bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

1 # Conservation and flexibility in the gene regulatory
2 # landscape of Heliconiine butterfly wings

3 Joseph J Hanly[1,2,3*], Richard WR Wallbank[1,2], W Owen
4 McMillan[2], Chris D Jiggins[1,2]

5 1. Department of Zoology, University of Cambridge, Downing St., Cambridge
6 CB2 3EJ, UK
7 2. Smithsonian Tropical Research Institute, Gamboa, Panama
8 3. The George Washington University, Washington DC, 20052, USA

9 * = corresponding author. joe.hanly@gmail.com

10 **Abstract**
11 **Background**
12 Many traits evolve by *cis*-regulatory modification, by which changes to non-coding
13 sequences affect the binding affinity for available transcription factors and thus
14 modify the expression profile of genes. Multiple examples of *cis*-regulatory evolution
15 have been described at pattern switch genes responsible for butterfly wing pattern
16 polymorphism, including in the diverse neotropical genus *Heliconius,* but the
17 identities of the factors that can regulate these switch genes have not been identified.
18 **Results**
19 We investigated the spatial transcriptomic landscape across the wings of three closely
20 related butterfly species, two of which have a convergently-evolved, co-mimetic
21 pattern, the other having a divergent pattern. We identified candidate factors for
22 regulating the expression of wing patterning genes, including transcription factors
23 with a conserved expression profile in all three species, and others, including both
24 transcription factors and Wnt pathway genes, with markedly different profiles in each
25 of the three species. We verified the conserved expression profile of the transcription
26 factor homothorax by immunofluorescence, and showed that its expression profile
27 strongly correlates with that of the selector gene optix in butterflies with the
28 Amazonian forewing pattern element 'dennis'.
29 **Conclusions**

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

30 Here we show that, in addition to factors with conserved expression profiles like
31 homothorax, there are also a variety of transcription factors and signaling pathway
32 components that appear to vary in their expression profiles between closely related
33 butterfly species, highlighting the importance of genome-wide regulatory evolution
34 between species.

35

36 **Keywords:** *cis*-regulation, *Heliconius,* butterfly, transcription factor, homothorax,
37 gene expression, Wnt signaling, transcriptomics

38

39

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

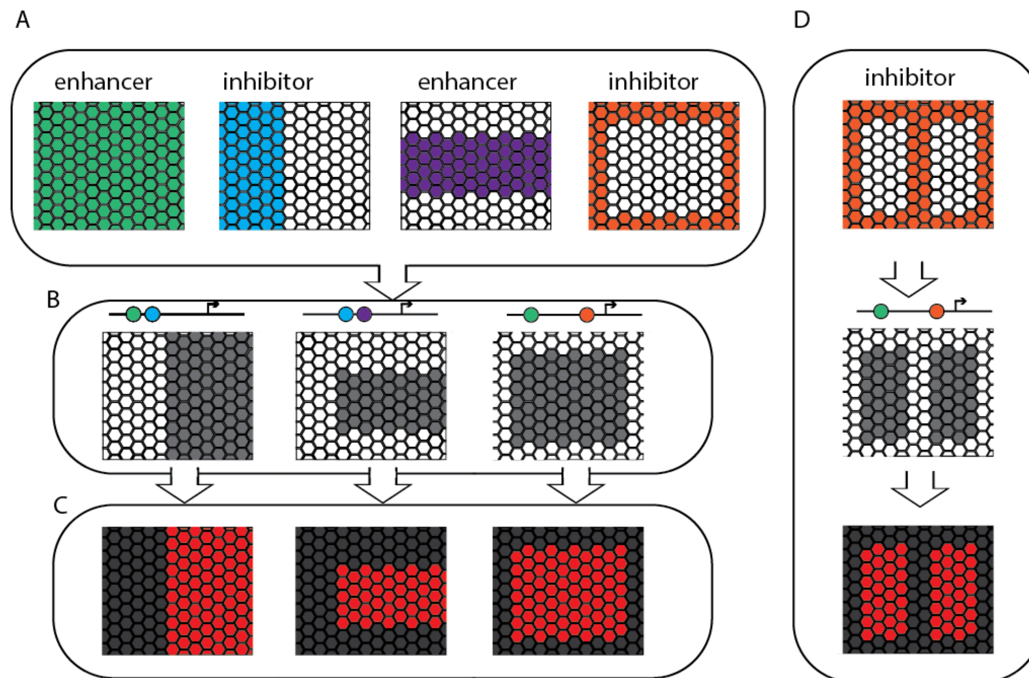
40 **Background**

41 A major challenge in evolutionary developmental biology is to understand how
42 modifications to gene expression can lead to biological diversity. In particular,
43 modulation of *cis*-regulatory elements has been repeatedly identified as the material
44 source of polymorphism and divergence in physiology, behaviour, pigmentation
45 patterns and morphological structures; Gephebase, a database of genotype-phenotype
46 relationships, identifies 323 such examples of *cis*-regulatory evolution in diverse
47 eukaryote clades (Martin and Orgogozo, 2013). These elements must function by
48 differential binding of regulatory factors, and so to understand the evolution of gene
49 regulatory networks, we must first identify which regulatory factors are present and
50 able to perform this function in a given spatial and temporal context.
51
52 The Lepidoptera make up ~18% of described animal diversity and have a vast array of
53 wing patterns, both within and between species. Recent advances in genetics,
54 genomics and experimental methods have begun to uncover the underlying genetic
55 and developmental basis of lepidopteran wing pattern variation (Jiggins et al., 2017,
56 Livraghi et al., 2017). One of the most diverse and well-studied groups are the
57 *Heliconius* butterflies, and it is now understood that much of the variation in wing
58 pattern in this group results from regulatory evolution at just three genes, *optix, WntA*
59 and *cortex* (Reed et al., 2011, Martin et al., 2012, Nadeau et al., 2016). CRISPR/Cas9
60 mutagenesis has shown that *optix* and *WntA* are also involved in the patterning of
61 wings in multiple butterfly lineages (Mazo-Vargas et al., 2017, Zhang et al., 2017). At
62 each of these loci there is a large diversity of complex regulatory alleles, controlling
63 expression patterns across the wing surface during development (Van Belleghem et
64 al., 2017, Enciso-Romero et al., 2017, Wallbank et al., 2016). This regulatory
65 diversity that generates the extraordinary variation in wing patterns nonetheless acts
66 against a background of a highly conserved regulatory factors that underlie insect
67 wing development. Consistent with this idea, earlier candidate gene studies have
68 shown that many patterning factors previously identified in *Drosophila* wing
69 development show similar patterns of expression on butterfly wings (Table 1).
70
71 If this is in fact the case, then the pattern variation we observe in *Heliconius* and other
72 butterflies could be generated by the differential “readout” of a highly conserved set

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

73 of transcription factors that effectively "prepattern" the wing (figure 1A). These
74 conserved expression patterns could then provide input to the regulatory elements of
75 pattern switch genes like *optix* (figure 1B), and in turn, modifications of these
76 elements could lead to production of wing pattern diversity (figure 1C), a mechanism
77 that allows for the gain of novel phenotypes with the avoidance of deleterious
78 pleiotropic effects (Prud'homme et al., 2007). This hypothesis has previously been
79 examined at the within-species level in *Heliconius* through transcriptomics (Hines et
80 al., 2012) Key examples of this mode of regulatory evolution have been described in
81 evolution of melanic patterns in *Drosophila* species. For example, the protein
82 Engrailed is a deeply conserved component that specifies the posterior compartment
83 in arthropod segmentation (Patel et al., 1989a, Patel et al., 1989b). Some *Drosophila*
84 species have a melanic spot on the anterior tip of the wing, which is sculpted in part
85 by repression of the *yellow* gene by *en*, on the posterior boundary (Gompel et al.,
86 2005). Here, *cis*-regulatory evolution at the *yellow* locus locked onto the conserved
87 spatial information encoded by *en*. The expression and function of the *en* gene did not
88 change, rather a new regulatory connection was established that modified the
89 expression of yellow to generate novel diversity. Likewise, expression of
90 pigmentation genes in the thorax of *D. melanogaster* is repressed by the expression of
91 the transcription factor *stripe*, which specifies flight muscle attachment sites. The
92 shape of this thoracic element is thus constrained by factors that specify flight muscle
93 pattern (Gibert et al., 2018). On the other hand, there are also examples of *Drosophila*
94 pigmentation evolving by modification to the *trans*-regulatory landscape, by which
95 conserved components of the regulatory landscape themselves change in expression
96 profile to affect downstream melanin synthesis domains. For example; *Dll* in *D.*
97 *biarmipes* and *D. prolongata* has gained additional expression profiles to its usual
98 peripheral pattern, in correlation with melanic elements (Arnoult et al., 2013). This
99 mode of *trans*-regulatory divergence could also play a role in the evolution of
100 butterfly wing patterns (Figure 1D). No evidence has been found for this at the
101 intraspecific level, but could occur between more distantly related species for which
102 genetic mapping is not
103 possible.

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.



104

105 **Figure 1 - Hypothetical mechanisms of wing pattern development and evolution**.
106 In this model, a set of prepatterning factors (A) are expressed early in the developing
107 wing, may pattern the development of the wing, and do not vary in their expression
108 profiles in different morphs or species (see Table 1). These factors feed in to the
109 regulation of the wing pattern switch genes and shape their expression profiles
110 accordingly, for example in *Heliconius* the transcription factor optix (B), which
111 causes scale cells that would otherwise develop to be melanic to express
112 ommochrome pigments (C). It is also possible that changes to the expression of wing
113 pattern switch genes like optix could be caused by changes in expression of
114 prepatterning factors (D).

