



Research



Cite this article: Philson CS, Martin JGA, Blumstein DT. 2025 Multilevel selection on individual and group social behaviour in the wild. *Proc. R. Soc. B* **292**: 20243061. <https://doi.org/10.1098/rspb.2024.3061>

Received: 20 December 2024
Accepted: 24 February 2025

Subject Category:
Evolution

Subject Areas:
behaviour, evolution

Keywords:
multilevel selection, social behaviour, social networks, individual social position, group social structure

Author for correspondence:
Conner S. Philson
e-mail: cphilson@ucla.edu

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.7712226>.

Multilevel selection on individual and group social behaviour in the wild

Conner S. Philson^{1,2,3}, Julien G. A. Martin⁴ and Daniel T. Blumstein^{1,2}

¹Department of Ecology & Evolutionary Biology, UCLA, Los Angeles, CA 90095, USA

²Rocky Mountain Biological Laboratory, Crested Butte, CO 81224, USA

³Natural Reserve System, UCSB, Santa Barbara, CA 93106, USA

⁴Department of Biology, University of Ottawa, Ottawa, Ontario, K1N 9A4, Canada

id CSP, 0000-0002-5974-347X; JGAM, 0000-0001-7726-6809; DTB, 0000-0001-5793-9244

How phenotypes are shaped by multilevel selection—the theoretical framework proposing natural selection occurs at more than one level of biological organization—is a classic debate in biology. Though social behaviours are a common theoretical example for multilevel selection, it is unknown if and how multilevel selection acts on sociality in the wild. We studied the relative strength of multilevel selection on both individual behaviour and group social structure, quantified with social networks and 19 years of data from a wild, free-living mammal, the yellow-bellied marmot (*Marmota flaviventris*). Contextual analysis (exploring the impact of individual and group social phenotypes on individual fitness, relative to each other) revealed multilevel selection gradients in specific fitness and life history contexts, with selection for group social structure being just as strong, if not stronger, than individual social behaviour. We also found antagonistic multilevel selection gradients within and between levels, potentially explaining why increased sociality is not as beneficial or heritable in this system compared with other social taxa. Thus, the evolutionary dynamics of hierarchical or nested biological traits should be assessed at multiple levels simultaneously to tell a more accurate and comprehensive story. Overall, we provide empirical evidence suggesting that multilevel selection acts on social relationships and structures in the wild and provide direct evidence for a classic, unanswered question in biology.

1. Introduction

The evolution of sociality is a central question in biology [1–3]. Classic explanations for the emergence of social group living focus on the individual fitness costs and benefits incurred via, for example, increased resource competition or predator detection and avoidance [1–3]. This has helped develop a general theory for the selection pressures of social group living. However, within social groups, individuals vary in their pattern of social interactions with others and this individual variation, in part, drives differences in emergent group social structure [2,4–8]. In turn, group structures feed back to influence patterns of individual social interactions [7,8].

Despite these two levels of social organization being nested and thus non-independent, both can be quantified as discrete phenotypic levels of biological organization [2,5–9]. These individual and group social phenotypes both have fitness and population dynamic consequences in wild populations [5,6,10–15] and affect health and human ageing [12,16]. While the drivers and consequences of individual social behaviour and group social structure have been explored independently, if and how these two nested social phenotypes are under selection relative to each other remain largely unknown.

Multilevel selection is a theoretical framework positing natural selection simultaneously occurs at multiple levels of biological organization, and at levels other than only the gene [17–22]. The theory of multilevel selection has undergone many transformations over the years through healthy debate [17–19,21,23–26] and has been suggested as a driving force in the emergence of multicellular organisms [25,27], human cultural evolution [22] and the structure of entire ecosystems [28,29].

The modern multilevel selection argument proposes that individual fitness may be impacted by an individual's phenotype and by the group's phenotype simultaneously [9,21,25,30]. This modern argument is not suggesting groups have fitness and that there is variation among groups (e.g. as done with the Price–Hamilton approach [31,32]). Instead, individuals are the result of selection on both individual and group phenotypes, as observed in a variety of morphological phenotypes in plants [33] and animals [34–36]. For example, a flower's fitness may be impacted by how tall its stem is but also impacted by how tall the stems around it are—so both the flower's individual phenotype and the phenotype of its group impact the individual flower's fitness. In the context of social behaviour, individual and group social phenotypes also both have documented fitness consequences for individuals too, when either level is explored independently, as seen in some birds [10], fishes [37], insects [9] and mammals [13–15]. Because individuals are impacted by both their individual and their group's phenotype, a multilevel selection framework can tease apart what selection is on the individual phenotype and what selection is on the group phenotype.

Furthermore, the individual fitness consequences between the two phenotypic levels, when explored independently or simultaneously, do not always align [9,15]. Because the levels do not always align and because the levels feed back to influence each other, natural selection at one level will likely affect the other (either directly or indirectly [38,39]). Selection at one level may increase or inhibit selection at the other, and thus have a stronger impact on the overall evolutionary response [39]. Furthermore, when exploring selection on either level independently, selection on a higher or lower level may be detected at the level being explored, leading to inaccurate or biased evolutionary predictions or false positives. Alternatively, no selection may be detected despite selection occurring at a level not being explored, leading to false negatives and inaccurate evolutionary expectations.

Therefore, for nested traits like social behaviour, selection should not be quantified at either level independently but should be quantified simultaneously at all relevant levels to partition selection among the nested levels. Thus, multilevel selection may be an important mechanism for the evolution and maintenance of the nested individual social behaviours and group social structures that emerge from group living.

Despite social behaviour being a common theoretical example in making the case for multilevel selection [17–19,21–26], prior work has not explored the presence and strength of multilevel selection for social behaviours in wild, free-living populations. Thus, the decades-old debate of individual versus group selection for social behaviours [40], and more recently multilevel selection [20,26], remains largely theoretical and not empirical [25]. If multilevel selection acts on sociality in the wild is an open question in evolutionary biology that has the potential to inform our understanding of the evolutionary origins of social relationships and structures that underpin social lives across species.

To address if multilevel selection acts on sociality in the wild, we used 19 years of continuous social, fitness and life history data on a well studied, wild, free-living population of yellow-bellied marmots (*Marmota flaviventris*). A harem polygynous rodent, yellow-bellied marmots live in colonies and socially interact in summer before hibernating over winter. The evolutionary origins of yellow-bellied marmot group living may be attributable to predator avoidance and/or reflect limited opportunities to disperse to establish new colonies [41]. Marmot social groups are composed of mostly kin, with occasional non-kin immigration (most often a yearling or adult male [41]). Individual social behaviour has a genetic basis [42] and individual behaviour and group structure are environmentally mediated [41–45]. Individuals also experience fitness consequences based on both their individual and group social phenotypes [13–15,46–50], namely, more social individuals have increased summer survival owing to antipredator benefits [50] and are more philopatric potentially because of the costs of dispersal [46], but have decreased hibernation survival [48], reproductive success [47] and longevity [49] potentially owing to the time and energetic costs of social interactions. Individuals residing in more socially connected groups also experience decreased individual reproductive success [14], again owing to energetic costs, but increased individual winter survival [15] as a result of potential social hibernation and thermoregulatory benefits. Increased sociality being largely associated with individual fitness costs is largely unique for social mammals [12], but nevertheless shows that sociality has consequences in this species.

While exploring nested levels of organization independently may lead to inaccurate estimation of selection, these studies exploring the levels of sociality independently provide a strong background to ask informed questions about the importance of multilevel selection for social traits. Thus, with identified genetic and environmental drivers of social phenotypic variance and varying fitness consequences for individuals (when explored independently), this system provides an ideal opportunity to estimate multilevel selection for individual and group social phenotypes in the wild to more accurately partition selection among the nested levels of organization and ultimately obtain a better understanding of the evolution of sociality.

