unfailing love.

Dhanashree Khanale

Chapter 1

Introduction

Arthropods that originated from Urbilaterians (the common protostome-deuterostome ancestor), show remarkable diversity in terms of their body plan and the habitat they inhibit. There are over one million described and yet to be described species under this phylum (Basset et al., 2012). It is subdivided into four major subphyla Chelicerata, Myriapoda, Crustacea and Hexapoda (to which the group insect belongs) (Regier et al., 2010). Evolution of diversity in the morphology has played an important role in their successful adaptation to various habitats. Since a defined set of genes control particular development processes which give rise to morphological structures, diversity in the morphology has been attributed to changes in the gene regulatory networks. Understanding diversity in the morphology hence demands for understanding the genetic basis underlying these changes in the body plan.

1.1. *Drosophila melanogaster* as a model system

A good model system should be easier to manipulate experimentally with existing tools, should have a well known genetic structure and a small enough life cycle so that experimental manipulations can be studied for many generations. More importantly, the studies should have wider implications that can be extrapolated to understanding of life processes in a wide range of organisms including humans.

Thomas Hunt Morgan in the early 1900's first used *Drosophila* as a model organism to study transmission of genetic information. Since then *Drosophila* has been a favourite model system for thousand of researchers across the globe. *Drosophila* is an insect belonging to the order Diptera and has one pair of wings and one pair balancing organs named-halteres. *Drosophila* is easy to culture in the laboratory, has a short life span of about 10 days at room temperature and a well characterized genetic structure (180 Mb) and many of the genes involved in development in vertebrates were first identified in *Drosophila*. Many spontaneous and induced mutations are available and a large collection of mutations showing variety of useful phenotypes exists for the system. RNAi lines for about 85% of the genes are available and continuous improvements to the genetic tools such as UAS-

GAL4 system (Brand and Perrimon, 1993), FLP / FRT technique (Xu and Rubin, 1993) make *Drosophila* even more beneficial for research. Balancer chromosomes are another reason why *Drosophila* becomes a great system for genetic research. These chromosomes have multiple inversions and deletions and sequence so scrambled that they are no longer capable of undergoing homologous recombination during meiosis. These chromosomes also often carry recessive lethal markers themselves and hence provide a means for constructing true breeding stocks for lethal mutations in which only flies that are doubly heterozygous for the balancer and lethal mutation bearing homolog survive (Greenspan, 2004).

1.1.1 Genome of *Drosophila melanogaster*

Fly genome was sequenced in 2000 and has been curated at the Flybase database. Drosophila contains 4 pairs of chromosomes, X/Y, and three autosomes labelled as 2, 3 and 4. The fourth chromosome is only about one-fifth as large as the others and is often neglected (Greenspan, 2004). This 139.49 million base pair sequence has been annotated and contains about 15,771 genes according to genome project reports (NCBI genome database).

1.1.2 Life cycle of *Drosophila*

Drosophila, belonging to the subclass Endopterygota is a holometabolous insect i.e. they undergo a complete metamorphosis which means that the adult form looks entirely different from the larval form. There are four stages of development: egg, larva, pupa and adult. The first stage of fertilized egg lasts for about 24 hrs at 25° C. At the end of embryogenesis, embryo starts dividing syncytially and gets organized into segments and precursor cells that later give rise to larval muscles, nervous system and structures such as heart, salivary glands and imaginal discs in appropriate segments. A group of 20 precursor cells known as organ primordiathat are formed during embryogenesis undergo divisions and get reorganized into imaginal discs during the larval stage. Larvae then molt twice and through three successive developmental stages or instar stages, enter the pupa stage. Pupa stage lasts for about 5 days at 25° C. During the pupa stage, larva undergoes complete metamorphosis and adult fly ecloses at the end of it. The entire life cycle takes about 10-12 days at 25° C.

1.1.3 Early patterning of the *Drosophila embryo*

Products of about 50 maternal genes get activated after fertilization and these molecules deposited by the mother, define the initial anterior-posterior and dorsal-ventral polarity in the *Drosophila* embryo more or less simultaneously. Maternal gene products define spatial distribution of proteins and RNA and these then activate zygotic genes at particular positions. Zygotic genes set up a hierarchy of gene activity to specify the segmented body plan. Gap genes are the first to be activated amongst the segmentation genes and mutations in these genes cause loss of large chunks of segments. Graded distributions of gap genes then activate pair rule genes such as *even-skipped* and *fushi-tarazu* which act at every other segment and this delimits the formation of 14 parasegments in the *Drosophila* embryo (Figure 1). These parasegments are the fundamental and independent units of segmentation and they later gain their own identity. The final segments formed, are out of register with the parasegments by about half a segment. Each segment is thus made up of posterior part of one parasegment and anterior part of the next. Pair-rule genes then activate segment-polarity genes which control patterning within each segment e.g. *engrailed* defines posterior compartment of each segment. Mutations in polarity genes cause polarity reversal at the segment level (L. Wolpert, 2011; Nusslein-Volhard and Wieschaus, 1980). A combined activity of Gap genes, pair rule genes and segment polarity genes activate homeotic genes and these act at the final step of the patterning hierarchy and define segmental identities in *Drosophila*. In case of absence of hoemotic genes, a mix of thoracic and cephalic pattern is observed and no morphologically diverse segments develop (Struhl, 1982).
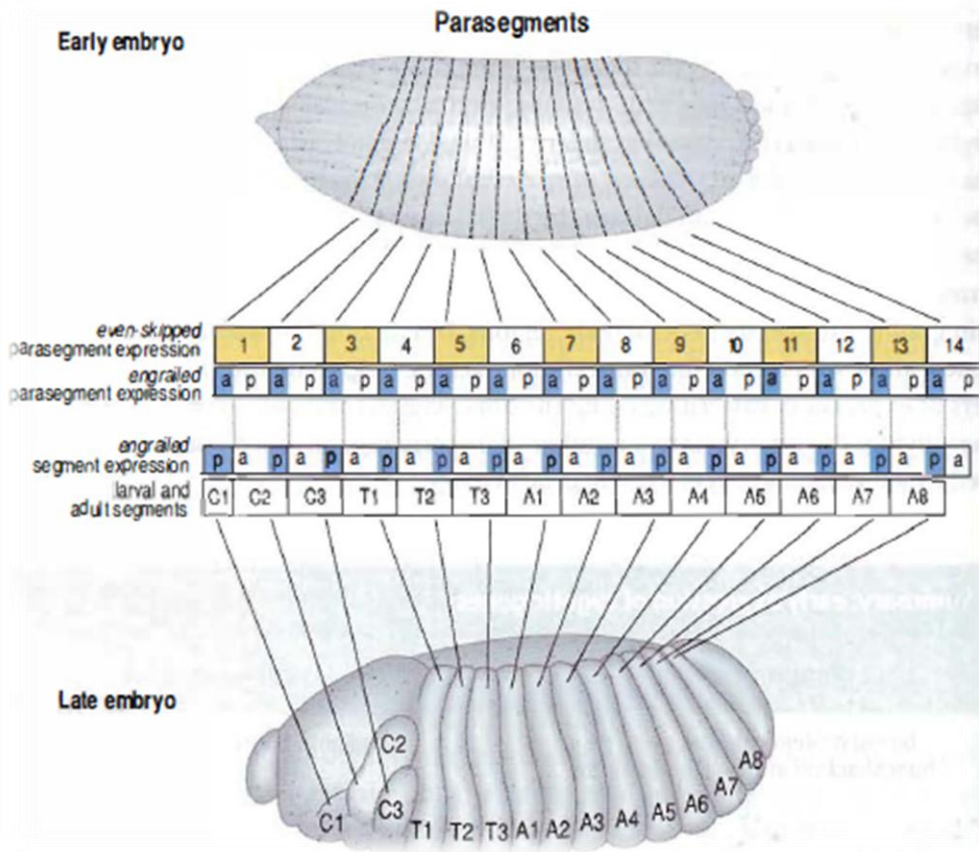
Figure 1. Early patterning of the *Drosophila* embyo; figure courtesy: 'Pinciples of development' by Wopert

1.2 Hox genes and development of the body plan

Urbilaterians are considered to be the first animals that show a segmented body plan and it is proposed that this ancestor had at least seven Hox genes. Animals belonging to all the major phyla such as annelids, arthropods and chordates that originated from Urbilaterians show a striking feature in terms of the body plan design which is construction from repeating structures (S. B. Carroll, 2004). Bateson coined a term 'Homeosis' meaning 'similar condition' and it is associated with formation of similar body parts at different locations (McGinnis W, 1994). Genes associated with such homeotic mutations are called homeotic genes. For example, in *Cbx* mutants, wing is transformed into a haltere in *Drosophila* (White and Akam, 1985). All the homeotic genes encode for transcription factors and contain a conserved DNA binding domain known as homeodomain or homeobox and hence the common accepted name for the genes is homeobox or Hox genes.

1.2.1 Organization of Hox genes

All the eight Hox genes in *Drosophila* are found on the third chromosome distributed into two clusters: the one closer to the centromere is named as Antennapedia-complex (ANT-C) and the distal one is called Bithorax compkex (BX-C).DNA fragments containing Homeodomain of the genes *Antp* and *Ubx* were found to cross react with these two clusters. Later, these were also found to cross react with DNA from earthworms, chickens and humans. This was the first concrete proof showing that across different phyla, similar patterning mechanisms might exist (McGinnis et al., 1984). The ANT-C codes for five Hox genes named *labial, proboscipedia, Deformed, Sex combs reduced* and *Antennapedia* that control the identity of head, mouthparts and the first two thoracic segments T1 and T2. Segments posterior to the T2 are defined by gene products of the BX-C. BX-C codes for three Hox genes namely: *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*) and *Abdominal-B* (*Abd-B*) (Sanchez-Herrero et al., 1985).

Hox genes exhibit colinearity (Figure 2) in their expression meaning they are expressed in the same order as they are located on the chromosome (Lewis, 1978). The first gene of the BX-C complex, *Ubx* shows expression starting from the posterior part of T2 segment and is extended into the abdominal segments. Although the expression persists beyond the anterior compartment of the first abdominal segment, it overlaps with the expression of *abd-A* and *Abd-B*and due to posterior prevalence of Hox genes; Ubx has no functional role in the development of abdominal segments. The abd-A imparts identity to the first four abdominal segments while the remaining four are specified by Abd-B. Vertebrate Hox genes show not only spatial colinearity but also temporal colinearity; most anterior genes are expressed first followed by the posterior ones in a sequential manner (reviewed in (Duboule, 1998)). In case of mouse and humans, Hox complexes are present on four clusters, on four different chromosomes. Sequence analysis of mammalian Hox genes shows that they arose from a single cluster of three Hox genes and got diverged and duplicated during the course of evolution (Carroll et al., 1995).
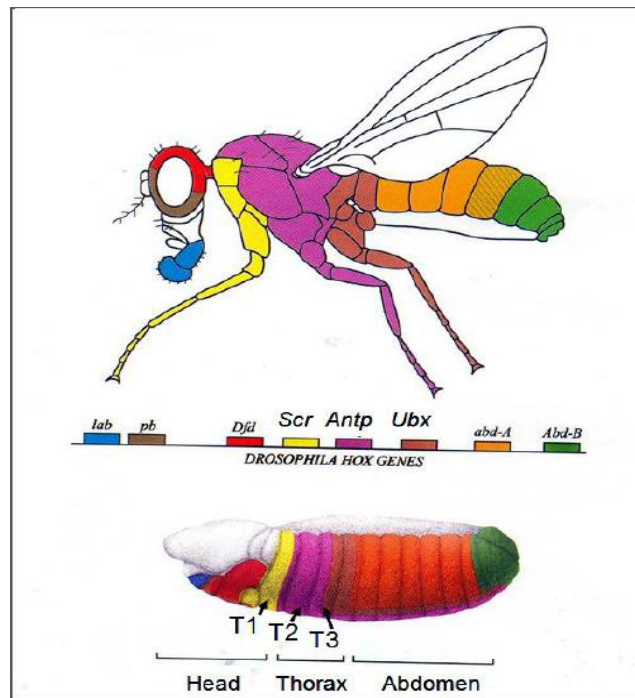
Figure 2. Hox genes and the colinearity principle

1.2.2 Function of Hox genes

Hox genes are transcription factors that act as master regulators of patterning along the A/P axis (Garcia-Bellido, 1977). Deletion of *Scr, Antp, Ubx, abd-A and Abd-B* removes the entire Hox input from abdominal and thoracic segments. The resulting pattern of cuticle consists of only thoracic elements which resembles second thoracic segment (T2), with some cephalic structures in posterior compartments (Struhl and White, 1985), which could be due to de-repression of some head selector genes such as *Dfd*. This result suggests that ground state pattern is thoracic (with legs on each segment). This also suggests that during development, all metamers start with same thoracic ground plan, but later diverge as Hox genes get activated.

Interestingly, ectopic expression of either Ubx, abd-A or to some extent Abd-B cause similar homeotic transformations i.e., antenna to leg and wing to haltere in *Drosophila*. Abd-A can fully and Abd-B can partially substitute Ubx for haltere development (Casares et al., 1996) and either of them can repress *Dll*. Studies indicate that while many Hox genes share common downstream targets among them, they also regulate subsets of unique downstream targets which help specifying segmental identity (Mann et al., 2009).

Suppression of wing fate and specification of haltere fate in *Drosophila* by the Hox gene *Ultrabithorax* (*Ubx*) (Figure 3) is a classical example of Hoxregulation of serial homology, which has served as a paradigm for understanding Hox gene function (Lewis, 1978). In *Drosophila,* wings and halteres are the dorsal appendages of the second and third thoracic segments (T2 and T3), respectively. As discussed earlier, T2 is the ground state and in the T3 segment this ground state is modulated to produce clubbed shaped balancing organs known as halteres. Removal of Ubx from T3 segment leads to homeotic transformation due to which entire T2 segmental morphology is repeated in T3 segment. This loss of Ubx function in T3 thus results in a four winged fly (Lewis, 1978), whereas, gain of Ubx in T2 segment leads to homeotic transformation of wing-to-halteres (Cabrera et al., 1985; Lewis, 1978; White and Akam, 1985). Another example of homeotic transformation is how Antennapedia (Antp) represses Homothorax (Hth) to specify leg discs (Casares and Mann, 1998). Antp is not normally expressed in wild type antennal discs.However ectopic expression of Antp in antennal discs leads to repression of Hth, subsequently leading to homeotic transformation of Antennal identity in to leg identity (Casares and Mann, 1998; Dong et al., 2000; Struhl, 1981) (figure 4).
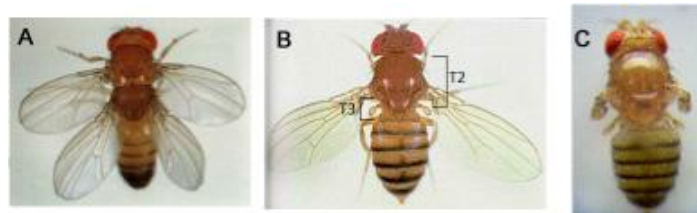
Figure 3. A four winged fly (A), instead of a normal two winged wild type fly (B) arises when the function of one hox gene Ultrabithorax is supressed. On the other hand a mutant fly with four halteres is formed when Ultrabithroax is ectopically expressed in the second thoracic segment (C).
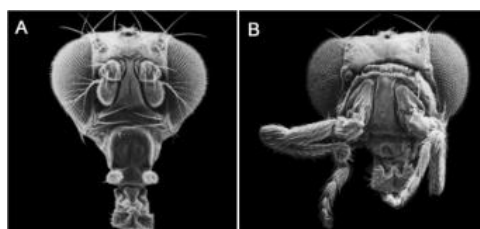
Figure 4. Ectopic expression of Antp in antennal discs leads to transformation of antennae into legs in the adult fly, a process known as homeotic transformation.

1.3 Development of flight appendages in *Drosophila*

Imaginal discs are the sets of undifferentiated, mitotic cells that are carried by larva of fruit fly during development. The wing and haltere (the dorsal second and third thoracic appendage of *Drosophila*) are formed from two such imaginal discs- the wing and haltere imaginal discs. Two organizing centres located at the boundary of antero posterior and dorso ventral compartments are responsible for shaping the wing blade. The wing imaginal disc consists of about 20 cells when it is formed during embryonic development. These cells divide and ultimately the wing imaginal disc in the late third instar larva consists of 75000 cells. The wing and leg discs of *Drosophila* have a common precursor which later separate as result of dorsal segregation of the wing disc. The common precursor is established at the A-P boundary within the mesothroacic segment and consists of Engrailed (En) expressing and non expressing cells. It therefore seems that the A- P boundary is inherited from the embryos and is maintained in the larvae.

The mature late third instar larval disc is a flattened sac with two distinct surfaces- a thin peripodial membrane and a thicker folded disc epithelium. The disc epithelium makes most of the wing blade and hinge and also the body of the adult fly. During pupation, the wing disc everts, folding upon itself to form the apposed dorsal and ventral epithelia of the wing blade. Once the wing disc everts, the basal sides of the dorsal and ventral surface of wing epithelia come together and proveins from as broad gaps or lacunae. Approximately 6-8 hours APF, the wing secretes an apical cuticle. Soon after this, the wing inflates and the dorsal and the ventral surfaces are again apart. The vein morphology reappears by the apposition of the dorsal and ventral surfaces around 18-30 hours APF. The pupal cuticle detaches itself from the apical surface and final adult cuticle is secreted at around 36 hours APF (Blair, 2007). Cells in the posterior compartment communicate with the adjoining cells in the anterior compartment by through localized expression of secreted molecule called Hh (Lee et al., 1992). Hh activity in posterior cells is required for development of both the compartments in imaginal discs. Hh can induce patterning only in the anterior compartment even though it is expressed in the entire posterior compartment. It induces expression of Dpp in a narrow stripe of cells near the compartment boundary (A-P) (Nellen et al., 1994). Dpp in turn organises

patterning and controls growth symmetrically in both the compartments. Hh also turns on patched (ptc), knot (kn) and en. Dpp acts as a morphogen and turns on vestigial (vg), optomoter blind (omb) and spalt (sal) at successively higher thresholds. These genes are responsible for patterning wing blade in various ways along the A- P axis.

The DV boundary too serves as an organizing center that controls growth and specifies spatial pattern along the dorsal ventral axis. Dl and Ser serve as the ligands of Notch and activate their receptor at the DV boundary (Rebay et al., 1991). Activation of Notch at the boundary leads to activation of Wg and Cut (Ct) at the margin. Wg acts a morphogen and diffuses in either direction of the D- V boundary to shape the wing blade. Wg also refines its expression and promotes the expression of Dl and Ser on either side of its expression domain to create a positive feedback loop that maintains Notch signalling and Wg expression. Expression of Ct gene product at the margin is responsible for maintenance of Wg expression as well as preventing the Wg target genes from responding to it at the margin. Wg targets Vg, Distalless (Dll), Achaete (Ach) at increasingly higher thresholds.

Haltere is a balancing organ found in all dipterans in the third thoracic segment by modification of the wing fate, affected by Ubx. Ubx functions at multiple levels in hierarchy of wing disc development and promotes the formation of halteres by suppression of various wing patterning genes like *wg* and *cut* (Galant et al., 2002; Prasad et al., 2003; Shashidhara et al., 1999; Weatherbee et al., 1998). Various components of important signalling pathways that are responsible for shaping the wing (Dpp, wnt, Notch, EGFR and Hedgehog signalling pathway) have been found to be targets of Ubx by microarray studies and also in studies that have focussed on understanding the role of these pathways in shaping the halteres.

1.4 Impetus for the present work

Hox genes are master control genes, function by regulating the expression of downstream target genes. It is now widely accepted that evolution at the level of a family of highly conserved (from insects to human) genes popularly known as Hox genes has led to the diversity in animal body plan that we see now. Suppression of wing fate and specification of haltere fate in *Drosophila* by the Hox gene

Ultrabithorax (*Ubx*) is a classical example of Hox regulation of serial homology, which has served as a paradigmfor understanding Hox gene function (Lewis, 1978).

Hox gene *Ultrabithorax (Ubx)* which is expressed in the third thoracic segment in *Drosophila* suppresses wing development and specifies haltere development in *Drosophila*. The molecular mechanism by which Ubx recognizes and regulates its targets continues to mystify the researchers. This is of particular interest because all the Hox proteins have very similar 60 amino acid long DNA-binding homeodomains and when these are swapped between two Hox proteins, they show similar sequence specificity in vitro. The observations certain chromatin factors serve as docking platforms for Ubx to recognize and bind its targets (Agrawal et al., 2011). This is consistent with the observations that Hox proteins acquire target-specificity by binding to cofactors (Slattery et al., 2011).

When compared across insects, diversity in the wing morphology and number is evident (Figure 5). While the insects *Apis mellifera* (order Hymenoptera),*Bombyx mori* (order Lepidoptera) and *Tribolium custaneum* (order Coleoptera)possess of two pairs of wings which is a characteristic of ancestral pterygotes, the fruit fly *Drosophila melanogaster* has its hind wing modified to a more specialized organ known as haltere which is a balancing organ in these flies. Wings are dorsal structures while legs develop ventrally. The dorsal morphology of the third thoracic segment varies across the insects. *Apis mellifera* has two sets of wings and the hind wings are slightly smaller in size compared to the fore wings and they also have a different venation pattern. In the silkmoth, *Bombx mori*, fore wings and hind wings are similar. The red flour beetle, *Tribolium custaneum* has two sets of wings where the fore wings are modified into more specialized wings type called elytra while the hind wings remain more ancestral type.
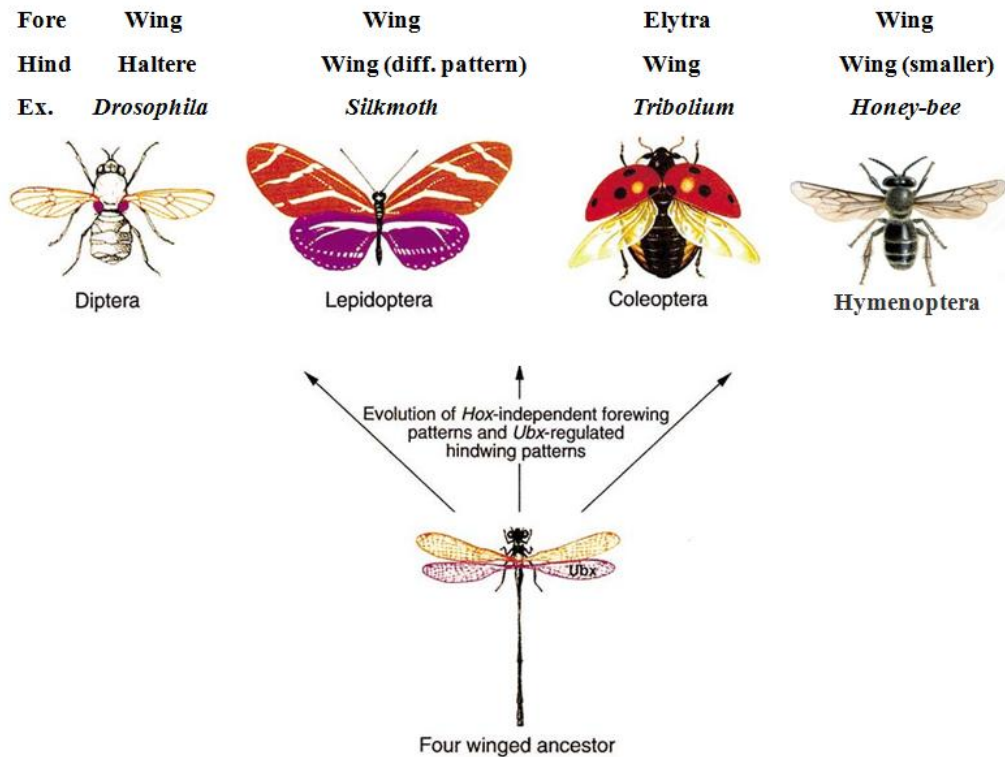
Figure 5. Diversity in the wing appendages in different insects

1.3.1 Ubx expression in various insects and its function

During the course of evolution (Figure 6); the expression and function of Ubx has remained conserved in various insects. However, the third thoracic segment which is shaped by Ubx has been modified to various degrees in different insects. Though the dorsal second thoracic organ is generally a wing in most insects, the dorsal third thoracic organ morphology varies a lot when compared to its second thoracic counterparts- a haltere in fruit fly while almost a similar hind wing in dragon flies.