115

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

116 Here, we begin to test these ideas by using comparative transcriptomic sequencing in
117 two well-described species in the genus *Heliconius*, as well as an outgroup species
118 *Agraulis vanillae*. The two *Heliconius*, *H. erato* and *H. melpomene*, are co-mimics
119 that diverged from each other between 10-12 mya (Kozak et al., 2015), and recently
120 co-diversified into around 25 different co-occurring wing pattern types. In contrast, *A.*
121 *vanillae*, which diverged from *Heliconius* roughly 25 mya, is largely monomorphic
122 across its extensive geographic distribution. Linkage and association studies of pattern
123 variation in *Heliconius* and other butterflies have repeatedly identified non-coding
124 regions as primary candidates for the loci of evolution (Kunte et al., 2014, Iijima et
125 al., 2018, Gallant et al., 2014, Wallbank et al., 2016, Van Belleghem et al., 2017). We
126 hypothesize that these candidate regulatory elements allow for regulatory coupling to
127 upstream patterning transcription factors in the wing. In order to understand the
128 upstream spatial information that provides an input to butterfly wing patterning, we
129 need to understand spatial patterns of gene expression in the developing wing,
130 building on a primarily gene-by-gene, candidate driven approach, as in many previous
131 studies which have primarily used factors known from *Drosophila* wing development
132 (Table 1). In particular, these data help us determine which transcription factors show
133 consistent spatial expression profiles in different species and are therefore candidate
134 constituents of a conserved developmental landscape, and which transcription factors
135 show variable patterns and are therefore candidates for the causative regulators of
136 pattern differences. In addition, following on from the discovery that WntA is a key
137 patterning gene in *Heliconius* and *Agraulis* (Mazo-Vargas et al 2017), we also
138 characterised the expression of *Wnt* pathway constituents in all three species. Our
139 results highlight both strong conservation and striking flexibility in the gene
140 regulatory landscape in the early wing development in *Heliconius*.

141

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

142

a) Gene	Species	Reference
Ultrabithorax	*J. coenia*	(Warren et al., 1994, Weatherbee et al., 1999)
apterous	*J. coenia, B. anynana*	(Carroll et al., 1994, Prakash and Monteiro, 2017)
invected	*J. coenia*	(Carroll et al., 1994)
engrailed	*P. rapae, B. anynana, Saturnia pavonia, Antheraea polyphemus*	(Monteiro et al., 2006)
scalloped	*J. coenia*	(Carroll et al., 1994)
wingless	*V. cardui*, *A. vanillae*, *J. coenia*, *Battus philenor, B. mori*, *M. sexta*, *Z. morio, B. anynana, P. rapae*	(Macdonald et al., 2010, Monteiro et al., 2006)
cut	*V. cardui*, *A. vanillae*, *J. coenia*, *Battus philenor, B. mori*, *M. sexta*, *Z. morio*	(Macdonald et al., 2010)
hedgehog	*J. coenia*	(Keys et al., 1999)
cubitus interruptus	*J. coenia*	(Keys et al., 1999)
patched	*J. coenia*	(Keys et al., 1999)
achaete-scute	*J. coenia*	(Galant et al., 1998)
Notch	*H. erato, H. melpomene, A. vanillae P. rapae, M. sexta, V. cardui, J. coenia, B. anynana*	(Reed and Serfas, 2004, Reed, 2004)
Optix	*Heliconius spp, J. coenia, V. cardui,*	(Reed et al., 2011, Martin et al., 2014)
WntA	*Heliconius spp, Limenitis arthemis, A. vanillae, J. coenia,*	(Martin and Reed, 2014, Gallant et al., 2014)
doublesex	*Papilio*	(Kunte et al., 2014)
aristaless1 & aristaless2	*J. coenia, Limenitis arthemis, Spodoptera ornithogalli, Ephestia kuehniella*	(Martin and Reed, 2010)
Distal-less	*H. erato, H. melpomene, A. vanillae P. rapae, M. sexta, V. cardui, J. coenia, B. anynana, Saturnia pavonia, Antheraea polyphemus*	(Reed and Serfas, 2004)
spalt major	*V. cardui, P. rapae, P. oleracea, Colias philodice, C. eurytheme, B. anynana*	(Zhang and Reed, 2016, Monteiro et al., 2006, Stoehr et al., 2013)
decapentaplegic	*J coenia*	(Carroll et al., 1994)
Ecdysone receptor	*J coenia*	(Koch et al., 2003)
pSmad	*B. anynana, P. rapae*	(Monteiro et al., 2006)
Antennapedia	*B. anynana, Heteropsis iboina, Pararge aegeria* and *Melanargia galathea, Lasiommata megera, J. coenia, M. cinxia, Inachis io, Caligo memnon*	(Saenko et al., 2011)
cortex	*H. melpomene, H. numata*	(Nadeau et al., 2016)
armadillo	*B. anynana*	(Connahs et al Biorxiv 2017)
BarH1	*Colias*	(Woronik et al Biorxiv 2018)

143 Table 1. A summary of previous single gene expression studies in butterfly wing
144 development.

145

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

146 **Results**

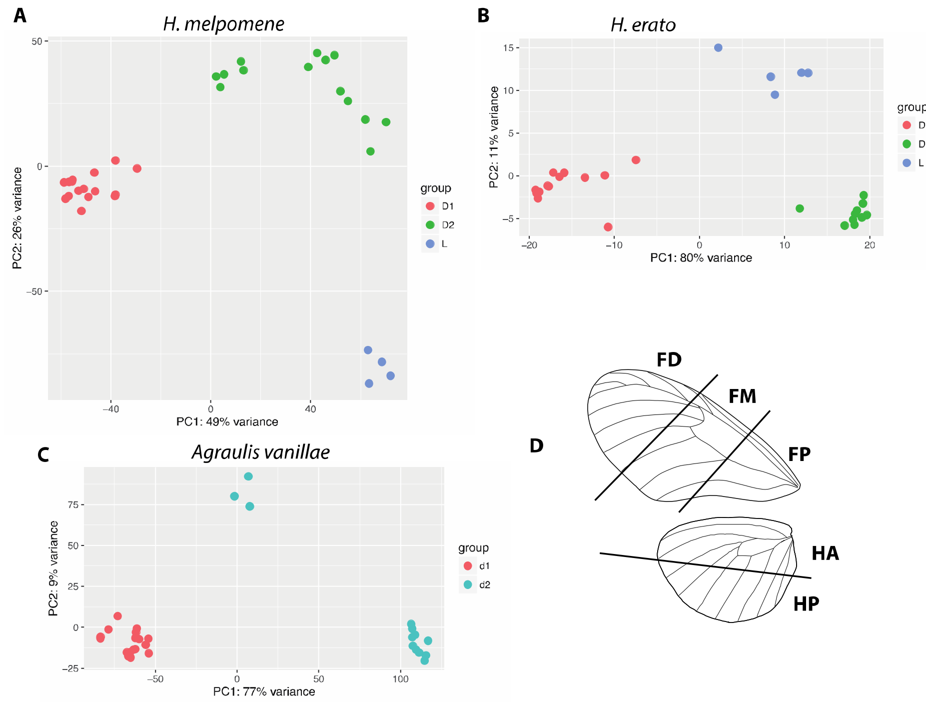
147 To gain insights into the regulatory landscape that organizes butterfly wing early
148 patterning, we conducted transcriptomic analysis of 110 samples representing whole
149 larval wings from *H. melpomene* and *H. erato*, and pupal wings dissected into 5
150 sections from *H. melpomene, H. erato* and *A. vanillae*. Between 10-24 million reads
151 were sequenced per sample, and the average percentage of reads per sample that did
152 not map was 11.8% (Table S1), compared to a previous RNAseq study in *H.*
153 *melpomene* in which more than 50% of reads failed to map (Walters et al., 2015). All
154 samples passed quality controls, and could be included in the differentia expression
155 analyses.

156 PCA analysis showed clustering of samples within each species by stage (Figure 2).
157 Sample clustering by compartment is clear in Day 1 samples, (Figure S1), and in
158 *Agraulis* and *H. melpomene*, distinct clusters for forewing and hindwing are also
159 present, indicating that there is sufficient detectible differential expression between
160 wing sections to determine spatial expression differences across the wing. At Day 2
161 clustering by compartment is not evident, and there is some clustering by individual,
162 indicating that there is less differential expression across the wing at this later time
163 point.

164 Differential expression analysis revealed 209 differentially expressed genes between
165 *H. melpomene* larval forewings and hindwings, versus 77 in *H. erato*. In total, 28 of
166 these genes are differentially expressed in both species (Table S2). This includes the
167 transcription factor *Ubx*, the notch pathway repressor and microtubule binding protein
168 *pigs*, and 9 genes with no homology to known transcripts. At day 1, we identified
169 2848 genes differentially expressed between the five compartments in *H. melpomene,*
170 1713 in *H. erato* and 1780 in *A. vanillae.* 617 transcripts were differentially expressed
171 in all three species at day 1 (Table S3). At day 2, 319 transcripts were differentially
172 expressed in *H. melpomene*, 2663 in *H. erato* and 167 in *A. vanillae*, with no genes
173 differentially expressed in all three species, and 30 genes differentially expressed in a
174 pair of species (Table S4), including the pigmentation genes *Ddc* and *tan*.

175

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.



176

177 **Figure 2:** Principal component analyses of RNA samples for each species clustered
178 by stage, with the exception of three samples of *Agraulis vanillae* form the day 2
179 stage, which formed a separate cluster. D shows the dissection scheme used for tissue
180 collection; FP=Proximal Forewing, FM=Medial Forewing, FD=Distal Forewing,
181 HA=Anterior Hindwing, HP=Posterior Hindwing.

