*Statement of authorship*: JM, KR, ON, JV, BR and SG designed the experiment. All authors contributed to the fieldwork. KR, JV and ON did the statistical analysis. KR led the writing and all authors contributed to it and accepted the final version.

*Data accessibility statement:* Should the manuscript be accepted, the data supporting the results will be archived in the University of Jyväskylä data repository JYX.

*Article title*: **Geographic mosaic of selection by avian predators on hindwing warning colour in a polymorphic aposematic moth**

*Authors*:
Katja Rönkä[1,2,3]*, katja.ronka@helsinki.fi, +358445551560
Janne K. Valkonen[1], janne.k.valkonen@jyu.fi
Ossi Nokelainen[1], ossi.nokelainen@jyu.fi
Bibiana Rojas[1], bibiana.rojas@jyu.fi
Swanne Gordon[1,4], swanne.gordon@wustl.edu
Emily Burdfield-Steel[1,5], e.r.burdfieldsteel@uva.nl
Johanna Mappes[1], johanna.mappes@jyu.fi

[1]*University of Jyväskylä, Department of Biological and Environmental Science, Jyväskylä, Finland*

[2]*Helsinki Institute of Life Sciences, University of Helsinki, Helsinki, Finland*

[3]*Organismal and Evolutionary Biology Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki*
[4]*Department of Biology, Washington University in St. Louis, St. Louis, Missouri, United States of America*
[5]*Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands*

*\*Lead Contact (Corresponding author)*

*Running title:* Geographic mosaic of selection on warning colour

*Keywords***:** frequency-dependent selection, aposematism, colour polymorphism, signal variation, signal convergence, predators, predator-prey interactions, wood tiger moth, *Arctia plantaginis*

1 **Abstract**

2 Warning signals are predicted to develop signal monomorphism via positive frequency-

3 dependent selection (+FDS) albeit many aposematic systems exhibit signal

4 polymorphism. To understand this mismatch, we conducted a large-scale predation

5 experiment in four locations, among which the frequencies of hindwing warning

6 coloration of aposematic *Arctia plantaginis* differ. Here we show that selection by avian

7 predators on warning colour is predicted by local morph frequency and predator

8 community composition. We found +FDS to be strongest in monomorphic Scotland,

9 and in contrast, lowest in polymorphic Finland, where different predators favour

10 different male morphs. +FDS was also found in Georgia, where the predator community

11 was the least diverse, whereas in the most diverse avian community in Estonia, hardly

12 any models were attacked. Our results support the idea that spatial variation in predator

13 and prey communities alters the strength or direction of selection on warning signals,

14 thus facilitating a geographic mosaic of selection.

15

**Introduction**

The survival strategy of aposematism, wherein prey use warning signals that predators learn to associate with their unprofitability and subsequently avoid, has stimulated biological studies for centuries (Ruxton *et al.* 2018; Skelhorn *et al.* 2016; Merrill *et al.* 2015; Mappes *et al.* 2005; Cott 1940; Poulton 1890; Wallace 1867). In aposematism, prey benefit from lowered costs of predator education by carrying a common signal, while predators reduce risks by not attacking defended prey. This results in selection for local similarity in warning signals, a view that has been corroborated by theoretical approaches (e.g. Aubier & Sherratt 2015; Sherratt 2008; Mallet & Joron 1999; Müller 1878), laboratory experiments (e.g. Rowland *et al.* 2007; Lindström *et al.* 2001; Greenwood *et al.* 1989), and field studies (e.g. Dell'aglio *et al.* 2016; Chouteau *et al.* 2016; Borer *et al.* 2010; Kapan 2001; Mallet & Barton 1989). Nevertheless, phenotypic variation and polymorphism in aposematic organisms are widespread in nature (e.g. frogs: Rojas 2017; Siddiqi *et al.* 2004; newts: Beukema *et al.* 2016; Mochida 2011; butterflies: Merrill *et al.* 2015; moths: Brakefield & Liebert 1985; bumblebees: Plowright & Owen 1980; beetles: Bocek & Bocak 2016; Brakefield 1985; locusts: Nabours 1929; myriapods: Marek & Bond 2009; nudibranchs: Winters *et al.* 2017), which requires an evolutionary explanation.

Given that the association between prey warning signal and defence should be learned by each generation of predators (Mappes *et al.* 2014), the benefit of signal sharing depends on how often predators encounter the signal. The encounter rate then depends on both the frequency (Heino *et al.* 1998; Müller 1879) and density (Endler & Rojas 2009; Rowland *et al.* 2007; Sword 1999; Müller 1879) of prey carrying the signal. Thus, it is expected that selection on aposematism is positively frequency-dependent (+FDS), with predators avoiding the most common warning signal in a

41  locality (Ruxton *et al.* 2018; Chouteau *et al.* 2016; Comeault & Noonan 2011; Chouteau

42  & Angers 2011; Sherratt 2008).

43      On the other hand, several mechanisms have been proposed to counterbalance

44  selection for signal monomorphism and facilitate warning colour polymorphism

45  (reviewed in Briolat *et al.* 2018). For example, temporally and spatially varying

46  interspecific interactions can result in geographically variable patterns of

47  polymorphism (McLean & Stuart-Fox 2014), particularly when coupled with limited

48  amounts of gene flow between differentially selected populations (e.g. Gordon *et al.*

49  2015; Aubier & Sherratt 2015; Merilaita 2001). Often these mechanisms are thought to

50  act simultaneously, or alternate in time or space (Stevens & Ruxton 2012; Gray &

51  McKinnon 2007; Mallet & Joron 1999) creating a geographic mosaic of selection

52  (Thompson 2005). Although both theoretical (e.g. Holmes *et al.* 2017; Gordon *et al.*

53  2015; Aubier & Sherratt 2015) and experimental work (e.g. Aluthwattha *et al.* 2017;

54  Willink *et al.* 2014) have identified several mechanisms that allow multiple morphs to

55  persist, there is no conclusive evidence from the field and the relative importance of

56  different selective agents is not well understood (Chouteau *et al.* 2016; Stevens &

57  Ruxton 2012). Alas, there is little empirical evidence as to the role of predator

58  communities on local or global morph frequencies of aposematic prey.

59      The variation in the degree of warning colour polymorphism shown by the wood

60  tiger moth (*Arctia plantaginis*) across the western Palearctic provides an excellent

61  system to study how warning signal variation is maintained in the wild (Hegna *et al.*

62  2015). At a local scale, predator community structure (Nokelainen *et al.* 2014) and

63  sexual selection (Gordon *et al.* 2015; Nokelainen *et al.* 2012) have been found to alter

64  the direction of selection on white and yellow male morphs, but no previous studies

65  have addressed selection on a wide geographic scale and including *A. plantaginis*

66    females, which are commonly red. We exposed artificial moths representing the three

67    hindwing colour morphs (white, yellow, red), to local predators in a field experiment

68    spanning across four countries, while monitoring the abundance and community

69    structure of local predator species. We tested whether 1) selection by predators favours

70    the locally common morph; 2) the community structure of avian predators is associated

71    with the predation pressure on different morphs; and 3) there is variation in the direction

72    or strength of selection among populations, matching the local morph frequencies.

