Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c5096034.

THE ROYAL SOCIETY
PUBLISHING

The evolution of startle displays: a case study in praying mantises

Marta Vidal-García[1,2], James C. O'Hanlon[3], Gavin J. Svenson[4] and Kate D. L. Umbers[5,6]

[1]Ecology and Evolution, Research School of Biology, Australian National University, Canberra, ACT 0200, Australia
[2]Department of Cell Biology and Anatomy, University of Calgary, Calgary, Alberta, Canada
[3]School of Environmental and Rural Science, University of New England, Armidale, New South Wales, Australia
[4]Department of Invertebrate Zoology, Cleveland Museum of Natural History, Cleveland, OH, USA
[5]School of Science, and [6]Hawkesbury Institute for the Environment, Western Sydney University, Penrith, NSW 2751, Australia

MV-G, 0000-0001-7617-7329; JCO, 0000-0002-7382-5543; GJS, 0000-0003-0207-6458; KDLU, 0000-0002-9375-4527

Anti-predator defences are typically regarded as relatively static signals that conceal prey or advertise their unprofitability. However, startle displays are complex performances that deter or confuse predators and can include a spectacular array of movements, colours and sounds. Yet, we do not fully understand the mechanisms by which they function, their evolutionary correlates, or the conditions under which they are performed and evolve. Here, we present, to our knowledge, the first phylogenetically controlled comparative analyses of startle displays including behavioural data, using praying mantises as a model system. We included 58 species that provide a good representation of mantis diversity and estimated the strength of phylogenetic signal in the presence and complexity of displays. We also tested hypotheses on potential evolutionary correlates, including primary defences and body size. We found that startle displays and morphological traits were phylogenetically conserved, whereas behavioural traits were highly labile. Surprisingly, body size was not correlated with display presence or complexity in phylogenetically controlled analyses. Species-rich clades were more likely to exhibit displays, suggesting that startle displays were probably involved in lineage diversification. We suggest that to further elucidate the conditions under which startle displays evolve, future work should include quantitative descriptions of multiple display components, habitat type, and predator communities. Understanding the evolution of startle displays is critical to our overall understanding of the theory behind predator–prey dynamics.

1. Background

Perhaps because they are a matter of life or death, anti-predator defences have long captured human attention and emergent research has provided deep insight into the underlying evolutionary processes [1]. Concealment through camouflage strategies, such as crypsis and masquerade, help prey to avoid detection and identification by predators, whereas aposematism—advertisement of unprofitability such as toxicity or weaponry—typically provides protection by predator learned avoidance [2]. Camouflage and aposematic strategies have received the bulk of research attention and they and their evolution is usually investigated independently, viewed as alternative anti-predator strategies with independent evolutionary histories.

Despite being considered in isolation, there are at least two clear ways in which camouflage and aposematism are combined in anti-predator defence: distance dependence and startle displays [3–6]. Distance-dependent colour patterns are those known to be cryptic to predators from far away and aposematic at close range, suggesting that the same colour pattern can be camouflaging and aposematic depending on predator viewing distance [3,5,6]. For example, in dyeing poison frogs (*Dendrobates tinctorius*), their colour pattern is salient to

predators at close range, but pattern elements are not discernible at a greater distance which causes a camouflaging effect [5]. Empirical data on such dual-strategy colour patterns are rare and so macroevolutionary patterns in distance-dependent changes in anti-predatory function are as yet difficult to identify. On the other hand, many species are known to defend against predators using startle displays, complex performances in which prey may swap between camouflage and aposematism among others [4,7]. These displays are widespread across multiple clades, and taxonomic diversity of species exhibiting startle displays spans at least three animal Phyla in 17 Orders (K. D. L. Umbers 2020, personal observation). Startle displays can incorporate conspicuous colours, light, sounds, postures and/or compounds when they perform their defence [8] and among species, even closely related ones, startle displays are often highly variable. For example, in *Catocala* moths, sympatric species' forewing crypsis is convergent whereas hindwing colour patterns are divergent and anomalous with regard to predator's previous encounters [9–11]. The great variability in startle displays stands in stark contrast with predictions of stabilizing selection in camouflage and aposematism and suggests the potential for different evolutionary drivers [12].

The charismatic nature of startle displays is at odds with our poor understanding of the mechanisms by which they function, their evolutionary correlates, and the conditions under which they are performed and evolve [4,13]. Phylogenetic comparative methods provide powerful opportunities to develop and test hypotheses on the evolution of complex traits like startle displays and their components [14]. Studies on size, shape and colour traits are possible through the assessment of large numbers of museum specimens, including fossils, where phylogenetic hypotheses are available [15]. However, compared to our understanding of body size, shape and colour, the evolutionary history of behaviour is often obscured. Descriptive behavioural data are typically relatively time-consuming to collect and difficult to classify [16]. In addition, behavioural traits are ephemeral. They can be difficult to observe or induce, are only exhibited temporarily, and only by live animals, meaning that presence and especially absence data can be especially difficult to obtain. However, behavioural traits cannot be discounted, because they can reveal the function and have strong fitness consequences. They can promote or constrain evolutionary change by, for example, compensating for limitations of morphological traits (e.g. genital size) [17] and/or physiological traits (e.g. thermoregulation) [18], or by reducing the availability of traits to selection (e.g. hidden colour patterns) [4,19,20]. It is thus unsurprisingly impossible to fully understand the evolution of traits that include behavioural components without behavioural data.

Startle displays lend themselves to comparative behavioural analyses because they are complex, discrete, multicomponent performances that can be highly diverse among taxa. Evolutionary patterns of startle displays have been recently investigated on the sounds and hidden colours of several insect clades and have found different patterns when accounting for phylogeny [21–23]. Such comparative studies are very valuable for generating and testing hypotheses on the evolution of colour patterns but unfortunately it remains difficult to draw conclusions about the complete startle display strategy without behavioural data. A few studies have aimed to provide evidence on the feasibility of different hypotheses on the evolution of startle displays [24]. Some of these prior studies

have tested the evolutionary sequence of the components of startle displays such as the Startle-First hypothesis [4,8,24,25]. The Startle-First hypothesis suggests that the movement component of a display can deter predators on its own, and the subsequent acquisition of a colour pattern component could enhance its effect. Studies such as Holmes *et al.* [24] provide critical empirical behavioural evidence to support tests of the Startle-First hypothesis at a broader phylogenetic scale. However, this hypothesis only comments on the order in which components of the displays may have arisen over evolutionary time, and theory on how startle displays evolve is in its infancy.

