Microevolutionary consequences of social structure in wild spotted hyenas

Kasha Strickland[1*], Oliver P. Höner[2,3], Larissa S. Arantes[4,5], Joel L Pick[1], Kenneth Aase[6], Loeske E. B. Kruuk[1], Eve Davidian[3]

1. Institute for Ecology and Evolution, School of Biological Sciences, University of Edinburgh, UK
2. Department of Evolutionary Ecology, Leibniz Institute for Zoo and Wildlife Research, Berlin, Germany
3. Evolutionary Anthropology Team, Institute of Evolutionary Sciences of Montpellier (ISEM), University of Montpellier, CNRS, IRD, Montpellier, France
4. Department of Evolutionary Genetics, Leibniz Institute for Zoo and Wildlife Research, Berlin, Germany
5. Berlin Center for Genomics in Biodiversity Research (BeGenDiv), Berlin, Germany
6. Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway

* corresponding author: kasha.strickland@ed.ac.uk

Social structure - arising from non-random associations and interactions among conspecifics - is a defining feature of most animal populations, yet evolutionary theory typically assumes genetic and social homogeneity. This disconnect limits our ability to predict how natural populations evolve. We combined nearly 30 years of behavioural, life-history, and genomic data from wild spotted hyenas (*Crocuta crocuta*) in Tanzania's Ngorongoro Crater to test how social structure and male-biased dispersal shape genetic structure and the rate of adaptive evolution. Genome-wide analyses revealed subtle but consistent genetic differentiation among clans, reflecting cryptic population genetic structure. These differences were best explained by asymmetric dispersal between clans rather than geographic distance, indicating that social processes drive population stratification. Individuals with more immigrant ancestry had higher fitness, demonstrating adaptive benefits of gene flow and no evidence for selection against immigrants. Finally, additive genetic variance in fitness differed among clans, showing that evolutionary potential is unevenly distributed across the population. Together, these findings reveal how social structure and non-random dispersal generate hidden genetic structure and result in heterogenous rates of adaptive evolution. Our results underscore the need to integrate social structure in evolutionary models to better predict microevolutionary dynamics in the wild.

Keywords: dispersal, social behaviour, social structure, genetic variance, fitness, gene flow

Introduction

Social behaviours have evolved across a wide range of animal species and result in diverse forms of social structure, ranging from transient pairwise associations to highly stable family groups and multilevel cooperative breeding systems (Clutton-Brock 2021; He *et al.* 2019). Such structures are widespread and are characterised by non-random, differentiated interactions or associations within populations, resulting in the uneven distribution of individuals and groups across space and time. This structuring ultimately affects how individuals move, interact, and reproduce, with far-reaching implications for how microevolutionary processes such as gene flow, genetic drift, and natural selection operate (Kurvers *et al.* 2014). Yet, despite decades of research into the evolution of social behaviour (Alexander 1974; Clutton-Brock 2002; Hamilton 1964; Snyder-Mackler *et al.* 2020), key theoretical models used in evolutionary biology typically ignore the social structure that defines most animal populations (Fisher 1930; Robertson & Lewontin 1968; Walsh & Lynch 2018). Improving our understanding of how social structure and evolution intersect will be key to building more accurate and generalisable models of microevolution in natural populations.

Social structure in a population is an emergent property of individual or group level behaviours, such as mating, grouping, and dispersal (Kurvers *et al.* 2014). Together, these behaviours mediate the relationship between social and genetic structure in populations. For instance, sex-biased dispersal and kin-biased interactions are common across social animals (Gardner *et al.* 2001; Li & Kokko 2019; Morinay *et al.* 2025; Pereira *et al.* 2023) and jointly shape how genetic variation is distributed across space and time. Male- or female-biased dispersal is typically considered an evolutionary response to the risk of inbreeding (Perrin & Mazalov 2000; Pusey 1987), but dispersal decisions are also shaped by social and ecological context. For instance, dispersal may occur together with kin, or be guided by mate choice, resource availability, or environmental conditions (Clobert *et al.* 2009; McPeek & Holt 1992; Peniston *et al.* 2024). These non-random patterns of movement create spatiotemporal heterogeneity in gene flow, potentially generating cryptic genetic structure within populations (Parreira & Chikhi 2015). Although the role of dispersal and social structure in shaping genetic structure is well recognised in behavioural ecology and population genetics (Bowler & Benton 2005; Clutton-Brock & Lukas 2012; Parreira & Chikhi 2015), empirical evidence linking these processes to contemporary evolutionary dynamics remains rare, partly due to the challenges of collecting sufficient data on dispersal in wild populations. Moreover, while social structure likely has important consequences for microevolutionary processes such as genetic drift and natural selection (Frean *et al.* 2013; Kurvers *et al.* 2014; Waples 2010), substructures are rarely accounted for. Instead, populations are often treated as genetically homogeneous units, overlooking the evolutionary impact of social structure.