In the fore winged ancestor, the dragon fly, *Ubx* is expressed in both T2 and T3 segments and it develops similar forewings and hindwings. In *Apis mellifera*, foreqwings and hindwings differ slightly in size and the venation pattern but *Ubx* is expressed in both forewings and hindwings buds although it is higher in the hindwing buds (Phd thesis of Naveen Prasad). Even in case of *Bombyx mori*, the silkmoth, both forewing and hindwing patterns are similar although *Ubx* is expressed only in the hindwings. The floor beetle, *Tribolium cusatneum* shows an interesting phenomenon; in this insect, Ubx which is expressed only in the T3 segment, is

required to 'despecialize' the development of wing form and thus repress the elytron fate of the T3 segment (Tomoyasu et al., 2005). *Ubx* expression is limited to the peripodial membrane of the wingdisc and is present throughout the haltere discs in case of *Drosophila*. This has been summarised in the figure 7.
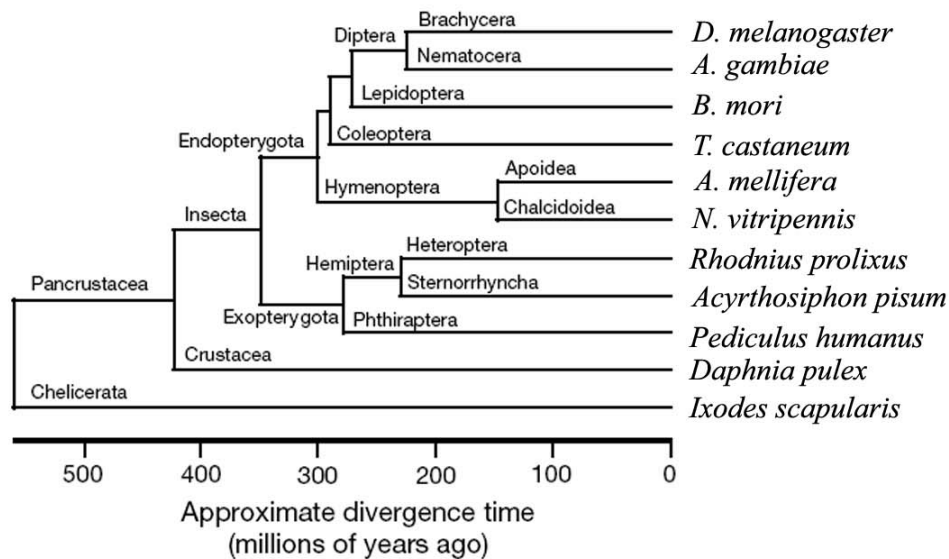
Figure 6. Evolution of insects from Endopterygotes; Figure courtesy: Porcellli et al., 2007
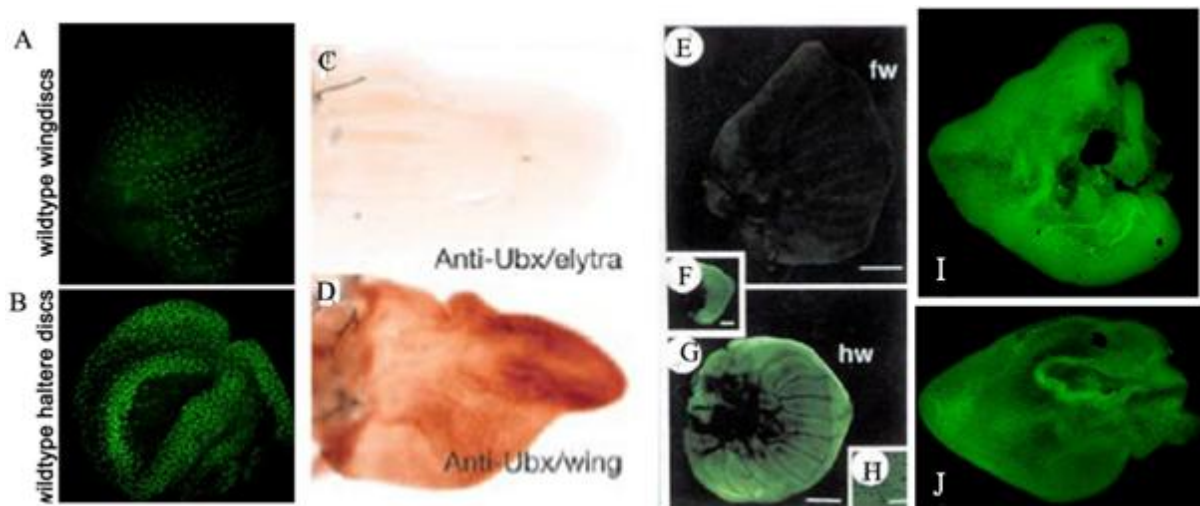
Figure 7. Ubx expression in thoracic segments of various insects

Various wing patterning genes like *wg* and *cut*and *vestigial* are differentially regulated in wing and haltere discs of *Drosophila* while they are not differentially regulated in forewing and hindwings of other insects. (Galant et al., 2002; Prasad et al., 2003; Shashidhara et al., 1999; Weatherbee et al., 1998) (Phd thesis of Naveen

Prasad). Question of our interest is how Ubx is able to differentially regulate these genes in case of *Drosophila* and not in other insects. We aim at understanding to what extent Ubx function is conserved among these insects over past 300 million years of evolution. Are the differences in the Ubx sequences from these insects significant to recruit different cofactors and thus regulate genes differentially in case of *Drosophila*? And/or the differences in the enhancer elements of the genes (where Ubx binds) are responsible for the differential regulation.

1.3.3 Objectives of the study

1. Generating transgenic *Drosophila* for full-length Ubx cDNA derived from *Apis*, *Bombyx* and *Tribolium*. The cDNA would be cloned downstream of an inducible promoter (employing GAL4-UAS system) that helps to achieve tissue- and temporal-specificity in over-expressing Ubx.

2. Over-expressed heterologous Ubx would be tested for its ability to (a) induce wing-to-haltere transformation, (b) replace endogenous Ubx and (c) suppress or enhance specific targets of *Drosophila* Ubx during haltere development.

3. Certain targets of Ubx in Apis that are also targets of Ubx in *Drosophila*, show differential expression pattern between wing and haltere in *Drosophila*, but not between forewing and hindwing in *Apis*. We would identify regulatory regions of one or two of such genes in Apis and express the same in transgenic *Drosophila* and test their ability to drive a reporter gene. This may help us understand the reason for differential regulation of corresponding genes in*Drosophila*: is it due to Ubx and its cofactors in *Drosophila*? and/or due to specific differences in the sequences of the regulatory regions?

This exercise is expected to provide insights molecular changes at the level of Hox protein and its targets during the evolution of wing morphology in insects.

Chapter 2

Materials and Methods

2.1 Routine Fly methods

2.1.1 Fly Husbandry

Canton-S strain of *Drosophila melanogaster* was used as a wild type strain. Fly culture was maintained on the regular cornmeal/sucrose agar at 25° C (Ashburner, 1989). Virgin females and males used for the crosses were not more than 10 days old. Virgin females can be identified by pale pigmentation and a dark spot visible on their abdomen also known as the meconium. Males can be identified by the presence of sex combs on the first pair of legs and dark bands on the abdomen.

FM7a was used to balance mutations/P insertions on the first chromosome, *CyO* as the second chromosome balancer and *TM3Sb* or *TM3B* as the third chromosome balancer. GAL4-UAS system (Brand and Perrimon, 1993) was used for targeted mis-expression of gene products.

2.1.2 Fly stocks used in the study

2.1.2.1 GAL4 drivers

vg-GAL4 (Simmonds et al., 1995); *sd*-Gal4 (Ayeni et al., 2014)

2.2.2.2 Other fly Lines

UAS-*Ubx*$_D$(gift from Savita Singh); UAS-Ubx$_D$-FLAG (gift from Harsha T);$\frac{\text{UAS}-Ubx_A-\text{FLAG}}{\text{CyO}}$(line1); $\frac{\text{UAS}-Ubx_A-\text{FLAG}}{\text{CyO}}$(line2); $\frac{\text{UAS}-Ubx_A-\text{FLAG}}{\text{TM3Sb}}$(line3) (generated in this study)

Vg-Gal4;$\frac{\text{UAS}-Ubx_A-\text{FLAG}}{\text{CyO}}$ line; $\frac{vg-\text{Gal4}}{\text{CyO}}$; $\frac{\text{Quadvglacz}}{\text{TM3B}}$ (Gift from Savita Singh)

2.2.2.3 Enhancer-GFP constructs

Avgqenb/CyO line-1; Avgqenb/TM6Tb line-2; Avgqena/CyO line-1 (generated in this study); Dcutena/CyO line 2; Dcutena/CyO line 3; Dcutena/TM3Sb line 1; Dcutena/TM3Sb line 4; Awgena/CyO line1; Awgenb/TM3Sb line1; Dwgena/TM3Sb line1(generated by Naveen Prasad)

2.1.3 Generation of Transgenic flies

2.1.3.2 Preparation of DNA for Injection

The gene/DNA sequence of interest was cloned between the terminal repeats of the P-element based vector (pUAST) containing a wild type copy of the eye colour gene, *white+*. The purified plasmid DNA (using Qiagen Tip-10) to be injected was mixed in milli-Q water (pH above 7) to a final concentration of 3-4μg/μl and sent for injections to NCBS fly facility, Bangalore. The DNA mix was injected into the poleplasm of 0-1 hour old *yw* embryos (also known as *w1118*) with helper plasmid carrying transposon, by standard procedures.

2.1.3.3 Selecting the transformants

While *yw* flies have white eye color, the transgenics generated for gene of interest in the P-element based vector that contains a wild type copy of the eye color gene *white+* would have red eye color. The embryos microinjected with the gene of interest were maintained at 18°C until the adult flies eclosed. Freshly eclosed adult flies were crossed to wild type *yw* flies and the progeny of this cross was scored for the presence of the eye color marker. Red-eyed flies were collected and stocks of these transgenics were generated by balancing to mark the location of chromosomal insertion in these flies.

2.1.4 Larval Dissections

Larval dissections were performed to obtain wing, haltere imaginal discs for immuno-histochemistry. Wandering third instar larvae were collected in a cavity block containing chilled Phosphate Buffered Saline (PBS, pH 7.4; Sigma, USA) solution. After a wash in the PBS, two third of the larvae was cut out and the anterior one third was turned inside out with the help of a pair of fine needles (B.D. Insulin syringes). The larval tissue was either fixed in 1% or 4% formaldehyde solution or immediately after flipping the larvae inside-out discs were isolated and used for staining.

2.1.5 Imaginal Disc Staining

Immunohistochemical staining was performed essentially as described by Patel et al.(1989).

Wandering third instar larvae were collected in a dissecting dish with PBS (pH7.4). Larvae were cut into half with the help of a pair of forceps and scissors and the anterior part, which contains all imaginal discs was turned inside out. The fat body, gut and salivary glands were removed carefully without damaging the discs. The clean anterior larval body with the discs attached to it was transferred to a microfuge tube with PBS on ice. Usually 10 to 20 larvae were dissected. The dissected larvae were fixed for 20 minutes in PBS (pH7.4) with 4% paraformaldehyde. After fixation the larvae were rinsed 5-6 times with PBS and blocked for 1 hour in PBTx (PBS + 0.1% TritonX-100 (Merck) + 0.5% BSA (Sigma)). Incubation with primary antibody in PBTx was done overnight at 4 ^{o}C or 3hrs at room temperature. Afterwards the larvae were washed with PBTx for a period of two hours, changing the wash buffer every 15-30 minutes. Incubation with the secondary antibody in PBTx was done for 2 hours at room temperature. After this incubation, 4 washes with PBTx of 15 minutes each were carried out. After the last wash, all PBTx was removed and the larvae were covered with mounting medium (PBS + 80% glycerol (anhydrous, Merck). The larval bodies were stored in mounting medium at 4°C overnight to allow the tissue to be saturated with the medium. The next day discs were detached from the rest of the larval body, collected in a drop of mounting medium on a glass slide and covered with a cover slip.

2.1.6 Antibodies used in this study

Primary antibodies used were rabbit anti-N-terminal Ubx$_D$ 1:2000(Agrawal et al. 2011); rabbit anti-flag 1:1000 (kind gift from Girish Ratnaparkhi); mouse Anti Wingless 1:500 (monoclonal from DSHB / (Brook and Cohen, 1996)); mouse Anti Cut 1:50 (monoclonal from DSHB, (Blochlinger et al., 1993)); mouse Anti β-gal 1:1000 (from DSHB). Secondary antibodies used were Goat Anti-mouse Alexa 488 1:1000 (Molecular Probes, Invitrogen) 1:1000.; Goat Anti-rabbit Alexa 488 1:1000 (Molecular Probes, Invitrogen); Goat Anti-rabbit Alexa 568 1:1000 (Molecular Probes, Invitrogen).

2.1.7 Microscopy and image formatting

All the images were taken on the Apotome fluorescence microscope and Image J software was used to crop and put scale bars on the images.

2.2 Cloning of full length Ubx cDNA (tagged to FLAG) from from *Apis*, *Bombyx* and *Tribolium*in pUAST-FLAG vector

Full length Ubx cDNA (tagged to FLAG) from *Apis*, *Bombyx* and *Tribolium* were subcloned into pUAST-flag vector was done and were sent for P-element transgenic insertions to NCBS fly facility, Bangalore after confirming their sequences. pUAST-flag (size ~ 9000 bp) is a P-element based vector. P elements are transposons present in the fly genome. They carry inverted repeats at their ends and carry a transposase gene which is needed for their movement. These P-elements have been modified artificially so that they can used to integrate a gene of interest (cloned into them) into the fly genome. Provided the transposase enzyme externally, the gene of integrates into the genome. pUAST-FLAG vector has five tandemly arrayed GAL4 binding sites, an hsp70 TATA promoter, a FLAG tag and multiple cloning sites containing unique restriction sites for NotI, EcoRI, BgIII, XhoI, KpnI. It also has a white gene marker that can be used to score transgenic flies. The FLAG tag is tagged to N-terminal of each Ubx protein (Brand and Perrimon, 1993).

1. Ubx full length cDNA from *Apis mellifera* cloned into pUAST-FLAG vector:

The full length cDNA was previously cloned into pGEM-T Easy vector. It was then subcloned into pUAST-FLAG. The cloning strategy was to amplify the Ubx gene from pGEM-T Easy vector with the sites for suitable restriction enzymes incorporated in it (which do not cut the Apis Ubx gene).

Following primers were designed and used for the amplification of Ubx:

Forward Primer: 5' ACGG GGTACC ATGAACTCGTATTTTGAGCAGACTGC 3'
Reverse Primer: 5' AGGC TCTAGA CTAGTTGGCCCCCTCCGG 3'

The restriction sites used were: **KpnI: ggtac|c and XbaI: t|ctaga**

Total size of the insert = 992 bp

Thirty rounds of PCR amplification of the Ubx gene were done using pfu polymerase using extension at 72 °C for 45 seconds and the annealing temperature used was 56° C.After the amplification, the insert and the pUAST-FLAG vector were cut with KpnI and XbaI sequentially (one after the other) and then ligated using T4 DNA ligase. Vector transformed DH5 alpha colonies were screened using primers specific

to pH-stinger vector. Vector was purified using Quiagen- miniprep kit and sent for sequencing. Sequencing results and the blast analysis are appended in the results section. Sequenced clones were sent for injections and seven transgenic lines have been received for the same.

2. Ubx full length cDNA from *Bombyx mori* cloned into pUAST-FLAG vector:

Primers were synthesized for amplifying the Ubx gene of *Bombyx* from its cDNA. These primers were sent to CDFD, Hyderabad and full length Ubx cDNA amplicon was requested. Full length cDNA of *BombyxUbx* was obtained from CDFD, Hyderabad. Using this amplicon as a template, an insert containing Ubx sequence and suitable restriction sites was created using PCR. Following primer sequences were used for the amplification. Thirty rounds of PCR amplification of the Ubx gene were done using pfu polymerase using extension at 72 °C for 45 seconds and the annealing temperature used was 58° C.

Forward Primer: 5' ACGG GGTACC ATGAACTCTTACTTCGAGCAGGGTG 3'
Reverse Primer: 5' ATGC TCTAGA TTAATGTTCGGGGTGTCCCTGG 3'

The restriction sites used were: **KpnI: ggtac|c and XbaI: t|ctaga**

Total size of the insert = 764 bp

After the amplification, the insert and the pUAST-FLAG vector were cut with KpnI and XbaI sequentially (one after the other) and then ligated using T4 DNA ligase. Vector transformed DH5 alpha colonies were screened using primers specific to pH-stinger vector. Vector was purified using Quiagen- miniprep kit and sent for sequencing. Sequencing results and the blast analysis are appended in the results section. Sequenced clones were sent for injections and nine transgenic lines have been received for the same.

3. Ubx full length cDNA from *Tribolium custaneum* cloned into pUAST-FLAG vector:

The full length cDNA was previously cloned into a pGEX vector. It was then subcloned into pUAST-FLAG. This particular full length clone is known as the 'a' isoform of Tribolium custaneum Ubx. This isoform lacks a DSMTF amino acid sequence at position 213 of the protein sequence which is otherwise present in the

'b' isoform of TcUbx. This isoform has been used in earlier studies and it has been shown that it is able to induce same phenotypes as that of *Drosophila* Ubx when expressed ectopically in the embryonic ectoderm. It is able to transform segmental identity from thoracic to abdominal and also represses the activity of Distal-less enhancer which is directly regulated by *Drosophila* Ubx (Galant et al., 2002). So we hope that function of the protein is not affected by this deletion.

The cloning strategy here was to amplify the Ubx gene from the pGEX vector with the sites for suitable restriction enzymes incorporated in it (which do not cut the Tribolium Ubx gene).

Following primers were designed and used for the amplification of Ubx:

Forward Primer: 5' ACGG <u>GAATTC</u> TATGAACTCTTACTTCGAGCAGAGC 3'
Reverse Primer: 5' AGGC <u>TCTAGA</u> CTAATTCGGGTCCACTTGTGCG 3'

The restriction sites used for this were **EcoRI g|aattc** and **XbaI t|ctaga**

Total size of the insert = 948 bp

Thirty rounds of PCR amplification of the Ubx gene were done using pfu polymerase using extension at 72 °C for 45 seconds and the annealing temperature used was 53° C.

After the amplification, the insert and the pUAST-FLAG vector were cut with EcoRI and XbaI sequentially (one after the other) and then ligated using T4 DNA ligase. Vector transformed DH5 alpha colonies were screened using primers specific to pH-stinger vector. Vector was purified using Quiagen- miniprep kit and sent for sequencing. Sequencing results and the blast analysis are appended in the results section. Sequenced clones were sent for injections and seven transgenic lines have been received for the same.

4. Cloning of vestigial quadrant enhancer in pHstinger:

Vestigial quadrant enhancer in *Drosophila* is reported to be of 891 bp in length. One such corresponding region in the 4th intron of *Apis mellifera* which was bound by Ubx was found to be of 700 bp in length. Out of these 700 base pairs, 576 base pairs' region was found highly conserved amongst the species of Honey bees. We

suspected this sequence to be the analogue of vestigial quadrant enhancer in *Drosophila*(The in-silico analysis and the primer designing were done by Naveen Prasad).

A 2014 bp region flanking the highly conserved 576 bp region was first cloned into pGEMT-Easy vector. Then from there, the 576 bp region of interest was cloned into pH-stinger vector which was then sent for injection. pH-stinger is a transgenic reporter construct which has GFP as the reporter gene (Barolo et al., 2004).

First, the genomic DNA from the posterior half of the body of *Apis mellifera* was extracted using standard phenol-chloroform extraction method. The 2014 bp region was then amplified from the genomic DNA using following primer sequences.

Forward primer: 5' TTCGTCGCTCGAATTCACCA 3'

Reverse Primer: 5' TTACGGCGGTTTGCTTTTGG 3'

Size of the amplicon= 2014 bp

PCR was performed using high fidelity Phusion DNA polymerase, extension carried out at 72° C for 1 min 30 sec and the annealing temperature used was 58° C. The PCR product was then gel purified and ligated into pGEMT –Easy vector using pGEMT-Easy TA cloning kit. Positive colonies were selected using the standard blue white colony screening process.

This pGEMT-Easy clone was then used to amplify the 576 bp region as mentioned above. Following primers were used for amplifying the region of interest with the sites for suitable restriction enzymes incorporated into it.