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

182 ***Transcription factors***

183 In order to investigate the nature of the regulatory landscape of the developing wing,
184 we focused on the expression of the 237 identified transcription factor orthology
185 groups. No expression was detected in any species for 37 of these genes; of the 200
186 TFs that were expressed, 6 were differentially expressed in all three species, 16 were
187 differentially expressed in 2 species, and an additional 31 were differentially
188 expressed in one species (Figure 3). This confirms the presence of TFs that are
189 expressed in a patterned way across the proximal-distal axis of the wing.

190 To examine the relationships between spatial domains of expression between the
191 species, expression profiles were clustered into 5 classes by similarity. A total of 20
192 TFs shared the same expression profile in all three species and an additional 21 shared
193 the same expression profile in two species, with an additional 10 factors showing a
194 different expression profile in all three species. Of the 37 factors which were
195 differentially expressed in either *H. melpomene* or *H. erato*, 26 (70%) were expressed
196 in the same pattern, whereas of the 27 factors differentially expressed in *Agraulis,*
197 only 13 TFs (48%) were differentially expressed in the same pattern between *Agraulis*
198 and one species of *Heliconius,* indicating that while some factors have conserved
199 ancestral expression profiles, others vary in their expression along the proximal-distal
200 axis.

201 A number of TFs have expression profiles that match known profiles either from
202 immunohistochemistry of butterfly wings or by analogy with gene expression in
203 *Drosophila* wings (Table 1). This includes *Ultrabithorax* (*Ubx*), expressed only in the
204 hindwing; *homothorax* (*hth*), expressed only in the proximal forewing and anterior
205 hindwing; *distal-less* (*dll*), expressed in an increasing gradient from proximal to
206 distal; and *mirror (mirr)*, expressed in the proximal forewing and anterior hindwing
207 (Figure 3, c.f. Table 1). The recapitulation of these expression profiles confirms the
208 presence of conserved expression of genes involved with wing pattern specification
209 between insects, and also serves as validation that our experimental design can detect
210 differential expression of transcription factors, which are typically expressed at
211 relatively low levels.

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

212 Multiple additional factors with conserved expression profiles along the proximal-
213 distal axis were identified, including *brinker,* a negative-regulator of Dpp signaling
214 (Campbell and Tomlinson, 1999); *bric a brac 2* (*bab2*), part of a proximal-distal gene
215 regulatory module linked to abdominal pigmentation pattern in *Drosophila* (Rogers et
216 al., 2013)*; ventral veins lacking (vvl),* linked to vein development in *Drosophila*, here
217 highly expressed in the proximal forewing, and previously suggested as a candidate
218 wing pattern regulator gene in *H. erato* (Van Belleghem et al., 2017, de Celis et al.,
219 1995); *Hr38,* a hormone receptor upregulated in the medial forewing; and *shaven* (*sv*),
220 related to the development of sensory structures and upregulated in the distal forewing
221 (Kavaler et al., 1999). These factors serve as additional candidate pre-pattern
222 regulators in heliconiine wings.

223 Several transcription factors are differentially expressed in all three species, but with
224 different profiles. *bunched (bun)* (Dpp pathway, (Treisman et al., 1995))*, glial cells*
225 *missing (gcm)* (related to neurogenesis, (Jones et al., 1995)) and *jun-related antigen* (a
226 transcription factor in the JNK-pathway, (Kockel et al., 1997)) are expressed in the
227 same profiles in *Heliconius* but in a different profile in *Agraulis*, whereas the
228 functionally undescribed TF *CG7786* is expressed in a similar profile in *H. erato* and
229 *Agraulis* but differently in *H. melpomene,* and the Wnt-pathway component factors
230 *sens* and *Sox102F* are differentially expressed in different patterns in *H. erato* and *H.*
231 *melpomene*. In all three species, Ecdysone-induced protein 74EF *Eip74EF* is
232 differentially expressed in a different pattern (Urness and Thummel, 1995). The
233 divergent expression of these factors in this set of species may indicate a role
234 downstream of the pre-patterning factors, for example as targets of Wnt signaling or
235 in scale cell differentiation.

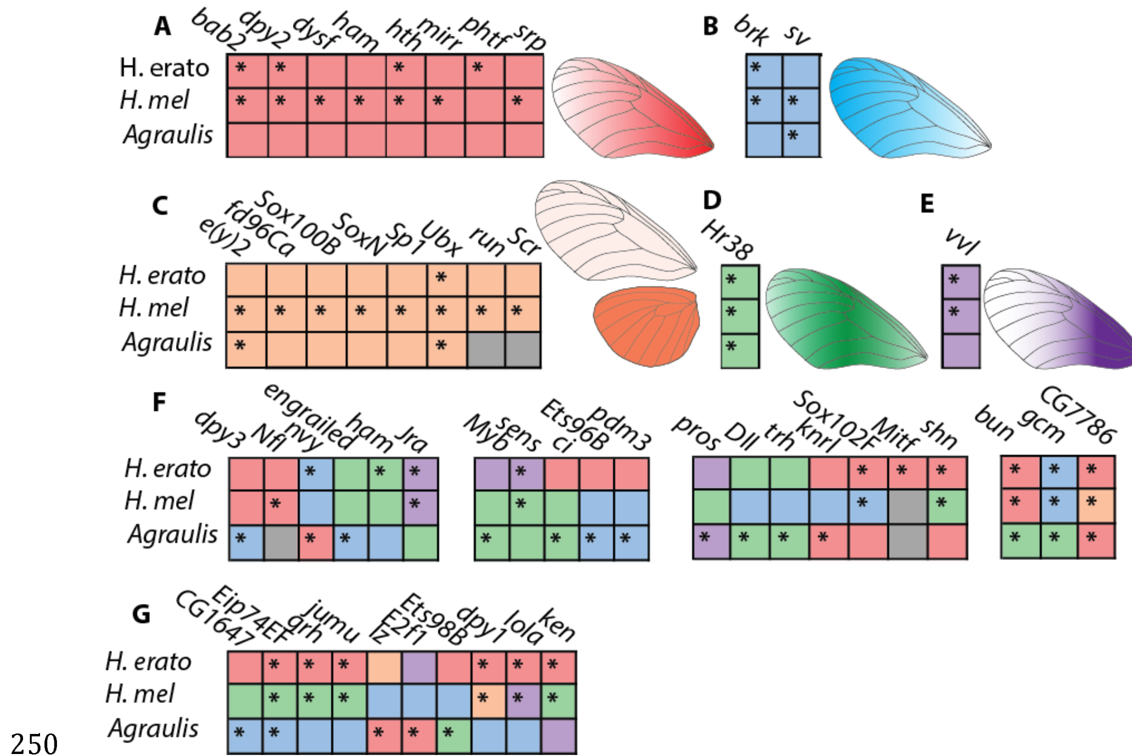
236 Several differentially expressed TFs are associated with the development of imaginal
237 discs generally, or are specifically associated with other imaginal discs, in particular
238 related to the eye and the genitals suggesting the possible evolution of novel functions
239 in the Lepidoptera. For example *bunched* (*bun*, eye development), *lozenge* (*lz,*
240 compound eye development and genital morphogenesis), *ken and barbie* (*ken,* genital
241 morphogenesis). Many differentially expressed TFs have known roles in neurogenesis
242 and the nervous system, including *senseless (sens*), *gcm*, *nervy* (*nvy*, axon guidance
243 and chetae morphogenesis). Other factors have specific associations with cuticle or

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

244 bristle development such as *nvy*, *grainy-head* (*grh*), as well as multiple copies of
245 *dumpy* (*dpy*). The factor *pdm3* is upregulated in *Agraulis* distal hindwing, and is a
246 hotspot for abdominal pigmentation evolution in *Drosophila* (Yassin et al., 2016).
247 This may implicate a novel or undescribed role for these factors in the pupal
248 development of insect wings, or specifically in the wings of butterflies.

249

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.



250

251 **Figure 3: Differential expression of transcription factors in day 1 pupae.**

252 Transcription factors are colour coded for their pattern of differential expression – A,
253 in red, indicates factors that are highly expressed in the proximal forewing and
254 expressed in a falling gradient in the medial and distal forewing; B, in blue, indicates
255 factors that are highly expressed in the distal forewing; C, in orange, indicates factors
256 that are highly expressed in the hindwing relative to the forewing; D, in green,
257 indicates factors that are highly expressed in the medial forewing; and E, in purple,
258 indicates factors that are highly expressed in the proximal forewing but low in the rest
259 of the forewing. Asterisks indicate genes which are significantly differentially
260 expressed; all depicted genes are differentially expressed in at least one species.

