

Original Article

# Clypeal patterning in the paper wasp *Polistes dominulus*: no evidence of adaptive value in the wild

Jonathan P. Green,<sup>a,b</sup> Ellouise Leadbeater,<sup>a</sup> Jonathan M. Carruthers,<sup>a,b</sup> Neil S. Rosser,<sup>a</sup> Eric R. Lucas,<sup>c</sup> and Jeremy Field<sup>a</sup>

<sup>a</sup>School of Life Sciences, University of Sussex, Falmer BN1 9QG, UK, <sup>b</sup>NERC Biomolecular Analysis Facility – Sheffield, Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield, South Yorkshire S10 2TN, UK, and <sup>c</sup>Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

Status signals function in a number of species to communicate competitive ability to conspecific rivals during competition for resources. In the paper wasp *Polistes dominulus*, variable black clypeal patterns are thought to be important in mediating competition among females. Results of previous behavioral experiments in the lab indicate that *P. dominulus* clypeal patterns provide information about an individual's competitive ability to rivals during agonistic interactions. To date, however, there has been no detailed examination of the adaptive value of clypeal patterns in the wild. To address this, we looked for correlations between clypeal patterning and various fitness measures, including reproductive success, hierarchical rank, and survival, in a large, free-living population of *P. dominulus* in southern Spain. Reproductive success over the nesting season was not correlated with clypeal patterning. Furthermore, there was no relationship between a female's clypeal patterning and the rank she achieved within the hierarchy or her survival during nest founding. Overall, we found no evidence that *P. dominulus* clypeal patterns are related to competitive ability or other aspects of quality in our population. This result is consistent with geographical variation in the adaptive value of clypeal patterns between *P. dominulus* populations; however, data on the relationship between patterning and fitness from other populations are required to test this hypothesis. **Key words:** clypeal pattern, fitness, *Polistes dominulus*, quality, status signalling, wild. [*Behav Ecol*]

## INTRODUCTION

Quantifying the adaptive value of traits is a central goal in behavioral ecology. A trait may be considered to have adaptive value if, in a given environment, its presence enhances fitness relative to that of a variant lacking the trait (Reeve and Sherman 1993). In the case of sexually selected traits, a trait's adaptive value lies in its effect on reproductive success, either via an increase in attractiveness to mates or via an increase in the ability to compete with same-sex rivals for reproductive benefits (Andersson 1994; Clutton-Brock 2007). Research into the adaptive value of sexually selected traits has generally proceeded in 3 phases. First, observational studies have sought to demonstrate correlations between the intensity of a trait and reproductive success (e.g., Siefferman and Hill 2003; Jensen et al. 2004). Second, where such

correlations are found, experimental manipulations of the trait have been performed to test the causality of the association (e.g., Veiga 1993; Petrie and Halliday 1994; Stapley and Keogh 2006). Third, where the impact of a trait on reproductive success has been demonstrated experimentally, an adaptive explanation is sought (e.g., the trait functions as a signal of quality to mates or rivals) (Andersson 1994). Thus, a trait may properly be considered an adaptation that has evolved via sexual selection only in the event that a positive effect of the trait on (lifetime) reproductive success is demonstrated.

In a number of species, individual competitive ability (termed resource-holding potential [RHP]) is known to correlate with aspects of body coloration, with asymmetries in body coloration predicting the outcome of contests over resources (Senar 1999; Whiting et al. 2003). Furthermore, the outcome of such contests can have an impact on reproductive success (e.g., via the attainment of breeding territories), and positive correlations between coloration and reproduction have been reported for some species (Whiting et al. 2003, and references therein). To test whether differences in coloration are directly responsible for variation in resource acquisition among individuals, experimental manipulations of body coloration have also been attempted. In some cases, such manipulations are seen to result in changes in an individual's success during contests (e.g., Qvarnström 1997; Whiting et al. 2003). Patterns of coloration in these species are thus considered to have evolved via sexual selection as an adaptation to competition and have been referred to as

---

Address correspondence to J.P. Green, who is now at Mammalian Behaviour and Evolution Group, University of Liverpool, Leahurst Campus, Neston CH64 7TE, UK. E-mail: [jpgreen@liv.ac.uk](mailto:jpgreen@liv.ac.uk).

E. Leadbeater is now at Institute of Zoology, Zoological Society of London, Regent's Park, London NW4 4RY, UK.

J.M. Carruthers is now at Rothamsted Research, West Common, Harpenden, Hertfordshire AL5 2JQ, UK.

N.S. Rosser is now at Department of Genetics, Evolution & Environment, University College London, Darwin Building, Gower Street, London WC1E 6BT, UK.

Received 3 September 2012; revised 28 November 2012; accepted 24 December 2012.

conventional signals or status signals (Maynard Smith and Harper 1988, 2003).

In the past decade, a number of studies have explored the potential adaptive value of clypeal (facial) patterns as status signals in the paper wasp *Polistes dominulus* (Tibbetts and Dale 2004; Cervo et al. 2008; Tibbetts and Lindsay 2008; Tibbetts et al. 2010; Green and Field 2011b). *P. dominulus* clypeal patterns are fixed throughout adult life and vary among females from a yellow clypeus (i.e., no black spots) to black spots of variable size and shape (Tibbetts and Dale 2004). The active signalling component within the pattern is argued to be the amount of disruption or “brokenness,” which increases as a function of the number and/or waviness of the black spots (Tibbetts and Dale 2004). In populations in the Northeastern United States, clypeal brokenness (hereafter “brokenness”) is positively correlated with body size and predicts the outcome of agonistic interactions between females in the lab (Tibbetts and Dale 2004). Wasps in these populations have also been found to avoid dead conspecifics whose clypeal patterns had been manipulated to increase the amount of brokenness, instead preferring to eat near conspecifics with less broken patterns (Tibbetts and Lindsay 2008; Tibbetts et al. 2010). Based on these findings, clypeal patterning has been suggested to have evolved as a status signal to facilitate rival assessment and thereby reduce the risk of escalated conflict during contests (Tibbetts and Dale 2004; Tibbetts and Lindsay 2008; Tibbetts et al. 2010; see also Tibbetts and Izzo 2010).

In spite of evidence from multiple lab studies that *P. dominulus* clypeal patterns are important in rival assessment, it remains to be shown that individuals with greater brokenness enjoy greater fitness payoffs in the wild than those with a less broken phenotype. Hinting at this possibility are data from a wild US population that show a positive correlation between brokenness and date of emergence among offspring (Tibbetts 2006). Date of emergence is an important predictor of nutritional state, which suggests that brokenness may therefore reflect aspects of condition linked to nutritional status (Tibbetts 2006; see also Tibbetts and Curtis 2007; Tibbetts 2010). In the lab, behavioral observations point to a competitive advantage for females with highly broken patterns (Tibbetts and Dale 2004; Tibbetts and Shorter 2009; Tibbetts, Izzo, et al. 2011). However, the strength of this result remains unclear due to inconsistencies in the relationship between brokenness and body size between studies (Tibbetts and Dale 2004; Tibbetts and Curtis 2007) and variation in the relationship between brokenness and dominance between years (Tibbetts, Izzo, et al. 2011). Moreover, the extent to which behavioral interactions staged in the lab over short time periods and in the absence of fitness-enhancing resources reflect the nature of competitive interactions in the wild is also unclear (Cervo et al. 2008). Attempts to address the functional significance of clypeal patterns in wild populations are further complicated by contradictory evidence for status signalling from populations in the United States and Europe. Using the same experimental approach as Tibbetts and Lindsay (2008), Green and Field (2011b) found no evidence for rival assessment based on clypeal patterning among females in a Spanish population, whereas in an Italian population, Cervo et al. (2008) found no correlation between brokenness and either hierarchical rank or survival in colonies in the lab. These findings may represent a genuine divergence in signal value between populations, possibly as a consequence of population differences in clypeal pattern variability (Tibbetts, Skaldina, et al. 2011; Green et al. 2012). Quantifying the relationship between brokenness and fitness in the wild thus represents a vital step in determining in the first instance whether brokenness has any adaptive

value in the wild, and if so, how and why this varies between populations.