To quantify the social phenotype, we used affiliative social interactions from 172 social groups comprising 723 unique individuals across 19 years to construct social networks [4–6] and calculated a pair of analogous individual and group social network measures for four core social traits (table 1, [9]). Each measure is commonly used across studies in human and non-human animal social network studies [4–6]. The analogous pairs are not strictly required to assess multilevel selection, but because multilevel selection could occur between the social traits, this approach allows clear evaluation of the relative importance of both individual- and group-level selection for specific social traits. To evaluate the distinct contributions of individual social behaviour and group social structure to four annual individual fitness correlates (summer survival, hibernation survival, if a female weaned offspring, and how many offspring a mother had), we used a contextual analysis that uses a partial regression to partition selection among levels [9,30,51]. Contextual analysis defines individual-level selection as the impact that the individual phenotype has on individual fitness and defines group-level selection as the impact that group phenotype has on

individual fitness. Despite the inherent non-independence of the individual and group social phenotypes, contextual analysis partitions among the two phenotypic levels relative to each other.

Because individual social behaviour and group social structure are nested, if we find selection was only observed at one phenotypic level, we could interpret that level as under direct selection while interpreting the other level as under indirect selection [30]. This would not support the presence of direct multilevel selection for social behaviour. If direct selection was observed for both social phenotypes, even across different social traits or contexts, we would have suggestive evidence of multilevel selection for sociality in the wild. Because, when exploring the individual and group levels independently, more social individual and group social phenotypes are mostly costly for individual fitness in this system [13–15,46–48,50], we predicted negative selection (i.e. selection for less sociality) for the connectivity, closeness and clustering social traits at both levels. We predicted positive selection (i.e. selection for being more social) for breakability traits at both phenotypic levels given more socially embedded individuals and less breakable groups (into two or more social groups) have some individual fitness benefits (philopatry [46], hibernation survival [15]). Lastly, we predicted the strength of selection for individual social behaviour would be stronger than selection for group social structure given the former has more known fitness consequences in this system (when explored independently) and because of the common critique of multilevel selection that the strength of selection for individuals is stronger than for groups [24,52].

2. Material and methods

(a) Study system

We used a 19 year dataset (2003–2022) on wild, free-living yellow-bellied marmots (*M. flaviventris*) studied at and around the Rocky Mountain Biological Laboratory in Colorado (38°57' N, 106°59' W; ca 2895 m above sea level). Yellow-bellied marmots are hibernating, harem polygynous, facultatively social ground-dwelling squirrels with matrilineal colony structures. It is important to note that these marmots are not cooperative breeders. This population is active for around 5 months annually (mid-April to mid-September). Mating soon after emergence from hibernation, new pups emerge, and yearlings disperse in late June to early July. Annually, most males and about half of females disperse as yearlings, typically resulting in movement out of our study area [41]. Marmots were studied annually at seven colony sites spread across 5 km at the bottom of the valley. Colonies are grouped into higher and lower elevation sites (four are at higher and three are at lower elevation sites). Higher elevation sites are approximately 166 m higher and experience harsher weather conditions [43,44,53].

Throughout the active season, marmots were repeatedly live-trapped, and their social behaviour was systematically quantified from 2003 to 2021. Recapture rate is above 86% in all colonies for all sex and age classes considered [54]. In addition to unique ear tags, all individuals were marked with unique non-toxic dye marks on their dorsal pelage to allow accurate identification. Marmots were weighed when trapped and these data were used to predict 1 June and 15 August body mass (to estimate early and late season body condition) via a best linear unbiased predictions (BLUPs) model [55,56]. June mass reflects the energy trade-off between leftover energy from hibernation and available energy for spring reproduction, whereas August mass reflects gain in fat mass during the active season and predicts overwinter survival [41,55,56]. Data used in our BLUPs consisted of 25 979 observations across 4330 individuals and 58 years. Using BLUPs can lead to higher rates of Type 1 error mainly owing to a decrease in the effect-associated standard error and not an increase in the absolute value of the effect [57,58]. Martin & Pelletier [59] showed that BLUPs can make accurate body mass predictions when there are on average greater than three measurements of body mass per individual. Our data have a mean of 5.99 observations per individual (range: 1.0–107.0; median = 4.0). Mixed with our relatively large dataset used to produce the BLUP [60] and previous sociality research in this system using this BLUP [13,15,44], these factors facilitate the accuracy and reliability of our body mass BLUP. Importantly, mass is included in the model to correct for potential confounding effects and is not a part of the hypothesis of sociality and multilevel selection. The estimate associated with mass should be unbiased using BLUPs adequately to correct for the mass effect; however, the associated probability should be interpreted with care.

(b) Social networks

Detailed behavioural observational methodology and the ethogram are outlined in Blumstein *et al.* [46]. For social interactions, the initiator, recipient, location, time and type of each interaction were recorded, with most interactions (79%) occurring between identified individuals. The remaining 21% of interactions could not be identified because of the interacting individuals' posture or visual obstructions. We excluded these interactions from our data, which should not significantly influence our estimates of social structure [61,62]. We only included adults and yearlings because only these cohorts were present in spring, when social interactions were most common, and because pups emerge mid-season and primarily interact with their mother and each other [63]. We eliminated transients by excluding individuals observed or trapped fewer than five times in a given year [47–49,64]. Only interactions in April, May and June (ca 2.5 months time frame is when marmots emerge from hibernation/mate to when pups emerge from natal burrows) were used because this is when most social interactions occur and when we have the highest resolution of observation data (the growth of vegetation begins to impair observations as the summer progresses). Lastly, we focused on affiliative interactions (e.g. allogrooming, greeting, play) because they relate to fitness on both the individual and group levels [13–15,47–50] and affiliative interactions constituted 88% of all social interactions.

Marmots share space with a subset of all possible individuals within their colony area. We therefore defined social groups based on space-use overlap per year (two individuals observed using the same burrow or seen/trapped at the same location and

Table 1. The four social traits and corresponding analogous pair of individual and group-level social network measures are used to quantify the individual and group social phenotypes.

social trait	individual social trait	group social trait
connectivity	degree: number of social relationships	density: proportion of possible social relationships that are observed
closeness	closeness: number of social links to access all other individuals in the group	inverse average path length ^a : mean social distance between all individuals in the group
breakability	embeddedness: connectedness in their cluster and group	inverse cut points ^a : number of social relationships that if broken/lost result in two or more separate social groups
clustering	clustering coefficient: proportion of an individual's direct social partners that also interact (i.e. local transitivity)	transitivity: proportion of connected triads actualized

^aInverse values of cut points and average path length were used so that increase in any of the eight variables reflected an increase in sociality.

time within 1 day intervals). Simple-ratio pairwise association indices based on colony space-use overlap [65] were calculated with SOCPROG (version 2.9; [66]) and run through the random walk algorithm MapEquation [67–69] to assemble association indices and identify social groups (network isolates within an association index). While MapEquation assigns each individual to only one social group (per year in our case), this can exclude key social connections, such as those with adult males. Because adult males often mate with females from multiple matriline and have important interactions with members of multiple groups, we added adult males to each social group for which they had at least one social interaction with a member of that group to enable more accurate social network measures. However, each year, a male's network measures and fitness were only calculated from his originally assigned group.