The praying mantises are a cosmopolitan, monophyletic group comprising nearly 2500 species in over 425 genera [26]. Mantises are famous for performing visually impressive startle displays, some of which include sound production. Many mantis species' displays, and lack thereof, have been described in the literature, making them an ideal system with which to examine the evolution of startle displays. We gathered descriptions of mantis displays from both the literature and experts in mantis behaviour and conducted phylogenetic analyses across 58 different praying mantis genera (one species per genus, approximately 13% of total diversity at the genus level). Our main goal was to quantify the macroevolutionary patterns of startle displays in mantises. Specifically, we addressed the following four questions using our behavioural and morphological dataset in a phylogenetic comparative framework:

First, is there a correlation between the presence of startle displays and the species' type of primary defence? Our hypothesis was that the type of primary defence is correlated with the presence of startle displays. We predicted that display presence and complexity would be more likely to be associated with 'special resemblance' (putative masquerade [27]; *sensu* Skelhorn [8,21]) than 'general resemblance' (putative crypsis) [28,29]. Our rationale was that 'special resemblance' and complex startle displays both require a lot of morphological changes and that both may evolve in response to very strong predation pressure.

Second, is startle display presence or complexity correlated with body size or other morphological traits (e.g. relative wing length)? Our hypothesis was that the presence and/or complexity of startle displays is positively correlated with overall body size. There are three rationales for this hypothesis: (i) larger species may be a more profitable meal, (ii) larger species may be poorer at escaping and therefore require a more potent defence, and (iii) larger startle displays may be more effective deterrents. Similarly, with regard to wing length, we predicted a negative correlation between display complexity and forewing length relative to body length. For example, a species with very short wings may benefit from a more complex startle display because it cannot fly to escape. Because female mantises are typically larger than males, we also asked whether there are differences in displays between the sexes.

Third, is there a difference between the macroevolutionary patterns of behavioural and colour display components as predicted by the Startle-First hypothesis? We assessed the fit of the Startle-First hypothesis in explaining the evolutionary trajectory of startle displays in the praying mantises. Following Blomberg *et al.*'s [14] findings that behavioural traits generally exhibit less phylogenetic signal than other traits, our hypothesis was that strength of the phylogenetic signal would differ between behavioural and colour pattern components. We predicted that if the Startle-First hypothesis is true, (i) behaviour

and colour traits would be evolutionarily decoupled in some species, particularly in less-recently derived species, and (ii) that behavioural components would show greater lability than colour patterns as is the case in many different taxa, such as in the exaggerated displays of bowerbirds [30].

Fourth, is the presence and complexity of startle displays correlated with species richness? Our hypothesis was that startle display presence and complexity is not randomly distributed among taxa with respect to species richness of a clade, and that speciose clades would be more likely to diversify. We predicted that species in species-rich clades have more complex displays than those in more isolated clades as species 'innovate' and 'elaborate' (*sensu* Endler [22]) to occupy different ecological niches [22]. Ecological character displacement resulting from interspecific competition [23] and ecological and phenotypic differentiation has been demonstrated in several groups [22,31,32]. Furthermore, increases in the complexity of startle displays may also increase survival through the anomaly, differences in the appearance of prey that share a predator population may decrease the probability that predators can learn to expect any one type of display [10,11]. The praying mantises provide a unique opportunity to learn about the macroevolutionary patterns of both the behavioural and morphological components of startle displays.

2. Methods

(a) Startle display and primary defence data
To generate our dataset, we searched Web of Science and Google Scholar for all published descriptions of praying mantis startle displays. In total, we gathered behavioural descriptions for 58 praying mantis species, with each species representing a different genus across the Mantodea, representing approximately 13% of mantis diversity at the genera level. Please see the electronic supplementary material Methods and Results for further details.

(b) Body size and shape data
To compile the dataset of mantis body size, we took size data from seven publications and, in addition, directly measured 294 specimens from 49 species kept in the collections of the Cleveland Museum of Natural History and the National Museum of Natural History (Smithsonian Institution). We provide further details in the electronic supplementary material Methods and Results.

(c) Mantis phylogeny and timetree estimation
A total of 93 praying mantis taxa were selected from previous studies [26], and one newly added species was included in our phylogenetic tree estimation (electronic supplementary material, table S1). We estimated and generated: (i) a partitioned maximum likelihood (ML) tree, (ii) a mixed model Bayesian tree, and (iii) a time-calibrated tree (timetree) using BEAST v1.8.3 [33]. Details on tree estimation and taxa selection are provided in the electronic supplementary material Methods and Results.

(d) Phylogenetic comparative analyses
We used both the consensus timetree and the posterior distribution of 10 000 randomly sampled trees to address our questions on macroevolutionary patterns. We pruned the Mantodea timetree to match the available data for each of the analyses to genus level (electronic supplementary material, figure S2), using the R package *ape* [34]. We then estimated their evolutionary distinctiveness using the equal splits metrics [35] with the *evol.distinct* function in *picante* [36]. Evolutionary distinctiveness

can be described as a measure of the genetic isolation of certain taxa in a phylogenetic tree [35].

(i) Phylogenetic signal and ancestral state reconstruction
We estimated the degree of phylogenetic signal in behavioural traits, colour pattern displays and body size and shape variables, to determine whether each of these traits is phylogenetically conserved or it shows a certain degree of evolutionary lability. To estimate the degree of phylogenetic signal in startle display and its component, as well as the type of primary defence (figure 1), we fitted Pagel's lambda models [37] using the function *fitDiscrete* in *geiger* [38]. We expected that species in which these categorical traits are phylogenetically conserved would have lambda values closer to 1 (i.e. have evolved under a Brownian motion (BM) model of phenotypic evolution), whereas species with phylogenetically labile traits would have values closer to 0. To determine whether there had been more or fewer evolutionary transitions for each display component than expected by chance, we also compared the number of evolutionary transitions for each display component against a randomization process of that character across the phylogeny [39]. We also estimated the strength of the phylogenetic signal on the presence of display and complexity of display across the posterior distribution of 10 000 randomly sampled trees, to account for phylogenetic uncertainty in our data. For body size measures, i.e. continuous morphological variables, in our dataset, we estimated the degree of the phylogenetic signal using Blomberg's K-statistic with *phylosig* in *phytools* [40] and incorporating sampling error as per [41], for both males and females. This allowed us to know whether behavioural display traits, colour pattern displays, and body size and shape exhibit different macroevolutionary patterns, and determine whether display traits are more labile than body size and shape traits.

Finally, we estimated ancestral states in the complexity of display using a continuous-time Markov chain model (Mk model) with the R package *ape* [34], to estimate the distribution of display state changes through time. We used three different models to reconstruct the ancestral states of this trait: (i) a one-parameter equal rates model (ER), (ii) a symmetric model of reconstruction (SYM), and (iii) an all-rates different matrix model (ARD). This allowed us to estimate the most likely ancestral state of deimatic displays, as well as the most likely point in time when deimatic displays appeared in mantises.

