INTRODUCTION

Insect societies are characterized by a strong asymmetry in the access to reproduction. In the most classic examples of social insects, such as in honeybees, many ant species and termites, only one (the queen) or a couple (queen and king) of individuals incessantly reproduce in each colony, while all the other individuals (from dozens to millions) will never reproduce in their lifespan, being completely sterile (WILSON, 1971). Many social insect species, such as paper and hover wasps, halictine bees and bumblebees, on the contrary, show a more primitive form of eusociality in which many/all individuals, inside a colony, retain fertility, being thus able to develop ovaries, mate and lay eggs (WILSON, 1971; MICHERER, 1974). In these cases the access to reproduction may widely vary from species to species and from colony to colony, with different individuals showing different levels of access to reproduction (KELLER and REEVE, 1994; SHERMAN et al., 1995). For example, reproduction is almost monopolized in the social wasp Polistes dominula, where the dominant breeder may lay 99% of the eggs (QUELLER et al., 2000), while it is less skewed in some hover wasps species, such as Parischnogaster striatula, where the percentage of wasps which are mated and have developed ovaries can be up to 80% in a colony (BARACCHI et al., 2013).

Understanding the origins and evolution of reproductive skew, as well as its evolutionary and ecological consequences, it’s obviously a keystone aim of social evolution studies (e.g. VEHNRENCAMP, 1983; REEVE and NONACS, 1992; RATNIK and VISSCHER, 1989; KELLER and REEVE, 1999; WEST-EBERHARD, 2003). The first step to be fulfilled is to describe and to measure the amount of reproductive skew, which in turn requires having reliable tools to measure the individual share of reproduction.

Reproductive skew in social insects can be evaluated by recording/counting every egg laying event through behavioural observation (e.g. LIEBIG et al., 2005) or, a posteriori, by genotyping the brood (e.g. QUELLER et al., 2000, but see CINI et al., 2013 for possible limitations). Apart from these quite expensive (in terms of energy and/or time) methods, reproductive skew is usually evaluated by measuring the degree of ovarian development, i.e. by looking at different morphological features, such as number, size, and development of mature and immature oocytes (reviewed in CINI et al., 2013). This approach however has been recently shown to be reliable only when considering the physiological investment in reproduction, while it is rather inconsistent when evaluating the actual individual egg laying rate (CINI et al., 2013).

In this paper I assess the reliability of another physiological trait, the presence and size of yellow bodies. Yellow bodies (hereafter YB, also known as corpora lutea) are the
remains of nurse cells deposited at the base of the ovarioles whenever an egg is laid (Billen, 1985). Their usefulness in indicating previous egg laying activity has been recognized since decades. Evidence supports the assumption that, on average, individuals with high egg laying activity (such as queens) have very evident, dark and big YB, while individuals not laying eggs do not have YB (e.g. DiETEMANN and PEETERS, 2000; D’ETTORE et al., 2004), thus supporting the use of this trait to roughly infer past reproductive activity. Presence and size of YB has also been used, however, at a higher resolution and in a semi-quantitative way, for example to understand which one of several individuals may have had previous higher egg laying activity than other individuals or to infer whether an individual had or had not laid eggs in the past (e.g. Turillazzi, 1988; DiETEMANN and PeETERS, 2000; HOLLDOBLER et al., 2002; FOITZIK et al., 2011; Yagi and HASEGAWA, 2012). Moreover, a central feature of the use of YB is the time frame of past reproductive activity which is considered. While it is clear that a long history of oviposition (such as years, in honeybee or ants) will be likely mirrored by the presence of corresponding YB, it is not clear at all how much reliable YB can be in tracking short term variation in reproductive activity and thus how much these can reflect the egg laying history on the short term. Most of the primitive eusocial species have indeed an annual cycle, with the reproductive phase lasting only some months (especially in temperate species). Moreover, conflicts over the reproduction are usually confined to specific period of the colonial cycle, such as the founding phase, which may last only a few weeks. Understanding whether YB can reliably indicate previous egg laying activity on the short term is thus of overwhelming importance. As far as we know, the reliability of YB in such short term situations has never been formally tested.