From these spatially defined groups, directed and weighted social interaction matrices were constructed from affiliative social interactions for each group each year with the R (version 4.2.0; [70]) package 'igraph' (version 1.4.2 [67]). These affiliative social interaction matrices consisted of 42 369 social interactions between 1294 individuals (338 of whom were observed across multiple years). This operationalization produced 180 social groups with group sizes ranging from 2 to 35 individuals with a mean of 7.65 ± 5.92 (mean \pm s.d.). Individuals had an average of 66.23 ± 90.72 social interactions per year, ranging from 1 to 694. Within social groups, social interactions averaged 447.35 ± 653.18 , ranging from 2 to 4118.

We calculated (using 'igraph') four pairs of analogous individual and group social network measures to quantify the independent contributions of the individual and group phenotypes (table 1). These analogous network measure pairs quantify four core traits in human and non-human animal social networks (including yellow-bellied marmots) at the two levels of social network organization: connectivity, closeness, breakability and clustering [4,71]. Degree (how many social partners an individual has [4]) and density (proportion of possible social connections that are observed in a group [72]) are paired to quantify social connectivity. Closeness (social distance between all other individuals in the group [4]) and average path length (mean social distance between all individuals in the group [72]) quantify social closeness. Embeddedness (individual connectedness based on their direct and indirect relationships with their cluster and group [73]) and cut points (number of social links that if broken or lost result in two or more social groups of at least two individuals [74]) quantify social breakability. Local clustering coefficient (proportion of an individual's direct social partners that are also social partners; i.e. local transitivity [75]) and transitivity (proportion of possible social triads that are observed in a group; i.e. global transitivity [76]) quantify social clustering. After scaling (see §2d), we flipped the sign for average path length and cut points, so all social network measures' values correspond to increased sociality. This was done to facilitate interpretation and presentation as the directional slope in figure 1 can be interpreted as the direction of selection for all measures. We thus refer to inverse average path length and inverse cut points throughout.

The reliability of the social network measures is facilitated by our regular observations of marmot social groups (mean n observations per individual across years = 28.81, range of each year = 6.79–75.14) and low rate of unknown individuals involved in social interactions [61,62,77]. Social group size (measured as social network size) is associated with many group-level social network measures ([4]; e.g. density, cut points) that may mediate the strength of selection. Because density, inverse average path length and transitivity are already 'standardized' based on their equations, we manually standardized inverse cut points by social group size so that all four group-level measures account for social group size in their calculation [4]. This facilitates comparison between our group-level measures and with previous research in this system [13,14].

(c) Fitness measures

Summer survival was defined as when individuals were seen or trapped after 1 August or in subsequent years, and over-winter/hibernation survival was defined as those individuals having survived the summer being seen the following spring or in subsequent years [15]. For summer survival, only adults (>2 years old) were included in the analysis owing to uncertainty quantifying survival for yearlings because a majority dispersed [78]. Predation accounts for 98% of summer mortality [79], and poor body condition and winter snowpack are primary predictors of hibernation mortality [53,80]. Summer survival was paired with network measures from the current active season, and hibernation survival with network measures from the active season before winter.

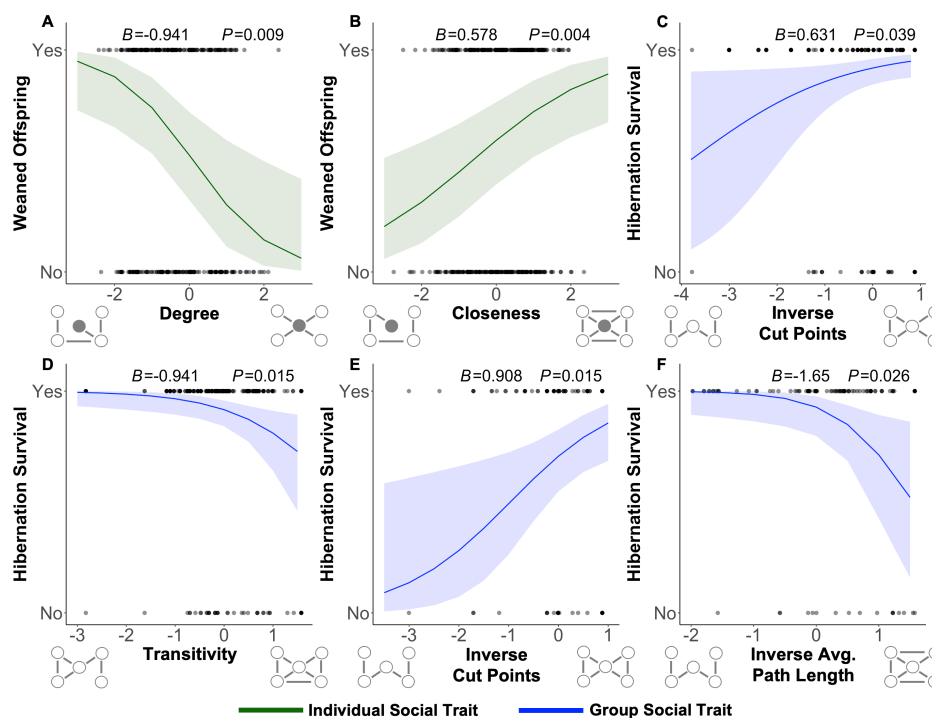


Figure 1. Selection gradients (plotted as marginal effects) for social traits representing the individual and group social phenotypes. (A,B) For the individual social phenotype (green), connectivity was under negative selection (A; fewer social partners) and closeness was under positive selection (B; less social distance from others) for adult female reproductive success. (C,D) For the group social phenotype (blue), adult females were under positive selection for breakability (C; more social ties required to break to split into two or more groups) and negative selection for clustering (D; lower proportion of connected transitive loop) for hibernation survival. (E) Group breakability was under positive selection for yearling male hibernation survival. (F) The group closeness trait (mean social distance between all individuals) was under negative selection for adult male hibernation survival. B represents the estimate, or strength of selection, and P represents the p-value.

Only female reproductive success is quantified because male reproductive success mostly depends on dominance, body condition and tenure length [81,82], which are difficult to quantify, and the smaller number of males in the population diminishes analysis power. We quantified two attributes of female reproductive success: (1) if a female successfully weaned offspring from the burrow, and (2) the number of offspring a mother weaned (if at least one pup was weaned) [14]. Behavioural observations and a comprehensive genetic pedigree [83,84] were used to assign offspring to mothers. This method does not account for pups that may have been born in the burrow but died before emergence (all pups are born in the burrow and emerge *ca* 30 days after birth [41]).

(d) Contextual analysis

We used contextual analysis, an extension of the Lande–Arnold selection analysis using partial regression to partition selection among levels (in this case, the individual and group social phenotypes; [30,51]). Contextual analysis is generalizable to multiple levels of organization provided that the lowest level included is the level at which fitness variation is being explored. Our contextual analysis differed from classic contextual analysis [30,38,85] since we used emergent group traits instead of using the mean of all individuals within a group.

