A micro-evolutionary change in target binding sites as a key determinant of Ultrabithorax function in *Drosophila*

Soumen Khan[1#], Saurabh J. Pradhan[1], Guillaume Giraud[2], Françoise Bleicher[2], Rachel Paul[2], Samir Merabet[2], LS Shashidhara[1,3#]

[1] Indian Institute of Science Education and Research (IISER), 411008, Pune, India

[2] IGFL, ENS Lyon, UMR5242, 32 Av. Tony Garnier, 69007 Lyon, France

[3] Ashoka University, Sonipat, 131029, Haryana, India

Correspondence: ls.shashidhara@iiserpune.ac.in
soumen.khan@students.iiserpune.ac.in

Keywords: Ultrabithorax, ChIP-seq, high affinity binding sites, molecular evolution

Abstract

Hox genes encode Homeodomain-containing transcription factors, which specify segmental identities along the anterior-posterior axis. Functional changes in Hox genes have been directly implicated in the evolution of body plans across the metazoan lineage. The Hox protein Ultrabithorax (Ubx) is expressed and required in developing third thoracic (T3) segments in holometabolous insects studied so far, particularly, of the order Coleoptera, Lepidoptera and Diptera. Ubx function is key to specify differential development of the second (T2) and T3 thoracic segments in these insects. While Ubx is expressed in the third thoracic segment in developing larvae of Hymenopteran *Apis mellifera*, the morphological differences between T2 and T3 are subtle. To identify evolutionary changes that are behind the differential function of Ubx in these two insects, which are diverged for more than 350 million years, we performed comparative analyses of genome wide Ubx-binding sites between *Drosophila* and *Apis*. Our studies reveal that a motif with a TAAAT core is a preferred binding site for Ubx in *Drosophila*, but not in *Apis*. Biochemical and transgenic assays suggest that in *Drosophila*, TAAAT core sequence in the Ubx binding sites is required for Ubx-mediated regulation of two of its target genes studied here. *CG13222*, a gene that is normally upregulated by Ubx and *vestigial* (*vg*), whose expression is repressed by Ubx in T3. Interestingly, changing the TAAT site to a TAAAT site was sufficient to bring an otherwise unresponsive enhancer of the *vg* gene from *Apis* under the control of Ubx in a *Drosophila* transgenic assay. Taken together, our results suggest an evolutionary mechanism by which

critical wing patterning genes might have come under the regulation of Ubx in the Dipteran lineage.

Introduction

The diversification of morphology and emergent function of appendages have been important features of animal evolution. This is best exemplified in the diversity that is seen in closely related species of Arthropoda. In Bilaterians, first discovered in *Drosophila melanogaster*, *Hox* genes regulate the development of serially homologous structures along the anterior posterior axis (Carroll 1995; Lewis 1978; McGinnis and Krumlauf 1992). Subsequently, large number of studies have shown that variations in Hox gene functions, due to changes in their (i) copy number, (ii) primary structure at protein levels, (iii) their expression patterns and (iv) regulation of downstream target genes, are critical for morphological diversification.

In *Drosophila melanogaster*, the Hox protein Ultrabithorax (Ubx) specifies the development of halteres in the third thoracic (T3) segment (for detailed review see Lewis, 1978 and Khan et al., 2020). Loss of function mutations of Ubx give rise to four winged flies with complete duplication of the T2 segment in the place of T3, while overexpression of Ubx in larval wing discs leads to wing to haltere transformations (Lewis, 1978; Castelli-Gair et al., 1990; Cabrera et al., 1985; White and Akam, 1985; White and Wilcox, 1985). While Ubx is expressed in hindwing primordia of different insect species (Carroll 1995; Prasad, Tarikere et al. 2016), the outcome of its expression is not the same in all insect groups. In the Coleopteran *Tribolium castaneum* and the Lepidopteran *Precis coenia* and *Bombyx mori* Ubx is expressed only in T3 segment as in Dipteran *Drosophila*. Functional studies have shown that Ubx is required for the suppression of elytra and the specification of hindwings in the T3 of Tribolium (Tomoyasu et al. 2005), it is required to specify differences in eyespot patterns between the hindwings and forewings in *Precis* (Weatherbee et al. 1999; Matsuoka and Monteiro 2021). However, in Hymenopterans such as *Apis mellifera,* wherein T2 and T3 are distinguished by a marginally smaller hindwing, Ubx is expressed in both forewing and hindwing primordia (Prasad, Tarikere et al. 2016). The expression, however, is stronger in the hindwing primordia (Prasad, Tarikere et al. 2016). Over-expression of Ubx derived from *A. mellifera or B. mori* or *T. castaneum* can suppress wing development and specify haltere fate in transgenic *D. melanogaster* (Prasad, Tarikere et al. 2016), suggesting that changes in Ubx at the protein

level may either have not contributed or marginally contributed to the evolution of its function. A study comparing genome-wide targets of Ubx in developing halteres of *D. melanogaster* and developing hindwings of *A. mellifera* and *B. mori* revealed that a large number of genes have remained common targets of Ubx in these three species (Prasad, Tarikere et al. 2016), which are diverged for nearly 350 million years. Among those common targets, a few wing patterning genes are differentially expressed in *D. melanogaster*, but not in *A. mellifera* or *B. mori* (Prasad, Tarikere et al. 2016).

To identify evolutionary changes in the regulatory sequences of targets of Ubx that may have brought some of the wing patterning genes under the regulation of Ubx, we performed comparative analyses of genome wide Ubx-binding motifs between *D. melanogaster* and *A. mellifera*. We find that while in *Drosophila*, a motif with a core TAAAT sequence (MATAAATCAY) (hereafter referred to as TAAAT motif) was enriched by Ubx in ChIP pulled-down sequences, the same was not observed in *A. mellifera*. Using a combination of in-vitro, cell culture and transgenic assays, our results reveal that the TAAAT motif is important (and bears greater relevance as compared to the TAAT motif) for the Ubx mediated-regulation of its target gene, CG13222, in *Drosophila*. We then analyzed the significance of the TAAAT motif in the Ubx mediated down-regulation of a pro-wing gene, *vestigia*l, in *Drosophila* as against the TAAT motif in its orthologue in *Apis*. In transgenic *D. melanogaster,* the enhancer of *vg* of *A. mellifera* drives reporter gene expression in both wing and halteres. We show that changing the TAAT motif to TAAAT motif was sufficient to bring this enhancer of *vg* from *Apis* under the regulation of Ubx. Taken together, our study suggests that evolutionary changes in the regulatory sequences in the targets of Ubx in dipteran lineages, such as from TAAT to TAAAT as core sequence in Ubx-binding motifs, may have been an important mechanism contributing to the evolution of wing to haltere appendages.

Results

Motif with a TAAAT core sequence is enriched for direct targets of Ubx in developing halteres in *D. melanogaster* but not for developing hindwings in *A. mellifera*

Previous studies indicate that despite the differences in the hindwing morphology, the Ubx protein is expressed in the third thoracic segment of both insects (as well as other insect lineages) (Fig. 1A) (Carroll 1995; Prasad, Tarikere et al. 2016). Additionally, Ubx is highly

97 conserved across *A. mellifera, B. mori, T. castaneum, Junonia coenia* and *D. melanogaster* at
98 the level of its DNA-binding homeodomain and the protein interaction domains, YPWM and
99 UbdA (Fig. S1 A). We have previously shown that overexpression of Ubx from different insect
100 species in the wing imaginal disc of transgenic *Drosophila* resulted in wing to haltere
101 transformation in a pattern similar to *Drosophila* Ubx (Prasad, Tarikere et al. 2016). A triple
102 mutant combination of alleles, *abx bx3* pbx, results in complete transformation of T3 to T2
103 (Lewis 1978). This four-wing phenotype is rescued when Ubx is over-expressed exogenously
104 during development using a GAL4-UAS system. This phenotype can also be rescued by
105 expressing Ubx derived from *Apis* or *Bombyx* (Merabet, S; unpublished results). This suggests
106 that differences in insect hindwing morphology is a consequence of events occurring
107 downstream of or in parallel to Ubx. Comparison of targets of Ubx between *D. melanogaster*,
108 and *A. mellifera* suggest that many functionally important genes are common to the two
109 lineages, although many key genes are differentially expressed between T2 and T3 in
110 *Drosophila*, but in *Apis* (Prasad, Tarikere et al., 2016; Fig. S1 B). This reconfirms that
111 morphological differences may lie at the levels of regulation of those targets.

