Intense group selection selects for ideal group compositions, but selection within groups maintains them

Jonathan N. Pruitt ^{a,*}, Charles J. Goodnight ^b, Susan E. Riechert ^c

^a *University of Pittsburgh, Department of Biological Sciences, Pittsburgh, PA, U.S.A.*

^b *University of Vermont, Department of Biology, Burlington, VT, U.S.A.*

^c *University of Tennessee Knoxville, Department of Ecology and Evolutionary Biology, Knoxville, TN, U.S.A.*

Article history:

Received 26 June 2016

Initial acceptance 10 August 2016

Final acceptance 3 October 2016

Available online xxx

MS. number: A16-00565R

* Correspondence: J. N. Pruitt, University of Pittsburgh, Department of Biological Sciences, Pittsburgh, PA 15260, U.S.A.

E-mail address: Agelenopsis@gmail.com (J. N. Pruitt).

A group's composition is important for its success. Colonies of the spider *Anelosimus studiosus* appear to have responded to this pressure by evolving the ability to maintain mixtures of docile versus aggressive individuals that help colonies avoid extinction. Here we demonstrate that colony extinction events unite the optimal group composition of all colony constituents, regardless of phenotype, with that of the colony as a whole. This is because colony extinction events explain the majority of individual mortality events in *A. studiosus*. Through within- and across-habitat colony manipulations, we further determined that reduction in reproductive output by individuals bearing overabundant phenotypes underlies the ability of colonies to adaptively regulate their compositions. When we experimentally created colonies with an overabundance of the docile or aggressive phenotype, individuals bearing the overabundant phenotype exhibited reduced reproductive output, which helped to move colony compositions back towards their site-specific optima. Colonies displaced from their native sites continued to recreate the patterns of reproductive output that characterized their site of origin, suggesting a genetic component to this trait. Individuals thus appear to adaptively cull their reproductive output depending on their phenotype and the composition of their colony. There is also considerable parent–offspring colony resemblance in the extent to which colonies can or do track their ideal compositions. This conveys a kind of collective heritability to this trait. Together, while group selection appears to be the principal driver of ongoing selection on group composition in *A. studiosus*, patterns of selection among individuals within groups appear to promote colonies' ability to track their ideal mixtures.

Keywords:

animal personality

behavioural syndrome

green-beard nepotism

frequency dependent selection

multilevel selection

temperament

Further, little evidence exists that selection on groups has promoted the evolution of any trait (Coyne, 2011)

Coyne's comment summarizes a view held by a surprisingly large number of evolutionary biologists. Proponents of multilevel selection argue that variation in the survival and reproductive output of groups can be a potent evolutionary force (Wade, 1978; Wilson, 1983). While there is some disagreement over the ideal definition of group selection (Gardner, 2015a, b; Goodnight, 2015), here we define group selection as selection caused by the differential performance of groups. While it was established long ago in laboratory studies that experimentally imposed group selection can lead to evolutionary changes (Wade, 1977, 1980), and that group selection is regularly detectable in both laboratory (Eldakar et al., 2009; Eldakar et al., 2010; Shaffer et al., 2016) and natural populations (Goodnight & Stevens, 1997; Gordon, 2013), the ability of group selection to drive adaptations in nature is still doubted by many (Eldakar & Wilson, 2011). This is, in part, because there has been a tradition of viewing group selection and inclusive fitness as incompatible theories (Wilson & Hölldobler, 2005; Nowak et al., 2010). This is incorrect. Kin selection theory argues that indirect fitness can be an important contributor to evolutionary change (Hamilton, 1964a, b). Group selection argues that differences among groups in their collective performance can be a potent selective force (Wilson, 1975). There is nothing incompatible with these two notions. Both theoretical frameworks can converge on the same set of equations (Queller, 1991, 1992), however, the kinds of questions addressed by either framework are quite different (Birch & Okasha, 2015; Wilson, 2015). The frameworks are therefore compatible with each other, although they are not equivalent.

The efficacy of group selection is further doubted because few experimental studies have been able to demonstrate that group selection is the force responsible for the evolution of adaptations seen in natural populations (e.g. local adaptation in colony traits; Pruitt & Goodnight, 2014; Shaffer et al., 2016). For most systems, demonstrating that an adaptation is due to group selection would be a daunting and

possibly impossible task, due to the intractable experimental time spans required for tracking colony life history (e.g. decades; Gordon, 2013), and the difficulty of generating and tracking large numbers of experimental groups in situ. A rigorous test of group selection would require that one be able to manipulate the trait or traits that are thought to be under group selection, to deploy experimentally manipulated groups into the field, and then to track group's long-term survival and reproductive output. This is simply impossible for most systems. Social spiders, on the other hand, are a practical model system for the study of group selection because group longevity is short (Aviles, 1986, 1993a; Henschel, 1998), group extinction rates are high (Negron et al., 1991; Pruitt, 2012; Pruitt & Modlmeier, 2015), the number of offspring colonies produced by focal colonies can be estimated (Aviles, 1993b; Pruitt, 2013) and groups rarely mix (Johannesen et al., 2009; Agnarsson et al., 2013). As a result, it is feasible to manipulate collective traits that are thought to be under group selection in the wild.

The facultatively social spider *Anelosimus studiosus* forms multifemale societies in northern portions of its range (Furey, 1998; Jones et al., 2007). These colder climates slow the development of offspring and increase the chance that spiderling groups will be orphaned before they can fledge (Jones & Parker, 2002). This landscape is therefore thought to select for cooperative alloparental care, where females adopt the orphaned offspring of their neighbours (Jones et al., 2007; Riechert & Jones, 2008). This species also exhibits a behavioural polymorphism, where females exhibit one of two relatively discrete (Pruitt & Riechert, 2009a) temporally stable, heritable ($h^2 = 0.66$) behavioural phenotypes: 'docile' or 'aggressive' (Pruitt et al., 2008; Pruitt & Goodnight, 2014). In *A. studiosus*, as in other social *Anelosimus*, aggressiveness is bimodally distributed and individual differences in aggressiveness are repeatable over individuals' adult lifetimes (Pruitt et al., 2011b; Pruitt et al., 2012). Although neither phenotype breeds true in *A. studiosus*, broods produced by docile dams mated with docile sires are predominantly of the docile phenotype, and the same is true of the aggressive phenotype (Pruitt & Riechert, 2009b; Pruitt & Goodnight, 2014).

