Comparative analysis of selected targets of Ubx from *Apis, Bombyx* and *Drosophila*

Thesis submitted towards the Partial fulfillment of

BS-MS dual degree programme

By

Afsah Hasan V P

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 "**Comparative analysis of selected targets of Ubx from *Apis, Bombyx* and *Drosophila***" towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents the research carried out by **Afsah Hasan V P** at Department of Biology, Indian Institute od Science Education and Research, Pune under the supervision of **Prof. L. S. Shashidhara** during the academic year 2015-2016.

Prof. L. S. Shasidhara

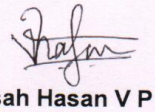
(Head of the Department of Biology)

Date: 28th March, 2016

Place: Pune

DECLARATION

I hereby declare that the matter embodied in the report entitled "**Comparative analysis of selected targets of Ubx from *Apis, Bombyx* and *Drosophila***" are the results of the investigations carried out by me at the Department of Biology, Indian Institute od 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.



Afsah Hasan V P

Date: 28th March, 2016

Place: Pune

Abstract

Hox genes are a group of genes that control the body plan of an organism along anterior posterior axis. Hox genes function as master control genes and they regulate the expression of downstream target genes. While insects such as *Apis mellifera* and *Bombyx mori* possess two pairs of wings, *Drosophila* has its hind wing modified into haltere. *Ultrabithorax* is a Hox family gene that functions to specify third thoracic segment in insects. In *Drosophila,* it suppresses the development of wing in T3 to specify the formation of haltere. *Hedgehog* (*hh*) and *vestigial* (*vg*) are targets of Ubx that are common to all the three species and are very important genes in the context of wing patterning and development. In this context, we investigate if the functions of *hh* and *vg* are conserved across the three insects and how proteins evolve independent of their functions. Quadrant enhancer of *vg* has been identified as a target of Ubx. *Vg* quadrant enhancer from *Apis* showed identical expression pattern between wing and haltere unlike its fly counterpart. Here, we investigate if quadrant enhancer of *vg* in *Bombyx* (which is also a target of Ubx) is differentially expressed or not. We test this hypothesis by generating transgenic *Drosophila* expressing GFP under the control of enhancer of *vestigial* from *Bombyx*. The construct is currently being tested for its expression. Also we made an attempt to make transgenic flies expressing *hh* and *vg* from *Apis* and *Bombyx* but could not succeed.

List of Figures

Acknowledgements

I take this opportunity to thank my mentor Prof. L S Shashidhara for his valuable, inspiring and encouraging guidance and for providing me with scientific freedom that enabled me to work. I would like to thank Dr. Gayathri Pananghat for her valuable inputs and encouragement throughout the study. I thank Dr. Farhat Habib for the help he provided me for the bioinformatic analysis.

I thank my colleagues Dr. Sreeharsha, Dhanashree and Dr. Swati for teaching me various techniques and helping guiding me throughout the work. I also thank my all other lab members for constant support during the study. I also thank my family and friends or providing me encouragement and showing faith on me.

Chapter 1
Introduction

1.1 *Drosophila* as a Model system

The fruit fly, *Drosophila Melanogaster* is the most extensively used and well-studied model organism. There are many characteristics that make *Drosophila* the ideal organism for studies regarding animal development, behavior, evolution, neurobiology, genetic diseases etc. The relationship between the genome of fly to the human genome make it a preferred model for many studies. About 75% of human disease genes have known homologs present in *Drosophila* (Lawrence et al, 2001). Also most of the basic mechanisms and pathways that control growth and development are conserved between the two species.

Drosophila was first used as a model organism by Thomas Hunt Morgan. It belongs to the Class Insecta and Order Diptera. Many practical features of this organism make it a user friendly model for scientists. *Drosophila* has a short life cycle of about 8-14 days depending on the environmental conditions. This provides the opportunity of producing and observing several generations of flies in a matter of months. Low number of chromosomes and relatively small genome is advantageous for genetic studies in *Drosophila*. In a short span of time, they produce large number of externally laid embryos. *Drosophila* can be genetically modified by different ways and the modifications can be maintained through balancer chromosomes.

Advanced genetic tools has been developed on the organism over the time. Path breaking tools such as UAS-GAL4 (Brand and Perrimon, 1993), FLP/FRT (Xu and Rubin, 1993) and RNAi knockdown have enhanced the possibilities of research using *Drosophila*. Invention of balancer Chromosomes made *Drosophila* the best system for genetic research as mutations can be maintained permanently without losing them by recombination. Balancer Chromosomes prevent crossing over between homologous chromosomes during meiosis and thus maintain any mutation in heterozygous conditions. It also carries a dominant marker to identify the presence of the mutation.

Similar to many other insects, *Drosophila* undergoes a four-stage life cycle: egg, larva, pupa and adult fly. At normal scenario, it completes the lifecycle in 12 days and starts laying eggs after 24 hours of emergence. The hatching of the eggs happen after 12-15 hours. The larva moults into a second instar in around 24-25 hours. It undergoes one more successive molts to become a third instar larva. Third instar larva feeds and pupate in a relatively drier area. In the 3-4 days of pupal stage, larva undergoes extensive reorganization of the body plan (morphogenesis) to become an adult fly.

1.2 Silkworm as a Lepidopteran model

Apart from its economic value, Silkworm (*Bombyx Mori*) is also a good model organism for studies on genetics and development. Genetic manipulation tools and availability of genetically homogenous inbred lines make them a valuable to genetic studies beside *Drosophila.* The genome of silkworm was sequenced in 2004 independently by Chinese and Japanese groups (Xia et al, 2004, Mita et al, 2004). *B. mori* has a large genome compared to *Drosophila* and contains 28 chromosomes. Two studies of *B. mori* genome are available online through two databases, namely Japanese Kaikobase (Shimomura et al, 2009) and Chinese SilkDB (Xia et al, 2004).

1.3 Early patterning of *Drosophila*

When a single celled embryo develops into an adult organism, it undergoes different levels of pattern formation. The initial process of pattern formation allows the organism to develop a well-organized body structure. Positioning of various body organs are brought out by this process of identifying spatial and temporal coordinates. This orchestration is executed by cellular and molecular mechanisms of regulating genes at each level of development. The first phase of polarization occurs at two axes named antero-posterior (AP) and dorso-ventral (DV) axis. Origin of antero-posterior specification comes from the polarity of the egg. The sequential expression of different sets of genes establishes the body plan along this axis. The Four different sets of genes functional during developmental along the AP axis are the gap genes, pair rule genes, segmentation genes and homeotic (Hox) genes.

Drosophila has 8 Hox genes, and they all are located on the third chromosome distributed in two clusters (figure 1.1): the one closer to the centromere is named as

Antennapedia complex (ANT-C) and the distal one is called Bithorax complex (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, respectively. 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).

