No conclusive evidence of group selection in spiders

Lena Grinsted[*1], Trine Bilde[2], James D. J. Gilbert[1,3]

[*]Corresponding author. Email: l.grinsted@sussex.ac.uk Phone: +44 1273 872757

[1]School of Life Sciences, University of Sussex, John Maynard Smith Building, Brighton BN1 9QG, UK

[2]Department of Bioscience, Aarhus University, Ny Munkegade 116, building 1540, 8000 Aarhus C, Denmark

[3]School of Biological, Biomedical and Environmental Sciences, University of Hull, Cottingham Road, Hull HU6 7RX, UK

Any field study showing convincing evidence of group selection [GS] would be a significant contribution to the field of evolutionary biology. Pruitt and Goodnight (2014)[1] [PG14] claim to provide such evidence in a 14-18 months field experiment on spiders. However, we contend that critical flaws in their predictions, assumptions, methods and interpretations undermine this claim. The data presented are unreliable and are equally consistent with GS and individual-level selection [ILS]; thus, the authors cannot credibly conclude that GS has produced the observed patterns.

1. Predictions: No critical test of GS

Evaluating GS involves, at a bare minimum, estimating and comparing both individual and group fitness, as stated by previous reviews[2,3] and performed by other studies[4,5]. Yet PG14 did not estimate individual fitness, and so cannot evaluate the relative importance of GS compared to ILS. The chosen species, *Anelosimus studiosus*, is solitary, rarely forms groups[6], and shows no evidence of reproductive restraint or skew within groups[7]. Thus, individual and group fitness are not expected to conflict and are generally confounded, emphasising how crucial it is, firstly, to formulate predictions capable of distinguishing ILS and GS explanations and, secondly, to estimate individual fitness.

Both of PG14's predictions could follow equally well from ILS as from GS: Prediction 1) "Compositions [i.e. within-group phenotypic frequencies] that approximate the normal mixtures that characterize each site will enjoy greater success". Merely demonstrating differential survival of groups does not allow the authors to distinguish successful groups from groups of successful individuals. PG14's "group trait" is a group size dependent behavioural polymorphism. Experimental changes in this "group trait" (i.e. manipulating group size and phenotype frequency) may directly affect within-group individual fitness just as well as whole-group fitness[8-11]. Specifically, creating experimental groups that deviate from locally stable polymorphisms may reduce mean *individual* fitness, rendering group extinction more likely. The prediction of differential group extinction can therefore result from ILS just as plausibly as from GS. Similarly ambiguous is Prediction 2) "Colonies should only be able to adaptively hone compositions when composed of native individuals". If 'Native colonies' can "adaptively" change phenotype frequencies over time, this may occur via several mechanisms, as PG14 mention (plasticity, phenotype-biased dispersal, etc.). Yet, any of these mechanisms may evolve by ILS, a possibility ignored by PG14.

2. Assumptions: Unreliable selection pressure

PG14's conclusions rest upon the assumption that 'naturally-occurring mixtures' (i.e. field phenotypic frequencies, PG14's Fig. 1a) represent consistent selection pressures across years. Yet, the years of measurement were patchy (2007-2014), differed among sites and often did not overlap (Table 1). Indeed, PG14 sampled significantly different phenotypic mixtures and group sizes among years at each site (mixtures: $p = 1.1 \times 10^{-5}$; group size: $p = 7.8 \times 10^{-14}$; Fisher's combined *p*-value across separate Kruskal-Wallis tests for each site) but ignored this variation and pooled dissimilar data. In four out of nine samples taken at high resource sites, the selection pressure was no different from zero (non-significant correlations between mixtures and group size; separate lm for each year at each site), and in two out of the three low resource sites the relationship between mixture and group size differed significantly from year to year (Norris Dam: $p = 0.0074$; Don Carter: $p = 0.017$; lm: interaction between log(group size) and year on phenotypic mixtures). Moreover, half of the sites had not been assessed for four to six years. These measurements cannot be assumed to represent consistent, current selection pressures.

