

Adaptive plasticity in extra-group breeding in a social mammal with high inbreeding risk

Supplementary material

SM1 Independence claims

Table S1. Table of the independence claims and their results from the piecewise structural equation model presented in the main text. Response variables are shown to the left of the '~' while explanatory variables are to the right and are separated by '+'. '*' indicates an interaction between two explanatory variables. The independence claim is between the response variables and the first explanatory variable listed in each row conditional upon the subsequent variables in the row.

Independence claim	Estimate	Std.Error	DF	<i>p</i> -value
Group size ~ in natal pack + group age	0.00	0.01	343.28	0.96
Group size ~ age + group age	0.00	0.00	169.25	0.48
Relatedness ~ age + group age + in natal pack	0.06	0.03	454.73	0.08
Relatedness ~ group size + group age + in natal pack	-0.02	0.04	273.17	0.56
Extra-group breeding ~ relatedness + group age + in natal pack + age + group size	0.30	0.18	NA	0.14
Maternal litter f ~ group age + extra-group breeding*relatedness	0.01	0.05	24.21	0.89
Maternal litter f ~ in natal pack + extra-group breeding*relatedness	0.28	0.15	84.96	0.07
Maternal litter f ~ age + extra-group breeding*relatedness	0.03	0.04	543.37	0.38
Maternal litter f ~ group size + group age + extra-group breeding*relatedness	0.05	0.04	251.45	0.18

SM2 MCMC model of extra-group breeding

This section describes how we refitted the extra-group breeding model, described in the main text and Table 3, using MCMCglmm (Hadfield, 2010). Refitting the model had two advantages; we could accurately assess the significance of the fixed effects and come to a more complete understanding of the estimated variance components, most interestingly individual repeatability and breeding event repeatability. Both of these advantages are inherent in any statistical framework that estimates the full posterior distribution of the model parameters.

The extra-group breeding model was refitted using MCMCglmm using the same fixed and random effect structure. Parameter expanded priors ($V=1$, $\nu=0.002$, $\alpha.\mu=0$, $\alpha.V=1000$) were used for each random effect and the residual variance was fixed to one as it is inestimable for binary models (Nakagawa and Schielzeth, 2010). Because the residual variance was fixed, the random effect variances can only be meaningfully interpreted after rescaling because absolute values of random effect variances can be completely different based on the value residual variance is fixed to. This was done following (Nakagawa and Schielzeth, 2010) for a binomial model with a logit link and additive overdispersion, as described in the main text.

The model was run for 5,265,000 iterations with a burn-in of 3000 and a thinning interval of 5000 iterations. Trace plots and Geweke plots were checked for each parameter to ensure good mixing of the chain and convergence. Autocorrelation and effective sample size were also checked to ensure sufficient sampling of the posterior distribution.

The fixed effect coefficients estimated by glmer and MCMCglmm showed good agreement (Figure S1). This agreement indicates that both models successfully converged on the correct set of model parameters.