112 We carried out ChIP-Seq for Ubx from the third instar haltere discs (GSE205177). This
113 method provided better resolution over earlier methods used by us (Agrawal et al., 2011) and
114 others (Choo et al., 2011; Slattery et al., 2011) to identify Ubx-specific binding motifs. We then
115 performed a comparative analysis of Ubx ChIP-Seq data generated from *D. melanogaster*
116 haltere imaginal discs to Ubx ChIP-seq data generated from *A. mellifera* hindwing discs
117 (reported in Prasad Tarikere et al., 2016).

118 De-novo motif analysis of ChIP-Seq data from *D. melanogaster* showed significant
119 enrichment for a motif with a TAAAT core sequence (MATAAATCAY) (henceforth referred to
120 as the TAAAT motif) as putative Ubx-binding site (Fig 1B). We observed that enrichment for
121 TAAAT motif is only in *Drosophila* and not in *Apis*, while both the ChIP-Seq datasets showed
122 enrichment for binding sites for other transcription factors such as GAGA, CTCF and Brk (Fig.
123 1B). We calculated the frequency of the TAAAT motif in the *Drosophila* dataset and found it
124 to be 1.7X in Ubx-pulled down sequences as compared to its prevalence in the whole genome.
125 Interestingly, the frequency of the TAAAT motif in the entire genome was 1.8 fold higher in
126 *A. mellifera* as compared to *D. melanogaster* (Fig. 1C), suggesting that the TAAAT motif is
127 selectively enriched in the Ubx bound regions only in *D. melanogaster* and not in A. *mellifera*.

128 We also observe, similar to previous results (Agrawal et al. 2011), that the canonical
129 TAAT motif is not enriched in Ubx-pulled down sequences, rather it was found under-
130 represented compared to its prevalence in the whole genome (Fig. 1C). This was found to be
131 true for both *Drosophila* and *Apis* datasets.
132
133 **The TAAAT motif is critical for Ubx mediated regulation of a target gene in *Drosophila***
134 **halteres**
135 Our results suggest that the TAAAT motif is enriched in Ubx-pulled down sequences from
136 haltere imaginal discs of *Drosophila*, but not from developing hindwings of *A mellifera*. Earlier
137 studies have reported that the TAAAT motif is bound by the Ubx-Extradenticle-Homothorax
138 complex (Slattery et al. 2011; Sánchez-Higueras et al. 2019). However, neither Extradenticle
139 (Exd) nor Homothorax (Hth) are expressed or required for the development of the haltere
140 capitellum (Casares and Mann 2000). In this context, we sought to determine the functional
141 significance of Ubx-binding to the TAAAT motif during haltere development.
142 We first compared Ubx ChIP-seq data from *Drosophila* halteres to RNA seq data
143 (GSE205352) for *Drosophila* wing and haltere imaginal discs to identify those direct targets
144 that are upregulated or downregulated in the haltere. We then compared the frequency of
145 the TAAAT motif in Ubx bound response elements of putative direct targets that are
146 upregulated or downregulated between wing and haltere discs. We found the frequency of
147 the TAAAT motif in Ubx bound response elements in the upregulated category is marginally
148 higher (1.4 times) as compared to the downregulated category (Table S2) . This suggests that
149 while the TAAAT motif may be preferred site for Ubx to bind on the enhancers of its targets,
150 Ubx may use this motif to regulate both upregulated and downregulated targets in the
151 haltere.
152 We next evaluated the affinity and functional importance of Ubx binding to the TAAAT
153 motif in one of the well characterized targets, *CG13222*, which is upregulated by Ubx in
154 developing halteres (Mohit et al. 2006; Hersh et al. 2007). The "*edge*" enhancer of *CG13222*
155 has two TAAT motifs at 17 bp apart (CG13222_WT, Fig 2A). Hersh et al. (2007) have named
156 the two motifs as site 1 and site 2. In their transgenic assays, site1 appears to be critical for
157 Ubx-mediated regulation of *CG13222*, while site2 is dispensable. Interestingly, the site1 in the
158 *edge* enhancer, which is critical for its regulation by Ubx, has a TAAAT motif which overlaps
159 with the TAAT motif (Fig. 2A). Their analysis of site1 involved mutating both TAAT and TAAAT

160 motifs, which leads to loss of enhancer driven reporter gene expression (Hersh et al. 2007).
161 We mutated site1 in such a way that the mutant enhancers carry either the TAAAT motif
162 (CG13222_M1_A) or the TAAT motif (CG13222_M1_B) (Fig. 2A).

163 To examine the relative affinities of Ubx binding to the TAAAT and TAAT sites by
164 Electrophoretic Mobility Shift Assay (EMSA), we used a 21bp long probe covering site 1 of
165 *edge* enhancer of *CG13222* (nucleotide stretch in the grey box in Fig. 2A). We find that Ubx
166 can bind to the wild type probe derived from the *CG13222 edge* enhancer. Mutations in the
167 TAAAT motif in site1 of *CG13222* lead to a substantial loss of this binding (probes
168 CG13222_M1_A, Fig. 2B). Conversely, mutating the TAAT motif alone had no significant
169 effects on binding of Ubx to the probe (CG13222_M1_B, Fig S2 A).

170 We also observed that another Hox protein, Scr, was unable to bind to the wildtype
171 probe *CG13222*. On the other hand, both Ubx and Scr bind to a 29bp probe derived from the
172 *sal* enhancer (Galant et al. 2002) containing TAAT sites (Fig 2B). Perhaps, the TAAAT motif
173 provides higher specificity of binding for Ubx as compared to other Hox proteins.

174 Next we examined the importance of the TAAAT motif in gene regulation using
175 Luciferase reporter assays in S2 cells and haltere imaginal discs in transgenic *Drosophila*. For
176 these functional assays, we used a 512bp long region spanning the *edge* enhancer of *CG13222*
177 as reported by Hersh et al. (2007), which has both site1 and site2. Consistent with *CG13222*
178 being a target of Ubx upregulated during haltere development, in the presence of Ubx a
179 luciferase expression construct driven by the enhancer of the gene was significantly
180 upregulated (2.8 folds) in S2 cells (Fig. 2C). Using a similar experimental design, we observed
181 a significant decrease in Ubx-dependent expression when TAAAT alone is mutated (construct
182 CG13222_M1_A in Fig. 2C) compared to the wildtype *CG13222* enhancer.

183 We next generated transgenic strains of *D. melanogaster* carrying the wild type or
184 mutant enhancers (as explained above) cloned upstream of a GFP reporter. The wildtype
185 *CG13222* enhancer showed GFP expression at the posterior edge of the haltere imaginal discs
186 consistent with earlier reports (Fig. 2D) (Mohit et al. 2006; Hersh et al. 2007). We first
187 replicated the mutation described in Hersh et al. 2007 wherein both the TAAT and TAAAT
188 motifs are mutated (CG13222_M1) and found that the reporter gene expression is
189 significantly reduced when driven by the mutant enhancer (Fig. 2E). We then analysed the
190 CG13222_M1_A mutant where the TAAAT motif is specifically mutated while the TAAT motif
191 is maintained intact and observed a significant loss of reporter expression (Fig. 2F, 2H).