Forward primer: 5' GC TCTAGA CTTCTCGCGAGAAACGAGAGGC 3'
Reverse primer: 5' CG GGATCC GTGGACAGTGACGAGGACACG 3'

Restriction enzymes used for the digestion were: **XbaI t|ctaga and BamHI g|gatcc**

Total size of the insert: 576 bp

Thirty rounds of PCR amplification were carried out using high fidelity Phusion DNA polymerase, extension was carried out at 72 °C for 30 seconds and the annealing temperature used was 58° C. After the amplification, both pHstinget vector and the

insert were cut by using XbaI and BamHI restriction enzymes (double digestion method). The vector and the insert were then ligated using T4 DNA ligase enzyme. Vector transformed DH5 alpha colonies were screened using primers specific to pH-stinger vector. Vector was purified using Quiagen- miniprep kit and sent for sequencing. Sequencing results and the blast analysis are appended in the results section. Sequenced clones were sent for injections and three transgenic lines were received for the same.

Chapter 3

Results

3.1 Full length Ubx cDNA from *Apis mellifera* cloned into pUAST-FLAG

Figure 8. (Left) shows full length Ubx cDNA from *Apis* amplified from a previous pGEMT-Easy vector, lane1 is 100 bp step up ladder from Bngalore Genei, lane2 is the amplicon; (Right) shows Ubx cDNA from *Apis* cloned into pUAST-FLAG, lane1 is 1kb step up DNA ladder from Bangalore Genei, lane 2 is the construct; (0.7% Agarose gels were used for running the DNA)

3.2 Full length Ubx cDNA from *Bombyx mori* cloned into pUAST-FLAG
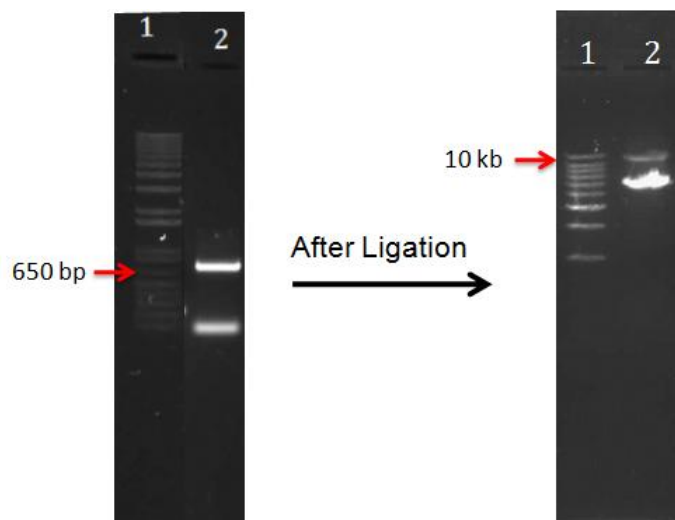
Figure 9. (Left) shows full length Ubx cDNA from *Bombyx* amplified from *Bombyx* cDNA, lane1 is 1kb+ DNA ladder from Invitrpgen, lane2 is the amplicon; (Right) shows Ubx cDNA from

Bombyx cloned into pUAST-FLAG, lane1 is 1 kb step up DNA ladder from Bangalore Genei, lane2 is the construct (0.7% Agarose gels were used for running the DNA)

3.3 Full length Ubx cDNA from *Tribolium custaneum* cloned into pUAST-FLAG
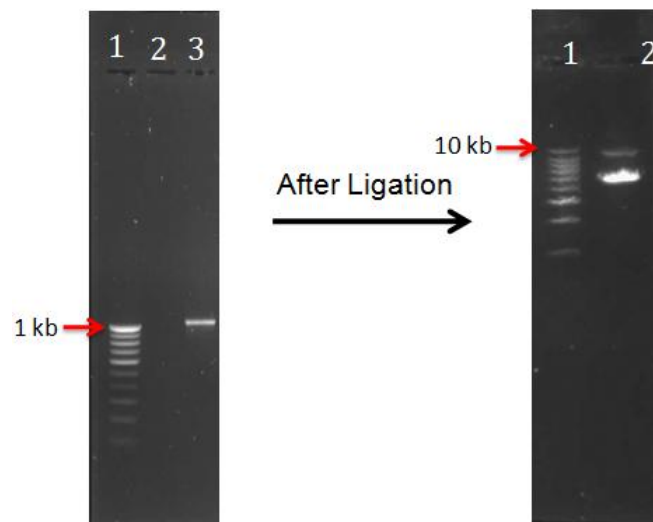
Figure 10. (Left) shows full length Ubx cDNA from *Tribolium* amplified from the previous clone, lane1 is 100 bp DNA ladder from Bangalore Genei, lane2 is the amplicon; (Right) shows Ubx cDNA from *Tribolium* cloned into pUAST-FLAG, lane1 is 1 kb step up DNA ladder from Bangalore Genei, lane2 is the construct (0.7% Agarose gels were used for running the DNA)

3.4 *Vestigial* quadrant enhancer from *Apis mellifera* cloned into pHstinger
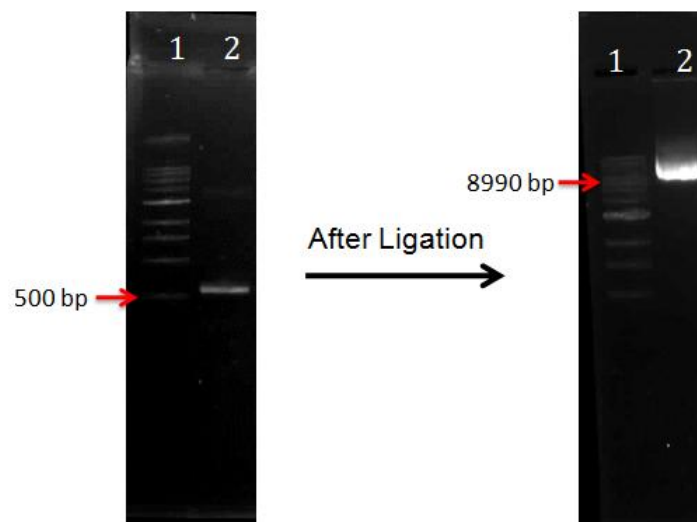
Figure 11. (Left): shows full length Ubx cDNA from *Tribolium* amplified from the previous clone, lane1 is 100 bp DNA ladder from Bangalore Genei, lane2 is the amplicon; (Right): shows Ubx cDNA from *Tribolium* cloned into pUAST-FLAG, lane1 is 1 kb step up DNA ladder from Bangalore Genei, lane2 is the construct (0.7% Agarose gels were used for running the DNA)

3.5 Adult phenotypes induced by over-expression of Ubx from *Apis* and *Bombyx*

Ubx from *Apis* and *Bombyx* were tested for its ability to induce wing to haltere transformation. Three independent transgenic lines (mentioned in the materials and methods section) for *Apis* Ubx and one for *Bombyx* Ubx were used for the same purpose. Males homozygous for the *Apis* Ubx gene were collected and crossed to virgin females of two wing specific Gal4 drivers namely *vg*-Gal4 and *sd*-Gal4 and maintained at 25° C.

Figure 12 (D, E and F) shows adult phenotypes when Ubx from *Apis* was over-expressed in the T2 thoracic segment of *Drosophila* using *vg*-Gal4 driver (T2 and T3 segments are focussed in the pictures). Figure 12 G shows over-expression phenotype given by *Bombyx* Ubx. These phenotypes are compared with the adult phenotypes given by over-expression of Ubx from *Drosophila* using the same Gal4 driver (Figure 12 B & C).Figure 5 B and C show over-expression phenotypes given by Ubx_D and by Ubx_D tagged to the FLAG tag. We can see that over-expression of Ubx from *Apis* and *Bombyx* is able to repress wing development in *Drosophila*, similar to the repression caused by Ubx from *Drosophila* itself. Variation in the degree of repression is seen in this case which might depend on the degree of over-expression of the Ubx protein. The penetrance is almost 100% in case of both the Ubx. *Drosophila* Ubx (n= 53, N= 60), *Apis* Ubx (n=64, N=64), *Bombyx* Ubx (n=67, N=67).

Figure 13 (D, E & F) shows adult phenotypes shown by over-expression of Ubx from *Apis mellifera* induced using *sd*-Gal4 for three independent transgenic lines. Figure 13 G shows adult phenotypes shown by over-expression of Ubx from *Bombyx*. These phenotypes are compared with the phenotypes given by over-expression of Ubx from *Drosophila* using the same Gal4 driver (B & C). Figure 13 (B and C) respectively show the over-expression phenotypes given by Ubx_D and by Ubx_D tagged to the FLAG tag. Both the Ubx cause almost a complete loss of the wing appendage. Observed phenotype is stronger than that observed in case of *vg*-Gal4. Penetrance is 100% for all the three the Ubx: *Drosophila* Ubx (n= 50, N= 50), *Apis* Ubx (n= 54, N= 54), *Bombyx* Ubx (n= 48, N= 48).
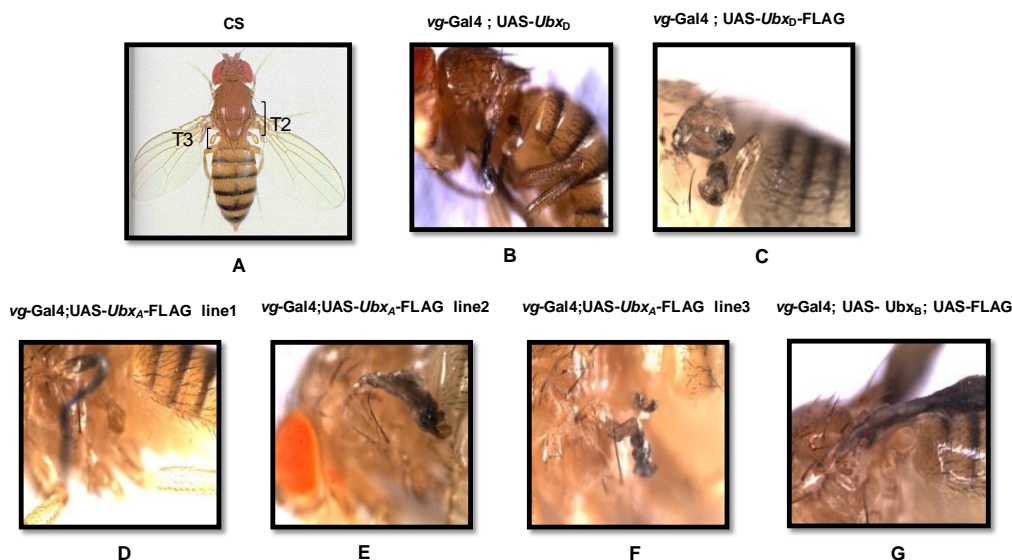
Figure 12. Adult phenotypes induced by over-expression of Ubx using *vg*-Gal4: A shows a wild type fly; D, E and F show over-expression phenotypes for three independent *Apis* Ubx transgenic lines. G shows over-expression phenotype for *Bombyx* Ubx. B and C show over-expression phenotypes for Ubx from *Drosophila*.
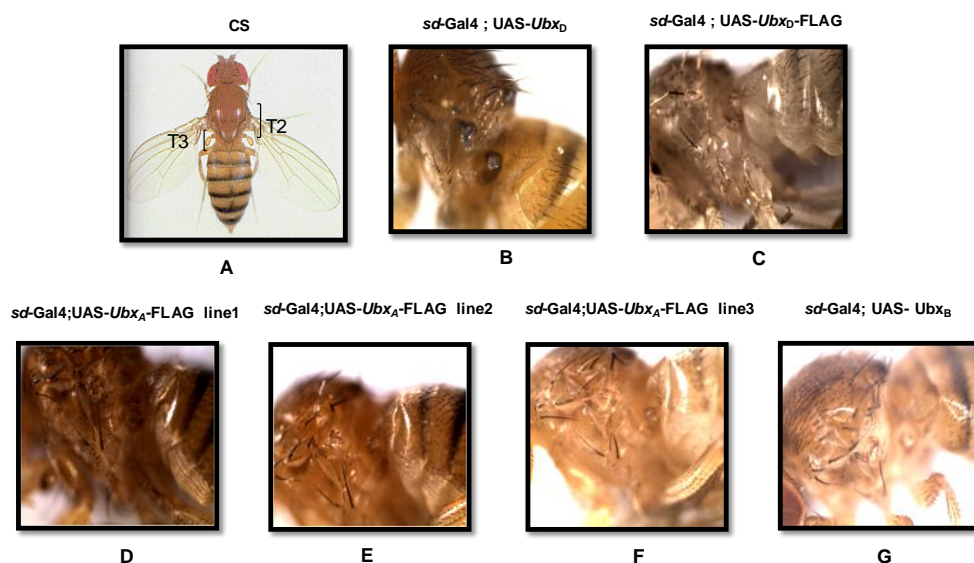
Figure 13. Adult phenotypes given by over-expression of Ubx using *sd*-Gal4: A shows a wild type fly; D, E and F show over-expression phenotypes for three independent *Apis* Ubx transgenic lines. B and C show over-expression phenotypes for Ubx from *Drosophila*. G shows overexpression phenotype shown by *Bombyx* Ubx

3.6 Suppression of specific targets of Ubx$_D$ during haltere development by Ubx from *Apis*

wingless and *cut* are direct targets of *Drosophila* Ubx during haltere development. These targets are differentially regulated by Ubx$_D$ in the wing and the haltere of *Drosophila* but not in the case of forewing and hindwing of *Apis*. As explained earlier,

wingless and *cut* are repressed in the posterior compartment of haltere by Ubx_D; while expressed in the entire D-V boundary in wing discs. *cut* is expressed in a similar manner in both the forewing and hindwing buds of *Apis* although Ubx is present in both. Also, *wingless* is not regulated differentially in the forewing and hindwing of *P. Coenia*. To test the ability of Ubx_A and Ubx_B to suppress/enhance these specific targets of Ubx_D, Ubx_A and Ubx_B were over-expressed in the wing imaginal discs of *Drosophila* using the wing specific Gal4s *vg*-Gal4 and *sd*-Gal4.

Over-expression of Ubx from *Drosophila* has been shown to repress *wingless* and *cut* expression in the posterior compartment of wing imaginal discs (Shashidhara et al., 1999). Figure 14 shows over-expression of Ubx_D and Ubx_D-FLAG in the D-V boundary of wing imaginal discs of *Drosophila* and its effect on *wingless* expression. A FLAG tagged Ubx_D was used as a control to eliminate the possibility that FLAG tag is in any way not changing the function of the protein. *wingless* is repressed in the posterior compartment of wing discs when Ubx from *Drosophila* is over-expressed in the wing disc D-V boundary using *vg*-Gal4 (as shown by the white arrows in figure 14). Penetrance is almost 100% in both the cases: *Drosophila*-Ubx (n=8, N=11), *Drosophila*-Ubx-FLAG (n= 20, N=20).

Figure 14. On the left hand side, *wingless* expression in the wing disc and haltere discs of *Drosophila* in wild type background; Panel A: Ubx_D over-expression using *vg*-Gal4 (red), *wingless* expression (green) and the merger; Panel B: Ubx_D-FLAG over-expression using *vg*-Gal4 (red), *wingless* expression (green) and the merger. White arrows show repression of *wingless* in the posterior wing disc.

Figure 15 shows over-expression of Apis-Ubx-FLAG (line1, 2 & 3) in the D-V boundary of wing imaginal discs of *Drosophila* and its effect on *wingless* expression. *Wingless* is repressed in the posterior compartment of wing discs when Ubx from *Apis mellifera* is over-expressed in the wing disc D-V boundary using *vg*-Gal4 (as shown by the white arrows in figure 15). Anti-N-terminal-Ubx_D antibody was used to stain these discs to eliminate the possibility that the phenotype might be due to over-expression of *Drosophila* Ubx itself. As we can see in figure 15, wild type Ubx_D expression is seen when these wing discs are stained using anti-N-terminal-Ubx_D antibody. Hence we conclude that the phenotype observed and the *wingless* repression is indeed an effect of over-expression of Ubx_A. Penetrance is 100% in all the three cases (n= 26, N=26).

Figure 15. Panel A: Apis-Ubx-FLAG (line1) over-expression using *vg*-Gal4 (anti-FLAG, red), *wingless* expression (green) and the merger; Panel B: Apis-Ubx-FLAG (line2) over-expression using *vg*-Gal4 (anti-FLAG, red), *wingless* expression (green) and the merger. Panel C: Apis-Ubx-FLAG (line3) over-expression using *vg*-Gal4 (anti-FLAG, red), *wingless* expression (green) and the merger; 4th column in the figure shows staining of these discs done using antibody against *Drosophila* Ubx; White arrows show repression of *wingless* in the posterior wing disc.

3.7 Suppression *wingless* by Ubx from *Bombyx*

Figure 16 shows over-expression of *Bombyx*-Ubx-FLAG in the D-V boundary of wing imaginal discs of *Drosophila* and its effect on *wingless* expression. *Wingless* is repressed in the posterior compartment of wing discs when Ubx from *Bombyx* is over-expressed using *vg*-Gal4 (n= 15, N=17) and there is complete loss of wingless in case of over-expression by sd-Gal4 drivers (n= 13, N= 14) (as shown by the white arrows in figure 16).
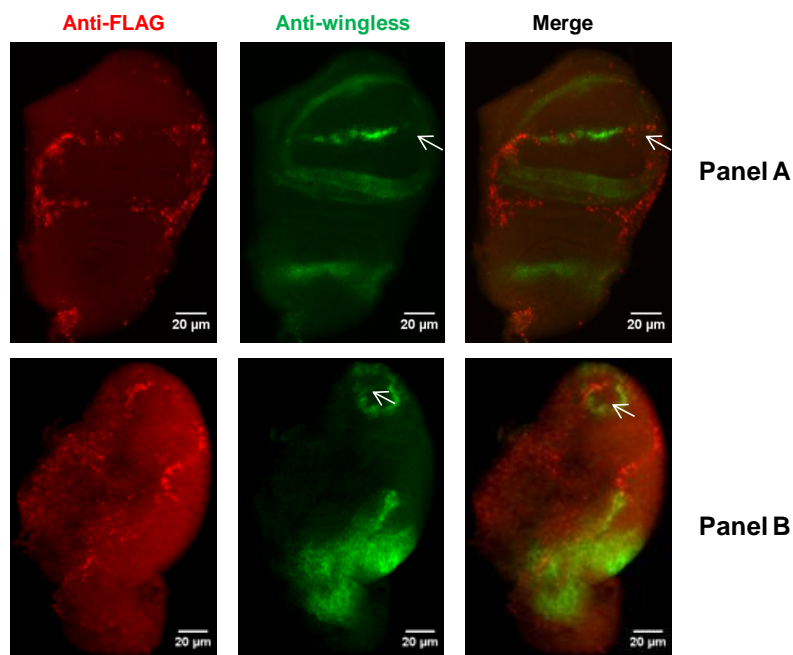
Figure 16. Panel A: *Bombyx*-Ubx-FLAG over-expression using *vg*-Gal4 (anti-FLAG, red), *wingless* expression (green) and the merger; Panel B: *Bombyx*-Ubx-FLAG over-expression using *sd*-Gal4 (anti-FLAG, red), *wingless* expression (green) and the merger; White arrows show repression of *wingless*

Similar experiments were performed using *sd*-Gal4 for the over-expression of Ubx. Figure 17 shows over-expression of Ubx$_D$ and Ubx$_D$-FLAG in the wing pouch of wing imaginal discs of *Drosophila* and its effect on *wingless* expression. *Wingless* is repressed from the entire D-V boundary of wing discs when Ubx from *Drosophila* is over-expressed in the wing pouch using *sd*-Gal4 (as shown by the white arrows in figure 17). Similar to what happened in case of adult phenotypes, stronger repression of *wingless* was observed when Ubx$_D$ was over-expressed using *sd*-Gal4.
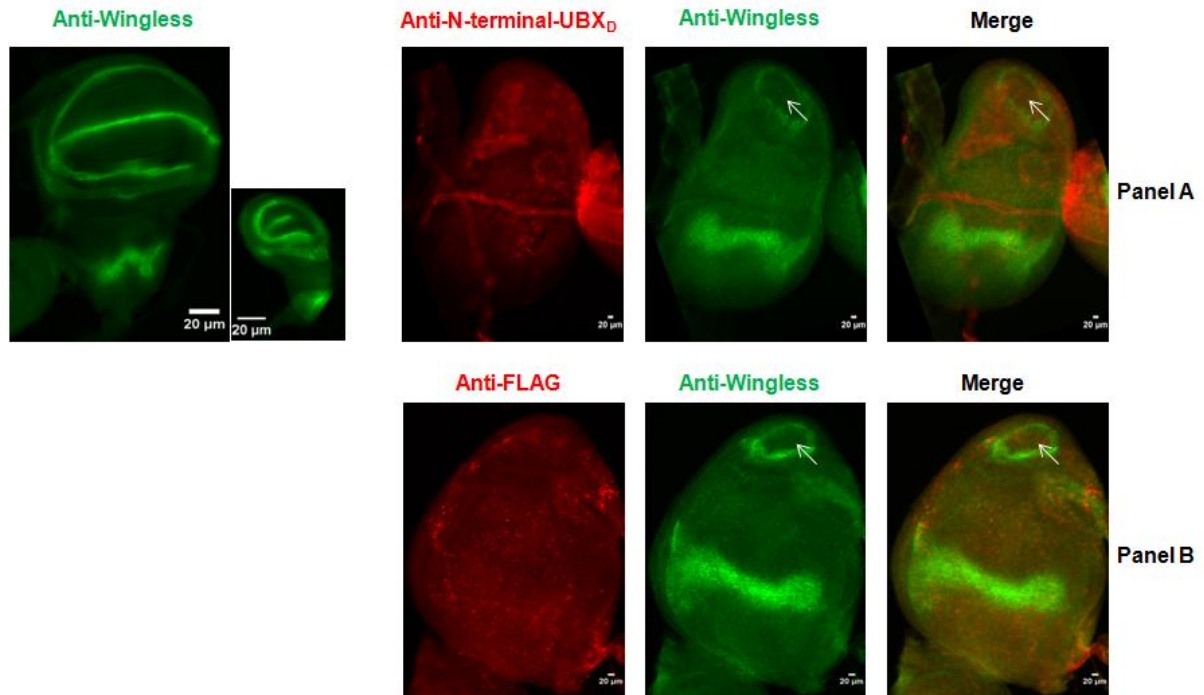
Figure 17. On the left hand side, *wingless* expression in the wing disc and haltere discs of *Drosophila* in wild type background; Panel A: Ubx$_D$ over-expression using *sd*-Gal4 (red), *wingless* expression (green) and the merger; Panel B: Ubx$_D$-FLAGover-expression using *sd*-Gal4 (red), *wingless* expression (green) and the merger. White arrows show repression of *wingless* from the entire D-V boundary.

