

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

Research article

**A field test for frequency-dependent selection on mimetic colour
patterns in *Heliconius* butterflies**

Authors: Patricio A. Salazar
Martin Stevens
Robert T. Jones
Imogen Ogilvie
Chris D. Jiggins

Affiliation: Department of Zoology, University of Cambridge, Downing Street,
Cambridge CB2 3EJ, United Kingdom

Running title: Predation on *Heliconius* butterflies

Corresponding author: Patricio A. Salazar

Current Address: Centro para la Investigación de la Biodiversidad y el Cambio
Climático - BioCamb
Universidad Tecnológica Indoamérica
Machala y Sabanilla
1703175 – Quito, Ecuador
Phone: +593 9 8655 9252
E-mail: patriciosalazar@uti.edu.ec / patoafan@gmail.com

26 **A field test for frequency-dependent selection on mimetic colour**
27 **patterns in *Heliconius* butterflies**

28 Patricio A. Salazar, Martin Stevens, Robert T. Jones, Imogen Ogilvie and Chris D. Jiggins

29 Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, United
30 Kingdom

31 **ABSTRACT**

32 Müllerian mimicry, the similarity among unpalatable species, is thought to evolve by frequency-
33 dependent selection. Accordingly, phenotypes that become established in an area are positively
34 selected because predators have learnt to avoid these forms, while introduced phenotypes are
35 eliminated because predators have not yet learnt to associate these other forms with
36 unprofitability. We tested this prediction in two areas where different colour morphs of the
37 mimetic species *Heliconius erato* and *H. melpomene* have become established, as well as in the
38 hybrid zone between these morphs. In each area we tested for selection on three colour
39 patterns: the two parental and the most common hybrid. We recorded bird predation on
40 butterfly models with paper wings, matching the appearance of each morph to bird vision, and
41 plasticine bodies. We did not detect differences in survival between colour morphs, but all
42 morphs were more highly attacked in the hybrid zone. This finding is consistent with recent
43 evidence from controlled experiments with captive birds, which suggest that the effectiveness of
44 warning signals decreases when a large signal diversity is available to predators. This is likely to
45 occur in the hybrid zone where over twenty hybrid phenotypes coexist.

46 **Key words:** *Heliconius*, frequency-dependent selection, Müllerian mimicry, artificial prey, false-
47 head hypothesis

48 **INTRODUCTION**

49 Müllerian mimicry evolves because two or more unpalatable species gain mutual advantage
50 from advertising a common display. In particular, mimicry reduces the number of individuals
51 that need to be killed per species before a predator learns to identify them as distasteful (Joron
52 & Mallet, 1998; Mallet & Joron, 1999; Sherratt, 2008). The species pair *Heliconius erato* Linnaeus,
53 1758 and *H. melpomene* Linnaeus, 1758 is a famous case of Müllerian mimicry. These distantly
54 related species are widely distributed in Central and South America, and despite their great
55 geographical variation in wing colour pattern they show almost identical colour patterns
56 wherever they co-occur. The persistence of numerous mimetic races of *H. erato* and *H.*
57 *melpomene* seems paradoxical, because these intra-specific variants can easily hybridize in
58 nature, a process that promotes the mixing of these divergent populations into a single hybrid
59 pool. Furthermore, frequency-dependent selection for mimicry should favour the existence of
60 one single colour pattern. Therefore, in order to maintain their colour pattern distinctiveness,
61 phenotypes that have reached a high frequency in a particular area are expected to be positively
62 selected, while novel colour morphs should tend to be eliminated. Frequency-dependent
63 selection in *Heliconius* is thought to be mainly caused by bird predators, as birds have excellent
64 colour vision (Cuthill, 2006) and are well known to discriminate tasteful from distasteful prey
65 (e.g. Brower, *et al.*, 1971; Langham, 2004; Ihalainen *et al.*, 2007).

66 Experiments with captive birds have shown that known *Heliconius* predators avoid familiar
67 colour patterns, but repeatedly attack novel forms (Langham, 2004), and field studies using
68 mark recapture experiments have revealed that mimetic (locally established) colour patterns
69 have higher survival than non-mimetic (introduced) patterns (Benson, 1972; Mallet & Barton,
70 1989; Kapan, 2001). However, there is no direct evidence that these survival differences are
71 caused by birds (Mallet & Barton, 1989). Moreover, mark-recapture studies do not tease apart
72 the potentially confounding effects of local adaptations, other than colour pattern, on the
73 survival of trans-located individuals (Mallet & Barton, 1989; Kapan, 2001). In this study, we

74 attempted to overcome these limitations by using artificial prey, a method that had successfully
75 been used to test for aposematism or mimicry in various organisms (e.g. Brodie III, 1993;
76 Comeault & Noonan, 2011; Valkonen *et al.*, 2011), including *Heliconius* and other Lepidoptera
77 models (Cuthill *et al.*, 2005; Stevens *et al.*, 2008; Finkbeiner *et al.*, 2012; Merrill *et al.*, 2012). Our
78 *Heliconius* models were created with paper wings that resembled real butterflies to bird vision,
79 and plasticine bodies that allowed us to identify predation events by the marks left on the
80 plasticine (Figure 1). Moreover, because the only difference between models is their colour
81 pattern, this method allowed us to test for colour pattern effects while controlling for the effects
82 of potentially confounding factors that exist when studying living butterflies (e.g. toxin levels,
83 behavioural differences, or physiological adaptations to different thermal environments).

84 In addition, we investigated the pattern of bird attacks on the butterfly models by recording the
85 location of each attack on the plasticine body. We were interested in this aspect of bird
86 predatory behaviour because there is evidence that birds tend to target their prey's head. For
87 instance, several butterfly species, especially Lycaenids, have evolved head mimicking structures
88 in the posterior end of their hind wings, which have been shown to deflect bird attacks away
89 from the head (Wourms & Wasserman, 1985). In our case, confirming that bird marks tend to
90 occur on the butterflies' head would further support the expectation that the artificial (and real)
91 prey are preyed upon by birds.

92 In this study, we focused on the *H. erato* and *H. melpomene* mimicry system in eastern Ecuador,
93 South America. Here, in the upper Pastaza river valley, above 1100m in elevation, the endemic
94 colour pattern race *H. e. notabilis* coexists with its co-mimic *H. m. plesseni*. Further east into the
95 Amazonian lowlands, below 500m, a different mimetic pair of *H. erato* and *H. melpomene* races
96 occurs (*H. e. lativitta* and *H. m. malleti*). At mid-elevation, between 500 and 1100m, the races
97 within each species hybridize across an approximately 38km-wide zone (Figure 2A). Models of
98 three *H. erato* colour patterns were exposed to predation in three zones: the two inhabited by
99 pure races and the hybrid zone (Figure 2B). In each zone we exposed models of the locally

100 established or most frequent form, as well as two introduced colour patterns (one of a pure race
101 and one hybrid; Table 1). Contrary to predictions, no differences in attack rates were detected
102 between the different aposematic patterns tested. However, higher overall predation rates were
103 recorded in the hybrid zone, which is characterised by a diverse prey community, and a strong
104 tendency for avian predators to target the prey's head was observed across all studied zones.