(ii) Trait correlation
We tested for evidence of correlated evolutionary changes between the presence of startle display and type of primary defence using the function *fitPagel* [42] in *phytools* [40]. This statistical method estimates the change rates for binary traits across a phylogenetic tree and tests whether the evolutionary patterns of two discrete traits are correlated. We tested for correlated evolutionary patterns across all of the combinations of binary display traits. We also performed phylogenetic ANOVAs (phyloANOVAs) between morphological traits and startle displays, to discern whether specific startle displays are correlated with specific body shape patterns. All phyloANOVAs were performed, as per [43] using the *phylANOVA* function in *phytools* [40]. Finally, we also compared the probability of a species having a display using two separate generalized linear models, one for males and one for females, without phylogenetic corrections. We performed that with display presence as a binary response variable to average body length as the fixed effect using the 'glm' function in the R package *lme4*.

(iii) Phylogenetic distinctiveness
We estimated phylogenetic distinctiveness using the equal splits metrics [35] with the *evol.distinct* function in *picante* [36]. We performed phyloANOVAs between each display trait
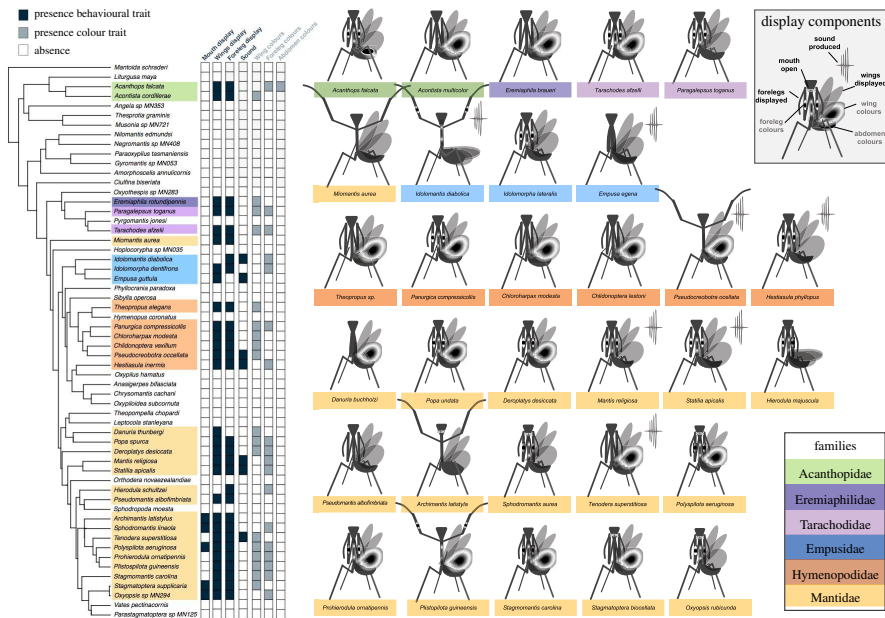
Figure 1. Timetree of mantis species with available behavioural data. The presence and absence of different behavioural components are noted in black and grey, respectively. Colour shading on tree tip labels indicate the Mantodea family to which they belong. (Online version in colour.)

(both behavioural and colour pattern displays) and evolutionary distinctiveness. We also fitted phylogenetic generalized least-squares (PGLS) models to test for a correlation between phylogenetic distinctiveness and continuous traits (including morphological traits, i.e. body size and shape), using the *gls* function in *nlme* [44]. We used two different models of phenotypic evolution to estimate which correlation structure fitted the data best: Brownian Motion (BM) and Ornstein–Uhlenbeck (OU), using the functions *corBrownian* and *corMartins* from the *ape* package [34], respectively. We estimated the expected covariance under a BM model with the parameter $\gamma = 1$, and for the OU correlation matrix, we used a fixed parameter $\alpha = 7.1$ (best fit). We calculated McFadden's adjusted R-squared (McF R^2_{adj}) by dividing the log-likelihood of each model (minus the number of estimated parameters, which was $k = 1$ for all of our PGLS models) by the log-likelihood of a null model without the predictor (phylogenetic distinctiveness in all the other PGLS models), and then subtracting this value from 1 [45]. We used Akaike information criterion (AIC; [46]) to select either a BM or OU model, with lower AIC values depicting a better fit (we did not use AICc because $n < 40$). All morphological variables were log-transformed for each PGLS analysis, with the exception of the sexual dimorphism ratios on morphological variables. These tests allowed us to determine whether the presence and the complexity of startle displays, as well as body size and shape traits, were correlated with phylogenetic distinctiveness to determine whether speciose taxa are more likely to exhibit deimatic displays.

3. Results

(a) Startle display and primary defence data
We found reliable descriptions of the defensive behaviour and morphology of 58 species of mantis in 58 genera,

approximately 13% of extant genera. Of the available descriptions, 31 species were reported to perform a startle display when provoked and 27 to not perform such a display despite the same level of provocation. Please see the electronic supplementary material, Methods and Results for further details.

(b) Body size data
We illustrate mantis phenotypic variation in both males (m; $n = 165$) and females (f; $n = 129$) for body length ($m = 14$–118 mm; $f = 11$–127 mm), pronotum length ($m = 2$–47 mm; $f = 2$–52 mm) and forewing length ($m = 10$–62 mm; $f = 4$–68 mm) in the electronic supplementary material, figure S1.

(c) Mantis phylogeny and time tree estimation
Both the partitioned ML and the Bayesian analyses recovered a topology (likelihood score: -150062.114978) with high bootstrap support (BS) values (≥ 80) across most terminal level nodes and low BS values (less than 80) across most backbone nodes (electronic supplementary material, figure S2), which is consistent with prior studies [26,47,48]. Further details are provided in the electronic supplementary material Methods and Materials.

(d) Phylogenetic comparative analyses
(i) Phylogenetic signal and ancestral state reconstruction
Tests of phylogenetic signal in the presence of a startle display recovered a lambda value of 1, indicating that closely related taxa are more likely to have a display than distantly related taxa. The complexity of the display also showed moderate

phylogenetic signal ($l = 0.571$). However, when accounting for phylogenetic uncertainty by testing the strength of phylogenetic signal across of 10 000 trees, the phylogenetic signal was moderately strong for the complexity of display ($\lambda = 0.780$; electronic supplementary material, figure S3), and very strong for the presence of display ($\lambda = 0.996$; electronic supplementary material, figure S3). Both display presence and complexity also have fewer evolutionary transitions than would be expected by chance, again supporting the hypothesis that displays are phylogenetically conserved (electronic supplementary material, table S2). In terms of the seven display components, three of the four behavioural traits—wings display, arms display and sound—had a very low phylogenetic signal ($l = 0$), indicating that these traits are labile (electronic supplementary material, table S2). On the other hand, the colour patterns—wing colours, arm colours, abdomen colours—and mouth display showed moderately strong phylogenetic signal ($l > 0.805$). There were no significant differences in evolutionary transition in comparison to a randomization process for any of the individual display components, even though the number of transitions was lower than expected by chance in some cases (electronic supplementary material, table S2). Mantis primary defence (putative crypsis or putative masquerade) showed strong phylogenetic signal ($l = 1$), but no significant differences in the number of transitions compared to that expected by chance (electronic supplementary material, table S2). We found a strong phylogenetic signal for the body size variables we measured, but a low phylogenetic signal in all the dimorphism ratios (electronic supplementary material, table S3).

