The relationship between mimetic imperfection and phenotypic variation in insect colour patterns

Graham Holloway\textsuperscript{1*}, Francis Gilbert\textsuperscript{2} and Amoret Brandt\textsuperscript{1}

\textsuperscript{1} Division of Zoology, AMS Building, University of Reading, Reading, Berkshire RG6 6AJ, UK
\textsuperscript{2} Department of Life Sciences, University of Nottingham, Nottingham NG7 2RD, UK

Many hoverflies (Syrphidae) mimic wasps or bees through colour or behavioural adaptations. The relationship between phenotypic variation in colour pattern and mimetic perfection (as determined by pigeons) was investigated in three species of Mullerian mimics (\textit{Vespula} spp.) and 10 Batesian hoverfly mimics, plus two non-mimetic species of flies. Four predictions were tested: (i) Batesian mimics might be imperfect because they are in the process of evolving towards perfection, hence there should be a positive relationship between variation and imperfection; (ii) some Batesian mimics are imperfect because they do not have the appropriate genetic variation to improve and have evolved to be as good as possible, hence there should be no differences between species, all displaying a low level of variation; (iii) very common hoverflies should show the highest levels of variation because they outnumber their models, resulting in high predation and a breakdown in the mimetic relationship; and (iv) social wasps (\textit{Vespula}) have such a powerful defence that anything resembling a wasp, both Mullerian and perfect Batesian mimics, would be avoided, resulting in relaxed selection and high variance. Poor mimics may still evolve to resemble wasps as well as possible and display lower levels of variation. The data only provided support for the fourth prediction. The Mullerian mimics, one of the most perfect Batesian mimics, and the non-mimetic flies displayed much higher levels of variation than the other species of Batesian mimics.

Keywords: mimetic imperfection; colour variation; mimicry; Batesian; Mullerian; Syrphidae

1. INTRODUCTION

Chemically defended or otherwise noxious insects often have bright warning, or aposmatic, colour patterns (e.g. Wickler 1968); well-known examples include many of the Hymenoptera. Rather than evading predators, aposmatically coloured insects flaunt themselves, encouraging predators to learn that the coloration they carry is associated with a sting, bad taste or other unpleasant physiological reaction. The existence of bright aposmatic coloration has provided a fascinating challenge to biologists to explain how something so conspicuous could evolve in noxious animals (Sheppard 1975; Schuler & Hesse 1985; Huheey 1988; Guilford 1990; Mallet & Joron 2000; Rowe 2001). Many of these insect species have evolved colour patterns that resemble each other and in doing so have become Mullerian mimics (Muller 1878). As all Mullerian mimics are defended in some way, the colour pattern serves to reinforce the message to potential predators that species with these patterns should be avoided. It has been argued that sympatric Mullerian mimics sharing a particular colour pattern should evolve to resemble each other very closely to minimize the chance of mistakes being made by the predator (Ackery & Vane-Wright 1984).

The colour patterns adopted by Mullerian models have also evolved in palatable insects, dubbed Batesian mimics after Bates (1862) who first described the phenomenon. The function of Batesian mimetic colour patterns is to trick the predator into believing that all insects bearing an aposmatic colour pattern are noxious and should be avoided. Consequendy, it has also been argued that the signal from Batesian mimics should resemble the signal from their Mullerian models very closely in order for the mimicry to be advantageous (Mappes & Alatalo 1997). Many hoverflies (Diptera: Syrphidae) are Batesian mimics of either wasps (e.g. \textit{Vespula} spp.), honeybees (\textit{Apis mellifera}) or bumble-bees (\textit{Bombus} spp.). On the one hand, within the Syrphidae, there is a range of mimetic perfection (at least to the human eye) (Waldbauer 1988; Howarth et al. 2000). For example, \textit{Eristalis tenax} and \textit{Sericomyia silentis} resemble honeybees and social wasps, respectively, very closely and the two main colour morphs of \textit{Volucella bombylans} are excellent mimics of red-tailed (\textit{B. lapidarius}) and buff-tailed (\textit{B. terrestris}) bumble-bees (see Stubbs & Falk 1983). The degree of mimetic perfection often extends beyond colour pattern to include morphology and behaviour (Nicholson 1927; Gaul 1952; Golding & Edmunds 2000). On the other hand, there are several hoverfly species that do not resemble their supposed model very closely (Edmunds 2000; Howarth et al. 2000).