In this study, we investigated the relationship between brokenness and quality in a Spanish population of *P. dominulus* under free-living conditions in the wild. To do this, we first explored the relationship between an individual’s brokenness and fitness, estimated as its reproductive success. In common with other temperate *Polistes* species, reproduction in *P. dominulus* is limited to a single nesting season. For other annual insect species, survival over the breeding season is known to be an important determinant of reproductive success (Lopez-Vaamonde et al. 2009; Rodríguez-Muñoz et al. 2010). However, in the case of *P. dominulus*, the complex social interactions between individuals are also expected to have an important impact. A primitively eusocial species, *P. dominulus*, exhibits considerable flexibility in behavioral and reproductive strategies (Reeve 1991). Mated females emerge from diapause in the spring and found colonies, either singly or, more commonly in the Spanish population, in small groups. Reproduction by singly-nesting females (lone foundresses) is a product of survival and fecundity (Liebert et al. 2005); however, reproduction in groups of cofoundresses is determined in the first instance by hierarchical rank in the group, with the dominant individual producing the vast majority of offspring (Queller et al. 2000; Leadbeater et al. 2011). Consequently, cofoundress reproduction will depend not only on individual attributes such as survival and fecundity but also on group attributes, specifically the relative RHP of other group members, which is thought to play a role in shaping the initial hierarchy (Röseler 1991).

Estimates of foundress reproductive success are also complicated by the change in the reproductive value of offspring across the season (Leadbeater et al. 2011). Although all *P. dominulus* females are capable of mating, the majority of female offspring produced earlier in the season do not reproduce but instead remain on the nest as workers. In contrast, the majority of late-emerging females mate and overwinter to found colonies the following spring (Reeve 1991; Leadbeater et al. 2011). When estimating reproductive success, it is therefore necessary to consider both the number of offspring produced and the time in the season in which they are produced. Adopting a season-wide view of reproduction is also important if the reproductive output of subordinate foundresses is to be quantified accurately. Although the original dominant foundress in a group produces the majority of offspring, her position in the hierarchy may be inherited on her death, allowing former subordinates the opportunity to reproduce (Leadbeater et al. 2011). Indeed, inheritance of the dominant position has been shown to account for the majority of reproduction enjoyed by subordinates and occurs most commonly later in the season, after worker emergence (Leadbeater et al. 2011).

A positive correlation between reproductive success and brokenness would provide strong support for brokenness being an adaptation; however, it would not reveal how brokenness impacts on fitness (Grafen 1988). Numerous aspects of an individual’s quality contribute to its fitness, with complex interactions and trade-offs between individual quality components (Hunt et al. 2004). Previous research indicates that brokenness is most likely to reflect aspects of quality relating to RHP (Tibbetts and Dale 2004; Tibbetts and Shorter 2009; Tibbetts, Izzo, et al. 2011). RHP itself is a complex and multifaceted trait (Parker 1974) and is poorly understood in *P. dominulus*, with potential determinants including body size (Tibbetts and Shorter 2009; Green and Field 2011a; but see Cant et al. 2006) and juvenile hormone titer (Tibbetts, Izzo, et al. 2011). Consequently, when testing for an association between brokenness and RHP in

the wild, it is more straightforward to look for correlations between brokenness and the payoffs that individuals gain through competition. For this reason, in addition to looking at reproductive success, we also explored the relationship between brokenness and rank in the foundress hierarchy, which is closely linked to reproduction and thought to be determined through competition between cofoundresses (Reeve 1991). Finally, given previous findings linking brokenness to individual condition (Tibbetts and Curtis 2007; Tibbetts 2010), we tested for an association between brokenness and survival during the nesting phase, which is likely to vary in part with individual condition and/or underlying quality. Overall, then, our aim was to quantify the relationship between brokenness and measures of fitness in a natural population of *P. dominulus*, and, in doing so, determine the adaptive value of brokenness in the wild.

## MATERIALS AND METHODS

### Study population and nest censuses

We studied *P. dominulus* populations at 3 nearby rural sites close to Conil de la Frontera (Cádiz Province, Spain) from February to July 2008. The sites consisted of arable and pastoral farmland, with nests occurring primarily on hedges of prickly pear cactus (*Opuntia* sp.). Nests were located during the nest-founding phase in late February–early March and subsequently monitored throughout the entire nesting season until dispersal of reproductives in July. At the start of nest founding, a night census of all nests was carried out between 0500 and 0800 when foundresses were cold and inactive. All foundresses present on the nest were collected and given a unique paint mark on the thorax using Humbrol enamel paint. Additionally, the tarsus from one of the middle legs was clipped with scissors and stored in 1 mL absolute ethanol for use later in genotyping. Tarsal sampling has been used in other studies (Liebert et al. 2005; Leadbeater et al. 2010) and allows a foundress' reproductive success to be estimated in the event of her death or disappearance during the season. Two further night censuses were carried out, one in late March and a third and final census in late April at the end of the nest-founding phase. Any new foundresses recorded at these censuses were marked and tarsal-sampled as above. The sequence of field methods performed, together with their timing in relation to the colony cycle, is given in Figure 1.

### Lone foundress reproductive success

Fifty-six lone foundresses were identified from the second night census in late March. The reproductive output of lone foundresses was recorded as the number of pupae in the nest in the late preworker emergence phase. For those nests that did not survive to produce pupae (due to predation or abandonment), a pupal count of zero was recorded.

### Cofoundress reproductive success

#### Field methods

Cofoundresses' reproductive success was estimated by genotyping foundresses and pupae from 38 cofoundress groups. These groups were among a larger sample genotyped for a separate study investigating the frequency of nest inheritance by subordinate foundresses (see Leadbeater et al. 2011). Nests were randomly assigned to particular collection dates at the start of the season. The first set of nests was collected in late April; pupae on these nests represented the first offspring to be produced in the season. Subsequent nest collections were timed so that the oldest pupae on collected nests had pupated

immediately following the previous collection date. Thus, the 5 collection dates spanned the entire offspring production period (for details, see Leadbeater et al. 2011). Collections were made on the following dates (the phase of the colony cycle in which pupae were laid as eggs is indicated in parenthesis): 1) 23 April (nest founding), 2) 21 May (preworker emergence), 3) 18 June (worker emergence), 4) 4 July (late worker—reproductive phase), and 5) 17 July (reproductive phase) (Figure 1). Nests plus associated wasps were collected at night and stored at  $-80^{\circ}\text{C}$  at the University of Cádiz.

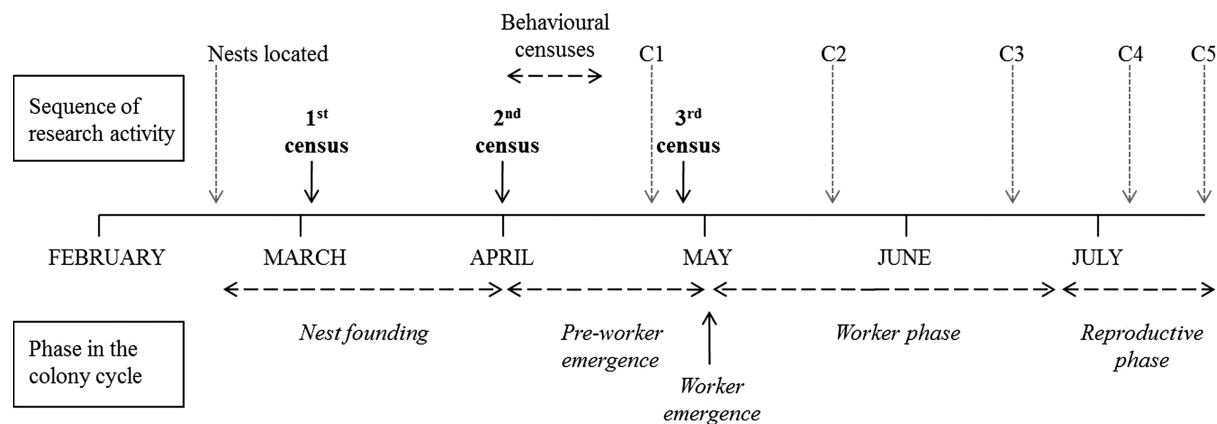
Here, we report the reproductive success of 252 foundresses from the 38 collected cofoundress groups in which  $\geq 1$  foundress had clypeal spots. Between 6 and 9 groups were collected at each collection date. Group size was estimated as the number of foundresses present on the nest at the second night census for nests in the first collection. For subsequent collections, group size was estimated as the number of foundresses that were present at both the second and third night censuses. Group sizes based on these estimates ranged from 2 to 16 foundresses (mean  $\pm$  SE =  $6.63 \pm 0.52$  foundresses).