The ancestral state reconstruction showed phylogenetic clustering of presence and complexity of display, with the ancestral state being no display, and the different displays appearing approximately 60 Ma (electronic supplementary material, figure S4). The model of highest likelihood being the all-rates different matrix model (logLik$_{ER}$ = -81.60; logLik$_{SYM}$ = -72.04; logLik$_{ARD}$ = -65.00; electronic supplementary material, figure S4).

(ii) Trait correlation

We did not find any evidence that the evolutionary patterns of display traits are correlated (electronic supplementary material, table S4). We found significant results between display traits and some body shape traits for some of the phylo-ANOVAs performed, mostly related to forewing length (forewing length dimorphism and sound, female forewing length and wings display, male forewing length and wings display, male forewing length and presence of display, male forewing length and complexity of display (electronic supplementary material, table S5). Independent of phylogenetic history, the probability of display presence was higher for larger animals in both males and females, but the effect sizes were very small (0.03 and 0.02, respectively; electronic supplementary material, table S7 and figure S5).

(iii) Phylogenetic distinctiveness

phyloANOVAs performed on evolutionary distinctiveness were significant for the presence of display and display complexity ($F_{1,57} = 14.527$, $p = 0.003$ and $F_{3,57} = 4.969$, $p = 0.013$, respectively; figure 2). For the PGLS analyses, we found no strong relationship between any morphological variable and evolutionary distinctiveness (electronic supplementary

material, table S6), for either models following BM or OU. Also, we found no clear pattern in the model of phenotypic evolution of this correlation, as there was a mix of models being best fitted by BM, while others were better fitted by an OU covariance structure.

4. Discussion

Startle displays are spectacular performances to deter or confuse predators, which can greatly increase biological fitness. This study provides, to our knowledge, the first comprehensive picture of the evolution of startle displays and their components in a phylogenetic framework. We analysed the available descriptions of 58 praying mantis species' startle displays to address four questions on their evolution. Our data provided the following insights, each of which we discuss in further detail, and in instances where our hypotheses were not supported we briefly speculate on explanations to guide future hypothesis testing. First, neither display presence nor complexity was correlated with the type of primary defence. Second, we found a weak effect of body size on the presence of display independent of phylogeny, but when accounting for phylogeny, we did not find evidence that body size traits were correlated with the presence or complexity of startle displays or their components. Third, we found a moderate phylogenetic signal in display presence, display complexity and the presence of hidden colour patterns. We also found a strong phylogenetic signal in body size measures, but a weak signal in our derived measures of sexual dimorphism. By contrast, we found a weak phylogenetic signal in three of the four behavioural display components—the use of wings, the use of forelegs and the production of sound. Fourth, species displays in species-rich clades were more complex seemingly through lability in behavioural rather than colour pattern traits. Finally, our data do not provide a strong test of the Startle-First hypothesis because all but four species' displays include both behavioural and colour pattern components.

(a) Is there a correlation between the presence of startle displays and the species' type of primary defence?

Our data do not support our hypothesis that startle displays are positively associated with 'special resemblance' (putative masquerade) [27]. This could be owing to ecological factors such as variability in predation pressure, the defences of sympatric prey species, other advantages in resembling environmental objects (such as floral simulation), or because our data are subjective from a human perspective and binary, thus reducing confidence in our conclusions. Secondary defences like startle displays are used when primary defences fail [4,19,49] so our results could mean that species with greater complexity in their displays are so because they are more often identified as prey and attacked [50]. To test this hypothesis, data on the predator–prey dynamics of many species is required, specifically the relative vulnerability of different morphologies to predation—a formidable task at this phylogenetic scale.

The decoupling of primary and secondary defence in mantises could be driven by species living in sympatry, though unfortunately the data to test this hypothesis is not available
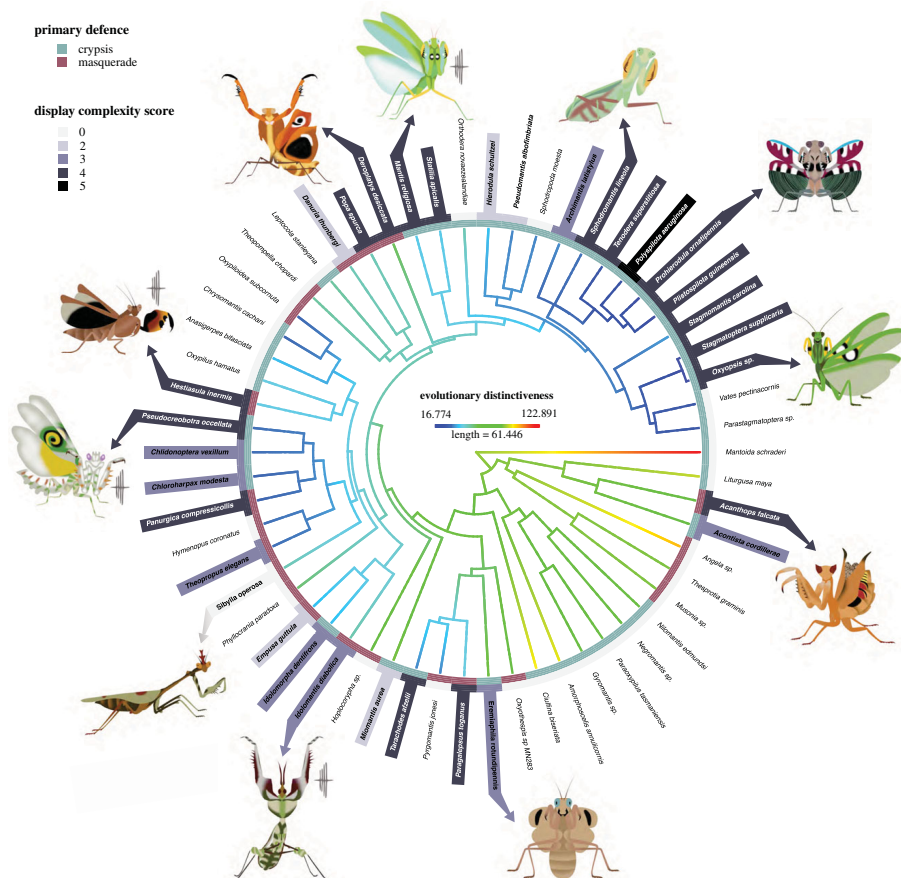
Figure 2. Timetree of mantises with available behavioural data depicting evolutionary distinctiveness and primary defence type (crypsis or masquerade) and display complexity score for each species. (Online version in colour.)