73    Variable selection pressure is one of the main candidate mechanisms for the

74    maintenance of polymorphism. By our work, we provide the best-documented case to

75    date of a geographic mosaic of selection on warning signals at broad spatial scales.

76

77    **Material and methods**

78    *Study system*

79    Adult wood tiger moths, *Arctia plantaginis* (Erebidae: Arctiinae; formerly *Parasemia*;

80    see Rönkä *et al.* 2016  for classification), show conspicuous warning colours and

81    possess a chemical defence fluid, which contains pyrazines (Burdfield-Steel *et al.*

82    2018b; Rojas *et al.* 2017) and is a deterrent to avian predators (Burdfield-Steel *et al.*

83    2018a; Rojas *et al.* 2017) . Their warning coloration varies throughout their Holarctic

84    distribution, but local polymorphism is common too (Hegna *et al.* 2015). In the western

85    Palearctic male hindwing colour is either white or yellow, or varies more continuously

86    between yellow and red as seen in females. We selected four study locations that

87    represent the colour variation continuum from monomorphic to polymorphic *Arctia*

88    *plantaginis* populations in the western Palearctic (Figure 1). For the purposes of this

89    study, we consider both sexes to belong to the white, yellow or red morph based on

90    their hindwing colour, and simplify the continuous hindwing coloration of females and

91    Georgian males into two classes: yellow and red (categorized by human eye in 6 grades

92    as in Lindstedt *et al.* 2011), here grades 1-2 are determined yellow and 3-6 red).

93    Accordingly, Scotland is monomorphic with yellow males and females, Georgia is

94    mostly red with 4.3% of males being yellow, and Estonia and Finland are polymorphic

95    with all females caught between 2013-2015 classified as red, and males as either white

96    or yellow (Figure 1).

97        Wood tiger moths are widespread but often low in numbers. Therefore, colour

98    morph frequencies were calculated by population based on annual surveys using

99    pheromone traps and netting between 2009 and 2014 in Scotland, and 2013-2015 in

100   Georgia, Estonia and Finland. Morph frequencies for white and yellow males, and

101   yellow or red males in Georgia, were calculated as the average frequencies from all

102   data available. Morph frequencies for yellow and red females were based on netting

103   data only, as the pheromone traps only lure male moths. Because our dataset was thus

104   biased towards male moths, we corrected the morph frequencies according to a sex ratio

105   of 45 females to 156 males, based on a mark-release-recapture study spanning two years

106   in Central Finland (Gordon *et al*. unpublished data). This sex ratio was used, as it is

107   likely to depict the detectability of each morph more accurately than an even 1:1 sex

108   ratio. The higher frequency of males to females is supported by two observations: male

109   wood tiger moths live longer and fly more actively than females, and the adult sex ratio

110   immediately after eclosion is slightly biased towards males even in laboratory

111   conditions (K. Suisto, personal communication). The concluding morph frequencies

112   (Figure 1A) are consistent with museum samples (Hegna *et al.* 2015) and laboratory

113   stocks originating from the four study populations (Central-Finland, Estonia, Scotland

114   and Georgia).

115

116    *Predation experiment*

117    To estimate the attack risk of white, yellow and red hindwing colour morphs by local

118    predators in the wild we used artificial moth models, resembling real *A. plantaginis*

119    morphs. Models with plasticine (Caran D'Ache Modela 0259.009 Black) bodies

120    attached to printed waterproof (Rite in the Rain ©, JL Darling Corporation, Tacoma,

121    WA, USA) paper wings were prepared following methods described in Nokelainen *et*

122    *al*. (2014). Models were constructed using pictures of one white moth hindwing and

123    two forewings, one with a typical European pattern and another with a typical

124    Caucasian (Georgian) pattern, which were copied and assembled in GIMP 2.8.16

125    SOFTWARE (GNU Image manipulation program) to create six models representing

126    the white, yellow and red morphs in Europe and Georgia (Figure 1B). A locally

127    common forewing type was used to reduce potential novelty effect caused by the

128    forewing pattern (Hegna & Mappes 2014). Resemblance of the artificial models to the

129    real moths was verified by taking measurements of reflectance from the black and

130    coloured areas of real moth wings and printed wings with a Maya2000 Pro spectrometer

131    (Ocean Optics) using a PX-2 Pulsed Xenon Light Source (Ocean Optics) for

132    illumination and adjusting the model colours with Gimp (2.8.16) to match the natural

133    wing colour as closely as possible with a calibrated (HP Colour LaserJet CP2025)

134    printer (see spectral curves of hindwing colour in Rönkä *et al*. 2018), where identical

135    models were used). As our study focused on the hindwing coloration, all other variables

136    such as wing size and pattern were kept constant.

137        We set up 60 predation transects across the four study populations (15 in each

138    country) in open, semi-open and closed natural habitats where the wood tiger moth and

139    its potential avian predators were known or presumed to occur. The predation transects

140    were set at least 500 m apart to avoid birds having overlapping territories between the

141 transects. Along each 900 m transect 20 white, 20 yellow and 20 red artificial moth

142 models were set individually every 15 meters using a randomized block design, so that

143 two models of the same colour would never be next to each other. Models were pinned

144 directly on natural vegetation, either to green leaves large enough to support their

145 weight, or to tree trunks, as visibly as possible. All models were left in the field for a

146 maximum of 6 days (2-6 days, 4 days on average), during the *A. plantaginis* flight

147 season in 2014 (May 31st – July 6th in Estonia, May 26th – July 6th in Finland, June

148 15th – July 30th in Scotland and July 12th – August 3rd in Georgia). Predation events

149 were recorded every 24 hours except for days of heavy rain (as birds were likely not

150 active). For practical reasons (i.e. accessibility of mountain roads and weather

151 conditions) the protocol was modified in Georgia. The 20 white, 20 yellow and 20 red

152 models were set every 10 m totalling up to 600 m, left in the field for 3 consecutive

153 days (72 h), and checked only once.

154  Attacks were recorded based on imprints on the plasticine body and fractures in

155 the wings (see Supplemental Experimental Procedures). Only clear avian attacks were

156 included in the analyses (Supplemental Table 1). Missing and attacked models were

157 replaced with a new model of the same colour to ensure constant morph frequency

158 during the experiment. Excluding or keeping consecutive attacks on the replaced

159 models in the analyses did not markedly change the outcome, reported here (Table 1)

160 for the dataset including replaced models (4004 observations) and for the dataset

161 including original models only (3600 observations; in Supplemental Table 2).

162 Therefore, we kept the replaced models in for all of the analyses, as it increased the

163 sample size.

164

165

166 *Measures of predator community*

167 To estimate the abundances of different insect-feeding birds, which are the most likely

168 predators of wood tiger moths, we counted birds belonging to the orders Passeriformes

169 and Piciformes (Supplemental Table 3). These counts were done once, either before or

170 during the predation experiment, along the predation lines using a modified transect

171 count method (see Nokelainen *et al.* 2014). Bird species observed only in one transect

172 (out of 60), or clearly not adapted to prey on moths (e.g. crossbills), were excluded from

173 analysis. Observations were done within 25 m from the middle of the transect in calm

174 weather between 6 am -1 pm, when birds were most active. Shannon-Wiener diversity

175 index (Figure 3) was calculated using R package 'vegan' 2.5-6 (Oksanen *et al.* 2013).