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

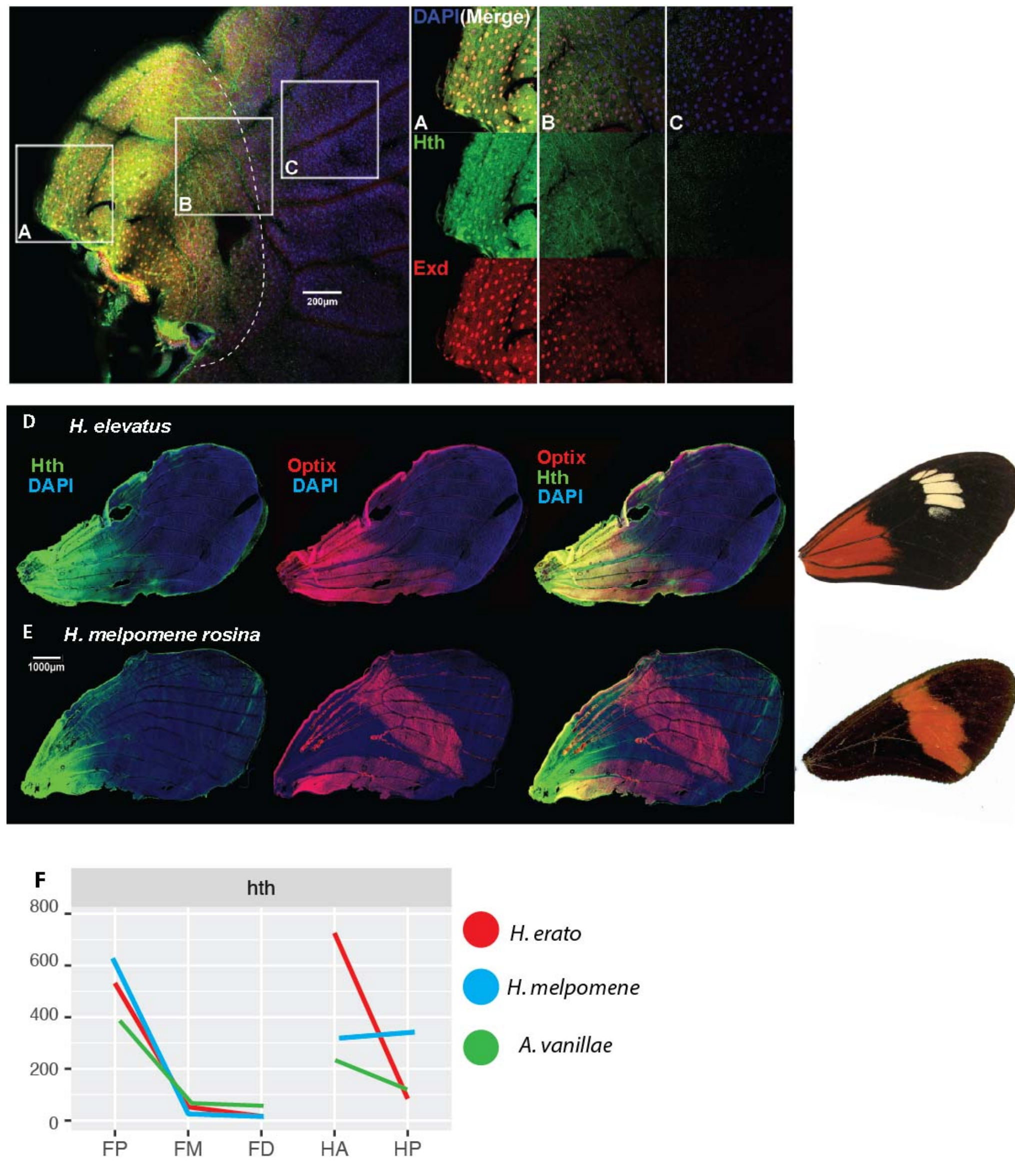
261 ***homothorax***

262 The gene *homothorax (hth)* was upregulated in the proximal forewing in all three
263 species. This replicates its previously detected expression in early pupal wings of *H.*
264 *erato* (Hines et al., 2012). To confirm patterning of the hth protein, pupal butterfly
265 wings were fluorescently stained against an anti-hth antibody raised against
266 *Drosophila* hth. The DNA binding homeodomain of *Heliconius*, *Danaus* (Monarch
267 butterfly), *Tribolium* and *Drosophila* hth are highly conserved (Figure S3). Staining
268 with anti-Hth highlighted a gradient of hth from the basal to the medial region of the
269 wing in *H. melpomene*. Expression was most strongly detected in presumptive scale
270 cell nuclei (Figure 4)

271 Some *Heliconius* wing patterns show a red patch in the proximal forewing that
272 correlates with this *hth* expression domain, known as the *dennis* patch. This patch of
273 red scales is known to be specified by the *optix* gene, and we therefore hypothesised
274 that the evolution of this patch might have arisen through a novel regulatory link
275 between *hth* and *optix*. To explore this possibility, we co-stained wings for both *hth*
276 and *optix* from taxa both with and without the *dennis* patch. In *H. elevatus* which has
277 the *dennis* phenotype, hth expression was strongly coincident with Optix. In contrast
278 in *H. melpomene rosina* which has a medial red forewing band, there was no
279 coincident expression of hth with Optix, except in the region of forewing and
280 hindwing overlap where Optix expression is known to be widely conserved among
281 butterflies. Importantly, *hth* expression was detected from the start of pupation up to
282 and during optix expression at 12-60 hours post-pupation. Together the coincident
283 timing, position and nuclear staining suggest that hth is a potential interacting factor
284 of the *dennis* enhancer in *Heliconius* butterflies. hth was not coincident with the
285 *dennis* bar in the hindwing, however, suggesting other factors are also involved
286 (Figure S4).

287

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.



288

289 **Figure 4: Immunohistochemistry shows pattern of Hth expression in the**
290 **butterfly wing is replicated by RNAseq analysis.** Immunohistochemistry confirmed
291 the expression of Homothorax in a proximal-distal gradient across the basal third of
292 the *Heliconius* wing, in larvae (A-C) and pupae (D, E) of *Heliconius* butterflies. A-C
293 highlight three regions along the proximal-distal axis of the larval wing, showing
294 coincident expression of Homothorax and its co-factor Extradenticle. D shows
295 Homothorax expression in a region coincident with the expression of Optix in a
296 *dennis-ray* butterfly, *Heliconius elevatus*. The same expression pattern of Hth is
297 conserved in a red-banded butterfly (E), but is not associated with Optix expression.
298 All *Heliconius* show Optix expression in the overlapping fore and hindwing region,

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license.

299 associated with wing coupling scales as documented previously (Martin et al, 2014).
300 The expression profile observed in pupal wings here recapitulates the levels of *hth*
301 transcript observed in the RNAseq analysis of all three species examined here (F).

302

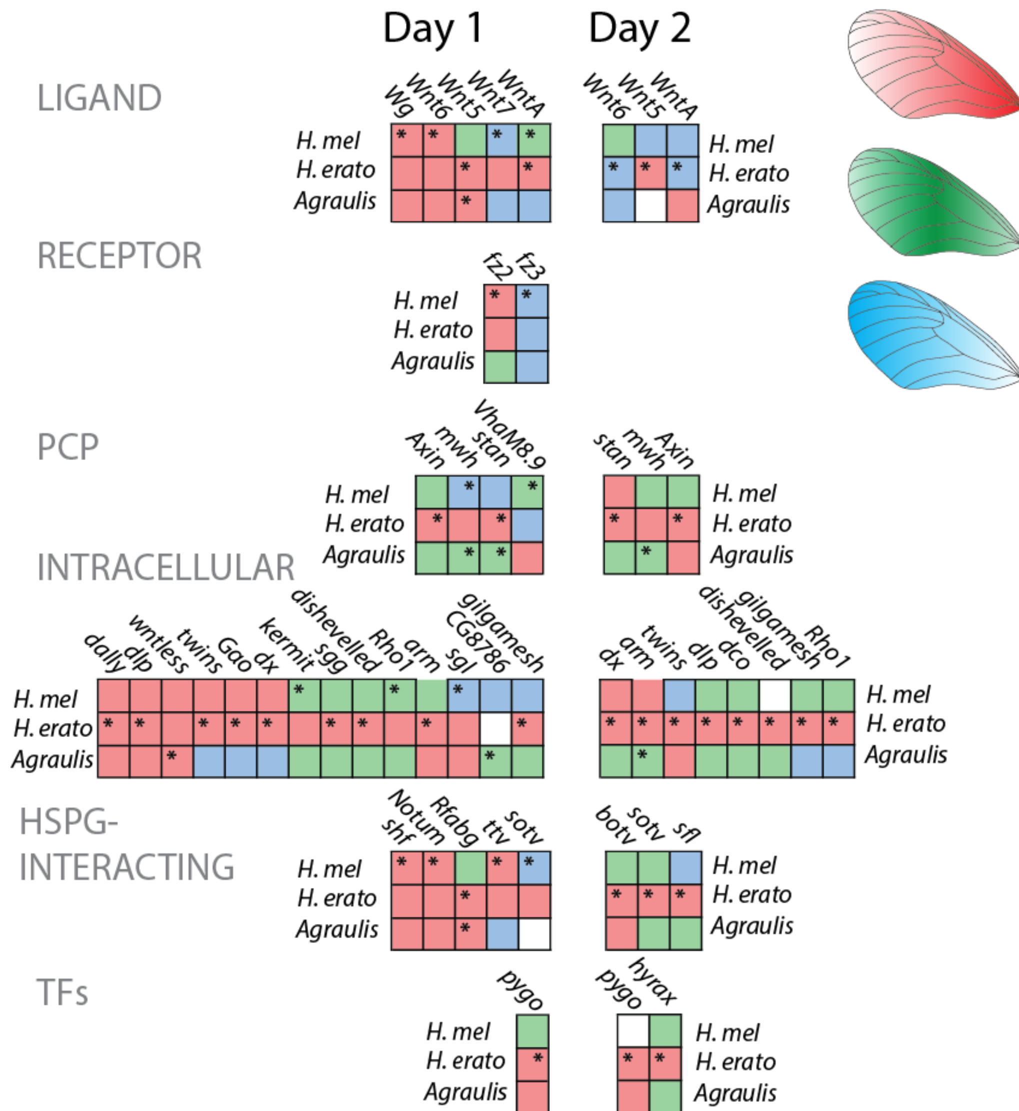
bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

303 ***Wnt pathway***

304 Recent studies have highlighted the importance of the ligand WntA in butterfly wing
305 patterning, so we next focused on the expression domains of other Wnt pathway
306 constituents. We identified 52 Wnt pathway constituents in the genomes of our three
307 species (Table S5). Expression profiles were split into three groups based on
308 similarity. In *H. erato* pupal wings, most Wnt pathway genes showed a very similar
309 pattern of highest expression in the proximal forewing and lower expression in the
310 medial and distal forewing, whereas in *H. melpomene* and *A. vanillae,* a variety of
311 different expression profiles for these genes were observed (Figure 5A). The variance
312 in Wnt pathway gene expression is reflected in the prototypical Wnt target
313 transcription factor *senseless* (Figure 3). In contrast to the transcription factors, of the
314 32 genes differentially expressed at day 1, only 8 (25%) had shared expression
315 profiles in all three species, with an additional 4 (12.5%) sharing a profile between the
316 *Heliconius* species. At day 2, further divergence in expression profiles between
317 species occurred, with none of 19 differentially expressed genes sharing an expression
318 profile between all three species.