105 **MATERIALS AND METHODS**

106 **Prey stimuli**

107 Artificial prey were paper 'butterflies' printed with specific patterns to resemble the different
108 pure and hybrid forms tested in this study, along with a plasticine body (Figure 1). Stimuli wings
109 were made from waterproof paper (HP LaserJet Tough Paper, Palo Alto, USA) printed with the
110 different colour patterns of the real butterflies using a Hewlett Packard LaserJet 2605dn printer
111 at 300 dpi. The colour patterns used were based on digital photographs of real butterfly
112 specimens (four specimens per colour-pattern), taken with a Fujifilm IS Pro UV-sensitive digital
113 camera with a quartz CoastalOpt ultraviolet (UV) lens (Coastal Optical systems). It is well known
114 that birds see a wider colour spectrum than humans do (Cuthill, 2006), and stimuli printed from
115 uncalibrated printers rarely correspond closely to real object colours, especially to non-human
116 vision. Colours printed on the paper wings were therefore calibrated so that they stimulated the
117 photoreceptors of potential avian predators in a similar way to real wings. Specifically,
118 appropriate colours were selected so that when printed they produced photon catch values for
119 each cone type that fell within the range of values for the corresponding colour patches on the
120 real butterflies (as with past work on camouflage; e.g. Cuthill *et al.*, 2005; see also
121 Supplementary Material). All colours present on the studied phenotypes were closely
122 reproduced, except for the white patches of the highlands' race (*H. e. notabilis*), whose
123 reflectance in the UV range it was not possible to reproduce accurately (Figure 3; Supplementary
124 Material, Figure S1).

125 **Experimental design for frequency-dependent selection test**

126 We conducted field predation experiments with the artificial butterflies using a modified version
127 of a protocol well established for testing the survival value of camouflage and eyespot markings
128 with artificial prey (e.g. Cuthill *et al.*, 2005; Stevens *et al.*, 2008). Here, prey items are pinned out
129 at low density along non-linear transects and monitored for predation. We applied a fully
130 crossed factorial design, with two study factors (colour pattern and zone), each having three
131 levels of variation (3 colour patterns x 3 zones). If the hypothesis of frequency-dependent
132 selection was to be supported, each colour pattern was expected to survive best in the zone
133 where it reaches the highest frequency, while introduced (non-mimetic) colour patterns should
134 have lower survival. We carried out experimental trials in three replicate sites per zone (Figure
135 2B), each separated by at least 4 km. In each site, 80 models per pattern were set and exposed to
136 predation for 72 hours. One model per pattern was placed every 10-13m along accessible
137 pathways in areas where the studied *Heliconius* species were observed. Models were checked for
138 signs of attack every 24 hours. In order to optimize the methodology, we first carried our pilot
139 experiments between December 2009 and March 2010, and the actual test for frequency-
140 dependent selection took place between July and September 2010. Details of how potential
141 confounding factors were controlled for are described in the Supplementary Material.

142 **Analysis of predation marks data**

143 We tested the hypothesis of frequency-dependent selection as part of a set of competing
144 hypotheses that could explain the predation marks data. We followed a model selection
145 approach based on the Akaike information criterion (AIC; Burnham & Anderson, 2002). The data
146 consisted of time-to-mark observations for every model that was placed in the field. Models that
147 remained intact for the full 72h of the experimental trials were considered right-censored (i.e.
148 the time to mark is unknown, but is certainly longer than 72h). Data for marked models, in
149 contrast, were treated as interval-censored (e.g. when a model was found marked at 24h, it is
150 known that the mark occurred at some point in the interval 0-24h). Treating time-to-mark data

151 as interval-censored allowed us to use of all observations in the analysis, including those in
152 which a model was not found in one or more censuses but was re-sighted later; in this latter case,
153 the interval when the mark is known to occur is longer than the standard 24h, but remains
154 informative.

155 As first step in the analysis, we developed a set of competing hypotheses that could explain the
156 information contained in the data, and formalized them as statistical models. The experimental
157 design included the possible effect of the following factors: colour pattern, zone and
158 experimental site (nested within zone). Since the hypothesis of interest predicts that each tested
159 colour pattern should be the least attacked in the zones where they constitute the most frequent
160 phenotypes, we expected the statistical interaction between colour pattern and zone to be
161 strongly supported by the data. Thus, the final set of competing statistical models included
162 different combinations of the experimental design factors, including the interaction between
163 colour pattern and zone.

164 Statistical models were fitted in Minitab® (Minitab Inc., 2010), using the ‘Regression with life
165 data’ tool. In order to model the hazard function for the time-to-mark data (i.e. the function that
166 describes the probability of a butterfly model being marked through the course of the
167 experiment), we assessed the goodness-of-fit of several probability distributions before fitting
168 any of the competing models. We decided to use different probability distributions for bird and
169 non-bird marks. For bird marks, we assumed an exponential distribution—which presupposes a
170 constant hazard rate—because, given the small number of bird marks recorded, there was not
171 enough information in the data to justify a more complex function. For non-bird marks, in
172 contrast, we assumed a Weibull distribution—which can fit either increasing, constant, or
173 decreasing hazard rates—because the number of non-bird marks recorded was large enough to
174 give signs of variation in hazard rates for butterfly models affected by different factors, and the
175 fitting algorithm for models that assumed the more restrictive exponential distribution rarely
176 converged.

177 Once the statistical models were fitted, we computed the AIC for each of the competing models,
178 and ranked the models accordingly. The relative support in the data for each competing model
179 was assessed on the basis of their Akaike weights (w_i), and the evidence ratios obtained by
180 dividing the weight of the best-ranked model by the weight of each competing model (w_1/w_i).
181 Akaike weights range from zero to one, and can be interpreted as an approximate probability of
182 each model being the actual best-ranked model, if the model selection exercise would be
183 repeated with a different but equivalent data set (Burnham & Anderson, 2002). Finally, in order
184 to visualize the survival patterns of the butterfly models through the course of the experimental
185 trials, we reconstructed survival curves based on non-parametric Turnbull estimates of survival
186 probabilities for interval-censored data (Turnbull, 1976; Minitab Inc., 2010).

187 **Analyses of marked body sections**

188 In order to test the hypothesis that birds were more likely to attack the head of the plasticine
189 bodies rather than any section at random, we made all plasticine bodies the same length (3cm)
190 and recorded the body section (or sections) where attack marks occurred. We considered three
191 equally sized sections: anterior, middle and posterior, and tested whether the frequency of
192 marks occurring in one or more sections was larger than expected if all sections would be
193 equally targeted, using a G test.

194 Several organisms marked the models. However, bird marks were easily distinguished from
195 non-bird marks as they typically consisted of one single—usually triangular—dent, as opposed
196 to many scratches (Figure 1). Moreover, we directly observed attacks from ants, grasshoppers
197 and sweat bees in the field, and these explained the appearance of the majority of non-bird
198 marks. We performed analyses for both bird and non-bird marks, as it is reasonable to expect
199 different attack patterns for birds, which are visual predators, than for other organisms (e.g.
200 insects) that most probably attacked the models without regard to the colour pattern they
201 exhibited. Models marked by either birds or non-birds were taken away and not replaced.