176

177 *Statistical analyses*

178 To investigate how local predator community affects the direction and strength of

179 selection on wood tiger moth morphs, we constructed generalised linear mixed models.

180 Because the artificial moths were presented to predators over a different number of days

181 in each transect, the attack risk (attacked or not) within a day exposed was used as the

182 response variable for all analyses, modelled with a binomial distribution and a logit link

183 function. First, we tested whether predators select for wood tiger moth warning colours

184 in a frequency-dependent manner across populations (Figure 2). For this, we used local

185 morph frequency calculated from field monitoring data and its interaction with colour

186 morph as the explanatory variables in Model 1 (Table 1). Transect ID, nested within

187 country, was set as the random factor to account for the nested spatial structure of the

188 study design.

189　　　　To test for predator community composition effects, the dimensions of the bird

190 count data, consisting of 12 genera, was first reduced with a principal component

191    analysis using the R function 'princomp'. To avoid overparameterization, the main

192    effects of the first three resulting components (explaining 44.7 %, 33.7 % and 8.5 % of

193    the variation in predator community), and their three-way interactions with morph

194    colour and country, were included one by one as explanatory variables in three separate

195    GLMMs (Table 2). Country was included as an explanatory variable to test for local

196    differences in selection and thus transect ID alone was set as a random effect to each

197    model.

198        Statistical models were simplified using a backward stepwise deletion method

199    based on Akaike Information Criterion. Variables were excluded one by one from the

200    full models and the new model was accepted if the deletion reduced the AIC value more

201    than 2 units, until only main effects or significant interactions were left in each model.

202    All analyses were performed with R (RCoreTeam 2013) in RStudio 0.99.491 (RStudio

203    Team 2015), using the package *lme4* (Bates *et al.* 2015).

204

**Results**

*Positive frequency-dependent selection*

207    Altogether, we observed a total of 718 bird attacks on the 4004 artificial moths. The

208    relative attack risk of each colour morph was lower when the natural frequencies of the

209    respective morph were higher in relation to the others (Table 1, Figure 2). Also, the

210    morphs with intermediate local frequencies show corresponding levels of attack risk

211    (Figure 2). This effect did not depend on colour morph itself (Table 1), as expected if

212    the local predator avoidance depends more on local morph frequency than on morph

213    colour.

214

215 *Predator community*

216 The attacks were not evenly distributed across countries or transects (Figure 2C).

217 Predation pressure varied between and within countries, being highest in Scotland and

218 lowest in Estonia (Figure 2C). Georgia had the lowest amount of insect feeding birds

219 observed (2.1 per 100 meters) compared to Finland (2.6), Scotland (4.0) and Estonia

220 (4.4), respectively. Georgia also had the least diverse predator community measured

221 with Shannon-Wiener diversity index, whereas Estonia was most diverse, followed by

222 Scotland (Figure 3). Across countries, the three most commonly observed potential

223 predators included the common chaffinch, the willow warbler (replaced by green

224 warbler in Georgia) and the great tit (Supplemental Table 3), the latter of which was

225 observed to attack the artificial moths. The first three principal components (PC1, PC2

226 and PC3) that explained 44.7 %, 33.7 % and 8.5 %, respectively, captured 87.0 % of

227 variance in the predator community data. PC1 was dominated by Sylvidae (warblers),

228 Fringillidae (finches) and Muscicapidae (flycatchers), which loaded in the negative end,

229 whereas the positive end of the axis was loaded with Paridae (tits) (Figure 3). PC2 was

230 dominated by Paridae and PC3 with Fringillidae, Muscicapidae and Troglotydidae (the

231 Eurasian wren) (see Supplemental Table 4 for factor loadings).

232

233 *Significant association between predator community structure and selection*

234 A consecutive analysis, where the effect of predator community on the attack risk of

235 each moth colour morph was addressed, revealed a significant three-way-interaction

236 between moth colour, country and PC1 (Model 2, Table 2a, Figure 3). This significant

237 interaction means that the variation in predator community structure captured by PC1

238 is associated with predation pressure on different colour morphs, but the direction of

239    the association is different between countries (i.e. between local communities). PC2

240    and PC3 were not significantly associated with predation pressure (Table 2C and 2D).

241

242    **Discussion**

243    Our experiment is among the first experimental approaches integrating community-

244    level interactions into the study of selection on warning signals (Aluthwattha *et al.*

245    2017; Nokelainen *et al.* 2014; Valkonen *et al.* 2012; Mochida 2011), and the first to do

246    so on such a large geographical scale. With a wide-ranging field experiment spanning

247    populations varying in their degree of polymorphism, we demonstrate that local bird

248    predators avoid locally common morphs, but also that both the strength and direction

249    of selection on warning colour varies geographically. We found that changes in local

250    predator communities drive geographic variation in selection despite positive

251    frequency-dependence. Local predator-prey interactions are thus contributing to the

252    maintenance of both geographic variation and local polymorphism in warning signals.

253        Local avian predators appear as a key in driving warning colour evolution,

254    which can take different evolutionary trajectories over a geographic scale. Here, the

255    predator community had a significant, but different effect to attack risk towards each

256    morph in different countries. The first component from the principal component

257    analysis, explaining 44.7 % of the variation in the abundances of insectivorous birds in

258    different families, significantly affected estimated risk of attack. However, it did so

259    differently towards each morph in the different countries. Interpreting the component

260    loadings and model estimates (Table 2, Figure 3), the Paridae (e.g. tits) and Prunellidae

261    (consisting of only one species, the dunnock, *Prunella modularis*) selected for different

262    morphs in different countries. Our results corroborate the predator community effects

263    found by Nokelainen *et al.* (2014), as we also found that in Finland the yellow morph

264     is favoured in communities characterised by Paridae whereas the white morph is

265     favoured in communities characterised by Prunellidae. In contrast, an opposite effect

266     was found in Scotland where the yellow morph dominates, suggesting that local

267     predators can select for different colours in different populations.

268           Our experiment showed that across countries locally dominating colour morphs

269     were attacked least, as predicted by +FDS. Thus, warning signal efficacy is enhanced

270     with increasing frequency of similarly signalling individuals as predicted due to the

271     number-dependence of predator learning and memorisation. Nonetheless, we found

272     geographic variation in the strength of predator-induced selection. Comparison with

273     previous experiments in those study areas that overlap (Nokelainen et al. 2014) also

274     reveal temporal differences. We found high overall predation pressure in Scotland

275     where the yellow morph was in favour compared to other study locations. Although

276     Nokelainen *et al*. (2014) did not detect positive frequency dependency, they also found

277     much higher overall attack rates in Scotland compared to Southern Finland and Estonia.

278     On the other hand, Nokelainen *et al.* (2014) found that yellow males were significantly

279     less attacked than white males in Southern Finland, whereas in our study the yellow

280     morph tended to have more attacks than the other morphs. Interestingly, the frequency

281     of yellow and white morphs varies in Southern Finland in a biannual cycle (Galarza *et*

282     *al.* 2014), and the yellow morph was more common during Nokelainen *et al*.'s (2014)

283     study, whereas in contrast the white was more common during our experiment,

284     suggesting again that the locally most common morphs have advantage. Temporal

285     fluctuations in local predator-prey interactions could therefore plausibly explain why

286     estimates of predation pressure on different colour morphs conducted in different years

287     have varied.