Contextual analysis scales the phenotypic levels being explored because natural selection is relative, and thus it is the distribution of values, not the absolute values, of a trait that may reveal multilevel selection. Contextual analysis is sensitive to the scale of standardization and should be based on the biological and ecological processes that generate selection in the context of the study system [9,86]. Thus, we mean-variance standardized individual-level social network measures at the scale of each social group within each year (i.e. intra-group). Because group-level selection inherently operates across groups, we standardized group-level social network measures across all social groups across all 19 years (i.e. inter-group). We further mean-variance standardized group size on the global scale across all social groups across all years. Overall, the model can be expressed as:

$$w_{jkl} = \beta_0 + \beta_{w, \Delta P_i} \Delta P_{i,jkl} + \beta_{w, \Delta P_g} \Delta P_{g,jkl} + e_{jkl}$$

where w_{jkl} is the relative fitness of individual j in group k in year l , P_i are the social traits of an individual (individual social phenotype) and P_g are the social traits of the social group (group social phenotype). $\Delta P_{i,jkl}$ is the deviation of social trait for individual j from the mean of its group k in year l . $\Delta P_{g,jkl}$ is the deviation of the social group trait for group k of individual j in year l from the overall mean of social group trait across all groups and all years. $\beta_{w, \Delta P_i}$ and $\beta_{w, \Delta P_g}$ are the among-individual and among-group selection gradients. β_0 and e_{jkl} are the intercept and residual terms, respectively [35,51].

Owing to fundamental differences in life history strategies between age classes (adult and yearling) and sexes in this system [41], we fitted separate models for each cohort. Each model included the fitness measure as the response variable, the four pairs of analogous social network traits, social group size and valley location (higher or lower elevation) as fixed effects to account for known environmental and social variation. Individual ID and year were fitted as random effects to further account for environmental and demographic variation. Eight models were fitted in total, and the final models, after correcting fit for multicollinearity, met their respective assumptions. In R (version 4.2.0 [70]), models were fitted with 'lme4' (version 1.1-33 [87]) and assumptions were checked with the packages 'car' (version 3.1-2; [88,89]) and 'DHARMA' (version 0.4.6; [90]).

The set of models (adult females only) for both measures of reproductive success additionally included June body mass as a fixed effect given body condition's importance for reproductive success in this system [41]; neither model had multicollinearity issues. The model for whether or not a female weaned a litter was fitted with a binomial distribution with a bobyqa optimizer of 10 000 function evaluations and had 363 observations across 157 unique individuals. The model for the number of offspring (if a litter was weaned) was fitted with a Poisson distribution and had 191 observations across 98 unique individuals (electronic supplementary material, table S1).

The four sets of models (one for each age–sex cohort) for hibernation survival additionally included August body mass as a fixed effect [41]. These models were fitted with a binomial distribution and a bobyqa optimizer of 10 000 function evaluations. Multicollinearity was an issue between the network traits and thus the degree–density analogous pair was removed from all four models. With 19 years of data, the yearling male and female models had 119 and 209 observations, respectively (electronic supplementary material, table S2). The adult male model had 109 observations across 58 unique individuals and the adult female model had 324 observations across 134 unique individuals across 19 years (electronic supplementary material, table S3).

The two models for summer survival (male and female adults only) additionally included June body mass and a predation index as fixed effects [41]. The predation index was calculated by dividing the number of predators seen in a colony by the time spent observing that colony for that year [91]. These models were fitted with a binomial distribution and a bobyqa optimizer of 10 000 function evaluations. Both of these models had variance inflation factor (VIF) >5 for density. While contextual analysis accounts for the inherent non-independence of the individual and group social phenotypes by partitioning among the two phenotypic levels relative to each other, we used extra caution and took a more conservative approach by removing the degree–density analogous pair and running the models again. The summer model included 19 years of data with 138 observations across 80 unique individuals for males, and for females had 363 observations across 157 unique individuals (electronic supplementary material, table S4).

Because of the standardization approach we used as part of this contextual analysis, model estimates represent the magnitude of the selection gradient. We also controlled for multiple comparisons by calculating false discovery rate (FDR)-adjusted *p*-values [92–94] based on eight comparisons for the eight models, which still showed multilevel selection occurred (despite losing statistical significance for some group measures after the FDR adjustment; electronic supplementary material, table S5).

In summary, our contextual analysis (i) accounts for the inherent non-independence of the individual and group social phenotypes by partitioning among the two phenotypic levels relative to each other, and (ii) controls for other biologically relevant variables that, if excluded, could otherwise cause spurious associations. If we find significant evidence of selection, it is within the context of the biology of this system, not in isolation. Our contextual analysis permits the quantification of independent contributions to individual fitness at different social phenotypic levels, and between age–sex cohorts, while accounting for natural variation in environmental and demographic factors across a nearly two-decade dataset of a wild, free-living mammal.

3. Results and discussion

(a) Context-dependent multilevel selection

Our results revealed quantitative differences in the presence, strength and direction of multilevel selection for individual and group social phenotypes dependent on sex, age and fitness context (figure 1).

Individual connectivity was under negative selection (i.e. fewer direct social partners) and individual closeness was under positive selection (i.e. socially closer to all others in the group, directly and indirectly) for reproductive success in adult females (figure 1A,B; electronic supplementary material, table S1). This could represent an evolutionary balance in this system between the time and energetic costs of social relationships for reproductive success [47], while still maintaining closer social distance to others to maximize anti-predator benefits (e.g. hearing alarm calls), allowing more time and resources to be devoted to reproduction [50]. Given each species' time budget and selective pressures are unique, what individual social traits are under selection, and in what way, will be context-dependent based on the biology of that system. For example, in more gregarious species, more aligned selection across individual social traits may be observed.

Group clustering was under negative selection (i.e. there were fewer connected transitive loops) for adult female hibernation survival (figure 1D; electronic supplementary material, table S3), showing selection for adult females residing in less socially connected and clustered social groups. Group closeness was also under negative selection (i.e. larger social distance between all individuals) for adult male hibernation survival (figure 1F; electronic supplementary material, table S3). This shows selection for adult males residing in groups where all individuals are more socially distant. Yellow-bellied marmots are facultatively social, meaning that, compared with more gregarious species, marmots are more socially flexible and sometimes solitary [41]. Thus, marmots may not benefit from residing in more connected social groups, as supported by this negative selection for social structures that may facilitate increased rates of social interactions. However, group breakability was under positive selection

(i.e. more social ties required to break a group into two or more separate groups) for both adult female (figure 1C) and yearling male (figure 1E) hibernation survival (electronic supplementary material, tables S2 and S3). That is, selection favours individuals residing in more structurally cohesive and less fragmentable social groups because these groups may facilitate synchrony in hibernation onset [95] and because the breaking of social groups may reduce the number of individuals engaging in social hibernation, reducing subsequent thermoregulatory benefits and increasing the risk of death in hibernation (as seen in the alpine marmot (*Marmota marmota*) [95]). Collectively, this suggests group structures that facilitate phenological timing and thermoregulatory benefits are under positive selection, while structures that foster increased rates of interaction are under negative selection. The particular type of group social structure matters.

Compared with exploring the nested levels of social organization independently, which may lead to an inaccurate estimation of selection, this multilevel selection analysis revealed a more comprehensive picture of selection for sociality in this system, namely only the two group-level breakability (i.e. inverse cut points) results were found previously [15]. This difference could be attributable to multiple factors, such as different-sized datasets and statistical approaches across the studies; the individual social phenotype and reproductive success analysis was conducted over a decade ago [47]. Now with a 19 year dataset and a standardized statistical approach, we have a more accurate estimate of selection for sociality in this system. While exploring the levels of sociality independently provided a strong background to ask informed questions about the importance of multilevel selection for social traits, this multilevel selection approach, partitioning selection among the nested levels of organization, reveals a more informed and representative view of natural selection in this system, facilitating more accurate evolutionary predictions about sociality in this system.