Figure 18 shows over-expression of Apis-Ubx-FLAG (line1, 2 & 3) in the pouch region of wing imaginal discs of *Drosophila* and its effect on *wingless* expression. *Wingless* is repressed in the entire D-V boundary of wing discs when Ubx from *Apis* is over-expressed using *sd*-Gal4 (as shown by the white arrows in figure 18). Anti-N-terminal-Ubx$_D$ antibody was used to stain these discs to eliminate the possibility that the phenotype might be due to over-expression of *Drosophila* Ubx itself. As we can see in figure 18, wild type Ubx$_D$ expression is seen when these wing discs are stained using anti-N-terminal-Ubx$_D$ antibody. Therefore, the disc phenotype observed and the *wingless* repression is indeed an effect of over-expression of Ubx$_A$. Penetrance is 100% in all the three cases (n= 18, N= 18).

Ubx from *Apis mellifera* was also tested for its ability to repress *cut* as done by *Drosophila* Ubx during haltere development. Figure 19 shows over-expression of Ubx$_D$ and Ubx$_D$-FLAG in the D-V boundary of wing imaginal discs of *Drosophila* and its effect on *cut* expression. *cut* is repressed from the entire D-V boundary of wing discs when Ubx from *Drosophila* is over-expressed using *vg*-Gal4 (as shown by the white arrows in figure 19).
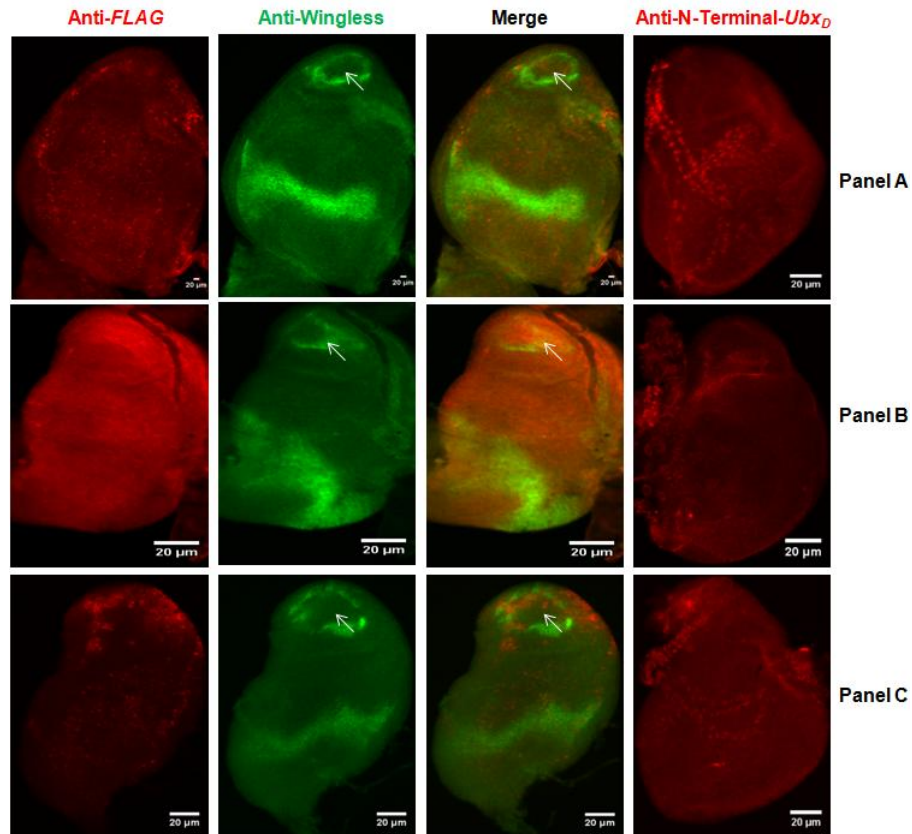
Figure 18. Panel A: *Apis*-Ubx-FLAG (line1) over-expression using *sd*-Gal4 (anti-FLAG, red), *wingless* expression (green) and the merger; **Panel B:** *Apis*-Ubx-FLAG (line2) over-expression using *sd*-Gal4 (anti-FLAG, red), *wingless* expression (green) and the merger. **Panel C:** Apis-Ubx-FLAG (line3) over-expression using *sd*-Gal4 (anti-FLAG, red), *wingless* expression (green) and the merger; 4th Column in the figure shows staining of these discs done using antibody against *Drosophila* Ubx; White arrows show repression of *wingless* from the entire D-V boundary.

Figure 20 shows over-expression of *Apis*-Ubx-FLAG (line1, 2 & 3) and *Bombyx*-Ubx-FLAG in the D-V boundary of wing imaginal discs of *Drosophila* and its effect on *cut* expression. FLAG-staining has not worked very well in these cases but owing to the disc phenotype, it is expected that Ubx$_A$ is getting over-expressed in these discs. *cut* is repressed from the entire D-V boundary of wing discs when Ubx from *Apis mellifera* is over-expressed using *vg*-Gal4 (as shown by the white arrows in figure 20) in case of line 1 and line 2. In case of *Apis*-Ubx-FLAG line 3 and *Bombyx*-Ubx, *cut* repressed only in the posterior compartment of the wing disc. Anti-N-terminal-Ubx$_D$ antibody was used to stain these discs to eliminate the possibility that the phenotype might be due to over-expression of *Drosophila* Ubx itself. As we have seen in figure 18, wild type Ubx$_D$ expression is seen when these wing discs are stained using anti-N-terminal-Ubx$_D$ antibody. Hence we can say that the disc

phenotype observed and the *cut* repression is indeed an effect of over-expression of Ubx$_A$. Penetrance is 100% in all the four cases: *Drosophila*-Ubx (n=10, N=13), *Drosophila*-Ubx-FLAG (n= 14, N= 14), Apis-Ubx (n= 21, N= 21), *Bombyx*-Ubx (n= 6, N= 8).
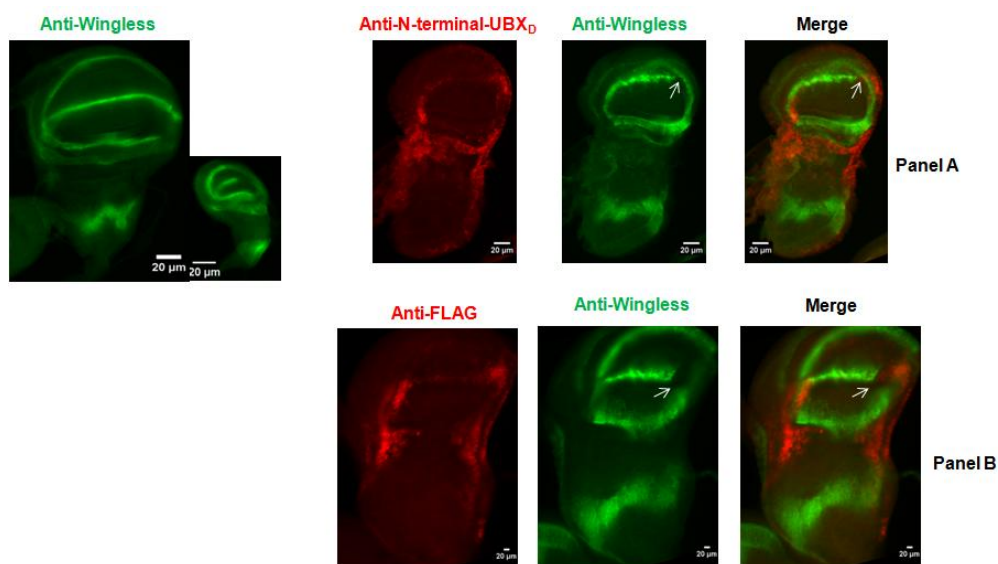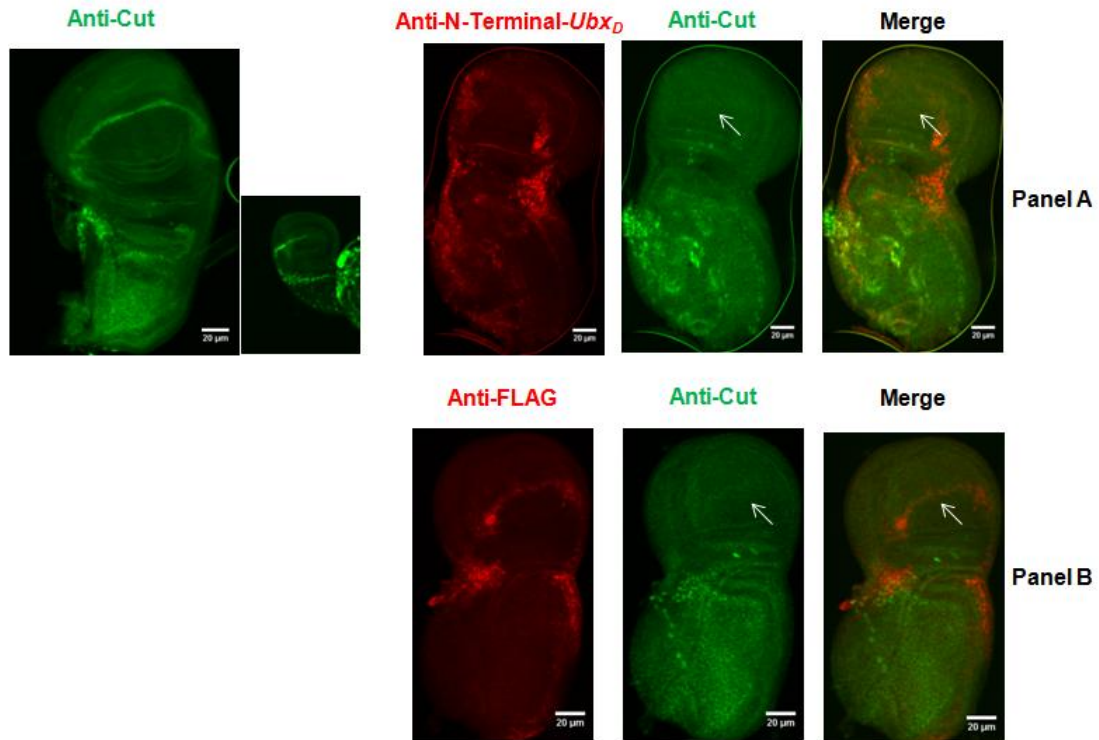

Figure 19. On the left hand side, *cut* expression in the wing disc and haltere discs of *Drosophila* in wild type background; Panel A: Ubx$_D$ over-expression using *vg*-Gal4 (red), *cut* expression (green) and the merger; Panel B: Ubx$_D$-FLAG over-expression using *vg*-Gal4 (red), *cut* expression (green) and the merger. White arrows show repression of *cut* from the entire D-V boundary.
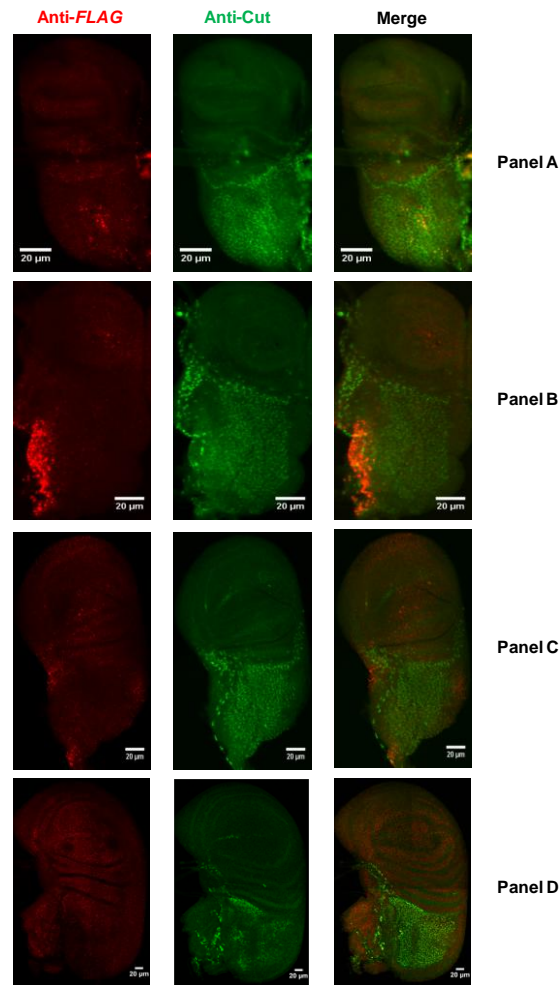
Figure 20. On the left hand side, *cut* expression in the wing disc and haltere discs of *Drosophila* in wild type background; Panel A: Ubx$_A$ over-expression using *vg*-Gal4 (red), *cut* expression (green) and the merger. White arrows show repression of *cut* from the entire D-V boundary.

Similar experiments were performed using *sd*-Gal4 for the over-expression of Ubx. Figure 21 shows over-expression of Ubx$_D$ in the wing pouch of wing imaginal discs of *Drosophila* and its effect on *cut* expression. *cut* is repressed from the entire D-V boundary of wing discs when Ubx from *Drosophila* is over-expressed in the wing pouch using *sd*-Gal4 (as shown by the white arrows in figure 21).
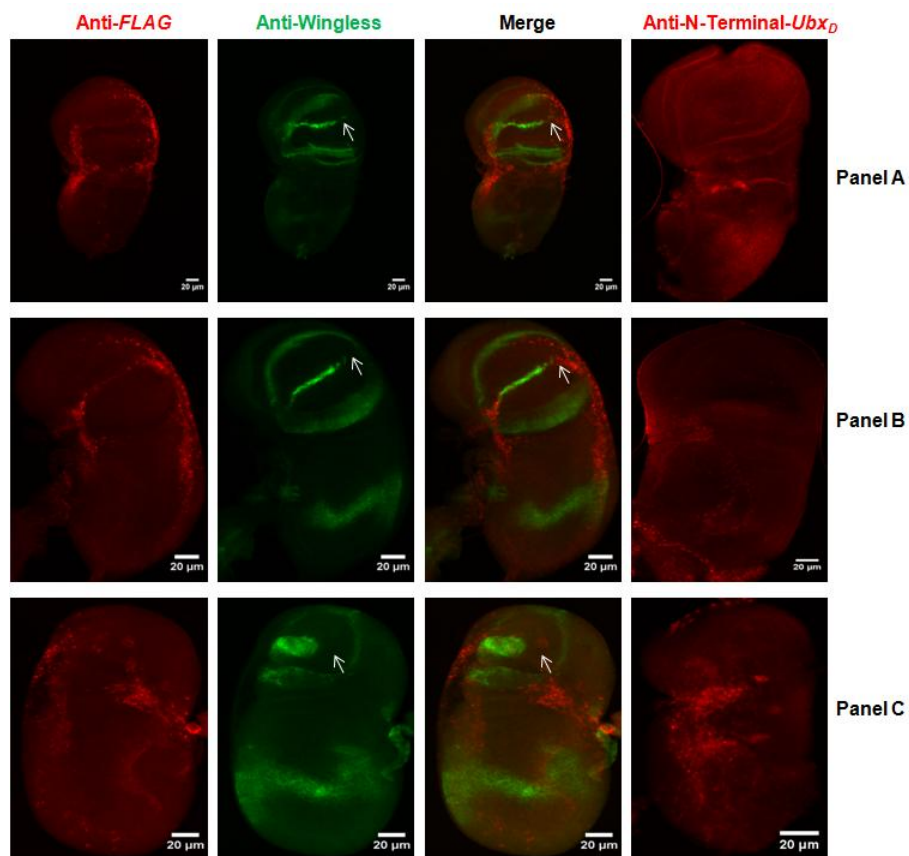

38
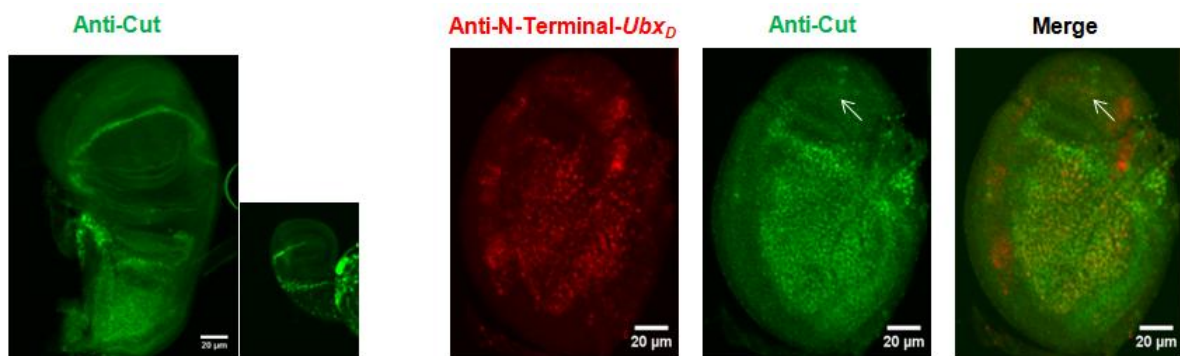
Figure 21. Panel A: Apis-Ubx-FLAG (line1) over-expression using *vg*-Gal4 (anti-FLAG, red), *cut* expression (green) and the merger; Panel B: Apis-Ubx-FLAG (line2)over-expression using *vg*-Gal4 (anti-FLAG, red), *cut* expression (green) and the merger. Panel C: Apis-Ubx-FLAG (line3) over-expression using *vg*-Gal4 (anti-FLAG, red), *cut* expression (green) and the merger; White arrows show repression of *cut* in the wing discs.

Figure 22 shows over-expression of *Apis*-Ubx-FLAG (line1, 2 & 3) in the D-V boundary of wing imaginal discs of *Drosophila* and its effect on *cut* expression. FLAG-staining has not worked very well in these cases but owing to the disc phenotype, it is expected that Ubx_A is getting over-expressed in these discs. *cut* is repressed from the entire D-V boundary of wing discs when Ubx from *Apis mellifera* is over-expressed using *sd*-Gal4 (as shown by the white arrows in figure 22). Anti-N-terminal-Ubx_D antibody was used to stain these discs to eliminate the possibility that the phenotype might be due to over-expression of *Drosophila* Ubx itself. As we have seen in figure 18, wild type Ubx_D expression is seen when these wing discs are stained using anti-N-terminal-Ubx_D antibody. Hence we can say that the disc phenotype observed and the *wingless* repression is indeed an effect of over-expression of Ubx_A. Penetrance is almost 100% in all the three cases (n= 17, N= 17).
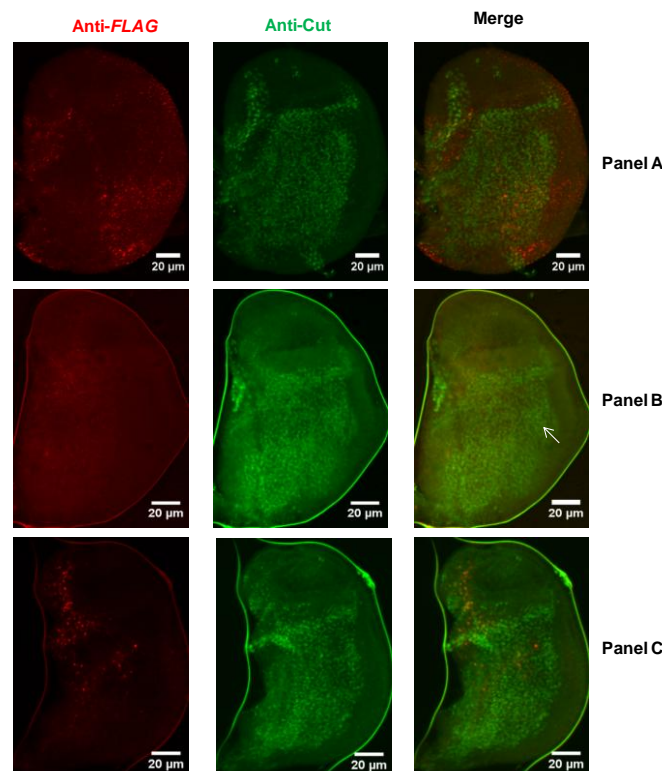
Figure 22. Panel A: *Apis*-Ubx-FLAG (line1) over-expression using *sd*-Gal4 (anti-FLAG, red), *cut* expression (green) and the merger; Panel B: Apis-Ubx-FLAG (line2) over-expression using *sd*-Gal4 (anti-FLAG, red), *cut* expression (green) and the merger; Panel C: Apis-Ubx-FLAG (line3) over-expression using *sd*-Gal4 (anti-FLAG, red), *cut* expression (green) and the merger; White arrows show repression of *cut* from the entire D-V boundary.