288    All morphs were attacked at equally low levels in Estonia, which implies spatial

289    variation in the strength of selection or even locally relaxed natural selection on the

290    warning signal. The low predation pressure is not explained by a low number of

291    predators, as there were more insectivorous birds in Estonia than in any other study site

292    (Supplemental Table 3). The bird community composition in Estonia differed from the

293    other countries though, suggesting that the strength of selection was lower in diverse

294    communities characterized by Sylvidae, Fringillidae, Muscicapidae, Turdidae,

295    Troglotydidae and Oriolidae, as opposed to when Paridae (e.g. tits) characterized the

296    community. Other properties of the predator community that can affect the strength of

297    selection on warning signals include the relative abundance of naïve vs. experienced

298    predators (Mappes *et al.* 2014), predators' capacity to learn many different signals

299    (Beatty *et al.* 2004), broad generalisation between the morphs (Sherratt 2008; Balogh

300    & Leimar 2005), conflicting selection by different predators (Nokelainen *et al.* 2014;

301    Valkonen *et al.* 2012), and the spatial arrangement of predators in relation to prey

302    (Endler & Rojas 2009).

303    In temperate regions, most insectivorous birds are migratory and prey

304    population sizes are highly variable due to interseasonal weather variability. This is

305    likely to cause variation in the relative abundances of naïve predators across the

306    breeding season. Furthermore, local seasonal communities are continuously changing,

307    altering the direction and/or strength of selection on warning signals (Mappes *et al.*

308    2014). Siepielski *et al.* (2013) reviewed directional selection on phenotypes, and found

309    that selection tends to vary more in strength than in direction between populations, with

310    most of their examples coming from mid-latitudes in the northern hemisphere. Most

311    experimental evidence of +FDS in the wild, however, comes from tropical systems

312    (Comeault & Noonan 2011; Chouteau & Angers 2011; Mallet & Barton 1989), where

313   the prey and predator community composition is temporally less variable (Mittelbach

314   *et al.* 2007). In such communities, strong +FDS can lead to very accurate mimicry

315   between warning coloured prey, whereas in more variable conditions, higher levels of

316   variation and polymorphism can be maintained.

317        The paradoxical maintenance of local polymorphism despite +FDS could thus

318   be explained by spatial and temporal variation in morph survival combined with

319   individuals migrating between the subpopulations (Gordon *et al.* 2015; McLean &

320   Stuart-Fox 2014; Joron *et al.* 1999). Differences in the level of population isolation,

321   and thus gene flow between them, could explain part of the geographic variation in

322   wood tiger moth warning colours. Population genetic evidence is both for and against

323   this hypothesis: the red-dominated Georgian subspecies *A. p. caucasica* occuring in the

324   mountains of Caucasus is genetically isolated to some degree from the rest of the

325   Western Palearctic samples (Rönkä *et al.* 2016; Hegna *et al.* 2015). However, although

326   the monomorphic yellow population in Scotland is also remote, it clusters together with

327   the Finnish and Estonian samples based on both nuclear and mitochondrial genes

328   (Rönkä *et al.* 2016) and microsatellites (Hegna *et al.* 2015), indicating no restrictions

329   on gene flow. The long-term co-existence of multiple morphs and the low genetic

330   differentiation among polymorphic populations in Finland, Estonia and the Alps with

331   yearly variation in genetic structure (Galarza *et al.* 2014) do suggest a role for gene

332   flow along with varying predation pressure in maintaining local populations at different

333   frequencies.

334        As recently noted by several authors (e.g. Skelhorn *et al.* 2016; Chouteau *et al.*

335   2016; Nokelainen *et al.* 2014), more experimental work is needed to clarify predator-

336   prey interactions at the community level in order to understand how selection is driving

337   the evolution of warning signals in diverse natural ecosystems. Our experiment is so

338    far the most comprehensive analysis showing how spatio-temporal variation in

339    predator-prey communities affects the maintenance of within-species variation and

340    evolutionary pathways to biodiversity.

341

356

357

**References**

Aluthwattha, S.T., Harrison, R.D., Ranawana, K.B., Xu, C., Lai, R. & Chen, J. (2017). Does spatial variation in predation pressure modulate selection for aposematism? *Ecol. Evol.*, 7, 7560–7572.

Aubier, T.G. & Sherratt, T.N. (2015). Diversity in Müllerian mimicry: The optimal predator sampling strategy explains both local and regional polymorphism in prey. *Evolution*, 69, 2831–2845.

Balogh, A.C.V. & Leimar, O. (2005). Müllerian mimicry: an examination of Fisher's theory of gradual evolutionary change. *Proc. Biol. Sci.*, 272, 2269–2275.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R.H.B., Singmann, H., *et al.* (2015). lme4: Linear mixed-effects models using Eigen and S4, 2015. *R package version*, 1.

Beatty, C.D., Beirinckx, K. & Sherratt, T.N. (2004). The evolution of müllerian mimicry in multispecies communities. *Nature*, 431, 63–66.

Beukema, W., Nicieza, A.G., Lourenço, A. & Velo-Antón, G. (2016). Colour polymorphism in Salamandra salamandra (Amphibia: Urodela), revealed by a lack of genetic and environmental differentiation between distinct phenotypes. *J. Zoolog. Syst. Evol. Res.*, 54, 127–136.

Bocek, M. & Bocak, L. (2016). Species limits in polymorphic mimetic Eniclases net-winged beetles from New Guinean mountains (Coleoptera, Lycidae). ZooKeys 593: 15–35.

Borer, M., Van Noort, T., Rahier, M. & Naisbit, R.E. (2010). Positive frequency-dependent selection on warning color in Alpine leaf beetles. *Evolution*, 64, 3629–3633.

Brakefield, P.M. (1985). Polymorphic Müllerian mimicry and interactions with thermal melanism in ladybirds and a soldier beetle: a hypothesis. *Biol. J. Linn. Soc. Lond.*, 26, 243–267.

Brakefield, P.M. & Liebert, T.G. (1985). Studies of colour polymorphism in some marginal populations of the aposematic jersey tiger moth Callimorpha quadripunctaria. *Biol. J. Linn. Soc. Lond.*, 26, 225–241.

Briolat, E.S., Burdfield-Steel, E.R., Paul, S.C., Rönkä, K.H., Seymoure, B.M., Stankowich, T., *et al.* (2018). Diversity in warning coloration: selective paradox or the norm? *Biol. Rev. Camb. Philos. Soc.*

Burdfield-Steel, E., Brain, M., Rojas, B. & Mappes, J. (2018a). The price of safety: food deprivation in early life influences the efficacy of chemical defence in an aposematic moth. *Oikos*.

Burdfield-Steel, E., Pakkanen, H., Rojas, B., Galarza, J.A. & Mappes, J. (2018b). De novo Synthesis of Chemical Defenses in an Aposematic Moth. *J. Insect Sci.*, 18.

Chouteau, M. & Angers, B. (2011). The role of predators in maintaining the geographic organization of aposematic signals. *Am. Nat.*, 178, 810–817.

Chouteau, M., Arias, M. & Joron, M. (2016). Warning signals are under positive frequency-dependent selection in nature. *Proc. Natl. Acad. Sci. U. S. A.*, 113, 2164–2169.