Populations of *A. studiosus* at higher latitudes also exhibit characteristic differences in their docile:aggressive ratios that vary based on a site's resource levels and as colonies grow (Fig. 1) (Pruitt &

Goodnight, 2014). These site-specific colony compositions are, at least in part, maintained by colony extinction events. The phenotypic compositions that beget success at high-resource sites cause colony extinction at low-resource sites and vice versa. Experimental colonies that exhibit mixtures unlike those of natural colonies at each site quickly go extinct. *Anelosimus studiosus* colonies appear to have responded to this selection pressure by evolving the ability to adjust suboptimal phenotypic ratios in ways that promote colony success (Pruitt & Goodnight, 2014). When colonies' compositions are experimentally altered, colonies are able to shift their behavioural compositions back towards those that characterize their site of origin and facilitate group success. This tendency is maintained even when colonies are reared in contrasting environments for multiple generations, suggesting a genetic component to the mechanisms that allow colonies to adjust their compositions. Thus, group selection appears to have played a role in driving adaptive population-level differences in group composition and the mechanisms that govern it. The mechanism(s) by which these alterations are achieved in *A. studiosus* colonies is (are) presently unknown (Grinsted et al., 2015; Smallegange & Egas, 2015). Although group selection has been detected in several systems (Chang & Sih, 2013; Gordon, 2013), experiments with *A. studiosus* were the first to show that group selection differs across sites and that colonies appear to have responded to this selection.

The first paper on group selection in this system (Pruitt & Goodnight, 2014) received considerable attention but was critiqued for three primary reasons (summarized in Table 1). (1) In the original paper we failed to account for selection within colonies and the role that it might play in guiding the dynamics of the system. This oversight is arguably problematic because some researchers have suggested that simultaneously accounting for individual selection and group selection is required in order to implicate group selection as a guiding evolutionary force. Although we disagree with this assertion, we feel it valuable to address this critique (Grinsted et al., 2015). (2) In our first paper (Pruitt & Goodnight, 2014), we referred to local adaptation in colonies' phenotypic ratios and the mechanisms that guide them as 'group-level adaptations'. Following Gardner's (2014) definition, this implies that the optimal phenotypic ratio for colonies is different from that for the individuals within them and, to be deemed a

‘group-level adaptation’, the compositions seen in nature would need to match the optimum for colonies (Gardner & Grafen, 2009; Gardner, 2014). We did not rigorously evaluate this possibility in our prior paper. In contrast to the first critique, this second critique is important because, if there is disagreement over the ideal composition for individual spiders with that of the group, then this could lead to conflict over control of group membership and composition (Lichtenstein & Pruitt, 2015). Finally, (3) we did not test whether colonies’ tendency to track their ideal mixture is transmitted down colony lineages (Grinsted et al., 2015; Smallegange & Egas, 2015). While transmission of this trait down colony lineages is not required in order for group selection to act on this trait, vertical transmission of some kind is necessary for an effective response to group selection (Darwin, 1872; Boyd & Richerson, 2010). We attempt to address each of these critiques in this paper.

First, we consider here the role of selection among individuals within colonies in guiding the ability of colonies to adjust their phenotypic composition. In particular, we test the hypothesis that once a behavioural phenotype is overabundant within a colony, individuals of that phenotype will tend to reproduce less. This, in turn, will guide the mixture of phenotypes back towards the site-specific optimum. In other words, we ask whether the compositions that beget colony success are also colonies’ equilibrium compositions based on selection among individuals within colonies. The concept of equilibrium selection notes that social interactions within groups can result in multiple equilibria that are locally stable but can differ in their ‘group-level’ properties (Boyd & Richerson, 1990, 2009; Samuelson, 1997). We therefore examine here whether the compositions favoured by group selection are also an equilibrium composition at each site.

Second, we assess whether the ideal group composition for docile or aggressive individuals differs based on within- and among-colony selection. This is important because if the optimal group composition differs for individuals of either phenotype, this could lead to conflict over control of group membership and thwart the ability of a colony to maintain its ideal mixture (Cote et al., 2012; Farine et al., 2015). Identifying differences in the optimal group composition for individuals versus colonies is a prerequisite for labelling group composition as a ‘group-level adaptation’ (Gardner & Grafen, 2009;

Gardner, 2015a). Given the potential for colony extinction events to determine the majority of individual fitness in this system, we predicted that the optimal composition for individuals would be the same as that of the colony.

Third, we assess whether colonies' tendency to adaptively adjust their phenotypic composition runs in colony lineages; parent–offspring colony resemblance in collective traits is argued to be a necessary precondition for an evolutionary response to group selection (Grinsted et al., 2015). If there is such lineage resemblance, this would also provide an independent line of evidence that group composition and the mechanisms that govern it are genetically influenced traits. We predicted that the tendency to regulate group composition would be transmitted down colony lineages in *A. studiosus*.

Historic debates over methodology, rhetoric and the relative importance of individual versus group selection have led to much confusion in our field (Foster, 2009; Okasha, 2010; Birch & Okasha, 2015; Biernaskie & Foster, 2016). We propose that by refocusing the literature's dialogue back on real biological systems using ecologically motivated data, we may be able to shift the literature back towards the goal of enhancing our understanding of natural systems. This practice also promises to allow a greater diversity of scientists, both empiricists and theoreticians, to actively participate in scientific dialogue and to make independent empirically motivated decisions regarding the viability of group selection as an evolutionary force. To this end, the *A. studiosus* system and the data herein are well positioned to address these objectives.

METHODS

Experimental Colony Formation

We created an array of experimental colonies in order to assess the role of colony extinction events in driving the ideal group compositions of individuals and groups, and to assess the effects of differential reproductive output in permitting colonies to track their ideal compositions. We created 34

artificial colonies of a range of sizes (2–27 adult females) and compositions (11–93% aggressive individuals) and established them at each of two riparian study sites: one high-resource site (Little River, TN, U.S.A.) and one low-resource site (Clinch River, TN, U.S.A.) (68 colonies total). Individuals used to construct each experimental colony were all taken from the same source colony; we therefore never created experimental colonies of mixed origin in order to preserve natural levels of within-colony relatedness (Duncan et al., 2010). Each source colony was used to make only one experimental colony. These colonies were deployed in March 2014. Twenty-two of these colonies per site were composed of individuals that were collected, marked and redeployed back at their site of origin (‘native treatment’). The other 12 colonies per site were composed of foreign individuals, which were collected at one site and deployed at a foreign site with contrasting resource levels (i.e. Clinch River individuals to Little River and vice versa) (‘foreign treatment’). All of the resident spiders were individually marked using modelling paint atop their cephalothorax. We subsequently monitored colony survival and the individual reproductive output and survival of two randomly selected docile females and two randomly selected aggressive females per colony. If only one individual of either phenotype was present within the colony, then we monitored only one individual of that phenotype.