3. Methods: Unreliable group trait

PG14 infer GS by comparing parental [P] and grand-offspring [F2] generations based on parental traits alone: they compare P-phenotypic compositions of P-colonies with P- (not F2) compositions of surviving F2-colonies (PG14's Fig. 1c is identical to Fig. 1b, minus extinct colonies: F2-colonies are depicted with their grandparents' compositions). It stretches credibility to assume that past compositions are visible to selection but present compositions are not. Indeed, 'Foreign colonies' changed to display F2-compositions in a pattern *opposite* to the assumed selection pressure (Fig. 1; PG14's Fig. 2). These changes mean that F1-compositions presumably also differed from P-compositions (and were visible to selection during that generation) but F1 was not assessed (Fig. 1).

Changes *within* generations were also not considered (Fig. 1). Around egg hatching[12], colonies peak in size, after which mortality and dispersal decrease colony size. Phenotypic composition cannot be considered a stable "group trait" when its proposed selective advantage is a function of group size, and group size changes nonlinearly over time. Compounding this, compositions of P and F2 were apparently measured at different developmental stages (Fig. 1): a serious flaw, as individual phenotypes are affected by reproductive status[13]. PG14 compared sexually mature females (P) with grandchildren (F2) that presumably were juvenile, mixed-gender, and receiving maternal care (Fig. 1; mothers die off in October[12]; juvenile spiders are unsexable).

4. Interpretations: No evidence of evolution

None of their findings supports PG14's puzzling claim to have demonstrated a "marked evolutionary response to GS". Rather, after two generations, surviving 'Foreign colonies' *failed* to change phenotypic compositions in site-appropriate ways (instead changing to express compositions appropriate for their original site), suggesting a lack of genetic change over the experiment (Fig. 1). PG14 argue that this constitutes evidence that mechanisms for adjusting compositions are locally adapted due to historical GS, but provide no justification for this claim: while they provide data suggesting phenotypes themselves may be partially heritable, there is no evidence that this "adjustment mechanism" has undergone genetic change and, again, no attempt

to reject ILS as an explanation. ‘Foreign colonies’ may revert to their native phenotype compositions without genetic change, for example due to persistent maternal or epigenetic effects, either of which may respond to ILS or GS. The evolutionary mechanisms shaping population-level differences, whether in phenotype frequencies or the means by which these change over time, simply have not been addressed in this paper.

We stress that we would welcome any field study demonstrating that GS causes genetic change over generations in ways inconsistent with ILS. Given recent high-profile exchanges over the relative importance of GS[14,15], such a paper would be a significant contribution to the field. Unfortunately, PG14 is not such a study.

Figures and tables

Table 1. Collection years for each field site. Data from years marked with 'X' were pooled within sites to create the naturally-occurring mixtures (Fig. 1a in PG14) used to infer site-specific selection pressures. Grey shaded years indicate the years during which the experiment took place.

		Collection Years							
	Site	2007	2008	2009	2010	2011	2012	2013	2014
High Resource Sites	Melton Hill	X	X	X		X	X		
	Little River	X	X						
	Moccasin Creek							X	X
Low Resource Sites	Norris Dam	X	X	X		X	X		
	Clinch River		X		X				
	Don Carter			X	X				