Violations of the assumption of homogeneity may be especially problematic when estimating the rate of adaptive evolution in structured populations (Barton & Clark 1990). Fisher's fundamental

theorem of natural selection relates the rate of adaptive evolution to the additive genetic variance in fitness (Fisher 1930), but this model assumes a genetically unstructured, homogenous population. If genetic variation is unevenly distributed among social groups, and if selection varies among these groups due to ecological or social heterogeneity, then the rate of adaptive evolution may differ across the population (Bonnet *et al.* 2022; Montiglio *et al.* 2013). Moreover, gene flow between subunits such as family groups may enhance or reduce fitness, depending on whether immigrants introduce beneficial genetic variation or disrupt local adaptation (Räsänen & Hendry 2008; Wolak & Reid 2017). By ignoring these complexities, standard evolutionary models may under- or overestimate the capacity for microevolutionary change (Pujol *et al.* 2018). The extent to which we can predict and observe evolutionary change in the wild therefore depends on integrating the social environment in which wild animals live to empirical and theoretical research.

To examine how social structure and non-random dispersal impact microevolutionary processes, we analysed 29 years of behavioural, life-history and genomic data from a well-characterised population of spotted hyenas (*Crocuta crocuta*) living in Tanzania's Ngorongoro Crater. This population is part of the wider hyena metapopulation of the Greater Serengeti-Mara ecosystem (Davidian et al. 2016) and has been studied on an almost-daily basis since 1996. The resulting dataset includes detailed behavioural and life-history data from over 3200 individuals and a genetic pedigree spanning a maximum of nine generations. Spotted hyenas are large, social carnivores that form multilevel fission–fusion societies called clans, each consisting of multiple matrilines and up to 130 individuals (East & Hofer 2001; Frank 1986; Holekamp *et al.* 2012; Kruuk 1966; Wemmer 1973). Within clans, social structure is shaped by female-dominated dominance hierarchies, and dispersal is strongly male-biased: females are almost always philopatric, while most males emigrate from their natal clan to breed (Davidian *et al.* 2016; Höner *et al.* 2007). Male dispersal is non-random whereby males choose recipient clans with the highest number of young females (Höner *et al.* 2007), and closely related males may choose the same clan (Davidian & Höner 2022). Hyenas mate promiscuously within clans, but extra-group paternity is rare (Davidian *et al.* 2016), and the dominance hierarchy within each clan results in strong reproductive skew in both males and females (Engh *et al.* 2002; Holekamp *et al.* 1996, 2012; Höner *et al.* 2010). Crucially, previous work identified biologically meaningful levels of additive genetic variance in reproductive fitness in the population as a whole (Bonnet *et al.* 2022), suggesting ongoing contemporary adaptive evolution.

We leveraged this exceptionally rich dataset, comprising nearly three decades of detailed demographic and behavioural data, a multi-generation pedigree, and genome-wide SNP data from over 1100 individuals, to test how social structure and male-biased dispersal influence genetic structure and evolutionary potential in the Ngorongoro hyena population. Individuals in the Ngorongoro population occupy one of eight resident clans whose territories vary along key ecological axes such as group size (e.g., min = 22, max = 90 individuals in a clan in June 2024), habitat type, territory size, exposure to pastoralism and sex ratio (Fig 1A, (Dheer *et al.* 2022)). We

addressed three specific questions: (1) does social structure generate detectable genetic differentiation among clans? (2) does non-random dispersal contribute to this genetic structure? and (3) do rates of adaptive evolution, measured as additive genetic variance in fitness traits, vary among clans? By addressing these questions, we provide empirical evidence that social structure can shape key microevolutionary processes and argue for the integration of social processes into evolutionary models of natural populations.

