Evolutionary Ecology Research, 2010, **12**: 169–188

Bayesian analysis of split sex ratios: methods and application to the ant *Aphaenogaster rudis*

David Lubertazzi and Eldridge S. Adams

Department of Ecology and Evolutionary Biology,
The University of Connecticut, Storrs, Connecticut, USA

ABSTRACT

Objective: We develop the use of Bayesian logistic regression models for the analysis of sex ratio data for organisms with variable brood sizes.

Background: Ant populations often exhibit 'split sex ratios', whereby most colonies produce strongly male-biased or female-biased broods. Furthermore, colony-level sex ratios may depend on brood size or other non-normally distributed covariates. The lack of fit of such data to the assumptions of traditional statistical models can produce inaccurate estimates of population sex ratios.

Methods: Using the freeware program OpenBUGS, we demonstrate the Bayesian approach by analysing the effects of food supplements on sex ratio expression in the ant *Aphaenogaster rudis*. Posterior sampling facilitates model checking and improves the estimation of population-level sex ratios for non-standard models.

Results: Bayesian logistic regression modelling with a random effects term is effective for analysing ant sex ratio data. The distribution of sex ratios among colonies of *A. rudis* was bimodal with larger broods tending to be more male-biased. Posterior sampling confirmed that a random-effects model is consistent with the observed split sex ratios. The population-level sex ratio was estimated taking into account that brood size was greater for colonies with male-biased sex ratios. The investment sex ratio was 0.80 (proportion female investment: 95% credibility interval = 0.69 to 0.88), consistent with worker control of the sex ratio. We found no evidence to suggest that experimental food supplements affected brood size or sex ratio expression.

Keywords: ant sex ratios, *Aphaenogaster rudis*, Bayesian logistic regression models, hierarchical models, split sex ratios.

INTRODUCTION

The application of sex ratio theory to social Hymenoptera established ants as an important focal group for sex ratio research (Craig, 1980; Nonacs, 1986b; Frank, 1987; Boomsma *et al.*, 1995; Hardy, 2002). Trivers and Hare (1976) showed that queens and workers in the same colony are in conflict

---

Correspondence: D. Lubertazzi, Museum of Comparative Zoology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA. e-mail: lubertazzi@gmail.org

Consult the copyright statement on the inside front cover for non-commercial copying policies.

© 2010 David Lubertazzi

over their preferred investment in male and female reproductive offspring, setting the stage for many subsequent studies examining the patterns and mechanisms of sex ratio expression in ants (e.g. Pearcy and Aron, 2006; Sorvari and Hakkarainen, 2007; Wiernasz and Cole, 2009; reviewed by Bourke and Franks, 1995; Crozier and Pamilo, 1996). Two types of statistical analyses are particularly important in this work: comparison of predicted to observed population sex ratios, and estimation of the dependence of colony sex ratios on intrinsic or extrinsic factors, including experimental treatments. However, both types of analysis are complicated by the high degree of variation in colony sex ratios found in many ant populations.

Ants often exhibit extreme colony-level sex ratio biases such that many colonies produce reproductive offspring that are mostly male or mostly female (Grafen, 1986; Nonacs, 1986a). These populations are said to express 'split sex ratios' (see Fig. 1). In these species, applying traditional statistical methods yields confidence intervals for the population average sex ratio that are often so wide as to be uninformative (see Bourke and Franks, 1995, pp. 160–161). Furthermore, standard methods to account for high between-colony variation in sex ratios may yield biased estimates of the relationship between the sex ratio and other variables, such as experimental treatments, colony size, and food availability. Analytical approaches are needed that can accurately describe sex ratio expression despite the wide variation among colonies.

In this paper, we describe the use of Bayesian logistic regression models for analysing sex ratio data, and we show how to estimate these models using the freeware program OpenBUGS (Thomas *et al.*, 2006). This approach builds on maximum likelihood logistic regression methods previously used to analyse sex ratio data (Schall, 1991; Kruuk *et al.*, 1999; Wilson and Hardy, 2002). Bayesian regressions are widely used in some areas of ecology (McCarthy, 2007), but to our knowledge have not previously been applied to analysis of ant sex ratios. We explain the methods in some detail and illustrate their use through an extended example.

Logistic regressions are a form of generalized linear model (GLM) with a binomial error term and a logit link function (Pampel, 2000). These models can be appropriate for proportional data bounded by 0 and 1 and can be used to test whether specific variables, such as colony size or colony investment, influence sex ratio expression. However, logistic regressions

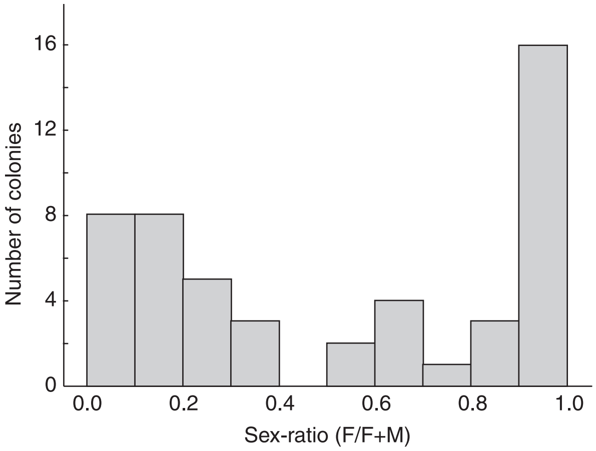

Fig. 1. Colony-level numeric sex ratios. The distribution illustrates a bimodal pattern which is typical of split sex ratios produced by many ant species.

typically require modification when there is substantial variation in brood size among groups (Krackow and Tkadlec, 2001). Without such modification, the observed degree of variation across groups (colonies or broods) typically exceeds the predicted degree of variation (Schall, 1991), a phenomenon known as 'overdispersion'.

Bayesian methods offer a number of pragmatic advantages in the analysis of sex ratios, which become more pronounced as the complexity of the model increases. Implementation of Bayesian models typically relies on Markov chain Monte Carlo (MCMC) methods (Gill, 2002; Gelman *et al.*, 2004). The availability of software, such as OpenBUGS, that automatically implements MCMC techniques has greatly increased the accessibility of Bayesian methods and, by extension, the ability to fit non-standard models. In this paper, we exploit the computational capacity of OpenBUGS to incorporate two modifications to logistic regression analysis of colony sex ratios. First, a random error term is included at the level of the colony to account for high variation in sex ratio among groups (see also Schall, 1991; Krackow and Tkadlec, 2001). Second, the colony sex ratio may be dependent on a covariate that itself has an unknown distribution. For example, the sex ratio may vary with brood size – the number of reproductive offspring produced by the colony. Estimating the population sex ratio then requires modelling the joint distribution of the covariate and the sex ratio. The necessary steps are readily implemented in OpenBUGS, yielding summaries that are easily interpreted.

We illustrate this approach by the analysis of sex ratio expression in a naturally occurring population of the ant *Aphaenogaster rudis* Enzmann, a species that exhibits split sex ratios. We tested experimentally whether colony sex ratios were affected by food availability, as in some ant populations (Deslippe and Savolainen, 1995; Morales and Heithaus, 1998; Aron *et al.*, 2001; Bono and Herbers, 2003). We use Bayesian methods to address three questions. First, can adding group-level random effects to a logistic regression model account for the split sex ratios seen in some ant populations? Second, does colony sex ratio depend on covariates, such as brood size or experimental additions of food? Third, given a non-normal distribution of colony sex ratios, or dependence on a covariate, how can one estimate credibility intervals for the population sex ratio?