192 However, the pattern of GFP expression patterns driven by CG13222_M1 and
193 CG13222_M1_A enhancers were similar, suggesting that the TAAAT motif is specifically
194 required for Ubx-mediated activation of the *CG13222* enhancer in *Drosophila* halteres.
195 As described earlier, the *edge* enhancer of *CG13222* has a second TAAT motif (site2 in
196 Fig 2A), which is redundant for Ubx-mediated upregulation of *CG13222* in the haltere discs
197 (Hersh et al. 2007). We also attempted to examine the effect of mutating the TAAT motif to
198 a TAAAT motif (in site2, construct CG13222_M2_A), which results in two high-affinity binding
199 motifs in the *edge* enhancer of *CG13222*. We observed a significant reduction in the luciferase
200 activity (Fig. S2 B), perhaps, due to negative effect of two high-affinity binding sites for Ubx in
201 close proximity. However, the *Drosophila* transgenic line carrying the same mutation showed
202 significantly stronger expression of the GFP reporter (Fig 2G, 2H) compared to the wildtype
203 *CG13222* enhancer. Enhancer driven expression was also observed in ectopic regions in the
204 posterior compartment (red arrows in Fig. 2G), suggesting that the presence of an additional
205 TAAAT motif may bring the otherwise unresponsive site2 under Ubx regulation.
206
207 **The TAAAT motif causes Ubx mediated repression of an orthologous enhancer of the**
208 ***vestigial* gene from *A. mellifera* in a transgenic *Drosophila* assay**
209 We had previously reported that several wing patterning genes such as the *vestigial* (*vg*) gene
210 are expressed in both fore- and hindwings primordia in *A. mellifera*. On the contrary, in
211 *Drosophila*, *vg* is a direct target of Ubx (Suppl. 2C) and is downregulated in the developing
212 halteres (Galant et al. 2002; Hersh and Carroll 2005). Of the two well characterized enhancers,
213 the quadrant enhancer of *vg* (henceforth named as *quad-vg*) is differentially expressed
214 between the wing and haltere imaginal discs. In a transgenic *Drosophila* assay, an enhancer
215 of *vg* from *Apis* (henceforth named as *Apis-vg*) equivalent to *quad-vg* showed similar patterns
216 and levels of expression between wing and haltere discs (Prasad, Tariekre et al. 2016),
217 suggesting that *Drosophila* Ubx too cannot repress the *Apis-vg* enhancer. We evaluated the
218 role of TAAAT in *Drosophila* as against TAAT in *Apis* in the regulation of the *quad-vg* and *Apis-*
219 *vg* enhancers in transgenic assays.
220 We first scanned the entire 850bp sequence of the *quad-vg* enhancer and the 575bp
221 sequence of the *Apis-vg* enhancer for Ubx binding motifs, TAAAT and TAAT. We observed a
222 25bp cassette containing both TAAT and TAAAT motifs in the *quad-vg* enhancer (Fig 3A). In
223 the Apis-vg enhancer, we found only a single TAAT motif and did not find any TAAAT motif

224 (Fig 3A). In an attempt to test the significance of this difference in Ubx binding motifs between
225 the two enhancers, we generated several *Drosophila* transgenics carrying the wild-type and
226 mutated *quad-vg* and *Apis-vg* constructs (mutations in TAAT or in TAAAT motifs) upstream of
227 a GFP reporter (Fig 3A, Suppl. Fig 2D, Suppl. Fig 3A). As reported earlier, we observed that the
228 wild type *quad-vg* enhancer drives expression of the GFP reporter in the wing imaginal discs,
229 albeit at much lower levels (compared to earlier reported *quad-vg* -lacZ reporter), but not in
230 the haltere imaginal discs (Suppl. Fig 2E). However, owing to complete loss of enhancer
231 readout in both wing and haltere imaginal discs in the transgenic carrying mutations in the
232 TAAAT motif of the *quad-vg* enhancer (Suppl. Fig 2E), we were unable to conclude the role
233 of the TAAAT motif or the TAAT motif in Ubx mediated regulation of the *quad-vg* enhancer.
234 We next analysed the reporter GFP expression in wing and haltere imaginal discs of
235 *Drosophila* transgenics carrying the wild type and mutated *Apis-vg* enhancers. As reported
236 earlier (Prasad, Tarikere et al., 2016), we observed similar expression of GFP in both wing and
237 haltere imaginal discs (Fig. 3B). Interestingly, in a chimeric enhancer (*Apis-vg_M2*), where the
238 25bp Ubx-binding cassette from the *Drosophila quad-vg* enhancer replaced its counterpart in
239 the *Apis-vg* enhancer (Suppl. Fig. 3A (iii)), we observed differential expression of the GFP
240 reporter between wing and haltere imaginal discs (Suppl. Fig. 3B). Its expression was much
241 lower in the haltere pouch compared to the GFP reporter driven by the wildtype *Apis-vg*
242 enhancer. The loss of reporter GFP expression in the pouch region was consistent across all
243 mutant *Apis-vg* constructs which bore a TAAAT motif but was not observed in constructs
244 where the TAAAT motif was absent (Suppl Fig 3A (vi), 3B). More importantly, changing only
245 the TAAT motif of the *Apis-vg* enhancer to a TAAAT motif (*Apis-vg_M1*), was sufficient for the
246 repression of the GFP reporter in the pouch (Fig. 3C, 3D,Suppl. Fig 3B) of haltere discs without
247 affecting its expression in the wing pouch (Suppl. Fig 3C).
248 Taken together, our results reveal that a microevolutionary change in the enhancer
249 sequence, specifically modifying the TAAT motif to a TAAAT motif, may have brought certain
250 critical wing-patterning genes, such as the pro-wing gene *vg*, under the regulation of Ubx
251 during the evolution of dipterans.
252
253 **Discussion**
254

255 Genetics, cell and molecular biology methods have established, unequivocally, the role of Hox
256 genes as master control genes in regulating segment-specific developmental pathways and
257 diversification of body plans during evolution. However, precise mechanism by which
258 orthologous Hox proteins mediate differential development of a specific morphological
259 feature in different species is still largely unknown. Main reasons this has eluded systematic
260 study include their highly conserved DNA-binding domains across all species.

261 In *Drosophila*, the Ubx protein is expressed in the T3 segment and specifies the
262 development of the haltere by activating or repressing a number of genes at various
263 hierarchical levels of wing patterning (Weatherbee et al. 1998; Shashidhara et al. 1999). Ubx
264 is also expressed in the T3 segment of other insect species and thought to specify different
265 fates in each of them. For example, in Coleopterans, Ubx specifies the development of
266 hindwings instead of elytra and in Lepidopterans it specifies differences in eyespot patterns
267 between the forewing and hindwings. In more ancestral Hymenopterans such as *Apis*
268 *mellifera*, which have two pairs of wings, hindwings are marginally smaller than the
269 forewings. In A. mellifera, Ubx is expressed in both forewing and hindwing primordia,
270 although its expression is stronger in T3 (Prasad, Tarikere et al. 2016). It is possible that
271 divergence of Coleopterans, Lepidopterans and Dipterans involve both suppression of Ubx
272 expression in T2 and Ubx acquiring ability to regulate various developmental pathways in T3.