202 **RESULTS**

203 **Test for frequency-dependent selection on colour patterns**

204 We recorded 56 bird and 667 non-bird marks in a total of 2160 models exposed to predation.
205 The low number of avian attacks compared to models probably reflects the fact that many birds
206 will have learnt to avoid the real butterflies. Bird marks did not support the hypothesis of
207 frequency-dependent selection for mimicry. None of the tested colour patterns survived best in
208 the zones where they are the most abundant phenotypes (Figure 4A). Moreover, none of the
209 statistical models that incorporate the interaction between colour pattern and zone were well
210 supported. In fact, the simplest model containing the interaction parameter was ranked third,
211 with an Akaike weight of 0.006, which is respectively eight and five times less supported than
212 the first and second ranked models (Table 2). Furthermore, colour pattern is in general a poor
213 predictor of the survival of butterfly models, as the survival curves for different colour patterns
214 tend to overlap in all zones (Figure 4A), and the statistical models containing the factor colour
215 pattern are not consistently ranked among the best supported (Table 2). The zone factor, in
216 contrast, has a much stronger effect. The survival of all butterfly models, irrespective of colour
217 pattern, is several times lower in the hybrid zone as compared to the parental zones (Figure 4B),
218 and the first two ranked statistical models, which incorporate the zone factor, have a combined
219 Akaike weight (w_{1+2}) of 0.93. This implies approximately thirteen times more support for these
220 first two models than for all other contested models (Table 2). Importantly, even though the first
221 ranked model incorporates the parameter 'colour pattern' in addition to 'zone', the Akaike
222 weights for each of the first two models are not very different. Thus, there is no evidence that
223 one of the two best-ranked models has more support in the data than the other, and
224 consequently the zone factor alone is a good predictor of the survival data.

225 For non-bird marks the results are different. Neither colour pattern nor zone (nor their
226 interaction) are good predictors of the survival of butterfly models. All colour patterns showed
227 similar survival within each zone (Figure 5A). However, attack rates were highly variable

228 between sites within zones (Figure 5B). In fact, the parameter 'experimental site' (nested within
229 zone) is the only well supported predictor of the non-bird marks data set. The first ranked
230 model, which is the most reduced model containing the site parameter, has a much larger Akaike
231 weight compared to all other candidate models ($w_1=0.87$; Table 2).

232 **Analysis of marked body sections**

233 There was a strong non-random distribution of bird attack marks on the body segments
234 ($G_6=33.04$, $p<0.001$). Bird marks were more likely to affect the most anterior or the
235 anterior+middle sections of the body, rather than any section or combination of sections at
236 random. In contrast, non-bird marks were more evenly distributed across body sections
237 ($G_6=562.89$, $p<0.001$, Figure 6).

238 **DISCUSSION**

239 **Is there frequency-dependent selection for mimicry?**

240 Unexpectedly, we found no evidence that bird attacks were associated with wing colour pattern.
241 Furthermore, none of the locally established or most abundant colour patterns were favoured
242 over introduced patterns (Figure 4A). Overall, and contrary to predictions, there was no
243 statistical evidence that colour pattern had any effect on bird attack rates (Table 2). The inability
244 to detect a colour pattern effect was probably due to a lack of statistical power, as the effect size
245 of differential predation on colour patterns may be small, and therefore it would be necessary to
246 gather a larger sample of informative data (i.e. bird marks) to have the power to detect such
247 effects. Small colour pattern effects are perhaps not surprising, since all the tested colour
248 patterns are aposematic and will therefore tend to be avoided. This is supported by the fact that
249 we only recorded 56 avian attacks despite putting out a total of 2160 models. It is worth noticing,
250 however, that this attack rate is within the range reported in other studies with artificial prey
251 (Table 3), including at least one experiment with artificial *Heliconius* butterflies that did detect
252 differences in bird predation associated with colour pattern (Merrill *et al.*, 2012; Table 3). Hence,

253 bird marks were not particularly infrequent in our study, but still not enough to detect any
254 colour pattern effects that may exist. Importantly, the fact that we did not detect colour pattern
255 effects does not mean that the hypothesis of frequency-dependent selection for mimicry is
256 rejected. There is simply no evidence in our data that the locally established patterns were
257 favoured in the zones where they are the most frequent phenotypes, but similarly there is no
258 evidence that introduced colour patterns were favoured either.

259 Nevertheless, our results could have other implications that have rarely been explored. It is
260 possible that hybrid forms may only be at a disadvantage when they differ substantially in
261 appearance from locally occurring abundant pure forms. It is well known that predators,
262 including birds, can generalise their response in avoiding familiar aposematic prey towards
263 novel or unfamiliar prey of similar appearance (e.g. Aronsson & Gamberale-Stille, 2008;
264 Ihalainen *et al.*, 2012). However, how and when generalisation works is poorly understood, and
265 research aiming to understand generalisation behaviour has generally been held back because
266 few studies have analysed warning coloration from the perspective of the predator's visual and
267 cognitive systems rather than a human perspective (Stevens & Ruxton, 2012). In addition, most
268 work has focussed on one aspect of prey appearance at a time (e.g. colour) and rarely considered
269 the entire appearance of the prey item, which may be crucial in generalisation. It is also possible
270 that predators use key diagnostic features in generalising between aposematic species (e.g. the
271 presence of red and black patches along with characteristic wing shapes). Therefore, in this
272 study the avian predators could have been very effective in generalising the appearance of the
273 different morphs and associating key features with unpalatability. This is an area that needs
274 much more research and could have significant implications for the evolution of mimicry and the
275 fitness of novel colour forms as a whole. Although we did not detect a colour pattern effect on
276 bird marks, there was a strong effect of 'zone', with higher attack rates observed within the
277 hybrid zone (Figure 4B). This observation offers a possible insight into the behaviour of bird
278 predators when exposed to a diversity of colour patterns. It has been shown experimentally that

279 the effectiveness of aposematic signals decreases when a larger diversity of signals is available
280 to a bird predator (Ihalainen *et al.*, 2012). In the hybrid zone, there are at least 10 hybrid
281 patterns of *H. erato* and at over 10 more of *H. melpomene*. This large diversity of aposematic
282 signals can potentially explain the lower survival of all tested colour patterns in the hybrid zone,
283 because the predator bird community may have a harder time learning all these different
284 patterns and identifying them all as distasteful, and attack them indiscriminately.

285 Finally, an additional important effect detected was a random effect of the variation between
286 experimental sites on non-bird marks (Figure 5B, Table2). This effect was explained by
287 anecdotal observations about the local abundance of some non-bird 'predators', notably sweat
288 bees and grasshoppers, which varied widely among experimental sites. Thus, it is clear that
289 accidental (non-visual) predators, such as insects, attack the butterfly models randomly, without
290 respect to colour pattern or zone.