3.7 Repression of vestigial-quadrant enhancer from *Drosophila* by Ubx$_A$

The gene *vestigial* encodes for a nuclear protein which acts as key player in the development of wings in *Drosophila*. Loss of function mutants of vestigial show defects in wing development while its ectopic expression gives rise to wing tissue (Kim et al., 1996). Two types of enhancers have been identified for the gene *Vestigial* in *Drosophila melanogaster*, one is vestigial boundary enhancer and the other one is vestigial quadrant enhancer. Vestigial boundary enhancer is not repressed in the haltere while the quadrant enhancer is repressed in the haltere in *Drosophila* (Shashidhara et al., 1999). *Vestigial* quadrant enhancer is found on the 4th intron of *Drosophila*, it is a unique and wing specific enhancer and also a direct target of Ubx in *Drosophila*. It has been shown that mis-expression of Ubx$_D$ in the wing disc D-V boundary causes non-cell antonomous repression of *vg*-QE (Shashidhara et al., 1999).To further test the ability of Ubx from *Apis mellifera* to repress the target genes of Ubx$_D$, expression of a lacz construct for Drosophila's vestigial quadrant enhancer was checked in the over-expressed Ubx$_A$ background. For this purpose, virgin females of $\frac{vg-\text{Gal4}}{\text{CyO}}$; $\frac{\text{Quadvglacz}}{\text{TM3B}}$ line were crossed to the males homozygous for Apis-Ubx-FLAG (from line1). Here Quadvglacz stands for *Drosophila* vestigial quadrant enhancer fused to lacz reporter. To check whether *vg*-Gal4 was indeed present on the second chromosome, adult phenotypes were scored after the crosses. Reduction in the wing size was observed after Ubx$_D$ and Ubx$_D$-FLAG got over-expressed (figure 23 A & B); similar reduction was observed when Ubx$_A$ got over-expressed in the progeny (figure 23 C).
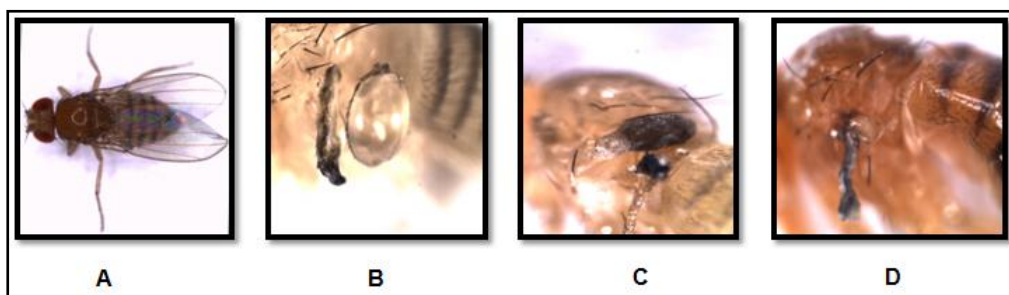
Figure 23. Ubx over-expression phenotypes scored in case of B: $\frac{vg-\text{Gal4}}{\text{CyO}}$; $\frac{\text{Quadvglacz}}{\text{TM3B}}$ **X UAS Ubx$_D$; C:** $\frac{vg-\text{Gal4}}{\text{CyO}}$; $\frac{\text{Quadvglacz}}{\text{TM3B}}$ **X UAS-Ubx$_D$-FLAG; C:** $\frac{vg-\text{Gal4}}{\text{CyO}}$; $\frac{\text{Quadvglacz}}{\text{TM3B}}$ **X UAS-Ubx$_A$-FLAG (line1); A shows a** $\frac{vg-\text{Gal4}}{\text{CyO}}$; $\frac{\text{Quadvglacz}}{\text{TM3B}}$ **female**

Figure 24 shows effect of ectopic expression of Ubx_D and Ubx D-FLAG on the expression of *Drosophila* quadrant vestigial enhancer. As we can see from the Panel A of figure 24, Ubx_D when over-expressed in the wing disc D-V boundary represses the expression of quadrant vestigial enhancer to some extent (with a frequency of 2 discs out of 70 in total); repression seems to be stronger in case of Ubx_D-FLAG over-expression (with a frequency of 3 discs out of 80 in total). Although I do not have the pictures here, most of the discs that have Ubx_D and Ubx_D-FLAG over-expression, show no lacz staining (frequency being ~20 discs out of 80 in total in both the cases).

In case of Ubx_A-FLAG (line1) over-expression in the D-V boundary, we observed that all the discs that have Ubx_A-FLAG over-expression show no lacz staining (frequency of 23 discs out of 75). Assuming that ideally half of these discs should have the quadrant vestigial enhancer expression, we can say that Ubx from *Apis* is able to completely repress the expression of the enhancer.
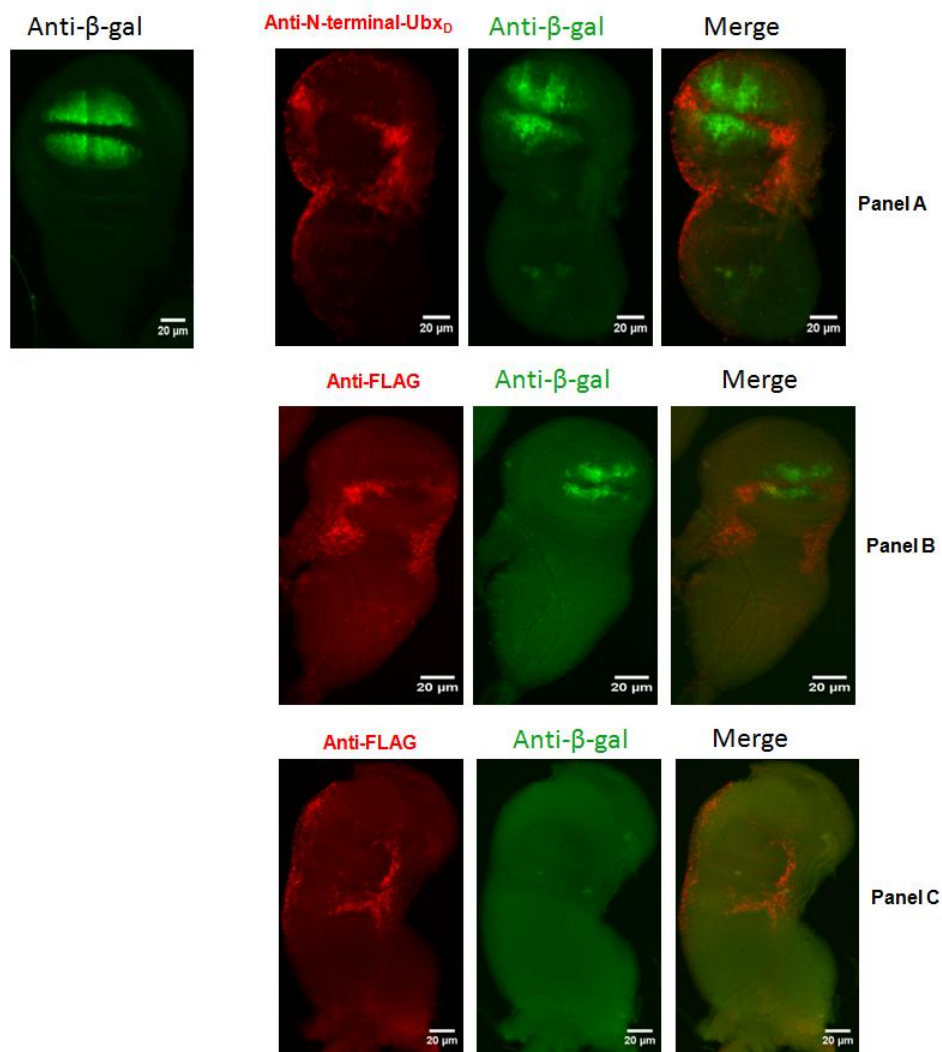
Figure 24. Panel A: Ubx_D over-expression using *vg*-Gal4 (anti-N-terminal-Ubx_D, red), *vg* quadrant enhancer expression (lacz, green) and the merger; Panel B: Ubx_D-FLAG (line2)over-

expression using *vg*-Gal4 (anti-FLAG, red), *vg* quadrant enhancer expression (lacz, green)and the merger. Panel C: Apis-Ubx-FLAG (line1)over-expression using *vg*-Gal4 (anti-FLAG, red), *vg* quadrant enhancer expression (lacz, green)and the merger

3.8 Reporter construct of *vg*-QE enhancer from *Apis*

As mentioned in the methods section, *vg*-QE from *Apis mellifera* was cloned into UAS-stinger GFP transformation vector. To check the reporter gene activity of the construct, wing, haltere, leg and eye imaginal discs from the third instar larvae were stained using rabbit anti-GFP antibody (figure 24). GFP expression seen in the hinge and notum regions of wing disc is similar *Apis* wingless expression but some expression seen in the pouch region resembles to quadrant *vestigial* gene to some extent. Haltere discs also showed enhancer expression in the pouch which resembles *vestigial* expression. Expression in the eye and leg discs also suggests *wingless* expression pattern. Since *Apis* wingless (sent by Naveen Prasad) and vestigial-QE constructs were both sent for transgenic injections, there could have been mix up between these constructs at the NCBS fly facility. To rectify this, we need to perform PCR amplification of the construct from genomic DNA of these flies and confirm the sequences again. If this construct is indeed of *Apis vestigial*-QE then expression in the haltere pouch suggests differences in the enhancers of this gene from *Apis* and *Drosophila* since *vg*-QE from the fruit fly shows no expression in the haltere. Sequence homology analysis of *vg*-QE from *Apis* and *Drosophila* is added in the appendix section, the two sequences show no significant similarity.
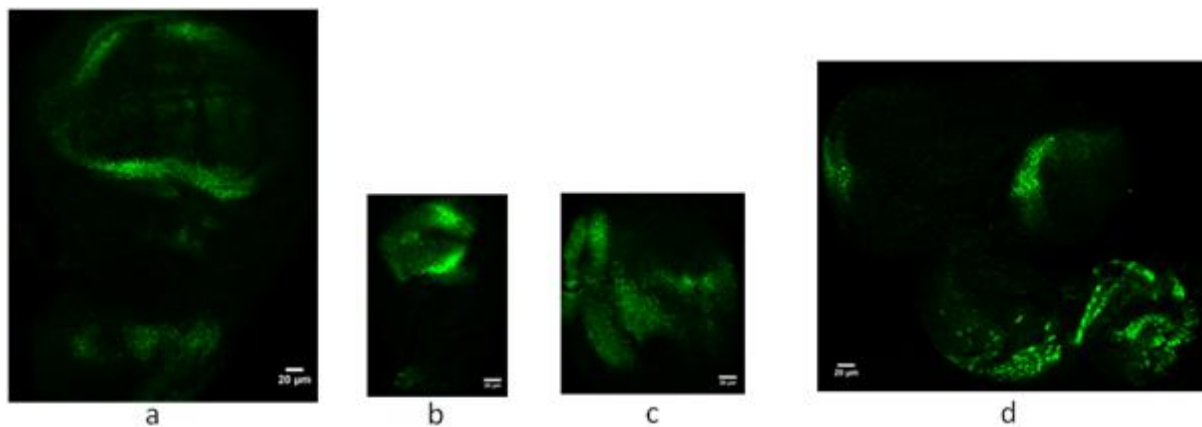
Figure 24. *vg*-QE enhancer from *Apis mellifera* cloned into UAS-Stinger GFP transformation vector. GFP staining for the wing (a), haltere (b), leg (c) and eye (d) imaginal discs shown here.

Chapter 4

Discussion

Over-expression of Ubx from *Apis* using *vg*-Gal4 and *sd*-Gal4 showed suppression of wing development in *Drosophila*. The phenotypes observed were stronger in case of *sd*-Gal4 which can be explained in a way that *sd*-Gal4 is expressed from earlier stages than *vg*-Gal4; although both the Gal4s show expression before the segmentation is complete. Reduction in the size of the wing pouch and without effecting the notum region suggests wing to haltere transformation.The fact that *Apis* Ubx is able to repress *wingless* only from the posterior compartment of the wing disc suggests that these results are not mere effects of over-expression of a heterologous protein in the wing disc of *Drosophila*. It is indeed the ability of Apis Ubx to induce homeotic transformation.

Nevertheless, there is a possibility that probably Ubx from different insects do not bind to the promoters of these genes in other insects but genome wide CHIP assays of Ubx performed earlier in the lab (by Naveen Prasad and T Harsha) suggest that Ubx does indeed bind to promoters in at least *Apis* and *Bombyx*. So interestingly, it seems that Ubx protein itself has not evolved amongst the diverse insect groups. An interesting observation has come up from the comparative in-silico analysis of targets of Ubx from the insects *Apis mellifera* and *Drosophila melanogaster* (carried out by Abhijeet Awadhiya) and it suggests that the targets that are common to Ubx from both the insects in case of hindwing development are evolving at a faster rate than the targets that are special to the insect orders. This suggests that the diversity in the T3 morphology must be due to changes downstream to Ubx.

In order to understand the functional similarity between Ubx orthologs, a comparison of Onychoporan Ubx and *Drosophila* Ubx was carried out in an earlier study. Onychophora is a sister phylum of arthropoda and the protein Ubx has been diverging in sequence for about more that 520 million years. Ectopic expression of onychophoran Ubx carried out almost the same functions in *Drosophila melanogaster* as its own Ubx. However, some of the embryonic functions of Ubx in *Drosophila* couldn't be mimicked by the Onychophora Ubx (Grenier and Carroll, 2000). However, most of the embryonic functions of Ubx in *Drosophila* were mimicked by other insect Ubx like *Tribolium* and *Junonia*.To understand the

43

evolution of Ubx protein in the insect lineage when compared to the oncyphoran Ubx, a study was carried out which compared the Ubx proteins of various insects and an onychophoran, Akanthokara kaputensis. Ubx proteins in the insect lineage have acquired a QA domain towards their C terminal end over the course of evolution. This domain present in all the insects plays an important role in carrying out some of the repressive functions of Ubx in the embryonic stage (Galant et al., 2002). As seen from the sequence homology analysis of the Ubx proteins from *Apis*, *Bombyx* and *Tribolium* (appendix section 5), this domain is conserved in these Ubx proteins too which can account for proteins' repressive functions. Over-expression of Ubx derived from butterfly was also found sufficient to induce wing-to-haltere transformations in *Drosophila* (Grenier and Carroll, 2000) and these studies support the results obtained here.

Chapter 5

Conclusions and future work

The Hox protein, Ultrabitorax brings about the development of haltere by repressing some of the most important wing patterning genes that come under its direct regulation in *Drosophila melanogaster*. Our results show that Ubx protein from *Apis mellifera* is also able to suppress wing development in *Drosophila melanogaster* when over-expressed in the T2 thoracic segment of *Drosophila*. It also possesses the ability to repress some of the most important wing patterning genes like *wingless* and *cut* and *vestigial* in the same way Ubx from *Drosophila melanogaster* does which means Ubx from Apis is able to induce wing to haltere transformation in *Drosophila*. Thus we can conclude that the ability of Ubx to differentially regulate the target genes in the wing and the haltere of *Drosophila* and its inability to do so in the forewing and hind wing of *Apis mellifera* is not attributed to the structural differences in the Ubx proteins from these insects. Thus, the target selection of Ubx during through recruitment of specific cofactors during the haltere development in *Drosophila* is more likely due to differences in the cis-regulatory codes of these genes where the Ubx protein.

Future work involves testing Ubx from *Bombyx mori* and *Triboluim custaneum* to induce wing to haltere transformations in *Drosophila* and their ability to repress *wingless*, *cut* and *vestigial*. Also to check how well Ubx from Apis is able to replace endogenous Ubx of *Drosophila*, a rescue experiment is being planned. We would over-express *Apis* Ubx in the T3 segment of *Drosophila* (that carries three known mutations on the BX-C which completely transform the T3 segment from haltere to wing) and check its ability to induce wing to haltere transformation in *Drosophila*.

Future aims of this project involves comparing the target and cofactor specificity of Ubx from all the four insects which will give a deeper understanding of the differential regulation phenomenon and also shed light on protein-protein interactions conserved over 250 million years of divergence in these insects. Enhancer elements of the differentially regulated genes from *Apis*, *Bombyx* and *Tribolium* will be cloned into a reporter construct and regulation of their expression will be checked in the over-expressed *Drosophila* Ubx background. Expression and regulation of orthologous enhancers from *Drosophila* will be checked in over-

expressed heterologous Ubx background. This approach might answer questions over what differences in the sequences of orthologous cis-regulatory elements are responsible for causing the differential regulation by Ubx.

Chapter 6

References

1. Agrawal, P., Habib, F., Yelagandula, R., and Shashidhara, L.S. (2011). Genome-level identification of targets of Hox protein Ultrabithorax in Drosophila: novel mechanisms for target selection. Sci Rep *1*.

2. Ashburner, M. (1989). Drosophila. A laboratory handbook., Second edn (Cold Spring Harbor Laboratory).

3. Ayeni, J.O., Varadarajan, R., Mukherjee, O., Stuart, D.T., Sprenger, F., Srayko, M., and Campbell, S.D. (2014). Dual phosphorylation of cdk1 coordinates cell proliferation with key developmental processes in Drosophila. Genetics *196*, 197-210.

4. Barolo, S., Castro, B., and Posakony, J.W. (2004). New Drosophila transgenic reporters: insulated P-element vectors expressing fast-maturing RFP. BioTechniques *36*, 436-442.

5. Basset, Y., Cizek, L., Cuénoud, P., Didham, R.K., Guilhaumon, F., Missa, O., Novotny, V., Ødegaard, F., Roslin, T., Schmidl, J.*, et al.* (2012). Arthropod Diversity in a Tropical Forest. Science *338*, 1481-1484.

6. Blair, S.S. (2007). Wing Vein Patterning in Drosophila and the Analysis of Intercellular Signaling. Annual Review of Cell and Developmental Biology *23*, 293-319.

7. Blochlinger, K., Jan, L.Y., and Jan, Y.N. (1993). Postembryonic patterns of expression of cut, a locus regulating sensory organ identity in Drosophila. Development *117*, 441-450.

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

9. Brook, W.J., and Cohen, S.M. (1996). Antagonistic Interactions Between Wingless and Decapentaplegic Responsible for Dorsal-Ventral Pattern in the Drosophila Leg. Science *273*, 1373-1377.

10. Cabrera, C.V., Botas, J., and Garcia-Bellido, A. (1985). Distribution of Ultrabithorax proteins in mutants of Drosophila bithorax complex and its transregulatory genes. Nature *318*, 569-571.

11. Carroll, S.B., Weatherbee, S.D., and Langeland, J.A. (1995). Homeotic genes and the regulation and evolution of insect wing number. Nature *375*, 58-61.

12. Casares, F., Calleja, M., and Herrer, E.S. (1996). Functional similarity in appendage specification by the Ultrabithorax and abdominal-A Drosophila HOX genes. The EMBO Journal *15*, 3934–3942.

13. Casares, F., and Mann, R.S. (1998). Control of antennal versus leg development in Drosophila. Nature *392*, 723-726.

14. Dong, P.D., Chu, J., and Panganiban, G. (2000). Coexpression of the homeobox genes Distal-less and homothorax determines Drosophila antennal identity. Development *127*, 209-216.

15. Duboule, D. (1998). Vertebrate Hox gene regulation: clustering and/or colinearity? Current Opinion in Genetics & Development *8*, 514-518.

16. Galant, R., Walsh, C.M., and Carroll, S.B. (2002). Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. Development *129*, 3115-3126.

17. Garcia-Bellido, A. (1977). Homoeotic and atavic mutations in insects. American Zoologist *17*, 613-629.

18. Greenspan, R.J. (2004). Fly pushing: the theory and practice of Drosophila genetics (Cold Spring Harbor Laboratory Press).

19. Grenier, J.K., and Carroll, S.B. (2000). Functional evolution of the Ultrabithorax protein. Proceedings of the National Academy of Sciences *97*, 704-709.

20. Kim, J., Sebring, A., Esch, J.J., Kraus, M.E., Vorwerk, K., Magee, J., and Carroll, S.B. (1996). Integration of positional signals and regulation of wing formation and identity by Drosophila vestigial gene. Nature *382*, 133-138.

21. L. Wolpert, C.T. (2011). Principles of Development, Fourth edn (Oxford).

Lee, J.J., von Kessler, D.P., Parks, S., and Beachy, P.A. (1992). Secretion and localized transcription suggest a role in positional signaling for products of the segmentation gene hedgehog. Cell *71*, 33-50.

22. Lewis, E.B. (1978). A gene complex controlling segmentation in Drosophila. Nature *276*, 565-570.

23. Mann, R.S., Lelli, K.M., and Joshi, R. (2009). Chapter 3 Hox Specificity: Unique Roles for Cofactors and Collaborators. In Current Topics in Developmental Biology, P. Olivier, ed. (Academic Press), pp. 63-101.

24. McGinnis, W., Garber, R.L., Wirz, J., Kuroiwa, A., and Gehring, W.J. (1984). A homologous protein-coding sequence in drosophila homeotic genes and its conservation in other metazoans. Cell *37*, 403-408.

25. McGinnis W, K.M. (1994). The molecular architects of body design. Scientific American *270*, 58-61.

26. Nellen, D., Affolter, M., and Basler, K. (1994). Receptor serine/threonine kinases implicated in the control of Drosophila body pattern by decapentaplegic. Cell *78*, 225-237.

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

28. Prasad, M., Bajpai, R., and Shashidhara, L.S. (2003). Regulation of Wingless and Vestigial expression in wing and haltere discs of Drosophila. Development *130*, 1537-1547.

29. Rebay, I., Fleming, R.J., Fehon, R.G., Cherbas, L., Cherbas, P., and Artavanis-Tsakonas, S. (1991). Specific EGF repeats of Notch mediate interactions with Delta and serrate: Implications for notch as a multifunctional receptor. Cell *67*, 687-699.

30. Regier, J.C., Shultz, J.W., Zwick, A., Hussey, A., Ball, B., Wetzer, R., Martin, J.W., and Cunningham, C.W. (2010). Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. Nature *463*, 1079-1083.

31. S. B. Carroll, J.K.G., S. D. Weatherbee (2004). From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design, Second edn (Wiley).

32. Sanchez-Herrero, E., Vernos, I., Marco, R., and Morata, G. (1985). Genetic organization of Drosophila bithorax complex. Nature *313*, 108-113.

33. Shashidhara, L.S., Agrawal, N., Bajpai, R., Bharathi, V., and Sinha, P. (1999). Negative Regulation of Dorsoventral Signaling by the Homeotic Gene Ultrabithorax during Haltere Development in Drosophila. Developmental Biology *212*, 491-502.

34. Simmonds, A.J., Brook, W.J., Cohen, S.M., and Bell, J.B. (1995). Distinguishable functions for engrailed and invected in anterior-posterior patterning in the Drosophila wing. Nature *376*, 424-427.

35. Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I., Zhou, T., Rohs, R., Honig, B., Bussemaker, Harmen J.*, et al.* (2011). Cofactor Binding Evokes Latent Differences in DNA Binding Specificity between Hox Proteins. Cell *147*, 1270-1282.

36. Struhl, G. (1981). A gene product required for correct initiation of segmental determination in Drosophila. Nature *293*, 36-41.

37. Struhl, G. (1982). Genes controlling segmental specification in the Drosophila thorax. Proceedings of the National Academy of Sciences of the United States of America *79*, 7380-7384.

38. Struhl, G., and White, R.A.H. (1985). Regulation of the Ultrabithorax gene of drosophila by other bithorax complex genes. Cell *43*, 507-519.

39. Tomoyasu, Y., Wheeler, S.R., and Denell, R.E. (2005). Ultrabithorax is required for membranous wing identity in the beetle Tribolium castaneum. Nature *433*, 643-647.

40. Weatherbee, S.D., Halder, G., Kim, J., Hudson, A., and Carroll, S. (1998). Ultrabithorax regulates genes at several levels of the wing-patterning hierarchy to shape the development of the Drosophila haltere. Genes & Development *12*, 1474-1482.

41. White, R.A.H., and Akam, M.E. (1985). Contrabithorax mutations cause inappropriate expression of Ultrabithorax products in Drosophila. Nature *318*, 567-569.

42. Xu, T., and Rubin, G.M. (1993). Analysis of genetic mosaics in developing and adult Drosophila tissues. Development *117*, 1223-1237.