#### Laboratory methods

With the above sampling regime, it was not possible to quantify the reproductive success of individual foundresses over their entire lifetimes; instead, we examined reproductive success at 5 time points spanning the entire nesting season. Reproductive success was estimated by counting the number of pupae produced by each foundress in a group. Genomic DNA was extracted for each foundress from tarsal samples collected during night censuses. Any unmarked foundresses collected with the nest (i.e., foundresses joining after the last night census) were also genotyped. All pupae on nests were genotyped, except for nests containing  $>30$  pupae at collection, for which a random sample of 30 pupae were genotyped. The genotyping protocol used was identical to that described in Leadbeater et al. (2010). Foundresses and pupae were genotyped at 9 microsatellite markers (*Pbe128TAG*, *Pdom1*, *Pdom2*, *Pdom7*, *Pdom20*, *Pdom22*, *Pdom25*, *Pdom127b*, and *Pdom140*) as described in Leadbeater et al. (2011). Polymerase chain reaction products were genotyped on a 48-capillary ABI3730 DNA Analyzer at the NERC Biomolecular Analysis Facility at Sheffield (NBAF-S). Allele assignment was performed using GeneMapper v3.4 (Applied Biosystems). Results of tests for linkage disequilibrium, deviation from Hardy–Weinberg equilibrium, and heterozygote deficiency for this population using the 9 microsatellites are reported in Lengronne et al. (2012). In each case, no disequilibrium or deficiency was found ( $P > 0.05$ ).

Maternity assignment was performed using the software KINGROUP v.2.9 (Konovalov et al. 2004). To assign offspring to a particular mother in the group, we first used the genetic data to determine the sex of the offspring. In the Hymenoptera, females are diploid, whereas males are haploid. Individuals that were homozygous at every locus were therefore classed as male and those that were heterozygous at one or more loci as female. Given the observed heterozygosities at the 9 loci (Leadbeater et al. 2010, 2011), the probability of a female being homozygous at all 9 loci (and therefore of being wrongly identified as a male) was calculated as  $9.94 \times 10^{-8}$ . We used the Full Sibship Reconstruction procedure to partition female offspring within nests into full-sister groups (for details of the procedure, see Leadbeater et al. 2010). We assumed single mating for all potential mothers (meaning that no individual could be the mother of more than 1 sister group) (Queller et al. 2000; Strassmann 2001). The maternity of male offspring was determined individually for each male by comparing its genotype with those of the adults on the nest.





**Figure 1**

Sequence of field methods used in relation to timing of events in the colony cycle. During the 3 night censuses (in bold), foundresses were painted, photographed, and tarsal-sampled. C1–C5 show the 5 nest collection times. Timing of colony events are approximate and vary from nest to nest depending on age and development. Offspring production is split into a worker phase and a reproductive phase based on data from Leadbeater et al. (2011).

Genotypes were obtained for a total of 514 female and 229 male pupae. Of the female pupae genotyped, it was not possible to assign 11 individuals (1.8%) to an individual foundress, on account of alleles shared with  $\geq 1$  foundress in the group. Of the male pupae genotyped, this proportion was much higher (36%). If males with genotypes common to more than 1 foundress in a group are excluded, male offspring production will consequently be biased toward those foundresses with rarer genotypes (i.e., those more distantly related to other group members). For this reason, we excluded all male offspring from the analyses of cofoundress reproduction. Reproductive success was thus estimated as the number of female pupae produced by each cofoundress in a group.

### Hierarchical rank in cofoundress groups

Within *P. dominulus* cofoundress groups, the propensity to engage in cooperative behaviors such as foraging varies with inheritance rank, with individuals of lower rank spending a greater time away from the nest provisioning than higher-ranking individuals (Cant and Field 2001). Foraging effort (estimated as time spent away from the nest) can thus be used as a rough proxy for a wasp's rank within the hierarchy (Cant and Field 2001). By recording the proportion of daytime censuses in which individuals were off the nest, we were able to infer the rank of 97 foundresses in 18 cofoundress groups in which  $\geq 1$  foundress had clypeal spots. Group sizes ranged from 2 to 14 foundresses (mean  $\pm$  SE =  $5.39 \pm 0.64$ ). Groups were censused between 1100 and 1600 on sunny days when wasps were able to fly. For each group, a minimum period of 45 min separated each census. Censuses were performed for a minimum of 2 consecutive days (mean =  $11.77 \pm 0.67$  censuses per nest) in the 2 weeks following the second night census.

### Foundress survival

Survival was recorded of 118 foundresses in 22 cofoundress groups, in which  $\geq 1$  foundress had clypeal spots. Group sizes ranged from 2 to 13 foundresses (mean  $\pm$  SE =  $5.36 \pm 0.6$ ). Survival was estimated over a 1-month period during the pre-emergence phase of the colony cycle by recording whether foundresses present at the second night census were still on the nest by the third night census.

### Morphological measurements

During night censuses, photographs of the clypeus were taken for all foundresses that had one or more clypeal spots. Wasps were held still and photographed using a Nikon D80 digital camera (with macro lens) from a fixed distance and under constant lighting conditions. Brokenness scores for foundresses with clypeal spots were calculated using the image analysis software MATLAB v.7.1 (The MathWorks Ltd, Cambridge, UK). First, we converted digital images of the clypeus into a standard  $60 \times 30$  pixel bitmap. Following Tibbetts (2010), we then calculated brokenness as the standard deviation of the number of black pixels in each vertical column of the bitmap from pixels 6 to 55 along the horizontal gradient of the clypeus. Pixels 1–5 and 56–60 were excluded as these contained black pixels corresponding to the border of the clypeus rather than to the patterns on the clypeus. Multiple photographs were taken from a sample of 24 individuals to assess the reliability of pattern measurements in the field. Repeatability, calculated using the concordance correlation coefficient (Zar 1999), was high ( $r = 0.96$ , 95% CI: 0.91–0.98). Foundresses without clypeal spots received a brokenness score of zero.

Although the primary focus of this study was brokenness, we also explored the relationship between body size and foundress quality. Although studies have shown that size predicts RHP during usurpation contests (Tibbetts and Shorter 2009; Green and Field 2011a), attempts to assess the importance of body size in the wild have met with mixed results: in some populations, size predicts hierarchical rank (Cervo et al. 2008), whereas in others, it does not (Zanette and Field 2009). However, the relationship between size and reproductive success remains to be explored. Body size was estimated by measuring the width of the first metatarsus at the widest point, measured under a dissecting microscope at  $\times 64$  magnification. Tarsus width was positively correlated with head width (Pearson's  $r = 0.5$ ), a commonly used measure of body size in other studies (e.g., Tibbetts and Dale 2004; Cervo et al. 2008).

### Statistical analysis

#### *Lone foundress reproductive success*

In general, the reproductive output of lone foundresses was very low. Forty-seven of the 56 foundresses failed to produce

pupae and the remaining 9 foundresses produced between 1 and 12 pupae (mean =  $4.33 \pm 1.22$ ). The high number of zero counts resulted in zero inflation of the data, which we therefore analyzed using a generalized linear model (GLM) with a zero-inflated, negative binomial error structure (“ZINB” model; [Zuur et al. 2009](#)). This model assumes that the zero counts are the result of 2 distinct processes: failure of the nest to survive to the pupal stage (modelled by a binomial distribution) and failure of those nests that survive to produce any pupae (modelled by a negative binomial distribution). Because only a small number of lone foundresses (14) had clypeal spots, and hence a brokenness score  $>0$ , presence/absence of clypeal spots was used as a proxy for brokenness in the analysis, following [Tibbetts and Lindsay \(2008\)](#). Individuals with clypeal spots have a higher brokenness than those without spots (see Introduction). Presence of clypeal spots, body size, and site were fitted as explanatory variables in the binomial part of the model, and the presence of spots and body size were the explanatory variables in the negative binomial part of the model.