In this paper I combined behavioural observations and physiological measurements in the temperate social wasp *Polistes dominula*, in order to determine whether the presence and the size of YB can be used to infer the relative egg laying rate on the short term. The great and dynamic variability in reproductive skew at both the colonial and population levels (see below) qualifies this species as a very suitable model to assess the reliability of the tools used for the assessment of reproductive skew.

**MATERIALS AND METHODS**

The samples used in this experiment derived from a previous comparative investigation on the various indexes of ovarian development (CINI et al., 2013), to which further physiological investigations were performed to assess the presence of YB and measure their size.

**STUDIED SPECIES**

*P. dominula* colonies are found in spring by one or more mated foundresses (associative foundations), which exit from the diapause and found new nests. In associative foundations, a linear hierarchy is soon established among the various females (PARDI, 1942; 1996): one foundress will become the dominant individual and will monopolize reproduction, while the others will become subordinate and will engage in foraging and provisioning (PARDI, 1948; Turillazzi, 1980; DAPPORTO and Turillazzi, 2010; QUELLER et al., 2000). The first workers emerge at the end of May and after some weeks the worker-reared brood will develop into male and female reproductives, which leave the colonies and mate. While males die soon after, mated females aggregate together in large clusters in sheltered places, where they will wait the following founding season in Spring (REEVE, 1991).

Contrarily to advanced eusocial species, such as many ants and bees species, *P. dominula* show a slight caste differentiation, with all individuals that are fully fertile, e. g. able of developing ovaries, mating and laying eggs (REEVE et al., 1998; STRASSMANN et al., 2004; DAPPORTO et al., 2005; LIEBIG et al., 2005). A large variability in reproductive investment among individuals is thus present throughout the whole life-cycle, from the initial founding phase in Spring (in which foundresses have different degrees of access to reproduction according to their rank in the hierarchy) to the worker phase and to the overwintering period (PARDI, 1948; DAPPORTO et al., 2005; DAPPORTO et al., 2006; DAPPORTO and Turillazzi, 2010; TIBBETTS et al., 2011). Moreover, individual access to reproduction may change very rapidly, according to variation in the colonial social environment. For example, the disappearance of the breeder, a decrease in its fertility or its replacement by an heterospecific usurper trigger development of ovaries, egg laying and even mating in subordinates (STRASSMANN et al., 2004; LIEBIG et al., 2005; MONNIN et al., 2009; CINI et al., 2014).

**SAMPLE COLLECTION AND EXPERIMENTAL PROCEDURES**

Twelve *P. dominula* colonies were collected in the surroundings of Florence (Italy) during Spring 2008, before the emergence of workers. They were transferred to the laboratory and reared in cubic glass cages of 15 cm side, with ad libitum water, fly maggots, sugar, and filter paper for nest building. The experiment started when each colony had at least five workers. All wasps have been individually marked with Humbrol modellism colors on thorax and wings. In order to trigger egg laying in a large part of individuals I removed each day, for 22 consecutive days, about half of the immature brood on each nest. This treatment is known to evoke ovarian development and egg laying in both subordinate foundresses and workers (LIEBIG et al., 2005; MONNIN et al., 2009; CINI et al., 2013). As this treatment requires at least 5-6 days in order to evoke ovarian development and egg laying (MONNIN et al., 2009), behavioural observations to record egg laying started only thirteen days after the beginning of the brood removal treatment. Observations took place during the central hours of the day (between 10 AM and 3 PM) and colonies have been observed for about 22 h each (one colony was observed for about 14 h). Ovipositions were defined and recorded as the typical egg-laying postures that lasted 1 min or more (MONNIN et al., 2009; CINI et al., 2013). All individuals were stored at -20 °C at the end of the observation period.