Colonies were checked for eggcases every other day throughout the late spring and early summer 2014. When eggcases were observed, we verified visually that each eggcase was viable, weighed it using a portable balance and marked it with a unique colour combination that matched that of its mother. This protocol was viable because once an *A. studiosus* female produces an eggcase, she seizes it within her chelicerae and guards it until it hatches (Brach, 1977). Eggcases are rarely guarded by nonmaternal females, unless eggcases are very near hatching, and never within a week of parturition (Supplementary 1). Eggcase mass is highly correlated with the number of eggs therein (Pearson correlation: $r_{63} = 0.81$, $P < 0.0001$) and the number of offspring that emerge ($r_{58} = 0.82$, $P < 0.0001$), but not the mass of individual eggs ($r_{63} = 0.15$, $P = 0.23$) or individual spiderling mass ($r_{58} = -0.01$, $P > 0.94$) (Supplementary 2). Expanded methods on colony establishment are available in our online supplementary materials

(Supplementary 3) and in previous papers (Pruitt, 2012, 2013). Singleton mortality events were tracked and evidenced by the presence of marked carcasses in the web.

We also noted instances of colony extinction, defined as a mass mortality event resulting in the apparent death of all colony constituents. Colonies were monitored for eggcase production and extinction events every other day until early July 2014, and every 2–4 weeks thereafter until November 2014. The vast majority of colony extinction events occurred after colonies produced eggcases but before those eggcases hatched (see Results). This pattern allowed us to assess the degree of within-colony selection on female behavioural tendency, even for colonies that eventually collapsed. Such data are important because they help provide clues as to how colony compositions would have shifted had they not gone extinct; even under the harshest of circumstances, some colonies with poor-fitting compositions do manage to circumvent extinction (Pruitt & Goodnight, 2014).

In *A. studiosus*, selection among individuals within colonies is thought to be primarily driven by food limitation (Pruitt & Riechert, 2009a) or over fights for favourable microhabitats within colonies (Pruitt et al., 2011a), whereas group selection is primarily driven by variation in colony extinction risk (Pruitt, 2012; Pruitt & Goodnight, 2014). Colony extinctions are mainly driven by eggcase cannibalism at low-resource sites and invasion by foreign heterospecific spiders at high-resource sites (Pruitt, 2012; Pruitt & Goodnight, 2014). At low-resource sites, the majority of eggcase cannibalism occurs too late in the season for females to disperse and initiate their own colonies (Pruitt & Goodnight, 2014). This is because females that produce eggcases late in the season typically lose their broods to autumn cold snaps (Jones et al., 2007; Riechert & Jones, 2008). Thus, the evolutionary pressures that have caused rampant eggcase cannibalism to evolve remain elusive.

Policing Behaviour

For a random subset of 16 experimental colonies (eight colonies per site) we conducted six 20 min field observations per colony where we noted all instances of agonistic behaviour towards our focal

females (two docile and two aggressive females per colony). We focused on native colonies only for these observations. Two of these observation periods were performed during daylight hours, two in the early evening immediately after sunset, and two during early morning twilight hours. The latter two time intervals correspond to the peak activity periods in this species (Watts et al., 2015). Examples of agonistic behaviour included biting or chasing focal females or contacts with colony members that resulted in raised forelegs, which constitutes an aggressive posture in these spiders (Pruitt et al., 2008; Pruitt & Riechert, 2009b). We then tallied the total number of agonistic interactions directed towards individuals of the ‘overabundant’ versus ‘rare’ phenotype and tested for an association between the rate of agonistic behaviours directed towards spiders and their overabundant/rare phenotype distinction. Behavioural phenotypes were deemed to be ‘overabundant’ versus ‘rare’ based on whether the experimental colony resided above or below the natural size/composition relationship at each site (Fig. 1), which also corresponds to the size/composition relationship that maximizes success at each site (Pruitt & Goodnight, 2014).

Parent–Offspring Colony Resemblance

To test for resemblance between parent and offspring colonies’ tendency to track their ideal compositions, we deployed 12 experimental colonies of each of two size/composition combinations at each site: colonies of 12 females including five aggressive females and colonies of 15 females including 11 aggressive females (48 colonies total). Only native individuals were used to create these colonies. Here again, we never mixed individuals from different source colonies. We deployed these colonies into the field in May 2013, recorded the eggcase mass of two colony members of each behavioural phenotype per colony (as above), and reassessed the behavioural composition of each colony the following year (May 2014). This procedure permitted us to assess whether the relative reproductive output of docile versus aggressive individuals could predict shifts in group composition across generations. Eleven of our

colonies collapsed over the course of the year; all 11 colonies collapsed after females produced their eggcases but before any eggcases hatched.

In May 2014 we collected one egg-guarding female per remaining colony (33 females total). New *A. studiosus* colonies are typically formed by singleton females that disperse from their natal webs in order to found new colonies (Brach, 1977; Riechert & Jones, 2008; Pruitt, 2012). Thus, our single-female removal procedure was designed to resemble a natural colony-founding event. We subsequently reared the offspring that emerged from each female's eggcase on a diet of termite workers until they reached maturity in a common garden (~8–10 weeks later). After their third instar, spiderlings were reared individually in 59 ml deli cups in order to limit heterogeneity in social interactions over their development. Upon reaching maturity, we determined the behavioural phenotype of each individual spider, and mated each female to a randomly selected male from the same source colony. This breeding design emulates the within-colony mating system of most social spiders (Riechert & Roeloffs, 1993; Agnarsson et al., 2006; Lubin et al., 2009); *A. studiosus* exhibits a mixed strategy of both inbreeding and outbreeding (Duncan et al., 2010). We then created sib-group colonies that matched the size/composition combination of their parental colonies in the generation before (12 females 42% aggressive or 15 females 73% aggressive). These groups were thus composed of grandoffspring of the original experimental colonies. We failed to obtain sufficient numbers of offspring to create these compositions for three sib groups, resulting in the successful establishment of 30 grandoffspring colonies reared in laboratory.