#### *Cofoundress reproductive success*

Cofoundress reproductive success was estimated as the number of female pupae each foundress produced. The data were analyzed in 2 ways, following the approach of [Zanette and Field \(2009\)](#). Group productivity is strongly correlated with group size in our study population ([Cant and Field 2001](#); [Shreeves et al. 2003](#); [Leadbeater et al. 2011](#)). The reproductive successes of the cofoundresses within a group of a given size are unlikely to be independent, so that partitioning of reproduction approximates a zero-sum game. That is, when one individual produces an offspring, there is effectively one fewer offspring available for production by other individuals. For this reason, offspring produced by different individuals within a group cannot be considered as independent data points. In order to test for a relationship between brokenness and reproductive success, we therefore compared the observed mean within-nest correlation between brokenness and number of pupae with that obtained using simulated groups in which the number of pupae was randomly reallocated among group members. To do this, we first calculated Spearman's rank correlation coefficient ( $\rho$ ) between brokenness and the number of pupae within each group and then calculated the mean correlation across groups. We then recalculated  $\rho$  after the number of pupae was randomly permuted in each group to obtain a simulated mean correlation. This procedure was repeated 10 000 times to obtain a null distribution of means. The proportion of simulated means equal to or greater than the observed mean was used as an estimate of the probability ( $P$ ) of obtaining the observed correlation by chance. Correlations were considered to be significant if  $P < 0.05$ . The same procedure was then used to evaluate the relationship between reproductive success and body size.

Because size and brokenness are examined separately, the above analysis does not control for any correlations between brokenness and body size or for effects of group size and site. We therefore analyzed the combined effect of these variables on reproductive success using GLM. The analysis was divided into 2 parts: factors affecting whether or not a cofoundress produced pupae and factors affecting how many pupae were produced among foundresses who succeed in producing pupae. In the majority (33/38) of cofoundress groups, only a single individual produced pupae; however, in 5 groups, more than 1 foundress produced pupae. In the first step, we tested whether brokenness and/or size affected the likelihood of producing pupae, using a GLM with binomial errors. Because reproduction of individuals in a group is not independent, we sampled 1 foundress at random from each group. Whether or not pupae were produced was the binary response variable

and brokenness, size, and group size were fitted as explanatory variables. To test whether potential effects of brokenness and size vary over the season, we also included interactions between collection date and brokenness and size as additional explanatory variables. This procedure (sampling followed by GLM) was repeated 2000 times, each time using a new randomly sampled subset of foundresses, in order to determine how frequently significant results were obtained at the 95% confidence interval. Explanatory variables were considered significant if their associated  $P$  values within the maximal model are  $<0.05$  in at least 10% of the resampling runs, following [Zanette and Field \(2009\)](#).

In the second part, we examined the factors accounting for variation in the number of pupae produced among wasps that successfully produced pupae. To avoid the issue of non-independence of pupae counts for foundresses in the same group, we included only the most productive foundress from each group (i.e., the dominant reproductive) in the analysis. Data were analyzed in a GLM, using quasi-Poisson errors to account for overdispersion in the data. Number of pupae was the response variable, and brokenness, size, group size, site, and collection date were the explanatory variables. As before, we also fitted interactions between collection data and brokenness and size and collection date.

#### *Hierarchical rank*

Rank within the hierarchy was inferred from a foundress' foraging effort, which we estimated as the proportion of daytime censuses in which that individual was away from the nest. Within a group, the amount of time spent by each wasp foraging is unlikely to be independent of the activities of its nest mates. To examine the influence of brokenness on foraging effort, we therefore compared the observed mean within-nest correlation (Spearman's  $\rho$ ) between brokenness and foraging effort with that obtained using simulated groups in which foraging effort was randomized using the method described above. The same procedure was also used to evaluate the relationship between foraging effort and body size.

As in the analysis of cofoundress reproductive success, the combined effects of brokenness and size on foraging effort were analyzed in a GLM, using quasi-binomial errors to account for overdispersion in the data. Because the foraging efforts of individuals in a group are not independent, we again sampled 1 foundress at random from each group. Proportion of time off the nest was the response variable and brokenness, body size, group size, and site were fitted as explanatory variables. This procedure (sampling followed by GLM) was repeated 2000 times in order to determine how frequently significant results were obtained at the 95% confidence interval. As before, explanatory terms were considered to be significant if their associated  $P$  values are  $<0.05$  in at least 10% of the resampling runs.

In the above analyses of cofoundress rank and reproductive success, we calculated brokenness and size relative to that of other cofoundresses in the group. Relative brokenness was calculated as a foundress' brokenness divided by the mean brokenness of all cofoundresses in the group (relative size was calculated in the same way). Relative values were used because a foundress' hierarchical rank, together with any accompanying reproductive benefits, are thought to be determined at least in part by asymmetries in RHP between cofoundresses during group formation (see Introduction). However, absolute trait values may also be important in determining reproductive success if the number of offspring produced also relates to factors such as condition or fecundity. We, therefore, repeated the analysis of cofoundress reproductive success using absolute values of brokenness and size. Below we report the results for relative values of size and brokenness; using absolute values does not alter the results.

### Foundress survival

To determine whether foundress survival was related to clypeal patterning, we used a generalized linear mixed model (GLMM) with binomial errors where survival (yes or no) was the binary response variable; brokenness, body size, and site were the explanatory variables; and “nest” was fitted as a random factor to control for similar conditions experienced by foundresses within the same group. Due to the high rates of disappearance and nest abandonment among cofoundress groups at Site 3, foundress survival was examined at Sites 1 and 2 only.

All statistical analyses were performed in R v. 2.9.2 (R Development Core Team 2009). Where GLM or GLMM was used to analyze whole data sets (lone foundress and dominant cofoundress reproductive success and foundress survival), model simplification proceeded by backwards deletion of nonsignificant terms until further removals led to a significant ( $P < 0.05$ ) increase in deviance, as assessed by  $F$  values for quasi-binomial errors and  $\chi^2$  values for binomial and negative binomial errors. In cases where GLM was accompanied by random sampling of data points (cofoundress rank and reproductive success), the mean  $t$  values and  $P$  values associated with each variable in a full model containing all explanatory variables across 2000 runs are reported following Zanette and Field (2009). Unless otherwise stated,  $\alpha = 0.05$  for all analyses. Effects of covariates such as site and group size and of interactions between explanatory variables are included in the Results only if statistically significant.

## RESULTS

### Population brokenness

Altogether, 3133 foundresses were marked in the period between nest founding and worker emergence. The proportions of wasps with clypeal spots at the 3 sites were 13.4% (Site 1), 21.5% (Site 2), and 26% (Site 3). Morphological measurements were collected for a total of 804 foundresses, of which 190 had clypeal spots. These data were collected from 56 lone foundresses and 748 cofoundresses in 103 groups (including the subset of 38 groups that were collected for genotyping). Body size did not differ significantly between wasps with and without clypeal spots (mean  $\pm$  SE =  $482.24 \pm 2.19$  vs.  $477.69 \pm 1.42$   $\mu\text{m}$ ; Mann–Whitney  $U$  test,  $P = 0.17$ ). Across all foundresses, body size was not significantly correlated with brokenness (Spearman's  $r = 0.05$ ,  $P = 0.16$ ). Among foundresses with clypeal spots, there was a significant negative correlation between body size and brokenness, that is, smaller wasps had more broken patterns (Spearman's  $r = -0.22$ ,  $P = 0.003$ ).

### Lone foundress reproductive success

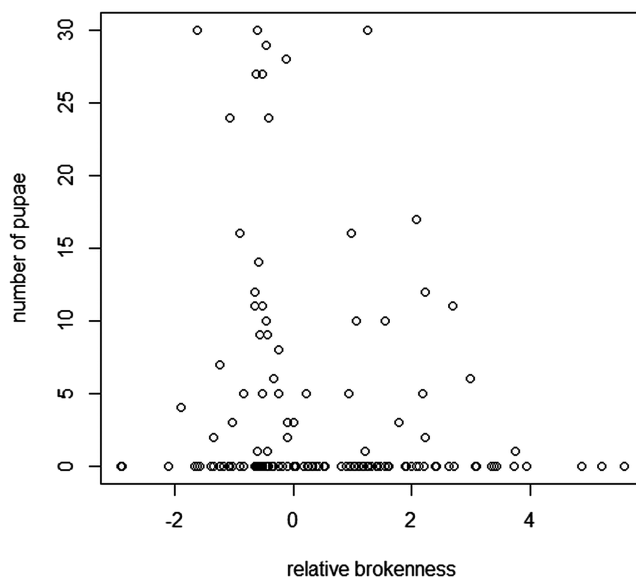
Reproductive success was estimated for 56 lone foundress, of which 14 had clypeal spots. Lone foundress nests experienced a high failure rate compared with cofoundress nests, with no lone foundress nests surviving to produce workers. Reproductive output was measured as the number of pupae in the nest in the late preemergence phase. A lone foundress' brokenness did not predict the number of pupae she produced. Nests founded by wasps with clypeal spots were no more likely to survive to the pupal stage than nests founded by wasps without spots ( $\chi^2_1 = 0.05$ ,  $P = 0.83$ ), and among nests surviving to the pupal stage, foundresses with clypeal spots did not produce more pupae than foundresses without spots ( $\chi^2_1 = 0.07$ ,  $P = 0.79$ ). There was likewise no significant effect of body size on the reproductive output of lone foundresses. In particular, size did not predict survival to the pupal stage ( $\chi^2_1 = 1.05$ ,  $P = 0.31$ ) or the productivity of surviving nests ( $\chi^2_1 = 1.66$ ,  $P = 0.20$ ).