Results somewhat align with our *a priori* predictions of negative selection for less connected, close and clustered social traits at both phenotypic levels and positive selection for less breakable traits at both phenotypic levels. However, individual closeness was under positive selection, contrary to our prediction. Neither social phenotypic level was under selection for summer survival (electronic supplementary material, table S4) or in female yearlings (electronic supplementary material, table S2), also contrary to our *a priori* predictions. Because the two social phenotypic levels are scaled, model estimates represent the magnitude of the selection gradient. Across all traits, age classes, and both sexes, the individual social phenotype had a mean selection gradient of 0.76 (0.26 s.e.) and the group social phenotype a mean of 1.03 (0.43 s.e.). This suggests that selection is just as strong, if not stronger, for the group social phenotype, contrary to our *a priori* prediction and to the classic criticism of multilevel selection. Collectively, these results support the importance of a multilevel selection framework to better understand the evolution and maintenance of sociality.

(b) Antagonistic multilevel selection on sociality

These results suggest not only that multilevel selection may be occurring in the wild for individual social behaviour and group social structure, but also that there is antagonistic multilevel selection *within* each of the two phenotypic levels and *between* the two phenotypic levels [9,33,34]. For adult females, there is both positive and negative selection within the individual (between connectivity and closeness) and within the group social phenotype (between closeness, breakability and clustering). There is also antagonistic selection between the two social phenotypes for adult females: positive selection for more social individuals potentially being counteracted by selection for less social groups, and *vice versa*. Antagonistic multilevel selection *within* and *between* social phenotypes may flatten or constrain the overall selection pressure for both social phenotypes [20,33], which could explain why sociality is not as highly heritable ($h^2 = 0.11$ [42]) or ubiquitously beneficial in this system as it is in other, more gregarious systems [12], such as humans (*Homo sapiens*; $h^2 = 0.47$ [96]) and rhesus macaques (*Macaca mulatta*; $h^2 = 0.84$ [97]).

In a captive population of forked fungus beetles (*Bolitotherus cornutus*) with a fixed social group size, Costello *et al.* [9] showed positive selection for more connected individual social phenotypes in males and negative selection for more connected group social phenotypes in females within the context of reproductive success, suggesting sexually antagonistic multilevel selection (as we also observe, but across different fitness contexts). That study quantified individual and group social phenotypes with an analogous social network trait and contextual analysis approach similar to ours here, but without quantifying the independent contributions of selection for multiple analogous pairs in the same model. Combined with our results, this shows multilevel selection for social phenotypes is present in Coleoptera and Rodentia. This evidence of multilevel selection in the wild and across species provides further impetus for experimental work to manipulate the composition of individual and group social phenotypes to further disentangle these nested, inherently non-independent, structures and to nail down causality and eliminate unmeasured variables [39]. For example, future work should leverage wild populations to identify, and experimental populations to confirm, the presence of multilevel selection and the factors that may mediate the strength and context of multilevel selection. Factors such as relatedness, phenotypic plasticity, individual and group age, group composition and sex ratios, and environmental factors (such as resource distribution) could increase or decrease the strength and direction of multilevel selection in other species. Future work should also consider nonlinear selection, in order to explore disruptive, stabilizing and correlational selection within and between levels of sociality. We did not have sufficient data to appropriately account for nonlinear selection in all models in this current study (for the models that did converge, the biological interpretation of the results was quantitatively similar to the results presented here).

(c) Multilevel adaptation in response to multilevel selection

Because the contextual analysis regressed individual fitness against the individual and group social phenotypes (i.e. partitioning selection among the two levels relative to each other), our results of selection at both levels add support to the distinct

and discrete nature of individual behaviour and group social structure despite being nested levels of biological organization [8–10,13–15,18,20,21,37]. Given we now show selection at multiple phenotypic levels in addition to heritability [42] for social behaviour in this system, this allows the potential for an evolutionary response. Evidence for multilevel adaptation in response to multilevel selection would require demonstration of intergenerational changes in the units of replication that underlie social behaviours [51]. For example, in fruit flies (*Drosophila melanogaster*), the gene *CG14109* influences betweenness centrality [98], a key property of social networks that indicates an individual's importance in facilitating disease spread, communication and cohesion within the network. Future work should leverage genomic tools to link individual and group phenotypes with candidate genes, like *CG14109*, and to track intergenerational changes in the units of replication that underlie social behaviours.

Given sociality inherently involves interactions between varying phenotypes, indirect genetic effects likely play a role in the evolution of sociality [39,99,100]. Furthermore, in these nested levels of biological organization, selection at one level may accelerate or constrain evolution at another level [39], which is further complicated by the antagonistic multilevel selection within and between social phenotypes we show here. The next steps to better understanding the evolution of societies across species should identify the genetic architecture and correlations underpinning multilevel selection, incorporating both direct and indirect genetic effects in a unified empirical framework [99].

(d) Group living and social phenotypes

Interestingly, we found no selection for either social phenotype in the context of summer survival. Predation accounts for 98% of summer mortality in this system [79] and the evolutionary origins of yellow-bellied marmot group living partly involve predator avoidance [41]. This suggests selection driving group living can be, and is in some cases, different from the subsequent selection pressures driving how social individuals are and the structure of the social group. Again, work in more gregarious species where the evolutionary origins of group living may be attributable to grooming, heat retention, or resource acquisition could find that the selection for group living and the individual and group social phenotypes are more aligned.

A common argument against group selection (and thus multilevel selection) used since the 1960s focuses on differences in individual lifespan versus group duration [24,52,101]: where individual lifespans are shorter than group duration, selection proceeds faster at the level of the individual. However, in yellow-bellied marmots, dispersal, mortality and births ensure that groups restructure annually. Individuals surviving to adulthood have a mean lifespan of 4.07 years and are thus members of multiple social groups in their lifetime [41]. In this population, individual lifespan is longer than group turnover and thus this argument does not apply. Indeed, the same logic would predict stronger selection for the group phenotype, which we found.

4. Conclusion

By leveraging a 19 year dataset from a wild, free-living mammal and contextual analysis to partition selection relatively among two nested levels of sociality, we provide evidence that suggests that multilevel selection is a potentially important and overlooked force in the evolution of sociality. While we have identified multilevel selection, we have not demonstrated evolution in response to this selection, which will require additional work. We recognize that experimental tests are required to formalize causality and eliminate unmeasured variables affecting fitness and social phenotypes beyond what our models controlled for.

Nevertheless, our detection of multilevel selection in the wild provides the impetus to develop these multilevel selection experiments across systems. Considering levels of sociality independently may be reductive and lead to Type I or Type II errors; when we do not incorporate a multilevel selection perspective, we may incorrectly attribute natural selection to the levels of organization not actually under direct natural selection. Instead, sociality can be quantified as a nested, multidimensional phenotype of which different levels can be selected on and change with age and sex. Compared with exploring levels of sociality independently and given contextual analysis partitions selection among phenotypic levels relative to each other, multilevel selection may more comprehensively represent how natural selection acts on the social interactions and structures that arise from group living. We provide empirical evidence to spark further study in this classic debate in biology.

Ethics. This study uses a long-term dataset. Marmots were originally studied under the research protocol ARC 2001-191-01 (approved by the UCLA Animal Care Committee on 13 May 2002 and renewed annually) and protocols approved by the Rocky Mountain Biological Laboratory, and trapped under permits issued annually by the Colorado Department of Parks and Wildlife (TR-917).

Data accessibility. Data and code to reproduce analyses are available at [102].

Supplementary material is available online [103].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. C.S.P.: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, supervision, validation, visualization, writing—original draft, writing—review and editing; J.G.A.M.: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, writing—review and editing; D.T.B.: conceptualization, data curation, funding acquisition, investigation, project administration, resources, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed herein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This work was supported by the UCLA, American Society of Mammalogists, Animal Behaviour Society, Rocky Mountain Biological Laboratory (RMBL), the Natural Sciences and Engineering Research Council of Canada, the University of Ottawa, National Geographic Society, and the U.S. National Science Foundation (NSF IDBR-0754247 and DEB-1119660 and 1557130 to D.T.B.; DBI 0242960, 07211346, 1226713, and 1755522 to RMBL).