We maintained these colonies in laboratory for 8 weeks and recorded the eggcase masses of two randomly selected docile and two randomly selected aggressive females per colony. We then compared the relative reproductive outputs of docile versus aggressive individuals in the parental (field-reared) generation versus the grandoffspring (laboratory-reared) generation.

If the regulatory mechanisms used by colonies to manipulate their compositions are genetically influenced and reproduction-based, then (1) we would predict transgenerational resemblance in the reproductive output of docile versus aggressive phenotypes within colonies. And, (2) we would predict a tight association between the relative reproductive output of docile versus aggressive individuals within

each colony in the field and their observed shift in phenotypic composition across generations. Both predictions can be evaluated with this experiment.

Statistical Methods

Data were inspected for normally distributed residuals and heteroscedasticity prior to analysis. We broke our various LMM analyses down into six modular components. The effect tests of interest are noted in the main body of our text but the complete outputs of our models and full model descriptions are provided in our supplementary materials (Supplementary 4). The ‘parameter estimates’ shown below are the estimates of the model parameters and, for each parameter, a corresponding *t* test for the hypothesis that the parameter value is zero. All of our statistical analyses were run using JMP 12.0 (SAS Institute, Cary, NC, U.S.A.).

RESULTS AND DISCUSSION

We first tested whether the relationship between colony size, composition and individual reproductive success (eggcase mass*individual survival) differed for docile versus aggressive individuals when accounting for within-colony selection alone in native colonies: ((egg mass*individual survival)/(average of egg mass*individual survival until the production of their eggcases)). At both sites, we observed that both phenotypes enjoy a rare-type advantage within their colonies (Fig. 2): Clinch River (aggressive phenotype: parameter estimate (frequency of aggressive phenotype) = -0.95 ± 0.26 , $t_{18.16} = -3.68$, $P = 0.0017$; docile phenotype: parameter estimate = 0.98 ± 0.26 , $t_{19.56} = 3.8$, $P = 0.0012$) and Little River (aggressive phenotype: parameter estimate = -0.81 ± 0.21 , $t_{19.16} = -3.9$, $P = 0.001$; docile phenotype: parameter estimate = 0.95 ± 0.24 , $t_{14.67} = 3.92$, $P = 0.0014$). Thus, based on selection among individuals within colonies, the compositions that maximized the reproductive success of either phenotype were diametrically opposed (Fig. 2). It would appear, therefore, that it is in the interest of an individual female

to keep a low representation of its phenotype and large numbers of the opposite phenotype within its colony.

We also found that the ideal group size differed based on an individual's behavioural phenotype: Clinch River (aggressive phenotype: parameter estimate (group size) = -0.02 ± 0.008 , $t_{17.81} = -2.55$, $P = 0.02$; docile phenotype: parameter estimate = 0.02 ± 0.008 , $t_{19.56} = 2.31$, $P = 0.03$) and Little River (aggressive phenotype: parameter estimate = 0.016 ± 0.007 , $t_{19.1} = 2.21$, $P = 0.039$; docile phenotype: parameter estimate = -0.02 ± 0.008 , $t_{10.64} = -2.67$, $P = 0.02$) (see Supplementary 5 for raw eggcase mass data). Aggressive individuals were more successful in smaller colonies and docile individuals performed better in larger colonies at Clinch River (Fig. 2). At Little River, the pattern was reversed. We argue that such differences in optimal group composition and group size based on selection among individuals within colonies could generate conflict over control of group membership, size and composition, unless there is a stronger selective force that overrides it.

We next accounted for colony extinction events in our analysis, which were a major driver of individual fitness at both sites. Across both sites, 63% of colonies went extinct during the course of our study. Of those colony extinction events, 86% (37/43) occurred after females had produced eggcases but before those eggcases hatched, and 9.3% (4/43) occurred in the 2 weeks following the emergence of spiderlings from their eggcases but before spiderlings fledged. Colony extinction events explain 85% of individual survival (LMM: $F_{1,65.1} = 1055.14$, $P < 0.0001$) and individual survival explains 86% of the relative fitness of individuals (LMM: $F_{1,24.9} = 1644.52$, $P < 0.0001$).

While accounting for these extinction events, we recalculated individuals' reproductive success relative to that of the whole population ((egg mass*individual survival*colony survival)/(average of egg mass*individual survival*colony survival for the whole population)). Within each site, when accounting for within- and between-colony selection, the size/composition combination that maximized reproductive success was nearly identical for docile and aggressive individuals (Clinch River: (all interaction terms with behavioural phenotype in LMM) $P > 0.38$; Little River: all $P > 0.74$; Supplementary 1, Fig. 2). Notably, the ideal size/composition relationships differed between sites, as documented in previous

studies (LMM: composition*site: $F_{1,160} = 35.66$, $P < 0.0001$; composition*colony size*site: $F_{1,160} = 5.97$, $P = 0.016$; Fig. 2, see Supplementary 5 for raw colony extinction data). Thus, although both docile and aggressive individuals appear to enjoy a rare-type advantage based on within-colony selection alone, these patterns vanish when one accounts for selection among colonies. In essence, intense group selection (i.e. differential colony extinction) unites the ideal size/composition of both phenotypes, overwhelming the more subtle selection among individuals occurring within colonies. This, in turn, should diminish conflict over control of group membership, since it is in the reproductive interests of both phenotypes and the colony as a whole to maintain these ideal compositions. It is conceptually important to note that although it is in the fitness interests of all colony constituents to maintain certain high-performing group compositions, this does not necessarily mean that individuals can or will adopt strategies that, in aggregate, produce these compositions. Yet, when we observe natural colonies, we see that they do consistently approximate their ideal compositions (Fig. 1). This leads us to ask: how are these ideal mixtures achieved?