291 **Do birds target the head?**

292 Our data clearly support the hypothesis that birds attack the head section of the body, rather
293 than any section at random. Other studies have also found a tendency for birds to attack head-
294 like structures. For instance, in a large sample of wild-caught Lycaenids, Robbins (1981) found
295 that species with a conspicuous false-head, in the anal angle of the hind-wings, showed five
296 times more damage from predator attacks on their hind-wings than species that lack the false-
297 head. Likewise, Wourms & Wasserman (1985) showed experimentally that head-like structures,
298 painted on the hind-wings of a butterfly that naturally lacks a false-head, increased the chances
299 that blue jays would attack the hind-wing area, away from the real head. Furthermore, Wourms
300 & Wasserman (1985) also found that the deflected attacks increased the chance that the
301 butterflies would escape during the necessary prey handling that happens after the attack. In
302 both studies, the head-like structures that fooled predators included features that are suggestive
303 of a butterfly head (e.g. false antennae or eye-spots). In our study, however, the plasticine bodies
304 did not have any feature that would suggest that one of the ends was meant to be the head

305 (Figure 1). Therefore, in our models birds must recognize the head section by the orientation, as
306 suggested by wing shape and colour pattern. It can be argued that the metallic reflection of the
307 pin, placed on the boundary between the anterior and middle sections, could have attracted
308 birds to this area. Even though we cannot rule out this possibility, the fact that the frequency of
309 bird marks that affected the most anterior section only, without affecting the pin area, is almost
310 as large as the frequency of marks that affected the pin (23.2% vs. 37.5%), suggests that the pin
311 did not influence the body section targeted by birds. These results also strongly indicate that our
312 artificial prey were effective: birds attacked our models as if they were real butterflies.

313 **On the use of artificial *Heliconius* to study bird predation and natural selection**

314 In spite of our negative results with respect to colour pattern effects, we still recommend the use
315 of artificial *Heliconius* to study bird predation on these butterflies. In fact, a recently published
316 study has successfully used *H. erato* models to detect lower predation rates over butterflies that
317 roost gregariously compared to butterflies that roost solitarily (Finkbeiner *et al.* 2012).
318 Furthermore, another study, performed in Panama around the same time as the study here
319 reported, succeeded in detecting differences in bird predation associated with colour pattern,
320 using models constructed in exactly the same way as those used in our experiment (Merrill *et al.*
321 2012).

322 In general, studies with artificial prey are complementary to mark recapture experiments, such
323 as those performed by Benson (1972), Mallet & Barton (1989) and Kapan (2001), because
324 artificial prey provide different technical advantages. For instance, predators can be easily
325 recognized by the marks that they leave in the models, which allow researchers to keep a record
326 of predation events, instead of just inferring these events from individual disappearances. Also,
327 models allow researchers to test for the effects of morphological characters, such as colour
328 pattern or body shape, while controlling for other potentially confounding factors. In the present
329 study, for instance, it was necessary to control for the possible effects of adaptations associated
330 with the altitudinal gradient across the hybrid zone (e.g. variation in thermal tolerance). Finally,

331 there is usually little constraint in the number of models that can be exposed to predation, and
332 therefore it is possible to attain larger sample sizes than those attained when using live
333 specimens. Despite these advantages, there are a few considerations that we suggest should be
334 taken into account in future studies. In particular, *Heliconius* aposematic patterns are likely to
335 reduce the chances of the models being attacked, increasing the effort required to gather a large
336 enough sample of bird marks. It seems important, in retrospect, to focus on gathering a large
337 number of predator marks rather than favouring the total number of models exposed to
338 predation. In the present study, for instance, replacing models that were accidentally marked by
339 non-bird ‘predators’—such as ants, grasshoppers or sweat bees—rather than taking these
340 models away from the experiment, would have probably increased the chances of recording
341 more bird marks. A further desirable addition would be to include non-aposomatic (either
342 palatable or fake) colour-patterns as control treatments, since they provide information about
343 overall rates of attack and the degree of aposematism of the models, and/or whether the
344 artificial prey appear realistic to predators.

345 Overall, our study has shown that the use of artificial *Heliconius* is helpful in exploring aspects of bird
346 predatory behaviour for which there is little evidence of their importance in the field. In particular, the
347 effects of prey community structure on birds’ discriminatory behaviour that selects for Müllerian
348 mimicry had been elegantly demonstrated in captivity (Ihalainen *et al.*, 2012), but not backed up by
349 field observations yet. Our discovery that the rate of bird attacks is substantially larger in a
350 morphologically diverse prey community, such as a hybrid zone, is an important piece of evidence in
351 support of the idea that a complex aposematic prey community can impact the patterns of bird
352 predation in nature, and—as a consequence—can impact the evolution of Müllerian mimicry as well.
353 According to Ihalainen *et al.*’s (2012) experiment in the lab, accurate mimicry is more likely to evolve
354 in simpler prey communities. Hence, our study has opened new opportunities to address questions
355 about the interactions between *Heliconius* butterflies and their bird predators.

356 **ACKNOWLEDGEMENTS**

357 We would like to thank Richard Merrill and Steven Shaak for their collaboration during the
358 colour calibration of *Heliconius* models. Nicola Nadeau and Agata Surma assisted in the field
359 during the pilot experiments, while Carlos Robalino and Elizabeth Heras did so during the large-
360 scale experiment. Jens Töniges and Robby Flemisch, from the Curiquingue Foundation (Tena),
361 and Carmen Luzuriaga, from the Pindo-Mirador Biological Station (Universidad Tecnológica
362 Equinoccial), provided great logistical support while we worked in their field stations. Santiago
363 Villamarín and Marco Altamirano, from the Ecuadorian Museum of Natural Sciences, sponsored
364 and helped us in obtaining research, collection and export permits from the Ecuadorian Ministry
365 of Environment. The Ministry of Environment, in turn, through its provincial offices in Napo and
366 Pastaza, provided us with the aforementioned permits. Finally, Patricia Salazar provided
367 constant logistical support during the whole fieldwork period in Ecuador. Financial support
368 came primarily from a Leverhulme Trust award to C.D.J., the University of Cambridge Trusts to
369 P.A.S, the Balfour Fund (Zoology Department at Cambridge) to I.O. MS was supported by a
370 Biotechnology and Biological Sciences Research Council David Phillips Research Fellowship
371 (BB/G022887/1).

372

372 **LITERATURE CITED**

- 373 Aronsson, M. & Gamberale-Stille, G. 2008. Domestic chicks primarily attend to colour, not
374 pattern, when learning an aposematic coloration. *Anim. Behav.* **75**: 417–423.
- 375 Benson, W.W. 1972. Natural selection for Müllerian mimicry in *Heliconius erato* in Costa Rica.
376 *Science* **176**: 936–939.
- 377 Brodie III, E.D. 1993. Differential avoidance of coral snake banded patterns by free-ranging avian
378 predators in Costa Rica. *Evolution* **47**: 227–235.
- 379 Brower, L.P., Alcock, J. & Brower, J.V.Z. 1971. Avian feeding behaviour and the selective
380 advantage of incipient mimicry. In: *Ecological Genetics and Evolution* (R. Creed, ed.), pp.
381 261–274. Blackwell Scientific, Oxford.
- 382 Burnham, K.P. & Anderson, D.R. 2002. *Model selection and multimodel inference: a practical*
383 *information-theoretic approach*, 2nd edn. Springer, New York, USA.
- 384 Comeault, A.A. & Noonan, B.P. 2011. Spatial variation in the fitness of divergent aposematic
385 phenotypes of the poison frog, *Dendrobates tinctorius*. *J. Evol. Biol.* **24**: 1374–1379.
- 386 Cuthill, I. C., Stevens M., Sheppard, J., Maddocks, T., Párraga, C.A. & Troscianko, T.S. 2005.
387 Disruptive coloration and background pattern matching. *Nature* **434**: 72–74.
- 388 Cuthill, I. C. 2006. Color perception in birds. In: *Bird coloration. I. Mechanisms and measurement*
389 (G. E. Hill & K. J. McGraw, eds.), pp. 3–40. Harvard University Press, Cambridge, MA, USA.
- 390 Endler, J.A. & Mielke, P.W. 2005. Comparing entire colour patterns as birds see them. *Biol. J. Linn.*
391 *Soc.* **86**: 405–431.
- 392 Finkbeiner, S.D., Briscoe, A.D. & Reed, R.D. 2012. The benefit of being a social butterfly:
393 communal roosting deters predation. *P. Roy. Soc. Lond. B* **279**: 2769–2776.
- 394 Hart, N.S., Partridge, J.C., Cuthill, I. C. & Bennett, A.T. 2000. Visual pigments, oil droplets, ocular
395 media and cone photoreceptor distribution in two species of passerine bird: the blue tit
396 (*Parus caeruleus* L.) and the blackbird (*Turdus merula* L.). *J. Comp. Physiol. A* **186**: 375–387.