Appendix: Sequencing results and their blast analysis

1. Full length *Apis mellifera Ubx* cDNA cloned into pUAST-flag vector:

>pUAST-flag forward primer:

```
AAAAAATATGACCGTCGCTAGCGAGCTAAGCAATAAACAAGCGCAGCTGAACAAGCTAAACAATCTGCAGT
AAAGTGCAAGTTAAAGTGAATCAATTAAAAGTAACCAGCAACCAAGTAAATCAACTGCAACTACTGAAATC
TGCCAAGAAGTAATTATTGAATACAAGAAGAGAACTCTGAATAGGGAATTGGGAATTGACGCAAATGGGCG
GTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTGATCTACCATG
GACTACAAAGACGATGACGACAAGCTTGCGGCCGCGAATTCATCGATAGATCTCGAGGGTACCATGAACTC
GTATTTTGAGCAGACTGCGGGTGGCTTCTACGGAAGCCACCACCATCAGACAGGAGCCGCCAGTCAGCATC
ATGATCCAGCCACGGCAGCCGCCTATCGAAGTTTCCCCCTCGGCCTCGGTATGTCACCGTACGCGTCCACC
CAACACCATCATCACACCTCCTCGTCGTTGGGCATACACCCGGGCGGTGGGACGAACACGAGGCCGCCCCA
GGATTCGCCGTACGATGCGAGCGTCGCGACGGCTTGCAAGCTTTATTCGACGACGCCCGAGGCAACTGGCC
ACACGACATCCTCGTATTCGACCACAGCGGCCAAGGACTGTAAGCAACAGGATCAAGCATCGGCGCATCAG
AACGGTTACGCCGCAGTGATGGCAGCTGCCGCCGTCAAGGACGTGTGGCAATCGGCTACCTCGGGGGCGAA
CAGCCAGAGCAATTCGGTGGTTCGCCCATCGGCGTGCACCCCGGAAGGGACGAGGGTTGGTAGCTACGGTG
GTCTCGTAGGCGGCGATCCGGCATCGAGTCCCGGCAACAACAGTTCCTCGAGGTCCCTCACGTCGTCCTGG
AACACCTGCAGTTTGAACTCGTCCGCGAGCCAACCGGTTGCCACGCAACTACATCAGCAACCCAGCAACCA
TACGTTCTACCCCTGGATGGCTATAGCAGGAGCGAACGGAATGCGCAGGCGCGGCCGCCAGACCTATACGC
GCTACCAGACGCTCGAACTGGAAAAAGGAATTCCACCCGAACCACTACCTCACTAGGGCGGAGGCGGATCG
AGATGGCACCACTCCTCTTGCCTGACGGAACGGCAGATCAAAGTCTGGGTTCCAAAAATCGGCGGATGAAA
CTTGAAAAAGGGAATTCACGGGGTATCCAAGGAGCTTGAACCAAACCGGAGAAAACCAGGGGCCAAGGCC
CN
```

>Reverse complement of pUAST-flag reverse primer:

```
GGCTTCCAAAGGCGGATGCCGGCAAAGCTTGCGGCCGGGGGATTTCATCGATAGATTTCGGAGGGTACCATGAAC
TCGTATTTTGAGCAGACTTCGGGGTGGCTTTTACGGAAGCCACCACCCTTCAGACAGGAGCCGCCAGTCAGCATC
ATGATCCAGCCACGGCAGCCGCCTATCGAAGTTTCCCCCTCGGCCTCGGTATGTCACCGTACGCGTCCACCCAAC
ACCATCATCACACCTCCTCGTCGTTGGGCATACACCCGGGCGGTGGGACGAACACGAGGCCGCCCCAGGATTCGC
CGTACGATGCGAGCGTCGCGACGGCTTGCAAGCTTTATTCGACGACGCCCGAGGCAACTGGCCACACGACATCCT
CGTATTCGACCACAGCGGCCAAGGACTGTAAGCAACAGGATCAAGCATCGGCGCATCAGAACGGTTACGCCGCAG
TGATGGCAGCTGCCGCCGTCAAGGACGTGTGGCAATCGGCTACCTCGGGGGCGAACAGCCAGAGCAATTCGGTGG
TTCGCCCATCGGCGTGCACCCCGGAAGGGACGAGGGTTGGTAGCTACGGTGGTCTCGTAGGCGGCGATCCGGCAT
CGAGTCCCGGCAACAACAGTTCCTCGAGGTCCCTCACGTCGTCCTGGAACACCTGCAGTTTGAACTCGTCCGCGA
GCCAACCGGTTGCCACGCAACTACATCAGCAACCCAGCAACCATACGTTCTACCCCTGGATGGCTATAGCAGGAG
CGAACGGAATGCGCAGGCGCGGCCGCCAGACCTATACGCGCTACCAGACGCTCGAACTGGAGAAGGAATTCCACA
CGAACCACTACCTCACTAGGCGGAGGCGGATCGAGATGGCACACTCGCTCTGCCTGACGGAACGGCAGATCAAGA
TCTGGTTCCAGAATCGGCGGATGAAGCTGAAGAAGGAGATACAGGCGATCAAGGAGCTGAACGAACAGGAGAAGC
AGGCGCAGGCGCAGAAGGCAGCGGCAGCAGCGGCCGCGGCTGCGCATCAGCAGCAAGCGGCCGGTGGGACCGGAC
GGGCCCAATTTN
```

>Reverse Complement of Ubx gene reverse primer:

```
TTCTGAATTGGGGAATTGGGAAATTGAAGCAAAATGGGCCGGTAAGCGTTGTTCCGGTCGGGAGGTCTTATTTAA
GCCAGAGCTCGGTTTAGTGAACCCTTCAGAATTGATCTTCCCATGGACTCCAAAGACGATGACGCCAAGCTTGCG
GCCGCGAATTCATCGATAGATCTCGAGGGTACCATGAACTCGTATTTTGAGCAGACTGCGGGTGGCTTCTACGGA
AGCCACCACCCTTCAGACAGGAGCCGCCAGTCAGCATCATGATCCAGCCACGGCAGCCGCCTATCGAAGTTTCCC
CCTCGGCCTCGGTATGTCACCGTACGCGTCCACCCAACACCATCATCACACCTCCTCGTCGTTGGGCATACACCC
GGGCGGTGGGACGAACACGAGGCCGCCCCAGGATTCGCCGTACGATGCGAGCGTCGCGACGGCTTGCAAGCTTTA
TTCGACGACGCCCGAGGCAACTGGCCACACGACATCCTCGTATTCGACCACAGCGGCCAAGGACTGTAAGCAACA
GGATCAAGCATCGGCGCATCAGAACGGTTACGCCGCAGTGATGGCAGCTGCCGCCGTCAAGGACGTGTGGCAATC
```

```
GGCTACCTCGGGGGCGAACAGCCAGAGCAATTCGGTGGTTCGCCCATCGGCGTGCACCCCGGAAGGGACGAGGGT
TGGTAGCTACGGTGGTCTCGTAGGCGGCGATCCGGCATCGAGTCCCGGCAACAACAGTTCCTCGAGGTCCCTCAC
GTCGTCCTGGAACACCTGCAGTTTGAACTCGTCCGCGAGCCAACCGGTTGCCACGCAACTACATCAGCAACCCAG
CAACCATACGTTCTACCCCTGGATGGCTATAGCAGGAGCGAACGGAATGCGCAGGCGCGGCCGCCAGACCTATAC
GCGCTACCAGACGCTCGAACTGGAGAAGGAATTCCACACGAACCACTACCTCACTAGGCGGAGGCGGATCGAGAT
GGCACACTCGCTCTGCCTGACGGAACGGCAGATCAAGATCTGGTTCCAGAATCGGCGGATGAAGCTGAAGAAGGA
GATACAGGCGATCAAGGAGCTGAACGAACAGGAGAAGCAGGCGCAGGCGCAGAAGGCAGCGGCAGCAGCGGCCGC
GGCTGCGCATCAGCTACTAAGAGTGCTGATCNTT
```

Blast analysis using pUAST-flag forward primer:

Apis mellifera ultrabithorax (Ubx), mRNA
Sequence ID: ref|NM_001168700.1|Length: 993Number of Matches: 1

Related Information
Gene-associated gene details
UniGene-clustered expressed sequence tags
Map Viewer-aligned genomic context
Range 1: 1 to 902GenBankGraphics Next Match Previous Match

Alignment statistics for match #1

Score	Expect	Identities	Gaps	Strand
1528 bits(827)	0.0	888/915(97%)	14/915(1%)	Plus/Plus

```
Query  348   ATGAACTCGTATTTTGAGCAGACTGCGGGTGGCTTCTACGGAAGCCACCACCATCAGACA
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     ATGAACTCGTATTTTGAGCAGACTGCGGGTGGCTTCTACGGAAGCCACCACCATCAGACA

Query  408   GGAGCCGCCAGTCAGCATCATGATCCAGCCACGGCAGCCGCCTATCGAAGTTTCCCCCTC
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    GGAGCCGCCAGTCAGCATCATGATCCAGCCACGGCAGCCGCCTATCGAAGTTTCCCCCTC

Query  468   GGCCTCGGTATGTCACCGTACGCGTCCACCCAACACCATCATCACACCTCCTCGTCGTTG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121   GGCCTCGGTATGTCACCGTACGCGTCCACCCAACACCATCATCACACCTCCTCGTCGTTG

Query  528   GGCATACACCCGGGCGGTGGGACGAACACGAGGCCGCCCCAGGATTCGCCGTACGATGCG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181   GGCATACACCCGGGCGGTGGGACGAACACGAGGCCGCCCCAGGATTCGCCGTACGATGCG

Query  588   AGCGTCGCGACGGCTTGCAAGCTTTATTCGACGACGCCCGAGGCAACTGGCCACACGACA
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241   AGCGTCGCGACGGCTTGCAAGCTTTATTCGACGACGCCCGAGGCAACTGGCCACACGACA

Query  648   TCCTCGTATTCGACCACAGCGGCCAAGGACTGTAAGCAACAGGATCAAGCATCGGCGCAT
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301   TCCTCGTATTCGACCACAGCGGCCAAGGACTGTAAGCAACAGGATCAAGCATCGGCGCAT

Query  708   CAGAACGGTTACGCCGCAGTGATGGCAGCTGCCGCCGTCAAGGACGTGTGGCAATCGGCT
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361   CAGAACGGTTACGCCGCAGTGATGGCAGCTGCCGCCGTCAAGGACGTGTGGCAATCGGCT

Query  768   ACCTCGGGGGCGAACAGCCAGAGCAATTCGGTGGTTCGCCCATCGGCGTGCACCCCGGAA
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421   ACCTCGGGGGCGAACAGCCAGAGCAATTCGGTGGTTCGCCCATCGGCGTGCACCCCGGAA

Query  828   GGGACGAGGGTTGGTAGCTACGGTGGTCTCGTAGGCGGCGATCCGGCATCGAGTCCCGGC
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

52

```
Sbjct   481    GGGACGAGGGTTGGTAGCTACGGTGGTCTCGTAGGCGGCGATCCGGCATCGAGTCCCGGC

Query   888    AACAACAGTTCCTCGAGGTCCCTCACGTCGTCCTGGAACACCTGCAGTTTGAACTCGTCC
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   541    AACAACAGTTCCTCGAGGTCCCTCACGTCGTCCTGGAACACCTGCAGTTTGAACTCGTCC

Query   948    GCGAGCCAACCGGTTGCCACGCAACTACATCAGCAACCCAGCAACCATACGTTCTACCCC
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   601    GCGAGCCAACCGGTTGCCACGCAACTACATCAGCAACCCAGCAACCATACGTTCTACCCC

Query   1008   TGGATGGCTATAGCAGGAGCGAACGGAATGCGCAGGCGCGGCCGCCAGACCTATACGCGC
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   661    TGGATGGCTATAGCAGGAGCGAACGGAATGCGCAGGCGCGGCCGCCAGACCTATACGCGC

Query   1068   TACCAGACGCTCGAACTGGAAAAAGGAATTCCACCCGAACCACTACCTCACTAGGGCGGA
               ||||||||||||||||||||||   |||||||||||| |||||||||||||||||| ||||||
Sbjct   721    TACCAGACGCTCGAACTGGAG-AAGGAATTCCACACGAACCACTACCTCACTA-GGCGGA

Query   1128   GGCGGATCGAGATGGCACCACTC-CTCTTGCCTGACGGAACGGCAGATCAAAGTCTGGGT
               |||||||||||||||||| |||| ||| |||||||||||||||||||||||  |||||| |
Sbjct   779    GGCGGATCGAGATGGCAC-ACTCGCTC-TGCCTGACGGAACGGCAGATCAAGATCTGG-T

Query   1187   TCCAAAAATCGGCGGATGAAACTTGAAAAAGGGAATTCACGGGGTATCCAAGGAGCTTG
               ||||  |||||||||||||| ||| || || | || | || || | ||| |||||||| |
Sbjct   836    TCCAG-AATCGGCGGATGAAGCT-GAAGAAGGAGA-TACA-GGCG-ATC-AAGGAGCT-G

Query   1247   AACCAAACCGGAGAA    1261
               ||| || | ||||||
Sbjct   889    AACGAA-CAGGAGAA    902
```

2. Full length *Bombyx mori Ubx* cDNA cloned into pUAST-flag:

>pUAST-flag forward primer:

```
AGGGGGGAAACGTCGCTAGCGATGCTAAGCAAATAAACAAGCGCAGCTGAACAAGCTAAACAATCTGCAGTAAAG
TGCAAGTTAAAGTGAATCAATTAAAAGTAACCAGCAACCAAGTAAATCAACTGCAACTACTGAAATCTGCCAAGA
AGTAATTATTGAATACAAGAAGAGAACTCTGAATAGGGAATTGGGAATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTGATCTACCATGGACTACAAAGACGATG
ACGACAAGCTTGCGGCCGCGAATTCATCGATAGATCTCGAGGGTACCATGAACTCTTACTTCGAGCAGGGTGGTT
TTTACGGGGCCCATGGAGTGCACCAGGGCGGCGGTGGTGGAGACCAGTACCGCGGCTTCCCTCTGGGCCTCACGT
ATGCACAGCCACACGCTTTGCACCAGCCTCGTCCTCAGGATTCACCGTACGACGCGTCTGTCGCGGCGGCCTGCA
AGCTCTATGCTGGAGAGCAGCAATATCCTAAAGCAGATTGTTCAAAGCCAGGCGGTGAGCAGCAGAATGGCTATG
GTGGGAAAGAAGCCTGGGGCTCAGGTCTGGGAGCACTAGTGAGGCCGGCAGCATGCACTCCTGAAGCTCGATACA
GTGAGTCGTCAAGTCCTGGTAGAGCGCTTCCGTGGGGCAACCAGTGTGCACTTCCGGGATCAGCAGCATCAGCCG
CGCAGCCAGTGCACCAGCAGCCTACTAACCACACTTTCTACCCTTGGATGGCCATAGCAGGAGCGAACGGCCTCA
GGAGACGAGGAAGACAAACCTACACTAGATATCAAACGCTAGAATTAGAGAAAGAGTTCCACACGAACCACTACC
TTACGCGAAGGAGACGCATAGAGATGGCGCACGCGTTGTGCCTCACGGAGAGGCAAATCAAAATATGGTTCCAGA
ACCGAAGGATGAAGTTAAAGAAAGAGATCCAGGCTATAAAGGAGTTGAACGAGCAGGAAAAACAGGCGCAGGCGC
AAAAGGCGGCAACGGCTGCTGCGGCGGCCGCGGCTGCTGCCCAGGGACACCCCGAACATTAATCTAAAGGATCTT
TGTGAAGGAACCTTACTTCTGTGGTGGGACAAAATTGGACAAACTACCTACAGAGATTTAAAGCTCTAAGGGAAA
TATAAAATTTTTAAGGGGAAAATGGGTAAAACAACGGATTCAAATTGGTTGGGGATTTTAAAATCCCCACCTTNG
GAACTGATGAAATGGGAACATTGGTGGAAATGCCTTTAATAGGAGAAAACCTGTTTTTGCCCAAAAAAATGGCCT
TCTAGGGGTTAGAAGGGCTCGGGGGGACTTTCAATTTTTCTCTCTCCCAAAAAAAAAAAAAAGGGAAAAAACCCC
CAGAGTTTCTTCCCAAAATGGGGAATTTTTTGGACCCCGGTGTGTTAAAAAAAAACACTCGCGGTTTTTGGTTTA
TCCCACCAAAGAGAAAATTC
```

>Reverse complement of pUAST-flag reverse primer

```
GTTTTATCCACCCAACAACGCGCCGCGCCCACCCAAAAGGAAAAAAGACGCCCCCCCTTTAAAAATATAAAAAAT
TGGGGGTGTTCTGCACAAAACAACTCCCCTTATTTGTTCGTGTCCCAAATTCGGGGGCAGAGAGAAAATATTT
TGGGGCTTGGTTTTAGTTTTTTTGGGAGAACCCGGGTTAGTGTTGGGTGTTACACCTTCGCACCCTCTCCGGGGG
TGTTAAAAAAACCCCAATAAATTAATTAGGGCGAAAATATTTTTTAAAGTTTTTTTTTTTATAATTTGCGGATGCG
CAAAAGCTGGTGCGCCTCCAACCCGTTGGGAAGTTTTCCGGGATCCAAAGGTTCCAACCGTCGCAAGTCGGGAGT
ACTGTCCTTCGGAGGAGGATTCATTTCTTCCCGAAGGGGATACTTGTCTTCCGAACGGGAGTACTGTCCTTCCGA
GCGGGAATACTTTCCTTCCGAGCGGGAGATCTTAGCGAGCGCCGGAAGTATTATTAGAGGCGCTTTGGTTTAGGG
AGGGACAATTCAATTCAAACCAGCCAAAGTGAACCAGGTCGTTAAGCGAAAGGTTAAGCAAATTAACCAAGCGCA
GCTGAACAAGTTAAACAATTTGCAGTAAAGTGCAAGTTAAAGTGAATCAATTAAAAGTAACCAGCAACCAAGTAA
ATCAACTGCAACTACTGAAATCTGCCAAGAAGTAATTATTGAATACAAGAAGAGAACTCTGAATAGGGAATTGGG
AATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGA
ATTGATCTACCATGGACTACAAAGACGATGACGACAAGCTTGCGGCCGCGAATTCATCGATAGATCTCGAGGGTA
CCATGAACTCTTACTTCGAGCAGGGTGGTTTTTACGGGGCCCATGGAGTGCACCAGGGCGGCGGTGGTGGAGACC
AGTACCGCGGCTTCCCTCTGGGCCTCACGTATGCACAGCCACACGCTTTGCACCAGCCTCGTCCTCAGGATTCAC
CGTACGACGCGTCTGTCGCGGCGGCCTGCAAGCTCTATGCTGGAGAGCAGCAATATCCTAAAGCAGATTGTTCAA
AGCCAGGCGGTGAGCAGCAGAATGGCTATGGTGGGAAAGAAGCCTGGGGCTCAGGTCTGGGAGCACTAGTGAGGC
CGGCAGCATGCACTCCTGAAGCTCGATACAGTGAGTCGTCAAGTCCTGGTAGAGCGCTTCCGTGGGGCAACCAGT
GTGCACTTCCGGGATCAGCAGCATCAGCCGCGCAGCCAGTGCACCAGCAGCCTACTAACCACACTTTCTACCCTT
GGATGGCCATAGCAGGAGCGAACGGCCTCAGGAGACGAGGAAGACAAACCTACACTAGATATCAAACGCTAGAAT
TAGAGAAAGAGTTCCACACGAACCACTACCTTACGCGAAGGAGACGCATAGAGATGGCGCACGCGTTGTGCCTCA
CGGAGAGGCAAATCAAAATATGGTTCCAGAACCGAAGGATGAAGTTAAAGAAAGAGATCCAGGCTATAAAGGAGT
TGAACGAGCAGGAGAAACAGGCGCAGGCGCAGAAGGCGGCAGCGGCTGCTGCGGCGGCCGCGGCTGCTGCCCAGG
ACTCCTGAAAC
```

>Reverse complement of Ubx gene reverse primer:

```
NATAATTTACCAAAACAAGGACGCGCGCCGCGGNNCTTTGAGGGAGCCGCCCCCCCGATAGAAAACAATAATGTG
GGGTTCCCGCGGCAAGAGCTTTCCCTTTTCTTGTCTCCCTAATAGGGGGGGAAGAAAGATGTTTTTTGGGGCTT
TTTGAGTTTTTGGGAGACACCCGAAATAAGTTTTTGTAAACCCCTCCCACCTGCCCGGGGGGTAGGATAAACTCA
AAAATATTATTGGCGGAAAATTTTTTAAGGTTTTTTTTTTAAAATTTGGGATTACCAAAAGTTGGGCGTATCCA
ACCGGTTGGGAACTTTCCGGATCCAAAGTTGCAAGCCTCCAGGTCGGAGTACTGTCTTCCGAGCGGAGTACTGTC
TTCCGAACGGAGTACTGTCTTCCGAACGGAGTACTGTCTTCCGAACGGAGTACTGTCTTCCGAGCGGAGATTCTA
GCGAGGCCCGGAGTATAAATAGAGGCGCTTCGTTTACGGAGGGACAATTCAATTCAACCAAGCAAAGTGAACACG
TCGGTAAGCGAAAGCTAAGCAAATAAACAAGCGCAGCTGAACAAGCTAAACAATCTGCAGTAAAGTGCAAGTTAA
AGTGAATCAATTAAAAGTAACCAGCAACCAAGTAAATCAACTGCAACTACTGAAATCTGCCAAGAAGTAATTATT
GAATACAAGAAGAGAACTCTGAATAGGGAATTGGGAATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGT
CTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTGATCTACCATGGACTACAAAGACGATGACGACAAGCT
TGCGGCCGCGAATTCATCGATAGATCTCGAGGGTACCATGAACTCTTACTTCGAGCAGGGTGGTTTTTACGGGGC
CCATGGAGTGCACCAGGGCGGCGGTGGTGGAGACCAGTACCGCGGCTTCCCTCTGGGCCTCACGTATGCACAGCC
ACACGCTTTGCACCAGCCTCGTCCTCAGGATTCACCGTACGACGCGTCTGTCGCGGCGGCCTGCAAGCTCTATGC
TGGAGAGCAGCAATATCCTAAAGCAGATTGTTCAAAGCCAGGCGGTGAGCAGCAGAATGGCTATGGTGGGAAAGA
AGCCTGGGGCTCAGGTCTGGGAGCACTAGTGAGGCCGGCAGCATGCACTCCTGAAGCTCGATACAGTGAGTCGTC
AAGTCCTGGTAGAGCGCTTCCGTGGGGCAACCAGTGTGCACTTCCGGGATCAGCAGCATCAGCCGCGCAGCCAGT
GCACCAGCAGCCTACTAACCACACTTTCTACCCTTGGATGGCCATAGCAGGAGCGAACGGCCTCAGGAGACGAGG
AAGACAAACCTACACTAGATATCAAACGCTAGAATTAGAGAAAGAGTTCCACACGAACCACTACCTTACGCGAAG
GAGACGCATAGAGATGGCGCACGCGTTGTGCCTCACGGAGAGGCAAATCAAAATATGGTTCCAGAACCGAAGGAT
GAAGTTAAAGAAAGAGATCCAGGCTATAAAGGAGTTGAACGAGCAGGAGAAACAGGCGCAGGCGCAGAAGGCGGC
AGCGGCTGCTGCGGCGCCGCGCTGCCGCCC
```

Blast analysis using pUAST-flag forward primer:

Bombyx mori ultrabithorax (Ubx), mRNA
Sequence ID: ref|NM_001114160.1|Length: 765Number of Matches: 1
Related Information
Gene-associated gene details

UniGene-clustered expressed sequence tags
Range 1: 1 to 765GenBankGraphics Next Match Previous Match
Alignment statistics for match #1

Score	Expect	Identities	Gaps	Strand
1397 bits(756)	0.0	762/765(99%)	0/765(0%)	Plus/Plus