Figure 1

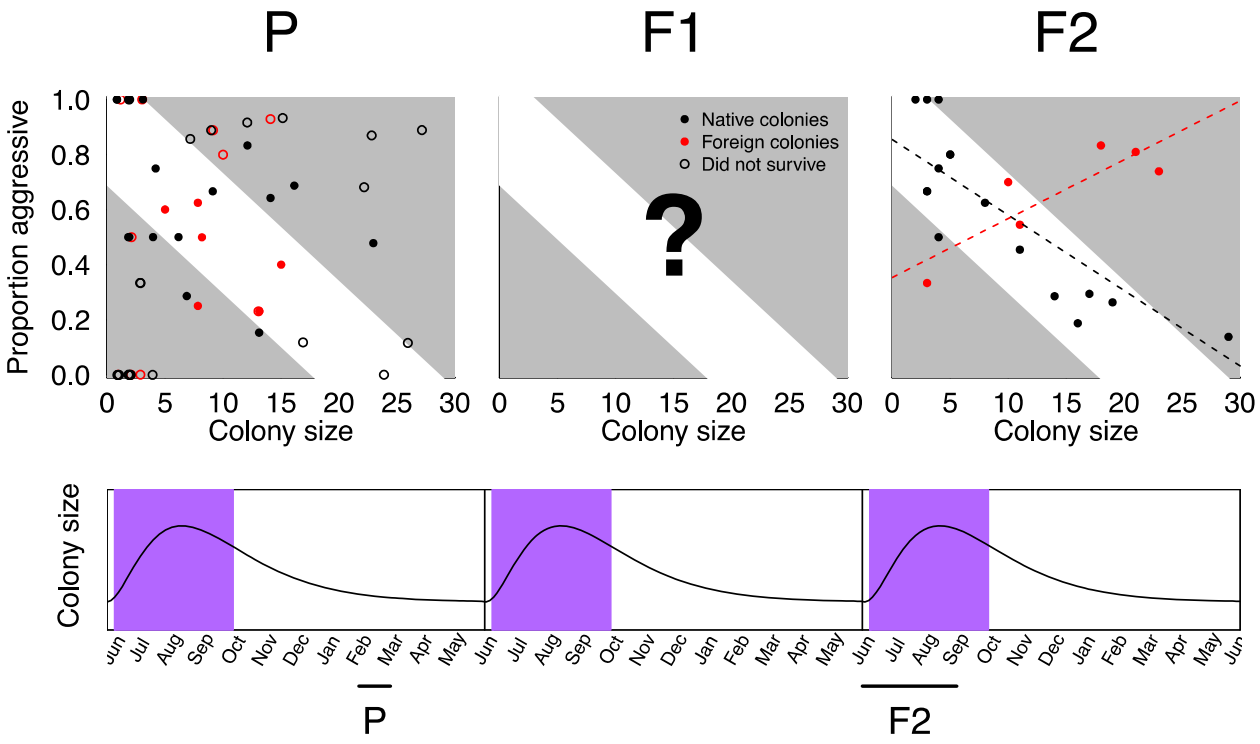


Figure 1: Overview of methods and results from PG14

Upper panel: Distribution of experimental colonies placed in all of six field sites in the parental [P] generation; the missing information of the next generation [F1]; and the distribution of final compositions of the grand-offspring generation [F2]. Phenotypic compositions, i.e. proportions of aggressive individuals in each colony, are plotted against colony sizes. We present data only from one low-resource field site, 'Don Carter', to illustrate the setup. Black dots represent 'Native colonies' (created with spiders collected at Don Carter); red dots represent 'Foreign colonies' (spiders collected at high-resource field site 'Moccasin Creek'). Full circles of both colours in P are colonies that were still alive in F2 (equal to PG14's Fig. 1c); empty circles are colonies that had gone extinct by F2. The white band represents the proposed selection pressure at that field site: a regression line fitted on phenotypic compositions and colony sizes of 'naturally-occurring' colonies at Don Carter (based here on colonies of sizes up to 30; its thickness chosen arbitrarily). Dotted lines in F2 represent regressions of the final F2-compositions of the surviving colonies: 'Native' (black) versus 'Foreign' (red). Surviving

colonies had P-compositions close to the white 'selection band', but F2-compositions differed according to site of origin: 'Native' F2-compositions were close to the selection band while 'Foreign' F2-compositions followed a positive regression, dissimilar to the selection band. Setup and results were similar in the two additional low-resource sites while the three high-resource sites showed opposite trends (i.e. selection bands were positive regression lines while the 'Foreign' F2-regressions showed negative correlations). *Note that although 'Foreign colonies' end up opposite to the proposed selection pressure, PG14 still conclude that their proposed selection pressure was supported.* **Lower panel** shows how the size of an *A. studiosus* colony is expected to vary within years with a peak around egg hatching during summer. The period of maternal care is marked in purple. Black lines marked with 'P' and 'F2' indicate at which point in the life cycle PG14 performed the behavioral assays to determine the phenotypic compositions of colonies. *Note that group sizes and phenotypic compositions of P and F2 were measured at different points, comparing sexually mature females (P) with juvenile grandchildren (F2) during maternal care at a stage where offspring sex cannot be determined.*

REFERENCES

1 Pruitt, J. N. & Goodnight, C. J. Site-specific group selection drives locally adapted group compositions. *Nature* **514**, 359-362, doi:10.1038/Nature13811 (2014).