Acknowledgements. We thank the ‘marmoteers’ who participated in collecting the long-term data, the Rocky Mountain Biological Laboratory (RMBL) for providing field facilities and funding critical to long-term research, and the editor and reviewers for their helpful comments.

References

- Alexander RD. 1974 The evolution of social behavior. *Annu. Rev. Ecol. Syst.* **5**, 325–383. (doi:10.1146/annurev.es.05.110174.001545)
- Wilson EO. 1975 *Sociobiology: the new synthesis*. Cambridge, MA: Harvard University Press.
- Hinde RA. 1976 Interactions, relationships and social structure. *Man* **11**, 1–17. (doi:10.2307/2800384)
- Wasserman S, Faust K. 1994 *Social network analysis: methods and applications*. Cambridge, UK: Cambridge University Press.
- Kurvers RHJM, Krause J, Croft DP, Wilson ADM, Wolf M. 2014 The evolutionary and ecological consequences of animal social networks: emerging issues. *Trends Ecol. Evol.* **29**, 326–335. (doi:10.1016/j.tree.2014.04.002)
- Croft DP, Darden SK, Wey TW. 2016 Current directions in animal social networks. *Curr. Opin. Behav. Sci.* **12**, 52–58. (doi:10.1016/j.cobeha.2016.09.001)
- Kappeler PM. 2019 A framework for studying social complexity. *Behav. Ecol. Sociobiol.* **73**, 13. (doi:10.1007/s00265-018-2601-8)
- Cantor M *et al.* 2021 The importance of individual-to-society feedbacks in animal ecology and evolution. *J. Anim. Ecol.* **90**, 27–44. (doi:10.1111/1365-2656.13336)
- Costello RA, Cook PA, Brodie ED, Formica VA, Brodie III ED. 2023 Multilevel selection on social network traits differs between sexes in experimental populations of forked fungus beetles. *Evolution* **77**, 289–303. (doi:10.1093/evolut/qpac012)
- Royle NJ, Pike TW, Heeb P, Richner H, Kölliker M. 2012 Offspring social network structure predicts fitness in families. *Proc. R. Soc. B* **279**, 4914–4922. (doi:10.1098/rspb.2012.1701)
- Fisher DN, McAdam AG. 2017 Social traits, social networks and evolutionary biology. *J. Evol. Biol.* **30**, 2088–2103. (doi:10.1111/jeb.13195)
- Snyder-Mackler N *et al.* 2020 Social determinants of health and survival in humans and other animals. *Science* **368**, eaax9553. (doi:10.1126/science.aax9553)
- Philson CS, Todorov SM, Blumstein DT. 2022 Marmot mass gain rates relate to their group's social structure. *Behav. Ecol.* **33**, 115–125. (doi:10.1093/beheco/arab114)
- Philson CS, Blumstein DT. 2023 Group social structure has limited impact on reproductive success in a wild mammal. *Behav. Ecol.* **34**, 89–98. (doi:10.1093/beheco/ara102)
- Philson CS, Blumstein DT. 2023 Emergent social structure is typically not associated with survival in a facultatively social mammal. *Biol. Lett.* **19**, 20220511. (doi:10.1098/rsbl.2022.0511)
- Sapolsky RM. 2004 Social status and health in humans and other animals. *Annu. Rev. Anthropol.* **33**, 393–418. (doi:10.1146/annurev.anthro.33.070203.144000)
- Wilson DS, Sober E. 1994 Reintroducing group selection to the human behavioral sciences. *Behav. Brain Sci.* **17**, 585–608. (doi:10.1017/s0140525x00036104)
- Okasha S. 2005 Multilevel selection and the major transitions in evolution. *Philos. Sci.* **72**, 1013–1025. (doi:10.1086/508102)
- Wilson DS, Wilson EO. 2007 Rethinking the theoretical foundation of sociobiology. *Q. Rev. Biol.* **82**, 327–348. (doi:10.1086/522809)
- Goodnight C. 2013 On multilevel selection and kin selection: contextual analysis meets direct fitness. *Evolution* **67**, 1539–1548. (doi:10.1111/j.1558-5646.2012.01821.x)
- Gardner A. 2015 The genetical theory of multilevel selection. *J. Evol. Biol.* **28**, 305–319. (doi:10.1111/jeb.12566)
- Wilson DS, Madhavan G, Gelfand MJ, Hayes SC, Atkins PWB, Colwell RR. 2023 Multilevel cultural evolution: from new theory to practical applications. *Proc. Natl Acad. Sci. USA* **120**, e2218222120. (doi:10.1073/pnas.2218222120)
- Wynne-Edwards VC. 1962 *Animal dispersion in relation to social behaviour*. Edinburgh, UK: Oliver & Boyd.
- Maynard Smith J. 1964 Group selection and kin selection. *Nat. New Biol.* **201**, 1145–1147. (doi:10.1038/2011145a0)
- Okasha S. 2006 *Evolution and the levels of selection*. Oxford, UK: Oxford University Press.
- Nowak MA, Tarnita CE, Wilson EO. 2010 The evolution of eusociality. *Nature* **466**, 1057–1062. (doi:10.1038/nature09205)
- Yu L, Boström C, Franzénburg S, Bayer T, Dagan T, Reusch TBH. 2020 Somatic genetic drift and multilevel selection in a clonal seagrass. *Nat. Ecol. Evol.* **4**, 952–962. (doi:10.1038/s41559-020-1196-4)
- van Vliet S, Doebeli M. 2019 The role of multilevel selection in host microbiome evolution. *Proc. Natl Acad. Sci. USA* **116**, 20591–20597. (doi:10.1073/pnas.1909790116)
- Johnson NC, Gibson KS. 2021 Understanding multilevel selection may facilitate management of arbuscular mycorrhizae in sustainable agroecosystems. *Front. Plant Sci.* **11**, 627345. (doi:10.3389/fpls.2020.627345)
- Goodnight CJ, Schwartz JM, Stevens L. 1992 Contextual analysis of models of group selection, soft selection, hard selection, and the evolution of altruism. *Am. Nat.* **140**, 743–761. (doi:10.1086/285438)
- Hamilton WD. 1964 The genetical evolution of social behaviour. II. *J. Theor. Biol.* **7**, 17–52. (doi:10.1016/0022-5193(64)90039-6)
- Price GR. 1972 Extension of covariance selection mathematics. *Ann. Hum. Genet.* **35**, 485–490. (doi:10.1111/j.1469-1809.1957.tb01874.x)
- Weinig C, Johnston JA, Willis CG, Maloof JN. 2007 Antagonistic multilevel selection on size and architecture in variable density settings. *Evolution* **61**, 58–67. (doi:10.1111/j.1558-5646.2007.00005.x)
- Eldakar OT, Wilson DS, Dlugos MJ, Pepper JW. 2010 The role of multilevel selection in the evolution of sexual conflict in the water strider *Aquarius remigis*. *Evolution* **64**, 3183–3189. (doi:10.1111/j.1558-5646.2010.01087.x)
- Fisher DN, Boutin S, Dantzer B, Humphries MM, Lane JE, McAdam AG. 2017 Multilevel and sex-specific selection on competitive traits in North American red squirrels. *Evolution* **71**, 1841–1854. (doi:10.1111/evo.13270)
- Formica V, Donald H, Marti H, Irgebay Z, Brodie E III. 2021 Social network position experiences more variable selection than weaponry in wild subpopulations of forked fungus beetles. *J. Anim. Ecol.* **90**, 168–182. (doi:10.1111/1365-2656.13322)
- Solomon-Lane TK, Pradhan DS, Willis MC, Grober MS. 2015 Agonistic reciprocity is associated with reduced male reproductive success within harem social networks. *Proc. R. Soc. B* **282**, 20150914. (doi:10.1098/rspb.2015.0914)
- Heisler IL, Damuth J. 1987 A method for analyzing selection in hierarchically structured populations. *Am. Nat.* **130**, 582–602. (doi:10.1086/284732)
- Bijma P, Wade MJ. 2008 The joint effects of kin, multilevel selection and indirect genetic effects on response to genetic selection. *J. Evol. Biol.* **21**, 1175–1188. (doi:10.1111/j.1420-9101.2008.01550.x)
- Leigh Jr EG. 2010 The group selection controversy. *J. Evol. Biol.* **23**, 6–19. (doi:10.1111/j.1420-9101.2009.01876.x)
- Armitage KB. 2014 *Marmot biology: sociality, individual fitness, and population dynamics*. Cambridge, UK: Cambridge University Press.
- Lea AJ, Blumstein DT, Wey TW, Martin JGA. 2010 Heritable victimization and the benefits of agonistic relationships. *Proc. Natl Acad. Sci. USA* **107**, 21587–21592. (doi:10.1073/pnas.1009882107)