THE BAYESIAN MODEL

The formulation of our analytical approach owes much to two studies that advocate the use of logistic regression to analyse sex ratio data (Boomsma and Nachman, 2002; Wilson and Hardy, 2002). Many of the general points we make concerning logistic regression are presented in greater detail in these two sources. To keep the description of the model general, we refer here to male and female offspring. However, as in most studies of sex ratio expression in ants and other social Hymenoptera, only the reproductive offspring (queens and males) are counted. Non-reproductive female workers are not considered when quantifying sex ratios (Trivers and Hare, 1976).

A logistic regression models the number of female offspring produced in the *i*th colony, f_i, as binomially distributed conditional on the total number of offspring produced, s_i, and a colony-specific proportion of females, p_i:

$$f_i \sim \text{binomial}(p_i, s_i) \tag{1}$$

The logit of p_i, which is equal to $\ln(p_i/(1 - p_i))$, is assumed to have a linear relationship with the model's explanatory variables. Thus

$$\text{logit}(p_i) = \alpha + \beta_1 * x_1 + \beta_2 * x_2 + \ldots \beta_z * x_i \tag{2}$$

where α is the intercept, $x_1, x_2, \ldots x_i$ are possible explanatory variables, and $\beta_1, \beta_2, \ldots \beta_z$ are the variable-specific coefficients.

Ant colony sex ratios are typically overdispersed compared with the predictions of logistic regressions (Boomsma and Nachman, 2002). This mismatch between the model and the data can produce biased estimates of the effects of explanatory variables and inaccurate 95% confidence or credibility intervals (Williams, 1982). A possible remedy is to estimate a scaling factor that adjusts the variance for samples drawn from a binomial distribution (McCullagh and Nelder, 1989; Boomsma and Nachman, 2002). However, this method is not recommended when the number of sexual offspring produced by different colonies (or individual litters, clutches, etc.) is highly variable (Krackow and Tkadlec, 2001). This suggests that the use of a conventional logistic regression GLM, with or without a scaling factor, may be inappropriate for the analysis of the split sex ratio data that are commonly produced by studies on ants.

Overdispered sex ratio data have been analysed with a generalized linear mixed model (GLMM) with a random error term (Williams, 1982; Schall, 1991; Kruuk *et al.*, 1999; Krackow and Tkadlec, 2001). Random error models explicitly account for the high variance of group-level sex ratio expression. A Bayesian hierarchical model with a group-level random effects term (Gelman *et al.*, 2004) is similar to a GLMM with a random error term. The random-effects model retains the assumption that the number of females in the *i*th colony is binomally distributed given p_i and s_i but allows the p_i values to be more variable than expected under a traditional logistic regression model. Expression (2) becomes

$$\text{logit}(p_i) = \alpha + \beta_1 * x_1 + \ldots + \beta_z * x_z + \varepsilon_i \tag{3}$$

where ε_i is normally distributed with a mean of 0 and an unknown standard deviation, σ.

The product of a Bayesian analysis is a posterior probability distribution ('posterior') for the quantities of interest, such as regression coefficients. A Bayesian posterior is found by combining a likelihood function (defined by the logistic regression model) with prior probability distributions ('priors') assigned to each parameter. Priors are either based on previous theoretical and empirical findings (informed priors) or constructed to be vague (flat or uninformative priors) when alternative values are thought to be nearly equally likely.

The MCMC methods estimate model parameters by repeatedly sampling from the posterior distribution, as determined by the likelihood function, priors, and data. Credibility intervals describe the uncertainty of posterior estimates. For example, the 95% credibility interval contains the true value of an estimated quantity with a probability equal to 0.95, given the model and the data. The 95% credibility interval can be estimated to any degree of precision by MCMC sampling.

Possible colony-level explanatory variables in our field study (described in the next section) include the number of reproductive offspring produced (s, hereafter referred to as 'brood size'), site (L or N), treatment (food supplement or control), queen presence (present or absent), and the number of workers. We evaluated models with different combinations of the explanatory variables and the random error term (Table 1). Because we had little prior information concerning the values of the model parameters (α and β_i), each was assigned a vague normal prior with a large variance.

Computations were carried out in OpenBUGS 3.0.3 (Thomas *et al.*, 2006) (for the sex ratio analysis code, see online Appendix A at www.evolutionary-ecology.com/data/2506.pdf). The goals of our analysis were to estimate the effects of the explanatory variables on colony sex ratio expression and to estimate the population sex ratio. Modelling runs generated estimates of each model's parameters, and 95% credibility intervals for these estimates.

Table 1. Parameter estimates (means and 95% credibility intervals) for alternative logistic regression models

Model	α	*S*	QU	SITE	TRE	WOR
α, S, QU, SITE, TRE, WOR, E	1.64 (−1.48 to 4.89)	−0.03 (−0.06 to −0.01)	−0.65 (−4.03 to 2.62)	0.99 (−1.73 to 3.94)	−2.05 (−4.80 to 0.45)	0 (0 to 0.01)
α, TRE, E	1.34 (−0.18 to 2.99)				−1.98 (−4.58 to 0.51)	
α, WOR, E	0.12 (−2.90 to 3.14)					0 (0 to 0.01)
α, E	0.58 (−0.64 to 1.89)					
α, S, TRE, WOR, E	1.31 (−1.42 to 4.08)	−0.03 (−0.05 to −0.01)			−2.10 (−4.56 to 0.15)	0 (0 to 0.01)
α, S, WOR, E	0.84 (−1.89 to 3.52)	−0.03 (−0.05 to −0.01)				0 (0 to 0.00)
α, S, TRE, E	2.88 (1.08 to 4.91)	−0.02 (−0.04 to −0.01)			−1.79 (−4.13 to 0.42)	0 (0 to 0.00)
α, S, E	2.21 (0.57 to 3.98)	−0.03 (−0.04 to −0.01)				
α, S, QU, SITE, TRE, WOR	0.77 (0.54 to 1.00)	−0.02 (−0.02 to −0.01)	−0.28 (−0.49 to −0.06)	−0.08 (−0.28 to 0.12)	−0.39 (−0.57 to −0.21)	0 (0 to 0.00)
α, S, TRE, WOR	0.67 (0.46 to 0.89)	−0.02 (−0.02 to −0.01)			−0.45 (−0.62 to −0.28)	0 (0 to 0.00)
α, S	0.80 (0.65 to 0.96)	−0.01 (−0.02 to −0.01)				
α, WOR	−0.12 (−0.30 to 0.07)					0 (0 to 0.00)
α, TRE	−0.50 (−0.59 to −0.40)				−0.42 (−0.58 to −0.27)	

Note: Models that include the error term (E) are all of the form $\mathrm{logit}(p_i) = \alpha + \beta_1 * x_1 + \beta_2 * x_2 + \ldots + \varepsilon_i$, and those without the error term exclude ε_i.

The models are arranged according from the lowest to highest DIC scores. α = alpha, S = number of reproductive offspring, QU = queen present in nest, SITE = L or N, TRE = treatment group, WOR = number of workers in a colony.