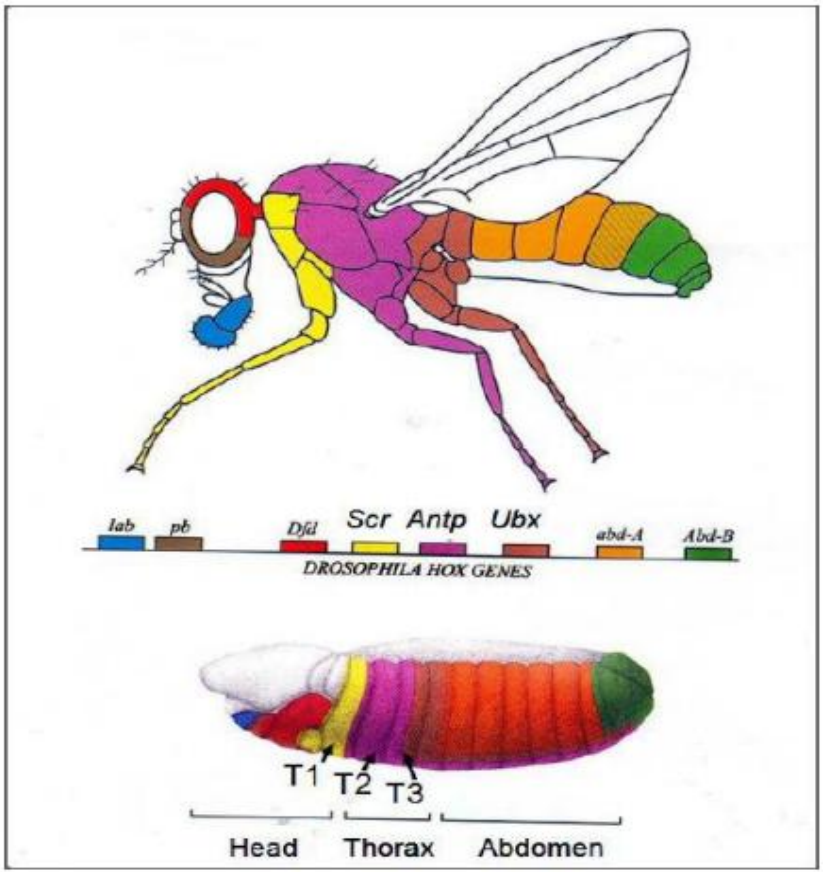


Figure 1.1: Hox genes in the *Drosophila* melanogaster (Sadava, D.et al.)

Hox genes function as master control genes and they regulate the expression of downstream target genes. Evolution of this family of genes may have contributed

significantly to the diversity in the body plan of organisms we see today. While insects such as *Apis mellifera* and *Bombyx mori* possess two pairs of wings, *Drosophila* has its hind wing modified into haltere (figure 1.2). *Ultrabithorax* is a Hox family gene that functions to specify third thoracic segment in insects. In *Drosophila,* it suppresses the development of wing in T3 to specify the formation of haltere.

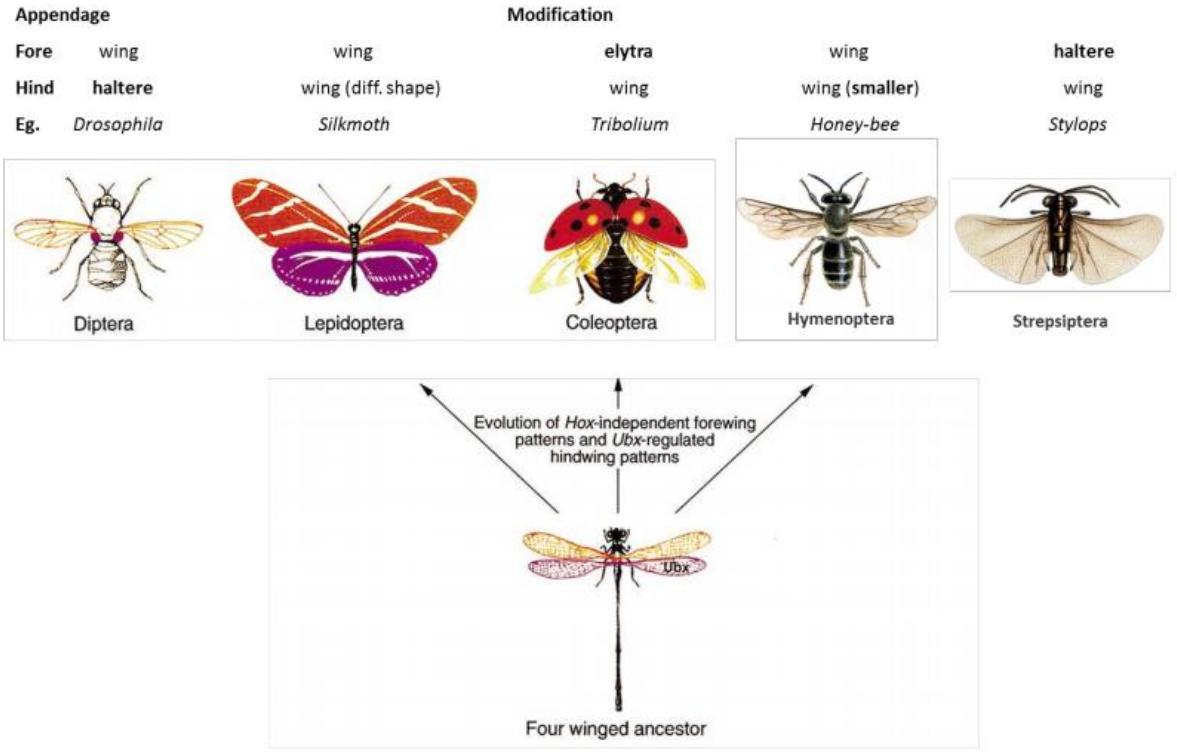


Figure 1.2: Diversity in the appendage evolution of various Insect species (modified from Carrol et al, 2000, Kathirthamby et al, 2006, Omlet.us 2010)

In order to understand the evolution of Ubx function to specify haltere development in the T3 segment, direct targets of Ubx have been identified by Chromatin Immunoprecipitation in three species *Drosophila, Bombyx* and *Apis* (Agarwal et al, 2011, Choo et al, 2011). Amongst the targets that are common to all the three species (which are diverged for nearly 300 million years) are *hedgehog* (*hh*) and *vestigial* (*vg*). In *Drosophila*, both *hh* and *vg* have are very important genes in the context of wing patterning and development. While vg is a pro-wing gene and specifies wing identity to

epithelial cells, *hh* mediates patterning along the antero-posterior axis. It is possible that

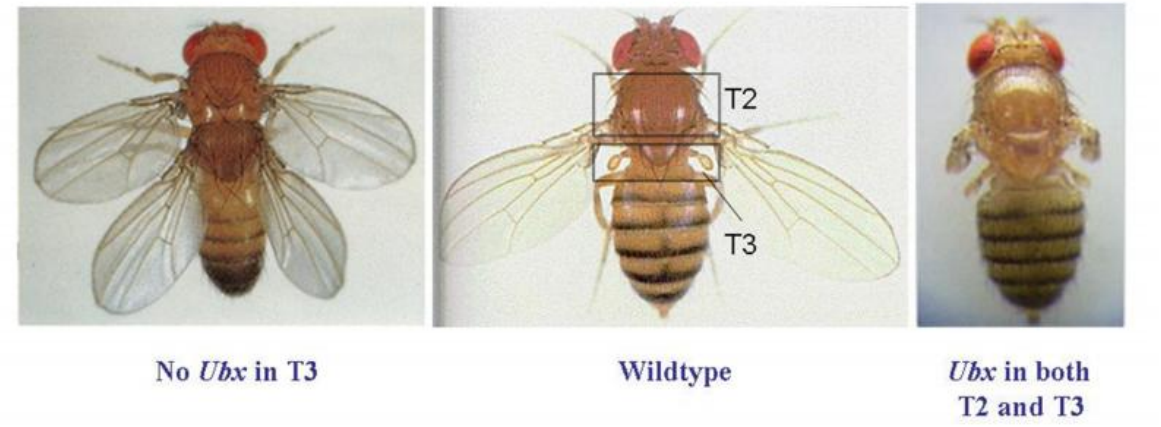


Figure 1.3: **Ubx suppresses wing development to specify haltere development.** Middle Panel shows a Wild type fly with two wings and two halteres. Suppression of hox gene Ubx is leads to Four-wing Fly (Left panel). Fly with four haltere arouse when Ubx is ectopically expressed in T2 segment. (Figure Courtesy Ed Lewis, 1963).

vg is differentially regulated between wing and haltere in *Drosophila*, it is expressed in identical patterns in the forewing and hindwing of *Apis* and *Bombyx* (Naveen et al., unpublished). *Drosophila vg* has two enhancers: D/V boundary enhancer and quadrant enhancer. Quadrant enhancer of *vg* has been identified as a target of Ubx in *Drosophila* and is expressed differentially between wing and haltere discs. Earlier results from our lab suggest that quadrant enhancer of *vg* is a target of Ubx in *Apis* too, although it is not differentially expressed between forewing and hindwing. Here, we investigate if quadrant enhancer of *vg* in *Bombyx* (which is also a target of Ubx) is differentially expressed or not.