Results

Population structure and genetic differentiation between clans

Using a panel of genome-wide SNPs (N = 27219) sequenced for 1181 individuals, we first ran a suite of analyses to characterise the genetic structure of the population, including estimating the extent of genetic differentiation between clans. Genomic PCA revealed that genetic variation was distributed continuously throughout the population. The first two axes (PC1 and PC2) of the PCA each explained a relatively low proportion (2%) of the genetic variation in the population, and individuals generally sat continuously across both of these axes. However, individuals that were born into the same clan clustered together across PC1 and PC2 (Figure 1C). This suggests that individuals born into the same clan were more genetically similar to each other than to individuals from other clans. Individuals born into the same clan also clustered together across a continuum of genetic variation in the third and fourth PCA axes (Figure S1), further suggesting that individuals that were from the same clan were more similar to each other than to individuals from other clans. Genetic differentiation between clans was estimated as pairwise F_{ST}. Pairwise F_{ST} between clans varied between 0.013 and 0.043 (Figure 1D). We also found evidence for isolation-by-distance between individuals born into the Crater, whereby individuals that lived further apart from each other were less genetically similar than those living closer (Figure 2D). Genetic variation, estimated as average heterozygosity across individuals, did not differ between clans (Table 1), although there was slightly more variation in individuals' heterozygosity (estimated as the standard deviation of individuals' heterozygosity) in the Triangle clan and less variation among individuals from non-Crater clans (Table 1). The Triangle clan also had the lowest inbreeding coefficient (F_{IS}) of all of the clans, indicating that they were marginally more heterozygous than expected (Table 1). In contrast, individuals from non-Crater clans had a marginally higher inbreeding coefficient (F_{IS}), indicating that they were slightly less heterozygous than expected (Table 1).

Impact of differential rates of dispersal among clans on genetic differentiation

We measured geographic distance and dispersal rates between clans to test whether these factors predicted F_{ST} between clans. Geographic distance between clans was estimated as the distance between the centroid of each clan's territory (see Figure 1A and methods). Dispersal rates between pairs of clans were estimated as the proportion of males that dispersed from one

clan to another clan. Dispersal was considered when males displayed a sexual interest in the females of the clan and engaged in social interactions with members of a new clan for at least 3 months. The majority of males dispersed to a new clan to breed, but a smaller proportion of males remained philopatric to their natal clan (see methods, Figure 1B). Dispersal rates between clans were highly heterogeneous across the population, ranging from 0% of males from Airstrip clan dispersing to Triangle clan (and vice versa), to 51% males from Engitati clan dispersing to Munge clan (Figure 1B). Dispersal rates were often very asymmetric in that the exchange of males between clans via dispersal was rarely reciprocal (Figure 1B). For instance, 21% of males from Triangle clan dispersed to the Lemala clan, but only 3% males from Lemala clan dispersed to the Triangle clan. Geographic distance between clans varied between 2768m (Lemala to Shamba clans) and 11976m (Forest to Munge clans) (Figure 1A). Geographic distance partly explained the probability of male dispersal, but the relationship was quite weak (MRQAP: β = -1.98x10^{-5}, p = 0.006, Figure S4), suggesting that other factors contribute to generating non-random rates of male dispersal between clans.

There were multiple lines of evidence to suggest that genetic differentiation between clans was driven by differential rates of male dispersal between clans. First, for all pairs of clans, observed F_{ST} was larger than that expected if males dispersed between clans at random (Figure 2A). Second, dispersal rates between clans were negatively correlated with pairwise F_{ST} between clans, as pairs of clans that exchanged fewer males were more genetically differentiated than those that had a higher rate of male dispersal (MRQAP: β = -0.036, p = 0.017, Figure 2B). This remained true even when controlling for the effect of geographic distance, which did not have an effect on F_{ST} when fitted in a model together with dispersal rates (MRQAP: β = -2.15 x 10^{-7}, p = 0.750, Figure 2C) or when estimated alone (MRQAP: β =5.05 x 10^{-7}, p = 0.379). This suggests that dispersal between clans, rather than geographic distance, was the mechanism through which genetic stratification emerges in the population. We also found that when males were assigned to their adult clans (i.e., the clan that they disperse into), F_{ST} between clans were lower than when assigned to their natal clan (as presented above), further suggesting that male dispersal acts to facilitate gene flow through the population (Figure S2). Nevertheless, dispersal rates between clans remained negatively associated with observed F_{ST} when calculated with males assigned to their adult clans (MRQAP: β = -0.028, p = 0.009, Figure S3).