We used posterior predictive checks to determine whether inclusion of random effects improves the ability of the model to account for the high frequency of colonies with extreme sex ratios. Predictive checks entail comparing observed data with data simulated under a particular model to detect discrepancies (Gelman *et al.*, 2004). Using the code shown in online Appendix B (www.evolutionary-ecology.com/data/2506.pdf), we simulated data under two models, differing only in whether colony-level random effects were included. The sample size (number of colonies) and brood sizes matched those of the observed data. For each model, 50,000 simulated data sets were generated using different draws from the posterior distributions of the model coefficients. The observed number of same-sex colonies was compared with the mean and the 95% credibility intervals of values generated under each model.

To compare alternative models, differing in terms of which explanatory variables are included, we used the deviance information criterion (DIC) (Spiegelhalter *et al.*, 2002), a Bayesian version of Akaike's information criterion (Akaike, 1973). Models are scored by their fit to the data (the deviance of the model) and penalized for the effective number of parameters (pD) estimated from the data. Spiegelhalter *et al.* (2002) proposed the DIC, which is equal to the sum of $\bar{\mathrm{D}}$ (the mean of the posterior deviance) and pD, as a criterion for model selection. Models with lower DIC scores are more strongly supported than models with higher scores. If the difference in DIC values for two models (ΔDIC) is greater than 2 or 3, the model with the larger DIC has less support, whereas if ΔDIC is greater than 10, the model with the larger DIC has essentially no support (McCarthy, 2007).

During the analysis of the data, we encountered conflicting indications concerning possible effects of brood size on colony sex ratio. The 95% credibility interval for the coefficient for brood size was consistently entirely negative (Table 1). This suggests that the sex ratio is more female-biased when broods are small, yet comparisons of DIC values did not support including brood size as an explanatory variable. We therefore conducted a numerical experiment, constructing 100 data sets in which there is a known negative relationship between brood size and sex ratio of the magnitude suggested by the data. Data sets were generated in R (R Development Core Team, 2009) using the model $\mathrm{logit}(p_i) = \alpha + \beta^*(\text{brood size})_1 + \varepsilon_i$, with the model parameters α, β, and σ set to the posterior means produced by analysis of the data. For each simulated data set, OpenBUGS was used to estimate DIC values for two random effects models, with and without brood size as an explanatory variable. This allowed us to evaluate whether comparison of DIC values provides an accurate method for model choice for this type of analysis. Each model was run for 60,000 iterations. A preconvergence 'burn-in' of 10,000 iterations was discarded. We examined how often two different assessment criteria, the DIC values and the 95% credibility intervals, correctly indicated that the sex ratio varies with brood size.

Estimating the population sex ratio, our second analytical goal, can be accomplished by sampling from the posterior probability distribution of the selected model. A working example, presented in a later section, is used to explain the details of this estimation procedure.

An important advantage of the Bayesian approach is the ease with which uncertainty in the population sex ratio can be summarized for non-standard models without making assumptions about the shapes of distributions. Thus, accurate 95% credibility intervals can be produced despite the presence of split sex ratios, highly unequal brood sizes, or dependence of the sex ratio on colony attributes. Calculating scaling factors, placing distributional restrictions on variance estimates, and making *post hoc* adjustments to parameter estimates are not necessary. By contrast, it is problematic to apply a traditional logistic model to ant sex ratio data, because the colony variances do not approximate a normal distribution (Schall, 1991; Boomsma and Nachman, 2002).

THE ANT EXPERIMENT

Aphaenogaster rudis is one of the most common ant species in the hardwood forests of southern New England. Mature colonies of this ant contain a single reproductive queen and from a few hundred to over a thousand workers (Headley, 1949; Talbot, 1951). Winged *A. rudis* reproductives are produced annually, eclose as adults in July, and leave the nest to mate in mid-August. Whole nests collected from late July up to the time mating occurs therefore contain all of a colony's reproductive offspring.

Sex ratio theory models the evolution of investment sex ratios, which are based on the amount of energy required to produce reproductive offspring. Investment ratios are not equivalent to numerical sex ratios if the cost of producing males and females is unequal. *Aphaenogaster rudis* males are smaller and cheaper to produce than females. Numerical sex ratios (F/F + M) for this species were converted to investment sex ratios according to the following formula:

$$I = 5.3 * F/[(5.3 * F) + M]$$

where I represents the proportional investment in female reproductives, F is the number of female reproductives (excluding the colony queen), and M is the number of males. The constant 5.3 is the production cost, c (see Boomsma, 1989; Bourke and Franks, 1995), as determined for *A. rudis* (Lubertazzi, 2006) where:

$$c = [(\text{average female dry weight})/(\text{average male dry weight})]^{0.7}$$

This adjusted production cost yields a more accurate estimate of reproductive investment than dry weight alone (Boomsma, 1989).

Two isozyme analyses (Crozier, 1973, 1974) and a microsatellite analysis (Lubertazzi, 2006) have shown that each *A. rudis* colony contains one singly-mated queen and that the workers do not reproduce. Under these conditions, hymenopteran sex ratio theory predicts an investment sex ratio of 3 : 1 for worker control and a 1 : 1 ratio for queen control (Trivers, 1974; Benford, 1978; Pamilo, 1982).

A previous study of *A. rudis* sex ratios provided support for the resource abundance hypothesis, which predicts a positive correlation between food availability and the investment sex ratio (Nonacs, 1986b; Crozier and Pamilo, 1996; Rosenheim *et al.*, 1996). Colonies given supplements of elaiosome-bearing seeds, a natural food resource (Heithaus, 1981), produced a greater number of reproductive offspring and had more female-biased sex ratios than control colonies (Morales and Heithaus, 1998). The treatment response was suggested to result from the quantity of food added rather than a particular nutrient contained in the elaiosomes (Bono and Heithaus, 2002).

We conducted a field experiment to test the influence of a food pulse on sex allocation and sex ratios in an eastern Connecticut *A. rudis* population. We divided a set of field colonies into two treatments: a control group, whose diet was not altered, and a group provided with supplementary protein for approximately 5 weeks in June and July of 2004. The resource abundance hypothesis predicts that food-supplemented colonies will produce more reproductive offspring and more female-biased sex ratios than control colonies.

In November 2003, 117 wooden nests were placed on the forest floor in the Mohegan State Forest (41°39′52″N, 72°5′8″W). Each nest was composed of two rectangular pine boards held together by screws. One of the two boards contained a hollowed-out area with a small external opening that the ants could use as a nesting chamber. *Aphaenogaster rudis* often nest in wood on the forest floor and readily use these artificial structures. In the spring, colonies move out of their underground overwintering nests and can find and move into the artificial nests. Additional details about the construction and use of these nests can be found in Lubertazzi (2006).

The study area was a secondary forest dominated by red maples (*Acer rubrum*) interspersed with oaks (*Quercus* spp.), conifers (*Tsuga canadensis* and *Pinus strobus*), and other less abundant hardwood tree species. The forest is situated on gently undulating terrain. Nests were positioned along nine transects located in two sites. Site L, containing five transects, crossed and surrounded a small ridge. Site N, with four transects, included a ridge-to-swale slope and a portion of a flatter swale bottom. Transects varied in length from 50 to 110 m, were haphazardly placed within each site, and were no closer than 20 m to any neighbouring transect. The wooden nests were placed at 10-m intervals along each transect; thus, nests were separated by considerably more than the 2-m foraging range of *A. rudis* workers (Lubertazzi, 2006).

The wooden nests were visually inspected two or three times per week, beginning in mid-May 2004, to determine which nests were occupied. Seventy-two nests (62%) were in use when the feeding treatments began on 23 June. Every other occupied nest along each transect was designated to receive food supplements ($n = 36$), with the remaining 36 nests serving as controls. Five additional nests occupied by the time of the second feeding (June 28) were subsequently added to the control group ($n = 41$).

