

Analysis Plan

Do red squirrels benefit from living near close kin?

Proposed authorship: Sam F. Walmsley, Stan Boutin, Ben Dantzer, Jeffrey E. Lane, David W. Coltman, Andrew G. McAdam.

Data analysis steps

1. Organize observations from long-term dataset: For all squirrels living on grid (Kloo and Sulfur) between 2003-2014¹, we will identify pairwise relationships among close kin ($r = 0.5$; mother, father, son, daughter, full sibling) using the KRSP pedigree. Records, indexed by individual and year, will include the age and fitness (survival, sex-specific ARS) of the focal individual, the genealogical relationship and territory distance to the “target” close kin, and whether the current year is a mast year. Territory distance will be measured as the linear distance between primary middens of each squirrel (including defended areas for individuals lacking middens). From these records, we will subset datasets for use in modelling based only on *nearest* close kin.

2. Exploratory analyses: We will plot the distribution of geographic distances between primary territories for each relationship type (one record per dyad), visualizing individual variation in proximity to close kin. For each fitness measure and genealogical category, the relationship between fitness and distance to kin will be described using generalized additive mixed effects models (GAMMs) to test for non-linearities.

3. Model fitness as a function of distance to nearest close kin: If visualizations and spline terms (Step 2) suggest that the relationships between distance and fitness can be represented linearly, we will use generalized linear mixed effects models (GLMMs) to estimate partial influences on fitness. Alternatively, we will develop parametric non-linear models to explain variation in fitness. Separate models will be fit using sex-specific survival and ARS as response variables and each will condition on mast year and effects of the age of the focal squirrel (linear and quadratic terms). Linear distances may be log-transformed, if there is right-skew in their distribution. We will also fit a global model with sex as an interaction with distance and relationship type to test for sex-based differences in survival effects. We will account for multiple observations of the same individual by estimating random intercepts around the focal squirrel’s ID. The distribution of records per individual and dyad will guide our decision to add a dyad-level random intercept *or* further thin the dataset to accommodate for the non-independence of repeated measures (sometimes bi-directional) of the same pairs of individuals. Analyses will be conducted in R version 4.0.3 [1].

Simplified equations explaining fitness as a function of distance to close kin, by relationship:

$$Survival_{female} \sim distance_{nearestCloseKin} \times relationship + mast_{Y/N} + age + (age)^2 + (1|focalID) \quad (1)$$

¹ Additional years may be included if new paternity data becomes available in time for analysis.

$$Survival_{male} \sim distance_{nearestCloseKin} \times relationship + mast_{Y/N} + age + (age)^2 + (1|focalID) \quad (2)$$

$$ARS_{female} \sim distance_{nearestCloseKin} \times relationship + mast_{Y/N} + age + (age)^2 + (1|focalID) \quad (3)$$

$$ARS_{male} \sim distance_{nearestCloseKin} \times relationship + mast_{Y/N} + age + (age)^2 + (1|focalID) \quad (4)$$

4. Consider differences between littermates and non-littermates: Following the approach outlined in steps 2 & 3, we will fit similar models to test whether any influences of close kin on fitness depend on early-life experience, i.e., if fitness benefits are only gained from living near littermates. We will classify each relationship as being between “littermate close kin” (i.e., full siblings in the same litter, mothers and offspring) or “non-littermate close kin” (i.e., full siblings from different litters, fathers and offspring).

Simplified equations explaining fitness as a function of distance to close kin littermates:

$$Survival_{female} \sim distance_{nearestCloseKin} \times littermate_{Y/N} + mast_{Y/N} + age + (age)^2 + (1|focalID) \quad (5)$$

$$Survival_{male} \sim distance_{nearestCloseKin} \times littermate_{Y/N} + mast_{Y/N} + age + (age)^2 + (1|focalID) \quad (6)$$

$$ARS_{female} \sim distance_{nearestCloseKin} \times littermate_{Y/N} + mast_{Y/N} + age + (age)^2 + (1|focalID) \quad (7)$$

$$ARS_{male} \sim distance_{nearestCloseKin} \times littermate_{Y/N} + mast_{Y/N} + age + (age)^2 + (1|focalID) \quad (8)$$

Expectation of large distance values & supplementary analyses

Given the low relatedness of red squirrel neighbourhoods, we expect there to be a sufficiently large range of territory distances to capture variation in a dyad’s ability to interact (even when considering nearest close kin). However, if individuals typically have a close relative nearby (e.g., < 130m), we may fit additional models incorporating records of individuals *without* close kin in the population. These individuals may be included in models comparing individuals with and without close kin in the population (across all relationships), or by including them in original models with large distance values (i.e., the maximum observed distance among close kin). Based on distance-fitness relationships identified by GAMMs, we may also test whether the *number* of close kin within an empirically informed distance threshold influences fitness.

Proposed figures

1. Boxplots or violin plots showing distributions of geographic distances between primary middens of mothers & offspring, fathers & offspring, and among full siblings.
2. Scatterplots showing overall relationships between fitness and distance with raw data and GAMM fits.
3. Modelled, partial influences of distance to close kin on fitness for any genealogical relationships of interest.
4. Modelled, partial influences of distance to close kin on fitness for the littermate vs. non-littermate comparison.
5. Graphical abstract to summarize main findings.

References

1. R Core Development Team. 2020 R: A language and environment for statistical computing.