319 The ligand WntA was differentially expressed in the *Heliconius* species in profiles
320 that correlate to that observed in Martin et al (2012), and multiple other Wnt ligands
321 were differentially expressed in the three species, including Wnt2 and Wnt6 in
322 *Heliconius,* which in other Nymphalid species have redundant expression profiles and
323 have been correlated with the discalis pattern elements, and which was also
324 previously reported in *H. erato* (Martin et al., 2014, Hines et al., 2012). Two Wnt
325 receptors, *fz2* and *fz3*, are expressed in opposition to each other; fz2 is the primary Wg
326 receptor in *Drosophila*, while fz3 plays an inhibitory role (Bhanot et al., 1996, Sato et
327 al., 1999). Wnt pathway components involved with planar cell polarity including
328 *multiple wing hairs (mwh)*, *starry night (stan),* and *Axin*, were differentially expressed
329 at both stages, along with other intracellular components of the Wnt pathway
330 including *armadillo/* β*-catenin* and *wntless*, a transmembrane factor required for Wnt
331 ligand secretion. In *H. erato,* two TFS downstream of Wnt signaling were
332 differentially expressed; *pygo* and *hyrax*.

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.



333

334 **Figure 5: Differential expression of Wnt pathway components in pupal**
335 **development.** Wnt pathway components are colour-coded for their pattern of
336 expression. Red indicates transcripts that were highly expressed in the proximal
337 forewing, green indicates transcripts that were highly expressed in the medial
338 forewing and blue indicates transcripts that were highly expressed in the distal
339 forewing. Asterisks indicate transcripts that were identified as significantly
340 differentially expressed. Note the low discordance of expression profile between
341 species, in contrast to the transcription factors indicated in Figure 3. HSPG = heparin
342 sulfate proteoglycan

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

343 **Discussion**

344 We have explored the patterns of gene expression both through development, across
345 evolutionary divergence and across the developing wing. Our results paint a vivid
346 picture of how wing patterns develop and evolve across the three butterfly species.
347 Broadly, *H. melpomene*, *H. erato* and *A. vanillae* share a common spatial
348 transcriptomic landscape in the developing wing with other Lepidoptera and with
349 *Drosophila*, implying the existence of a shared insect wing gene regulatory network
350 (cf. Table 1). The results for *H. erato* here also broadly replicate those from a
351 previous transcriptomic analysis there(Hines et al., 2012). However, our data also
352 highlight considerable flexibility in the transcriptional landscape. More than half of
353 the transcription factors that are differentially expressed across the developing wing
354 surface have different expression profiles within the heliconiines. This flexibility is
355 most evident within the Wnt signaling pathway constituents, where *H. erato* has a
356 derived pattern of strongly coordinated Wnt pathway gene expression, that is not seen
357 in the other two species, despite strong convergence in pattern between the two
358 *Heliconius* species studied. Recent work has shown genome-wide selection on
359 regulatory elements at the between-population level in *Heliconius (Lewis and Reed,*
360 *2018)*, and it is likely that in the ancestral linages of each species, many functional
361 changes could be accrued that would lead to many differences in patterns of gene
362 expression in the wing.

363 ***Transcription factors expressed in butterfly wings***

364 Transcription factors provide the physical interactions that lead to differential
365 regulation in gene regulatory networks. Several transcription factors that are known to
366 be involved in development of wings in *Drosophila* and *Junonia* were identified in
367 this experiment in their expected expression profiles, including *Ubx* and *hth*. Several
368 other factors were expressed in similar patterns in all three species in this experiment;
369 these additional factors could delineate the developmental morphospace along the
370 forewing proximal-distal axis and hindwing anteroposterior axis in pupal wings. We
371 identified an additional cohort of transcription factors with non-conserved expression
372 profiles between the three studied species; many more of these factors had shared
373 expression profiles between *H. melpomene* and *erato* than between *Agraulis* and
374 Heliconius. The two Heliconius species are more closely related to one another, but

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

375 are also convergent in their wing patterns, so we cannot currently disentangle whether
376 the share expression patterns are due to common ancestry or are convergent due to
377 shared selection pressures. Others were different in all three species, implying
378 developmental drift, or a lack of constraint, on the regulation of these factors. Such
379 factors have the potential to act as the substrate for functional diversification.

380 ***hth implicated in mimetic pattern evolution***

381 One of the most strongly and consistently differentiated transcription factors was hth,
382 and we therefore followed up on the expression patterns using immunohistochemistry.
383 This confirmed that the hth protein shows a conserved pattern of expression restricted
384 to the proximal wing region in all species examined. Furthermore, co-staining with
385 anti-optix demonstrated a strong correlation of hth with the expression of optix
386 protein in butterflies with the red proximal *dennis* patch. Expression of both proteins
387 was localized to scale cell nuclei, and spatial patterns were tightly correlated between
388 the two factors. In contrast, butterflies lacking the dennis patch showed a conserved
389 expression of hth, but no correlated expression of optix. hth a candidate regulator of
390 optix in dennis+ butterflies. A possible mechanism for the evolution of the dennis
391 pattern is therefore that an *optix* regulatory region gained transcription factor binding
392 sites for hth, allowing the development of a novel pattern without the requirement for
393 changes to the expression or function of hth itself, similar to the roles of *en* and *sr* in
394 the development of melanic patterns in *Drosophila* (Gompel et al., 2005, Gibert et al.,
395 2018). Alternatively, hth might regulate optix through intermediate factors, and a
396 number of other candidates upregulated in the proximal region are evident from our
397 results. Future analyses of the dennis regulatory element will be required to
398 determine the precise mechanisms of interaction with upstream regulators.

399 ***Wnt pathway variance implies different functions***

400 Variance in *WntA* expression in correlation with wing pattern has previously been
401 shown in many butterfly clades, including between races and species of *Heliconius*,
402 and in *Agraulis* (Martin and Reed, 2014, Martin et al., 2012). We found that other
403 Wnt pathway constituents also vary in their expression domains between species.

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

404 Surprisingly, expression patterns of Wnt pathway constituents were completely
405 different between the two *Heliconius* co- mimics. In particular, the Wnt pathway
406 constituents in *H. erato* were mainly expressed in a correlated pattern – high in the
407 proximal forewing, and low everywhere else, whereas the same factors were
408 expressed in a variety of patterns in *Agraulis* and *H. melpomene*.

409 Notably, the correlated pattern in *H. erato* closely mirrors the expression profile of
410 *WntA* in larval wing discs of *H. erato demophoon* from Panama, the pattern form used
411 in this study (Martin et al., 2012), while the *WntA* expression profile for Panamanian
412 *H. melpomene rosina* is notably different from its co-mimic. Here, expression is
413 present in the distal forewing as well as the proximal forewing, and the boundary of
414 proximal *WntA* expression does not correlate well with the proximal boundary of the
415 red pattern element in the adult wing.

416 Together, these differences in both WntA function and Wnt pathway component
417 expression could suggest that regulatory diversification of the Wnt pathway has
418 occurred between *Heliconius* species, implying that this aspect of the wing gene
419 regulatory network has diverged in the lineages leading to these mimetic forms,
420 requiring the utilization of different functional mechanisms for building an identical
421 wing pattern. This suggests extensive genome-wide regulatory divergence at the
422 between-species level between *Heliconius* butterflies. Alternatively, the differences in
423 expression of Wnt pathway components could entirely be a consequence of the
424 different expression profiles of WntA between the two species. Wnt ligands in
425 *Drosophila* effect their own autoregulation through modulation of receptor, ligand
426 and transcription factor expression levels (Cadigan et al., 1998, Schilling et al., 2014).
427 If WntA is capable of directly regulating the expression of Wnt pathway components,
428 this could explain all the differences we observed here while requiring between-
429 species regulatory divergence at just one genomic locus. Separating these two models
430 will require the functional exploration of the effects of WntA signaling.

431 **Conclusion**

432 Our understanding of the regulatory evolution of wing pattern in butterflies is
433 dependent on a clear picture of the expression of developmental factors around the
434 time of wing pattern specification. This study has provided a picture of gene

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

435 expression along one axis of developing wings in a manner unbiased by our
436 understanding of wing development in non-lepidopteran systems. At the within-
437 species level, we can broadly rule out the hypothesis that *trans*-regulatory factors
438 change their expression profiles in different pattern forms (i.e. Figure 1D) based on
439 genetic mapping, but we are not able to rule out this phenomenon at the between-
440 species level – it is likely that both processes play a role, either through selection or
441 drift. Our deeper understanding of factors that are expressed in the wing in correlation
442 with pattern elements will permit us to decode the regulatory linkages that lead to the
443 differential expression of pattern switch genes like *optix, WntA* and *cortex*, and it is
444 clear that we should look to both conserved and diverging regulatory factors as the
445 causative agents of *cis*-regulatory evolution.