Supplemental food consisted of ~1.5 g of tuna presented in a small (1 × 9 cm) glass test tube. This represented a generous protein addition for *A. rudis* (Southerland, 1988) and was much greater than the quantity of food used in a feeding experiment by Morales and Heithaus (1998). Food tubes were placed within 10 cm of a colony nest entrance and were quickly discovered by foragers from the targeted nest. Food was provided twice weekly. The first food supplement occurred on 23 June and the final (tenth) food supplement took place on 27 July. To ensure that disturbance was similar for treatment and control colonies, every nest was visited on each feeding day.

Colonies were collected between 6 and 12 August. The nests were picked up from the forest floor, sealed in plastic bags, and brought to the laboratory. The numbers of males, females, and workers and the presence or absence of a queen were recorded for each colony (Table 2). Voucher samples have been deposited in the Museum of Comparative Zoology at Harvard University.

DATA ANALYSIS

At the end of the experiment, 65 nests remained occupied. Thirty control colonies and 19 food-supplemented colonies produced reproductive offspring (Table 2). The distribution of colony-level numeric sex ratios was bimodal; in other words, the population was characterized by split sex ratios (Fig. 1).

Logistic regressions were fit with and without colony-level random effects [compare expressions (2) and (3)]. Using the tools OpenBUGS provides to assess whether the Markov chains have converged to the posterior probability distribution, we consistently found:

Table 2. Colony attributes of the nests from each treatment group

Colony	Number of workers	Queen	Females	Males
Food-supplemented colonies				
LB12	467	Y	9	14
LB10	397	Y	0	1
LA01	751	Y	7	19
LB07	501	Y	0	19
LA09	190	Y	61	0
NC01	838	Y	53	172
NC11	633	Y	0	27
NC08	354	Y	5	1
LE04	689	Y	31	0
LF17	386	N	12	53
LE06	701	Y	31	18
LF12	444	Y	0	69
NB04	675	Y	18	146
LE03	659	Y	24	112
LF08	348	Y	26	17
NA03	721	Y	15	0
NB09	618	Y	25	0
AA08	655	Y	0	51
BA02	1268	Y	26	141
LB05	317	Y	0	0
LB03	235	Y	0	0
LA05	145	Y	0	0
ND07	352	Y	0	0
ND02	231	Y	0	0
ND14	283	Y	0	0
ND10	599	Y	0	0
LF19	503	Y	0	0
LF14	292	Y	0	0
LZ02	528	Y	0	0
AA04	300	Y	0	0
Control colonies				
LB06	488	Y	17	0
LB09	543	Y	15	2
LA02	512	Y	46	36
LA04	190	N	9	2
LA06	282	N	41	3
LB01	645	Y	19	1
LB04	424	N	51	31
LA10	325	N	9	32
LA08	519	Y	13	0
NC14	89	N	1	0
ND05	432	N	45	0
ND08	427	Y	18	146
ND13	196	N	43	83

(*continued*)

Table 2.—(*continued*)

Colony	Number of workers	Queen	Females	Males
NC03	284	Y	8	51
LF18	303	N	32	14
LF15	445	Y	9	0
LE05	488	Y	11	123
LF09	225	N	15	35
LF10	595	Y	27	23
LF11	143	Y	0	1
LF13	430	Y	17	106
LF01	763	Y	12	3
LF06	605	Y	47	73
NB07	822	Y	0	46
NB05	854	Y	13	0
AA03	1117	Y	28	0
BA03	1022	Y	47	130
NA02	1043	Y	39	0
BA05	891	N	23	167
AA06	683	Y	20	0
LB11	25	N	0	0
LZ03	416	Y	0	0
LF04	338	Y	0	0
NA04	390	Y	0	0
BA01	920	Y	0	0

Note: The first letter of the colony ID indicates the site (L or N) and the Queen column indicates the presence (Y) or absence (N) of a queen.

quantile plots that were stabilizing by the end of the 10,000 step burn-in, trace plots showing convergence in posterior means when multiple chains were run from different initial values, and estimated Monte Carlo errors of less than 0.01%. Autocorrelation was detected in the posterior sampling results for some parameters in some models. These samples were rerun and thinned to varying degrees until the autocorrelation was removed, for example by retaining every other sample or every third sample. In sum, the posterior sampling behaved as expected for a properly characterized and coded model (Gill, 2002; Spiegelhalter *et al.*, 2002). Posterior estimates were obtained from 100,000 post burn-in samples.

The DIC values for models lacking the random error term were much higher (which equates to less support) than for models with the random error term (Table 3). Furthermore, whereas random effects models predicted frequencies of all-male or all-female colonies similar to the observed numbers, models lacking the random error term did not (Table 4).

For all models that included brood size as an explanatory variable, the 95% credibility interval for the coefficient was entirely negative (Table 1), showing that colony sex ratios are more male-biased in colonies that produce large broods (Fig. 2). The 95% credibility intervals of the coefficients for all other explanatory variables straddled zero (Table 1); thus, we cannot be confident that any of these variables, including the food supplements, affected colony sex ratio.

Table 3. DIC values for the models listed in Table 1

Model	DIC	$\bar{D}$	pD	ΔDIC
α, S, QU, SITE, TRE, WOR, E	199.04	160.47	38.6	0
α, TRE, E	199.79	160.67	39.1	0.75
α, WOR, E	199.82	160.57	39.2	0.78
α, E	200.56	161.14	39.4	1.52
α, S, TRE, WOR, E	200.86	161.86	39.0	1.82
α, S, WOR, E	201.68	162.33	39.3	2.64
α, S, TRE, E	201.96	162.59	39.4	2.92
α, S, E	202.61	163.00	39.6	3.57
α, S, QU, SITE, TRE, WOR	1073.87	1067.83	6.0	874.83
α, S, TRE, WOR	1076.54	1072.55	4.0	877.50
α, S	1108.46	1106.46	2.0	909.42
α, WOR	1511.97	1509.97	2.0	1312.93
α, TRE	1523.03	1521.03	2.0	1323.99

Note: Models that include the error term (E) are all of the form $\text{logit}(p_i) = \alpha + \beta_1 * x_1 + \beta_2 * x_2 + \ldots + \varepsilon_i$, and those without the error term exclude ε_i.

$\bar{D}$ = posterior mean of the deviance, pD = effective number of parameters in the model, ΔDIC = (model DIC − the lowest observed DIC score for all the models analysed), α = alpha, S = number of reproductive offspring, QU = queen present in nest, SITE = L or N, TRE = treatment group, WOR = number of workers.

Table 4. Observed number of colonies with all-male or all-female broods compared with values predicted under two models (means and 95% credibility intervals)

	Observed	Model 1	Model 2
All-male	7	5.7 (1 to 12)	0.93 (0 to 3)
All-female	13	11.4 (5 to 19)	2.2 (0 to 4)
Combined	20	17.1 (9 to 26)	3.15 (3 to 4)

Note: Model 1 was a logistic regression with random effects and with brood size as an explanatory variable; Model 2 was identical but omitted random effects.

Compared with the simplest random effects model, none of the explanatory variables led to a substantially different DIC value when added to the model. Furthermore, while several of these models had DIC values within two units of the lowest DIC, this was primarily because model complexity (pD) was similar, rather than because including these variables improved the model fit to the data ($\bar{D}$; Table 3). Therefore, in choosing the best approximating model for this data set, comparison of DICs provides little or no support for including any of these variables.