These studies would help us to understand the relative contribution of new genes coming under the regulation of Ubx vs changes in the function and expression patterns of genes that are targeted by Ubx in all the three species during the evolution of hindwing/haltere morphology. Also by looking at *vg* from related species would help us to understand how the proteins evolved independent of its regulatory elements.

1.4 Objectives of the study

1. Generating transgenic flies to express Vg and Hh from *Apis* and *Bombyx* in *Drosophila*.

We would use GAL4-UAS system to over-express these proteins in flies. We would compare the phenotypes caused by the over-expression of these genes in *Drosophila* system. This helps to understand if their function is similar or diverged.

Following transgenics would be generated as part of the study:

UAS-Vg (*Apis*), UAS-Hh (*Apis*), UAS-Vg (*Bombyx*), UAS-Hh (*Bombyx*)

We would design appropriate primers and isolate full-length cDNA from the respective insect larvae by RT-PCR and sub-clone the same into pUAS vector. The clone would be sequence-verified before using for generation of transgenic flies.

2. Generation of transgenic flies expressing GFP under the quadrant enhancer of *vg* from *Bombyx*. Extensive bioinformatics study carried out in our lab has identified a 600bp region within the 4^{th} intron of *vg* gene to be equivalent to quadrant enhancer of *vg* in *Drosophila*. This region would be isolated by PCR and cloned in pH-Stinger vector to raise transgenic flies. GFP expression pattern would indicate the properties of the enhancer, which would be compared to quadrant enhancer of *vg* from *Apis* and *Drosophila*.

3. A comparative GO (Gene Ontology) analysis of targets of Ubx was done before in the lab using ChIP-Seq data of *Bombyx* and *Apis* (Thesis work by Sreeharsha). To rule out the possibility of biasness in the analysis, we would do GO analysis of all the genes from *Drosophila*, *Apis* and *Bombyx*.

Chapter 2

Materials and Methods

2.1 Cloning of *vestigial* quadrant enhancer in pH-Stinger:

vestigial quadrant enhancer in *Drosophila* is reported to be of 891 bp in length. One such corresponding region in the *Bombyx Mori*, which was bound by Ubx was found to be of 700 bp in length. We suspected this sequence to be the analogue of *vestigial* quadrant enhancer in *Drosophila* based on the array of transcription factor binding sites present in this region (The in-silico analysis was done with Shreeharsha). A 726 bp region of interest was cloned into pH-stinger vector, which was then sent for injection to raise transgenic flies at NCBS facility. pH-stinger is a transgenic reporter construct which has GFP as the reporter gene (Barolo et al., 2004). The steps involved in cloning is detailed as follows.

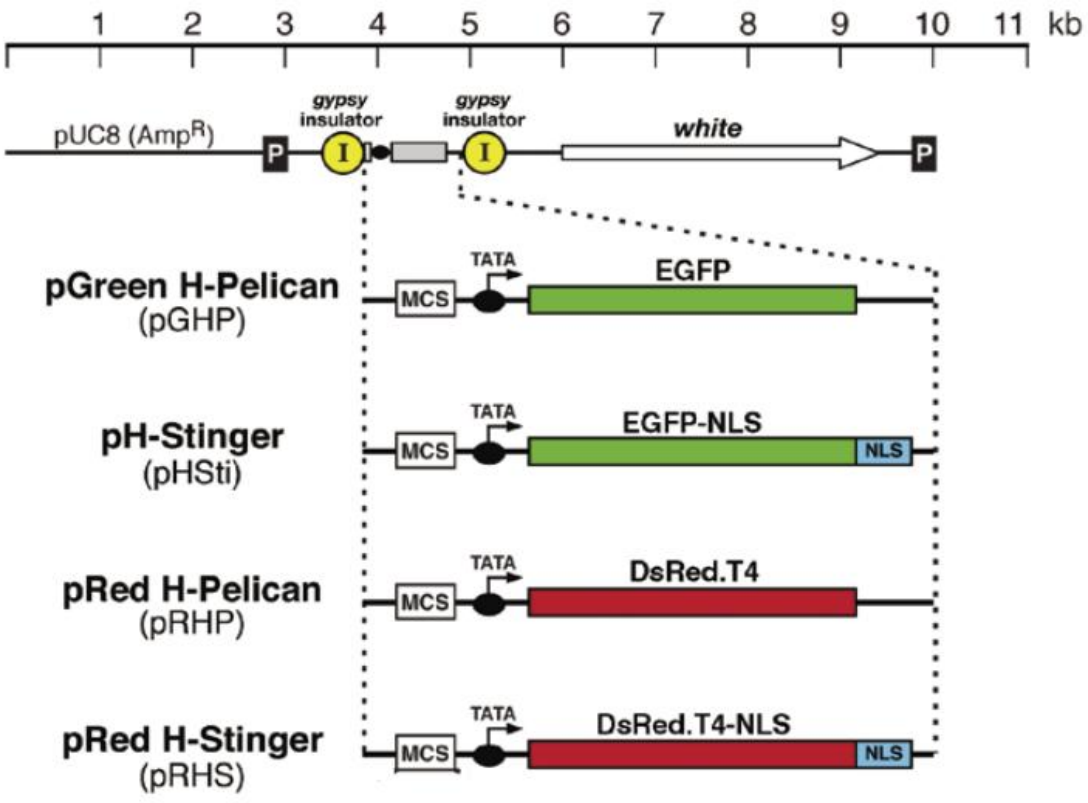


Figure 2.1: **Vector map of pH-Stinger**

First, the genomic DNA from the larvae of *Bombyx mori* was extracted using standard phenol-chloroform extraction method. The larva of *Bombyx mori* was selected as the tissue source for genomic DNA extraction. Tissue samples were ground using micro pestle. Routine phenol and chloroform extractions were performed to remove protein and phenol respectively. It was then treated with RNAse A and Proteinase K to get rid of RNA and protein. Precipitation with phenol /chloroform provided purified genomic DNA.