### Cofoundress reproductive success

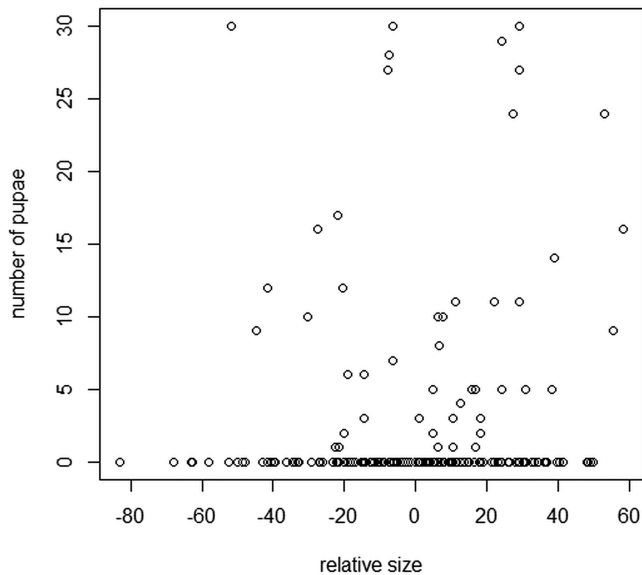
Reproductive success was estimated for 252 foundresses in 38 cofoundress groups, of which 78 had clypeal spots. Forty-seven foundresses (including both dominants and subordinates) successfully produced one or more pupae; the remaining 205 foundresses produced no pupae. Results of the randomization tests showed that the number of pupae a cofoundress produced was not significantly correlated with brokenness (mean  $\rho = 0.03$ ,  $P = 0.71$ ; Figure 2). The number of pupae was positively correlated with body size though this was not quite statistically significant (mean  $\rho = 0.15$ ,  $P = 0.08$ ; Figure 3).

Similar results were obtained when analyzing cofoundress reproductive success using GLM. Whether or not cofoundresses produced pupae was not predicted by brokenness, either across the season as a whole (mean  $t = 0.94$ ,  $P < 0.05$  for 4.45% simulations) or in a particular phase of the colony cycle (brokenness  $\times$  collection date: mean  $t = 0.90$ ,  $P < 0.05$  for 4.05% simulations). Similarly, in this analysis, body size did not predict whether or not a foundress produced offspring (mean  $t = 0.78$ ,  $P < 0.05$  for 3.25% simulations; size  $\times$  collection date: mean  $t = 0.83$ ,  $P < 0.05$  for 4.2% simulations). The likelihood of producing pupae was not predicted by collection date (mean  $t = 0.88$ ,  $P < 0.05$  for 3.1% simulations) but was negatively related to group size (mean  $t = 1.13$ ,  $P < 0.05$  for 11.4% simulations). The effect of group size is expected as larger groups contained more individuals that did not produce offspring, a consequence of the high skew in reproduction within cofoundress groups (see Leadbeater et al. 2011).

Within the 38 cofoundress groups, the proportion of dominant foundresses with clypeal spots did not differ significantly from the proportion of all cofoundresses with clypeal spots (Pearson's  $\chi^2_1 = 0.61$ ,  $P = 0.45$ ). Looking at reproductive success among the 38 dominant reproductives, the number of pupae produced was again not predicted by brokenness, either across the season as a whole ( $F_{1,35} = 0.27$ ,  $P = 0.61$ ) or in a particular phase of the colony cycle (brokenness  $\times$  collection date:  $F_{1,35} = 0.31$ ,  $P = 0.58$ ). Likewise, there was no significant effect of body size on the number of pupae produced



**Figure 2** Relative brokenness and reproductive success for 252 cofoundresses in 38 cofoundress groups.



**Figure 3**  
Relative body size and reproductive success for 252 cofoundresses in 38 cofoundress groups.

( $F_{1,35} = 0.05$ ,  $P = 0.83$ ; size  $\times$  collection date:  $F_{1,35} = 0.29$ ,  $P = 0.60$ ). The analysis showed a significant effect of collection date, with dominant reproductives on nests collected later in the season producing more pupae than those collected earlier in the season ( $F_{1,36} = 5.35$ ,  $P = 0.03$ ).

#### Hierarchical rank

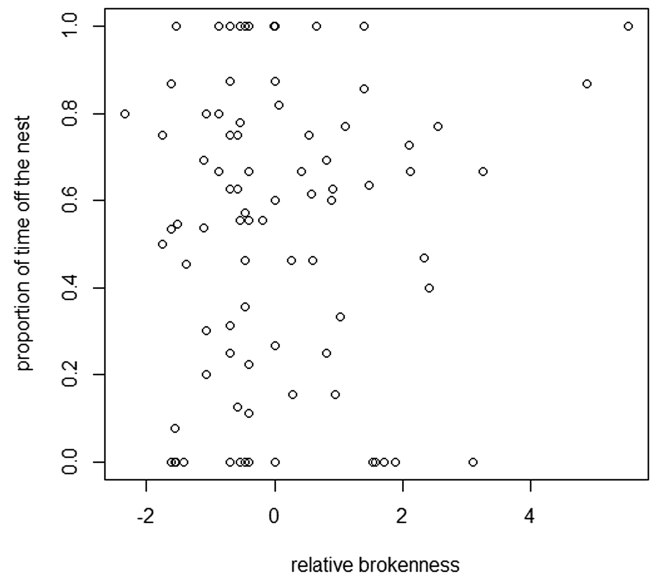
Hierarchical rank was estimated for 97 foundress, of which 38 had clypeal spots. Results of the randomization tests showed that rank was not correlated with brokenness (mean  $\rho = -0.03$ ,  $P = 0.42$ ; Figure 4) or with body size (mean  $\rho = -0.17$ ,  $P = 0.65$ ). Results of the GLM confirmed that rank was not predicted by a foundress' brokenness (mean  $t = 0.74$ ,  $P < 0.05$  for 1.05% simulations) or size (mean  $t = 0.74$ ,  $P < 0.05$  for 1.85% simulations).

#### Foundress survival

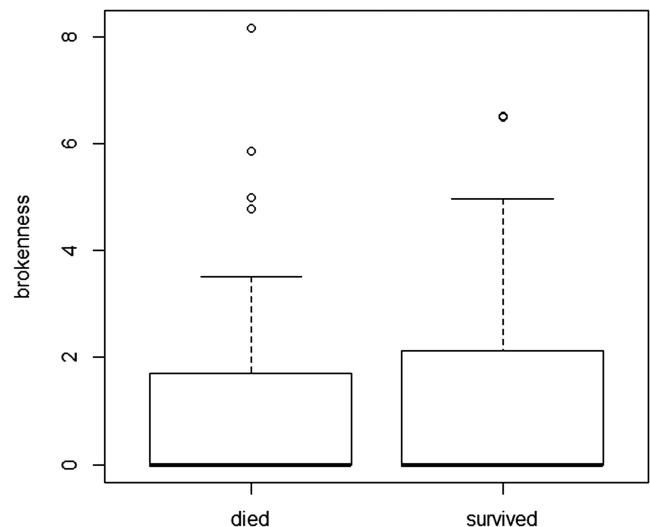
Survival over the preemergence phase was estimated for 118 foundress, of which 42 had clypeal spots. Survival was not significantly predicted by brokenness ( $\chi^2_1 = 0.65$ ,  $P = 0.42$ ; Figure 5) or body size ( $\chi^2_1 = 0.04$ ,  $P = 0.85$ ) but did vary between sites (77% vs. 51%,  $\chi^2_1 = 10.32$ ,  $P = 0.001$ ).