273 To identify the mechanisms governing the differential regulation of wing patterning
274 genes between T2 and T3 in *Drosophila*, we employed a comparative genomics approach and
275 identified that the TAAAT motif is enriched by Ubx specifically in the enhancers of targets of
276 Ubx in *Drosophila* halteres but not in *Apis* hindwings. Our functional studies reported here
277 suggest that Ubx binds to the TAAAT motif with higher affinity than to the TAAT motif and its
278 binding to former is critical for the up-regulation of *CG13222* expression in the haltere
279 imaginal discs.

280 An earlier report of ours (Prasad, Tarikere et al. 2016) and this report suggest that
281 unlike the *quad-vg* enhancer of *Drosophila*, an enhancer of *vg* gene of *A. mellifera* is not
282 differentially expressed, but drives GFP expression in both wing and haltere imaginal discs in
283 a transgenic assay. We show that changing the TAAT motif to TAAAT motif, dramatically,
284 brought the enhancer under the negative regulation of Ubx. The *Apis-vg* enhancer, thus,
285 allowed us to dissect the importance of TAAAT motif in dipteran evolution. As *vg* is a pro-wing
286 selector gene and can assign wing fate to any group of dorsal epithelial cells (Kim et al. 1996;

287 Klein and Arias 1998; Williams et al. 1991; Neumann and Cohen 1996), it is a critical target of
288 Ubx to specify haltere development. In this context, a microevolutionary change (TAAT to
289 TAAAT motif) in the enhancer of an ancestral *vg* may be a critical step in the evolution of
290 dipterans.

291 It is, however, very unlikely that a large number of convergent mutations would be
292 involved in this process, wherein multiple TAAT motifs are evolved to TAAAT motifs. A few
293 critical genes, such as *vg*, upstream of wing patterning pathways may have evolved to become
294 targets of Ubx through this route. Once regulatory networks of those genes are modulated
295 by Ubx, chromatin landscape of many other targets (Ubx may bind to these targets via TAAT
296 motifs) would change making them amenable for Ubx-mediated regulation and thereby
297 specifying the haltere fate. In addition, certain co-factors influencing Ubx-mediated
298 regulation of its targets irrespective of Ubx binds to TAAAT or TAAT motif cannot be ruled.
299

300 **Materials and Methods**

301 **1. ChIP sequencing:** Third instar wandering larvae were cut, inverted and fixed with 1.5%
302 PFA (ThermoFisher scientific) for 20 minutes at room temperature and subsequently quenched
303 (125mM Glycine solution, 1XPBS). Samples were washed twice with 1XPBS (Sigma) for
304 10minutes each and dissected for wing and haltere discs (500 number of wing discs and 1000
305 number of haltere discs were used per replicate). Samples were lysed with Cell Lysis Buffer
306 (10mM Tris-Cl pH 8, 10mM NaCl, 0.5% NP40, 1xPIC (Roche)) followed by mechanical
307 shearing for 10minutes on ice. After centrifugation at 2000rpm for 10mins at 4 °C, the was
308 pellet resuspended in Sonication buffer (50mM Tris-Cl pH 8, 1% SDS, 1% NP40, 10mM
309 EDTA, 1xPIC) followed by incubation on ice for 30 minutes. Sonication was carried out using
310 the Covaris S2 sonicator using settings (DC 20%, Intensity 5, Cycles of Burst 200, Time= 20
311 minutes). Samples were centrifuged at 14000rpm for 15mins at 4 °C and precleared using with
312 4ml of Magnetic A Beads (Diagenode) at 4 °C for 1 hour. Preclearing beads were removed and
313 2ml of Polyclonal anti-Ubx antibodies (Agrawal et al. 2011) were added with overnight
314 incubation at 4 °C. Magnetic A beads were added and incubated for 4 hours at 4 °C. The
315 supernatant was removed and beads washed twice with low salt buffer (20mM Tris-Cl pH 8,
316 150mM NaCl, 0.1 % SDS, 2mM EDTA, 1% TritonX-100), twice with High Salt buffer (20mM
317 Tris-Cl pH 8, 500mM NaCl, 0.1 % SDS, 2mM EDTA, 1% TritonX-100), twice with LiCl
318 buffer (10mM Tris-Cl pH 8, 250mM LiCl, 1% NP40, 1% Na-Deoxycholate, 500mM EDTA,
319), and once with TE buffer (10mM Tris-Cl pH 8, 500mM EDTA). Samples were eluted in

320 elution buffer (50mM Tris-Cl, 1mM EDTA, 1% SDS, 50mM NaHCO3). Samples were
321 decrosslinked followed by treatment with RNAse treatment at 37 degrees for 1hour and
322 Proteinase K treatment at 42 °C for 2 hours. DNA was purified using PCI purification and
323 quantified using the Qubit HS DNA quantification system.
324
325 For library preparation, equal amount of DNA (~2 ng) was used as an input for NEB
326 Ultra II DNA library prep kits (NEB #E7645). Number of cycles for amplification of adapter
327 ligated libraries were estimated by the qPCR before final amplification to avoid any bias arising
328 due to PCR amplification and indexing (NEB #E7350). Final amplified libraries were purified
329 twice, first with 1X followed by 0.8x volume of beads per sample using HiPrep PCR clean up
330 system (Magbio #AC-60050). Library concentration was determined using Qubit HS DNA kit
331 (Invitrogen #E7350) and average fragment size was estimated using DNA HS assay on
332 bioanalyzer 2100 (Agilent #5067-4626) before pooling libraries in equimolar ratio. Sequencing
333 reads (100bp PE) were obtained on the Hiseq 2500 V4 platform at Macrogen Inc, Korea.
334
335 **2. ChIP sequencing analysis:** For *D. melanogaster*, raw reads were trimmed using the
336 Trimmomatic software (command) and aligned to the dm6 genome (version BDGP6.28). Peak
337 calling was performed using MACS2 (maxdup=1, FDR 0.01) (Zhang et al. 2008) and high
338 confidence peaks (occurring in at least two biological replicates) were identified (https://ro-
339 che.info/articles/2018-07-11-chip-seq-consensus). For *Apis mellifera*, raw fastq reads were
340 downloaded from NCBI (GSE71847). Reads were aligned to the Amel_4.5 genome (version
341 4.5.47) and peak calling done using MACS2 (maxdup=1, FDR 0.01). High confidence peaks
342 occurring in both replicates were identified. Peaks were annotated to their nearest TSS using
343 the Homer software (annotatePeaks.pl) (Heinz et al. 2010). Gene ontology of targets was
344 performed using GO (Ashburner et al. 2000; Carbon et al. 2021) and overrepresented GO
345 terms plotted. The ChIPseeker program (Yu et al. 2015) was used to annotate peaks to
346 genomic regions. Genome tracks were made using Deeptools (normalizeUsing RPGC, binsize
347 1, minMappingQuality 10, smoothLength 150) (Ramírez et al. 2016) and visualized using the
348 IGV software (Robinson et al. 2011). Heatmaps were generated using Deeptools
349 (referencePoint TSS).
350