Numerous suggestions have been forwarded to explain why there are good and poor mimetic hoverflies (Edmunds 2000). Evolution towards mimetic perfection proceeds through a process known as normalizing selection, where phenotypes deviating from the most common form are selected out of the population through predation. This clearly implies that the most perfect mimics should display very low levels of phenotypic variation in colour pattern at the population level. However, no examination of the relationship between the degree of mimetic imperfection and level of phenotypic variation has been carried out. One possibility is that there should be a positive correlation between the two factors in Batesian mimics, with the most imperfect mimics displaying the highest level of...
variation (figure 1a). This could occur if mimics were imperfect only because they were in the process of evolving perfection (Edmunds 2000). Mullerian mimic colour patterns might show low levels of variation since these are likely to have had longer to converge (Mullerian patterns must have evolved first in order to be copied by Batesian mimics).

However, perfection can only be achieved if a population of Batesian mimics contains appropriate genetic variation. An imperfect mimic could have evolved to become as good as it can, prior to the appearance of novel mutations that offer the opportunity for further development. In this case there may be little correlation between degree of imperfection and phenotypic variation, although the levels of variation measured might be low irrespective of the degree of mimetic perfection (figure 1b).

It has been noted that many of the commonest hoverflies have imperfect mimic colour patterns. If mimics become very common (Sheppard 1959) or extend their range beyond that of their model (Carpenter & Ford 1933), the frequency of mistakes increases, making it more difficult for predators to associate the colour pattern with distastefulness, and mimicry breaks down. Hoverfly species such as Episyrphus balteatus and Syrphus ribesii are very common and are often assumed to be imperfect mimics of social wasps, or possibly Nomada spp. bees in the case of E. balteatus (Howarth 2000). Azmeh et al. (1998) noted that many of the very common hoverflies displaying imperfect mimicry are aphidophagous. Man-induced habitat changes have greatly increased the number of aphids available and consequently some hoverfly species have benefited. Azmeh et al. (1998) argue that as a result of the increase in abundance of certain aphidophagous hoverfly species, mimetic relationships have broken down in the manner proposed by Sheppard (1959). Hence a third possibility is that common, aphidophagous imperfectly mimetic hoverflies should display higher levels of phenotypic variation as a result of relaxed selection for perfection (figure 1c). Figure 1c illustrates two such species with higher levels of variation. No prediction can be made about where along the x-axis the species might occur, or which one of the two is likely to be more variable.

Finally, social wasps possess a powerful sting and it is possible that predators will avoid even imperfect mimics because of the degree of noxiousness of the model (Duncan & Sheppard 1965; Lindström et al. 1997). If this is the case, very noxious Mullerian mimics may not be under selective pressure to evolve to resemble each other very closely, which is in contrast to the proposal described by Ackery & Vane-Wright (1984). It follows that Batesian mimics that can easily be mistaken for social wasps could also display a similar level of phenotypic variation and still gain substantial protection (Edmunds 2000). Batesian mimics resembling social wasps less closely might show less phenotypic variation than social wasps and their good Batesian mimics, because they have evolved to become as good as possible with the genetic variation available to them. Therefore, a scenario similar to that depicted in figure 1d might be found.

The purpose of the current study was to examine levels of phenotypic variation in colour pattern in social wasps, a variety of supposed Batesian hoverfly mimics of social wasps and non-mimetic insects, in relation to the rank of

![Figure 1](image-url)
mimicry imperfection, to test the four hypotheses described above. Using operant conditioning methods, the hoverfly species studied here had been ranked previously by birds according to their similarity to a wasp model (Dittrich et al. 1993). Ranking using this system had several advantages over ranking by the human eye. For example, the most likely predators of hoverflies are birds and birds may not view insect colour patterns in the same way that humans do. Also, because the experimental birds had not experienced the discomfort of a sting from a social wasp, the Batesian mimetic colour patterns were ranked for perfection without noxiousness as a potential confounding factor.