We next address the role of differential reproductive output in promoting colonies' ability to hone their compositions. When we created artificial colonies with an overabundance of docile or aggressive individuals, individuals of the overabundant phenotype tended to produce smaller eggcases relative to colony members bearing the opposing phenotype (native colonies: top row of Fig. 3). Notably, whether a particular phenotype was overabundant for the interests of the colony depended on colony size at each site (Figs 1, 2), and the relative reproductive output of the overabundant phenotype tracked these shifts (Fig. 3). This pattern held for native colonies at both sites despite the fact that colonies at each site tracked opposite size/composition relationships (LMM: colony size*site: $F_{1,37} = 13.59$, $P = 0.0007$). Aggressive individuals reproduced more than docile individuals in small colonies at the Clinch River site and the opposite was true at the Little River site. As colonies grew larger, these patterns reversed at each site (Fig. 3; Clinch River (parameter estimate (group size) = -0.04 ± 0.02 , $t_{19} = -2.68$, $P = 0.01$; Little River (parameter estimate (group size) = 0.04 ± 0.02 , $t_{19} = 2.54$, $P = 0.02$). These patterns were projected to help native colonies restore a negative relationship between colony size and the proportion of aggressive

phenotype females at Clinch River site and the reverse relationship at Little River (Fig. 3). Thus, colonies exhibited patterns of selection among individuals within colonies that were projected to adjust phenotypic ratios back towards those that occurred naturally at each site (Fig. 1) and were advantageous for colony success (Fig. 2). The differential reproductive output of the docile and aggressive phenotypes therefore appeared to help native colonies maintain their optimal compositions.

There were large differences noted in the patterns of differential reproductive output in displaced foreign colonies. Their patterns of relative reproductive success more closely matched those of their site of origin, rather than native colonies at their new site. The relationship between the relative reproductive output of docile versus aggressive individuals and colony size differed between native colonies and foreign colonies at each site (ANOVA: Clinch River: colony size*site of origin (native versus foreign): $F_{1,26} = 12.02$, $P = 0.002$; Little River: colony size*site of origin (native versus foreign): $F_{1,26} = 12.63$, $P = 0.002$). However, these differences vanished when comparing colonies at home versus away (Clinch River: colony size*site of release (home versus away): $F_{1,27} = 1.10$, $P = 0.30$; Little River: colony size*site of release (home versus away): $F_{1,25} = 3.83$, $P = 0.061$). Thus, variation in the relative reproductive output of either phenotype actually caused foreign colonies to recreate the size/composition relationships that would have been advantageous had they remained at their site of origin (Figs 2, 3). These patterns of reproductive output were retained even when colonies were moved into contrasting environments, suggesting a genetic component to the trait.

We further detected transgenerational resemblance in colonies' tendency to track their ideal compositions (LMM: $F_{1,24.9} = 43.72$, $P < 0.0001$), and this relationship did not vary by site (LMM: site*project composition of parental colony: $F_{1,21.13} = 0.06$, $P = 0.80$). Parental colonies that tended to shift their compositions towards their naturally occurring size/composition relationship produced F2 descendant colonies with similar tendencies (Fig. 4). This resemblance also held for the small number of parental colonies that tended to shift their compositions in more unexpected directions (Fig. 4). The same pattern held for both sites despite the fact that parent colonies and F2 descendent colonies were reared in contrasting environments (field versus laboratory, respectively). We also detected an association between

our projected compositions of parental colonies based on the relative reproductive output of two docile and two aggressive individuals and the colonies' actual compositions in the following generation (LMM: $F_{1,24.64} = 40.52$, $P < 0.0001$). This outcome did not vary across sites (LMM: site*project composition of parental colony: $F_{1,20.08} = 1.09$, $P = 0.31$). This finding reinforces the inference that relative reproductive output of docile versus aggressive individuals is the primary mechanism of adjusting group compositions in situ. The fact that we detected transgenerational resemblance in ratio adjustments further suggests that the mechanisms colonies use to adjust their compositions are heritable and potentially genetically influenced. Thus, colony lineages with poor-performing ratio adjustment mechanisms are expected to be weeded out over time by group selection.

Although group selection is a significant selective force acting on group composition in *A. studiosus*, site-specific differences in selection among individuals within colonies appears to underlie colonies' ability to track their ideal group compositions. Patterns of selection among individuals within colonies were maintained when colonies were transplanted across sites (Fig. 3) and intercolony differences in selection among individuals within colonies were transmitted down colony lineages (Fig. 4). These findings suggest a genetic component to these dynamics. One may levy the criticism that these interpretations are based on only two sites, but we emphasize here that the size versus composition relationships observed at these two sites resemble those of many other sites with similar resource levels (Pruitt and Goodnight 2014; Goulet et al., 2016; Ingley et al., 2016). Importantly, patterns of selection among individuals within colonies also helped to move colonies towards the size/composition relationships that maximized colony success at their home sites (Figs 2, 3). Thus, the outcomes of selection among individuals within colonies were functionally adaptive for colonies as a whole and the individuals within them. We, therefore, propose that site-specific patterns of selection among individuals within colonies may have evolved as a kind of collective strategy to help groups circumvent extinction. This likely has both direct and indirect fitness benefits for all colony constituents (Hamilton, 1964a, b; Duncan et al., 2010). We notably failed to detect an association between rates of agonistic behaviour displayed towards nestmates and whether they exhibited an overabundant phenotype (all $P > 0.64$;

Supplementary 6), suggesting that policing is unlikely to play a large role in the observed patterns. These data are instead consistent with the hypothesis that colony members bearing an overabundant phenotype wilfully reduce their reproductive output in a manner that maximizes the survival of their own offspring, their kin and their colony. This inference is similar to claims made by Wynne-Edwards (1962) that some animals could evolve mechanisms of curtailing their own fitness because of the benefits that it provides their population.