Effect of group ancestry on fitness

When individuals disperse, they introduce new alleles to recipient groups and, if these alleles are associated with fitness-related traits, dispersal can affect the distribution of additive genetic (breeding) values for fitness (i.e., survival and reproduction) in the recipient group. These changes can modify the genetic variance available for selection. Consequently, dispersal can shape evolutionary trajectories not just through gene flow, but by altering the heritable basis of fitness (Reid & Arcese 2020; Wolak & Reid 2017). We examined whether dispersal between social groups had the potential to affect genetic variation for fitness by testing if genetic ancestry from

different clans, including from non-Crater clans, was associated with different breeding values for fitness. To do this, we used "genetic group animal models" described in (Muff *et al.* 2019) to estimate the effect of ancestry from different clans using two estimates of individual fitness, lifespan (in years) and lifetime reproductive success (LRS, total number of offspring), estimated for 1635 Crater-born individuals born between 1996 and 2015 together with the population pedigree of 3239 individuals with known natal clans (see methods for details). The number of individuals with non-zero genetic ancestry from different clans varied between 105 and 1318 (Lemala = 722, Munge = 877, Shamba = 105, Triangle = 117, Ngoitokitok = 346, Engitati = 413, Forest = 256, Airstrip = 840, non-Crater = 1318). Note that while we only analysed data from individuals born into Crater clans, a large proportion of these individuals had ancestry from immigrant (or non-Crater born) individuals. Accordingly, we found that individuals with a greater proportion of genetic ancestry from outside the Crater (i.e., from ancestors that immigrated into the Crater) had more offspring and longer lifespans (Table 2, Figure 3). Statistically, this was reflected in negative regression coefficients for the proportion of genetic ancestry from all Crater clans on fitness measures (Table 2), indicating that higher Crater ancestry was associated with lower fitness relative to individuals with greater immigrant ancestry, which served as the reference category (see Methods). As such, individuals that had more genetic ancestry from the Crater *relative* to outside the Crater had lower fitness (Table 2). Beyond the effect of immigrant genetic ancestry, proportion of genetic ancestry from Crater clans did not affect either LRS or lifespan (Table 2).

Estimated rate of adaptive evolution in clans

We estimated rates of adaptive evolution following the fundamental theorem of natural selection which states that the per generation change in mean fitness caused by selection is equivalent to the additive genetic variance (V_A) in fitness (Fisher 1930). We estimated whether the rate of evolution differed between clans using quantitative genetic analyses to estimate V_A in fitness per clan (Bonnet *et al.* 2022; Fisher 1930). As a baseline estimate of V_A, we first fit our genetic group model assuming homogeneous variance, and found that V_A was high and the posterior distributions were clearly distinct from zero for both LRS and lifespan (V_A^{LRS} = 0.379 (0.226 - 0.546), $V_A^{Lifespan}$ = 0.360 (0.285 - 0.443), Table 2, Figure 4). These estimates were similar, but slightly lower, than that found by (Bonnet *et al.* 2022) who found an estimate of V_A^{LRS} of 0.448 ($95\%_{CI}$ 0.147 – 0.811) in this population. We then fit genetic group models with heterogeneous V_A among clans. Despite the increased uncertainty due to sample sizes per clan, we found clear evidence for differences in V_A in fitness between some clans (Table 3, Figure 4). Following methods outlined by (Muff et al., 2019), we calculated the posterior distribution of differences in V_A for each pair of clans using the posterior distributions of variances, identifying pairs of clans for which 95% of the posterior of the difference did not overlap 0. In doing so, we identified pairs of clans whose V_A in fitness differed for 95% of the posterior distribution. For LRS, we found that seven pairs of clans had different estimates of V_A to each other (out of a possible 36 pairs, Figure S6) and for lifespan there were eight pairs of clans with different V_A (Figure S7). For both LRS and

lifespan, the Forest clan had higher V_A in fitness than four of the eight other clans in both traits (posterior median and 95% CI difference to Forest clan in V_A **LRS**: Lemala $2.186_{0.476-5.196}$, Triangle $2.063_{0.169-4.973}$, Engitati $1.986_{0.102-4.918}$, Munge $1.869_{0.036-4.909}$; **Lifespan**: Lemala $1.145_{0.481-2.133}$, Engitati $0.916_{0.191-1.893}$, non-Crater $0.839_{0.142-1.788}$, Munge $0.786_{0.049-0.719}$). Triangle clan had lower V_A in LRS than Ngoitokitok ($-1.177_{-2.71--0.054}$). Notably, Lemala clan had very low estimates of V_A for both LRS and lifespan (Table 2, Figure 4). This was further reflected in the Lemala clan having lower V_A than two clans for LRS (Forest $-2.186_{-5.196--0.476}$, Ngoitokitok $-1.239_{-2.822--0.394}$) and five clans for lifespan (Forest $-1.445_{-2.133--0.471}$, Triangle $-1.004_{-2.106--0.297}$, Ngoitokitok $-0.740_{-1.376--0.271}$, Airstrip $-0.546_{-0.953--0.186}$, Munge $-0.277_{-0.719--0.049}$). Residual variance (which reflects the remaining phenotypic variance which is not explained by parameters in the model) for LRS varied substantially between clans. The Lemala clan, which had the lowest V_A, had the lowest residual variance in LRS, whereas Engitati and Airstrip clans had the largest residual variance in LRS (Table 3, Figure S5). Residual variance for lifespan did not differ much between clans, although it was marginally smaller for Engitati clan and marginally larger for Lemala clan (Table 3, Figure S5).