2. MATERIAL AND METHODS

(a) The models and mimics

As far as possible, the same insect species used by Dittrich et al. (1993) were used in the present study. The hoverfly species used are thought to mimic social wasps (Vespa spp.) to varying degrees. Three species of wasp models were used: V. vulgaris, V. sylvestris and V. germanica. Rather than using slides, as Dittrich et al. (1993) did, samples of each model species were collected from the wild in the vicinity of the University of Reading, Berkshire, UK (Ordnance Survey grid reference SU7371) throughout 1995 and 1996, or obtained from the Natural History Museum in London. Ten mimetic hoverfly species were analysed: S. ribesii, Temnostoma vespariforme, Scaeva pyrastri, Chrysoxythus bicinctum, Chrysoxythus castaum, Helophilus pendulus, Epistrophe grossulariae, Volucella zonaria, Xanthogramma pedissequum and Ichyrosyrphus glaucius. Temnostoma vespariforme, C. castaum and I. glaucius are rare, or absent from around Reading, so measurements for these specimens were derived wholly from preserved specimens. Two non-mimetic flies were used as controls: Tachina fera and Eriothrix rufo maculata. Neither of these flies is believed to be mimetic, but both have orangy-red markings on the abdomen. Samples of both species were obtained from the Natural History Museum.

(b) Image analysis

After capture from the field, all insects were placed in a freezer at −18°C. The maximum time between capture and freezing never exceeded 2 h. The insects were later removed from the freezer, one at a time, to put their images into a Macintosh Quadra 840AV. The insects were illuminated with white light (two 100 W bulbs), and individually laid out carefully under a JVC TK-1270 camera that interfaced directly with the computer. The images were captured using a Data Translation DT 2255 Quick Capture Board (www.datatranslation.com) and stored on hard disk and DDS2 tapes. Length calibrations were performed using a micrometer placed under the camera so that a known distance could be converted into number of pixels, and were carried out only at the start of each session, thereafter the capture magnification was altered each time. Length calibrations were stored along with each captured image so that, during measurement, linear distances were displayed in mm and areas as mm². The images were captured onto an area of 393 216 (768 x 512) pixels, which almost filled a 17 inch screen. The magnification was adjusted so that the image of each insect was as large as possible within this area.

Image analysis was carried out using NIH IMAGE (v. 1.52). The first two large tergites (2 and 3) from the insects provided the most accurate measurements of the proportion of yellow pigmentation and together accounted for most of the dorsal surface area of the abdomen in the Syrphidae. Tergites further along the abdomen are smaller and curved downwards more, making focusing difficult. The two non-mimetic flies had large squamae that obscured some of tergite 2 on many of the specimens. In these flies, only the area of pigmentation on tergite 3 was measured. Tergites 2 and 3 were analysed separately for surface area and the proportions of these areas that were yellow (or otherwise coloured). As the specimen ages, the brightness of the yellow pigmentation often fades, but the extent of the colouration does not change. Some frame-grabbers are known to distort colour measurements. The frame-grabber used here (DT 2255 Quick Capture Board) does not cause such colour distortions, but, although unlikely, it was not known whether automated area measurement was also subject to certain distortions. Consequently, the areas to be measured were delineated by hand which, as well as being reliable, was also more accurate (although more time consuming) than an automated procedure.

Our estimation of the degree of resemblance between model and mimic was derived from the work reported by Dittrich et al. (1993) and Green et al. (1999). Dittrich et al. (1993) trained pigeons to peck at one type of insect (model/ non-mimic) and not at its opposite (non-mimic/model) and then presented the birds with a range of images of hoverfly mimics. The pigeons pecked at these hoverfly images at various rates. The rank order of these peck rates was used to estimate resemblance and the same rank order is used in the present study. Cuthill & Bennett (1993) criticized this ranking because the presentations used colour slides, which lack UV components to which pigeons are sensitive. They claimed that the pigeons may have seen these images very differently as compared with real insects. However, this view is invalid since the ranking is the same with real insects (Green et al. 1999) and there are no UV components in wasp or hoverfly colour patterns (Nickol 1994; Green et al. 1999).