**PHYSIOLOGICAL MEASUREMENTS**

Physiological analyses were performed on 76 wasps: the 42 wasps that had been observed laying at least one egg and a random subset of 34 wasps among the 103 that have never been observed laying an egg. Individual abdomens were dissected under a stereomicroscope, and for each individual I noted the presence/absence of yellow bodies. The inner reproductive apparatus of individuals in which yellow bodies were present and in which all ovaries were intact has been photographed and the area occupied by yellow ovarioles was measured using Image J 1.47v and normalized on the overall ovarian area (all ovaries, regardless the presence of yellow bodies).
RELIABILITY OF YELLOW BODIES AS INDEXES OF EGG LAYING ACTIVITY IN THE PRIMITIVELY EUSOCIAL... 185

Statistical analyses

In order to evaluate the reliability of YB presence as a dichotomous index, i.e. to infer whether an individual has had or not a reproductive activity, I a) compared egg laying rate between wasps with YB and wasps without YB using Mann Whitney Test; b) compared the percentages of wasps with YB in groups of wasps with different egg laying activity, using Chi-square test (in order to avoid alpha inflation due to multiple comparisons, significant results have been calculated according to Step Up False Discovery Rate procedure, BENJAMINI and HOCHBERG (1995) and finally I c) assessed the reliability of YB in predicting egg/non egg layers with binary logistic regression, with presence/absence of YB as a dependent variable. Binary Logistic regressions were run for different values of egg laying threshold (threshold =1, 2, 3, 4, 5, 7, 10; which means that wasps were considered egg layers if they had laid at least 1, 2, 3, 4, 5, 7 or 10 eggs respectively). Percentage of overall correct classification, non egg layers correct classification and egg layers correct classification were computed and reported for each threshold value, to evaluate how performances changes according to changes in the considered egg laying threshold.

In order to assess the reliability of YB area as a quantitative index, i.e. whether the size of yellow bodies may allow inferring the relative egg laying rate of different individuals, I used a similar approach to that of CINTI et al. (2013). I a) evaluated whether YB areas varies (i.e. increases) as egg laying activity (number of eggs laid) increases, through the assessment of the correlation between the numbers of eggs laid and YB areas with Pearson product correlation; b) assessed whether the individual relative positions in egg laying were mirrored by relative positions in YB areas, by assessing with Mantel test the correlation between the matrix of Euclidean distances calculated on the number of eggs laid and the matrix calculated on YB areas; finally c) I calculated the mean error in rank assignment (± dev. st) varies from 4.23 ± 3.03 positions, their egg laying activity. Indeed, the mean error in rank assignment (± dev. st) varies from 4.23 ± 3.03 positions, for any of the different wasps inside a group, nor to rank individuals for their egg laying activity. Among egg layers and non egg layers, but the reliability of YB Area as egg laying activity increases with the increase of egg laying activities, it cannot be reliably used to infer relative levels of egg laying of different wasps inside a group, nor to rank individuals for their egg laying activity. Indeed, the mean error in rank assignment (± dev. st) varies from 4.23 ± 3.03 positions, for any of the different wasps inside a group, nor to rank individuals for their egg laying activity. Among egg layers and non egg layers, but the reliability decreases, due to another increase in the false negative percentage, i.e. more individuals are correctly classified as egg layers. Beyond a certain threshold however the performance again decreases, due to another increase in the false negative percentage (a drop in the ability to correctly classify egg layers).

RESULTS

In the 22 days of experiment 42 out of 179 wasps has been observed laying at least once (on average 3.83 ± 3.43 eggs, min=1, max=15). 20 wasps had visible YB while in 56 wasps it was not possible to detect YB.

a) Is the presence of YB a reliable tool to infer whether an individual has laid or not eggs?