Gene/sequence		**Primer Sequence**	**Restriction Enzymes**
vestigial Quadrant enhancer	Forward Reverse	5' GCCTCTAGATGTCAACTTTATTACTCGACATTACT 3' 5' AAGGATCCGTTCTTTTCGGCATATTACTTTTACA 3'	XbaI BamHI
Vestigial full length cDNA from *B.mori* (set1)	Forward Reverse	5' ATTAGGTACCATGGAGAGGTGTCGTGTGTTG 3' 5' TGGCTCTAGATTACCGTCTTACAGAACTATCCAAAC 3'	KpnI XbaI
Vestigial full length cDNA from (set2)	Forward Reverse	5' ATATGGTACCATGGCGGTGAGCTGCCCCGAGGT 3' 5'ATCCTCTAGATCAGAACCAGTACATGTCCTTGGACG 3'	KpnI XbaI
Hedgehog full length cDNA from *B.mori*	Forward Reverse	5' TAAGAATTCAATGAACCAGTGGCCGGGAGT 3' 5' CTGGGGTACCTTATCGATATCTATACGATGCTGGT 3'	EcoRI KpnI

Table 1: Primer Sequences and restriction enzymes of various genes /sequences

Primers were designed and synthesized for the corresponding genome region (Table 1). 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 an annealing temperature of 58° C. Amplicon of an expected size of DNA was obtained by the PCR. A gradient PCR was performed to estimate the optimum annealing temperature for PCR. After the amplification, both pH-stinger 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 Qiagen®-Miniprep kit and sent for sequencing. Sequencing results are appended at the results section. Sequenced clones were sent for injection to generate the transgenic line.

2.2 Cloning of Vestigial Full length cDNA from *Bombyx mori*:

A gene model of *vestigial* was identified from Silkworm genome database (gene ID: BGIBMGA012721). The gene predicted to be of 714 bp long. Primers were designed and synthesized for the amplification of gene (Table 1). Primers were sent to CDFD, Hyderabad for gene amplification from the cDNA library that they are maintaining.

Unfortunately, there was no Amplicon obtained from the PCR performed. A new set of primers were designed and synthesized for the predicted gene model from KAIKObase, a different genome database of *B.mori*. This database predicted a gene of length 810bp (gene ID: BMgn012720). PCR was performed at CDFD, Hyderabad and gene was amplified at expected size.

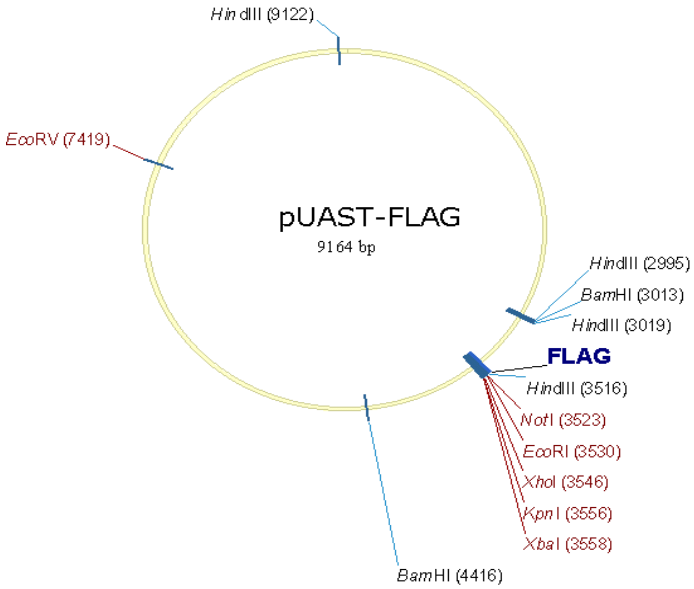


Figure 2.2: Vector map of pUAST-FLAG

To enable expression of the gene in *Drosophila*, the vector chosen is P-element based pUAST-FLAG (Figure 2.2) vector which could be regulated by GAL4. Vector pUAST-FLAG was

transformed and purified. Both vector and amplicon were digested with restriction enzyme KpnI and XbaI simultaneously (double digestion). T4 DNA ligase enzyme was used to ligate the vector and insert.

2.3 Cloning of Hedgehog full length cDNA from *Bombyx mori*:

A gene model of *hedgehog* was identified from Silkworm genome database (gene ID: BGIBMGA012535). The gene is of length 834 bp. Primers were synthesized for amplifying the *hh* gene of *Bombyx*. Full length cDNA of *Bombyx hh* was amplified from cDNA library synthesized from *Bombyx* hindwing. Thirty rounds of PCR amplification of the *hh* gene were done using Pfu polymerase. Extension was at 72 °C for 45 seconds and the annealing temperature used was 58°C. The restriction sites used were EcoRI and KpnI. Vector pUAST-FLAG (figure 2.2) was used for the ligation

2.4 Gene Ontology Analysis

Since the *Apis* and *Bombyx* genes cannot undergo GO analysis directly, they are categorized based on the functions of their *Drosophila* homologs. GO analysis were done for all the genes from *Drosophila*, *Apis* and *Bombyx.* Here, all the genes from both *Apis* and *Bombyx* which has a homolog in *Drosophila* are selected. GO analysis were done for this set of genes and for all *Drosophila* genes. The GO categories analysed in the previous study were selected and plotted.

5433 genes were selected for *Apis* and 5770 for *Bombyx,* which have homologs in *Drosophila*. GO analysis were done separately for each species. 7654 genes from *Drosophila* were identified for the analysis. Percentage of genes in each GO category were plotted for 3 species

Chapter 3

Results and discussion

3.1 Cloning of vestigial quadrant enhancer in pH-Stinger:

Amplicon of an expected size of DNA was obtained from the PCR (figure 3.1). Sequencing results confirmed the ligation of the exact sequence. Sequencing results and sequence alignment using BLAST are appended at the end of the chapter (Appendix 1). A single line of transgene was received for the vestigial quadrant enhancer. Testing of the transgenic line couldn't produce any GFP staining. This could be because of the divergence in the sequence of the *vestigial* quadrant enhancer from *Bombyx* to *Drosophila*. Ubx from *Drosophila* need to identify the enhancer sequence as a target to produce the expression of GFP.

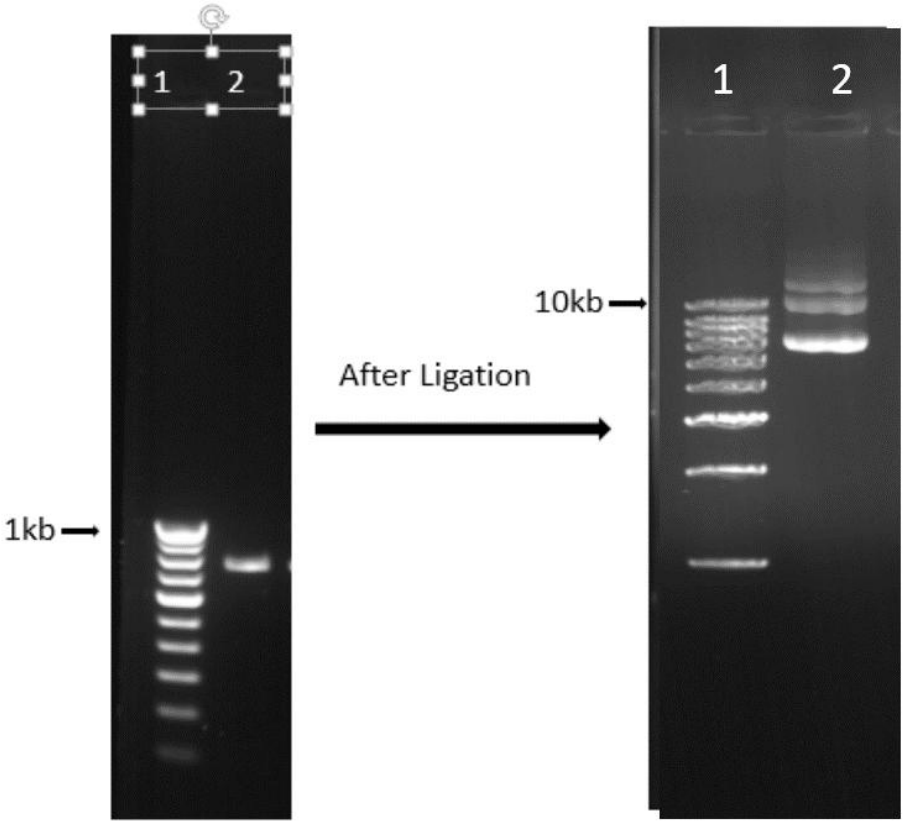


Figure 3.1: Cloning of enhancer element into pH-Stinger vector. Left panel shows the amplified vestigial enhancer from *Bombyx.* Lane 1 is 1kb ladder and lane 2 is the amplicon. Right shows the cloned construct of Enhancer element and pH-stinger vector**.**

3.2 *vestigial* Full length cDNA from *Bombyx mori*:

PCR was performed at CDFD, Hyderabad and gene was amplified at expected size. But there was also a primer dimer of size 200bp present along with the amplified gene (figure 3.2). This made the re-amplification difficult as Primer dimer was in higher concentration than the product. This was confirmed to be a primer dimer based on appropriate controls including only the primers of the reaction (Figure 3.3).