Comeault, A.A. & Noonan, B.P. (2011). Spatial variation in the fitness of divergent aposematic phenotypes of the poison frog, Dendrobates tinctorius. *J. Evol. Biol.*, 24, 1374–1379.

Cott, H.B. (1940). Adaptive coloration in animals.

Dell'aglio, D.D., Stevens, M. & Jiggins, C.D. (2016). Avoidance of an aposematically

408    coloured butterfly by wild birds in a tropical forest. *Ecol. Entomol.*, 41, 627–
409        632.
410   Endler, J.A. & Rojas, B. (2009). The spatial pattern of natural selection when
411        selection depends on experience. *Am. Nat.*, 173, E62–78.
412   Galarza, J.A., Nokelainen, O., Ashrafi, R., Hegna, R.H. & Mappes, J. (2014).
413        Temporal relationship between genetic and warning signal variation in the
414        aposematic wood tiger moth (Parasemia plantaginis). *Mol. Ecol.*, 23, 4939–
415        4957.
416   Gordon, S.P., Kokko, H. & Rojas, B. (2015). Colour polymorphism torn apart by
417        opposing positive frequency-dependent selection, yet maintained in space.
418        *Journal of Animal*.
419   Gray, S.M. & McKinnon, J.S. (2007). Linking color polymorphism maintenance and
420        speciation. *Trends Ecol. Evol.*, 22, 71–79.
421   Greenwood, J.J.D., Cotton, P.A. & Wilson, D.M. (1989). Frequency-dependent
422        selection on aposematic prey: some experiments. *Biol. J. Linn. Soc. Lond.*, 36,
423        213–226.
424   Hegna, R.H., Galarza, J.A. & Mappes, J. (2015). Global phylogeography and
425        geographical variation in warning coloration of the wood tiger moth (
426        Parasemia plantaginis ). *J. Biogeogr.*, 42, 1469–1481.
427   Hegna, R.H. & Mappes, J. (2014). Influences of geographic differentiation in the
428        forewing warning signal of the wood tiger moth in Alaska. *Evol. Ecol.*
429   Heino, M., Metz, J.A. & Kaitala, V. (1998). The enigma of frequency-dependent
430        selection. *Trends Ecol. Evol.*, 13, 367–370.
431   Holmes, I.A., Grundler, M.R. & Davis Rabosky, A.R. (2017). Predator Perspective
432        Drives Geographic Variation in Frequency-Dependent Polymorphism. *Am.*
433        *Nat.*, 190, E78–E93.
434   Joron, M., Wynne, I.R., Lamas, G. & Mallet, J. (1999). Variable Selection and the
435        Coexistence of Multiple mimetic forms of the Butterfly Heliconius numata.
436        *Evol. Ecol.*, 13, 721–754.
437   Kapan, D.D. (2001). Three-butterfly system provides a field test of müllerian
438        mimicry. *Nature*, 409, 338–340.
439   Lindstedt, C., Eager, H., Ihalainen, E., Kahilainen, A., Stevens, M. & Mappes, J.
440        (2011). Direction and strength of selection by predators for the color of the
441        aposematic wood tiger moth. *Behav. Ecol.*, 22, 580–587.
442   Lindström, L., Alatalo, R.V., Lyytinen, A. & Mappes, J. (2001). Strong antiapostatic
443        selection against novel rare aposematic prey. *Proc. Natl. Acad. Sci. U. S. A.*,
444        98, 9181–9184.
445   Mallet, J. & Barton, N.H. (1989). Strong natural selection in a warning-color hybrid
446        zone. *Evolution*, 43, 421–431.
447   Mallet, J. & Joron, M. (1999). Evolution of Diversity in Warning Color and Mimicry:
448        Polymorphisms, Shifting Balance, and Speciation. *Annu. Rev. Ecol. Syst.*, 30,
449        201–233.
450   Mappes, J., Kokko, H., Ojala, K. & Lindström, L. (2014). Seasonal changes in
451        predator community switch the direction of selection for prey defences. *Nat.*
452        *Commun.*, 5, 5016.
453   Mappes, J., Marples, N. & Endler, J.A. (2005). The complex business of survival by
454        aposematism. *Trends Ecol. Evol.*, 20, 598–603.
455   Marek, P.E. & Bond, J.E. (2009). A Müllerian mimicry ring in Appalachian
456        millipedes. *Proc. Natl. Acad. Sci. U. S. A.*, 106, 9755–9760.
457   McLean, C.A. & Stuart-Fox, D. (2014). Geographic variation in animal colour

458        polymorphisms and its role in speciation. *Biol. Rev. Camb. Philos. Soc.*

459  Merilaita, S. (2001). Habitat heterogeneity, predation and gene flow: colour
460        polymorphism in the isopod, Idotea baltica. *Evol. Ecol.*, 15, 103–116.

461  Merrill, R.M., Dasmahapatra, K.K., Davey, J.W., Dell'Aglio, D.D., Hanly, J.J.,
462        Huber, B., *et al.* (2015). The diversification of Heliconius butterflies: what
463        have we learned in 150 years? *J. Evol. Biol.*, 28, 1417–1438.

464  Mittelbach, G.G., Schemske, D.W., Cornell, H.V., Allen, A.P., Brown, J.M., Bush,
465        M.B., *et al.* (2007). Evolution and the latitudinal diversity gradient: speciation,
466        extinction and biogeography. *Ecol. Lett.*, 10, 315–331.

467  Mochida, K. (2011). Combination of local selection pressures drives diversity in
468        aposematic signals. *Evol. Ecol.*, 25, 1017.

469  Müller, F. (1878). *Ueber die Vortheile der Mimicry bei Schmetterlingen.*

470  Müller, F. (1879). Ituna and Thyridia: a remarkable case of mimicry in butterflies.
471        *Trans. Entomol. Soc. Lond*, 1879, 20–29.

472  Nabours, R.K. (1929). The Genetics of the Tettigidae (Grouse Locusts). In: *The*
473        *Genetics of the Tettigidae (Grouse Locusts)* (ed. Nabours, R.K.). Springer
474        Netherlands, Dordrecht, pp. 27–100.

475  Nokelainen, O., Hegna, R.H., Reudler, J.H., Lindstedt, C. & Mappes, J. (2012).
476        Trade-off between warning signal efficacy and mating success in the wood
477        tiger moth. *Proc. Biol. Sci.*, 279, 257–265.

478  Nokelainen, O., Valkonen, J., Lindstedt, C. & Mappes, J. (2014). Changes in predator
479        community structure shifts the efficacy of two warning signals in Arctiid
480        moths. *J. Anim. Ecol.*, 83, 598–605.

481  Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'hara, R.B., *et*
482        *al.* (2013). *Vegan: Community Ecology Package.* R package version 2.5-6.
483        https://CRAN.R-project.org/package=vegan

484  Plowright, R.C. & Owen, R.E. (1980). THE EVOLUTIONARY SIGNIFICANCE OF
485        BUMBLE BEE COLOR PATTERNS: A MIMETIC INTERPRETATION.
486        *Evolution*, 34, 622–637.

487  Poulton, E.B. (1890). *The colours of animals: their meaning and use, especially*
488        *considered in the case of insects*. Kegan Paul, Trench, Trubner.