446

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

447 **Methods**

448 ***Tissue sampling and dissection***

449 *Heliconius melpomene rosina* and *Heliconius erato demophoon* were collected from
450 stocks maintained at the Smithsonian Tropical Research Institute in Gamboa, Panama
451 in February and July 2014. Adults were provided with an artificial diet of
452 pollen/glucose solution supplemented with flowers of *Psiguria*, *Lantana* and/or
453 *Psychotria alata* according to availability. Females were provided with *Passiflora*
454 plants for egg laying (*P. menispermifolia* for *H. melpomene*, *P. biflora* for *H. erato*).
455 Eggs were collected daily, and caterpillars reared on fresh shoots of *P. williamsi*
456 (*melpomene*) or *P. biflora* (*erato)* until late 5th (final) instar, when they were
457 separated into individual pots in a temperature-monitored room, and closely observed
458 for the purpose of accurate developmental staging. *Agraulis vanillae* larvae were
459 collected from *P. edulis* located near the insectary in March 2014.

460 Pre-pupation larvae were identified for dissection. Late 5th instar larvae of *Heliconius*
461 undergo colour changes from white to purple on the last larval day, followed by an
462 additional change to pink-orange in the hours before pupation. Additionally, several
463 behavioural changes accompany the pre-pupation period; the larvae stop eating and
464 clear their digestive tract, then undertake a period of rapid locomotion and wandering
465 until they find an appropriate perch for pupation - preferably the underside of a leaf or
466 a sturdy twig - at which point they settle in place and produce a strong silk
467 attachment. Gradually, over a period of 30-120 minutes, they suspend themselves
468 form their perch in a J shape and then pupate. Larvae that were post-locomotion but
469 pre-J-shape were dissected in cold PBS and the wing discs removed.

470 Pupae were allowed to develop until 36h (+/- 1.5h), or to 60h (+/- 1.5h). These time
471 points are referred to as Day 1 and Day 2 throughout. In the hours immediately post-
472 pupation (Day 0), the pupal carapace is soft and the membranous structures of the
473 pupa are thin, weak, transparent and sticky, hence the effective dissection of unfixed,
474 intact pupal structures is very challenging at the earliest pupal time points.

475 Pupae were dissected in cold PBS. Wings were removed from the pupa and cleared of
476 peripodial membrane. The wings were then cut with microdissection scissors into 5

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

477 sections: forewing proximal, medial and distal, and hindwing anterior and posterior
478 (figure 2). The lacunae (developing veins) were used as landmarks for dissection.

479 Whole larval wing discs and pupal wing sections were transferred into RNAlater
480 (ThermoFisher, Waltham, MA) and kept on liquid nitrogen in Gamboa, then
481 transported to the UK on dry ice, and transferred to -80 C on arrival in Cambridge.

482 ***RNA extraction and sequencing***

483 RNA extraction was carried out using a standard hybrid protocol. Briefly, wing tissue
484 sections were transferred into Trizol Reagent (Invitrogen, Carlsbad, CA) and
485 disassociated using stainless steel beads in a tissue lyser. Chloroform phase extraction
486 was performed, followed by purification with the RNeasy kit (Qiagen, Valencia CA).
487 RNA was eluted into distilled water and treated with DNAseI (Ambion, Naugatuck
488 CT), then quantified and stored at -80 C. Left and right wings and wing sections were
489 pooled.

490 cDNA synthesis, library preparation and sequencing were carried out by Beijing
491 Genomics Institute (Beijing, China). Samples were sequenced at either 75 PE on
492 Illumina HiSeq 3000 or at 150 PE on Illumina HiSeq 4000.

493 ***Transcriptome assembly - Agraulis***

494 All paired end sequence data for *Agraulis* was assembled with the transcriptome
495 assembler Trinity (Haas et al., 2013). This generated 87,214 contigs. Next the Trinity
496 output was passed through TransDecoder (Haas & Papanicolaou, in prep), which
497 annotates the transcript contigs based on the likelihood that they contain reading
498 frames, and also based on similarity by BLAST of transcripts to reference assemblies,
499 in this case *H. erato* and *H. melpomene*. This annotation (a GFF3 annotation of the
500 Trinity contigs) contained 24,984 genes, which compares to 20,102 annotated genes
501 in *H. melpomene v2.1* and 13,676 in *H. erato v1*.

502 ***Mapping and quantification***

503 Reads were aligned with Hisat2 aligner to the genome of the respective species
504 (Rutledge et al., 2006, Kim et al., 2015). The highest percentage of unique mappings

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

505 was achieved using default parameters (Figure 3.4). Alignments were then quantified
506 using GFF annotations of each genome with HTSeqCount, union mode (Anders et al.,
507 2015). Genomes and annotations are publicly available at www.lepbase.org (Challis
508 et al., 2016).

509 ***Data analysis***

510 Statistical analysis of counts was carried out using the R package DESeq2 (Love et
511 al., 2014) using the following generalized linear model (GLM):

512 ~ *individual* + *compartment*

513 In larvae, the compartments were Forewing (FW) and Hindwing (HW). In pupae, the
514 compartments were as follows: Proximal Forewing (FP), Medial Forewing (FM),
515 Distal Forewing (FD), Anterior Hindwing (HA), Posterior Hindwing (HPo) (figure
516 2)).

517 ***Determining homology***

518 Homology between differentially expressed genes in the three species was determined
519 in two ways. First, a small percentage of genes have been assigned homologs by
520 comparison with other lepidopteran genomes on LepBase using InterProScan (Jones
521 et al., 2014). This allowed recovery of one-to-one homologs and gene families with
522 distinct insect lineages. However in a number of cases where similar copies of a gene
523 are present, for example the Wnt ligands, some genes were assigned to the incorrect
524 orthogroup and were manually curated. For the rest of the genes, as well as all genes
525 in *Agraulis*, amino acid sequences were reciprocally searched with BLASTp, and the
526 top hit was taken as the homolog (Altschul et al., 1997). Genes with no assigned
527 orthogroup were compared by pBLAST against the polypeptide library of *D.*
528 *melanogaster* genes retrieved from FlyBase, associating them with a FBpp number
529 and gene code based on homology with *Drosophila* genes (Gramates et al., 2017).

530 ***Immunohistochemistry***

531 Pupae were dissected 60-80 hrs after pupation in chilled PBS, or at an estimated 12hrs
532 before pupation for final instar larvae. Wings were fixed in 4% MeOH-free

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

533 formaldehyde (Pierce) in PBSTw (PBS plus 0.01% Tween-20 (Sigma)) on ice for
534 40mins. They were then washed and permeabilized 6 x 5mins in PBSTx (PBS plus
535 0.5% Triton-X (Sigma)) and blocked for 2 hrs room temperature in PBSTx plus 5%
536 goat serum (Sigma). Rabbit anti-Homothorax antibody (gift from Prof. Adi Salzberg,
537 Technion-Israel Institute of Technology) or Rat anti-Optix (gift from Prof. Robert
538 Reed, Cornell University) antibody was diluted to 1/1000 in PBSTx plus 0.5% goat
539 serum and applied to wings overnight at 4° C. Wings were washed 6x5mins in PBSTx
540 and incubated with 1/1000 goat anti-rabbit alexa-488 conjugated antibody (Abcam) in
541 PBSTx for 3hrs at room temperature. Wings were washed 4x5mins with PBSTx and
542 incubated 10mins in 1ug/ml DAPI (ThermoScientific) PBSTx. Wings were mounted
543 in Fluoromount-G (SouthernBiotech) and imaged using a Leica DM6000B SP5
544 confocal microscope. An average of 60-80 images were taken to cover the entire
545 wing, each a stack of 40 2 μm thick slices and composed of three channels: 408 nm,
546 488 nm, and 568 nm. Each stack was converted to single images of maximum
547 intensity using FIJI ImageJ 1.47m, and then split by channel. Images were then
548 compiled into a single image and formatted using Adobe Photoshop CS5.1.

549 Declarations

550 **Data availability:** The datasets generated and/or analysed during the current study are
551 available in the SRA repository, [PERSISTENT WEB LINK TO DATASETS]

552 **Ethics Approval and Consent to Participate:** Not applicable

553 **Consent for Publication:** Not applicable.

554 **Conflicts of interest:** The authors declare no conflict of interest.

555 **Funding:** This research was funded by a PhD Studentship from the Wellcome Trust to JJH, European
556 Research Council grant no. 339873 and Leverhulme trust grant no RPG-2014-167 to RWRW.

557 **Author contributions:** JJH, WOM and CDJ planned the research and wrote the paper, JJH performed
558 experiments and analysis, RWRW performed immunofluorescence and imaging.

559 **Acknowledgements:** We thank Lucas Brenes, Henry Arenas-Castro, Oscar Paneso and Elizabeth
560 Evans for assistance with insect rearing and tissue dissection, and the Beijing Genomics Institute for
561 library preparation and sequencing. We thank Arnaud Martin for comments on an earlier draft of this
562 text. We thank Adi Salzberg for use of the Hth antibody. We thank the Smithsonian Tropical Research
563 Institute for support for tissue collection and work with butterflies in Panama, and ANAM in Panama
564 for permission to collect butterflies.
565