Whereas the 95% credibility intervals consistently indicated a negative effect of brood size on colony sex ratio, DIC comparisons did not support including brood size as an explanatory variable. To explore this apparent discrepancy, we generated and analysed 100 simulated data sets incorporating a negative effect of brood size on colony sex ratio, as

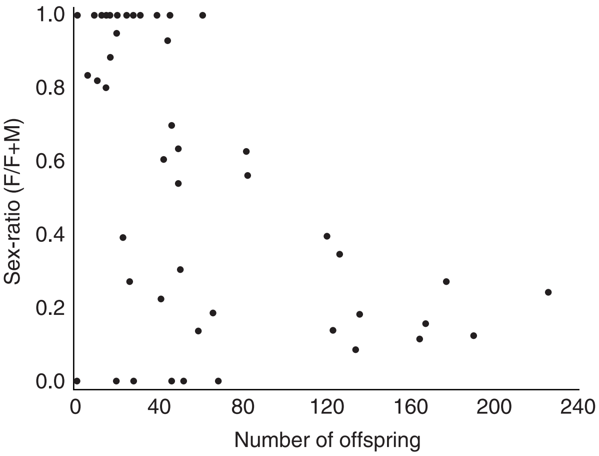

Fig. 2. Scatterplot of the number of reproductive offspring and the numeric sex ratio for each colony. Colonies that produced larger broods invested proportionally more in males than colonies producing fewer reproductive offspring.

described above. Using posterior means fitted from the data (Table 1), the model used for simulation was

$$\text{logit}(p_i) = \alpha + \beta * S_i + \varepsilon$$

where S_i is the brood size, $\alpha = 2.21$, $\beta = -0.025$, and the standard deviation for the random errors was 3.55. Each data set was analysed in OpenBUGS under two models: with and without brood size as an explanatory variable. When brood size was included, the estimated value of coefficient β averaged −0.026, which was very close to the true value. The 95% credibility interval for β was entirely negative in 74 of 100 cases; thus, examination of the posterior distribution of β usually correctly reveals the negative relationship between brood size and sex ratio. The DIC value declined only 0.48 on average (maximum: 2.91) when brood size was included, too little to indicate that the model including brood size was better supported. Therefore, comparison of DIC values for random effects models was not a reliable way of determining whether brood size affects colony sex ratio.

ESTIMATING THE POPULATION SEX RATIO

The analysis presented above indicates that colony sex ratio is inversely related to brood size. We illustrate the estimation of the population sex ratio for a model incorporating this relationship, and for a simpler random effects model lacking dependence on brood size, which may be appropriate for data on other species or populations.

The simplest model is:

$$\text{logit}(p_i) = \alpha + \varepsilon_i$$

$$\varepsilon_i \sim \text{normal}(\mu, \sigma)$$

When adding brood size, the model becomes:

$$\text{logit}(p_i) = \alpha + \beta_S * S_i + \varepsilon_i$$

$$\varepsilon_i \sim \text{normal}(\mu, \sigma)$$

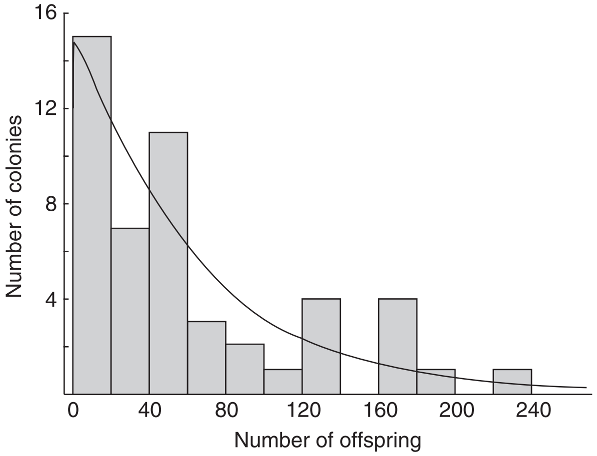

Fig. 3. Histogram of the total number of reproductive offspring produced by each colony. The line represents a gamma distribution fitted to the data.

Parameter estimates (posterior means) are α = 0.58 (95% CI: −0.64 to 1.89) in the simplest model and α = 2.2 (95% CI: 0.57 to 3.98) in the model with brood size, $\beta_S = -0.03$ (95% CI: −0.04 to −0.01), and $\sigma = 0.09$ (95% CI: 0.04 to 0.16). Because colony-level sex ratios depend on brood size, the distribution of brood size must be modelled to estimate accurately the population sex ratio and its 95% credibility interval. A frequency histogram of the number of reproductive offspring produced by each colony shows a positive skew that is approximated by a gamma distribution (Fig. 3). The gamma distribution ranges from 0 to infinity and is characterized by two parameters, which are termed the shape and the scale. Brood size must be a non-negative integer; thus, the distribution was modelled as:

$$s_i \sim \text{Poisson}(\lambda_i)$$

$$\lambda_i \sim \text{gamma}(\kappa, \theta)$$

where κ (shape) and θ (scale) were estimated from the data shown in Fig. 3, using the OpenBUGS code shown in online Appendix C (www.evolutionary-ecology.com/data/2506.pdf). With this model, one can estimate the population sex ratio by simulating sexual production in a population with a large number of colonies (e.g. 1000). The following steps were implemented by the OpenBUGS software (online Appendix C: www.evolutionary-ecology.com/data/2506.pdf).

For each of 100,000 simulations:

1. Values of the parameters α, β, κ, σ, and θ were drawn randomly from their posterior distributions.

2. Given these values, colony sex ratios were simulated as follows. For each colony i (1000 colonies in total):

(a) A number was drawn at random from gamma (κ, θ). This was taken to be λ_i, the Poisson mean.

(b) A number was drawn at random from Poisson (λ_i). The result was taken as the number of reproductive offspring produced.

(c) A value of ε was drawn at random from a normal distribution with mean equal to 0 and standard deviation equal to the estimated σ.
(d) The logit of p_i was calculated as

$$\alpha + \beta_S * S_i + \varepsilon_i$$

Therefore, the expected proportion of females is

$$p_i = \exp(\alpha + \beta_S * S_i + \varepsilon_i)/[1 + \exp(\alpha + \beta_S * S_i + \varepsilon_i)]$$

(e) The number of females (f_i) produced was drawn at random from a binomial distribution with the given s_i and p_i.
(f) The investment in females was calculated as $c * f_i$ (where c, the cost correction factor as described earlier, is 5.35).
(g) The total investment in reproductive offspring was calculated as:

$$\text{investment in females} + \text{number of males} = (c * f_i) + (s_i - f_i)$$

(the relative investment cost of each male is set as 1).

3. For the 1000 simulated colony values from step 2, sums were taken of the number of females, the number of reproductives, the investment in females, and the total investment in reproductives. The following values were stored:

(a) The mean number of reproductives produced by a reproductive colony.
(b) The numeric sex ratio, calculated as:

$$\sum f_i / \sum S_i$$

(c) The investment sex ratio, calculated as:

$$\sum c * f_i / \sum [(c * f_i) + (s_i - f_i)] * c$$

These steps were repeated for 100,000 steps of the Markov chain, following a burn-in of 15,000. The 95% credibility intervals were found by discarding the lowest 2.5% of values and the highest 2.5% values. The remaining lowest and highest values define the limits of the 95% credibility interval.