489  RCoreTeam. (2013). *R: A language and environment for statistical computing.* R
490        Foundation for Statistical Computing, Vienna, Austria.

491  Rojas, B. (2017). Behavioural, ecological, and evolutionary aspects of diversity in
492        frog colour patterns. *Biol. Rev. Camb. Philos. Soc.*, 92, 1059–1080.

493  Rojas, B., Burdfield-Steel, E., Pakkanen, H., Suisto, K., Maczka, M., Schulz, S., *et al.*
494        (2017). How to fight multiple enemies: target-specific chemical defences in an
495        aposematic moth. *Proc. Roy. Soc. B*, 284, 20171424.

496  Rönkä, K., De Pasqual, C., Mappes, J., Gordon, S. & Rojas, B. (2018). Colour alone
497        matters: no predator generalization among morphs of an aposematic moth.
498        *Anim. Behav.*

499  Rönkä, K., Mappes, J., Kaila, L. & Wahlberg, N. (2016). Putting Parasemia in its
500        phylogenetic place: a molecular analysis of the subtribe Arctiina (Lepidoptera)
501        : Molecular phylogeny of Arctiina. *Syst. Entomol.*, 41, 844–853.

502  Rowland, H.M., Ihalainen, E., Lindström, L., Mappes, J. & Speed, M.P. (2007). Co-
503        mimics have a mutualistic relationship despite unequal defences. *Nature*, 448,
504        64–67.

505  RStudio Team. (2015). *RStudio: integrated development for R* (Version 0.99..441)
506        [Computer software]. Boston, MA. Available from http://www.rstudio.org/.

507  Ruxton, G.D., Allen, W.L., Sherratt, T.N. & Speed, M.P. (2018). *Avoiding Attack:*

508          *The Evolutionary Ecology of Crypsis, Aposematism, and Mimicry*. Oxford
509               University Press.
510 Sherratt, T.N. (2008). The evolution of Müllerian mimicry. *Naturwissenschaften*, 95,
511               681–695.
512 Siddiqi, A., Cronin, T.W., Loew, E.R., Vorobyev, M. & Summers, K. (2004).
513               Interspecific and intraspecific views of color signals in the strawberry poison
514               frog Dendrobates pumilio. *J. Exp. Biol.*, 207, 2471–2485.
515 Siepielski, A.M., Gotanda, K.M., Morrissey, M.B., Diamond, S.E., DiBattista, J.D. &
516               Carlson, S.M. (2013). The spatial patterns of directional phenotypic selection.
517               *Ecol. Lett.*, 16, 1382–1392.
518 Skelhorn, J., Halpin, C.G. & Rowe, C. (2016). Learning about aposematic prey.
519               *Behav. Ecol.*, 27, 955–964.
520 Stevens, M. & Ruxton, G.D. (2012). Linking the evolution and form of warning
521               coloration in nature. *Proc. Biol. Sci.*, 279, 417–426.
522 Sword, G.A. (1999). Density-dependent warning coloration. *Nature*, 397, 217–217.
523 Thompson, J. N. (2005). *The geographic mosaic of coevolution*. University of
524               Chicago Press.
525 Valkonen, J.K., Nokelainen, O., Niskanen, M., Kilpimaa, J., Björklund, M. &
526               Mappes, J. (2012). Variation in predator species abundance can cause variable
527               selection pressure on warning signaling prey. *Ecol. Evol.*, 2, 1971–1976.
528 Wallace, A. (1867). *Proceedings of the Entomological Society of London* March 4th,
529               Ixxx–Ixxxi.
530 Willink, B., García-Rodríguez, A., Bolaños, F. & Pröhl, H. (2014). The interplay
531               between multiple predators and prey colour divergence. *Biol. J. Linn. Soc.*
532               *Lond.*, 113, 580–589.
533 Winters, A.E., Green, N.F., Wilson, N.G., How, M.J., Garson, M.J., Marshall, N.J., *et*
534               *al.* (2017). Stabilizing selection on individual pattern elements of aposematic
535               signals. *Proc. Biol. Sci.*, 284.
536

537 **Figures**
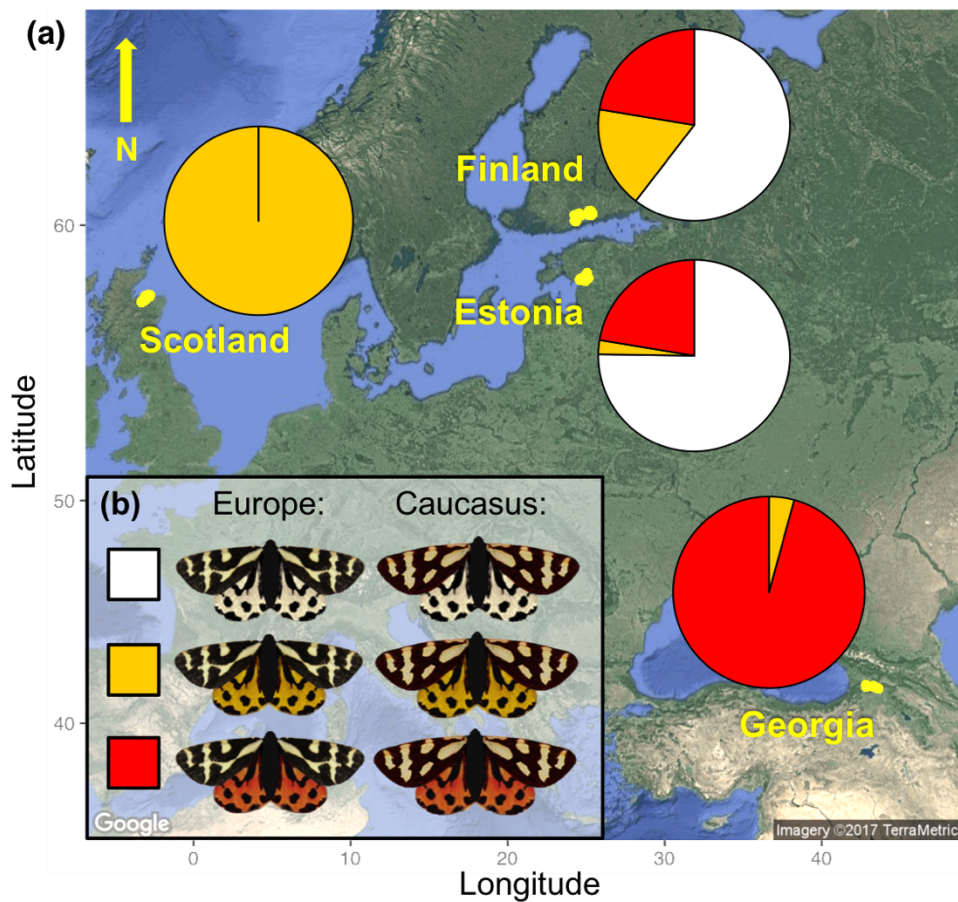


538

539

540 **Figure 1.** (a) Study populations and local morph frequencies and (b) moth models used

541 in Europe and Georgia representing a local forewing pattern and white, yellow or red

542 hindwing morph. At the monomorphic end of the monomorphic-polymorphic

543 continuum in Scotland, both sexes have yellow hindwing coloration. In Georgia red is

544 the dominating hindwing colour, but male coloration varies continuously towards

545 yellow. In Estonia, white hindwing colour dominates, as the males are almost

546 exclusively white and females red. In Southern Finland all three colour morphs are

547 present (white and yellow male morphs and females vary continuously from yellow to

548 red).