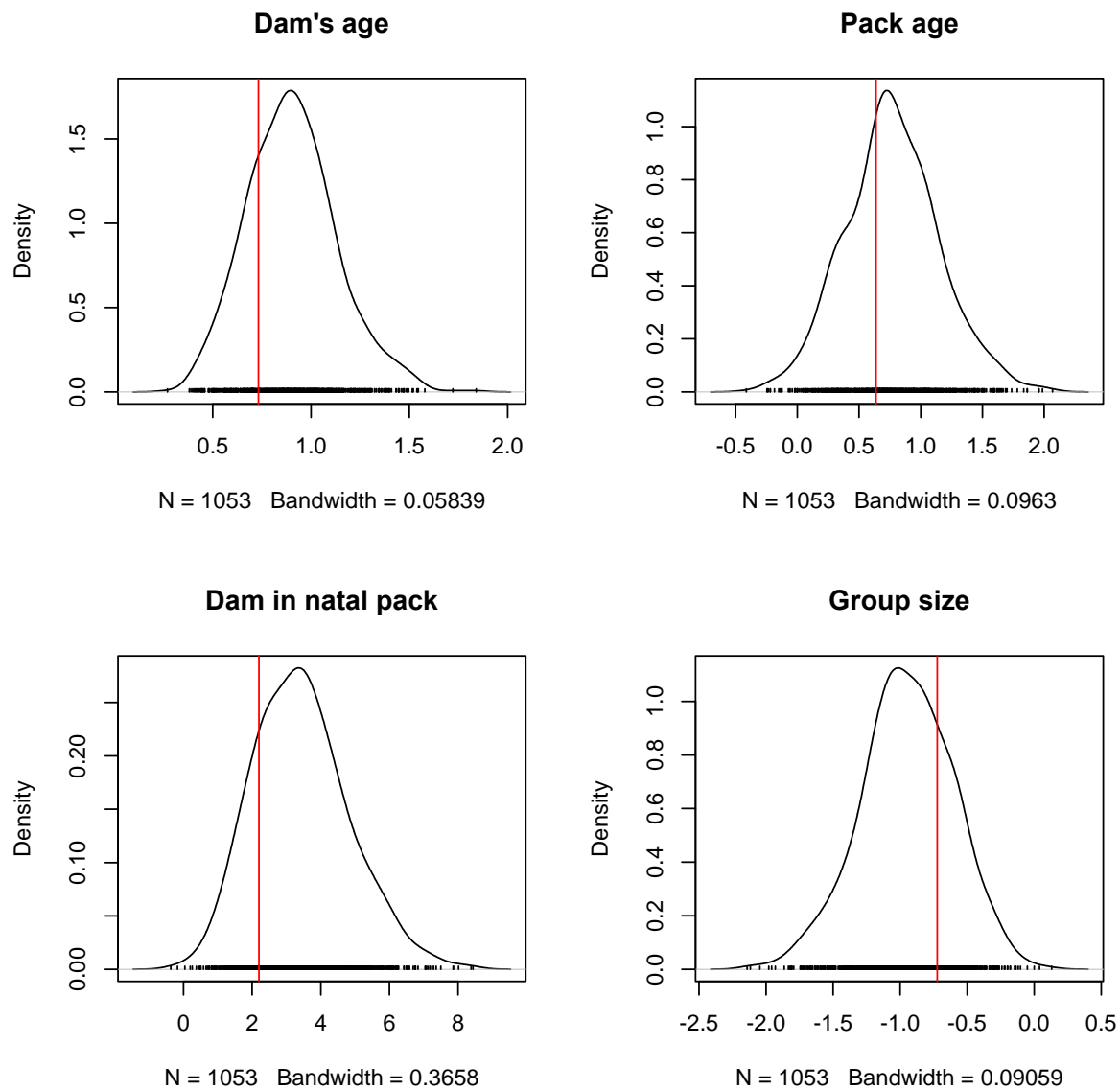


Figure S1. Density plots showing the posterior distribution of the fixed effect coefficients from the MCMCglmm extra-group breeding model. The red vertical line indicates the point estimate from the model fitted using lme4. The ticks along the x-axis indicate individual draws from the posterior distribution. The text below each panel indicates the number of draws from the posterior distribution and the bandwidth used when calculating the density.

The posterior distribution of the individual repeatability is asymmetrical and so difficult to describe using traditional point estimates and standard deviations. This illustrates one of the advantages of describing the complete posterior distribution of variance components as they are often asymmetrical. The posterior mode of individual repeatability is 0.09 but the 95% posterior density interval is bounded at zero as shown in Figure S2 (0-0.21). It is therefore somewhat unclear whether females are repeatable in their extra-group breeding tendencies. In contrast, females in the same social groups are clearly not consistent (posterior mode = 0.00, 95% posterior density interval = 0.00-0.23, Figure S2), whereas females giving birth to pups in the same communal litter do show significant repeatability (Figure S2, posterior mode = 0.50, 95% posterior density interval = 0.30-0.67). See the discussion in the main text for further details.