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

566 References
567

568 ALTSCHUL, S. F., MADDEN, T. L., SCHAFFER, A. A., ZHANG, J., ZHANG, Z., MILLER,
569 W. & LIPMAN, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation
570 of protein database search programs. *Nucleic Acids Res,* 25**,** 3389-402.
571 ANDERS, S., PYL, P. T. & HUBER, W. 2015. HTSeq--a Python framework to work
572 with high-throughput sequencing data. *Bioinformatics,* 31**,** 166-9.
573 ARNOULT, L., SU, K. F., MANOEL, D., MINERVINO, C., MAGRINA, J., GOMPEL, N. &
574 PRUD'HOMME, B. 2013. Emergence and diversification of fly
575 pigmentation through evolution of a gene regulatory module. *Science,*
576 339**,** 1423-6.
577 BHANOT, P., BRINK, M., SAMOS, C. H., HSIEH, J. C., WANG, Y., MACKE, J. P.,
578 ANDREW, D., NATHANS, J. & NUSSE, R. 1996. A new member of the
579 frizzled family from Drosophila functions as a Wingless receptor. *Nature,*
580 382**,** 225-30.
581 CADIGAN, K. M., FISH, M. P., RULIFSON, E. J. & NUSSE, R. 1998. Wingless
582 repression of Drosophila frizzled 2 expression shapes the Wingless
583 morphogen gradient in the wing. *Cell,* 93**,** 767-77.
584 CAMPBELL, G. & TOMLINSON, A. 1999. Transducing the Dpp morphogen
585 gradient in the wing of Drosophila: regulation of Dpp targets by brinker.
586 *Cell,* 96**,** 553-62.
587 CARROLL, S. B., GATES, J., KEYS, D. N., PADDOCK, S. W., PANGANIBAN, G. E.,
588 SELEGUE, J. E. & WILLIAMS, J. A. 1994. Pattern formation and eyespot
589 determination in butterfly wings. *Science,* 265**,** 109-14.
590 CHALLIS, R. J., KUMAR, S., DASMAHAPATRA, K. K. K., JIGGINS, C. D. & BLAXTER,
591 M. 2016. Lepbase: the Lepidopteran genome database. *BioRxiv***,** 056994.
592 DE CELIS, J. F., LLIMARGAS, M. & CASANOVA, J. 1995. Ventral veinless, the gene
593 encoding the Cf1a transcription factor, links positional information and
594 cell differentiation during embryonic and imaginal development in
595 Drosophila melanogaster. *Development,* 121**,** 3405-16.
596 ENCISO-ROMERO, J., PARDO-DIAZ, C., MARTIN, S. H., ARIAS, C. F., LINARES, M.,
597 MCMILLAN, W. O., JIGGINS, C. D. & SALAZAR, C. 2017. Evolution of novel
598 mimicry rings facilitated by adaptive introgression in tropical butterflies.
599 *Molecular Ecology,* 26**,** 5160-5172.
600 GALANT, R., SKEATH, J. B., PADDOCK, S., LEWIS, D. L. & CARROLL, S. B. 1998.
601 Expression pattern of a butterfly achaete-scute homolog reveals the
602 homology of butterfly wing scales and insect sensory bristles. *Curr Biol,* 8**,**
603 807-13.
604 GALLANT, J. R., IMHOFF, V. E., MARTIN, A., SAVAGE, W. K., CHAMBERLAIN, N. L.,
605 POTE, B. L., PETERSON, C., SMITH, G. E., EVANS, B., REED, R. D.,
606 KRONFORST, M. R. & MULLEN, S. P. 2014. Ancient homology underlies
607 adaptive mimetic diversity across butterflies. *Nat Commun,* 5**,** 4817.
608 GIBERT, J. M., MOUCHEL-VIELH, E. & PERONNET, F. 2018. Pigmentation pattern
609 and developmental constraints: flight muscle attachment sites delimit the
610 thoracic trident of Drosophila melanogaster. *Sci Rep,* 8**,** 5328.
611 GOMPEL, N., PRUD'HOMME, B., WITTKOPP, P. J., KASSNER, V. A. & CARROLL, S. B.
612 2005. Chance caught on the wing: cis-regulatory evolution and the origin
613 of pigment patterns in Drosophila. *Nature,* 433**,** 481-7.

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

614 GRAMATES, L. S., MARYGOLD, S. J., SANTOS, G. D., URBANO, J. M., ANTONAZZO,
615 G., MATTHEWS, B. B., REY, A. J., TABONE, C. J., CROSBY, M. A., EMMERT, D.
616 B., FALLS, K., GOODMAN, J. L., HU, Y., PONTING, L., SCHROEDER, A. J.,
617 STRELETS, V. B., THURMOND, J., ZHOU, P. & THE FLYBASE, C. 2017.
618 FlyBase at 25: looking to the future. *Nucleic Acids Res,* 45**,** D663-D671.
619 HAAS, B. J., PAPANICOLAOU, A., YASSOUR, M., GRABHERR, M., BLOOD, P. D.,
620 BOWDEN, J., COUGER, M. B., ECCLES, D., LI, B., LIEBER, M., MACMANES, M.
621 D., OTT, M., ORVIS, J., POCHET, N., STROZZI, F., WEEKS, N., WESTERMAN,
622 R., WILLIAM, T., DEWEY, C. N., HENSCHEL, R., LEDUC, R. D., FRIEDMAN, N.
623 & REGEV, A. 2013. De novo transcript sequence reconstruction from RNA-
624 seq using the Trinity platform for reference generation and analysis. *Nat*
625 *Protoc,* 8**,** 1494-512.
626 HINES, H. M., PAPA, R., RUIZ, M., PAPANICOLAOU, A., WANG, C., NIJHOUT, H. F.,
627 MCMILLAN, W. O. & REED, R. D. 2012. Transcriptome analysis reveals
628 novel patterning and pigmentation genes underlying Heliconius butterfly
629 wing pattern variation. *BMC Genomics,* 13**,** 288.
630 IIJIMA, T., KAJITANI, R., KOMATA, S., LIN, C. P., SOTA, T., ITOH, T. & FUJIWARA, H.
631 2018. Parallel evolution of Batesian mimicry supergene in two Papilio
632 butterflies, P. polytes and P. memnon. *Sci Adv,* 4**,** eaao5416.
633 JIGGINS, C. D., WALLBANK, R. W. & HANLY, J. J. 2017. Waiting in the wings: what
634 can we learn about gene co-option from the diversification of butterfly
635 wing patterns? *Philos Trans R Soc Lond B Biol Sci,* 372.
636 JONES, B. W., FETTER, R. D., TEAR, G. & GOODMAN, C. S. 1995. glial cells missing:
637 a genetic switch that controls glial versus neuronal fate. *Cell,* 82**,** 1013-23.
638 JONES, P., BINNS, D., CHANG, H. Y., FRASER, M., LI, W., MCANULLA, C.,
639 MCWILLIAM, H., MASLEN, J., MITCHELL, A., NUKA, G., PESSEAT, S., QUINN,
640 A. F., SANGRADOR-VEGAS, A., SCHEREMETJEW, M., YONG, S. Y., LOPEZ, R.
641 & HUNTER, S. 2014. InterProScan 5: genome-scale protein function
642 classification. *Bioinformatics,* 30**,** 1236-40.
643 KAVALER, J., FU, W., DUAN, H., NOLL, M. & POSAKONY, J. W. 1999. An essential
644 role for the Drosophila Pax2 homolog in the differentiation of adult
645 sensory organs. *Development,* 126**,** 2261-72.
646 KEYS, D. N., LEWIS, D. L., SELEGUE, J. E., PEARSON, B. J., GOODRICH, L. V.,
647 JOHNSON, R. L., GATES, J., SCOTT, M. P. & CARROLL, S. B. 1999.
648 Recruitment of a hedgehog regulatory circuit in butterfly eyespot
649 evolution. *Science,* 283**,** 532-4.
650 KIM, D., LANGMEAD, B. & SALZBERG, S. L. 2015. HISAT: a fast spliced aligner
651 with low memory requirements. *Nat Methods,* 12**,** 357-60.
652 KOCH, P. B., MERK, R., REINHARDT, R. & WEBER, P. 2003. Localization of
653 ecdysone receptor protein during colour pattern formation in wings of
654 the butterfly Precis coenia (Lepidoptera: Nymphalidae) and co-expression
655 with Distal-less protein. *Dev Genes Evol,* 212**,** 571-84.
656 KOCKEL, L., ZEITLINGER, J., STASZEWSKI, L. M., MLODZIK, M. & BOHMANN, D.
657 1997. Jun in Drosophila development: redundant and nonredundant
658 functions and regulation by two MAPK signal transduction pathways.
659 *Genes Dev,* 11**,** 1748-58.
660 KOZAK, K. M., WAHLBERG, N., NEILD, A. F., DASMAHAPATRA, K. K., MALLET, J. &
661 JIGGINS, C. D. 2015. Multilocus species trees show the recent adaptive
662 radiation of the mimetic heliconius butterflies. *Syst Biol,* 64**,** 505-24.