549
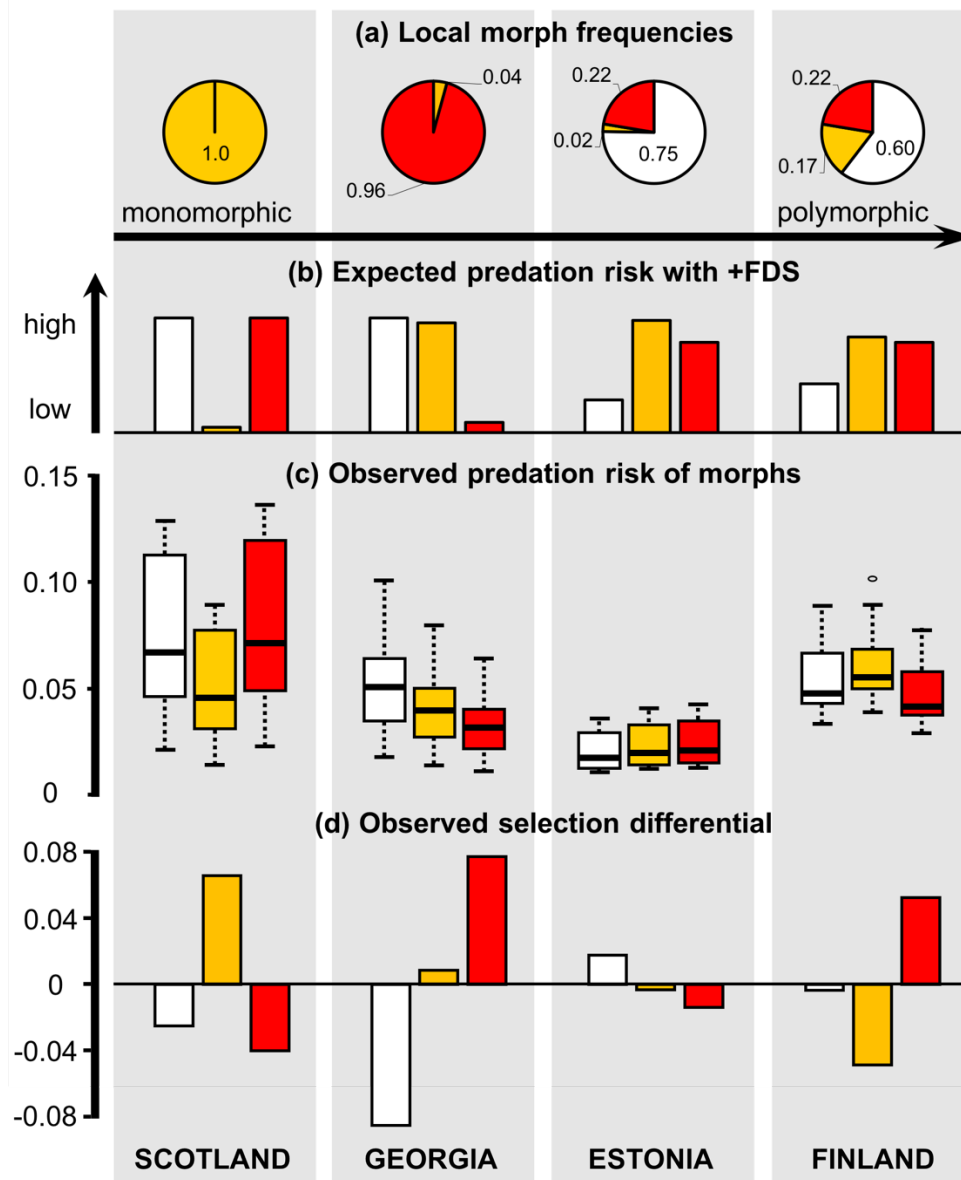
**Figure 2.** Wood tiger moth morph frequencies compared to expected and observed predation risk and selection differential by country. (a) Local morph frequencies calculated from annual monitoring data, (b) expected attack risk according to the +FDS hypothesis, where each morph is attacked according to its local frequency, (c) observed predation illustrated as GLMM estimates of daily attack risk for each morph by country and (d) observed difference in attacks per morph compared to a situation where all morphs would be attacked equally. Morph colours (white, yellow, red) as in Figure 1.

559



560
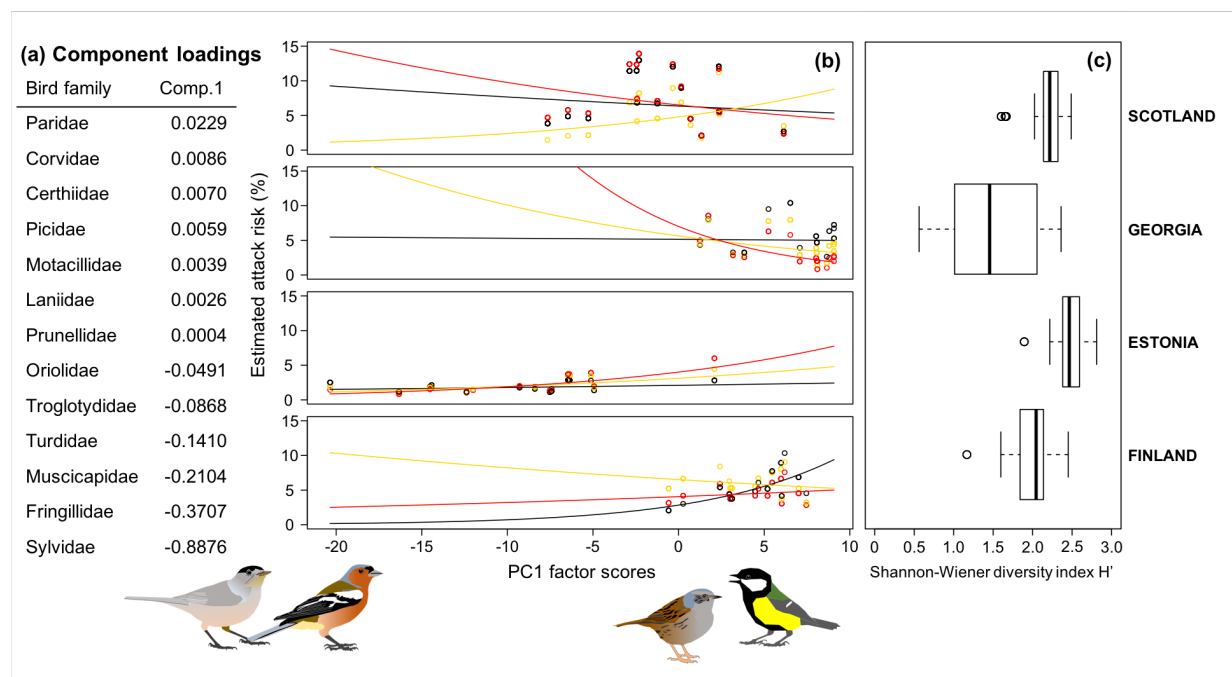
**Figure 3**. Community composition and diversity of insectivorous birds per population. Factor scores and loading of the first principal component describing 44.7% of the total variation of bird communities across countries. Panel (a) family level component loadings, panel (b) the three-way interaction effect of predator community on estimated attack risk of different colour morphs (black line corresponds to the white morph) at each transect illustrated by population, panel (c) Shannon-Wiener diversity indexes calculated per transect and plotted by population. The x-axes on both panels b and c represent factor scores for principal component 1.