Presence of YB is clearly associated with egg laying activity. Wasps with YB laid significantly more eggs than wasps without YB (number of eggs laid, mean ± dev. st. wasps with YB = 4.45 ± 0.97; wasps without YB = 1.29 ± 0.29 Mann Whitney, U=316.5, p=0.003, n=56 vs 20 vs, Fig. I, 1). However, the relationship between laying eggs and having yellow bodies is not strong. The likelihood of having yellow bodies sharply increases only in wasps having been observed laying from four to six eggs compared to wasps who laid from zero to three eggs, and reaches 80% in those wasps who laid from 7 to 15 eggs (Fig. I, 2, Chi square comparisons: 0 eggs vs 1-3 eggs $\chi^2=0.264$, df=1, p=0.607; 0 eggs vs 4-6 eggs $\chi^2=3.972$, df=1, p=0.046; 0 eggs vs 7-15 eggs $\chi^2=8.842$, df=1, p=0.003; 1-3 eggs vs 4-6 eggs $\chi^2=5.535$, df=1, p=0.019; 1-3 eggs vs 7-15 eggs $\chi^2=0.594$, df=1, p=0.001; 4-6 eggs vs 7-15 eggs $\chi^2=0.336$, df=1, p=0.460; significant comparisons are reported in Fig. I, 2).

Overall, the presence of YB does allow to distinguish among egg layers and non egg layers, but the reliability depends on the threshold set to consider an individual as an egg layer. If the threshold is equal to 1 (i.e. wasps are considered egg layers if they have been observed laying at least once) the presence of YB is not strictly associated with having laid an egg: the overall performance in classifying wasps is about 57%, with a high percentage (66%) of error II type, i.e. false negative, wasps classified as not having laid eggs according to YB who indeed has laid at least one egg. Overall performance increases as the considered egg-laying threshold increases, reaching a maximum for the thresholds 5 and 7 (overall performance 82% and 79% of correct assignments) (Fig. II) and this is due to a decrease in the false negative percentage, i.e. more individuals are correctly classified as egg layers. Beyond a certain threshold however the performance again decreases, due to another increase in the false negative percentage (a drop in the ability to correctly classify egg layers).

b) Does the size of YB allow inferring the relative egg laying rate of different individuals?

YB area increases as the egg laying activity increases, but this positive correlation between YB size and number of eggs laid holds only for dataset 1, i.e. when considering egg layers (defined as individuals who laid at least one egg), but not when considering only those wasps with YB (Table 1, Fig. III, I and 2). In addition, even if significantly correlated, YB area explains only about the 22 % of variation in egg laying activity (Table 1).

Relative positions of individuals according to their egg laying rate are not reliably mirrored by their relative positions based on yellow bodies area, for any of the datasets (Table 1). This suggests that, while YB area increases with the increase of egg laying activities, it cannot be reliably used to infer relative levels of egg laying of different wasps inside a group, nor to rank individuals for their egg laying activity. Indeed, the mean error in rank assignment (± dev. st) varies from 4.23 ± 3.03 positions, for any of the different wasps inside a group, nor to rank individuals for their egg laying activity. Among egg layers and non egg layers, but the reliability decreases, due to another increase in the false negative percentage (a drop in the ability to correctly classify egg layers).

DISCUSSION

The results show that the presence and size of YB are not fully reliable indexes of previous egg laying activity on the short term (few weeks) in the paper wasp Polistes dominula. In particular, while presence and size of YB are
clearly associated with egg laying, overall YB presence does not allow to reliably categorize individuals to the egg-laying/non egg-laying category, and YB size is not a precise quantitative index of previous oviposition.

According to the data, if YB presence is taken as an index to categorize egg layers and non egg layers, the mean error would be between 18 and 43 %. Logically, this error will be especially high when considering low thresholds for the discrimination among egg layers and non egg layers, i.e. when considering as egg layers also those individuals which laid very few eggs. As the threshold is set higher, YB turn to be more reliable but never reach high level performance (best performance of correct assignments is 82%).