## DISCUSSION

Previous research on rival assessment in *P. dominulus* has indicated that clypeal patterns function as signals of status, signalling a female's RHP to rivals during competition over resources (Tibbetts and Dale 2004; Tibbetts and Lindsay 2008; Tibbetts et al. 2010; but see Green and Field 2011b). To date, however, evidence for an adaptive function of *P. dominulus* clypeal patterns is limited to behavioral studies in the lab, and clear demonstrations of a link between clypeal patterning and quality in wild populations have been lacking. In this study, we present the first data on the relationship between clypeal patterning and fitness in a large, free-living population over the entire nesting cycle. Surprisingly, we found no correlation between clypeal patterning (brokenness) and



**Figure 4**  
Relative brokenness and foraging effort (estimated as the proportion of time off the nest) for 97 cofoundresses in 18 cofoundress groups.



**Figure 5**  
Brokenness versus survival in the preemergence phase for 118 cofoundresses in 22 cofoundress groups.

fitness across a range of contexts, including reproductive success, hierarchical rank, and survival. We consider each of these findings in turn below.

#### Brokenness and reproductive success

We found no relationship between brokenness and reproductive success among foundresses nesting alone or in groups. Lone founding was relatively rare and was associated with very low reproductive success, with the majority of lone foundresses failing to produce any pupae, which may have limited our ability to detect an effect of brokenness. Among lone foundresses, reproductive success is principally determined by survival and fecundity (Liebert et al. 2005); in the case of cofoundress, reproductive success is also expected to be influenced by the relative competitive ability of group members through the presumed effect of RHP on hierarchy



establishment (Röseler 1991). Given that brokenness is argued to be a signal of RHP, we might therefore expect that brokenness more strongly predicts reproductive success of cofoundresses. Surprisingly, however, this was not the case: brokenness was found to have no effect on the likelihood of a foundress producing offspring or on the number of offspring produced by foundresses who succeeded in reproducing.

### Brokenness and hierarchical rank

A previous study by Zanette and Field (2009) failed to identify morphological or behavioral determinants of rank among cofoundresses in the Spanish population though the authors did not explicitly consider the brokenness of foundress clypeal patterns. Nonetheless, the intense fighting that occurs during group establishment would appear to suggest that asymmetries in RHP between cofoundresses are important in hierarchy formation (e.g., Reeve 1991; Röseler 1991). If brokenness reflects RHP, it is therefore expected that brokenness and rank would be positively correlated; indeed, it has been suggested that the brokenness signal should be particularly important during the nest-founding phase when hierarchies are being established (Tibbetts and Lindsay 2008; Tibbetts, Izzo, et al. 2011). However, we found no evidence to support this. Rank in the hierarchy was not correlated with brokenness, indicating that brokenness does not reflect those aspects of quality that determine rank within cofoundress groups.

### Brokenness and survival

In *P. dominulus*, a female's lifespan across the nesting season will determine the number of offspring she produces where she is the dominant reproductive in a group or a lone foundress. For subordinates of a given rank, survival should be positively correlated with the chance of inheriting the dominant position. Though brokenness is primarily argued to reflect RHP, several studies have demonstrated a positive effect of larval nourishment on brokenness (Tibbetts and Curtis 2007; Tibbetts 2010), suggesting that brokenness may also reflect aspects of quality that predict survival. However, this was not the case: foundresses with a higher brokenness did not enjoy greater survival over the preemergence period.

### Body size and fitness in *P. dominulus*

In addition to brokenness, we also explored the relationship between body size and fitness. In a number of *Polistes* populations, size has been shown to correlate with dominance (*P. dominulus*: Pardi 1948; Turillazzi and Pardi 1977; Cervo et al. 2008; *P. fuscatus*: Noonan 1981; *P. metricus*: Dropkin and Gamboa 1981). Many authors have, therefore, assumed that size is an important determinant of RHP (e.g., Röseler 1991) and that size asymmetries help to shape the dominance hierarchy. The limited data available on the role of size in fighting ability in *Polistes* broadly support this idea (Tibbetts and Shorter 2009; Ortolani and Cervo 2010; Cini et al. 2011; Green and Field 2011a; but see Cant et al. 2006). In this study, however, we found that size was not significantly correlated with hierarchical rank, which is consistent with previous research on cofoundress hierarchies in this population (Zanette and Field 2009). Interestingly, in other *P. dominulus* populations in Italy, size and rank are positively correlated (Pardi 1948; Turillazzi and Pardi 1977; Cervo et al. 2008), indicating that population differences exist in the importance of size as a rank determinant though the reason for such differences is currently unclear. Our analysis did reveal a positive, though not quite significant, correlation between size and cofoundress reproductive success. The observation that

larger foundresses tended to produce more offspring suggests that large size might confer some kind of reproductive advantage within the Spanish population. For instance, although size does not appear to determine rank (nor, by extension, the identity of the dominant reproductive), it may be important in determining a dominant's productivity through an effect on condition or fecundity. A positive relationship between size and fecundity in particular is widespread among insects (Nylín and Gotthard 1998), and in *P. dominulus* may drive selection for large body size in females, though further work is required to explore this possibility.

A major, related finding of this study was that body size and clypeal pattern brokenness were not positively correlated. Indeed, among foundresses with clypeal spots, larger individuals had a lower brokenness than smaller individuals. In the original study of *P. dominulus* clypeal patterning in a New York population, Tibbetts and Dale (2004) reported a positive correlation between brokenness and size, which the authors cited as evidence that brokenness reflects RHP (Tibbetts and Dale 2004). Although significant, however, this correlation was very weak ( $r^2 = 0.028$ ) and more recent work by Tibbetts, Skaldina, et al. (2011) found no correlation between brokenness and size in a second US population. In Europe, there is evidence for a positive correlation in some populations (Tibbetts, Skaldina, et al. 2011) but not in others (Cervo et al. 2008; this study). Thus, although size may be a determinant of RHP (see above), brokenness is neither a strong nor consistent indicator of size in *P. dominulus*.

### The relationship between clypeal patterning and quality in *P. dominulus*

Despite evidence from a number of previous studies of a relationship between clypeal patterning and quality (e.g., Tibbetts and Dale 2004; Tibbetts 2006, 2010; Tibbetts and Curtis 2007; Tibbetts and Shorter 2009; Tibbetts and Banan 2010; Tibbetts, Izzo, et al. 2011), this study found no evidence of a relationship between brokenness and fitness in the wild. One possible explanation for the absence of a correlation between brokenness and rank and/or reproductive success is that our analyses were not sufficiently powerful to detect an effect of brokenness in the face of other factors that may impinge on an individual's survival and reproductive success. In studies such as this one, the amount of variability in the focal trait is likely to be an important limiting factor, with low trait variability reducing the power of tests to detect differences in fitness payoffs between individuals with different phenotypes (Grafen 1988). In the Spanish populations, only 15–20% of wasps have clypeal spots (Zanette 2007; this study), meaning that there is no variation in brokenness among cofoundresses in the majority of cofoundress groups (i.e., all have an entirely yellow clypeus). In this study, we attempted to maximize the amount of variation in brokenness in the analyses by including only cofoundress groups containing at least one individual with clypeal spots. An alternative solution to the problem of low variability would be to experimentally increase the amount of variation in brokenness by manipulating clypeal patterns in the field. However, attempts to manipulate clypeal patterns of free-living wasps (e.g., using paint) have so far met with no success (Green JP, Field J, personal observation). Moreover, even if the pattern was a signal of status, changes to the pattern alone may be insufficient to produce changes in fitness in the presence of antichecking mechanisms that are hypothesized to punish individuals whose clypeal patterns do not reliably reflect their true ability (Tibbetts and Dale 2004; Tibbetts and Izzo 2010).



The alternative explanation for our result is that the lack of a relationship between brokenness and quality in our population is a real phenomenon. In support of this idea, previous research by Green and Field (2011b) has shown that clypeal patterns are not used to assess rivals during competition over food in the Spanish population. A separate study by Cervo et al. (2008) on populations in central Italy also found no relationship between brokenness and hierarchical rank or survival in the lab. In both of these populations, variation in clypeal patterning is limited—as noted above, only 15–20% wasps in the Spanish population have clypeal spots, whereas this proportion is around 40% in the Italian populations (Cervo et al. 2008). An important consequence of the reduced variability in clypeal patterning is that its role in mediating competition is necessarily limited, consistent with the findings of Green and Field (2011b). This is because the majority of competitive interactions will occur between individuals without clypeal spots (i.e., with an entirely yellow clypeus), meaning that rival assessment based on clypeal patterning will frequently fail to reveal asymmetries in RHP between individuals (Green and Field 2011b). In accordance with this idea, the proportion of wasps with clypeal spots in the US population where signalling is thought to occur is much higher, at around 80% (Tibbetts and Dale 2004). This suggests that the level of variability in clypeal patterning at the population level may constrain the use of signals based on clypeal patterns and that differences in pattern variability between populations may be an important factor in maintaining variation in the signal value of clypeal patterns between populations (Green and Field 2011b; Tibbetts, Skaldina, et al. 2011; Green et al. 2012).