The estimated population investment sex ratio (posterior mean) for *A. rudis* was 0.86 (95% CI: 0.81 to 0.91; Table 5) for the model that included the random error term and 0.80 (95% CI: 0.69 to 0.88) for the model that included the error term and the variable 'brood size' (online Appendix C: www.evolutionary-ecology.com/data/2506.pdf).

Table 5. Brood size and population sex ratio estimates as determined by alternative logistic regression models

Model	ΔDIC	Mean brood size mean (95% CI)	Numeric sex ratio mean (95% CI)	Investment sex ratio mean (95% CI)
α, S	910	62.0 (46.3 to 82.5)	0.34 (0.26 to 0.42)	0.73 (0.65 to 0.79)
α, S, E	4.21	61.6 (46.0 to 82.4)	0.44 (0.30 to 0.58)	0.80 (0.69 to 0.88)
α, E	0.11	62.0 (46.4 to 83.0)	0.55 (0.44 to 0.66)	0.86 (0.81 to 0.91)

DISCUSSION

Hierarchical Bayesian models are well suited for analysing sex ratio data. Analysis of the *A. rudis* data illustrates that this statistical approach effectively manages three central goals of studies on ant sex ratios: (1) describing and analysing split sex ratios; (2) evaluating the dependence of colony sex ratios on intrinsic or extrinsic factors, including experimental treatments; and (3) comparing predicted and observed population sex ratios.

An important element of this method is the inclusion of a random effects term at the group or colony level. This can allow for more accurate modelling of non-normal distributions of colony sex ratios (Table 3), which are observed in many ant populations (e.g. split sex ratios; Fig. 1). Posterior predictive checks show that the observed number of colonies with extreme sex ratios (all-male or all-female broods) was surprising under models lacking random effects, but not under models that included random effects (Table 4). An alternative way to account for overdispersion is to estimate a scale parameter describing extra-binomial variation in the number of females produced by a colony. This is justified by assuming that the extra variation comes from non-independence of individual responses within groups, rather than variation in sex ratio among groups (Wilson and Hardy, 2002). Because the random effects model makes no such assumption and accepts instead that colony sex ratios truly are highly variable, we believe it is more biologically realistic.

When sex ratios vary widely among broods, including a random effects term can substantially alter the apparent influence of explanatory variables on the sex ratio. A model that includes random effects is not forced to account by other means for the extra-binomial variation created by split sex ratio expression. Omitting the random effects term causes a substantial reduction in estimates of coefficients quantifying the effects of explanatory variables (Table 1). Thus, failure to include a random effects term when it is appropriate may lead to underestimates of the effects of brood size, experimental treatments or other variables.

Comparison of observed and predicted population sex ratios is aided by accurate measures of uncertainty in population estimates. Sampling from the posterior distribution provides a convenient and flexible way to construct 95% credibility intervals for complex and non-standard models. For *A. rudis*, the population estimate of 0.80 proportional investment in female offspring is close to the worker optimum of 0.75, which is within the 95% credibility interval (0.69 to 0.88). However, the lower limit of the 95% credibility interval (0.69) is well above the 0.50 value expected under queen control. Thus, the data are consistent with worker control of sex ratio expression, but not with queen control, at least for the population and year studied.

In addition to illustrating the advantages of the modelling approach, analysis of the *A. rudis* sex ratio data produced several interesting results. Contrary to the predictions of the resource abundance hypothesis, we found no evidence that additional food influences brood size (Fig. 4) or colony sex ratios (Figs. 5 and 6). This result contrasts with *A. rudis* colonies in an Ohio forest where food supplementation, also done in the summer season but for a shorter duration, led to increased sex allocation and female-biased sex ratios (Morales and Heithaus, 1998). Nutrients found in elaiosome-bearing seeds may have a different influence on *A. rudis* sex ratios than the nutrients found in a protein-rich resource (tuna) or the sensitivity of sex ratio to food supply may depend on other variables. Clearly this problem requires further study.

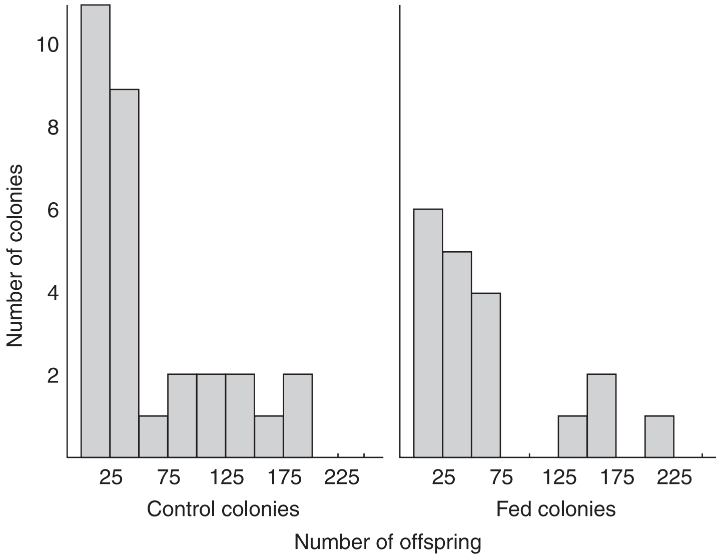

Fig. 4. The number of reproductive offspring produced by each colony in each treatment group. There was no evidence that supplemental food altered the sex ratio.

This result reinforces a growing body of empirical evidence showing that the influence of food resources on sexual production in ants is neither simple nor straightforward. These effects are known to vary from species to species, from population to population, and across space and time within a species (Buschinger and Pfeifer, 1988; Backus and Herbers, 1992; Herbers and Banschback, 1998; DeHeer *et al.*, 2001; Foitzik *et al.*, 2003).

Colonies with larger numbers of sexual offspring tended to have more male-biased sex ratios (Fig. 2). A similar relationship was found in the ant *Camponotus nipponicus* Wheeler, a species that also exists in populations of simple colonies; that is, colonies with one singly mated queen (Hasegawa, 1994). While larger broods were associated with greater investment in male offspring, the effect was not strong enough to account for split sex ratios. Instead, the importance of the random effects term suggests that unmeasured variables are responsible for producing the observed bimodal distribution of colony sex ratios.

Few studies (Hasegawa, 1994; Morales and Heithaus, 1998; Bono and Heithaus, 2002) have examined ant sex ratio expression or tested theory with an ant species that has such a simple social system. The lack of variation in breeding structure among colonies of *A. rudis* removes an important source of variation that could confound tests of the effects of other variables on sex ratio expression. With this study species, and using the analytical tools described in this paper, it is now easier to test how variables measured in experimental and observational studies influence sex ratio expression.

ACKNOWLEDGEMENTS

Charlie Henry, Peter Turchin, and Rob Colwell provided helpful suggestions that improved the manuscript. Support for this work was provided to David Lubertazzi from the graduate program in the Department of Ecology and Evolutionary Biology at the University of Connecticut.