for mantises. The primary defence of *Catocala* moths cryptic forewing patterning is convergent among sympatric species and the colour patterns on their hindwings—used in startle—are divergent [9]. Convergent crypsis suggests that for *Catocala* there is an optimal pattern for avoiding predator detection, and their divergent startle colours have been interpreted as supporting an anomaly hypothesis whereby because of their forewing convergence, predators cannot easily predict the hindwing colour and are startled [10,11]. As museum records are digitized, the addition of geographical information about mantis species distributions and thus the species with which they are sympatric, and their predator suite may allow comparisons with their defensive strategies and would be of great interest.

(b) Are startle displays correlated with body size and shape?

We predicted that displays would be present and complex in larger species because they may have a higher likelihood of detection and predation risk, be more worthwhile prey, and/or be poorer at escaping and thus receive a greater benefit from startle displays [20]. However, when controlling for phylogeny, we did not find evidence that body size and shape traits are correlated with display presence or complexity or components, though we found a small effect of body size on display presence independent of phylogeny. Previous studies on insects have found that the presence of hidden colours is positively correlated with body size in only Orthoptera and Phasmatidae when phylogeny was controlled but not in Mantidae, Saturniidae (Lepidoptera) or Sphigidae (Lepidoptera), which matches our results for mantises [20,51]. One possible explanation for a decoupling of body size and startle displays in mantises is that they have different evolutionary drivers. While predators are likely to be the main drivers of the evolution of startle displays, body size may be driven by many other factors including fecundity [52], dispersal and the mantis' own hunting strategy [53]. For example, in the orchid mantises

(Hymenopodinae: Hymenopodini), large female body size preceded the change in their morphology from mantis-like to flower-like seemingly to gain access to large pollinating insects as prey [53]. Conspecific males, on the other hand, remained ancestrally small and cryptic, affording them the ability to disperse and locate females.

(c) Do evolutionary patterns differ between behavioural and colour pattern display components?

We observed that colour traits are always present with behavioural traits in all but four species, and colour traits were much more phylogenetically conserved than behavioural traits (figure 1). The common appearance of behaviours and colours together could be caused by strong covariation between these traits, perhaps indicating an association between particular genes and multiple independent display traits [54]. Pleiotropic mutations could increase the potential for evolutionary diversification of these coupled traits [55], which would also explain the patterns of display complexity we observed in species-rich clades. However, the genetic underpinnings of behavioural traits are extremely complex, as they are usually regulated by complex genes, can be affected by multiple epistatic effects and are sensitive to environmental variation, making it difficult to discern trait effects [56,57].

Conversely, functional integration could be driving these patterns of concerted evolution between specific behavioural and colour components in startle displays [58]. Ligon *et al.* [58] report functional integration of visual and acoustic signals in the birds-of-paradise and suggest that the extreme diversification of 'courtship phenotype' is partly a consequence of the multicomponent and multimodal nature of their displays. An interesting further question for future studies is whether complex startle displays could have given rise to mantis radiations (see §4d immediately below).

Holmes *et al.* [24] showed that movement alone can deter predator attack in the absence of conspicuous coloration, depending on the speed of that movement [24], but that the addition of a conspicuous colour enhanced the display's anti-predator effect. The four species in our dataset that perform behaviour-only displays may be trading off between a display that is still effective at deterring predators but avoids the metabolic or conspicuousness costs of coloration. A wider sampling of the mantis phylogeny may reveal more behaviour-only displays—it is possible that the published descriptions of displays are skewed to the more spectacular species—and enable comparative analyses of the causation and survival value of displays with and without colour patches and the validity of the Startle-First hypothesis.

Our results do not lend clear support to the Startle-First hypothesis, nor do they convincingly negate its applicability to praying mantises. While we see some decoupling of behaviour and colour patterns, the four species that exhibit behaviour-only displays are scattered across the phylogeny and not concentrated in relatively less-derived clades (figures 1 and 2). Wider sampling is needed to test whether basal linages are more likely to perform behaviour-only displays, a lofty, but worthwhile goal. In addition, quantitative descriptions of displays such as colour pattern conspicuousness, size of the colour patch, speed and vigorousness of behavioural performance, are needed for more accurate measurement of display complexity. Of course, we cannot discount that the Startle-First hypothesis may simply be a poor explanation for the evolution of startle displays in mantises (at least). Quantifying the use of mantis raptorial forelimbs in defence and their variability across the Order may allow the Defence-First hypothesis to be tested in future studies. We were not able to test the Defence-First hypothesis here because we do not have data on the variation in mantis defences—their raptorial forelimbs. Specifically, we do not have evidence regarding the use in defence because their primary function is likely prey-capture [4].

Given that the vast majority of the descriptions in our dataset are from behaviour examined when triggered by humans or other ecological irrelevant stimuli, it is possible that some species deemed not to display, actually do display under real predatory attack. To confirm species true display behaviour, extensive field observation or systematic experiments are required. Similarly, the colour patterns described may have evolved for a different or dual functions, for example, flash coloration [59], sexual display [60] or predation [53]. Thus, much more behavioural data are required to fully understand the evolution of mantis colour patterns and their associated behavioural traits.

(d) Are startle displays correlated with species richness?

We found that species in species-rich clades were more likely to have a display and had more complex startle displays, whereas evolutionarily isolated (or phylogenetically distinct) species generally lacked or had less-complex displays. This is not surprising, as the display is probably a player in lineage survival and diversification. We speculate that closely related species, which typically share ecological niches and have similar morphologies can occupy different parts of the niche by adapting their behaviour, as opposed to a modification of body size, shape and colour [14]. The greater degree of evolutionary lability in behavioural traits adds support to this interpretation. Species displays should be divergent, especially if they are sympatric, as first suggested to explain the variation in hindwing colour patterns in the moth genus *Catocala* [9–11]. If phylogenetically proximal mantis species inhabit similar geographical areas and share predator populations, we predict that species' displays would diverge to reduce the possibilities for predators to expect their displays [10]. Alternatively, we may see complex startle displays in speciose clades because phylogenetic distinctiveness is not a good proxy for functional trait originality [61]. That is, if mantis displays are effective at increasing survival value, the presence of a display may promote speciation events and lower extinction rates, leading to higher rates of diversification in mantises with a higher display complexity.