Intraspecific variation in the function of sexually selected traits has been documented in a number of species (for a review, see Wilczynski and Ryan 1999). Although most of this research has focused on variation in intersexual signalling, there is some evidence that intrasexual signals can also vary in form and function between populations (e.g., Baird et al. 1997; Garamszegi et al. 2006). Given that differences in the variability of *P. dominulus* clypeal patterns between populations appear to be linked to the potential divergence in signalling between populations, what factors might be responsible for maintaining this difference in pattern variability? In view of the fact that the American populations have only become established in the last 30–40 years (Cervo et al. 2000), one factor promoting variation in clypeal patterning between European and American populations could be genetic drift, with populations passing through bottlenecks during the initial founding events. There is evidence from Tibbetts (2010) that the amount of black on the clypeus has a significant heritable component, which is consistent with the idea that initial differences in pattern variability between populations may be maintained or even amplified over generations. However, a study by Liebert et al. (2006) that compared levels of genetic variability in US and European populations has found evidence for several independent founding events, which implies that genetic bottlenecks may not have been particularly severe. Moreover, this hypothesis fails to account for the high level of variability in clypeal patterning reported in some other European populations, which is comparable to that seen in the United States (e.g., Ukraine: Rusina et al. 2006; Tibbetts, Skaldina, et al. 2011).

Alternative hypotheses for population differences in the variability of clypeal patterns have been proposed, including correlated selection on body size (Tibbetts, Skaldina, et al. 2011) and plasticity in the development of the clypeal pattern in response to environmental variation (Green et al. 2012). Results of a recent study by Green et al. (2012) support this latter idea by showing that expression of the pattern is sensitive to temperatures experienced during pupal development with more variable clypeal patterns among wasps reared at lower temperatures. This result

may point to a role for climate in generating and maintaining differences in the variability of clypeal patterns between different populations. However, it remains unclear to what extent these or other factors are responsible for the differences in clypeal patterning we observe among populations, and how this relates to the possible divergence in the signal value of clypeal patterns between populations and geographical locations.

## CONCLUDING REMARKS

The results that we have presented in this study provide no support for the hypothesis that the brokenness of the *P. dominulus* clypeal pattern functions as a signal of quality. Before seeking to explain this result, together with those of Green and Field (2011b) and Cervo et al. (2008), in terms of interpopulation variation in the signal value of patterns, it is first necessary to clarify the role of clypeal patterning in US populations. In US populations, there has not yet been a detailed analysis of the relationship between brokenness and fitness of the kind presented here. In the absence of this, it is not possible to determine unequivocally whether brokenness in these populations is truly an adaptation. To date, the majority of studies seeking to demonstrate an adaptive function for clypeal patterning have been based on experiments in the laboratory. Some of these studies have involved staged contests between wasps in the absence of a fitness-enhancing resource over which they can compete (Tibbetts and Dale 2004; Tibbetts and Izzo 2010; Tibbetts, Izzo, et al. 2011), whereas others have attempted to demonstrate signalling by testing receiver responses to manipulations performed on dead individuals (Tibbetts 2008; Tibbetts and Lindsay 2008; Tibbetts et al. 2010). Although such designs are useful for testing receiver responses to manipulations of conspecific phenotypes, they do not provide any information on the functional significance of the patterns in the wild. The results of the present study have found no evidence that clypeal pattern has adaptive value in a wild European population of *P. dominulus*. If we are to better understand this result, and its implications for status signalling in *P. dominulus* in general, an investigation into how brokenness relates to fitness in US populations is needed.

## FUNDING

Funding was from a GTA studentship from the University of Sussex (to J.P.G.) and Natural Environment Research Council grant NE/E017894/1 (to J.F.).

We are grateful to Juan Miguel Mancera Romero and members of his lab at the University of Cádiz for technical support and use of facilities. Genotyping was performed at the NBAF-S supported by the Natural Environment Research Council, UK (NERC). Terry Burke and Deborah Dawson kindly provided project support and comments on the manuscript, and Andy Krupa provided genotyping assistance. We also thank Thibault Lengronne for help in the field, Jasper van Heusden for assistance in the lab, Mike Cleese for help with the image analysis, Lorenzo Zanette for advice on performing the randomization tests, and 2 anonymous reviewers for useful comments.

**Handling editor:** Alexei Maklakov

## REFERENCES

- Andersson M. 1994. Sexual selection. Chichester (UK): Princeton University Press.
- Baird TA, Fox SF, McCoy JK. 1997. Population differences in the roles of size and coloration in intra- and intersexual selection in the collared lizard, *Crotaphytus collaris*: influence of habitat and social organization. *Behav Ecol*. 8:506–517.