2 Eldakar, O. T. & Wilson, D. S. Eight Criticisms Not to Make About Group Selection. *Evolution* **65**, 1523-1526, doi:10.1111/J.1558-5646.2011.01290.X (2011).

3 Marshall, J. A. R. Group selection and kin selection: formally equivalent approaches. *Trends Ecol Evol* **26**, 325-332, doi:10.1016/J.Tree.2011.04.008 (2011).

4 Rainey, P. B. & Rainey, K. Evolution of cooperation and conflict in experimental bacterial populations. *Nature* **425**, 72-74, doi:10.1038/Nature01906 (2003).

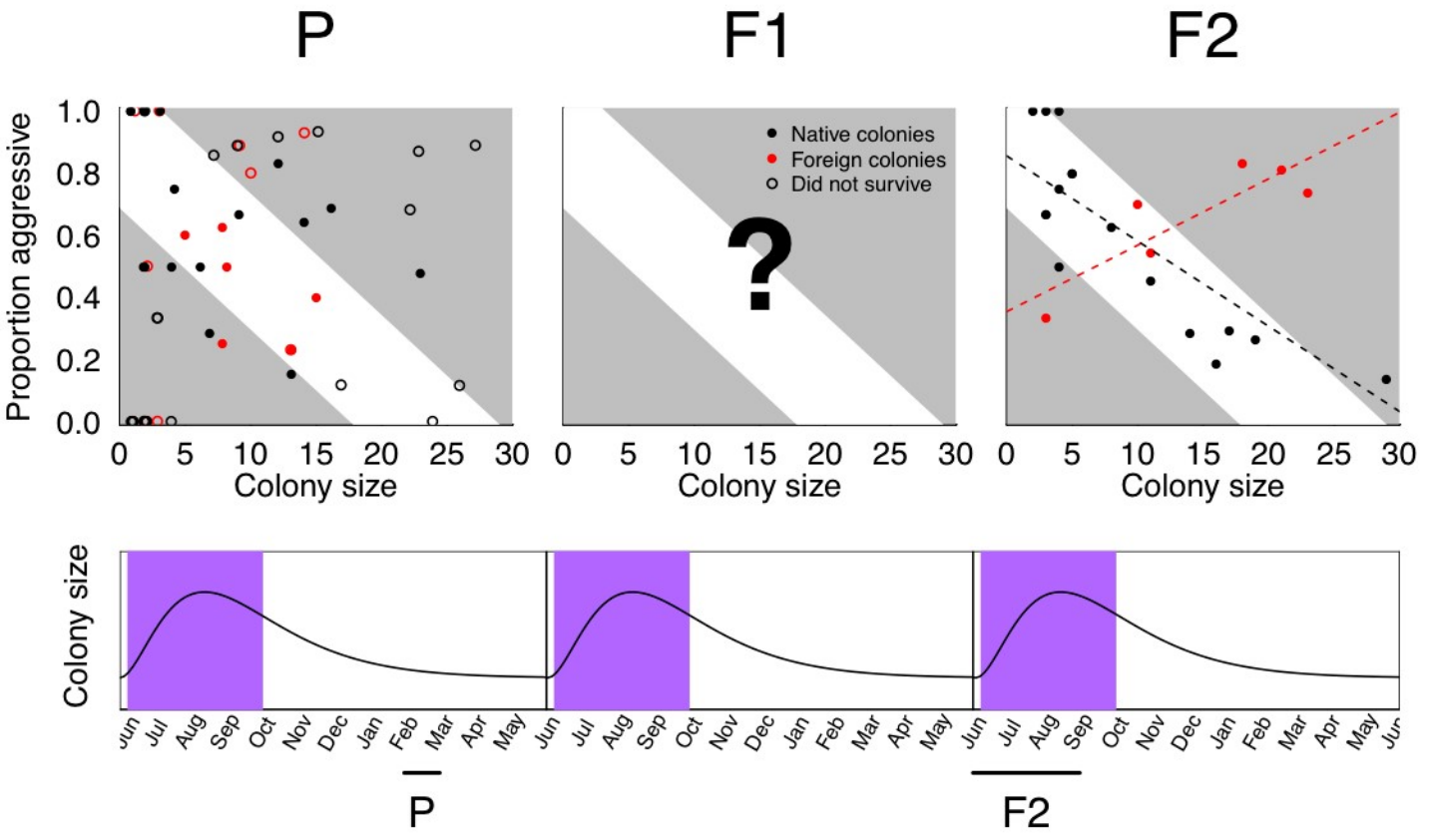
5 Kerr, B., Neuhauser, C., Bohannan, B. J. M. & Dean, A. M. Local migration promotes competitive restraint in a host-pathogen 'tragedy of the commons'. *Nature* **442**, 75-78, doi:10.1038/Nature04864 (2006).