569

570 **Tables**

571 **Table 1**. Positive frequency-dependency of the estimated attack risk. (a) The model

572 including only main effects of morph frequency and morph colour (underlined) was

573 selected because the reduction in AIC value compared to a full model with both main

574 effects and interaction between morph frequency and colour is higher than 2 ($\Delta$ AIC =

575 3.4). (b) Estimates of the best-fit model (Model 1a). Values of significance level <0.05

576 are bolded. $\Delta$ df denotes change in model degrees of freedom.

577

| (a) Model selection | $\Delta$ df | LRT | Pr(Chi) | model AIC |
|---|---|---|---|---|
| colour * morph frequency | | | | 3957.0 |
| colour + morph frequency | 2 | 0.1949 | 0.907 | 3953.2 |

The asterisk (*) denotes both main effects and interaction terms used.

578

**(b) Model 1a**

| Random effects | Variance | SD |
|---|---|---|
| transect within country | 0.3315 | 0.5758 |
| country | 0.1633 | 0.4041 |

| Fixed effects | Estimate | SE | Z-value | p-value |
|---|---|---|---|---|
| (Intercept): colour[w] | -3.0433 | 0.2275 | -13.376 | **<0.001** |
| colour[y] | -0.0923 | 0.0940 | -0.982 | 0.3259 |
| colour[r] | -0.0841 | 0.0925 | -0.909 | 0.3633 |
| morph frequency | -0.3728 | 0.1071 | -3.481 | **0.0005** |

579

580  **Table 2.** The interaction effect of predator community and location (country, C) on the

581  attack risk towards the wood tiger moth colour morphs. (a) Model selection starting

582  from the main effects, interactions and a three-way interaction between principal

583  component 1 (PC1), country (C) and colour morph (colour) as the explanatory

584  variables, with the best-fit model underlined. (b) Estimates from the selected model

585  (Model 2) with a significant three-way interaction of principal component 1, colour

586  morph and country. (c) Model selection for principal component 2. Principal

587  component 2 had no significant effects on attack risk, and thus estimates are not shown.

588  (d) Model selection for principal component 3. Principal component 3 had no

589  significant effects on attack risk, and thus estimates are not shown. Values of

590  significance level <0.05 are bolded. Δ df denotes change in model degrees of freedom.

591

| (a) Model selection with PC1 | Δ df | LRT | Pr(Chi) | model AIC |
|---|---|---|---|---|
| <u>PC1*colour*C</u> | | | | 3956.1 |
| PC1+colour+C+PC1:colour+PC1:C+colour:C | 6 | 14.35 | **0.026** | 3958.4 |

592  The asterisk (*) denotes both main effects and interaction terms used.

**(b) Model 2**

| Random effects | Variance | SD | | |
|---|---|---|---|---|
| transect | 0.2779 | 0.5272 | | |

| Fixed effects | Estimate | SE | Z-value | p-value[1)] |
|---|---|---|---|---|
| (Intercept): colour[w], C[Finland] | -3.556 | 0.438 | -8.126 | **<0.001** |
| colour[y] | 0.869 | 0.407 | 2.134 | **0.033** |
| colour[r] | 0.370 | 0.440 | 0.840 | 0.401 |
| PC1 | 0.139 | 0.088 | 1.590 | 0.112 |
| C[Estonia] | -0.309 | 0.621 | -0.498 | 0.619 |
| C[Georgia] | 0.618 | 0.647 | 0.955 | 0.340 |
| C[Scotland] | 0.843 | 0.477 | 1.766 | 0.077 |
| PC1 : colour[y] | -0.164 | 0.083 | -1.989 | **0.047** |
| PC1 : colour[r] | -0.115 | 0.089 | -1.299 | 0.194 |

| | | | | |
|---|---|---|---|---|
| C[Estonia] : colour[y] | -0.463 | 0.606 | -0.765 | 0.444 |
| C[Estonia] : colour[r] | 0.297 | 0.618 | 0.481 | 0.631 |
| C[Georgia] : colour[y] | -0.773 | 0.592 | -1.306 | 0.192 |
| C[Georgia] : colour[r] | -0.041 | 0.616 | -0.067 | 0.947 |
| C[Scotland] : colour[y] | -1.169 | 0.445 | -2.624 | **0.009** |
| C[Scotland] : colour[r] | -0.345 | 0.472 | -0.730 | 0.466 |
| PC1 : C[Estonia] | -0.123 | 0.097 | -1.273 | 0.203 |
| PC1 : C[Georgia] | -0.143 | 0.111 | -1.280 | 0.201 |
| PC1 : C[Scotland] | -0.159 | 0.103 | -0.551 | 0.121 |
| C[Estonia] : PC1 : colour[y] | 0.197 | 0.093 | 2.115 | **0.035** |
| C[Estonia] : PC1 : colour[r] | 0.176 | 0.099 | 1.781 | **0.075** |
| C[Georgia] : PC1 : colour[y] | 0.105 | 0.105 | 0.991 | 0.322 |
| C[Georgia] : PC1 : colour[r] | -0.033 | 0.113 | -0.294 | 0.769 |
| C[Scotland] : PC1 : colour[y] | 0.256 | 0.100 | 2.550 | **0.011** |
| C[Scotland] : PC1 : colour[r] | 0.091 | 0.102 | 0.892 | 0.372 |

593

| **(c) Model selection with PC2** | **Δ df** | **LRT** | **Pr(Chi)** | **model AIC** |
|---|---|---|---|---|
| PC2*colour*country | | | | 3962.6 |
| PC2+colour+C+PC2:colour+PC2:C+colour:C | 6 | 4.413 | 0.621 | 3955.0 |
| <u>PC2+colour+C+PC2:colour+colour:C</u> | 3 | 3.334 | 0.343 | 3952.4 |
| PC2+colour+C+colour:C | 2 | 3.923 | 0.141 | 3952.3 |
| PC2+colour+C | 6 | 16.737 | **0.010** | 3957.0 |
| colour*country | 1 | 0.545 | 0.460 | 3950.8 |

| **(d) Model selection with PC3** | **Δ df** | **LRT** | **Pr(Chi)** | **model AIC** |
|---|---|---|---|---|
| PC3*colour*country | | | | 3963.6 |
| PC3+colour+C+PC3:colour+PC3:C+colour:C | 6 | 5.190 | 0.520 | 3956.8 |
| <u>PC3+colour+C+PC3:colour+colour:C</u> | 3 | 1.400 | 0.706 | 3952.2 |
| PC3+colour+C+colour:C | 2 | 4.599 | 0.100 | 3952.8 |
| PC3+colour+C | 6 | 16.726 | **0.010** | 3957.5 |
| colour*country | 1 | 0.034 | 0.854 | 3950.8 |

594