- Cant MA, English S, Reeve HK, Field J. 2006. Escalated conflict in a social hierarchy. *Proc R Soc B*. 273:2977–2984.
- Cant MA, Field J. 2001. Helping effort and future fitness in cooperative animal societies. *Proc R Soc B*. 268:1959–1964.
- Cervo R, Dapporto L, Beani L, Strassmann JE, Turillazzi S. 2008. On status badges and quality signals in the paper wasp *Polistes dominulus*: body size, facial colour patterns and hierarchical rank. *Proc R Soc B*. 275:1189–1196.
- Cervo R, Zacchi F, Turillazzi S. 2000. *Polistes dominulus* (Hymenoptera, Vespidae) invading North America: some hypotheses for its rapid spread. *Insect Soc*. 47:155–157.
- Cini A, Bruschini C, Poggi L, Cervo R. 2011. Fight or fool? Physical strength, instead of sensory deception, matters in host nest invasion by a wasp social parasite. *Anim Behav*. 81:1139–1145.
- Clutton-Brock TH. 2007. Sexual selection in males and females. *Science*. 318:1882–1885.
- Dropkin JA, Gamboa GJ. 1981. Physical comparisons of foundresses of the paper wasp *Polistes metricus* (Hymenoptera: Vespidae). *Can Entomol*. 113:457–461.
- Garamszegi LZ, Rosivall B, Hegyi G, Szöllösi E, Török J, Eens M. 2006. Determinants of male territorial behavior in a Hungarian collared flycatcher population: plumage traits of residents and challengers. *Behav Ecol Sociobiol*. 60:663–671.
- Grafen A. 1988. On the uses of data on lifetime reproductive success. In: Clutton-Brock TH, editor. *Reproductive success*. London: University of Chicago Press. p. 454–471.
- Green JP, Field J. 2011a. Assessment between species: information gathering in usurpation contests between a paper wasp and its social parasite. *Anim Behav*. 81:1263–1269.
- Green JP, Field J. 2011b. Inter-population variation in status signaling in the paper wasp *Polistes dominulus*. *Anim Behav*. 81:205–209.
- Green JP, Rose C, Field J. 2012. The role of climatic factors in the expression of an intrasexual signal in the paper wasp *Polistes dominulus*. *Ethology*. 118:766–774.
- Hunt J, Bussière LF, Jennions MD, Brooks R. 2004. What is genetic quality? *Trends Ecol Evol*. 19:329–333.
- Jensen H, Sæther B-E, Ringsby TH, Tufto J, Griffith SC, Ellegren H. 2004. Lifetime reproductive success in relation to morphology in the house sparrow *Passer domesticus*. *J Anim Ecol*. 73:599–611.
- Kononov DA, Manning C, Henshaw MT. 2004. KINGROUP: a program for pedigree relationship reconstruction and kin group assignments using genetic markers. *Mol Ecol Notes*. 4:779–782.
- Leadbeater E, Carruthers JM, Green JP, van Heusden J, Field J. 2010. Unrelated helpers in a primitively eusocial wasp: is helping tailored towards direct fitness? *PLOS One*. 5:e11997.
- Leadbeater E, Carruthers JM, Green JP, Rosser N, Field J. 2011. Nest inheritance is the missing source of direct fitness in a primitively eusocial insect. *Science*. 333:874–876.
- Lengronne T, Leadbeater E, Patalano S, Dreier S, Field J, Sumner S, Keller L. 2012. Little effect of seasonal constraints on population genetic structure in eusocial paper wasps. *Ecol Evol*. 2:2615–2624.
- Liebert AE, Gamboa GJ, Stamp NE, Curtis TR, Monnet KM, Turillazzi S, Starks PT. 2006. Genetics, behavior and ecology of a paper wasp invasion: *Polistes dominulus* in North America. *Ann Zool Fenn*. 43:595–624.
- Liebert AE, Nonacs P, Wayne RK. 2005. Solitary nesting and reproductive success in the paper wasp *Polistes aurifer*. *Behav Ecol Sociobiol*. 57:445–456.
- Lopez-Vaamonde C, Raine NE, Koning JW, Brown RM, Pereboom JJM, Ings TC, Ramos-Rodriguez O, Jordan WC, Bourke AFG. 2009. Lifetime reproductive success and longevity of queens in an annual social insect. *J Evol Biol*. 22:983–996.
- Maynard Smith J, Harper DGC. 1988. The evolution of aggression: can selection generate variability? *Philos Trans R Soc B*. 319:557–570.
- Maynard Smith J, Harper DGC. 2003 *Animal signals*. Oxford series in ecology and evolution. New York: Oxford University Press.
- Noonan KM. 1981. Individual strategies of inclusive-fitness-maximizing in *Polistes fuscatus* foundresses. In: Alexander RD, Tinkle DW, editors. *Natural selection and social behaviour*. New York (NY): Chiron. p. 18–44.
- Nylin S, Gotthard K. 1998. Plasticity in life-history traits. *Ann Rev Entomol*. 43:63–83.
- Ortolani I, Cervo R. 2010. Intra-specific body size variation in *Polistes* paper wasps as a response to social parasite pressure. *Ecol Entomol*. 35:352–359.
- Pardi L. 1948. Dominance order in *Polistes* wasps. *Physiol Zool*. 21:1–13.
- Parker GA. 1974. Assessment strategies and the evolution of fighting behaviour. *J Theor Biol*. 47:223–243.
- Petrie M, Halliday T. 1994. Experimental and natural changes in the peacock's (*Pavo cristatus*) train can affect mating success. *Behav Ecol Sociobiol*. 35:213–217.
- Queller DC, Zacchi F, Cervo R, Turillazzi S, Henshaw MT, Santorelli LA, Strassmann JE. 2000. Unrelated helpers in a social insect. *Nature*. 405:784–787.
- Qvarnström A. 1997. Experimentally increased badge size increases male competition and reduces male parental care in the collared flycatcher. *Proc R Soc B*. 264:1225–1231.
- R Development Core Team. 2009. *R: a language and environment for statistical computing*. Vienna (Austria): R Foundation for Statistical Computing. ISBN 3-900051-07-0. Available from: <http://www.R-project.org> (Accessed 01 June 2010).
- Reeve HK. 1991. *Polistes*. In: Ross KG, Matthews RW, editors. *The social biology of wasps*. Ithaca (NY): Cornell University Press. p. 99–148.
- Reeve HK, Sherman PW. 1993. Adaptation and the goals of evolutionary research. *Q Rev Biol*. 68:1–32.
- Rodríguez-Muñoz R, Bretman A, Slate J, Walling A, Tregenza T. 2010. Natural and sexual selection in a wild insect population. *Science*. 328:1269–1272.
- Röseler P-F. 1991. Reproductive competition during colony establishment. In: Ross KG, Matthews RW, editors. *The social biology of wasps*. Ithaca (NY): Cornell University Press. p. 309–335.
- Rusina LY, Gilev AV, Skorokhod OV, Firman LA. 2006. Seasonal dynamics of the phenotypic structure of a population of the paper wasp *Polistes dominulus* (Christ) (Hymenoptera, Vespidae). *Entomol Rev*. 86:168–175.
- Senar JC. 1999. Plumage coloration as a signal of social status. *Proc Int Ornithol Cong*. 22:1669–1686.
- Shreeves G, Cant MA, Bolton A, Field J. 2003. Insurance-based advantages for subordinate co-foundresses in a temperate paper wasp. *Proc R Soc B*. 270:1617–1622.
- Siefferman L, Hill GE. 2003. Structural and melanin colouration indicate parental effort and reproductive success in male eastern bluebirds. *Behav Ecol*. 14:855–861.
- Stapley J, Keogh JS. 2006. Experimental and molecular evidence that body size and ventral colour interact to influence male reproductive success in a lizard. *Ethol Ecol Evol*. 18:275–288.
- Strassmann JE. 2001. The rarity of multiple mating by females in the social Hymenoptera. *Insect Soc*. 48:1–13.
- Tibbetts EA. 2006. Badges of status in worker and gyne *Polistes dominulus* wasps. *Ann Zool Fenn*. 43:575–582.
- Tibbetts EA. 2008. Resource value and the context dependence of receiver behaviour. *Proc R Soc B*. 275:2201–2206.
- Tibbetts EA. 2010. The condition dependence and heritability of signaling and nonsignaling color traits in a paper wasp. *Am Nat*. 175:495–503.
- Tibbetts EA, Banan M. 2010. Advertised quality, caste and food availability influence the survival cost of juvenile hormone in paper wasps. *Proc R Soc B*. 277:3461–3467.
- Tibbetts EA, Curtis TR. 2007. Rearing conditions influence quality signals but not individual identity signals in *Polistes* wasps. *Behav Ecol*. 18:602–607.
- Tibbetts EA, Dale J. 2004. A socially enforced signal of quality in a paper wasp. *Nature*. 432:218–222.
- Tibbetts EA, Izzo A. 2010. Social punishment of dishonest signalers caused by mismatch between signal and behaviour. *Curr Biol*. 20:1–4.
- Tibbetts EA, Izzo A, Huang ZY. 2011. Behavioural and physiological factors associated with juvenile hormone in *Polistes* wasp foundresses. *Behav Ecol Sociobiol*. 65:1123–1131.
- Tibbetts EA, Lindsay R. 2008. Visual signals of status and rival assessment in *Polistes dominulus* paper wasps. *Biol Lett*. 4:237–239.
- Tibbetts EA, Mettler A, Levy S. 2010. Mutual assessment via visual status signals in *Polistes dominulus* wasps. *Biol Lett*. 6:10–13.
- Tibbetts EA, Shorter JR. 2009. How do fighting ability and nest value influence usurpation contests in *Polistes* wasps? *Behav Ecol Sociobiol*. 63:1377–1385.
- Tibbetts EA, Skaldina O, Zhao V, Toth AL, Skaldin M, Beani L, Dale J. 2011. Geographic variation in the status signals of *Polistes dominulus* paper wasps. *PLOS One*. 6:e28173.

- Turillazzi S, Pardi L. 1977. Body size and hierarchy in polygynic nests of *Polistes gallicus* (L.) (Hymenoptera Vespidae). *Monit Zool Ital.* 11:101–112.
- Veiga JP. 1993. Badge size, phenotypic quality and reproductive success in the house sparrow: a study on honest advertisement. *Evolution.* 47:1161–1170.
- Whiting MJ, Nagy KA, Bateman PW. 2003. Evolution and maintenance of social status-signalling badges. In: Fox SF, McCoy K, Baird TA, editors. *Lizard social behaviour*. London: JHU Press. p. 47–82.
- Wilczynski W, Ryan MJ. 1999. Geographic variation in animal communication systems. In: Foster SA, Endler JA, editors. *Geographic variation in behaviour: perspectives on evolutionary mechanisms*. Oxford: Oxford University Press. p. 234–261.
- Zanette LRS. 2007. Group formation in a social wasp. London: University College London. p. 64.
- Zanette LRS, Field J. 2009. Cues, concessions and inheritance: dominance hierarchies in the paper wasp *Polistes dominulus*. *Behav Ecol.* 20:773–780.
- Zar JH. 1999. *Biostatistical analysis*. 4th ed. Upper Saddle River (NJ): Prentice Hall.
- Zuur AF, Ieno EN, Walker NJ, Savilev AA, Smith GM. 2009. *Mixed effects models and extensions in ecology with R*. New York (NY): Springer.