Conclusion

Group selection is a powerful force shaping colony composition in *A. studiosus* (Pruitt & Riechert, 2011; Pruitt, 2012, 2013; Pruitt & Goodnight, 2014). In *A. studiosus*, group selection arises in the form of frequent extinction events, which is the classic example of strong group selection (Lewontin, 1970). Although both the docile and the aggressive phenotype experience a rare-type advantage based on within-colony selection alone, colony extinctions effectively nullified these patterns (Fig. 2). Instead, when we accounted for selection among colonies, the ideal size/composition relationship for both phenotypes became indistinguishable (Fig. 2). This means that there is likely to be a selective advantage to any strategy that allows individuals to maintain their colonies at or near to these mixtures. Our results further indicate that patterns in selection among individuals within colonies help to maintain high-performing compositions. When we created experimental colonies with too many or too few of either phenotype, we observed a decrease in the relative reproductive output of the ‘overabundant’ phenotype and a concomitant increase in the reproductive output of the ‘rare’ phenotype. This caused colonies to move back towards their site-specific compositions. Data from our reciprocal transplant experiment further indicate that the reproductive patterns exhibited by colonies are locally adapted, because foreign colonies continued to maintain the reproductive patterns (and thus the optimal group compositions) of their site of origin. Had displaced colonies been at their home sites, these patterns would have been advantageous (Figs 2, 3). We also found that patterns of selection among individuals within colonies were

transmitted down colony lineages, meaning that these dynamics can (and possibly do) respond to group selection. We therefore argue that patterns of selection among individuals within colonies might themselves reflect a kind of collective adaptation (although not a ‘group-level adaptation’; Gardner & Grafen, 2009) in this species. Such equilibrium selection scenarios have been deemed possible by evolutionary models of group selection (Boyd & Richerson, 1990, 2009), but experimental verification of their presence has remained elusive until now.

Acknowledgments

We are indebted to Peter Nonacs, Noa Pinter-Wollman, Carl Nick Keiser, Colin Wright, James Lichtenstein, Andy Gardner, Sarah Bengston and three anonymous referees for their comments on previous versions of our manuscript. We thank Jessica Purcell, Alexander DeMarco, Nikki Renee and Fawn Armagost for assistance with data collection. Funding for this research was provided by National Science Foundation grant IOS 1352705 to J.N.P. J.N.P. designed the experiment, helped performed the experiment, and wrote the manuscript. C.J.G. assisted with data analyses and writing of the manuscript. S.E.R. helped design and perform the experiment and assisted with the writing of the manuscript.

References

- Agnarsson, I., L. Aviles, Coddington, J. A., & Maddison, W. P. 2006. Sociality in theridiid spiders: Repeated origins of an evolutionary dead end. *Evolution*, 60, 2342–2351.
- Agnarsson, I., Aviles, L., & Maddison, W. P. 2013. Loss of genetic variability in social spiders: Genetic and phylogenetic consequences of population subdivision and inbreeding. *J. Evol. Biol.*, 26, 27–37.
- Aviles, L. 1986. Sex-ratio bias and possible group selection in the social spider *Anelosimus eximius*. *Am.*

- Nat.*, 128, 1–12.
- Aviles, L. 1993a. Interdemic selection and the sex-ratio: A social spider perspective. *Am. Nat.*, 142, 320–345.
- Aviles, L. 1993b. Newly-discovered sociality in the neotropical spider *Aebutina binotata* Simon (Dictynidae). *J. Arachnol.*, 21, 184–193.
- Biernaskie, J. M., & Foster, K. R. (2016). *Ecology and multilevel selection explain aggression in spider colonies*. Manuscript submitted for publication.
- Birch, J., & Okasha, S. 2015. Kin selection and its critics. *Bioscience*, 65, 22–32.
- Boyd, R., & Richerson, P. J. 1990. Group selection among alternative evolutionarily stable strategies. *J. Theor. Biol.*, 145, 331–342.
- Boyd, R., & Richerson, P. J. 2009. Voting with your feet: Payoff biased migration and the evolution of group beneficial behavior. *J. Theor. Biol.*, 257, 331–339.
- Boyd, R., & Richerson, P. J. 2010. Transmission coupling mechanisms: Cultural group selection. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365, 3787–3795.
- Brach, V. 1977. *Anelosimus studiosus* (Araneae-Theridiidae) and evolution of quasi-sociality in theridiid spiders. *Evolution*, 31, 154–161.
- Chang, A. T., & Sih, A. 2013. Multilevel selection and effects of keystone hyperaggressive males on mating success and behavior in stream water striders. *Behav. Ecol.*, 24, 1166–1176.
- Cote, J., Fogarty, S., & Sih, A. 2012. Individual sociability and choosiness between shoal types. *Anim. Behav.*, 83, 1469–1476.
- Coyne, J. A. 2011. Can Darwinism improve Binghamton? *The New York Times Sunday Book Review*, 9 September.
- Darwin, C. 1872. *The origin of species* (6th London ed.). Chicago, IL: Thompson and Thomas.
- Duncan, S. I., Riechert, S. E., Fitzpatrick, B. M., & Fordyce, J. A. 2010. Relatedness and genetic structure in a socially polymorphic population of the spider *Anelosimus studiosus*. *Mol. Ecol.*, 19, 810–818.

- Eldakar, O. T., Dlugos, M. J., Pepper, J. W., & Wilson, D. S. 2009. Population structure mediates sexual conflict in water striders. *Science*, 326, 816–816.
- Eldakar, O. T., & Wilson, D. S. 2011. Eight criticisms not to make about group selection. *Evolution*, 65, 1523–1526.
- Eldakar, O. T., Wilson, D. S., Dlugos, M. J., & Pepper, J. W. 2010. The role of multilevel selection in the evolution of sexual conflict in the water strider *Aquarius remigis*. *Evolution*, 64, 3183–3189.
- Farine, D. R., Montiglio, P.-O., & Spiegel, O. 2015. From individuals to groups and back: The evolutionary implications of group phenotypic composition. *Trends Ecol. Evol.*, 30, 609–621.
- Foster, K. R. 2009. A defense of sociobiology. *Cold Spring Harbor Symp. Quant. Biol.*, 74, 403–418.
- Furey, F. E. 1998. Two cooperatively social populations of the theridiid spider *Anelosimus studiosus* in a temperate region. *Anim. Behav.*, 55, 727–735.
- Gardner, A. 2014. Group selection versus group adaptation. *Nature*, 524, E3–E4.
doi:10.1038/nature14596.
- Gardner, A. 2015a. The genetical theory of multilevel selection. *J. Evol. Biol.*, 28, 305–319.
- Gardner, A. 2015b. More on the genetical theory of multilevel selection. *J. Evol. Biol.*, 28, 1747–1751.
- Gardner, A., & Grafen, A. 2009. Capturing the superorganism: A formal theory of group adaptation. *J. Evol. Biol.*, 22, 659–671.
- Goodnight, C. J. 2015. Multilevel selection theory and evidence: A critique of Gardner, 2015. *J. Evol. Biol.*, 28, 1734–1746.
- Goodnight, C. J., & Stevens, L. 1997. Experimental studies of group selection: What do they tell us about group selection in nature? *Am. Nat.*, 150(Suppl.), S59–S79.
- Gordon, D. M. 2013. The rewards of restraint in the collective regulation of foraging by harvester ant colonies. *Nature*, 498, 91–93. doi:10.1038/nature12137.
- Goulet, C. T., Ingley, S. J., Scharf, I., & Pruitt, J. N. 2016. Thermal effects on survival and reproductive performance vary according to personality type. *Behavioral Ecology*. Advance online publication.
doi:https://doi.org/10.1093/beheco/arw084