Our ancestral state reconstruction of mantis display complexity showed high levels of phylogenetic clustering, and the timetree allowed us to estimate that defensive displays approximately appeared 60 Ma. The extinction of the dinosaurs and subsequent radiation of birds began with the end of the Cretaceous around 66 Ma which places the evolution of startle displays at a roughly similar time to the radiation of their likely predators. Alternatively, or in addition, this macroevolutionary pattern could have been driven by cold climatic periods in the Eocene and Oligocene [62]. Clavel & Morlon [62] show that past climate has driven

8

royalsocietypublishing.org/journal/rspb   *Proc. R. Soc. B* **287**: 20201016

body size evolution in birds and mammals as evolutionary rates are substantially higher during periods of cold in the Cenozoic, probably owing to changes in selective pressures [62]. Even though the effect of past climate on macroevolutionary patterns might be lower in insects than in endotherm vertebrates [63], temperature changes on endotherms could have affected insect defences and thus their morphological and behavioural traits. We urge caution in drawing conclusions around time estimates in the current dataset owing to the limited sampling and low deep-level support values (typical of many invertebrate phylogenies) that could have dramatic effects on both the topology and branch lengths across our timetree, thus affecting the ancestral state estimations. While these support value issues have persisted in Sangar-based analyses, newer genomic studies are supportive, in many cases, of both the topologies and estimated divergence dates derived from multi-loci datasets [64]. However, the timing of the appearance of startle displays in mantises in the era that gave rise to bird radiation and large body size in some animal taxa is a hypothesis worth pursuing in future studies, especially on species with well-known habitat usage and complete taxonomic coverage across well-resolved phylogenetic trees.

(e) Towards a better understanding of the evolution of deimatic displays

Our study is, to our knowledge, the first to test for evolutionary processes that might be driving both behavioural and morphological traits in startle displays. This study provided an opportunity to combine a large set of available behavioural descriptions with a robust phylogeny to further our understanding of startle displays in an evolutionary context. One caveat is that data were taken from the works of different researchers and from different papers, potentially giving rise to biases in the way different people measured the displays (but see the electronic supplementary material). For example, stimuli used to entice mantises to perform displays vary and but are often not ecologically relevant, e.g. physical stimuli applied by human experimenters (the 'science poke' [13,65]) or looming objects [66]. This leaves us with the question as to whether or not the descriptions truly reflect how and in response to what stimuli displays are used in the wild. Edmunds [60] pondered a similar problem when he noticed that although some species, such as two African mantises, *Hoplocorypha nigerica* and *Paramantis prasina*, had coloration on their forecoxae they could not be enticed into performing a display [60]. Similarly, Edmunds noted several species, such as the Madagascan marbled mantis (*Polyspilota aeruginosa*), the African mantis (*Sphodromantis lineola*) and Gambian spotted-eye flower mantis (*Pseudoharpax virescens*), had dorsal abdominal colours covered by their wings that were apparently not used in defensive displays, but rather were exposed during flight, perhaps functioning as flash coloration [21,60,67,68] or in intraspecific signalling [69].

There is exciting potential for gathering more biologically informed measures of primary defences in mantises (such as special resemblance and general resemblance) to expand our analysis of their association between primary and secondary defences. The addition of currently unavailable fine-scale data on the habitat and life-history characteristics of praying mantises would also provide deeper insight into what is driving the evolution of defensive displays, for example diversity of predator community and mantis hunting behaviour. Also, as always, more complete phylogenetic coverage, in particularly more evenly distributed species sampling across the Mantodea order will reveal whether our study shows the true evolutionary patterns. While this study represents the most comprehensive comparative dataset on praying mantis defensive displays, with more data, there is no doubt more complexity yet to be considered. Whole-body movement and rhythmical movements, for example, are commonly noted characteristics of deimatic displays—in some cases the animals simply orient to the predator (or threat) whereas others are more dynamic, e.g. Devil's flower mantis (*Idolomantis diabolica*) individuals perform repetitive back-and-forth actions and the European mantis (*Mantis religiosa*) stridulate repeatedly using their wings [60]. Additional complexity could also be scored in posturing, for example, many species can give either a 'frontal' or a 'lateral' display depending on the angle of the predator's attack whereas some seem to only give lateral or frontal displays. Also, the position of the forelegs during the display varies from flexed to outstretched [60] (figure 1). Such extreme signal complexity, and interspecies variability presents exciting future directions and highlights the praying mantises as a model group to understand the evolution of startle displays.

The complex associations between different selective pressures and covarying degrees of behavioural and morphological lability demonstrate the difficulties in identifying drivers of startle display evolution. Our results, as have those of many others, highlight the importance of using a phylogenetic framework when asking questions on broad behavioural patterns, but also the complexity associated with integrating behavioural data with phylogenetic comparative methods. Thus, our study sheds light on several new hypotheses that might explain the evolution of startle displays in praying mantises, and that it will serve as a basis for future empirical and manipulative research into this topic.

References

1. Alatalo RV, Mappes J. 1996 Tracking the evolution of warning signals. *Nature* **382**, 708–710. (doi:10.1038/382708a0)

2. Guilford T, Dawkins MS. 1991 Receiver psychology and the evolution of animal signals. *Anim. Behav.* **42**, 1–14. (doi:10.1016/S0003-3472(05)80600-1)

3. Tullberg BS, Merilaita S, Wiklund C. 2005 Aposematism and crypsis combined as a result of distance dependence: functional versatility of the colour pattern in the swallowtail butterfly larva. *Proc. R. Soc. B* **272**, 1315–1321. (doi:10.1098/rspb.2005.3079)

4. Umbers KDL, De Bona S, White TE, Lehtonen J, Mappes J, Endler JA. 2017 Deimatism: a neglected form of antipredator defence. *Biol. Lett.* **13**, 20160936. (doi:10.1098/rsbl.2016.0936)

5. Barnett JB, Michalis C, Scott-Samuel NE, Cuthill IC. 2018 Distance-dependent defensive coloration in the poison frog *Dendrobates tinctorius*, Dendrobatidae. *Proc. Natl Acad. Sci. USA* **115**, 6416–6421. (doi:10.1073/pnas.1800826115)

6. Honma A, Mappes J, Valkonen JK. 2015 Warning coloration can be disruptive: aposematic marginal wing patterning in the wood tiger moth. *Ecol. Evol.* **5**, 4863–4874. (doi:10.1002/ece3.1736)

7. Umbers KDL, Fabricant SA, Gawryszewski FM, Seago AE, Herberstein ME. 2014 Reversible colour change in Arthropoda. *Biol. Rev.* **89**, 820–848. (doi:10.1111/brv.12079)

8. Cott HB. 1940 *Adaptive coloration in animals*. London, UK: Methuen and Co. Ltd.

9. Sargent TD. 1978 On the maintenance of stability in hindwing diversity among moths of the genus *Catocala* (Lepidoptera: Noctuidae). *Evolution* **32**, 424–434. (doi:10.2307/2407609)

10. Schlenoff DH. 1985 The startle responses of blue jays to *Catocala* (Lepidoptera: Noctuidae) prey models. *Anim. Behav.* **33**, 1057–1067. (doi:10.1016/S0003-3472(85)80164-0)

11. Sargent TD. 1990 Startle as an anti-predator mechanism, with special reference to the undenting moths, (*Catocala*). In *Insect defences: adaptive mechanisms and strategies of prey and predators* (eds DL Evans, JO Schmidt), pp. 229–249. Albany, NY: University of New York Press.