When considering the possible mistakes in the classification process, two kinds of errors are possible: type I errors, i.e. classifying as egg layer an individual who never laid an egg, and the opposite type II error, i.e. classifying as non egg layer an individual that laid a certain number of eggs. The two errors have obviously different consequences on data interpretation; however, the use of the presence of YB as a criterion results in relevant levels of both types errors (on average 14 and 12 %).

The results also show that YB size only weakly reflects the individual egg laying rate. If one considers egg layers, YB area only explains the 22% of the total variance in egg laying activity, i.e. two individuals with very similar YB area can have very different egg laying history, and vice versa. If only those individuals with well developed YB are considered (thus possibly filtering out those individuals who for some reasons did not develop yellow bodies) reliability is even worse, and no correlation exists between YB area and egg laying rate.

To resume, this study shows that, on the short term and in species where moderate differences in egg laying are meant to be tracked, YB presence and size have limited reliability and should thus be considered as additional information, but cannot be safely used, on their own, to reliably infer moderate differences in egg laying activity. In particular, classifying egg layers on the basis of YB presence is correct only when dealing with moderate to big differences, i.e. when the aim is to distinguish great egg layers from non or very poor egg layers. Similarly, size of yellow bodies may reliably inform only on big differences among individual egg laying histories.

Why are YB not reliable indicators of egg laying activity? Possible explanations for this discrepancy may refer to limitations of the experimental setting or to inherent biological reasons.

A possible methodological explanation for the discrepancy between YB and egg laying is that behavioural observations did not cover the 24 hours, so that some egg events may have been missed. However, wasps are more
active in the central hours of the day and it is thus likely that a large part of egg laying events was indeed recorded. Moreover, egg laying rate in the central hours mirrors the overall egg laying activity in the 24 hours (CINI et al., 2013), so any temporal bias is expected. Finally, even if a few egg laying events would have gone unnoticed in my behavioural recordings, so that some wasps considered as not having laid eggs may have indeed laid at least one egg, this bias would affect essentially the error type 1, but would not affect error type 2, thus having a limited impact of the overall performance.

Apart from the possible methodological biases, it is rather likely that YB formation does not change linearly with egg laying, and rather requires a relevant egg laying activity to make YB accumulating in the ovaries and become visible. Indeed, my data show that the likelihood of having YB is significantly higher for wasps having laid from four to six eggs compared to those who laid from one to three eggs. This would make YB reliable only when comparing high with low levels of egg laying.

Another explanation may relate to the fact that part of the yellow bodies is indeed due to egg resorption (ovarian atresia) rather than to egg laying, as suggested for several invertebrates (e.g. REUNOV and CRAWFORD, 2010) and for the tropical hover wasp Parischnogaster serrei (TURILLAZZI, 1985). In this sense, trophic cells degeneration due to egg laying and oocytes atresia may mix up and thus confound YB assessment, making less reliable their use.

Finally, it is to be noted that the external validity (ALTAMANN, 1974) of these results is restricted to the time frame chosen for the experiment. I opted for a three weeks experiment, as this time frame fits with the period in which in many primitively eusocial species conflicts over direct reproduction are usually confined (such as annual temperate wasps of the genus Polistes, REEV, 1991). Consequently, my results do not apply to those social insect species in which individuals experience long histories of oviposition (such as years, as in honeybee or ants). On the contrary, as many primitive eusocial species indeed share the feature of reproductive totipotency and thus conflict, as well as a short time of colonial activity (annual), I believe that the results can be translated to many other primitively eusocial species.

ACKNOWLEDGMENTS

I wish to thanks Rita Cervo and Stefania Meconcelli for logistic support and useful discussions and Catherine F. Haut for the language revision.

REFERENCES

BARACCHI D., MAZZA G., CINI A., PETROCELLI I., BIN HASHIM R., TURILLAZZI S., 2013 - Social biology of