- 397 Ihalainen, E., Lindstrom, L. & Mappes, J. 2007. Investigating Mullerian mimicry: predator
398 learning and variation in prey defences. *J. Evol. Biol.* **20**: 780–791.
- 399 Ihalainen, E., Rowland, H.M., Speed, M.P., Ruxton, G.D. & Mappes, J. 2012. Prey community
400 structure affects how predators select for Mullerian mimicry. *P. Roy. Soc. Lond. B* **279**:
401 2099–2105.
- 402 Joron, M. & Mallet, J. 1998. Diversity in mimicry: paradox or paradigm? *Trends Ecol. Evol.* **13**:
403 461–466.
- 404 Kapan, D.D. 2001. Three-butterfly system provides a field test of müllerian mimicry. *Nature* **409**:
405 338–340.
- 406 Langham, G.M. 2004. Specialized avian predators repeatedly attack novel color morphs of
407 *Heliconius* butterflies. *Evolution* **58**: 2783–2787.
- 408 Langmore, N.E., Stevens, M., Maurer, G. & Kilner, R.M. 2009. Are dark cuckoo eggs cryptic in host
409 nests? *Anim. Behav.* **78**: 461–468. Elsevier Ltd.
- 410 Mallet, J. & Joron, M. 1999. Evolution of diversity in warning color and mimicry: Polymorphisms,
411 shifting balance, and speciation. *Annu. Rev. Ecol. Syst.* **30**: 201–233.
- 412 Mallet, J. & Barton, N.H. 1989. Strong natural selection in a warning-color hybrid zone. *Evolution*
413 **43**: 421–431.
- 414 Merrill, R.M., Wallbank, R.W.R., Bull, V., Salazar, P.A., Mallet, J., Stevens, M., *et al.* 2012. Disruptive
415 ecological selection on a mating cue. *P. Roy. Soc. Lond. B*, doi: 10.1098/rspb.2012.1968.
- 416 Minitab Inc. 2010. Minitab 16 statistical software. Minitab, Inc., State College, PA (USA).
- 417 Pfennig, D.W., Harcombe, W.R. & Pfennig, K.S. 2001. Frequency-dependent Batesian mimicry.
418 *Nature* **410**: 323.
- 419 Robbins, R.K. 1981. The “false head” hypothesis: predation and wing pattern variation of
420 Lycaenid butterflies. *Am. Nat.* **118**: 770–775.
- 421 Salazar, P.A. 2012. Hybridization and the genetics of wing colour-pattern diversity in *Heliconius*
422 butterflies. PhD thesis. University of Cambridge. Cambridge, UK.
- 423 Sherratt, T.N. 2008. The evolution of Müllerian mimicry. *Naturwissenschaften* **95**: 681–695.

- 424 Stevens, M., Hopkins, E., Hinde, W., Adcock, A., Connolly, Y., Troscianko, T., *et al.* 2007. Field
425 experiments on the effectiveness of “eyespot” as predator deterrents. *Anim. Behav.* **74**:
426 1215–1227.
- 427 Stevens, M., Hardman, C.J. & Stubbins, C.L. 2008. Conspicuousness, not eye mimicry, makes
428 “eyespot” effective antipredator signals. *Behav. Ecol.* **19**: 525–531.
- 429 Stevens, M. & Ruxton, G.D. 2012. Linking the evolution and form of warning coloration in nature.
430 *P. Roy. Soc. Lond. B* **279**: 417–426.
- 431 Turnbull, B. 1976. The empirical distribution function with arbitrarily grouped, censored and
432 truncated data. *J. Roy. Stat. Soc. B.* **38**: 290–295.
- 433 Valkonen, J., Niskanen, M., Björklund, M. & Mappes, J. 2011. Disruption or aposematism?
434 Significance of dorsal zigzag pattern of European vipers. *Evol. Ecol.* **25**: 1047–1063.
- 435 Vignieri, S.N., Larson, J.G. & Hoekstra, H.E. 2010. The selective advantage of crypsis in mice.
436 *Evolution* **64**: 2153–2158.
- 437 Wourms, M.K. & Wasserman, F.E. 1985. Butterfly wing markings are more advantageous during
438 handling than during the initial strike of an avian predator. *Evolution* **39**: 845–851.
- 439
- 440

440 **Table 1.** Location of experimental sites in Eastern Ecuador and frequency of each tested colour-pattern
 441 per experimental site (colour-pattern frequencies do not add up to 1 within sites, because other hybrid
 442 colour-patterns—less abundant than that tested in this study—could have been present at any particular
 443 site; see Salazar 2012, Chapter 3, for details).

Zone	Site name	Latitude	Longitude	Elevation (m)	Colour-pattern frequency		
					<i>notabilis</i>	hybrid	<i>lativitta</i>
Highlands	El Topo	S 1.398	W 78.1781	1322	1	0	0
	Mangayacu	S 1.4371	W 78.1229	1241	1	0	0
	Pindo-Mirador	S 1.4600	W 78.0728	1205	0.98	0	0
Hybrid	San Francisco de Llandia	S 1.3062	W 77.9098	833	0.03	0.67	0
	Km35 Antena	S 1.2910	W 77.8774	1055	0	0.67	0
	Punta San Cristobal	S 1.2908	W 77.8419	956	0	0.57	0.03
Lowlands	Y de Misahuallí	S 1.0614	W 77.6684	446	0	0	0.65
	San Pedro de Arajuno	S 1.0988	W 77.5845	376	0	0	0.76
	Río Pusuno	S 1.0157	W 77.5975	382	0	0	1.00

444

445

445 **Table 2.** Statistical model comparison based on AIC.