351 **3. RNA seq data generation and analysis:** Third instar wandering larvae of the Canton S
352 strain were collected, cut, inverted and dissected for wing and haltere discs in cold PBS.
353 Experiments were carried out in triplicates and samples were snap frozen in Trizol. Library
354 preparation and sequencing on Illumina Hiseq2000 was performed at Genotypic Technology,
355 Bangalore, India. Raw fastq files were obtained and used for analysis. Raw reads were
356 trimmed using Trimmomatic and aligned to the reference genome using the Hisat2 software
357 (Kim et al. 2015, 2) followed by read counts using the Htseq software (Anders et al. 2015).
358 Differentially expressed genes were identified using the edgeR software (Robinson et al.
359 2010) using 1.5 fold difference as the cut-off.
360
361 **4. Motif analyses and protein sequence comparison:** De-novo motif analysis was carried out
362 using Homer (findMotifsGenome.pl). The PWM for the TAAAT motif obtained from Homer
363 was converted to Transfac format using RSAT (Nguyen et al. 2018) and finally converted to
364 MEME format using transfac2meme command. Identification of binding motifs in enhancers
365 of *CG13222, quad-Vg* and *ApisVg* was performed using the MAST software from MEME suite
366 (Bailey et al. 2009).
367 Protein sequences of Ubx from different insect species were downloaded from Uniprot and
368 multiple sequence alignment performed using Clustal Omega (Madeira et al. 2019).
369 Alignments were visualised and processed using the Jalview software (Waterhouse et al.
370 2009).
371
372 5. **Frequency calculation of TAAAT motif in Ubx response elements:** Using a 1.5 fold
373 difference (FD) cut-off, genes that are upregulated or downregulated in the haltere imaginal
374 discs were identified. Ubx targets obtained from ChIP-seq were mapped to differentially
375 expressed genes to identify putative Ubx response elements that are upregulated,
376 downregulated or not-differentially expressed in the halteres. The PWM of the TAAAT motif
377 was generated and its frequency calculated in each group of Ubx response elements using the
378 FIMO software.
379
380 **6. Plasmids constructs, cloning and transgenic fly generation:** The full list of primers as well
381 as sequences of enhancers of the *CG13222, quad-Vg* and *ApisVg* is provided as supplementary
382 data (File S2). For Luciferase assays, site-directed mutagenesis was used to clone enhancer

383 constructs of the *CG13222* gene between Kpn1 and Nhe1 restriction sites, upstream of a
384 modified pGL3 vector containing a 5X Dorsal binding site. Metallothionein inducible pRMHa3
385 vectors containing Ubx were previously generated in the lab and empty pRMHa3 vector was
386 generated by excising the cloned Ubx sequence. For generating transgenics, all enhancer
387 constructs were cloned between the Nhe1 and Kpn1 restriction sites of the pH-stinger-attb (a
388 kind gift from Manfred Frasch) and injected into the attp40 site on the second chromosome
389 using a mini white screen at NCBS fly facility, Bangalore. All constructs were sequence verified
390 before use in functional assays. Insertion of various transgenes in attp40 site helped for better
391 comparison of wildtype and mutant transgenes of a given enhancer.
392
393 **7. Luciferase assays:** S2 cells were checked for contamination before performing Luciferase
394 assays. Cells were plated onto 24 well plates at a density of 3*10^5 cells per well 6 hours prior
395 to transfection. For every construct, either wild type or mutant, two sets of experiments were
396 designed; one well was co-transfected with the enhancer construct in pGL3 vector and the
397 empty pRMHa3 vector whereas the other well was co-transfected with the enhancer
398 construct and pRMHa3 vector containing Ubx. Renilla luciferase was used as an internal
399 control and co-transfected in all experiments. Transfection was carried out using the
400 Effectene transfection reagent and all experiments carried out in 3 technical replicates and at
401 least 3 biological replicates. 48 hours post transfection, sterile CuSO4 solution was used to
402 induce expression of Ubx at a final concentration of 500um and incubated for 24 hours. Cells
403 were harvested, pelleted down (1000rpm for 4mins) and 100ul of 1X Passive Lysis buffer
404 added and vortexed to dissolve the cell pellet. Cells were incubated for 15mins at room
405 temperature and further spun at 10,000rpm for 90secs to collect the supernatant. The
406 luminescence was measured with the Dual-glo Luciferase assay kit (Promega) on Ensight Plate
407 reader (Perkin Elmer). All readings were normalized to Renilla luminescence and datasets
408 compared using the Prism software.
409
410 **8. EMSA:** VC-Ubx and VC-Scr were cloned in the PcDNA3 vector and produced with the TNT-
411 T7-coupled in vitro transcription/translation system (Promega) for EMSAs, as previously
412 described (Hudry et al. 2012). Shortly, between 3 and 6 microliters of programmed lysate was
413 used for each protein (100 ng/ml of proteins were produced on average). The VC-Ubx and VC-
414 Scr fusion proteins were produced separately (0,5mg of each plasmid was used for the in

415 vitro transcription/translation reaction). Supershift against the VC fragment was performed
416 by adding the anti-GFP antibody after 15 minutes in the binding reaction. Each band shift
417 experiment was repeated at least two times.
418
419 **9. Immunostaining and Microscopy:** Wandering third instar larvae were cut and inverted in
420 cold PBS followed by fixation with 4% PFA for 20mins with gentle rocking. Samples were
421 washed thrice with 0.1% PBTX (0.1% Triton in PBS) for 10 minutes each, followed by one hour
422 of blocking at room temperature (blocking solution: 0.5% BSA in 0.1%PBTX). Samples were
423 incubated with primary antibodies (dilutions made in blocking solution) at 4 ^{o}C overnight
424 followed by washing with 0.1% PBTX for 10mins (x3) at room temperature. Samples were
425 incubated with secondary antibodies for one hour at room temperature, washed with 0.1%
426 PBTX for 10mins (x3) and finally with PBS before mounting. Wing and haltere imaginal discs
427 were dissected out, mounted using Prolong Gold Antifade (Invitrogen) and stored at 4
428 degrees. Antibodies used in the study are Rb-GFP (1:1000) (Invitrogen), Rb-Ubx (1:1000)
429 (Agrawal et al., 2011), m-Ubx (1:30) (DSHB), goat-Hth (1:100) (Santa Cruz Biotechnology),
430 Alexa Fluor 568 (1:1000) (Invitrogen) and Alexa Fluor 633 (1:1000) (Invitrogen). All images
431 were taken on Leica Sp8 system at Microscopy Facility, IISER Pune. Laser settings (Laser Power
432 (LP) and Gain (G)) used for measuring GFP fluorescence have been indicated for each image
433 panel. All images were processed using the ImageJ software and compiled using Microsoft
434 PowerPoint and Adobe Illustrator.
435
436 **Data availability**: Raw fastq files of ChIP sequencing for the Ubx protein in *Apis mellifera*
437 hindwings can be accessed from NCBI (GSE71847). Raw fastq files of ChIP sequencing for the
438 Ubx protein in Drosophila melanogaster halteres and RNA sequencing data for wing and
439 haltere imaginal discs can be accessed from NCBI (GSE205177 and GSE205352).
440
441 **Acknowledgments**
442 We thank G. Deshpande and members of the LSS and SM laboratories for critical input. We
443 thank S. Galande for providing reagents for Library preparation. This work was supported
444 primarily by an Indo-French Research grant from CEFIPRA to SM; a JC Bose Fellowship and
445 grant from the Department of Science and Technology, Government of India to LSS; and a
446 Council of Scientific & Industrial Research (CSIR) Fellowship to SK.

447

448 **Author contributions:**

449 SK carried out all fly experiments, image analyses, and wrote the MS.

450 SK and SP did ChIP-Seq and RNA-Seq and analyses.

451 GG, FB and RP did EMSA experiments.

452 LSS and SM conceived the project and wrote the MS.

453 We declare “no-conflict-of-interest”.

454

455 **References**

456 Agrawal P, Habib F, Yelagandula R, Shashidhara LS. 2011. Genome-level identification of
457 targets of Hox protein Ultrabithorax in Drosophila: novel mechanisms for target
458 selection. *Sci Rep* **1**: 1–10.

459 Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput
460 sequencing data. *Bioinformatics* **31**: 166–169.