Our estimate for the variation in the colour pattern was measured as the standard deviation of the proportion of paler pigmentation on any single tergite. All data were analysed using MINITAB for Windows 95 release 10.51 Xtra and MICROSOFT EXCEL for Windows 95 v. 7.0.

3. RESULTS

In total, 796 hoverflies, 276 wasps and 169 non-mimetic flies were analysed. The proportions of tergites 2 and 3 covered in pale pigmentation varied from over 0.2 to less than 0.7. It was therefore not considered necessary to arcsine transform the data prior to analysis. There was a high correlation between the amount of variation in the proportion of yellow pigmentation on tergite 2 and tergite 3 within each species (males: r = 0.732, females: r = 0.883). In order to obtain a single estimate of variation for each species for each sex, the standard deviation about the mean proportion of pale pigmentation over the two tergites combined was calculated. Table 1 shows the proportion of yellow pigmentation covering tergites 2 and 3 and the estimate of variation (standard deviation) about the mean. The mean within-wasp species abdominal colour pattern variation was greater than the variation shown by all of the mimetic fly species, except one (T. vespariforme). The probability of the wasp variation exceeding nine or more of the abdominal colour pattern variations of the 10 fly species is p < 0.01. Individual F-tests showed the mean wasp variation to be significantly greater than the variation
Table 1. Proportions of tergites 2 and 3 of a range of species of social wasps (workers) (M), mimetic hoverflies and non-mimetic flies covered with pale pigmentation and the standard deviation of the proportion of tergites 2 and 3 covered with pale pigmentation.
(The sample size \( n \) is given. Rank refers to the degree of perfection of mimicry of the wasp colour pattern according to Dittrich et al. (1993). Note: Chrysotoxum bicinctum has no yellow on tergite 3 and Vespula zonaria has no yellow on tergite 2. The large squamae of the non-mimetic (N) flies interfered with the measurement of tergite 2 pigmentation.)

<table>
<thead>
<tr>
<th>species</th>
<th>rank</th>
<th>( n )</th>
<th>sex</th>
<th>tergite 2</th>
<th>tergite 3</th>
<th>standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Vespula germanica</em></td>
<td>M</td>
<td>86</td>
<td>F</td>
<td>0.681</td>
<td>0.512</td>
<td>0.0618</td>
</tr>
<tr>
<td><em>Vespula zelaestris</em></td>
<td>M</td>
<td>91</td>
<td>F</td>
<td>0.578</td>
<td>0.379</td>
<td>0.0493</td>
</tr>
<tr>
<td><em>Vespula vulgaris</em></td>
<td>M</td>
<td>99</td>
<td>F</td>
<td>0.603</td>
<td>0.437</td>
<td>0.0858</td>
</tr>
<tr>
<td><em>Syrphus ribesi</em></td>
<td>1</td>
<td>36</td>
<td>M</td>
<td>0.339</td>
<td>0.424</td>
<td>0.0337</td>
</tr>
<tr>
<td><em>Tennostoma vespiforme</em></td>
<td>2</td>
<td>11</td>
<td>M</td>
<td>0.467</td>
<td>0.602</td>
<td>0.0721</td>
</tr>
<tr>
<td><em>Chrysotoxum cautum</em></td>
<td>3</td>
<td>44</td>
<td>M</td>
<td>0.400</td>
<td>0.592</td>
<td>0.0485</td>
</tr>
<tr>
<td><em>Helophilus pendulus</em></td>
<td>4</td>
<td>63</td>
<td>M</td>
<td>0.551</td>
<td>0.626</td>
<td>0.0534</td>
</tr>
<tr>
<td><em>Epistrophe grossulariae</em></td>
<td>5</td>
<td>38</td>
<td>M</td>
<td>0.430</td>
<td>0.562</td>
<td>0.0401</td>
</tr>
<tr>
<td><em>Xanthogramma pedissequum</em></td>
<td>6</td>
<td>55</td>
<td>M</td>
<td>0.225</td>
<td>0.244</td>
<td>0.0251</td>
</tr>
<tr>
<td><em>Chrysotoxum bicinctum</em></td>
<td>7</td>
<td>32</td>
<td>M</td>
<td>0.490</td>
<td>—</td>
<td>0.0399</td>
</tr>
<tr>
<td><em>Vespula zonaria</em></td>
<td>8</td>
<td>26</td>
<td>M</td>
<td>—</td>
<td>0.542</td>
<td>0.0484</td>
</tr>
<tr>
<td><em>Scaeva pyrastri</em></td>
<td>9</td>
<td>32</td>
<td>M</td>
<td>0.202</td>
<td>0.248</td>
<td>0.0314</td>
</tr>
<tr>
<td><em>Ichyrosyrphus glaucius</em></td>
<td>10</td>
<td>24</td>
<td>M</td>
<td>0.456</td>
<td>0.204</td>
<td>0.0428</td>
</tr>
<tr>
<td><em>Tachina fera</em></td>
<td>N</td>
<td>49</td>
<td>M</td>
<td>—</td>
<td>0.652</td>
<td>0.0374</td>
</tr>
<tr>
<td><em>Eriothrix rufulaculata</em></td>
<td>N</td>
<td>27</td>
<td>M</td>
<td>—</td>
<td>0.481</td>
<td>0.0826</td>
</tr>
</tbody>
</table>