```
Query  348   ATGAACTCTTACTTCGAGCAGGGTGGTTTTTACGGGGCCCATGGAGTGCACCAGGGCGGC
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     ATGAACTCTTACTTCGAGCAGGGTGGTTTTTACGGGGCCCATGGAGTGCACCAGGGCGGC

Query  408   GGTGGTGGAGACCAGTACCGCGGCTTCCCTCTGGGCCTCACGTATGCACAGCCACACGCT
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    GGTGGTGGAGACCAGTACCGCGGCTTCCCTCTGGGCCTCACGTATGCACAGCCACACGCT

Query  468   TTGCACCAGCCTCGTCCTCAGGATTCACCGTACGACGCGTCTGTCGCGGCGGCCTGCAAG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121   TTGCACCAGCCTCGTCCTCAGGATTCACCGTACGACGCGTCTGTCGCGGCGGCCTGCAAG

Query  528   CTCTATGCTGGAGAGCAGCAATATCCTAAAGCAGATTGTTCAAAGCCAGGCGGTGAGCAG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181   CTCTATGCTGGAGAGCAGCAATATCCTAAAGCAGATTGTTCAAAGCCAGGCGGTGAGCAG

Query  588   CAGAATGGCTATGGTGGGAAAGAAGCCTGGGGCTCAGGTCTGGGAGCACTAGTGAGGCCG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241   CAGAATGGCTATGGTGGGAAAGAAGCCTGGGGCTCAGGTCTGGGAGCACTAGTGAGGCCG

Query  648   GCAGCATGCACTCCTGAAGCTCGATACAGTGAGTCGTCAAGTCCTGGTAGAGCGCTTCCG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301   GCAGCATGCACTCCTGAAGCTCGATACAGTGAGTCGTCAAGTCCTGGTAGAGCGCTTCCG

Query  708   TGGGGCAACCAGTGTGCACTTCCGGGATCAGCAGCATCAGCCGCGCAGCCAGTGCACCAG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361   TGGGGCAACCAGTGTGCACTTCCGGGATCAGCAGCATCAGCCGCGCAGCCAGTGCACCAG

Query  768   CAGCCTACTAACCACACTTTCTACCCTTGGATGGCCATAGCAGGAGCGAACGGCCTCAGG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421   CAGCCTACTAACCACACTTTCTACCCTTGGATGGCCATAGCAGGAGCGAACGGCCTCAGG

Query  828   AGACGAGGAAGACAAACCTACACTAGATATCAAACGCTAGAATTAGAGAAAGAGTTCCAC
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481   AGACGAGGAAGACAAACCTACACTAGATATCAAACGCTAGAATTAGAGAAAGAGTTCCAC

Query  888   ACGAACCACTACCTTACGCGAAGGAGACGCATAGAGATGGCGCACGCGTTGTGCCTCACG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  541   ACGAACCACTACCTTACGCGAAGGAGACGCATAGAGATGGCGCACGCGTTGTGCCTCACG

Query  948   GAGAGGCAAATCAAAATATGGTTCCAGAACCGAAGGATGAAGTTAAAGAAAGAGATCCAG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  601   GAGAGGCAAATCAAAATATGGTTCCAGAACCGAAGGATGAAGTTAAAGAAAGAGATCCAG

Query  1008  GCTATAAAGGAGTTGAACGAGCAGGAAAAACAGGCGCAGGCGCAAAAGGCGGCAACGGCT
             |||||||||||||||||||||||||||| ||||||||||||||||||| |||||||||| |||||
Sbjct  661   GCTATAAAGGAGTTGAACGAGCAGGAGAAACAGGCGCAGGCGCAGAAGGCGGCAGCGGCT

Query  1068  GCTGCGGCGGCCGCGGCTGCTGCCCAGGGACACCCCGAACATTAA    1112
             |||||||||||||||||||||||||||||||||||||||||||||
Sbjct  721   GCTGCGGCGGCCGCGGCTGCTGCCCAGGGACACCCCGAACATTAA    765
```

55

3. Full length *Tribolium custaneum Ubx* cDNA cloned into pUAST-flag:

>pUAST-flag forward primer:

```
AAAACTAGTGACCGTCGCTAGCGAAAGCTAAGCAAATAAACAAGCGCAGCTGAACAAGCTAAACAATCTGCAGTA
AAGTGCAAGTTAAAGTGAATCAATTAAAAGTAACCAGCAACCAAGTAAATCAACTGCAACTACTGAAATCTGCCA
AGAAGTAATTATTGAATACAAGAAGAGAACTCTGAATAGGGAATTGGGAATTGACGCAAATGGGCGGTAGGCGTG
TACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTGATCTACCATGGACTACAAAGACG
ATGACGACAAGCTTGCGGCCGCGAATTCTATGAACTCTTACTTCGAGCAGAGCGGCTTCTACGGCAGCCACCACC
ACCAGAGCGGGTCGGTGGCGGGCCACCACCACGAGCAGTCGGCGGCGGCGGCGGCGGCCTACCGCTCCTTCCCGC
TGTCGCTCGGCATGTCCCCGTACGCCTCCAGCCAGCACCACCACCACCACCTGCAGGCGCGGCCCCCGCAGGACT
CGCCGTACGACGCCTCGGTGGCGGCCGCCTGCAAGCTCTACTCCTCCGAGGGCCAGCAGAACTCCAACTACTCCT
CCAACTCGAAGCCGGACTGCTCCAAAGGCAACGCCGACCAGAACGGATACGCCTCGGTGGTGGCGGCGGCCGCGG
TCAAGGACGTTTGGCAAAGTGCGACTTCTGGCGGTGGCGCTAATCTCACGAACAGTTTGACGGGGCCGGTCAGGC
CGGCGGCATGCACGCCGGACTCCAGGGTTGGCTACGGGTCGGTCGGGCTCGTCGGCGGAGATCCGGCCTCGAGTC
CGGGGGCGGCCGCAGGACGGACGGGCAACTCGCTCTCGTGGAATAACCCCTGCAGTATCAACTCGACCTCTTCGC
AGCCCGTTGGCACGCAGATACACCAGCAGACCAACCACACGTTTTACCCCTGGATGGCCATTGCAGGAGCGAATG
GTCTCCGAAGGCGAGGCCGACAGACGTATACCCGGTACCAGACGCTGGAGCTGGAAAAAGAGTTCCACACAAACC
ATTACCTGACACGGCGGCGGCGGATCGAAATGGCTCACGCACTGTGTCTTACCGAACGACAGATAAAAATCTGGT
TTCAGAATCGTCCCATGAAACTCAAGAAAGAGATCCCAGCGATCAAAGAACTCAACGAGCAAGAAAAACAAGCAC
AGGCTCAAAAAGGGGGGGCGGGAATTGCAGCCGTCCCCGCCAAATTGGACCCGAATTAGTCTAAAGGATCTTTGG
GAAGGAACCTTAATTCCTGGGGTGGGACAATTTGGCAAACCTCCTTCCGGGGTTTTAAGGCTCTAAGGGAATAAA
AAATTTTAAGGGGAAAGGGGTGAACCATCGGGTCCAATTGGTGTGGGGTTTAAATTCCCCCCTTGGGGAACTGAA
TTAGGGAAGAAGGGGGGGAAAGCCTTTTAAAGAGAAAACCCTGTTTCCCCAAAAAATGTCTCCTATGTTGAGAAG
GCTCGCGTGTCCCCCATTTTTCTCCTCCAAAAAAAAAAAAAAGAAGAACCCCCGTTTTCTCTCAAAATAATTTTT
TTAGGGGGGGGTGTTAAAAAAAAATTTTTGGTTTGTCTAAACAAAAAGAGNNAA
```

>Ubx gene forward primer:

```
GGGGGGACGTTTCTACGGCAGCCACCACCACCAGAGCGGGTCGGTGGCGGGCCACCACCACGAGCAGTCGGCGGC
GGCGGCGGCGGCCTACCGCTCCTTCCCGCTGTCGCTCGGCATGTCCCCGTACGCCTCCAGCCAGCACCACCACCA
CCACCTGCAGGCGCGGCCCCCGCAGGACTCGCCGTACGACGCCTCGGTGGCGGCCGCCTGCAAGCTCTACTCCTC
CGAGGGCCAGCAGAACTCCAACTACTCCTCCAACTCGAAGCCGGACTGCTCCAAAGGCAACGCCGACCAGAACGG
ATACGCCTCGGTGGTGGCGGCGGCCGCGGTCAAGGACGTTTGGCAAAGTGCGACTTCTGGCGGTGGCGCTAATCT
CACGAACAGTTTGACGGGGCCGGTCAGGCCGGCGGCATGCACGCCGGACTCCAGGGTTGGCTACGGGTCGGTCGG
GCTCGTCGGCGGAGATCCGGCCTCGAGTCCGGGGGCGGCCGCAGGACGGACGGGCAACTCGCTCTCGTGGAATAA
CCCCTGCAGTATCAACTCGACCTCTTCGCAGCCCGTTGGCACGCAGATACACCAGCAGACCAACCACACGTTTTA
CCCCTGGATGGCCATTGCAGGAGCGAATGGTCTCCGAAGGCGAGGCCGACAGACGTATACCCGGTACCAGACGCT
GGAGCTGGAAAAAGAGTTCCACACAAACCATTACCTGACACGGCGGCGGCGGATCGAAATGGCTCACGCACTGTG
TCTTACCGAACGACAGATAAAAATCTGGTTTCAGAATCGTCGCATGAAACTCAAGAAAGAGATCCAAGCGATCAA
AGAACTCAACGAGCAAGAAAAACAAGCACAGGCTCAAAAAGCGGCGGCGGCAGCTGCAGCCGTCGCCGCACAAGT
GGACCCGAATTAGTCTAGAGGATCTTTGTGAAGGAACCTTACTTCTGTGGTGTGACATAATTGGACAAACTACCT
ACAGAGATTTAAAGCTCTAAGGTAAATATAAAATTTTTAAGTGTATAATGTGTTAAACTACTGATTCTAATTGGT
TGTGTATTTTAAGATTCCAACCTATGGGAACTGATTGAATGGGAAGCAGTGGTTGGAATGCCTTTAATGGAGGAA
AACCTTGTTTTGCTCAGAAAAAAATGCCTTCTAGTGGATGATGGAGGCTACTGGTTGACTTTCAACAATTCTACT
CCTTCCAAAAAAAGAAGGGGAAAAGTAAAAAGACCCCAAGGGACTTTCCTTCCAGAATTGCTAAAGTTTTTTGA
AGCCAGGCCGGGGTTTAGGAATAAGAACTCTTTGCTTGGCTTTGCTATTTCACCCCAAAAAGGGAAAAAACTTGG
CTTCGCTATCCCAGGAAAATTTTGGGAAAAAATTTTTTAAGGTTAAGGGCCCTTGGAAAAGGGGATCTATATACC
CGCCCTACCCAATTTTTTGGGAAAGGTTTTCACCTGGGTTTAAAAAAACCTCCCCCACCCCTCCCCCTCGGAAACC
CGAAAAACAAAAAAGGAAGGGCAATTTTGGTGTTGGGTAACCTTGGTTTATGGGCCCCTTTTAAGGGGGGACAAA
AAAACAAAAAAAACCCCCCCAATTTTTTCCCAAAAAT
```

>Reverse Complement of pUAST-flag reverse primer:

```
TTGGGCCTCCAAAAAAGGTGGGGAACATTCCTGGGAACCAAAAGTTCCAACCCTCCGAAGTGGGGAAATCTTTCC
CTCCGAGGGGGGATTTCTCCTCCAGAGGGGGAGATCTTTCCTCCCGAAGGGGGATATTTTCCTTCCGAACGGAGATA
CTGTCCTTCCGAACGGGGAAATTTTACGGAGGCCCGGGAGTATAAATAAGAGCGCTTTCTGTTTCGGAAGCGCCAA
TTCAATTCCAACCAGGCAAAGTGGACACGGTCGTTAAGCGGAAGCTTAAGCAAATTAACCAAGGGCAGCTGGACA
AAGTTAAACAATTTTCCAGTAAAAGTCCAAGTTAAAGTGAATCAATTTAAAAGTACCCGCCACCCAAGTAAATC
```

```
AACTGCAACTACTGAAATCTGCCAAGAAGTAATTATTGAATCCAAGAAGAGAACTCTGAATAGGGAATTGGGAAT
TGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATT
GATCTACCATGGACTACAAAGACGATGACGACAAGCTTGCGGCCGCGAATTCTATGAACTCTTACTTCGAGCAGA
GCGGCTTCTACGGCAGCCACCACCACCAGAGCGGGTCGGTGGCGGGCCACCACCACGAGCAGTCGGCGGCGGCGG
CGGCGGCCTACCGCTCCTTCCCGCTGTCGCTCGGCATGTCCCCGTACGCCTCCAGCCAGCACCACCACCACCACC
TGCAGGCGCGGCCCCCGCAGGACTCGCCGTACGACGCCTCGGTGGCGGCCGCCTGCAAGCTCTACTCCTCCGAGG
GCCAGCAGAACTCCAACTACTCCTCCAACTCGAAGCCGGACTGCTCCAAAGGCAACGCCGACCAGAACGGATACG
CCTCGGTGGTGGCGGCGGCCGCGGTCAAGGACGTTTGGCAAAGTGCGACTTCTGGCGGTGGCGCTAATCTCACGA
ACAGTTTGACGGGGCCGGTCAGGCCGGCGGCATGCACGCCGGACTCCAGGGTTGGCTACGGGTCGGTCGGGCTCG
TCGGCGGAGATCCGGCCTCGAGTCCGGGGGCGGCCGCAGGACGGACGGGCAACTCGCTCTCGTGGAATAACCCCT
GCAGTATCAACTCGACCTCTTCGCAGCCCGTTGGCACGCAGATACACCAGCAGACCAACCACACGTTTTACCCCT
GGATGGCCATTGCAGGAGCGAATGGTCTCCGAAGGCGAGGCCGACAGACGTATACCCGGTACCAGACGCTGGAGC
TGGAAAAGAGTTCCACACAAACCATTACCTGACACGGCGGCGGCGGATCGAAATGGCTCACGCACTGTGTCTTA
CCGAACGACAGATAAAAATCTGGTTTCAGAATCGTCGCATGAAACTCAAGAAAGAGATCCAAGCGATCAAAGAAC
TCAACGAGCAAGAAAACAAGCACAGGCTCAAAAGCGGCGGCGGCAGCTGCAGCCGTCGCCGCACAATACCCAA
ATTTC
```

>Reverse Complement of Ubx gene reverse primer

```
AGGATTAGNAAAAAAANNNTTTTGGGGGGGGGGGCAGACAAAAANNNNNNNNNNNAAAAAGAAGAGAAAAAGAAA
NGGCGGGGCCCGCAGCCCACACCCANTGGGGGGAGAGAAAAAAAATATATTTGTGGGGGAAGAGAAAAATATTAT
TTCATCCCCCGGGCGGGTGGTAAACAACATAAATAAAGGGGGATATTTTAAATTTTTTAAAAATTTGCGGAAAC
AAAGTTGCGCTCCCAACGGGTGGACCTTTCCGGACCCAATTTCCCCCCCGAGTGGGGGAACTCTCCTCGAGGGGA
TATTTCTTCCGAGGGAGTACTTTCCTCCAGGGGGGAATCTTCCTCCGGAGGGGATTTTTCTCTCCGAGCGGAGAT
TCTAGCGAGGCCCGGGGGTTAAAATAGAGGCGTTTTGGTTNCGGAGCGCCAATTCAATTCAAACAAGCAAAAGTGA
CCAGGTCGTTAAGGGAAAGCTAAGCAAATAACCAAGCGCAGCTGACCAAGCTAAACAATTTGCAGTAAAGTCCAA
GTTAAAGTGAATCAATTAAAAGTACCCACCACCCAAGTAAATCAACTGCAACTACTGAAATCTGCCAAGAAGTAA
TTATTGAATCCAAGAAGAGAACTCTGAATAGGGAATTGGGAATTGACGCAAATGGGCGGTAGGCGTGTACGGTGG
GAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTGATCTACCATGGACTACAAAGACGATGACGAC
AAGCTTGCGGCCGCGAATTCTATGAACTCTTACTTCGAGCAGAGCGGCTTCTACGGCAGCCACCACCACCAGAGC
GGGTCGGTGGCGGGCCACCACCACGAGCAGTCGGCGGCGGCGGCGGCGGCCTACCGCTCCTTCCCGCTGTCGCTC
GGCATGTCCCCGTACGCCTCCAGCCAGCACCACCACCACCACCTGCAGGCGCGGCCCCCGCAGGACTCGCCGTAC
GACGCCTCGGTGGCGGCCGCCTGCAAGCTCTACTCCTCCGAGGGCCAGCAGAACTCCAACTACTCCTCCAACTCG
AAGCCGGACTGCTCCAAAGGCAACGCCGACCAGAACGGATACGCCTCGGTGGTGGCGGCGGCCGCGGTCAAGGAC
GTTTGGCAAAGTGCGACTTCTGGCGGTGGCGCTAATCTCACGAACAGTTTGACGGGGCCGGTCAGGCCGGCGGCA
TGCACGCCGGACTCCAGGGTTGGCTACGGGTCGGTCGGGCTCGTCGGCGGAGATCCGGCCTCGAGTCCGGGGGCG
GCCGCAGGACGGACGGGCAACTCGCTCTCGTGGAATAACCCCTGCAGTATCAACTCGACCTCTTCGCAGCCCGTT
GGCACGCAGATACACCAGCAGACCAACCACACGTTTTACCCCTGGATGGCCATTGCAGGAGCGAATGGTCTCCGA
AGGCGAGGCCGACAGACGTATACCCGGTACCAGACGCTGGAGCTGGAAAAAGAGTTCCACACAAACCATTACCTG
ACACGGCGGCGGCGGATCGAAATGGCTCACGCACTGTGTCTTACCGAACGACAGATAAAAATCTGGTTTCAGAAT
CGTCGCATGAAACTCAAGAAAGAGATCCAAGCGATCAAAGAACTCAACGAGCAAGAAAACAAGCACAGGCTCAA
AAAGCGGCGGCGGCAGCTGCAGCCGCGCC
```

Blast analysis using pUAST-flag reverse primer sequence:

Tribolium castaneum ultrabithorax (Ubx), mRNA
Sequence ID: ref|NM_001039408.1|Length: 945Number of Matches: 1

See 1 more title(s)

Related Information
Gene-associated gene details
UniGene-clustered expressed sequence tags

Map Viewer-aligned genomic context
Range 1: 1 to 930GenBankGraphics Next Match Previous Match
Alignment statistics for match #1

Score	Expect	Identities	Gaps	Strand
1620 bits(877)	0.0	915/930(98%)	15/930(1%)	Plus/Plus

```
Query  579   ATGAACTCTTACTTCGAGCAGAGCGGCTTCTACGGCAGCCACCACCACCAGAGCGGGTCG
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     ATGAACTCTTACTTCGAGCAGAGCGGCTTCTACGGCAGCCACCACCACCAGAGCGGGTCG

Query  639   GTGGCGGGCCACCACCACGAGCAGTcggcggcggcggcggcggcCTACCGCTCCTTCCCG
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    GTGGCGGGCCACCACCACGAGCAGTCGGCGGCGGCGGCGGCGGCCTACCGCTCCTTCCCG

Query  699   CTGTCGCTCGGCATGTCCCCGTACGCCTCCAGCCAGCACCACCACCACCACCTGCAGGCG
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121   CTGTCGCTCGGCATGTCCCCGTACGCCTCCAGCCAGCACCACCACCACCACCTGCAGGCG

Query  759   CGGCCCCCGCAGGACTCGCCGTACGACGCCTCGGTGGCGGCCGCCTGCAAGCTCTACTCC
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181   CGGCCCCCGCAGGACTCGCCGTACGACGCCTCGGTGGCGGCCGCCTGCAAGCTCTACTCC

Query  819   TCCGAGGGCCAGCAGAACTCCAACTACTCCTCCAACTCGAAGCCGGACTGCTCCAAAGGC
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241   TCCGAGGGCCAGCAGAACTCCAACTACTCCTCCAACTCGAAGCCGGACTGCTCCAAAGGC

Query  879   AACGCCGACCAGAACGGATACGCCTCGGTGGTGGCGGCGGCCGCGGTCAAGGACGTTTGG
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301   AACGCCGACCAGAACGGATACGCCTCGGTGGTGGCGGCGGCCGCGGTCAAGGACGTTTGG

Query  939   CAAAGTGCGACTTCTGGCGGTGGCGCTAATCTCACGAACAGTTTGACGGGGCCGGTCAGG
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361   CAAAGTGCGACTTCTGGCGGTGGCGCTAATCTCACGAACAGTTTGACGGGGCCGGTCAGG

Query  999   CCGGCGGCATGCACGCCGGACTCCAGGGTTGGCTACGGGTCGGTCGGGCTCGTCGGCGGA
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421   CCGGCGGCATGCACGCCGGACTCCAGGGTTGGCTACGGGTCGGTCGGGCTCGTCGGCGGA