461 Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight
462 SS, Eppig JT, et al. 2000. Gene Ontology: tool for the unification of biology. *Nat Genet*
463 **25**: 25–29.

464 Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009.
465 MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Res* **37**: 202–208.

466 Cabrera C V., Botas J, Garcia-Bellido A. 1985. Distribution of Ultrabithorax proteins in
467 mutants of Drosophila bithorax complex and its transregulatory genes. *Nature* **318**:
468 569–571.

469 Carbon S, Douglass E, Good BM, Unni DR, Harris NL, Mungall CJ, Basu S, Chisholm RL,
470 Dodson RJ, Hartline E, et al. 2021. The Gene Ontology resource: Enriching a GOld
471 mine. *Nucleic Acids Res* **49**: D325–D334.

472 Carroll SB. 1995. Homeotic genes and the evolution of arthropods and chordates. *Nature*
473 **376**: 479–485.

474 Casares F, Mann RS. 2000. A dual role for homothorax in inhibiting wing blade development
475 and specifying proximal wing identities in Drosophila. *Development* **127**: 1499–1508.

476 Castelli-Gair JE, Micol JL, Garcia-Bellido A. 1990. Transvection in the Drosophila
477 Ultrabithorax gene: A Cbx1 mutant allele induces ectopic expression of a normal
478 allele in trans. *Genetics* **126**: 177–184.

479 Galant R, Walsh CM, Carroll SB. 2002. Hox repression of a target gene: extradenticle-
480 independent, additive action through multiple monomer binding sites. *Development*
481 **129**: 3115–3126.

482 Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass
483 CK. 2010. Simple Combinations of Lineage-Determining Transcription Factors Prime
484 cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol Cell* **38**:
485 576–589.

486 Hersh BM, Carroll SB. 2005. Direct regulation of knot gene expression by Ultrabithorax and
487 the evolution of cis-regulatory elements in Drosophila. *Development* **132**: 1567–
488 1577.

489 Hersh BM, Nelson CE, Stoll SJ, Norton JE, Albert TJ, Carroll SB. 2007. The UBX-regulated
490 network in the haltere imaginal disc of D . melanogaster. **302**: 717–727.

491 Hudry B, Remacle S, Delfini MC, Rezsohazy R, Graba Y, Merabet S. 2012. Hox proteins
492 display a common and ancestral ability to diversify their interaction mode with the
493 pbc class cofactors. *PLoS Biol* **10**.

494 Khan S, Dilsha C, Shashidhara LS. 2020. Haltere development in D. melanogaster:
495 Implications for the evolution of appendage size, shape and function. *Int J Dev Biol*
496 **64**: 169–175.

497 Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory
498 requirements. *Nat Methods* **12**: 357–360.

499 Kim J, Sebring A, Esch JJ, Kraus ME, Vorwerk K, Magee J, Carroll SB. 1996. Integration of
500 positional signals and regulation of wing formation and identity by Drosophila
501 vestigial gene. *Nature* **382**: 133–138.

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

505 Lewis EB. 1978. A gene complex controlling segmentation in Drosophila. *Genes Dev Cancer*
506 *Life Work Edw B Lewis* 229–242.

507 Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter
508 SC, Finn RD, et al. 2019. The EMBL-EBI search and sequence analysis tools APIs in
509 2019. *Nucleic Acids Res* **47**: W636–W641.

510 Matsuoka Y, Monteiro A. 2021. Hox genes are essential for the development of eyespots in
511 *Bicyclus anynana* butterflies ed. P. Wittkopp. *Genetics* **217**: iyaa005.

512 McGinnis W, Krumlauf R. 1992. Homeobox genes and axial patterning. *Cell* **68**: 283–302.

513 Mohit P, Makhijani K, Madhavi MB, Bharathi V, Lal A, Sirdesai G, Reddy VR, Ramesh P,
514 Kannan R, Dhawan J, et al. 2006. Modulation of AP and DV signaling pathways by the
515 homeotic gene Ultrabithorax during haltere development in Drosophila. **291**: 356–
516 367.

517 Neumann CJ, Cohen SM. 1996. A hierarchy of cross-regulation involving Notch, wingless,
518 vestigial and cut organizes the dorsal/ventral axis of the Drosophila wing.
519 *Development* **122**: 3477–3485.

520 Nguyen NTT, Contreras-Moreira B, Castro-Mondragon JA, Santana-Garcia W, Ossio R,
521 Robles-Espinoza CD, Bahin M, Collombet S, Vincens P, Thieffry D, et al. 2018. RSAT
522 2018: Regulatory sequence analysis tools 20th anniversary. *Nucleic Acids Res* **46**:
523 W209–W214.

524 Prasad N, Tarikere S, Khanale D, Habib F, Shashidhara LS. 2016. A comparative genomic
525 analysis of targets of Hox protein Ultrabithorax amongst distant insect species. *Sci*
526 *Rep* **6**: 27885.

527 Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke
528 T. 2016. deepTools2: a next generation web server for deep-sequencing data
529 analysis. *Nucleic Acids Res* **44**: W160–W165.

530 Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP.
531 2011. Integrative genomics viewer. *Nat Biotechnol* **29**: 24–26.

532 Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential
533 expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140.

534 Sánchez-Higueras C, Rastogi C, Voutev R, Bussemaker HJ, Mann RS, Hombría JCG. 2019. In
535 vivo Hox binding specificity revealed by systematic changes to a single cis regulatory
536 module. *Nat Commun* **10**.

537 Shashidhara LS, Agrawal N, Bajpai R, Bharathi V, Sinha P. 1999. Negative Regulation of
538 Dorsoventral Signaling by the Homeotic Gene Ultrabithorax during Haltere
539 Development in Drosophila. *Dev Biol* **212**: 491–502.

540 Slattery M, Riley T, Liu P, Abe N, Gomez-Alcala P, Dror I, Zhou T, Rohs R, Honig B,
541 Bussemaker HJ, et al. 2011. Cofactor binding evokes latent differences in DNA
542 binding specificity between Hox proteins. *Cell* **147**: 1270–1282.

543 Tomoyasu Y, Wheeler SR, Denell RE. 2005. Ultrabithorax is required for membranous wing
544 identity in the beetle Tribolium castaneum. *Nature* **433**: 643–647.

545 Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. 2009. Jalview Version 2-A
546 multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**:
547 1189–1191.

548 Weatherbee SD, Halder G, Kim J, Hudson A, Carroll S. 1998. Ultrabithorax regulates genes at
549 several levels of the wing-patterning hierarchy to shape the development of the
550 *Drosophila* haltere. *Genes Dev* **12**: 1474–1482.

551 Weatherbee SD, Nijhout HF, Grunert LW, Halder G, Galant R, Selegue J, Carroll S. 1999.
552 Uitrabithorax function in butterfly wings and the evolution of insect wing patterns.
553 *Curr Biol* **9**: 109–115.

554 White RAH, Akam ME. 1985. Contrabithorax mutations cause inappropriate expression of
555 Ultrabithorax products in Drosophila. *Nature* **318**: 567–569.

556 White RAH, Wilcox M. 1985. Regulation of the distribution of Ultrabithorax proteins in
557 Drosophila. *Nature* **318**: 563–567.

558 Williams JA, Bell JB, Carroll SB. 1991. Control of Drosophila wing and haltere development
559 by the nuclear vestigial gene product. *Genes Dev* **5**: 2481–2495.

560 Yu G, Wang LG, He QY. 2015. ChIP seeker: An R/Bioconductor package for ChIP peak
561 annotation, comparison and visualization. *Bioinformatics* **31**: 2382–2383.