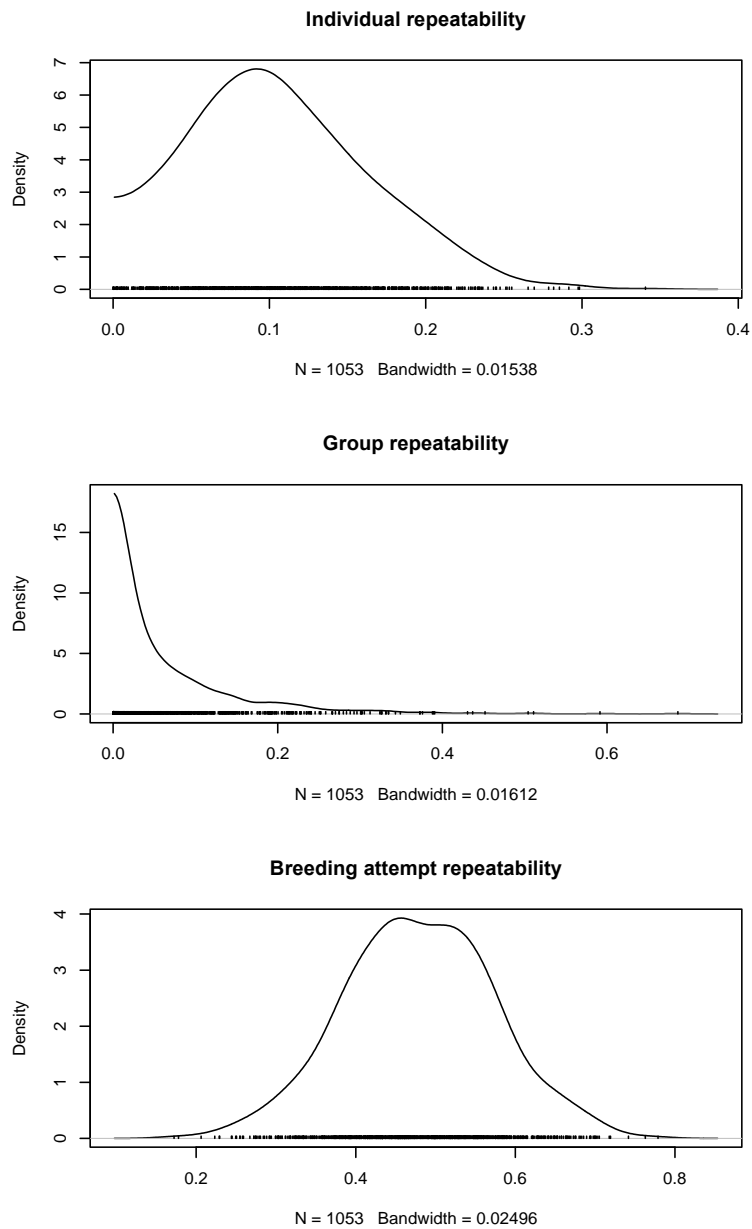


Figure S2. Density plots showing the posterior distribution of random effect repeatability for the MCMCglmm extra-group breeding model. The ticks along the x-axis indicate individual draws from the posterior distribution. The text below each panel indicates the number of draws from the posterior distribution and the bandwidth used when calculating the density.