- Grinsted, L., Bilde, T., & Gilbert, J. 2015. Questioning evidence of group selection in spiders. *Nature*, 524, E1–E3.
- Hamilton, W. D. 1964a. The genetical evolution of social behaviour. I. *J. Theor. Biol.*, 7, 1–16.
- Hamilton, W. D. 1964b. The genetical evolution of social behaviour. II. *J. Theor. Biol.*, 7, 17–52.
- Henschel, J. R. 1998. Predation on social and solitary individuals of the spider *Stegodyphus dumicola* (Araneae, Eresidae). *J. Arachnol.*, 26, 61–69.
- Ingle, S. J., Purcell, J., & Pruitt, J. N. 2016. Social context, but not individual personality, alters immigrant viability in a spider with mixed social structure. *Anim. Behav.*, 120, 153–161.
- Johannesen, J., Wickler, W., Seibt, U., & Moritz, R. F. A. 2009. Population history in social spiders repeated: Colony structure and lineage evolution in *Stegodyphus mimosarum* (Eresidae). *Mol. Ecol.*, 18, 2812–2818.
- Jones, T. C., & Parker, P. G. 2002. Delayed juvenile dispersal benefits both mother and offspring in the cooperative spider *Anelosimus studiosus* (Araneae : Theridiidae). *Behav. Ecol.*, 13, 142–148.
- Jones, T. C., & Riechert, S. E. 2008. Patterns of reproductive success associated with social structure and microclimate in a spider system. *Anim. Behav.*, 76, 2011–2019.
- Jones, T. C., Riechert, S. E., Dalrymple, S. E., & Parker, P. G. 2007. Fostering model explains variation in levels of sociality in a spider system. *Anim. Behav.*, 73, 195–204.
- Lewontin, R. C. 1970. The units of selection. *Annu. Rev. Ecol. Syst.*, 1, 1–18.
- Lichtenstein, J. L. L., & Pruitt, J. N. 2015. Similar patterns of frequency- dependent selection on animal personalities emerge in three species of social spiders. *J. Evol. Biol.*, 28(6), 1248–1256.
doi:10.1111/jeb.12651.
- Lubin, Y., Birkhofer, K., Berger-Tal, R., & Bilde, T. 2009. Limited male dispersal in a social spider with extreme inbreeding. *Biol. J. Linn. Soc.* 97, 227–234.
- Negron, J. A., Rodriguez, R., & Aviles, L. 1991. Fungicide evaluation for control of powdery mildew, *Leveillula taurica* (Lev.) Arn., of peppers in Puerto Rico. *Journal of Agriculture of the University of Puerto Rico*, 75, 119–123.

- Nowak, M. A., Tarnita, C. E., & Wilson, E. O. 2010. The evolution of eusociality. *Nature*, 466, 1057–1062.
- Okasha, S. 2010. Altruism researchers must cooperate. *Nature*, 467, 653–655.
- Pruitt, J. N. 2012. Behavioural traits of colony founders affect the life history of their colonies. *Ecol. Lett.*, 15, 1026–1032.
- Pruitt, J. N. 2013. A real-time eco-evolutionary dead-end strategy is mediated by the traits of lineage progenitors and interactions with colony invaders. *Ecol. Lett.*, 16, 879–886.
- Pruitt, J. N., Demes, K. W., & Dittrich-Reed, D. R. 2011a. Temperature mediates shifts in individual aggressiveness, activity level, and social behavior in a spider. *Ethology*, 117, 318–325.
- Pruitt, J. N., & Goodnight, C. J. 2014. Site-specific group selection drives locally adapted colony compositions. *Nature*, 28, 1248–1256.
- Pruitt, J. N., Iturralde, G., Aviles, L., & Riechert, S. E. 2011b. Amazonian social spiders share similar within-colony behavioural variation and behavioural syndromes. *Anim. Behav.*, 82, 1449–1455.
- Pruitt, J. N., & Modlmeier, A. P. 2015. Animal personality in a foundation species drives community divergence and collapse in the wild. *J. Ecol.*, 84(6), 1461–1468. doi:10.1111/1365-2656.12406.
- Pruitt, J. N., Oufiero, C. E., Aviles, L., & Riechert, S. E. 2012. Iterative evolution of increased behavioral variation characterizes the transition to sociality in spiders and proves advantageous. *American Naturalist*, 180, 496–510.
- Pruitt, J. N., & Riechert, S. E. 2009a. Frequency-dependent success of cheaters during foraging bouts might limit their spread within colonies of a socially polymorphic spider. *Evolution*, 63, 2966–2973.
- Pruitt, J. N., & Riechert, S. E. 2009b. Sex matters: Sexually dimorphic fitness consequences of a behavioural syndrome. *Anim. Behav.*, 78, 175–181.
- Pruitt, J. N., & Riechert, S. E. 2011. How within-group behavioral variation and task efficiency enhance fitness in a social group. *Proc. R. Soc. B: Biol. Sci.*, 278, 1209–1215.
- Pruitt, J. N., Riechert, S. E., & Jones, T. C. 2008. Behavioural syndromes and their fitness consequences