Query  1059  GATCCGGCCTCGAGTCCGGGGGCGGCCGCAGGACGGACGGGCAACTCGCTCTCGTGGAAT
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481   GATCCGGCCTCGAGTCCGGGGGCGGCCGCAGGACGGACGGGCAACTCGCTCTCGTGGAAT

Query  1119  AACCCCTGCAGTATCAACTCGACCTCTTCGCAGCCCGTTGGCACGCAGATACACCAGCAG
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  541   AACCCCTGCAGTATCAACTCGACCTCTTCGCAGCCCGTTGGCACGCAGATACACCAGCAG

Query  1179  ACCAACCACACGTTTTACCCCTGGATGGCCATTGCA---------------GGAGCGAAT
             ||||||||||||||||||||||||||||||||||||               |||||||||
Sbjct  601   ACCAACCACACGTTTTACCCCTGGATGGCCATTGCAGATTCTATGACTTTCGGAGCGAAT

Query  1224  GGTCTCCGAAGGCGAGGCCGACAGACGTATACCCGGTACCAGACGCTGGAGCTGGAAAAA
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  661   GGTCTCCGAAGGCGAGGCCGACAGACGTATACCCGGTACCAGACGCTGGAGCTGGAAAAA

Query  1284  GAGTTCCACACAAACCATTACCTGACACGGCGGCGGCGGATCGAAATGGCTCACGCACTG
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  721   GAGTTCCACACAAACCATTACCTGACACGGCGGCGGCGGATCGAAATGGCTCACGCACTG

Query  1344  TGTCTTACCGAACGACAGATAAAAATCTGGTTTCAGAATCGTCGCATGAAACTCAAGAAA
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  781    TGTCTTACCGAACGACAGATAAAAATCTGGTTTCAGAATCGTCGCATGAAACTCAAGAAA

Query  1404   GAGATCCAAGCGATCAAAGAACTCAACGAGCAAGAAAAACAAGCACAGGCTCAAAAAGCG
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  841    GAGATCCAAGCGATCAAAGAACTCAACGAGCAAGAAAAACAAGCACAGGCTCAAAAAGCG

Query  1464   GCGGCGGCAGCTGCAGCCGTCGCCGCACAA   1493
              ||||||||||||||||||||||||||||||
Sbjct  901    GCGGCGGCAGCTGCAGCCGTCGCCGCACAA   930
```

GATTCTATGACTTTC: DSMTF deletion at position 639 of the *Ubx* gene

4. Quadrant vestigial enhancer from *Apis mellifera* cloned into pH-stinger:

>*vestigial* gene forward primer:

```
GGGGGGGCCCTGGGGACGTGGAAGTGACGAGGACACGAAGAAAAAAAAAAAGAAAAAAAGAAAAAAAGAGGACG
CGATGGGTCGCGATATAGGGAAGGATGAGCGGGCTTTTTACCCACGATTTACATGATAAATCTCTTTGGAGGAGG
ATCGACTCGACATCAAAGTTGGAAGGCCGAGGAGTGCAGCGGGTTTGTTTGCGGCGGGCAGCGGAGAGGGAACGG
AGGCGGAATGGCGCGCAGGTAGCGCGCAAGAAAGGTCGGTCGGCCCCTCAGATGCTCGCAATTAATAAGAGCTTA
ATGGGGGCCGATATAGCGGTGCCACGCTCTCGGGACTCGTGCTTTCTTCGTTGTCCTCCCACCGAAATCATCGAT
GCCGACGACAAAAAGCCACGGCATGGCCGTAAACATTTCGCTCTCCCTCCTCCTCCACCTCCTCCCGTCCC
TCCCTCGACTCTGCTCCAACATATCCAAGACGCGTCGCCGCAAAAAGAGCCGCGCCGTTTTCATTCTCGCAGTCC
AACTTCCAGTTATATAAGAGTCGGGGAAAGCCTCTCGTTTCTCGCGAGAAGTCTAGACCAGATCTGCTGCAGCAT
GCTGTTGCCGAGCACAATTGATCGGCTAAATGGTATGGCAAGAAAAGGTATGCAATATAATAATCTTTTATTGGG
TATGCAACGAAAATTTGTTTCGTCAACGTATGCAATATTTTTTATTAAAAGAGGGTATGCAATGTATTTTATTAA
AAACGGGTATGCAATATAATAATCTTTTATTGGGTATGCAACGAAAATTTGTTTCGTCAAAGTATGCAATATTTT
TTATTAAAAGAGGGTATGCAATGTATTTTATTAAAAACGGGTATGCAATAAAAAATTATTTGGTTTCTCTAAAAA
GTATGCAGCACTTATTTTTTGATAAGGTATGCAACAAAATTTTACTTTGCCGAAAATATGCAATGTTTTTGCGAA
TAAATTCAACGCACACTTATTACGTGATGCAGCCAAGCTTGGCGAACAGTTCGGCTGGCGCGAGCCCCTGATGCT
CTTCGTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCCCTT
GGGGGTCGAATGGGCAAGTAGCCGGATCAAGCGTATGCACCCGCCGCATTGCTTCACCCATGAAGGATACTTTCT
CGGCAGGAGCAAGGTGAATTGACAGGGAATCCTGCCCCGGCACTCCGCCAATAGCAGCAAGTCCTTCCCGTTCAA
TGACAACTTCAAGACAAATTGCCAAAGGAACCCCTGTTGGGCAACCACAATAGCCGCCTTGCCTGCTTTGGATTT
ATCAAGGAACCGGAAGGGCGGCTTAGCCTAAGAAGTTTGGGCTTCCCAAATTTAAAAAAAAACTTAAAAAATTTC
CCCNNTATATTTTAGGTTTATAAAACCCCGGGCAGGTTGGGGTTAACAACATAATTCTGTCTCTTCACCATTAAA
AACCAAAAAATTTCTTCCCCCCTTTGGGAGGAAAAAAAAGGGGTTCCCCGCAGCCACACATATTTATTATGGGGG
GTCTCTTGGGGGTGGGGGCTTATTTTGTATTACCCGCCGGAGGAGAAAAACACAAAATTGTGTTTTTTCCACCCC
CCGGTCCGCACACAAAATNNTTAAACACCACAAAAACCTAGCCCCNNGTGGGT
```

Apis mellifera strain DH4 linkage group LG12, Amel_4.5, whole genome shotgun
sequence

Sequence ID: ref|NC_007081.3|Length: 11902654Number of Matches: 1

Related Information

Map Viewer-aligned genomic context

Range 1: 9131572 to 9132131GenBankGraphicsNext MatchPrevious Match

Score	Expect	Identities	Gaps	Strand
1022 bits(553)	0.0	558/560(99%)	1/560(0%)	Plus/Minus

Features:

protein vestigial-like isoform X2protein vestigial-like isoform X1

```
Query  18        GTGGA-AGTGACGAGGACACGAAGaaaaaaaaaaagaaaaaaagaaaaaaaGAGGACGC  76
                 ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  9132131   GTGGACAGTGACGAGGACACGAAGAAAAAAAAAAAGAAAAAAAGAAAAAAAGAGGACGC9132072


Query  77        GATGGGTCGCGATATAGGGAAGGATGAGCGGGCTTTTTACCCACGATTTACATGATAAAT  136
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  9132071   GATGGGTCGCGATATAGGGAAGGATGAGCGGGCTTTTTACCCACGATTTACATGATAAAT9132012


Query  137       CTCTTTGGAGGAGGATCGACTCGACATCAAAGTTGGAAGGCCGAGGAGTGCAGCGGGTTT  196
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  9132011   CTCTTTGGAGGAGGATCGACTCGACATCAAAGTTGGAAGGCCGAGGAGTGCAGCGGGTTT9131952


Query  197       GTTTGCGGCGGGCAGCGGAGAGGGAACGGAGGCGGAATGGCGCGCAGGTAGCGCGCAAGA  256
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  9131951   GTTTGCGGCGGGCAGCGGAGAGGGAACGGAGGCGGAATGGCGCGCAGGTAGCGCGCAAGA9131892


Query  257       AAGGTCGGTCGGCCCCTCAGATGCTCGCAATTAATAAGAGCTTAATGGGGGCCGATATAG  316
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  9131891   AAGGTCGGTCGGCCCCTCAGATGCTCGCAATTAATAAGAGCTTAATGGGGGCCGATATAG9131832


Query  317       CGGTGCCACGCTCTCGGGACTCGTGCTTTCTTCGTTGTCCTCCCACCGAAATCATCGATG  376
                 ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Sbjct  9131831   CGGTGCCACGCTCTCGGGACTCGTGCTTTCTTCGTTGTTCTCCCACCGAAATCATCGATG9131772


Query  377       CCGACGACAAAAAGCCACGGCATGGCCGTAAAACATTTCGctctccctcctcctcctcca  436
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  9131771   CCGACGACAAAAAGCCACGGCATGGCCGTAAAACATTTCGCTCTCCCTCCTCCTCCTCCA9131712


Query  437       cctcctcccgtccctccctcGACTCTGCTCCAACATATCCAAGACGCGTCGCCGCAAAAA  496
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  9131711   CCTCCTCCCGTCCCTCCCTCGACTCTGCTCCAACATATCCAAGACGCGTCGCCGCAAAAA9131652


Query  497       GAGCCGCGCCGTTTTCATTCTCGCAGTCCAACTTCCAGTTATATAAGAGTCGGGGAAAGC  556
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  9131651   GAGCCGCGCCGTTTTCATTCTCGCAGTCCAACTTCCAGTTATATAAGAGTCGGGGAAAGC9131592


Query  557       CTCTCGTTTCTCGCGAGAAG  576
                 ||||||||||||||||||||
Sbjct  9131591   CTCTCGTTTCTCGCGAGAAG  9131572
```

60

4. Sequence Homology analysis of Ubx from *Apis*, *Bombyx*, *Tribolium*:

```
Apis         MNSYFEQTAGGFYGSHHHQTG----AASQHHDP--ATAAAYRSFPLGLGMSPYASTQHHH 54
Tribolium    MNSYFEQS--GFYGSHHHQSGS---VAGHHHEQSAAAAAAYRSFPLSLGMSPYASSQHHH 55
Bombyx       MNSYFEQG--GFYGAHGVHQGG---GGGDQ----------YRGFPLGL---TYAQPHALH 42
Drosophila   MNSYFEQAS-GFYGHPHQATGMAMGSGGHHDQTASAAAAAYRGFPLSLGMSPYANHHLQR 59
             *******   ****        *      ...:          **.***.*   .**. : :

Apis         HTSSSLGIHPGGGTNTRPPQDSPYDASVATAC-KLYSTTP--------EATGHTTSSYST 105
Tribolium    HH-----------LQARPPQDSPYDASVAAAC-KLYSSE------------GQQNSNYSS 91
Bombyx       QP---------------RPQDSPYDASVAAAC-KLYAGE----------------QQYP- 69
Drosophila   TT-----------------QDSPYDASITAACNKIYGDGAGAYKQDCLNIKADAVNGYKD 102
                               ********:::** *:*.                     . *

Apis         TAAKDCKQQDQASAHQNGYAAVMAAAAVKDVWQSATSGANS-QSNSV---VRPSACTPEG 161
Tribolium    NSKPDCSK---GNADQNGYASVVAAAAVKDVWQSATSGGGANLTNSLTGPVRPAACTPD- 147
Bombyx       --KADCSK--PGGEQQNGYGG-------KEAWG---SGLGA--------LVRPAACTPEA 107
Drosophila   IWNTGGSNGGGGGGGGGGGGGAGGTGGAGNANGGNAANANGQNNPAGGMPVRPSACTPD- 161
                 . .:   ..   .* ..       :.     :. ..     ***:****:

Apis         TRVGSYG------------GLVGGDPASSPGNNSSS--------RSLTSSWN-TCSLNSS 200
Tribolium    SRVG-YGSV----------GLVGGDPASSPGAAAGR--------TGNSLSWNNPCSINST 188
Bombyx       RYSE----------------------SSSPGR---------------ALPWGNQCALPGS 130
Drosophila   SRVGGYLDTSGGSPVSHRGGSAGGNVSVSGGNGNAGGVQSGVGVAGAGTAWNANCTISGA 221
                               : * *                  .*.   *:: .:

Apis         ASQPVATQLHQQPSNHTFYPWMAIA-----------------------------G 226
Tribolium    SSQPVGTQIHQQ-TNHTFYPWMAIADSM-----------------------TFG 218
Bombyx       AAS-AAQPVHQQPTNHTFYPWMAIA-----------------------------G 155
Drosophila   AAQTAAASSLHQASNHTFYPWMAIAGKIRSDLTQYGGISTDMGKRYSESLAGSLLPDWLG 281
             ::. ..     :* :**********                               *

Apis         ANGMRRRGRQTYTRYQTLELEKEFHTNHYLTRRRRIEMAHSLCLTERQIKIWFQNRRMKL 286
Tribolium    ANGLRRRGRQTYTRYQTLELEKEFHTNHYLTRRRRIEMAHALCLTERQIKIWFQNRRMKL 278
Bombyx       ANGLRRRGRQTYTRYQTLELEKEFHTNHYLTRRRRIEMAHALCLTERQIKIWFQNRRMKL 215
Drosophila   TNGLRRRGRQTYTRYQTLELEKEFHTNHYLTRRRRIEMAHALCLTERQIKIWFQNRRMKL 341
             :**:*******************************:*******************

Apis         KKEIQAIKELNEQEKQAQAQKAAAAAAAAAHQQQAAGGGPEGAN 330
Tribolium    KKEIQAIKELNEQEKQAQAQKAAAAAAAVAAQVDPN-------- 314
Bombyx       KKEIQAIKELNEQEKQAQAQKAAAAAAAAAAAQGHPEH----- 254
Drosophila   KKEIQAIKELNEQEKQAQAQKAAAAAAAAAVQGGHLDQ----- 380
             **************************.*
```

4. Sequence Homology analysis of *vg*-QE from *Apis* and *Drosophila*:

```
Drosophila   --------------GGATAC--AAGTG-CAGGGACAC-ACGGGAAAATATTATGCTCTAAT 43
Apis         GGGGGGGCCCTGGGGACGTGGAAGTGACGAGGACACGAAGAAAAAAAAAAAAGAAAAAAA 60
                     ***     ***** * ****** * *  **** *  * *   **

Drosophila   GGAAGTTAGAA--------------CAATTTAG------ATTT-----CACTTCAATAAC 78
Apis         GAAAAAAAGAGGACGCGATGGGTCGCGATATAGGGAAGGATGAGCGGGCTTTTTACCCAC 120
             * **   ***              * ** ***       **      *  ** *  **

Drosophila   AAAATAAAAACTAGATC-------AAAAAAATTGTTTT-ATATTAA--TTATAA--CTAA 126
Apis         GATTTACATGATAAATCTCTTTGGAGGAGGATCGACTCGACATCAAAGTTGGAAGGCCGA 180
              *  ** *  ** ***       * *  ** *  * * *** **  **  **  * *

Drosophila   A------------CTTTTTCGCTTC---TAATTAAGA------------------CTCA 152
Apis         GGAGTGCAGCGGGTTTGTTTGCGGCGGGCAGCGGAGAGGGAACGGAGGCGGAATGGCGCG 240
                         ** ** ** *    *    ***                   * *

Drosophila   C---TATAGTTTAAACAAGATAAAAC-----------TACT----GTTGATTTGATTTTA 194
Apis         CAGGTAGCGCGCAAGAAAGGTCGGTCGGCCCCTCAGATGCTCGCAATTAATAAGAGCTTA 300
             *   ** *   **  *** *    *            * **    ** ** ** ***

Drosophila   GT-----TTGA------------ACACTCT--------GTACTTGTTTTA---------- 219
Apis         ATGGGGGCCGATATAGCGGTGCCACGCTCTCGGGACTCGTGCTTTCTTCGTTGTCCTCCC 360
             *      **            ** ****      ** *** **

Drosophila   ----GAATCTTC----------ATAGAGGAT--------------TAGGATAGTTTTCTCT 252
Apis         ACCGAAATCATCGATGCCGACGACAAAAAGCCACGGCATGGCCGTAAAACATTTCGCTCT 420
                 **** **          * *              ** * * ** ****

Drosophila   G--------TGCAGGACCGCACTATG------CGTCAACGCTG-----GCG-ATCGAA-- 290
Apis         CCCTCCTCCTCCTCCACCTCCTCCCGTCCCTCCCTCGACTCTGCTCCAACATATCCAAGA 480
                      * *   *** *    *        * ** ** ***     *  *** **

Drosophila   --CGTTATCGGGTCATAAATCGCCACGCTCTCTTCATT-----AGGCCAAA----AGGTG 339
Apis         CGCGTCGCCG---CAAAAAGAGCCGCGCCGTTTTCATTCTCGCAGTCCAACTTCCAGTTA 537
               ***     **   ** *** *** ***   * ******    ** ****    ** *

Drosophila   ---AAAGGTGGGGACAGGTAGTGCGGCTTTCGC------TTTTGA----AGCCGCTG--- 383
Apis         TATAAGAGTCGGGGAAAGCCTCTCGTTTCTCGCGAGAAGTCTAGACCAGATCTGCTGCAG 597
                **  ** ***  * *      ** * ****      * * **    * * ****

Drosophila   ---------GCTGA----AAGTGAAATGCTAACTG-----------AAGGAGCTGCAGAG 419
Apis         CATGCTGTTGCCGAGCACAATTGATCGGCTAAATGGTATGGCAAGAAAAGGTATGCAATA 657
                      ** **    ** ***   ***** **            ** *    ****

Drosophila   GAAATATCTTT----------------GGCAGTCTGCTT-------GTTTGCA-------- 449
Apis         TAATAATCTTTTATTGGGTATGCAACGAAAATTTGTTTCGTCAACGTATGCAATATTTTT 717
             **  ******                * * * ** **       ** ****

Drosophila   CATTAAT---GGGAATTCCACGGCCTCAGCTGAG--CGAGTGAGTGA-ATGGCAAT---- 499
Apis         TATTAAAAGAGGGTATGCAATGTATTTTATTAAAAACGGGTATGCAATATAATAATCTTT 777
             *****   *** ** * * *     * *  ** ** * *   * ** *** 

Drosophila   -----GGGATG--ATGACGACTTCTGGC----------------TTCT------------ 524
Apis         TATTGGGTATGCAACGAAAATTTGTTTCGTCAAAGTATGCAATATTTTTTTATTAAAAGAG 837
                  ** ***  * **  * ** *  *                   ** *

Drosophila   GGTACGGGA-----------AAAAAAAGGCACGCGGCA-----------TGGCATCCC--- 560
```

```
Apis        GGTATGCAATGTATTTTATTAAAAACGGGTATGCAATAAAAAATTATTTGGTTTCTCTAA 897
            ****  *   *         *****  ** * **    *          ***  ** *

Drosophila  ---GTATCCAGTATTCA------------GTATCCCA--------------GCCGCA--- 588
Apis        AAAGTATGCAGCACTTATTTTTTGATAAGGTATGCAACAAAATTTTACTTTGCCGAAAAT 957
               **** *** * *          **** * *              **** *

Drosophila  -----------------------CCA--GCACCATCATCA------------------- 603
Apis        ATGCAATGTTTTTGCGAATAAATTCAACGCACACTTATTACGTGATGCAGCCAAGCTTGG 1017
                                    **   ****   * ** *

Drosophila  --------TC------CGCA------------TCCTCATCC---TCATCCTCAT------ 628
Apis        CGAACAGTTCGGCTGGCGCGAGCCCCTGATGCTCTTCGTCCAGATCATCCTGATCGACAA 1077
                    **      ***            ** ** ***   ******* **

Drosophila  -------TCTCATC----------------------------------------GTC----- 638
Apis        GACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCCCTTGGGGGTCGAATG 1137
                   *   ****                                      ***

Drosophila  -------------ATCA---GCATGCAT--GGCACACT----CGCCAACG---------- 666
Apis        GGCAAGTAGCCGGATCAAGCGTATGCACCCGCCGCATTGCTTCACCCATGAAGGATACTT 1197
                         ****  * *****   * * ** *   * ** * *

Drosophila  TGCCGGC-----CAAAGT----TGTC---GAATC--GCGACGGCA----GCCAA--GCA- 705
Apis        TCTCGGCAGGAGCAAGGTGAATTGACAGGGAATCCTGCCCCGGCACTCCGCCAATAGCAG 1257
             *   ****      *** **     ** *   ***** **  *****    ***** ***

Drosophila  CAAA-------CATT----------------ACACATCCCCACAAGCACATCCA-----C 737
Apis        CAAGTCCTTCCCGTTCAATGACAACTTCAAGACAAATTGCCAAAGGAACCCCTGTTGGGC 1317
             ***        * **              *** ** *** * * **  *         *

Drosophila  ATCCACA---------------TCCTCCGACCTCGTC------------------CTCGTC 765
Apis        AACCACAATAGCCGCCTTGCCTGCTTTGGATTTATCAAGGAACCGGAAGGGCGGCTTAGC 1377
             * *****         * **   *  *   *  **                   **   *

Drosophila  C------------GCATCC-----TCAAAA----------------CTCCTCTCCA----- 788
Apis        CTAAGAAGTTTGGGCTTCCCAAATTTAAAAAAAAACTTAAAAAATTTCCCCNNTATATTT 1437
             *          ** ***     * ****                ***  *  *

Drosophila  ---GTT-----------------GTTGG----------ATATTTTTCTCTC--------- 809
Apis        TAGGTTTATAAAACCCCGGGCAGGTTGGGGGTTAACAACATAATTCTGTCTCTTCACCATT 1497
                ***                 *****          *** ** * ****

Drosophila  --GAATCGAG---------------TTTGGAA-------------------TGTTGTCA- 832
Apis        AAAAACCAAAAAATTTCTTCCCCCCTTTGGGAGGAAAAAAAAAGGGTTCCCCGCAGCCAC 1557
               ** * *                ***** *                 *  * ** **

Drosophila  ACGCA-----------------------AATGCAAGCCTAT------------------ 850
Apis        ACATATTTATTATGGGGGGTCTCTTGGGGGTGGGGGCTTATTTTGTATTACCCGCCGGAG 1617
             **  *                        **   ** ***

Drosophila  ---------GCAA------GTTTTTGGTATCTCCTG--CCGCTCA----ATTGC------ 883
Apis        GAGAAAAACACAAAATTGTGTTTTTTCCACCCCCCGGTCCGCACACAAAATNNTTAAACA 1677
                      ***      ******   * * ** *   **** **      **

Drosophila  ------------AGTTGCTA------ 891
Apis        CCACAAAAACCTAGCCCCNNGTGGGT 1703
                         **    *
```