SM3 Figures of within-individual changes in behaviour

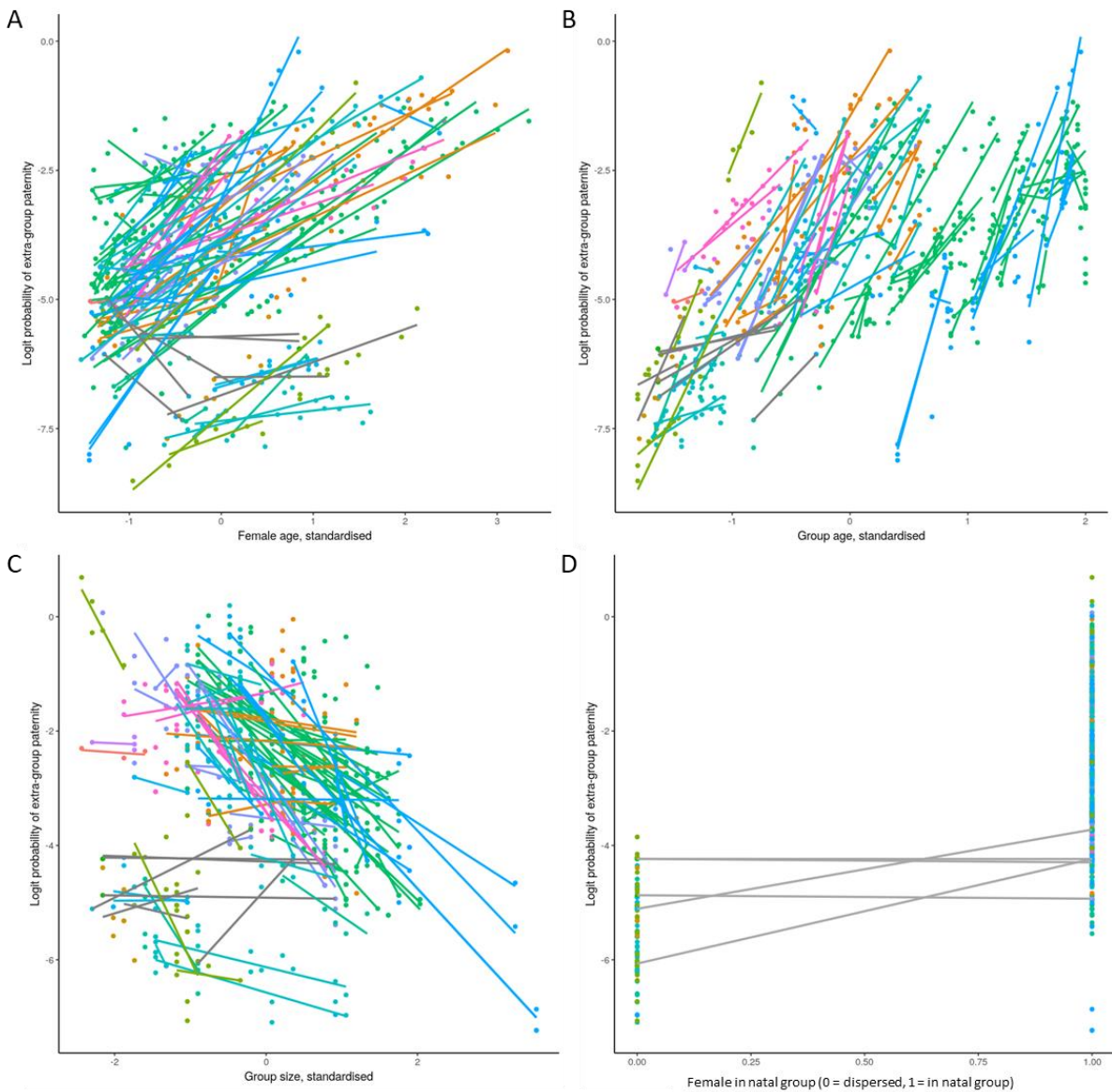


Figure S3. The change in logit probability of extra-group paternity against A) female age, B) group age, C) group size and D) whether or not the female was breeding in her natal group, after accounting for other variables in the model. Points are coloured by social group and each line is fit to data of a single female. Grey lines represent females that have moved between social groups over the course of the study. Plots of changes in the occurrence of

EGP with female age, group age and group size (A-C) show within-individual plasticity in the frequency of extra-group paternity. The plot of female dispersal (D) shows that only 5 females have bred in multiple groups over the course of our study, so there is limited evidence of within-individual plasticity in extra-group paternity in relation to female dispersal status.