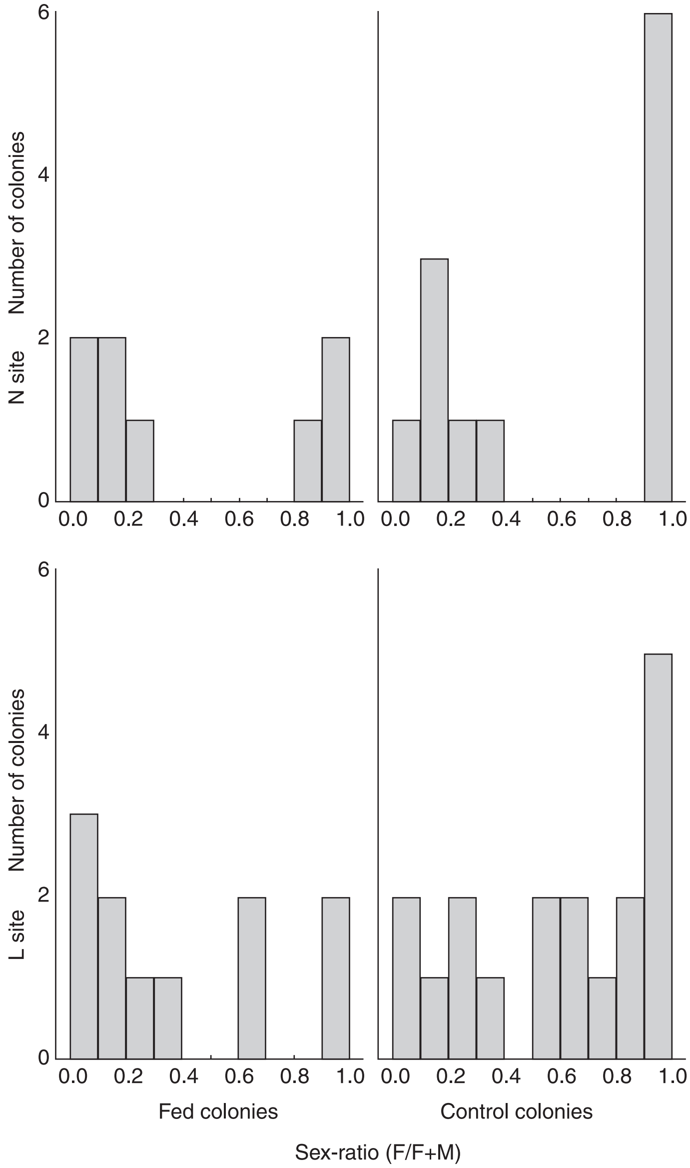

Fig. 5. Colony-level numeric sex ratios by treatment for each site.

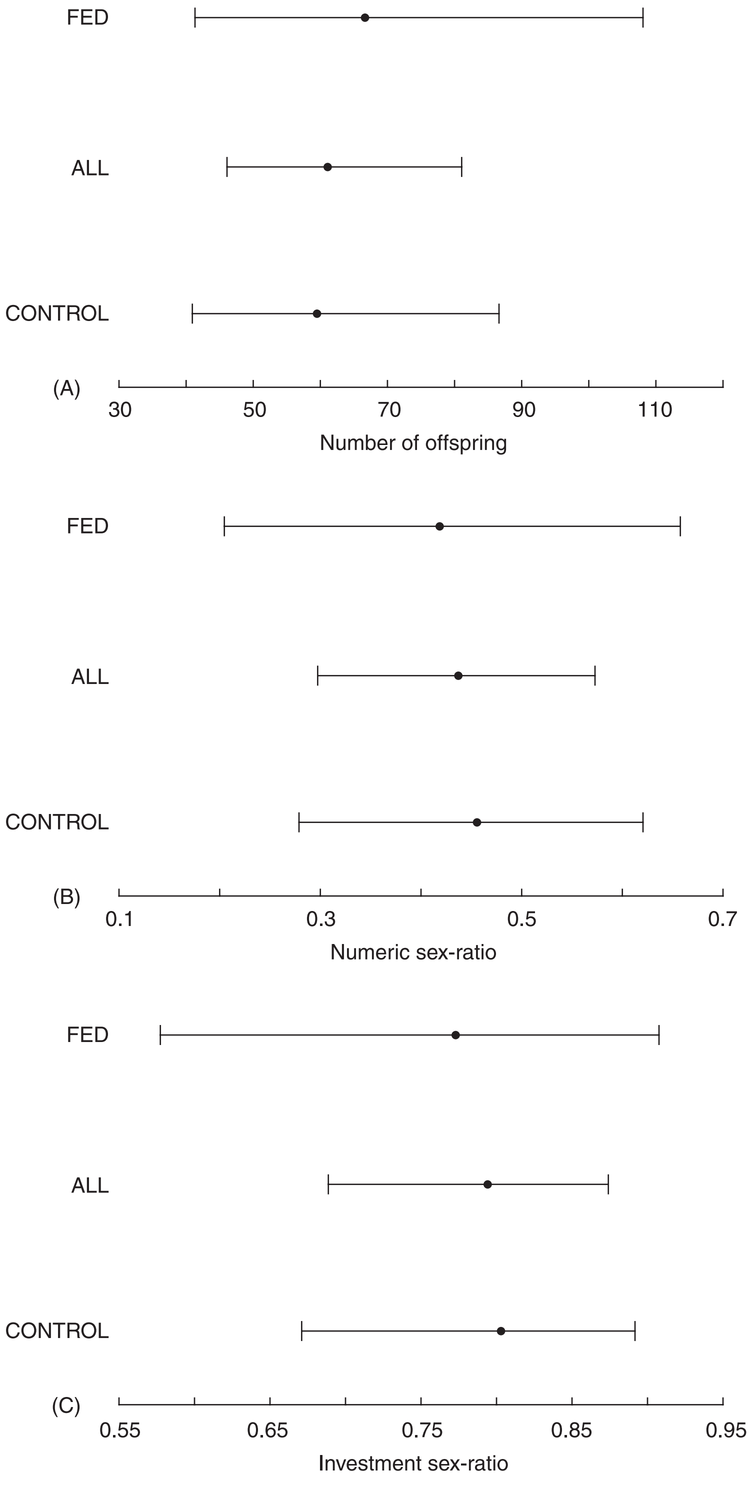

Fig. 6. Estimated means and 95% credibility intervals for sexual production values using the accepted model: (A) mean number of reproductive offspring produced per reproducing colony; (B) population numeric sex ratio; (C) population investment sex ratio. Each figure shows the control colony estimate (CONTROL, $n = 30$), the estimate for all the colonies combined from both treatment groups (ALL, $n = 39$), and the protein-supplemented colony estimate (FED, $n = 19$).

REFERENCES

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. In *2nd International Symposium on Information Theory* (B.N. Petrov and F. Csáki, eds.), pp. 267–281. Budapest: Akadémiai Kiadó.

Aron, S., Keller, L. and Passera, L. 2001. Role of resource availability on sex, caste and reproductive allocation ratios in the Argentine ant *Linepithema humile. J. Anim. Ecol.*, **70**: 831–839.

Backus, V.L. and Herbers, J.M. 1992. Sexual allocation ratios in forest ants: food limitation does not explain observed patterns. *Behav. Ecol. Sociobiol.*, **30**: 425–429.

Benford, F.A. 1978. Fisher's theory of the sex ratio applied to the social Hymenoptera. *J. Theor. Biol.*, **72**: 701–727.

Bono, J.M. and Heithaus, R.E. 2002. Sex ratios and the distribution of elaiosomes in colonies of the ant, *Aphaenogaster rudis. Insect. Soc.*, **49**: 320–325.

Bono, J.M. and Herbers, J.M. 2003. Proximate and ultimate control of sex ratios in *Myrmica brevispinosa* colonies. *Proc. R. Soc. Lond. B*, **270**: 811–817.

Boomsma, J.J. 1989. Sex-investment ratios in ants: has female bias been systematically overestimated? *Am. Nat.*, **133**: 517–532.