Bird marks							
Rank	Model †	$\log(\mathcal{L})$	K	AIC	Δ_i	w_i	w_1/w_i
1	Z + P	-300.81	5	611.62	0	0.5715	1
2	Z	-303.28	3	612.55	0.93	0.3586	1.59
3	Z + P + Z*P	-298.98	9	615.97	4.35	0.0651	8.78
4	Z + S(Z)	-301.99	9	621.97	10.35	0.0032	176.80
5	P	-309.19	3	624.39	12.77	0.0010	591.70
6	Z + P + Z*P + S(Z)	-297.66	15	625.32	13.70	0.0006	943.88
Non-bird marks							
Rank	Model	$\log(\mathcal{L})$	K	AIC	Δ_i	w_i	w_1/w_i
1	Z + S(Z)	-1905.39	10	3830.78	0	0.8684	1
2	Z + P + S(Z)	-1905.33	12	3834.65	3.87	0.1253	6.93
3	Z + P + Z*P + S(Z)	-1904.32	16	3840.64	9.86	0.0063	138.10
4	Z	-1946.54	4	3901.07	70.29	4.74E-16	1.83E+15
5	Z + P	-1946.47	6	3904.95	74.17	6.82E-17	1.27E+16
6	Z + P + Z*P	-1945.52	8	3907.04	76.26	2.40E-17	3.62E+16
7	P	-1978.81	4	3965.62	134.84	4.56E-30	1.91E+29

† Z: zone, P: colour pattern, Z*P: interaction between colour pattern and zone, S(Z): experimental site nested within zone

Column headings: $\log(\mathcal{L})$: minimized log-likelihood of model given the data, K : number of parameters, AIC: Akaike information criterion, Δ_i : AIC differences, w_i : Akaike weight, w_1/w_i : evidence ratio between best ranked model and each i th ranked model

446

447 **Table 3.** Summary of a sample of previous studies in which models were used to investigate natural selection on colour pattern.

Publication	Organism	Hypothesis	Total models	Bird marks		Non-bird marks		Exposure time (h)	Colour calibration	Support for hypothesis
				No.	Freq.	No.	Freq.			
Brodie III, 1993 (Experiment 1)	Coral snakes (<i>Micrurus nigrocintus</i>)	Aposematism	408	32	0.08	Not reported	-	48	No	Yes
Brodie III, 1993 (Experiment 2)	Coral snakes (<i>Micrurus</i> spp. and others)	Mimicry	2100	120	0.06	76	0.04	72	No	Yes
Pfennig <i>et al.</i> , 2001 (Experiment 1)	Coral snakes (<i>Micrurus fulvius</i>) and kingsnakes (<i>Lampropeltis triangulum</i>)	Batesian mimicry	480	25*	0.05	Not reported	-	672 (4 weeks)	No	Yes
Pfennig <i>et al.</i> , 2001 (Experiment 2)	Coral snakes (<i>Micruroides euryxanthus</i>) and kingsnakes (<i>Lampropeltis pyromelana</i>)	Batesian mimicry	720	49*	0.06	Not reported	-	336 (2 weeks)	No	Yes
Vignieri <i>et al.</i> , 2010	Beach mice (<i>Peromyscus polionotus</i>)	Crypsis	896	8	0.01	20	0.02	72	No	Yes†
Valkonen <i>et al.</i> , 2011 (Experiment 1)	Viper snakes (<i>Vipera</i> spp.)	Aposematism vs. disruptive coloration	900	Not reported	0.09	Not reported	0.07	44-47	No	Yes

448 **Table 3.** (Continuation)

Publication	Organism	Hypothesis	Total models	Bird marks		Non-bird marks		Exposure time (h)	Colour calibration	Support for hypothesis
				No.	Freq.	No.	Freq.			
Valkonen <i>et al.</i> , 2011 (Experiment 2)	Viper snakes (<i>Vipera</i> spp.)	Aposematism vs. disruptive coloration	360	Not reported	0.10	Not reported	0.04	42-49	No	Yes
Comeault & Noonan, 2011	Poison frogs (<i>Dendrobates tinctorius</i>)	Frequency-dependent selection for aposematism	1891	49	0.03	Not reported	-	72	No	Yes
Merrill <i>et al.</i> , 2012	<i>Heliconius</i> butterflies (<i>Heliconius cydno</i> and <i>H. melpomene</i>)	Frequency-dependent selection for mimicry	1440	58	0.04	Not reported	-	72	Yes	Yes
This study	<i>Heliconius</i> butterflies (<i>Heliconius erato</i>)	Frequency-dependent selection for mimicry	2160	56	0.03	667	0.31	72	Yes	No

* Predation marks were not bird marks but carnivore marks (e.g. black bears, bobcats, coyotes, foxes or raccoons)

† Hypothesis test was not only based on bird marks but also on mammals marks

NOTE: all these studies reported the occurrence of bird and non-bird marks; not all of them, however, reported non-bird marks numbers