- in a socially polymorphic spider, *Anelosimus studiosus*. *Anim. Behav.*, 76, 871–879.
- Queller, D. C. 1991. Group selection and kin selection. *Trends Ecol. Evol.*, 6, 64.
- Queller, D. C. 1992. Quantitative genetics, inclusive fitness, and group selection. *Am. Nat.*, 139, 540–558.
- Riechert, S. E., & Jones, T. C. 2008. Phenotypic variation in the social behaviour of the spider *Anelosimus studiosus* along a latitudinal gradient. *Anim. Behav.*, 75, 1893–1902.
- Riechert, S.E., & Roeloffs, R. 1993. Inbreeding and its consequences in the social spiders. In N. Wilmsen Thornhill (Ed.), *The natural history of inbreeding and outbreeding* (pp. 283–303). Chicago, IL: University of Chicago Press.
- Samuelson, L. 1997. *Evolutionary games and equilibrium selection*. Cambridge, MA: MIT Press.
- Shaffer, Z., Sasaki, T., Haney, B., Janssen, M., Pratt, S. C., & Fewell, J. H. 2016. The foundress's dilemma: Group selection for cooperation among queens of the harvester ant, *Pogonomyrmex californicus*. *Scientific Reports*, 6, 29828. doi:10.1038/srep29828.
- Smallegange, I. M., & Egas, M. 2015. Good for the group? Explaining apparent group-level adaptation. *Trends Ecol. Evol.*, 30, 379–381.
- Wade, M. J. 1977. Experimental-study of group selection. *Evolution*, 31, 134–153.
- Wade, M. J. 1978. Critical review of models of group selection. *Q. Rev. Biol.*, 53, 101–114.
- Wade, M. J. 1980. Group selection, population growth rate, and competitive ability in the flour beetles, *Tribolium* spp. *Ecology*, 61, 1056–1064.
- Watts, J. C., Ross, C. R., & Jones, T. C. 2015. Diel and life-history characteristics of personality: Consistency versus flexibility in relation to ecological change. *Anim. Behav.*, 101, 43–49.
- Wilson, D. S. 1975. Theory of group selection. *Proc. Natl. Acad. Sci. U. S. A.*, 72, 143–146.
- Wilson, D. S. 1983. The group selection controversy: History and current status. *Annu. Rev. Ecol. Syst.*, 14, 159–187.
- Wilson, D. S. 2015. *Does altruism exist? Culture, genes, and the welfare of others*. New Haven, CT: Yale University Press.

- Wilson, E. O., & Hölldobler, B. 2005. Eusociality: Origin and consequences. *Proc. Natl. Acad. Sci. U. S. A.*, 102, 13367–13371.
- Wynne-Edwards, V. C. 1962. *Animal dispersion in relation to social behaviour*. Edinburgh, U.K.: Oliver & Boyd.

Table 1

Summary of the main criticisms of Pruitt and Goodnight (2014) and how the data collected for the present study helps to address these critiques

Criticisms of Pruitt & Goodnight (2014)	Conclusions from this study
(1) Failure to account for individual selection within colonies and its influence on the <i>A. studiosus</i> system	Patterns of selection among individuals within colonies drive colonies' compositions towards equilibrium mixtures that promote colony survival. Evidence suggests that these dynamics are themselves the adaptive outcome of group selection
(2) Group composition referred to as a 'group-level adaptation'. By Gardner's (2014) definition, a group-level adaptation refers to a trait that approximates the optimal trait value for the group at the cost of the individuals within it. Thus, the interests of the individual and the group must be at odds	Colony extinction events drive the majority of individual fitness in <i>A. studiosus</i> . Therefore, the compositions that promote colony survival also maximize individual fitness. Group composition cannot, therefore, be referred to as a group-level adaptation following Gardner's (2014) definition. Instead, group composition is merely an adaptive collective trait that is honed by group selection. There is presently no widely accepted term for such a trait
(3) Failure to demonstrate that variation in the mechanisms governing group composition run in colony lineages, which is argued to be a precondition for an evolutionary response to group selection	We demonstrate here that subtle variation in colonies' tendency to adjust their compositions is transmitted down colony lineages for multiple generations. This provides further evidence that ratio adjustment mechanisms are genetically influenced and locally adapted in this species

Figure 1. The naturally occurring relationship between colony size and composition (ratio of aggressive:docile phenotypes) of *A. studiosus* at the two study sites: Clinch River (2008: ●; 2010: ●); Little River (2007: ●; 2008: ●).

Figure 2. Relative fitness of aggressive individuals (top row), docile individuals (middle row) and whole colonies (bottom row) as a function of group size (X axis) and group composition (proportion of colony members with aggressive phenotype, Y axis) for our two study sites, Clinch River and Little River. The data shown are for ‘native treatment’ colonies only. Panels labelled ‘within-colony selection’ account for fitness (egg mass*individual survival) relative to fellow colony members. Panels labelled ‘within- and between-colony selection’ account for colony extinction events and compare individuals’ fitness relative to individuals across the entire population. Warm colours indicate individuals or colonies that exhibited higher relative fitness in terms of survival and reproductive output. Patterns of selection among colonies are more apparent (blue versus red) because there are large differences in colonies’ survival and reproductive output in this system. In contrast, patterns of selection within colonies are more subtle because few individuals perish outside of colony extinction events, and thus, we are comparing relatively small differences in fitness (blue versus green).

Figure 3. Relative fitness of docile and aggressive individuals based on within-colony selection alone as a function of group size (X axis) and composition (proportion of colony members with aggressive phenotype, Y axis). Warm colours indicate that aggressive individuals outperform docile individuals, and cool colours indicate that docile individuals outperform aggressive individuals. The plots on the right depict the expected growth/composition trajectories of colonies as a result of the relative reproductive success of either behavioural phenotype (assuming heritability = 1) and the average colony growth at each site. ‘Natives’ were colonies that remained at their home sites; ‘foreigns’ were colonies that were transplanted across sites (Little River to Clinch River and vice versa). These native colonies are the same colonies depicted in Fig. 2.

Figure 4. Depiction of the shifts in colony size and composition. All colonies started at one of two size/composition combinations. The left panels (F1 Projected) depict the projected shift in colony size and composition from the parental to the F1 generation using data on the relative reproductive output of a subset of docile and aggressive colony members. The centre panels (F1 Outcome) depict the actual shifts in colony size and composition over the next generation. The right panels (F2 Projected) depict the projected shift in colony size and composition in grandoffspring colonies reared in the laboratory, again using data on the relative reproductive output of a subset of docile and aggressive colony members. Within each row, parental and offspring colonies share the same vector colour.

Figure 1

Figure 2

Figure 3

Figure 4