6 Grinsted, L. & Bacon, J. P. Animal Behaviour: Task Differentiation by Personality in Spider Groups. *Curr. Biol.* **24**, R749-R751, doi:10.1016/J.Cub.2014.07.008 (2014).

7 Jones, T. C., Pruitt, J. N. & Riechert, S. E. Reproductive success in a socially polymorphic spider: social individuals experience depressed reproductive success in isolation. *Ecological Entomology* **35**, 684-690, doi:10.1111/j.1365-2311.2010.01227.x (2010).

8 Moczek, A. P., Hunt, J., Emlen, D. J. & Simmons, L. W. Threshold evolution in exotic populations of a polyphenic beetle. *Evolutionary Ecology Research* **4**, 587-601 (2002).

9 Culumber, Z. W., Bautista-Hernandez, C. E., Monks, S., Arias-Rodriguez, L. & Tobler, M. Variation in Melanism and Female Preference in Proximate but Ecologically Distinct Environments. *Ethology* **120**, 1090-1100, doi:10.1111/Eth.12282 (2014).
10 Weider, L. J. *et al.* Genotype x environment interactions, stoichiometric food quality effects, and clonal coexistence in Daphnia pulex. *Oecologia* **143**, 537-547, doi:10.1007/S00442-005-0003-X (2005).
11 Fitzpatrick, M. J., Feder, E., Rowe, L. & Sokolowski, M. B. Maintaining a behaviour polymorphism by frequency-dependent selection on a single gene. *Nature* **447**, 210-U215, doi:Doi 10.1038/Nature05764 (2007).
12 Furey, F. E. Two cooperatively social populations of the theridiid spider Anelosimus studiosus in a temperate region. *Anim. Behav.* **55**, 727-735, doi:10.1006/anbe.1997.0648 (1998).
13 Watts, J. C., Ross, C. R. & Jones, T. C. Diel and life-history characteristics of personality: consistency versus flexibility in relation to ecological change. *Anim. Behav.* **101**, 43-49, doi:http://dx.doi.org/10.1016/j.anbehav.2014.12.020 (2015).
14 Abbot, P. *et al.* Inclusive fitness theory and eusociality. *Nature* **471**, E1-E4, doi:10.1038/Nature09831 (2011).
15 West, S. A., Griffin, A. S. & Gardner, A. Social semantics: how useful has group selection been? *J Evol Biol* **21**, 374-385, doi:10.1111/J.1420-9101.2007.01458.X (2007).


P
F1
F2
Proportion aggressive
1.0
0.8
0.6
0.4
0.2
0.0
0 5 10 15 20 25 30
Colony size
Native colonies
Foreign colonies
Did not survive
?
Colony size
Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May
P
F2