shown by one hoverfly species for males (*X. pedissequum*) and two hoverfly species for females (*E. grossulariae* and *S. pyrastri*), following multiple-test Bonferroni corrections (Rice 1989). The non-mimetic flies were also very variable in their colour pattern, especially *E. rufulaculata*. The amount of abdominal colour-pattern variation shown by the hoverflies varied considerably across the species examined. However, there was no tendency for one sex to show more yellow on the abdomen (tergite 2: paired-\( t = 0.36, \) d.f. = 9, not significant (n.s.); tergite 3: paired-\( t = 0.19, \) d.f. = 9, n.s.) or to be more variable about the mean (paired-\( t = 1.45, \) d.f. = 9, n.s.) than the other.

Figure 2 shows plots of standard deviation against resemblance rank (see table 1) for males (figure 2a) and females (figure 2b). The highest levels of variation in the mimetic insects (i.e. M and ranks 1–10 in figure 2) were found amongst the Müllerian mimics and the hoverfly species with the higher degrees of perfect mimicry of the social wasp colour pattern. For the mimetic insects, there was a negative correlation between rank and standard deviation for both males (\( r = -0.461, \) n.s.) and females (\( r = -0.353, \) n.s.), although neither value was significant at the 5% level. However, there was a significant correlation between the proportion of yellow pigmentation and standard deviation about the mean for both males and females (\( r = 0.632 \) and \( r = 0.573, \) \( p < 0.05 \) in both cases) and a weak correlation between the proportion of yellow pigmentation and rank (males: \( r = -0.38, \) females: \( r = -0.396, \) n.s. in both cases). The partial correlations between rank and standard deviation, removing the confounding effect of the proportion of yellow pigmentation, were \( r = -0.308 \) and \( r = -0.544 \) for males and females, respectively. Neither value was significant at the 5% level, but the female partial correlation was significant at \( p = 0.06. \)

4. DISCUSSION

It has been suggested that assemblages of sympatric Müllerian mimics should evolve to resemble each other very closely (Ackery & Vane-Wright 1984). This form of mimicry relies on reinforcement, so that all members of the Müllerian mimicry ring would benefit from the sampling of just one individual, as the predator would then avoid all similar colour patterns. Hoverflies, to the human eye, show a range of perfection of mimicry (Edmunds
Mimetic imperfection and variation in colour pattern

G. Holloway and F. Gilbert

0.12
0.08
0.04
0

(a)

(b)

rank degree of mimetic resemblance

Figure 2. Relationship between variation (standard deviation) in abdominal colour pattern (proportion of pale pigmentation) and degree of resemblance to three model Vespula wasp species (M) in various mimetic hoverfly species for (a) male and (b) female hoverflies. The rank of degree of resemblance (1–10) follows Dittrich et al. (1993). For comparison, the variation in colour pattern in two non-mimetic flies (N) is included.