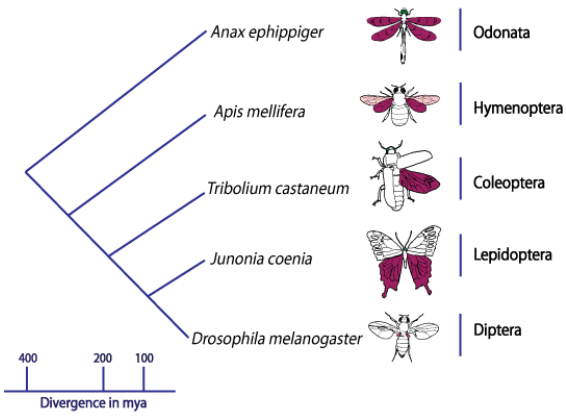
562 Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nussbaum C, Myers RM,
563 Brown M, Li W, et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*
564 **9**.

565

Figures and Figure legends

Figure 1: Motif with a TAAAT core sequence is enriched in direct targets of Ubx in developing halteres in *D. melanogaster* but not in developing hindwings in *A. mellifera*

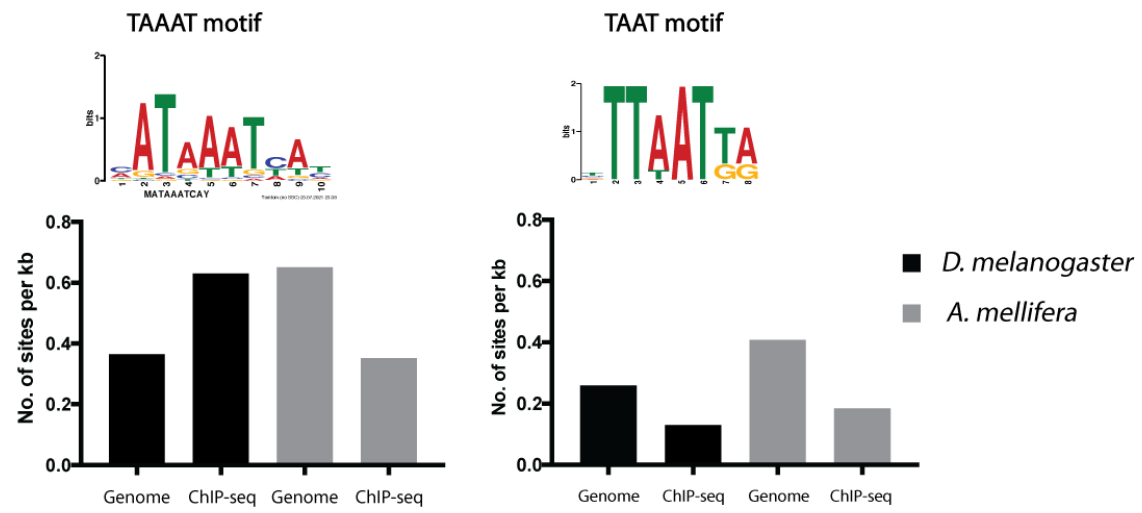
A



B

Enriched motif from de-novo analysis		p Value	
		D. melanogaster	*A. mellifera*
	TAAAT motif	1e-113	Not enriched
	CTCF	1e-39	1e-431
	Trl/GAGA	1e-60	1e-250
	Brk	1e-17	1e-17

C

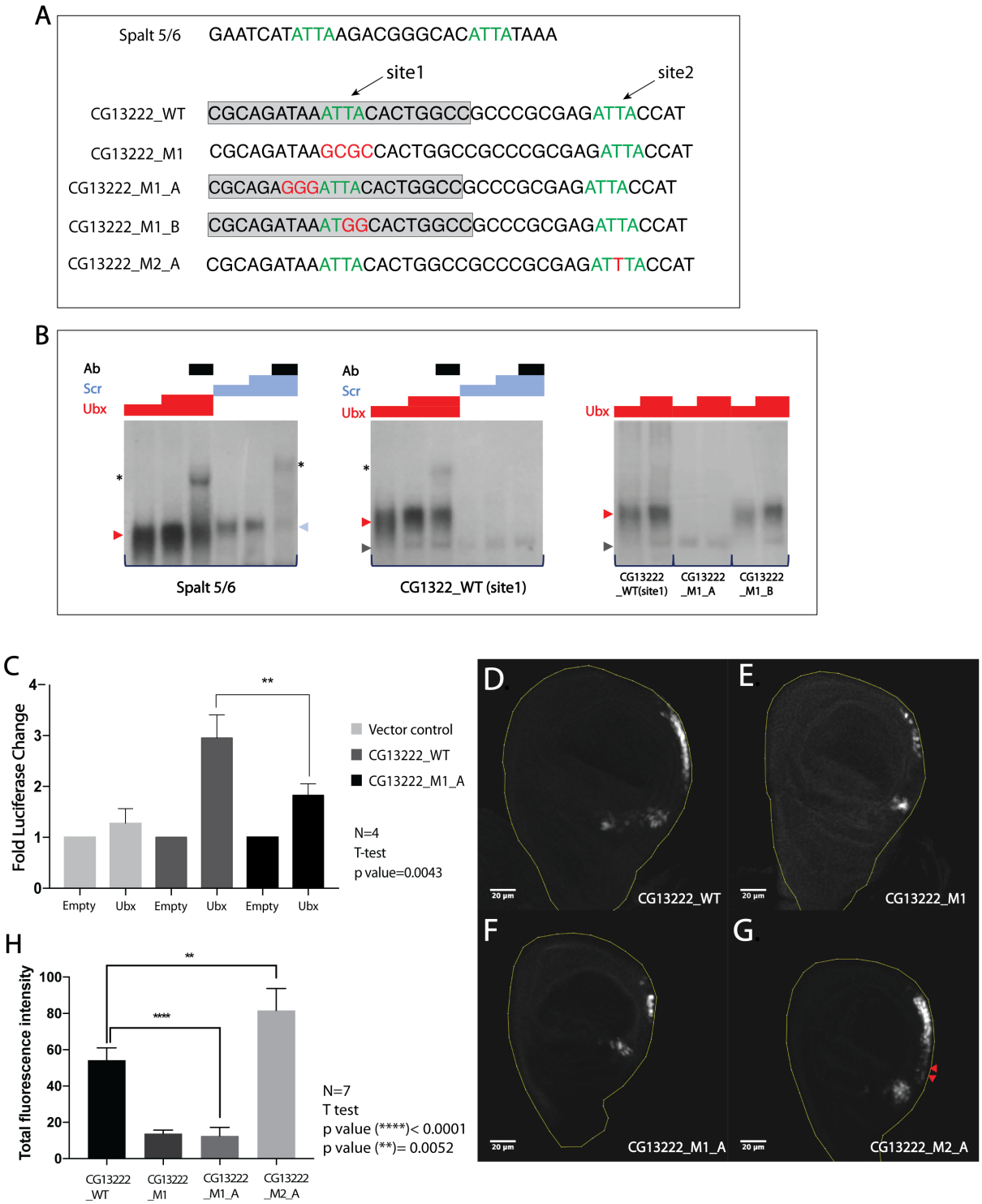


Legend Fig. 1:

A) While Ubx is required for haltere development in Dipterans such as *Drosophila melanogaster*, differences in hindwing morphology in the third thoracic segment of diverse insects is not correlated to the mere presence of Ubx.

B) De-novo motif analysis of Ubx ChIP-seq in *Apis* hindwings and *Drosophila* halteres reveal enrichment of the TAAAT motif specifically in *Drosophila* but not in *Apis*. Binding sites for transcription factors like Trl, CTCF and Brk are, however, enriched in both datasets.
C) FIMO analysis of the TAAAT motif reveals a 1.7fold enrichment of sites in *Drosophila* ChIP-seq sequences as compared to the entire genome. No such enrichment was found for the *Apis* ChIP-seq datasets. The TAAT motif was not enriched in either of the datasets.