12. Leimar O, Enquist M, Sillen-Tullberg B. 1986 Evolutionary stability of aposematic coloration and prey unprofitability: a theoretical analysis. *Am. Nat.* **128**, 469–490. (doi:10.1086/284581)

13. O'Hanlon JC, Rathnayake DN, Barry KL, Umbers KDL. 2018 Post-attack defensive displays in three praying mantis species. *Behav. Ecol. Sociobiol.* **72**, 176. (doi:10.1007/s00265-018-2591-6)

14. Blomberg SP, Garland T, Ives AR. 2003 Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* **57**, 717–745. (doi:10.1111/j.0014-3820.2003.tb00285.x)

15. Vidal-García M, Byrne PG, Roberts JD, Keogh JS. 2014 The role of phylogeny and ecology in shaping morphology in 21 genera and 127 species of Australo-Papuan myobatrachid frogs. *J. Evol. Biol.* **27**, 181–192. (doi:10.1111/jeb.12292)

16. Williams RL, Porter SK, Hart AG, Goodenough AE. 2012 The accuracy of behavioural data collected by visitors in a zoo environment: can visitors collect meaningful data? *Int. J. Zool.* **2012**, 1–13. (doi:10.1155/2012/724835)

17. Ptacek MB, Travis J. 1998 Hierarchical patterns of covariance between morphological and behavioural traits. *Anim. Behav.* **56**, 1044–1048. (doi:10.1006/anbe.1998.0874)

18. Huey RB, Hertz PE, Sinervo B. 2003 Behavioral drive versus behavioral inertia in evolution: a null model approach. *Am. Nat.* **161**, 357–366. (doi:10.1086/346135)

19. Umbers KDL, Lehtonen J, Mappes J. 2015 Deimatic displays. *Curr. Biol.* **25**, R58–R59. (doi:10.1016/j.cub.2014.11.011)

20. Kang C, Zahiri R, Sherratt TN. 2017 Body size affects the evolution of hidden colour signals in moths. *Proc. R. Soc. B* **284**, 20171287. (doi:10.1098/rspb.2017.1287)

21. Edmunds M. 1974 *Defence in animals: a survey of anti-predator defences*. London, UK: Longman.

22. Endler JA, Westcott DA, Madden JR, Robson T. 2005 Animal visual systems and the evolution of color patterns: sensory processing illuminates signal evolution. *Evolution* **59**, 1795–1818. (doi:10.1111/j.0014-3820.2005.tb01827.x)

23. Slatkin M. 1980 Ecological character displacement. *Ecology* **61**, 163–177. (doi:10.2307/1937166)

24. Holmes GG, Delferrière E, Rowe C, Troscianko J, Skelhorn J. 2018 Testing the feasibility of the startle-first route to deimatism. *Sci. Rep. UK* **8**, 10737. (doi:10.1038/s41598-018-28565-w)

25. Kang C-K, Moon H, Sherratt TN, Lee S-I, Jablonski PG. 2016 Multiple lines of anti-predator defence in the spotted lanternfly, *Lycorma delicatula* (Hemiptera: Fulgoridae). *Biol. J. Linn. Soc.* **1**, 115–124. (doi:10.1111/bij.12847)

26. Svenson GJ, Whiting MF. 2009 Reconstructing the origins of praying mantises (Dictyoptera, Mantodea): the roles of Gondwanan vicariance and morphological convergence. *Cladistics* **25**, 468–514. (doi:10.1111/j.1096-0031.2009.00263.x)

27. Skelhorn J, Rowland HM, Speed MP, Ruxton GD. 2010 Masquerade: camouflage without crypsis. *Science* **327**, 51. (doi:10.1126/science.1181931)

28. Endler JA. 1981 An overview of the relationships between mimicry and crypsis. *Biol. J. Linn. Soc.* **16**, 25–31. (doi:10.1111/j.1095-8312.1981.tb01840.x)

29. Skelhorn J, Rowland HM, Ruxton GD. 2010 The evolution and ecology of masquerade. *Biol. J. Linn. Soc.* **99**, 1–8. (doi:10.1111/j.1095-8312.2009.01347.x)

30. Kusmierski R, Borgia G, Uy A, Crozier RH. 1997 Labile evolution of display traits in bowerbirds indicates reduced effects of phylogenetic constraint.

Proc. R. Soc. Lond. B **264**, 307–313. (doi:10.1098/rspb.1997.0044)

31. Schluter D, Price TD, Grant PR. 1985 Ecological character displacement in Darwin's finches. *Science* **227**, 1056–1059. (doi:10.1126/science.227.4690.1056)

32. Schluter D, McPhail JD. 1992 Ecological character displacement and speciation in sticklebacks. *Am. Nat.* **140**, 85–108. (doi:10.1086/285404)

33. Drummond AJ, Rambaut A. 2007 BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214. (doi:10.1186/1471-2148-7-214)

34. Paradis E, Claude J, Strimmer K. 2004 APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290. (doi:10.1093/bioinformatics/btg412)

35. Redding DW, Mooers AØ. 2006 Incorporating evolutionary measures into conservation prioritization. *Conserv. Biol.* **20**, 1670–1678. (doi:10.1111/j.1523-1739.2006.00555.x)

36. Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO. 2010 Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**, 1463–1464. (doi:10.1093/bioinformatics/btq166)

37. Pagel M. 1999 The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Syst. Biol.* **48**, 612–622. (doi:10.1080/106351599260184)

38. Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W. 2008 GEIGER: investigating evolutionary radiations. *Bioinformatics* **24**, 129–131. (doi:10.1093/bioinformatics/btm538)

39. Bush SE, Weckstein JD, Gustafsson DR, Allen J, DiBlasi E, Shreve SM, Boldt R, Skeen HR, Johnson KP. 2016 Unlocking the black box of feather louse diversity: a molecular phylogeny of the hyper-diverse genus *Brueelia*. *Mol. Phylogenet. Evol.* **94**, 737–751. (doi:10.1016/j.ympev.2015.09.015)