Boomsma, J.J. and Nachman, G. 2002. Analysis of sex ratios in social insects. In *Sex Ratios: Concepts and Research Methods* (I.C.W. Hardy, ed.), pp. 93–111. Cambridge: Cambridge University Press.

Boomsma, J.J., Keller, L. and Nielsen, M.G. 1995. A comparative analysis of sex ratio investment parameters in ants. *Funct. Ecol.*, **9**: 743–753.

Bourke, A.F.G. and Franks, N.R. 1995. *Social Evolution in Ants*. Princeton, NJ: Princeton University Press.

Buschinger, A. and Pfeifer, E. 1988. Effects of nutrition on brood production and slavery in ants (Hymenoptera: Formicidae). *Insect. Soc.*, **35**: 61–69.

Craig, R. 1980. Sex ratio changes and the evolution of eusociality in the Hymenoptera: simulation and games theory studies. *J. Theor. Biol.*, **87**: 55–70.

Crozier, R.H. 1973. Apparent differential selection at an isozyme locus between queens and workers of the ant *Aphaenogaster rudis. Genetics*, **73**: 313–318.

Crozier, R.H. 1974. Allozyme analysis of reproductive strategy in the ant *Aphaenogaster rudis. Isozyme Bull.*, **7**: 18.

Crozier, R.H. and Pamilo, P. 1996. *Evolution of Social Insect Colonies*. Oxford: Oxford University Press.

DeHeer, C.J., Backus, V.L. and Herbers, J.M. 2001. Sociogenetic responses to ecological variation in the ant *Myrmica punctiventris* are context dependent. *Behav. Ecol. Sociobiol.*, **49**: 375–386.

Deslippe, R.J. and Savolainen, R. 1995. Sex investment in a social insect: the proximate role of food. *Ecology*, **76**: 375–382.

Foitzik, S., Stratz, M. and Heinze, J. 2003. Ecology, life history and resource allocation in the ant, *Leptothorax nylanderi. J. Evol. Biol.*, **16**: 670–680.

Frank, S.A. 1987. Variable sex ratio among colonies of ants. *Behav. Ecol. Sociobiol.*, **20**: 195–201.

Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. 2004. *Bayesian Data Analysis*. Boca Raton, FL: Chapman & Hall/CRC Press.

Gill, J. 2002. *Bayesian Methods: A Social and Behavioral Sciences Approach*. Boca Raton, FL: Chapman & Hall/CRC Press.

Grafen, A. 1986. Split sex ratios and the evolutionary origins of eusociality. *J. Theor. Biol.*, **122**: 95–121.

Hardy, I.C.W., ed. 2002. *Sex Ratios: Concepts and Research Methods*. Cambridge: Cambridge University Press.

Hasegawa, E. 1994. Sex allocation in the ant *Colobopsis nipponicus* (Wheeler). I. Population sex ratio. *Evolution*, **48**: 1121–1129.

Headley, A.E. 1949. A population study of the ant *Aphaenogaster fulva* ssp. *aquia* Buckley (Hymenoptera: Formicidae). *Ann. Entomol. Soc. Am.*, **42**: 265–272.

Heithaus, E.R. 1981. Seed predation by rodents on three ant-dispersed plants. *Ecology*, **62**: 136–145.

Herbers, J.M. and Banschback, V.S. 1998. Food supply and reproductive allocation in forest ants: repeated experiments give different results. *Oikos*, **83**: 145–151.

Krackow, S. and Tkadlec, E. 2001. Analysis of brood sex ratios: implications of offspring clustering. *Behav. Ecol. Sociobiol.*, **50**: 293–301.

Kruuk, L.E.B., Clutton-Brock, T.H., Albon, S.D., Pemberton, J.M. and Guinness, F.E. 1999. Population density affects sex ratio variation in red deer. *Nature*, **399**: 459–461.

Lubertazzi, D. 2006. *Sex ratio expression in ants: empirical studies with Aphaenogaster rudis*. Doctoral dissertation, University of Connecticut, Storrs, CT.

McCarthy, M. 2007. *Bayesian Methods for Ecology*. Cambridge: Cambridge University Press.

McCullagh, P. and Nelder, J.A. 1989. *Generalized Linear Models*. London: Chapman & Hall/CRC Press.

Morales, M.A. and Heithaus, R.E. 1998. Food from seed-dispersal mutualism shifts sex ratios in colonies of the ant *Aphaenogaster rudis. Ecology*, **79**: 734–739.

Nonacs, P. 1986a. Ant reproductive strategies and sex allocation theory. *Q. Rev. Biol.*, **61**: 1–21.

Nonacs, P. 1986b. Sex-ratio determination within colonies of ants. *Evolution*, **40**: 199–204.

Pamilo, P. 1982. Genetic evolution of sex ratios in eusocial Hymenoptera: allele frequency simulations. *Am. Nat.*, **119**: 638–656.

Pampel, F.C. 2000. *Logistic Regression: A Primer*. Thousand Oaks, CA: Sage.

Pearcy, M. and Aron, S. 2006. Local resource competition and sex ratio in the ant *Cataglyphis cursor. Behav. Ecol.*, **17**: 569–574.

R Development Core Team. 2009. *R: A Language and Environment for Statistical Computing*, 2nd edn. Vienna, Austria: R Foundation for Statistical Computing.

Rosenheim, J.A., Nonacs, P. and Mangel, M. 1996. Sex ratios and multifaceted parental investment. *Am. Nat.*, **148**: 501–535.

Schall, R. 1991. Estimation in generalized linear models with random effects. *Biometrika*, **78**: 719–727.

Sorvari, J. and Hakkarainen, H. 2007. Forest clearing and sex ratio in forest-dwelling wood ant *Formica aquilonia. Naturwissenschaften*, **94**: 392–395.

Southerland, M.T. 1988. The effects of temperature and food on the growth of laboratory colonies of *Aphaenogaster rudis* Emery (Hymenoptera: Formicidae). *Insect. Soc.*, **35**: 304–309.

Spiegelhalter, D.J., Best, N.G., Carlin, B.R. and van der Linde, A. 2002. Bayesian measures of model complexity and fit. *J. R. Stat. Soc. B Stat. Meth.*, **64**: 583–616.

Talbot, M. 1951. Populations and hibernating conditions of the ant *Aphaenogaster* (*Attomyrma*) *rudis* Emery (Hymenoptera: Formicidae). *Ann. Entomol. Soc. Am.*, **44**: 302–307.

Thomas, A., O'Hara, B., Uwe, L. and Sturtz, S. 2006. Making BUGS open. *R News*, **6**: 12–17.

Trivers, R.L. 1974. Parent–offspring conflict. *Am. Zool.*, **14**: 249–264.

Trivers, R.L. and Hare, H. 1976. Haplodiploidy and the evolution of the social insects. *Science*, **191**: 249–263.

Wiernasz, D.C. and Cole, B.J. 2009. Dioecy and the evolution of sex ratios in ants. *Proc. R. Soc. Lond. B*, **276**: 2125–2132.

Williams, D.A. 1982. Extra-binomial variation in logistic linear models. *Appl. Stat.*, **31**: 144–148.

Wilson, K. and Hardy, I.C.W. 2002. Statistical analysis of sex ratios: an introduction. In *Sex Ratios: Concepts and Research Methods* (I.C.W. Hardy, ed.), pp. 48–92. Cambridge: Cambridge University Press.