SM4 Predicting offspring heterozygosity

In the main paper we focused on pedigree inbreeding coefficient to quantify inbreeding as it is directly interpretable (0.25 is equivalent to breeding between full siblings) and generally explains more variation in fitness compared to heterozygosity estimates (Balloux et al., 2004; Nietlisbach et al., 2017; Slate et al., 2004). However, validation of our model predicting the average maternal litter inbreeding coefficient showed evidence of heteroscedacity, specifically an increase in residual variance with increasing relatedness values. This is likely because inbreeding coefficient does not follow a normal distribution but has a lower bound of zero. Here we fit an equivalent model using average maternal litter standardised multi-locus heterozygosity (sMLH) to estimate inbreeding and demonstrate that our results are not dependant on our measure of inbreeding.

Heterozygosity was calculated as sMLH using the R package *inbreedR* (Stoffel et al., 2016). sMLH is defined as the total number of heterozygous loci in an individual divided by the sum of average observed heterozygosities in the population for the same loci, with lower values representing less genetically diverse individuals. sMLH was calculated using the same panel of 35-43 microsatellites used to generate the pedigree, as detailed in the main text. No microsatellite loci were found to consistently deviate from Hardy-Weinburg equilibrium, and no pairs of loci were in linkage disequilibrium (Sanderson et al., 2015). All loci were manually checked for sex linkage by comparing a subset of male and female genotypes.

We investigated the effect of extra-group breeding and relatedness to the top ranked males on the average maternal litter heterozygosity. The model was fitted as a LMM in the R package lme4 (Bates et al., 2015). As in the model using inbreeding coefficients, we included an interaction between relatedness and the extra-group paternity status of the maternal litter because the relatedness between within-group individuals should not influence extra-group inbreeding. The random effect structure was also the same and included a random intercept for mother, mother's social group at conception and breeding event.

The significance of the interaction term was assessed through parametric bootstrapping. A reduced model was created by dropping the interaction term then both the reduced and full model were fit to the observed data to generate an observed likelihood statistic. The significance of this statistic was assessed by comparison to a distribution of similarly calculated likelihood statistics based on 1000 simulated datasets generated according to the reduced model. This procedure was conducted using the R package pbkrtest (Halekoh and Højsgaard, 2014).

Using sMLH as a measure of inbreeding did not qualitatively change our results compared to using pedigree inbreeding coefficients. As a female became more related to the top ranked males in her group, the heterozygosity of her within-group offspring decreased; this is equivalent to an increase in inbreeding (Figure S3). By contrast, the heterozygosity of a female's extra-group offspring was independent of her relatedness to top-ranked males, which was shown by a significant interaction coefficient approximately equal and opposite to the relatedness coefficient (relatedness $\beta = -0.22$, extra-group breeding $\beta = 0.35$, interaction $\beta = 0.25$, $p = 0.009$, Figure S3).

Average maternal litter sMLH was normally distributed and visual inspection of residual plots confirmed that the assumptions of this model were not violated. This supports the conclusion of our model based on inbreeding coefficients and implies that those results are not due to heteroscedasticity in the model.