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

663 KUNTE, K., ZHANG, W., TENGER-TROLANDER, A., PALMER, D. H., MARTIN, A.,
664 REED, R. D., MULLEN, S. P. & KRONFORST, M. R. 2014. doublesex is a
665 mimicry supergene. *Nature,* 507**,** 229-32.
666 LEWIS, J. J. & REED, R. D. 2018. Genome-wide regulatory adaptation shapes
667 population-level genomic landscapes in Heliconius. *Mol Biol Evol*.
668 LIVRAGHI, L., ARNAUD MARTIN, MELANIE GIBBS, NORA BRAAK, SAAD ARIF &
669 BREUKER., C. J. 2017. "CRISPR/Cas9 as the Key to Unlocking the Secrets of
670 Butterfly Wing Pattern Development and Its Evolution.". *Advances in*
671 *Insect Physiology*.
672 LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change
673 and dispersion for RNA-seq data with DESeq2. *Genome biology,* 15**,** 550.
674 MACDONALD, W. P., MARTIN, A. & REED, R. D. 2010. Butterfly wings shaped by a
675 molecular cookie cutter: evolutionary radiation of lepidopteran wing
676 shapes associated with a derived Cut/wingless wing margin boundary
677 system. *Evol Dev,* 12**,** 296-304.
678 MARTIN, A., MCCULLOCH, K. J., PATEL, N. H., BRISCOE, A. D., GILBERT, L. E. &
679 REED, R. D. 2014. Multiple recent co-options of Optix associated with
680 novel traits in adaptive butterfly wing radiations. *Evodevo,* 5**,** 7.
681 MARTIN, A. & ORGOGOZO, V. 2013. The Loci of repeated evolution: a catalog of
682 genetic hotspots of phenotypic variation. *Evolution,* 67**,** 1235-50.
683 MARTIN, A., PAPA, R., NADEAU, N. J., HILL, R. I., COUNTERMAN, B. A., HALDER, G.,
684 JIGGINS, C. D., KRONFORST, M. R., LONG, A. D., MCMILLAN, W. O. & REED,
685 R. D. 2012. Diversification of complex butterfly wing patterns by repeated
686 regulatory evolution of a Wnt ligand. *Proc Natl Acad Sci U S A,* 109**,** 12632-
687 7.
688 MARTIN, A. & REED, R. D. 2010. Wingless and aristaless2 define a developmental
689 ground plan for moth and butterfly wing pattern evolution. *Mol Biol Evol,*
690 27**,** 2864-78.
691 MARTIN, A. & REED, R. D. 2014. Wnt signaling underlies evolution and
692 development of the butterfly wing pattern symmetry systems. *Dev Biol,*
693 395**,** 367-78.
694 MAZO-VARGAS, A., CONCHA, C., LIVRAGHI, L., MASSARDO, D., WALLBANK, R. W.
695 R., ZHANG, L., PAPADOR, J. D., MARTINEZ-NAJERA, D., JIGGINS, C. D.,
696 KRONFORST, M. R., BREUKER, C. J., REED, R. D., PATEL, N. H., MCMILLAN,
697 W. O. & MARTIN, A. 2017. Macroevolutionary shifts of WntA function
698 potentiate butterfly wing-pattern diversity. *Proc Natl Acad Sci U S A,* 114**,**
699 10701-10706.
700 MONTEIRO, A., GLASER, G., STOCKSLAGER, S., GLANSDORP, N. & RAMOS, D.
701 2006. Comparative insights into questions of lepidopteran wing pattern
702 homology. *BMC Dev Biol,* 6**,** 52.
703 NADEAU, N. J., PARDO-DIAZ, C., WHIBLEY, A., SUPPLE, M. A., SAENKO, S. V.,
704 WALLBANK, R. W., WU, G. C., MAROJA, L., FERGUSON, L., HANLY, J. J.,
705 HINES, H., SALAZAR, C., MERRILL, R. M., DOWLING, A. J., FFRENCH-
706 CONSTANT, R. H., LLAURENS, V., JORON, M., MCMILLAN, W. O. & JIGGINS,
707 C. D. 2016. The gene cortex controls mimicry and crypsis in butterflies
708 and moths. *Nature,* 534**,** 106-10.
709 PATEL, N. H., KORNBERG, T. B. & GOODMAN, C. S. 1989a. Expression of engrailed
710 during segmentation in grasshopper and crayfish. *Development,* 107**,** 201-
711 12.

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

712 PATEL, N. H., MARTIN-BLANCO, E., COLEMAN, K. G., POOLE, S. J., ELLIS, M. C.,
713 KORNBERG, T. B. & GOODMAN, C. S. 1989b. Expression of engrailed
714 proteins in arthropods, annelids, and chordates. *Cell,* 58**,** 955-68.
715 PRAKASH, A. & MONTEIRO, A. 2017. apterous A Specifies Dorsal Wing Patterns
716 And Sexual Traits In Butterflies. *bioRxiv***,** 131011.
717 PRUD'HOMME, B., GOMPEL, N. & CARROLL, S. B. 2007. Emerging principles of
718 regulatory evolution. *Proceedings of the National Academy of Sciences of*
719 *the United States of America,* 104**,** 8605-8612.
720 REED, R. D. 2004. Evidence for Notch-mediated lateral inhibition in organizing
721 butterfly wing scales. *Dev Genes Evol,* 214**,** 43-6.
722 REED, R. D., PAPA, R., MARTIN, A., HINES, H. M., COUNTERMAN, B. A., PARDO-
723 DIAZ, C., JIGGINS, C. D., CHAMBERLAIN, N. L., KRONFORST, M. R., CHEN, R.,
724 HALDER, G., NIJHOUT, H. F. & MCMILLAN, W. O. 2011. optix drives the
725 repeated convergent evolution of butterfly wing pattern mimicry. *Science,*
726 333**,** 1137-41.
727 REED, R. D. & SERFAS, M. S. 2004. Butterfly wing pattern evolution is associated
728 with changes in a Notch/Distal-less temporal pattern formation process.
729 *Curr Biol,* 14**,** 1159-66.
730 ROGERS, W. A., SALOMONE, J. R., TACY, D. J., CAMINO, E. M., DAVIS, K. A., REBEIZ,
731 M. & WILLIAMS, T. M. 2013. Recurrent modification of a conserved cis-
732 regulatory element underlies fruit fly pigmentation diversity. *PLoS Genet,*
733 9**,** e1003740.
734 RUTLEDGE, W., GIBSON, R., SIEGEL, E., DUKE, K., JONES, R., RUCINSKI, D., NUNN,
735 G., TORRENCE, W. A., LEWELLEN-WILLIAMS, C., STEWART, C., BLANN, K.,
736 BELLETON, L., FINCHER, L., KLIMBERG, V. S., GREENE, P., THOMAS, B.,
737 ERWIN, D. & HENRY-TILLMAN, R. 2006. Arkansas Special Populations
738 Access Network perception versus reality--cancer screening in primary
739 care clinics. *Cancer,* 107**,** 2052-60.
740 SAENKO, S. V., MARIALVA, M. S. & BELDADE, P. 2011. Involvement of the
741 conserved Hox gene Antennapedia in the development and evolution of a
742 novel trait. *Evodevo,* 2**,** 9.
743 SATO, A., KOJIMA, T., UI-TEI, K., MIYATA, Y. & SAIGO, K. 1999. Dfrizzled-3, a new
744 Drosophila Wnt receptor, acting as an attenuator of Wingless signaling in
745 wingless hypomorphic mutants. *Development,* 126**,** 4421-30.
746 SCHILLING, S., STEINER, S., ZIMMERLI, D. & BASLER, K. 2014. A regulatory
747 receptor network directs the range and output of the Wingless signal.
748 *Development,* 141**,** 2483-93.
749 STOEHR, A. M., WALKER, J. F. & MONTEIRO, A. 2013. Spalt expression and the
750 development of melanic color patterns in pierid butterflies. *Evodevo,* 4**,** 6.
751 TREISMAN, J. E., LAI, Z. C. & RUBIN, G. M. 1995. Shortsighted acts in the
752 decapentaplegic pathway in Drosophila eye development and has
753 homology to a mouse TGF-beta-responsive gene. *Development,* 121**,** 2835-
754 45.
755 URNESS, L. D. & THUMMEL, C. S. 1995. Molecular analysis of a steroid-induced
756 regulatory hierarchy: the Drosophila E74A protein directly regulates L71-
757 6 transcription. *EMBO J,* 14**,** 6239-46.
758 VAN BELLEGHEM, S. M., RASTAS, P., PAPANICOLAOU, A., MARTIN, S. H., ARIAS, C.
759 F., SUPPLE, M. A., HANLY, J. J., MALLET, J., LEWIS, J. J., HINES, H. M., RUIZ,
760 M., SALAZAR, C., LINARES, M., MOREIRA, G. R. P., JIGGINS, C. D.,

bioRxiv preprint first posted online Feb. 5, 2019; doi: http://dx.doi.org/10.1101/541599. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

761 COUNTERMAN, B. A., MCMILLAN, W. O. & PAPA, R. 2017. Complex
762 modular architecture around a simple toolkit of wing pattern genes. *Nat*
763 *Ecol Evol,* 1.
764 WALLBANK, R. W., BAXTER, S. W., PARDO-DIAZ, C., HANLY, J. J., MARTIN, S. H.,
765 MALLET, J., DASMAHAPATRA, K. K., SALAZAR, C., JORON, M., NADEAU, N.,
766 MCMILLAN, W. O. & JIGGINS, C. D. 2016. Evolutionary Novelty in a
767 Butterfly Wing Pattern through Enhancer Shuffling. *PLoS Biol,* 14**,**
768 e1002353.
769 WALTERS, J. R., HARDCASTLE, T. J. & JIGGINS, C. D. 2015. Sex Chromosome
770 Dosage Compensation in Heliconius Butterflies: Global yet Still
771 Incomplete? *Genome Biol Evol,* 7**,** 2545-59.
772 WARREN, R. W., NAGY, L., SELEGUE, J., GATES, J. & CARROLL, S. 1994. Evolution
773 of homeotic gene regulation and function in flies and butterflies. *Nature,*
774 372**,** 458-61.
775 WEATHERBEE, S. D., NIJHOUT, H. F., GRUNERT, L. W., HALDER, G., GALANT, R.,
776 SELEGUE, J. & CARROLL, S. 1999. Ultrabithorax function in butterfly wings
777 and the evolution of insect wing patterns. *Curr Biol,* 9**,** 109-15.
778 YASSIN, A., DELANEY, E. K., REDDIEX, A. J., SEHER, T. D., BASTIDE, H., APPLETON,
779 N. C., LACK, J. B., DAVID, J. R., CHENOWETH, S. F., POOL, J. E. & KOPP, A.
780 2016. The pdm3 Locus Is a Hotspot for Recurrent Evolution of Female-
781 Limited Color Dimorphism in Drosophila. *Curr Biol,* 26**,** 2412-2422.
782 ZHANG, L. & REED, R. D. 2016. Genome editing in butterflies reveals that spalt
783 promotes and Distal-less represses eyespot colour patterns. *Nat Commun,*
784 7**,** 11769.
785 ZHANG, L. L., MAZO-VARGAS, A. & REED, R. D. 2017. Single master regulatory
786 gene coordinates the evolution and development of butterfly color and
787 iridescence. *Proceedings of the National Academy of Sciences of the United*
788 *States of America,* 114**,** 10707-10712.
789
790