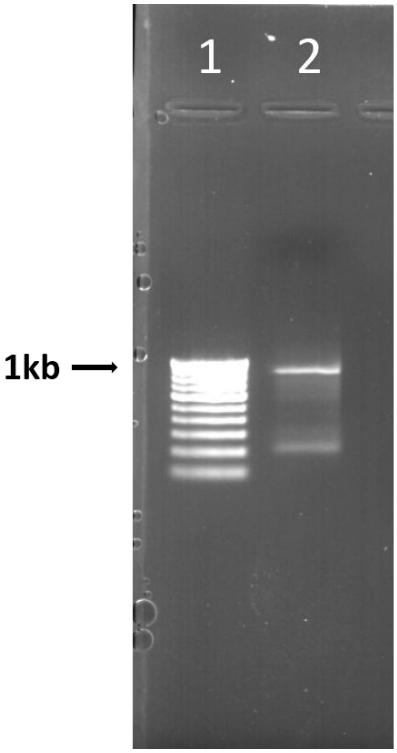


Figure 3.2: *B.mori vestigial* full length cDNA. Lane1 is 1 kb ladder and lane2 shows the Amplicon of size 810bp and primer dimer of size ~200bp.

We tried reducing this by decreasing the primer concentration in the PCR reaction. After trying various concentration of primers, 0.1 uM was finalized as the optimum concentration. Rest of the PCR reactions were carried out at this concentration. Reduced concentration of primers gave a slightly better result in PCR. To prevent the primer dimer formation, we used dimethyl sulfoxide (DMSO) as it prevents second structure formation of DNA. 5% DMSO was used to increase the yield of the reaction.

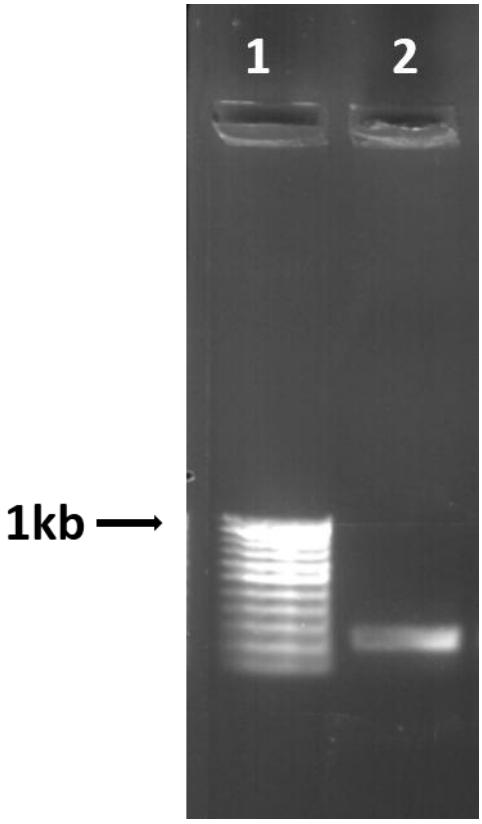


Figure 3.3: *B.mori vestigial* primer dimer. Lane1 is 1 kb ladder and lane2 shows the Primer Dimer from PCR without template. This shows that it is not an amplification.

The efficiency of heat shock transformation was less because of the huge size of the vector (9kb). So we decided to transform the ligation reaction using electro competent cells. Transformation of sample was performed at 2500V. Colonies were obtained and tested for transformed vector, but no positive colonies were found. This was repeated multiple times with different concentrations of sample, but was not successful. Hence, the problem could be in any of the earlier steps such as efficiency of restriction digestion of the vector and insert,

3.3 *hedgehog* full length cDNA from *Bombyx mori*:

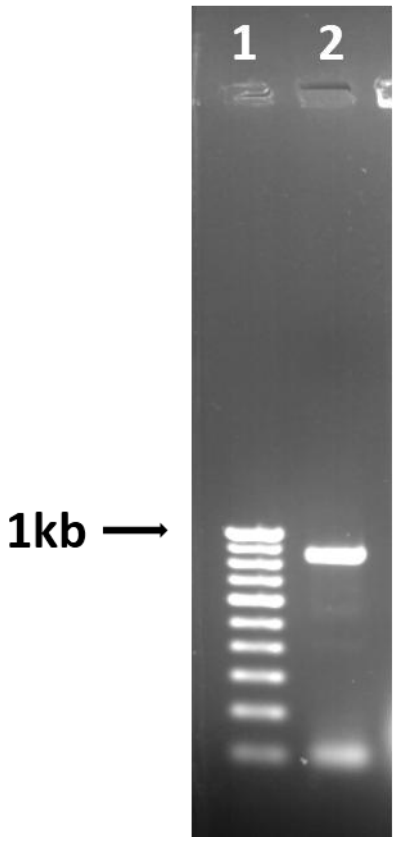


Figure 3.4: *Bombyx mori* hedgehog full length cDNA. Lane 1 is ikb ladder and lane 2 shows the amplified gene of size 834 bp.

An Amplicon of expected size was obtained from PCR (Figure 3.3). Ligation of this gene with the vector pUAST-FLAG was unsuccessful. Digestion of vector was confirmed and possible reason for failure in ligation could be inefficiency in the digestion of Insert sequences. Since the overhang of primers are small here, it reduces the efficiency of digestion of insert by enzymes. To overcome this, we increased the amount of insert DNA used in the digestion reaction. But this didn't make any difference in the end result. This was tried multiple times with various concentrations and remain unsuccessful.

3.4 Gene Ontology Analysis

From Figure 3.4 and Figure 3.5, it is visible that percentage of genes in each category is comparable. The dip in percentage of genes for *Drosophila* in figure 3.5 is accountable to the increase in total number of genes analysed.