43. Blumstein DT, Ozgul A, Yovovich V, Van Vuren DH, Armitage KB. 2006 Effect of predation risk on the presence and persistence of yellow-bellied marmot (*Marmota flaviventris*) colonies. *J. Zool.* **270**, 132–138. (doi:10.1111/j.1469-7998.2006.00098.x)
44. Maldonado-Chaparro AA, Martin JGA, Armitage KB, Oli MK, Blumstein DT. 2015 Environmentally induced phenotypic variation in wild yellow-bellied marmots. *J. Mammal.* **96**, 269–278. (doi:10.1093/jmammal/gyu006)
45. Pfau M, Degregori S, Johnson G, Tennenbaum SR, Barber PH, Philson CS, Blumstein DT. 2023 The social microbiome: gut microbiome diversity and abundance are negatively associated with sociality in a wild mammal. *R. Soc. Open Sci.* **10**, 231305. (doi:10.1098/rsos.231305)
46. Blumstein DT, Wey TW, Tang K. 2009 A test of the social cohesion hypothesis: interactive female marmots remain at home. *Proc. R. Soc. B* **276**, 3007–3012. (doi:10.1098/rspb.2009.0703)
47. Wey TW, Blumstein DT. 2012 Social attributes and associated performance measures in marmots: bigger male bullies and weakly affiliating females have higher annual reproductive success. *Behav. Ecol. Sociobiol.* **66**, 1075–1085. (doi:10.1007/s00265-012-1358-8)
48. Yang WJ, Maldonado-Chaparro A, Blumstein DT. 2017 A cost of being amicable in a hibernating marmot. *Behav. Evol.* **28**, 11–19. (doi:10.1093/beheco/arw125)
49. Blumstein DT, Williams DM, Lim AN, Kroeger S, Martin JGA. 2018 Strong social relationships are associated with decreased longevity in a facultatively social mammal. *Proc. R. Soc. B* **285**, 20171934. (doi:10.1098/rspb.2017.1934)
50. Montero AP, Williams DM, Martin JGA, Blumstein DT. 2020 More social female yellow-bellied marmots, *Marmota flaviventer*, have enhanced summer survival. *Anim. Behav.* **160**, 113–119. (doi:10.1016/j.anbehav.2019.12.013)
51. Lande R, Arnold SJ. 1983 The measurement of selection on correlated characters. *Evolution* **37**, 1210–1226. (doi:10.1111/j.1558-5646.1983.tb00236.x)
52. Williams GC. 1966 *Adaptation and natural selection*. Princeton, NJ: Princeton University Press.
53. Van Vuren DH. 1991 Predation on yellow-bellied marmots (*Marmota flaviventris*). *Am. Midl. Nat.* **145**, 94–100.
54. Ozgul A, Armitage KB, Blumstein DT, Oli MK. 2006 Spatiotemporal variation in survival rates: implications for population dynamics of yellow-bellied marmots. *Ecology* **87**, 1027–1037. (doi:10.1890/0012-9658(2006)87[1027:svsri]2.0.co;2)
55. Ozgul A, Childs DZ, Oli MK, Armitage KB, Blumstein DT, Olson LE, Tuljapurkar S, Coulson T. 2010 Coupled dynamics of body mass and population growth in response to environmental change. *Nature* **466**, 482–485. (doi:10.1038/nature09210)
56. Kroeger SB, Blumstein DT, Armitage KB, Reid JM, Martin JGA. 2018 Age, state, environment, and season dependence of senescence in body mass. *Ecol. Evol.* **8**, 2050–2061. (doi:10.1002/ece3.3787)
57. Hadfield JD, Wilson AJ, Garant D, Sheldon BC, Kruuk LEB. 2010 The misuse of BLUP in ecology and evolution. *Am. Nat.* **175**, 116–125. (doi:10.1086/648604)
58. Houslay TM, Wilson AJ. 2017 Avoiding the misuse of BLUP in behavioural ecology. *Behav. Ecol.* **28**, 948–952. (doi:10.1093/beheco/arx023)
59. Martin JGA, Pelletier F. 2011 Measuring growth patterns in the field: effects of sampling regime and methods on standardized estimates. *Can. J. Zool.* **89**, 529–537. (doi:10.1139/z11-018)
60. Dingemanse NJ, Moiron M, Araya-Ajoy YG, Mouchet A, Abbey-Lee RN. 2020 Individual variation in age-dependent reproduction: fast explorers live fast but senesce young? *J. Anim. Ecol.* **89**, 601–613. (doi:10.1111/1365-2656.13122)
61. Silk MJ, Jackson AL, Croft DP, Colhoun K, Bearhop S. 2015 The consequences of unidentifiable individuals for the analysis of an animal social network. *Anim. Behav.* **104**, 1–11. (doi:10.1016/j.anbehav.2015.03.005)
62. Sánchez-Tójar A, Schroeder J, Farine DR. 2018 A practical guide for inferring reliable dominance hierarchies and estimating their uncertainty. *J. Anim. Ecol.* **87**, 594–608. (doi:10.1111/1365-2656.12776)
63. Nowicki S, Armitage KB. 1979 Behaviour of juvenile yellow-bellied marmots: play and social integration. *Z. Tierpsychol.* **51**, 85–105. (doi:10.1111/j.1439-0310.1979.tb00674.x)
64. Fung H, Maldonado-Chaparro A, Blumstein DT. 2015 Are social attributes associated with alarm calling propensity? *Behav. Ecol.* **26**, 587–592. (doi:10.1093/beheco/aru235)
65. Cairns SJ, Schwager SJ. 1987 A comparison of association indices. *Anim. Behav.* **35**, 1454–1469. (doi:10.1016/S0003-3472(87)80018-0)
66. Whitehead H. 2009 SOCPROG programs: analysing animal social structures. *Behav. Ecol. Sociobiol.* **63**, 765–778. (doi:10.1007/s00265-008-0697-y)
67. Csardi G, Nepusz T. 2006 The igraph software package for complex network research. *Int. J. Commun. Syst.* **1695**, 1–9.
68. Rosvall M, Bergstrom CT. 2008 Maps of random walks on complex networks reveal community structure. *Proc. Natl Acad. Sci. USA* **105**, 1118–1123. (doi:10.1073/pnas.0706851105)
69. Rosvall M, Axelsson D, Bergstrom CT. 2009 The map equation. *Eur. Phys. J. Spec. Top.* **178**, 13–23. (doi:10.1140/epjst/e2010-01179-1)
70. R Development Core Team. 2023 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See <https://www.R-project.org/>.
71. Krause J, Lusseau D, James R. 2009 Animal social networks: an introduction. *Behav. Ecol. Sociobiol.* **63**, 967–973. (doi:10.1007/s00265-009-0747-0)
72. Opsahl T, Panzarasa P. 2009 Clustering in weighted networks. *Soc. Networks* **31**, 155–163. (doi:10.1088/1126-6708/2009/02/002)
73. Moody J, White DR. 2003 Structural cohesion and embeddedness: a hierarchical concept of social groups. *Am. Sociol. Rev.* **68**, 103–127. (doi:10.1177/000312240306800105)
74. Borgatti SP. 2006 Identifying sets of key players in a social network. *Comput. Math. Organ. Theory* **12**, 21–34. (doi:10.1007/s10588-006-7084-x)
75. Croft DP, James R, Krause J. 2008 *Exploring animal social networks*. Princeton, NJ: Princeton University Press.
76. Wey TW, Jordán F, Blumstein DT. 2019 Transitivity and structural balance in marmot social networks. *Behav. Ecol. Sociobiol.* **73**, 88. (doi:10.1007/s00265-019-2699-3)
77. Davis GH, Crofoot MC, Farine DR. 2018 Estimating the robustness and uncertainty of animal social networks using different observational methods. *Anim. Behav.* **141**, 29–44. (doi:10.1016/j.anbehav.2018.04.012)
78. Van Vuren D, Armitage KB. 1991 Duration of snow cover and its influence on life-history variation in yellow-bellied marmots. *Can. J. Zool.* **69**, 1755–1758. (doi:10.1139/z91-244)
79. Van Vuren DH. 2001 Predation on yellow-bellied marmots (*Marmota flaviventris*). *Am. Midl. Nat.* **145**, 94–100. (doi:10.1674/0003-0031(2001)145[0094:POYBMM]2.0.CO;2)
80. Armitage KB, Blumstein DT, Woods BC. 2003 Energetics of hibernating yellow-bellied marmots (*Marmota flaviventris*). *Comp. Biochem. Physiol. A* **134**, 101–114. (doi:10.1016/S1095-6433(02)00219-2)
81. Armitage KB. 1998 Reproductive strategies of yellow-bellied marmots: energy conservation and differences between the sexes. *J. Mammal.* **79**, 385–393. (doi:10.2307/1382969)
82. Huang B, Wey TW, Blumstein DT. 2011 Correlates and consequences of dominance in a social rodent. *Ethology* **117**, 573–585. (doi:10.1111/j.1439-0310.2011.01909.x)
83. Blumstein DT, Lea AJ, Olson LE, Martin JGA. 2010 Heritability of anti-predatory traits: vigilance and locomotor performance in marmots. *J. Evol. Biol.* **23**, 879–887. (doi:10.1111/j.1420-9101.2010.01967.x)
84. Olson LE, Blumstein DT. 2010 Applying the coalitionary-traits metric: sociality without cooperation in male yellow-bellied marmots. *Behav. Ecol.* **21**, 957–965. (doi:10.1093/beheco/arq094)
85. Goodnight CJ, Stevens L. 1997 Experimental studies of group selection: what do they tell us about group selection in nature? *Am. Nat.* **150**, S59–S79. (doi:10.1086/286050)
86. De Lisle SP, Svensson EI. 2017 On the standardization of fitness and traits in comparative studies of phenotypic selection. *Evolution* **71**, 2313–2326. (doi:10.1111/evo.13325)