449 **Figure 1**

Intact



Bird mark



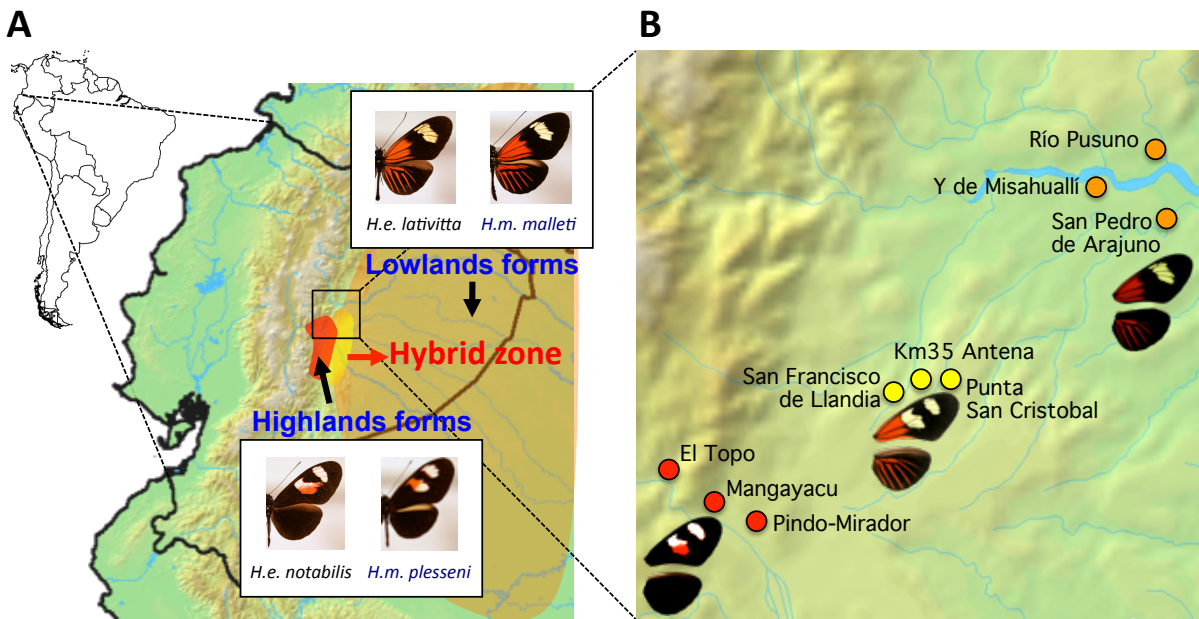
Non-bird mark



450

451

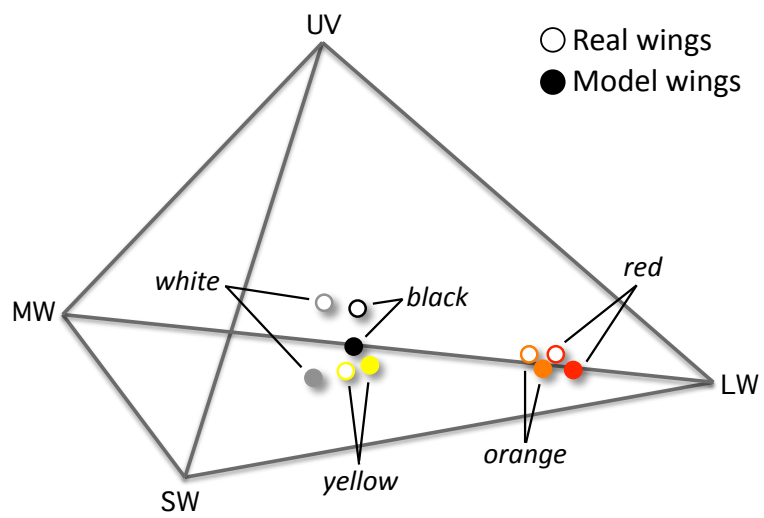
451 **Figure 2**



452

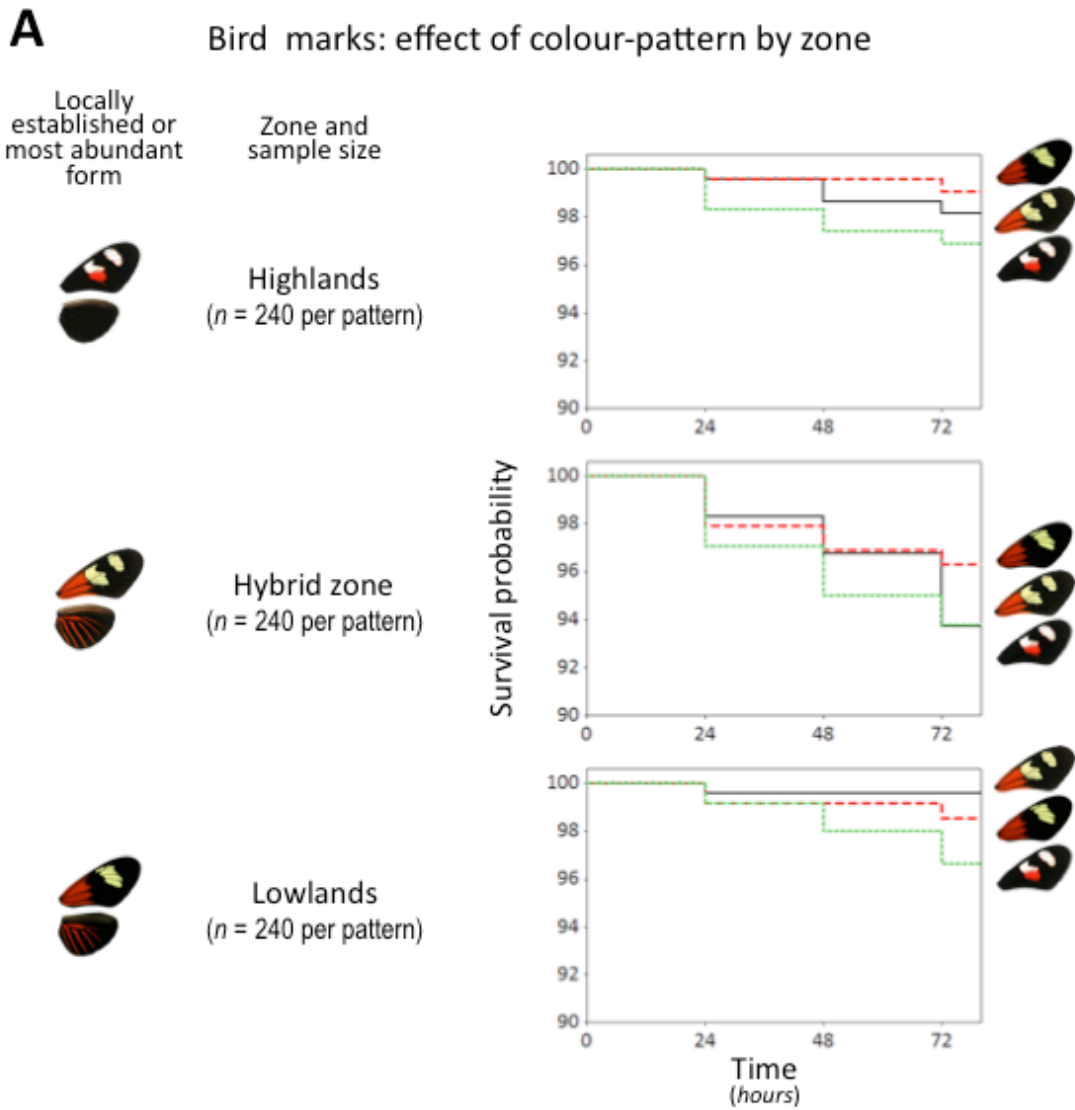
453

453 **Figure 3**

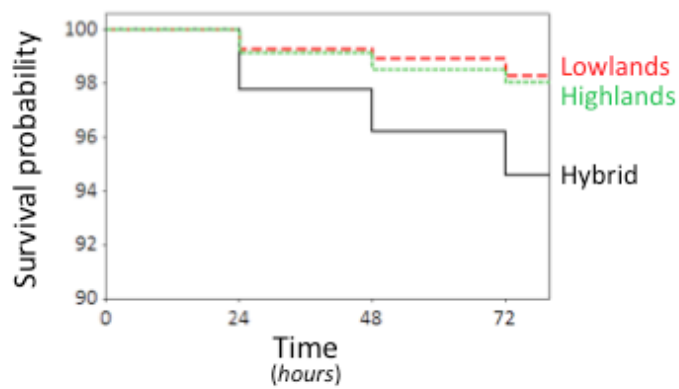


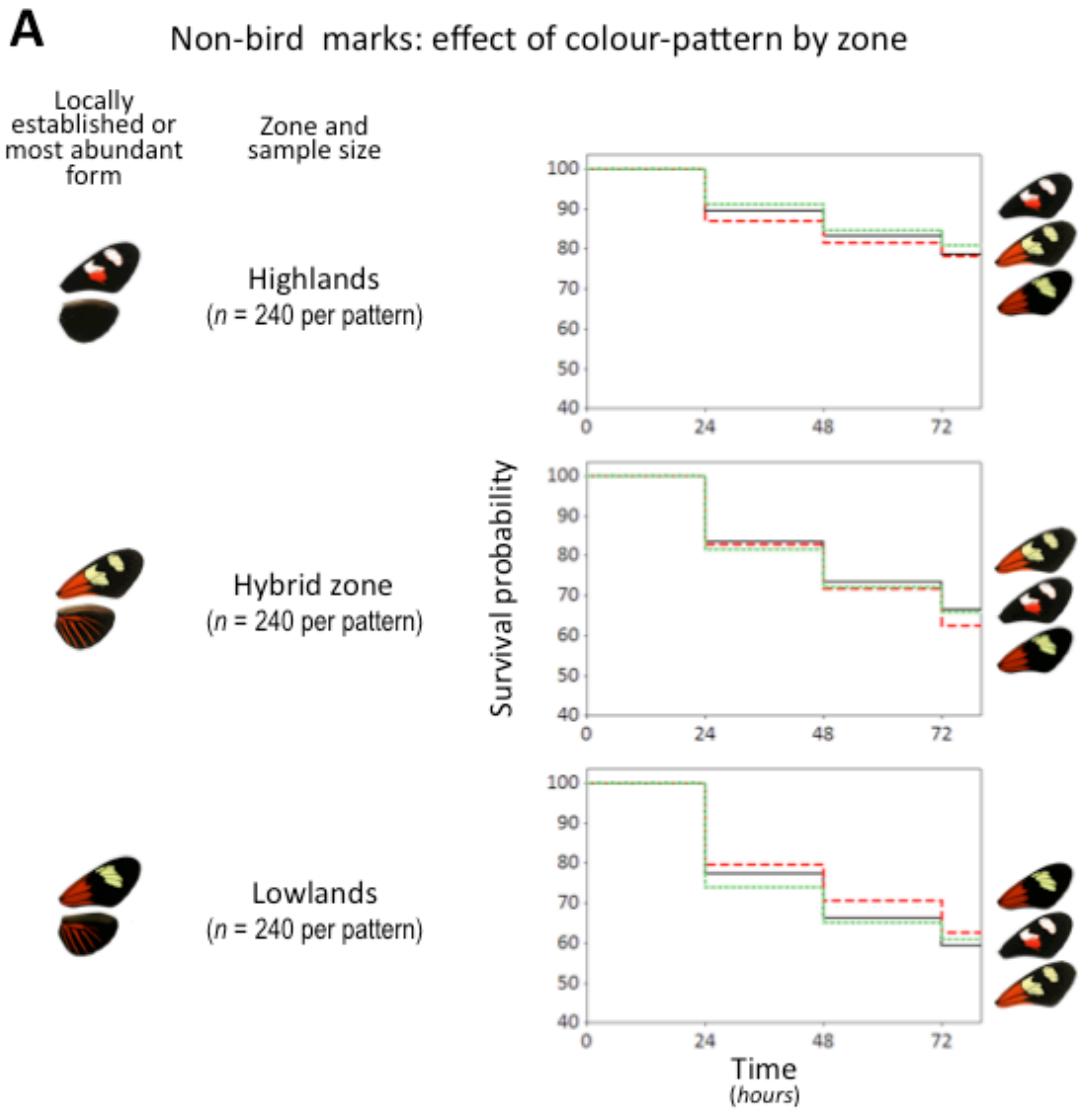
454

455

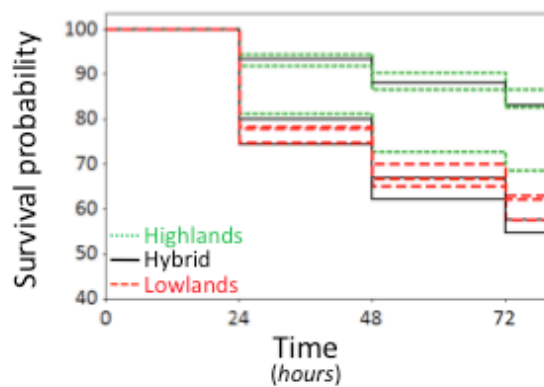


B Bird marks: effect of zone (all patterns combined)

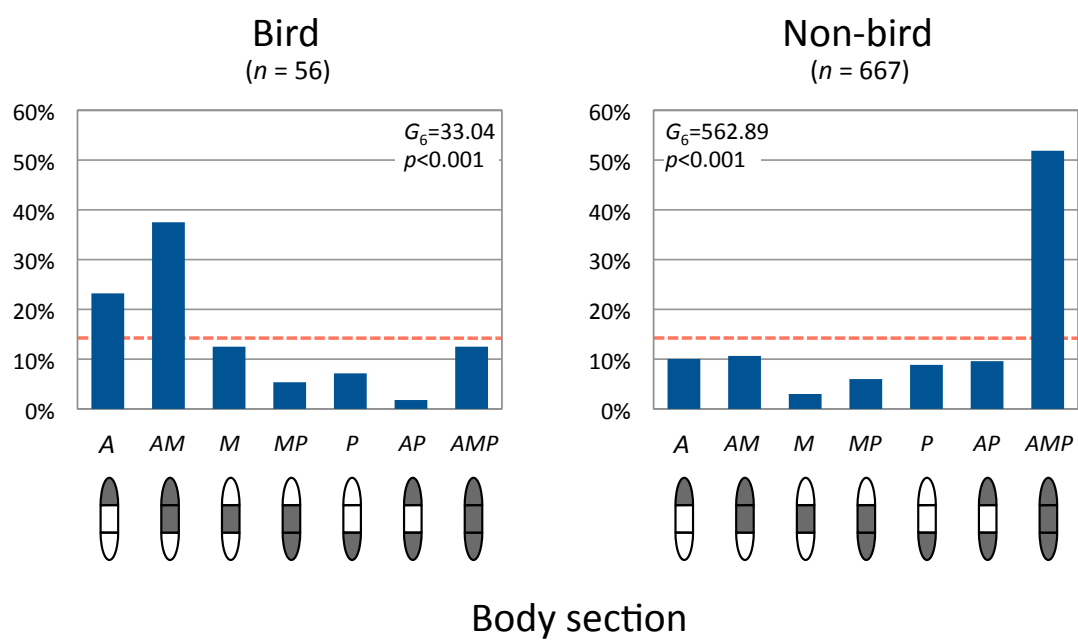




B Non-bird marks: effect of site (nested within zone)



459 **Figure 6**



460

461 **FIGURE LEGENDS**

462 **Figure 1.** Photographs of intact and marked models. Bird marks were easily identified because
463 they typically consisted of one single—usually triangular—dent, whereas non-bird marks
464 consisted of many scratches.

465 **Figure 2.** (A) Geographic location of the study area, and mimicry between *Heliconius erato* and *H.*
466 *melpomene*. Hybridization occurs between races within each species: *H. e. lativitta* hybridizes
467 with *H. e. notabilis*, and *H. m. malleti* hybridizes with *H. m. plesseni*. (B) Study sites (red, orange,
468 and yellow circles) and locally established or most abundant colour patterns of *H. erato*.

469 **Figure 3.** Location of colour patches of real wings (open circles) and paper models (filled
470 circles) in the tetrahedral colour space defined by the photon-catch values for each cone type of
471 the bird retina: ultraviolet (UV), short (SW), medium (MW) and long (LW) wavelength sensitive
472 cones. Notice that all colours in the models were reproduced very close to those of real wings.
473 The only exception was the white patches of the *H. erato* highlands' race (*H. e. notabilis*), whose
474 reflectance in the UV range it was not possible to reproduce (see Supplementary Material for
475 details).

476 **Figure 4.** Non-parametric survival curves over time for bird marks: (A) effect of colour pattern
477 in the survival of butterfly models within each zone, (B) effect of zone, disregarding colour
478 pattern.

479 **Figure 5.** Survival curves for non-bird marks: (A) effect of colour pattern in the survival of
480 butterfly models within each zone, (B) effect of experimental sites within zones (disregarding
481 colour pattern). High variation in survival between sites blurs any potential difference in
482 survival between zones.

483 **Figure 6.** Frequency of bird and non-bird marks affecting a particular body section, or
484 combination of sections. The body sections considered were A: anterior, M: middle and P:
485 posterior. The red dashed line indicates the expected frequency if each section or combination of
486 sections had been equally affected at random. Bird marks were more likely to affect either the
487 anterior or anterior+middle sections. Non-bird marks, in contrast, were more likely to affect all
488 body sections at the same time.

489