40. Revell LJ. 2012 phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223. (doi:10.1111/j.2041-210X.2011.00169.x)

41. Ives AR, Midford PE, Garland T. 2007 Within-species variation and measurement error in phylogenetic comparative methods. *Syst. Biol.* **56**, 252–270. (doi:10.1080/10635150701313830)

42. Pagel M. 1994 Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc. R. Soc. Lond. B* **255**, 37–45. (doi:10.1098/rspb.1994.0006)

43. Garland T, Dickerman AW, Janis CM, Jones JA. 1993 Phylogenetic analysis of covariance by computer simulation. *Syst. Biol.* **42**, 265–292. (doi:10.1093/sysbio/42.3.265)

44. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. 2020 nlme: linear and nonlinear mixed effects models. R package version 3.1-143. See https://CRAN.R-project.org/package=nlme.

royalsocietypublishing.org/journal/rspb Proc. R. Soc. B 287: 20201016

10

45. Long JS, Freese J. 2014 *Regression models for categorical dependent variables using Stata*, 3rd edn. College Station, TX: Stata Press. See https://www.stata.com/bookstore/regression-models-categorical-dependent-variables/.

46. Akaike H. 1974 *A new look at the statistical model identification*, pp. 215–222. New York, NY; Berlin, Germany: Springer.

47. Svenson GJ, Whiting MF. 2004 Phylogeny of Mantodea based on molecular data: evolution of a charismatic predator. *Syst. Entomol.* **29**, 359–370. (doi:10.1111/j.0307-6970.2004.00240.x)

48. Svenson GJ, Rodrigues HM. 2017 A Cretaceous-aged Palaeotropical dispersal established an endemic lineage of Caribbean praying mantises. *Proc. R. Soc. B* **284**, 20171280. (doi:10.1098/rspb.2017.1280)

49. Bateman AW, Vos M, Anholt BR. 2014 When to defend: antipredator defenses and the predation sequence. *Am. Nat.* **183**, 847–855. (doi:10.1086/675903)

50. Foster SA. 2013 Evolution of behavioural phenotypes: influences of ancestry and expression. *Anim. Behav.* **85**, 1061–1075. (doi:10.1016/j.anbehav.2013.02.008)

51. Loeffler-Henry K, Kang C, Sherratt TN. 2019 Consistent associations between body size and hidden contrasting color signals across a range of insect taxa. *Am. Nat.* **194**, 28–37. (doi:10.1086/703535)

52. Crespi BJ. 1989 Causes of assortative mating in arthropods. *Anim. Behav.* **38**, 980–1000. (doi:10.1016/S0003-3472(89)80138-1)

53. Svenson GJ, Brannoch SK, Rodrigues HM, O'Hanlon JC, Wieland F. 2016 Selection for predation, not female fecundity, explains sexual size dimorphism in the orchid mantises. *Sci. Rep.* **6**, 1–9. (doi:10.1038/srep37753)

54. Cheverud JM. 1996 Developmental integration and the evolution of pleiotropy. *Am. Zool.* **36**, 44–50. (doi:10.1093/icb/36.1.44)

55. Shaw KL, Ellison CK, Oh KP, Wiley C. 2011 Pleiotropy, 'sexy' traits, and speciation. *Behav. Ecol.* **22**, 1154–1155. (doi:10.1093/beheco/arr136)

56. Anreiter I, Sokolowski MB. 2018 Deciphering pleiotropy: how complex genes regulate behavior. *Commun. Integr. Biol.* **11**, 1–4. (doi:10.1080/19420889.2018.1447743)

57. Anholt RRH, Mackay TFC. 2004 Quantitative genetic analyses of complex behaviours in Drosophila. *Nat. Rev. Genet.* **5**, 838–849. (doi:10.1038/nrg1472)

58. Ligon RA, Diaz CD, Morano JL, Troscianko J, Stevens M, Moskeland A, Laman TG, Iii ES. 2018 Evolution of correlated complexity in the radically different courtship signals of birds-of-paradise. *PLoS Biol.* **16**, e2006962. (doi:10.1371/journal.pbio.2006962)

59. Bae S, Kim D, Sherratt TN, Caro T, Kang C. 2019 How size and conspicuousness affect the efficacy of flash coloration. *Behav. Ecol.* **30**, 697–702. (doi:10.1093/beheco/arz006)

60. Edmunds M. 1972 Defensive behaviour in Ghanaian praying mantids. *Zool. J. Linn. Soc.* **51**, 1–32. (doi:10.1111/j.1096-3642.1972.tb00771.x)

61. Pavoine S, Bonsall MB, Dupaix A, Jacob U, Ricotta C. 2017 From phylogenetic to functional originality: guide through indices and new developments. *Ecol. Ind.* **82**, 196–205. (doi:10.1016/j.ecolind.2017.06.056)

62. Clavel J, Morlon H. 2017 Accelerated body size evolution during cold climatic periods in the Cenozoic. *Proc. Natl Acad. Sci. USA* **114**, 4183–4188. (doi:10.1073/pnas.1606868114)

63. Clarke A. 2003 Costs and consequences of evolutionary temperature adaptation. *Trends Ecol. Evol.* **18**, 573–581. (doi:10.1016/j.tree.2003.08.007)

64. Misof B *et al.* 2014 Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346**, 763–767. (doi:10.1126/science.1257570)

65. Umbers KDL, Mappes J. 2015 Post-attack deimatic display in the mountain katydid (*Acripeza reticulata*). *Anim. Behav.* **100**, 68–73. (doi:10.1016/j.anbehav.2014.11.009)

66. Yamawaki Y. 2011 Defence behaviours of the praying mantis *Tenodera aridifolia* in response to looming objects. *J. Ins. Phys.* **57**, 1510–1517. (doi:10.1016/j.jinsphys.2011.08.003)

67. Loeffler-Henry K, Kang C, Yip Y, Caro T, Sherratt TN. 2018 Flash behavior increases prey survival. *Behav. Ecol.* **29**, 528–533. (doi:10.1093/beheco/ary030)

68. Murali G. 2018 Now you see me, now you don't: dynamic flash coloration as an antipredator strategy in motion. *Anim. Behav.* **142**, 207–220. (doi:10.1016/j.anbehav.2018.06.017)

69. Brown SM. 1986 Of mantises and myths: debunking a centuries-old 'truth' about animal behavior has made some researchers look closely at the process of scientific myth making. *BioScience* **36**, 421–423. (doi:10.2307/1310334)

70. Vidal-García M, O'Hanlon JC, Svenson GJ, Umbers KDL. 2020 Data from: The evolution of startle displays: a case study in praying mantises. Dryad Digital Repository. (https://doi.org/10.5061/dryad.4qrfj6q7c)