2000; Howarth et al. 2000) and Dittrich et al. (1993) further demonstrated this to be the case, using birds as predators and hoverfly mimics of social wasp colour patterns.

In the present study, four predictions were made concerning the relationship between the level of phenotypic variation in the mimetic colour pattern and the degree of mimetic perfection. Nonacs (1985) argued that selection pressure is highest on the prey type that suffers the highest mortality through predation and that high selection will increase resemblance of the mimic to the model. Directional selection such as this will erode genetic variation (Falconer 1989); consequently, those species showing the best mimetic resemblance should show the lowest variation about the mean (although there is not always a tight correlation between genetic and phenotypic variation (Maynard Smith 1989; Forbes et al. 1995)). Hence, the first prediction proposed that there should be very little variation about the mean in colour pattern in Müllerian mimics and that if Batesian mimics were all evolving towards perfection (Edmunds 2000), the more perfect the mimetic relationship, the lower the level of phenotypic variation (figure 1a). However, the level of phenotypic variation measured in the social wasps was high, much higher than most of the supposed Batesian mimics (figure 2). Furthermore, there was no evidence that as the degree of mimetic imperfection increased so did the level of phenotypic variation in colour pattern (although the non-mimetic flies did display high levels of variation). The data provide no support for the first prediction.

The second prediction stated that all the mimic insects should show similarly low levels of phenotypic variation because, through a process of normalizing selection, all had become as good a mimic as possible (figure 1b). Variation in mimetic perfection occurred because some species did not have the appropriate variation on which natural selection could act to induce improvements. Again, the data provided no support for the second prediction because there were clearly some species that displayed much higher levels of variation than others (figure 2).

The third prediction argued that very common aphidophagous hoverflies should show high levels of variation due to relaxed selection for the reasons stated by Azmeh et al. (1998). In the present study, two such species were included: S. ribesii and S. pyrastri (numbers 1 and 9 in figure 2). Both species displayed low levels of phenotypic variation for colour pattern.

The fourth prediction stated that very noxious Müllerian mimics, such as the social wasps examined here, might display high levels of variation in colour pattern. The rationale explained by Duncan & Sheppard (1965) was that the sting of a social wasp forms such a deterrent that predators would avoid anything resembling a social wasp, hence normalizing selection on the colour pattern would be relaxed, allowing more variation. It follows that the most perfect Batesian mimics might be subject to similar relaxed selection and consequently also display high levels of variation (see figure 1d). Figure 2 shows that the wasps were indeed highly variable for colour pattern, as was one of the better wasp mimics (T. vespiforme), although S. ribesii, also perceived to be a good mimic by pigeons (Dittrich et al. 1993), was not very variable. Most of the remaining Batesian mimics displayed lower levels of variation than the wasps and T. vespiforme, which might suggest that they had evolved to resemble a wasp colour pattern as closely as possible and the lack of genetic variation has prevented further progress. The fact that the non-mimetic flies were also highly variable provides further evidence that when a process of mimetic association is embarked on but the degree of resemblance remains poor, natural selection acts to reduce variation in colour pattern. In conclusion, the data appear to provide some support for the fourth prediction presented here.

The function of the striking yellow and black colorations of hoverflies and wasps has largely been attributed to mimetic relationships. There can be little doubt that mimicry has, indeed, been a powerful force shaping the evolution of these colour patterns. Mimicry has been a popular topic for research over the last 20 or more years and a number of models of the evolution of mimicry have dealt with the level of variation and the degree of perfection that might be expected in populations of mimic insects. The general conclusion has been that normalizing selection should be a significant force, resulting in low levels of phenotypic variation remaining. It is surprising therefore that, as far as we know, no other studies have addressed the issue of levels of variation in mimetic systems in detail. The results of the current study suggest that more empirical work needs to be performed, which
might refine our view of how mimetic relationships develop.

The image analysis equipment was purchased on NERC grant no. GR3/9463. We are grateful to the British Natural History Museum for allowing us to use some of their preserved specimens for the study.

REFERENCES