Figure 2: The TAAAT motif is critical for Ubx mediated regulation of a target gene in *Drosophila* halteres



Legend Fig. 2:

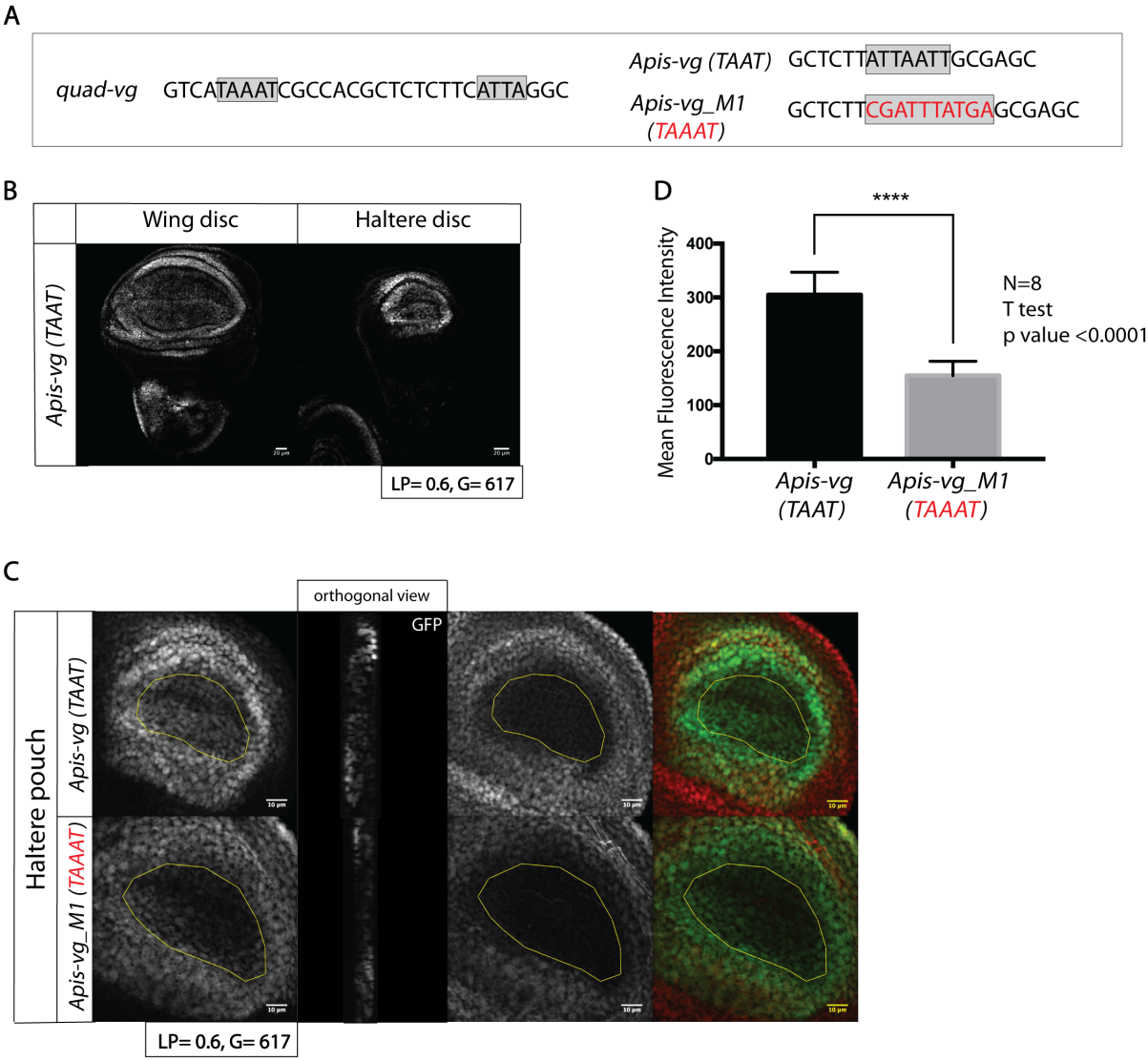
A) Sequence of part of *sal* enhancer (Spalt 5/6), *CG13222* enhancer and mutations of the *CG13222* enhancer (CG13222_M1_A, CG13222_M1_B) used as probes for EMSA. The wild type *CG13222* enhancer has two Ubx binding sites (shown in green) termed here as CG13222_WT (site1) and CG13222_WT (site2). Mutations were generated to specifically modify either the TAAT (CG13222_M1_B, CG13222_M2_A) or the TAAAT (CG13222_M1_A) motifs while keeping the other one intact in site1 of the *CG13222* enhancer. For Luciferase and Transgenic assays, the entire 512bp spanning the *CG13222* edge enhancer was used.

B) While both Ubx and Scr bind to TAAT motif of *sal*, only Ubx binds to TAAAT motif of CG13222_WT (site1). Binding of Ubx to CG13222_WT (site1) is severely reduced when the TAAAT motif is mutated to TAAT motif (CG13222_M1_A), whereas no such effect is seen on mutating the TAAT motif (CG13222_M1_B).

C) The *CG13222* enhancer drives Luciferase activity in a S2 cell culture system in presence of Ubx. Significant loss of reporter activity is observed in the mutant enhancer (CG13222_M1_A) where the TAAAT motif is affected. For statistical analysis, t-test was performed (two-tailed).

D-H). The *CG13222* enhancer drives reporter expression in the posterior edge of the third instar haltere imaginal discs (D). The *CG13222_M1* mutant where both the TAAT and TAAAT motifs are affected show a marked reduction in reporter expression (E). The *CG13222_M1_A* mutant carrying changes in the TAAAT site of the enhancer has a similar phenotype as the *CG13222_M1* mutant (F) and shows a significant reduction in GFP expression as seen from calculation of average fluorescence intensity (H). In the *CG13222_M2_A* enhancer where the TAAT is mutated to TAAAT in site2 results in significantly higher level of reporter expression (H) along with ectopic expression in some domains (shown by red arrows) (G). For statistical analysis, t-test was performed (two-tailed).

Figure 3: The TAAAT motif causes Ubx mediated repression of an orthologous enhancer of the *vestigial* gene from *A. mellifera* in a transgenic *Drosophila* assay



Legend Fig. 3:

A) Sequence of part of the 805bp enhancer of the *vestigial* gene in *Drosophila (quad-vg)* containing both TAAT and TAAAT motifs and part of the 575bp enhancer of the *vestigial* gene in *Apis (Apis-vg)*, showing the presence of a TAAT motif.

B) The Apis-vg enhancer drives similar expression of reporter GFP in hinge and pouch regions of wing and haltere imaginal discs.

C) The mutant enhancer (*Apis-vg_M1*) with introduced TAAAT motif shows reduced levels of GFP expression in the haltere pouch. Magnified images of the haltere pouch of *Drosophila* transgenics expressing GFP under *Apis-vg* and *Apis-vg_M1* enhancers. Note much reduced

GFP levels specifically in the haltere pouch of *Apis-vg_M1* transgene. Orthogonal views of the haltere imaginal pouch indicate a clear difference in GFP expression driven by the WT and the mutant enhancers. Hth staining, which is hinge-specific is used to demarcate the pouch region. **D)** Quantification of average fluorescence intensity in the haltere pouch for the *Apis-vg* and *Apis-vg_M1* transgenics. For statistical analysis, t-test was performed (two-tailed).