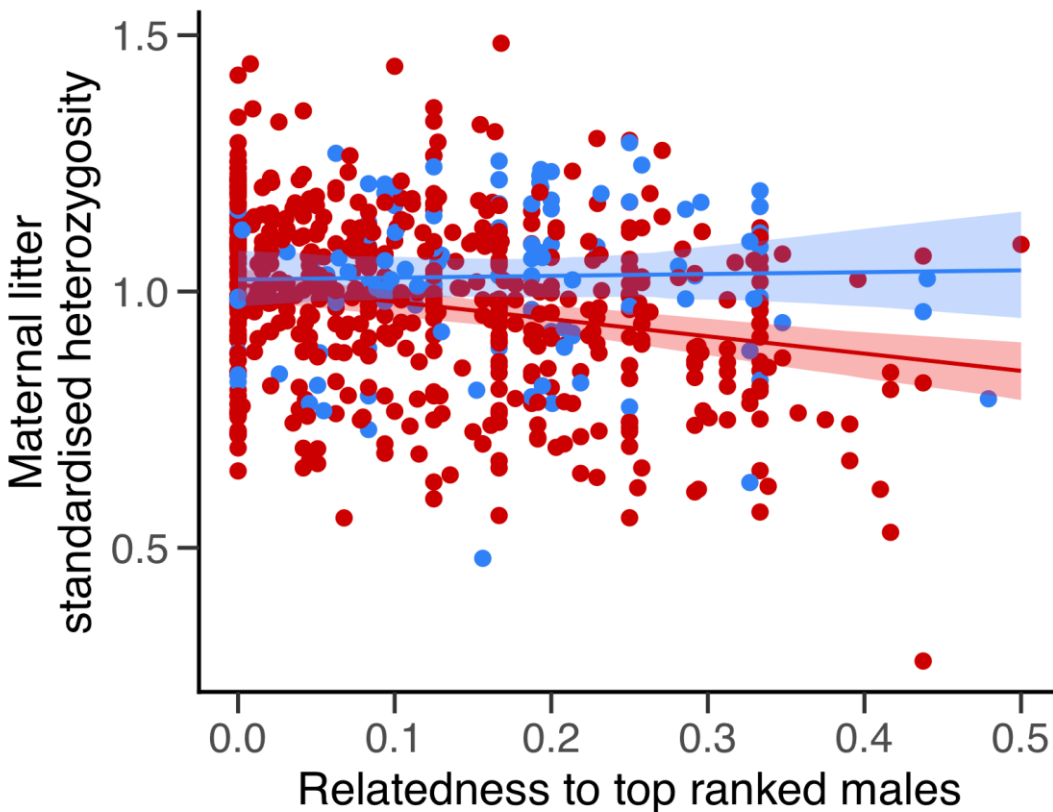


Figure S3. The average heterozygosity (sMLH) of maternal litters as a function of the mother's relatedness to the top ranked males in her social group and whether the maternal litter resulted from within-group breeding (red) or extra-group breeding (blue). Trend lines show the fitted model and the shaded regions show the 95% CI.

References

Balloux F, Amos W, Coulson TN, 2004. Does heterozygosity estimate inbreeding in real populations? *Molecular Ecology*. 13:3021-3031.

Hadfield JD, 2010. MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software*. 33:1-22.

Halekoh U, Højsgaard S, 2014. A kenward-roger approximation and parametric bootstrap methods for tests in linear mixed models—the R package pbkrtest. *Journal of Statistical Computation and Simulation*. 59:1-30.

- Nakagawa S, Schielzeth H, 2010. Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. *Biological Reviews*. 85:935-956.
- Nietlisbach P, Keller LF, Camenisch G, Arcese P, Reid JM, Postma E, 2017. Pedigree-based inbreeding coefficient explains more variation in fitness than heterozygosity at 160 microsatellites in a wild bird population. *Proceedings of the Royal Society of London, Series B: Biological Sciences*. 284:20162763.
- Sanderson JL, Wang J, Vitikainen EI, Cant MA, Nichols HJ, 2015. Banded mongooses avoid inbreeding when mating with members of the same natal group. *Molecular Ecology*. 24:3738-3751.
- Slate J, David P, Dodds KG, Veenvliet BA, Glass BC, Broad TE, McEwan JC, 2004. Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: theoretical expectations and empirical data. *Heredity* 93: 255-265.