87. Bates D, Maechler M, Bolker B, Walker S. 2003 *lme4: Linear mixed-effects models using 'Eigen' and S4*. R package version 1.1-10. See <http://CRAN.R-project.org/package=lme4>.
88. Fox J, Monette G. 1992 Generalized collinearity diagnostics. *J. Am. Stat. Assoc.* **87**, 178–183. (doi:10.1080/01621459.1992.10475190)
89. Fox J, Weisberg S. 2019 *An R companion to applied regression*, 3rd edn. Thousand Oaks, CA: Sage Publications.
90. Hartig F. 2022 DHARMA: residual diagnostics for hierarchical (multi-level/mixed) regression models. *R Package Version 0.4.6* (doi:10.32614/CRAN.package.DHARMA)
91. Blumstein DT, Sanchez M, Philson CS, Bliard L. 2023 Is flight initiation distance associated with longer-term survival in yellow-bellied marmots, *Marmota flaviventris*? *Anim. Behav.* **202**, 21–28. (doi:10.1016/j.anbehav.2023.05.013)
92. Shaffer JP. 1995 Multiple hypothesis testing. *Annu. Rev. Psychol.* **46**, 561–584. (doi:10.1146/annurev.ps.46.020195.003021)
93. Verhoeven KJF, Simonsen KL, McIntyre LM. 2005 Implementing false discovery rate control: increasing your power. *Oikos* **108**, 643–647. (doi:10.1111/j.0030-1299.2005.13727.x)
94. Pike N. 2011 Using false discovery rates for multiple comparisons in ecology and evolution. *Methods Ecol. Evol.* **2**, 278–282. (doi:10.1111/j.2041-210x.2010.00061.x)
95. Arnold W. 1993 Social evolution in marmots and the adaptive value of joint hibernation. *Verh. Dt. Zool. Ges.* **86**, 79–99.
96. Fowler JH, Dawes CT, Christakis NA. 2009 Model of genetic variation in human social networks. *Proc. Natl Acad. Sci. USA* **106**, 1720–1724. (doi:10.1073/pnas.0806746106)
97. Brent LJN, Heilbronner SR, Horvath JE, Gonzalez-Martinez J, Ruiz-Lambides A, Robinson AG, Skene JHP, Platt ML. 2013 Genetic origins of social networks in rhesus macaques. *Scient. Rep.* **3**, 1042. (doi:10.1038/srep01042)
98. Rooke R, Krupp JJ, Rasool A, Golemic M, Stewart M, Schneider J, Levine JD. 2024 The gene "degrees of kevin bacon" (*dokb*) regulates a social network behaviour in *Drosophila melanogaster*. *Nat. Commun.* **15**, 3339. (doi:10.1038/s41467-024-47499-8)
99. Radersma R. 2021 Estimating heritability of social phenotypes from social networks. *Methods Ecol. Evol.* **12**, 42–53. (doi:10.1111/2041-210x.13499)
100. McLean EM, Moorad JA, Tung J, Archie EA, Alberts SC. 2023 Genetic variance and indirect genetic effects for affiliative social behavior in a wild primate. *Evolution* **77**, 1607–1621. (doi:10.1093/evolut/qpaa066)
101. Wade MJ. 1977 An experimental study of group selection. *Evolution* **31**, 134–153. (doi:10.1111/j.1558-5646.1977.tb00991.x)
102. Philson CS, Martin JGA, Blumstein DT. 2024 Data for: Multilevel selection on individual and group social behaviour in the wild. *Open Science Framework* (doi:10.17605/OSF.IO/H5QSV)
103. Philson CS, Martin JG, Blumstein D. 2025 Supplementary material from: Multilevel selection on individual and group social behaviour in the wild. Figshare. (doi:10.6084/m9.figshare.c.7712226)