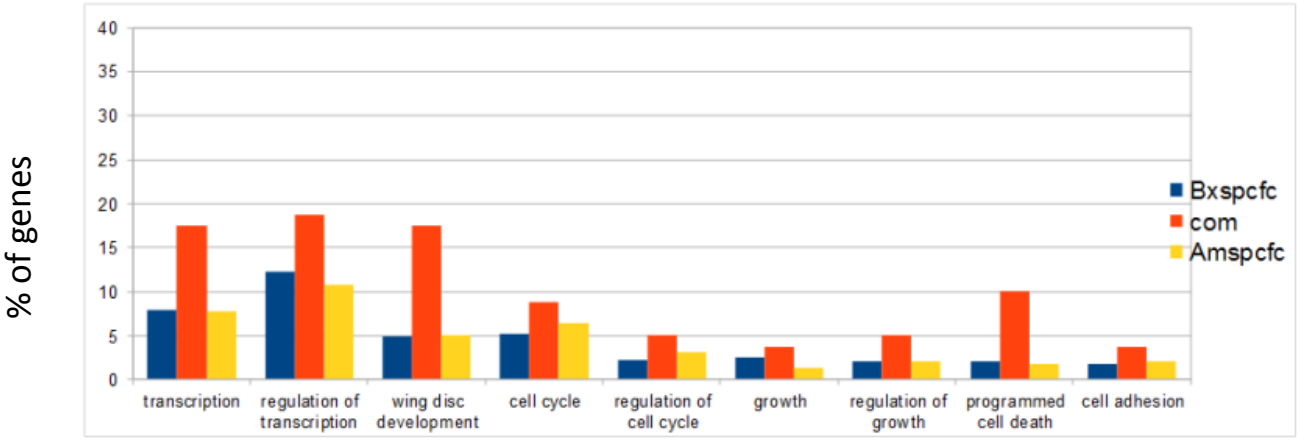


Figure 3.5: A comparative GO analysis of the targets of Ubx common to Bombyx and Apis (red) and the species-specific targets (blue: Bombyx and Yellow: Apis). (Work done by Shreeharsha)

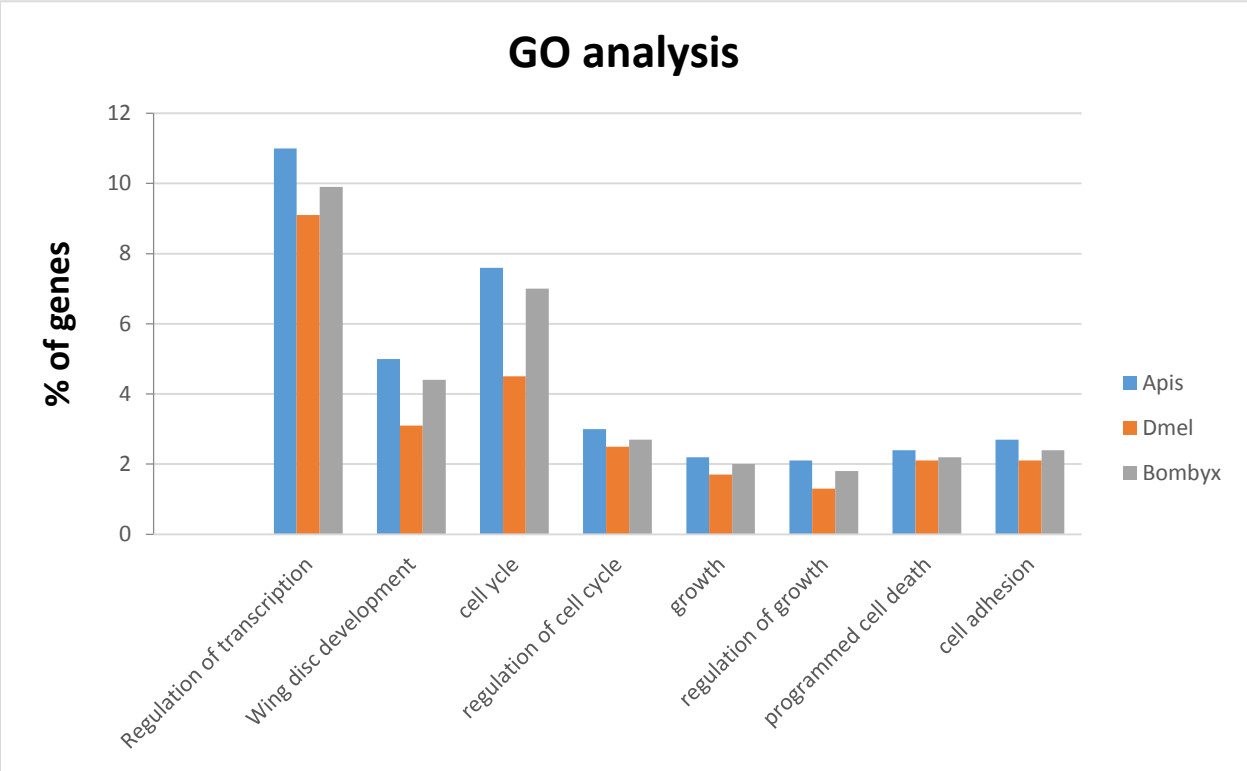


Figure 3.6: A comparative GO analysis of the genes from both *Apis* and *Bombyx* which has a homolog in *Drosophila* and all the genes from *Drosophila*.

These observations suggest that there is no specific developmental pathway/mechanism that is targeted by Ubx in Dipterans to specify haltere development. Additional analyses such as the mode of regulation of these targets by Ubx in different lineages may help to better understand evolutionary changes that have led to haltere development in T3 of Dipterans.

Conclusion

Putative enhancer sequence was cloned into pH-Stinger vector and integrated into fly genome. A transgenic *Drosophila* line expressing GFP under the quadrant enhancer of *vg* from *Bombyx* was generated. Testing of the same did not reveal any GFP staining and needs further verification. Also made attempts to make transgenic flies expressing *vg* and *hh* from *Apis* and *Bombyx* in *Drosophila.* We couldn't succeed in cloning the cDNA to pUAST-FLAG vector after many attempts. This need to be done with modifications to the standard techniques used here. Validation of gene ontology analysis of Targets of Ubx was done.

APPENDIX

Sequencing Result and Bioinformatic Analysis.

Putative *Bombyx vestigial* quadrant enhancer sequence:

nscaf3056:1450831_1451848 1018 Nucleotide Sequence:

TGTCAACTTTATTACTCGACATTACTCGGCCATGCAAAGTGAGACATTTTCACATCAGATTAATTAAATAGAATCTA
TTTAAATATTCCACATGAATTGTTGACGTTATGTTTTTAAAACGAATATGAATAGATATTTGTAATGTTGCATGTGTC
TGTGTCTACGAAACGTTTGTGGAGTTCAGCGTGTTCTCCTGTATTCGATATTCGGTGGACATTACTGTTTTAGAAAA
ACCGGCGGTCGCGAGCAAATTCTTAAGCTCCGAAGCAGGAATGCAGTCGAGCGGCTCGGCGCGCCCTTCCCGCCA
CATCACCGCATACTGTTACAGATCCTCTGATTTTTCGACTTATCTGATATTACTAAATCGACCACTAGATTGTTCTAA
GGCTTCCCCTAGACTTATCACGAAGTATGCAGACTGCTCGTTACGAAAGCTCTATGAAAGCCCTCGTGGCTGTTTG
ATTTTAATTTTCTTCAAACCAATATGAACGCACGATCATGGAATAAATTGAAACATACAGATCAAAGAAGGGGTTT
GATCCTCGAGGAGGAACTGTTTTCCTAATGAATGATGCACCTACCAACAATATTTGTTCCAACTGAATCGCTGACAT
TATTGCATTAATATTGAATATACACAATAAAAAAAAATGTGAGCCGTCACGGGACGCCTGGCAGATGTGAAACATC
GACATTATTTTGTAAAAGTAATATGCCGAAAAGAAC

Bombyx vestigial gene model from silkDB genome database:

Gene ID: **BGIBMGA012721**

Nucleotide sequence:

ATGGAGAGGTGTCGTGTGTTGGTACAGTTGAACGTGCAGAGCGGCGGCGTCAATGCGGGCAGCAGCG
AGGCGTTCGTGTCGGGGGCGCAGTCCCCGGGCTCTCCGCCGCACGCGCTCCCGGCGCACCGTCTGCGC
ACCAAGGAGGAGGATCTCTCTGCACACGGCCGAGCCAGCGCTGACGGAGGCGGTTCGTCCGAGTCGGA
GGGCGAGTGCGGCGGCGCGGGTCCGCGAGCGCGCGCGCAGTACGTCAGCGCCAACTGCGTCGTGTTC
ACGCACTACTCCGGGGACGTCGCCGCGGTCGTCGACGAGCACTTCGCCAGGGCGCTCTCGCTCGACAAG
ACTAAAGACGGCGTCCCGATGGTGTCGCGCAACCTGCCGGCGTCGTTCTTCAACGCGGCGGCGGTGGG
CGGCGCGGGCGGCGTGGAGCTGTACGAGTACGACCCGTGGCACCAGCACTACTCCGGCTACGGACACG
CGGCGCACCGCCACGCCGCCGAGTACCACGCGGCCGCCCAGGCGCACCACAACATGGCGGCGGCGGC
GGGCTACGGGCCGCTGCTGCTGCCCCGCGGCTCTCTGCACGCGCAGTACAAGCCGGTGGAGTGGGGCG
CACAGCACCACGCGGCGCACCACCTCGACCCCGCCGCGTGCTCGCCCTACTCTTATCCCGCAGTACCAGA
ACTAGTATGTTTGGATAGTTCTGTAAGACGGTAA

Bombyx vestigial gene model from KAIKObase:

Gene ID: BMgn012720:

Nucleotide sequence:

ATGGCGGTGAGCTGCCCCGAGGTGATGTACGGCGCCTACTATCCCTACCTGTACGGCCGCGGAGCCGC
GCGCTCCTTCCACCACGCGCCGCACTTCCAGTACGATCGGTTGAACGTGCAGAGCGGCGGCGTCAATGC
GGGCAGCAGCGAGGCGTTCGTGTCGGGGGCGCAGTCCCCGGGCTCTCCGCCGCACGCGCTCCCGGCGC
ACCGTCTGCGCACCAAGGAGGAGGATCTCTCTGCACACGGCCGAGCCAGCGCTGACGGAGGCGGTTCG
TCCGAGTCGGAGGGCGAGTGCGGCGGCGCGGGTCCGCGAGCGCGCGCGCAGTACGTCAGCGCCAACT
GCGTCGTGTTCACGCACTACTCCGGGGACGTCGCCGCGGTCGTCGACGAGCACTTCGCCAGGGCGCTCT
CGCTCGACAAGACTAAAGACGGCGTCCCGATGGTGTCGCGCAACCTGCCGGCGTCGTTCTTCAACGCGG
CGGCGGTGGGCGGCGCGGGCGGCGTGGAGCTGTACGAGTACGACCCGTGGCACCAGCACTACTCCGG
CTACGGACACGCGGCGCACCGCCACGCCGCCGAGTACCACGCGGCCGCCCAGGCGCACCACAACATGG
CGGCGGCGGCGGGCTACGGGCCGCTGCTGCTGCCCCGCGGCTCTCTGCACGCGCAGTACAAGCCGGTG
GAGTGGGGCGCACAGCACCACGCGGCGCACCACCTCGACCCCGCCGCGTGCTCGCCCTACTCTTATCCC
GCAGTACCAGGTCTGGAAGCGCAGGTGCAGGACACGTCCAAGGACATGTACTGGTTCTGA

Bombyx hedgehog gene model from silkworm genome database:

Gene ID: **BGIBMGA012535**

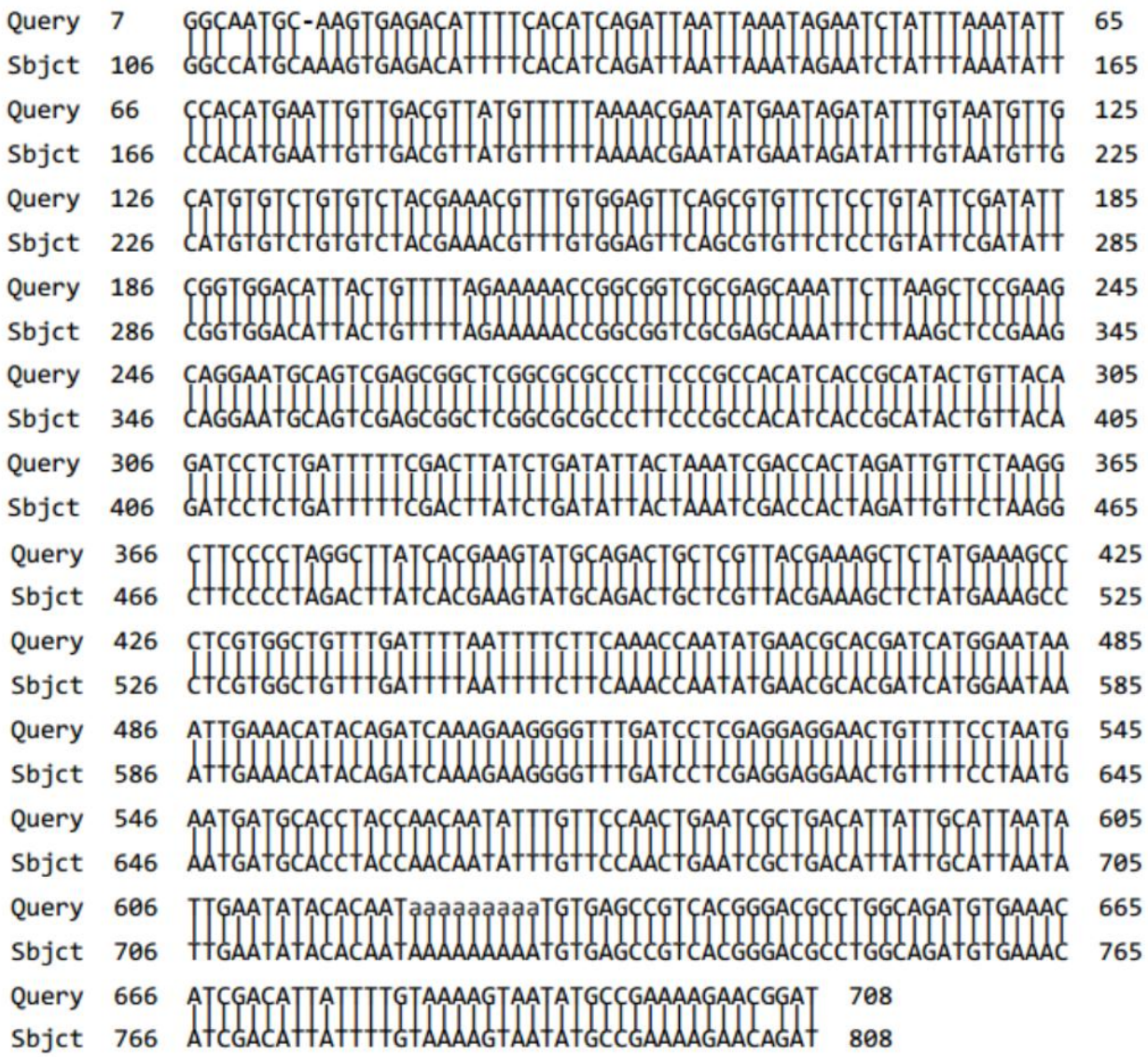
Nucleotide Sequence:

AATGAACCAGTGGCCGGGAGTGCGGCTCCGTGTGATCGAAGGTTGGGACGAAGAGAATAATCATCTAGACAAGTCTC
TCCACTACGAGGGTCGGGCTGTGGACTTAACCACCAGTGACCGAGACCGCAGCAAATACGGTATGCTGGCCCGCCTC
GCTGTCGAAGCCGGCTTTGACTGGGTCTTCTATGAAAGCCGGTCTTACATTCATTGTTCTGTTAAAACAGAATCATC
TGTGGGAACTGGTGCTGGCTGTTTCCCTGGCGGCTCCGTAGTTCACACTGAGAAAGGACCGAAGGATATAGCTGCCT
TGCAGAAGGGAGACAAGGTCTTAGCAGCTGATGATAATGGCATGATGATTTATTCTAAGGTGCTAACATTCATCGAT
CGTGATCCTAATGCAACTCGACAGTTCATCGAGATAACTGCGGAGAATGGTGTTGCGATCACAACGACGCCCTCACA
TTTGCTCCTTCTCGCTGCAGCTGACGGTTGGCGCGATGTCTTTGCAGCTAACATCAAAGAGGGAGACGTACTATTGA
CAAGAGGTCGAAGTGATGTTATGAGGCCCTCGAGAGTAACGAGAACTCGTATCCTCACGAAACGAGGTGTTTTTGCC
CCGCTTACCAAAGCCGGCACTATCATAGTGGATGACGTTCTGGCGTCATGCTACGCGCTAGTACGCAGTCATTCGCT
TGCGCATGTTGCGATGGCACCATTGCGATGGCTTTCCAGTTGGAGCACTTCTTCAGAAACGACCAAAGGCGTCCACT
GGTACGCGAGCGCGCTTTACTCTTTCGGCGACTACGTATTACCAGCATCGTATAGATATCGATAA

Sequencing result of Cloned *Bombyx vg* quadrant enhancer

Bombyx vg Q enhancer
Sequence ID: lcl|Query_245237 Length: 1018 Number of Matches: 1
Range 1: 106 to 808

Score	Expect	Identities	Gaps	Strand	Frame
1275 bits (690)	0.0()	699/703(99%)	1/703(0%)	Plus/Plus	


Query 7 GGCAATGC-AAGTGAGACATTTTCACATCAGATTAATTAAATAGAATCTATTTAAATATT 65
Sbjct 106 GGCCATGCAAAGTGAGACATTTTCACATCAGATTAATTAAATAGAATCTATTTAAATATT 165
Query 66 CCACATGAATTGTTGACGTTATGTTTTTAAAACGAATATGAATAGATATTTGTAATGTTG 125
Sbjct 166 CCACATGAATTGTTGACGTTATGTTTTTAAAACGAATATGAATAGATATTTGTAATGTTG 225
Query 126 CATGTGTCTGTGTCTACGAAACGTTTGTGGAGTTCAGCGTGTTCTCCTGTATTCGATATT 185
Sbjct 226 CATGTGTCTGTGTCTACGAAACGTTTGTGGAGTTCAGCGTGTTCTCCTGTATTCGATATT 285
Query 186 CGGTGGACATTACTGTTTTAGAAAAACCGGCGGTCGCGAGCAAATTCTTAAGCTCCGAAG 245
Sbjct 286 CGGTGGACATTACTGTTTTAGAAAAACCGGCGGTCGCGAGCAAATTCTTAAGCTCCGAAG 345
Query 246 CAGGAATGCAGTCGAGCGGCTCGGCGCGCCCTTCCCGCCACATCACCGCATACTGTTACA 305
Sbjct 346 CAGGAATGCAGTCGAGCGGCTCGGCGCGCCCTTCCCGCCACATCACCGCATACTGTTACA 405
Query 306 GATCCTCTGATTTTTCGACTTATCTGATATTACTAAATCGACCACTAGATTGTTCTAAGG 365
Sbjct 406 GATCCTCTGATTTTTCGACTTATCTGATATTACTAAATCGACCACTAGATTGTTCTAAGG 465
Query 366 CTTCCCCTAGGCTTATCACGAAGTATGCAGACTGCTCGTTACGAAAGCTCTATGAAAGCC 425
Sbjct 466 CTTCCCCTAGACTTATCACGAAGTATGCAGACTGCTCGTTACGAAAGCTCTATGAAAGCC 525
Query 426 CTCGTGGCTGTTTGATTTTAATTTTCTTCAAACCAATATGAACGCACGATCATGGAATAA 485
Sbjct 526 CTCGTGGCTGTTTGATTTTAATTTTCTTCAAACCAATATGAACGCACGATCATGGAATAA 585
Query 486 ATTGAAACATACAGATCAAAGAAGGGGTTTGATCCTCGAGGAGGAACTGTTTTCCTAATG 545
Sbjct 586 ATTGAAACATACAGATCAAAGAAGGGGTTTGATCCTCGAGGAGGAACTGTTTTCCTAATG 645
Query 546 AATGATGCACCTACCAACAATATTTGTTCCAACTGAATCGCTGACATTATTGCATTAATA 605
Sbjct 646 AATGATGCACCTACCAACAATATTTGTTCCAACTGAATCGCTGACATTATTGCATTAATA 705
Query 606 TTGAATATACACAATaaaaaaaaaTGTGAGCCGTCACGGGACGCCTGGCAGATGTGAAAC 665
Sbjct 706 TTGAATATACACAATAAAAAAAAAATGTGAGCCGTCACGGGACGCCTGGCAGATGTGAAAC 765
Query 666 ATCGACATTATTTTGTAAAAGTAATATGCCGAAAAGAACGGAT 708
Sbjct 766 ATCGACATTATTTTGTAAAAGTAATATGCCGAAAAGAACAGAT 808

References

1. Agrawal, P., Habib, F., Yelagandula, R., and Shashidhara, L.S. (2011). Genomelevel identification of targets of Hox protein Ultrabithorax in Drosophila: novel mechanisms for target selection. Sci Rep *1*.
2. 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.
3. 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.
4. 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.
5. 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.
6. 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.
7. 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.
8. Choo SW, White R and Russell S . (2011).Genome wide analysis of the binding of the Hox protein Ultrabithorax and the Hox co factor Homothorax in Drosophila . Plos One 6(4),1-14
9. Duboule, D. (1998). Vertebrate Hox gene regulation: clustering and/or colinearity? Current Opinion in Genetics & Development *8*, 514-518.
10. Lewis, E. B. (1963). Genes and developmental pathways. American Zoologist, 33-56.

11. 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.
12. Mita, K., Kasahara, M., Sasaki, S., Nagayasu, Y., Yamada, T., Kanamori, H., Abe, H. (2004). The Genome Sequence of Silkworm , Bombyx mori. Genome Biology, 35, 27–35.
13. 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.
14. Sanchez-Herrero, E., Vernos, I., Marco, R., and Morata, G. (1985). Genetic organization of Drosophila bithorax complex. Nature *313*, 108-113.
15. Sadava, D.et al. Life; the science of Biology, 8th ed.
16. 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.
17. Shimomura, M., Minami, H., Suetsugu, Y., Ohyanagi, H., Satoh, C., Antonio, B., Mita, K. (2004). KAIKObase , An integrated silkworm genome database and data mining tool, BMC Genomics. 8, 1–8.
18. 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.
19. 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.
20. Xia et al. (2004). Draft Sequence for the Genome of the Domesticated Silkworm (Bombyx mori). Science 306,1937-40.
21. Xu, T., and Rubin, G.M. (1993). Analysis of genetic mosaics in developing and adult Drosophila tissues. Development *117*, 1223-1237.