Fixed effect estimates from animal models suggest that males had lower fitness than females (Table 2). Individuals born into a higher social rank had higher fitness, both living longer and having more offspring overall (Table 2), which is a well-established phenomenon in spotted hyenas (Davidian *et al.* 2016; Hofer & East 2003; Holekamp *et al.* 1996). We did not find that the effect of social rank was different between the sexes (Table 2). We also found a negative relationship between the inbreeding coefficient, F, and both fitness measures, suggesting evidence for inbreeding depression (Table 2). There was evidence for a negative relationship with natal clan size for both fitness measures (Table 2), suggesting evidence for an effect of density in early life on fitness (Bailey *et al.* 2024). The fixed effect estimates were similar in models that fit either homogeneous or heterogeneous variances, and results of the latter can be found in Table S1. Variance estimates for all other random effects (i.e., not V_A) were fairly small (Table 3). There was very low variance in fitness associated with natal clan (Table 3), suggesting very little difference in the phenotypic means in fitness between clans.

Discussion

Social structures in wild animals emerge from the interaction between ecological and evolutionary processes (Székely *et al.* 2010). However, in addition to being shaped by evolution, social structure may influence microevolutionary processes by altering patterns of gene flow, modulating genetic drift, or structuring populations in ways that generate variation in natural selection (He *et al.* 2019). In a population of wild spotted hyenas, we combined nearly 30 years of behavioural, life-history, and fitness data with a new genome-wide SNP dataset to provide insight into how social structure may mediate microevolution. We found small but meaningful genetic differentiation among clans, driven by heterogeneous dispersal patterns. Furthermore, additive genetic variance of fitness (measured as lifespan and lifetime reproductive success)

varied among clans, suggesting that different social groups within the population may experience a distinct selective landscape.

While theory predicts that social structure should shape population genetic patterns, empirical tests remain limited, largely because dispersal is hard to observe directly (Clobert *et al.* 2009; Watts *et al.* 2011). In our study, genetic variation was distributed continuously across the population, but clans inhabiting the 250 km^2 Ngorongoro Crater were genetically differentiated to an extent comparable to metapopulations separated across island archipelagos (e.g., house sparrows, islands of Helgeland, Norway (Jensen *et al.* 2013)). Studies demonstrating genetic differentiation among social groups are rare. Previous work reported genetic differentiation between hyena clans (Watts *et al.* 2011), but this was based on very geographically distant clans. The patterns we observed likely reflects, in part, both the kinship inherent to hyena social groups (Estandía *et al.* 2025; Kruuk 1966) and reproductive skew in both sexes (Engh *et al.* 2002; Holekamp *et al.* 1996, 2012; Höner *et al.* 2010), which results in recent coalescence to common ancestors within clans. Our findings also elucidate the mechanisms by which social structure mediates genetic structure: while male dispersal facilitated gene flow, rates of dispersal were not predicted by geographic distance. Instead, local variation in clan density, sex ratio and kin structure, may underlie dispersal dynamics (Davidian *et al.* 2016; Davidian & Höner 2022). Furthermore, the tendency for male relatives to disperse together may limit genetic mixing and amplify differentiation (Yearsley *et al.* 2013). As such, variation in dispersal between clans appears to mediate the relationship between social and genetic structure in this population.

Dispersal can also shape evolutionary potential by introducing alleles that alter the additive genetic variance for fitness in recipient groups, provided these introduced alleles are associated with fitness-related traits. In our study, genetic ancestry from different Crater clans was not associated with differences in breeding values for fitness. This supports previous findings that suggest dispersal between clans does not confer a fitness advantage over philopatry (Davidian *et al.* 2016). However, ancestry from immigrants originating from outside the Crater was associated with higher fitness, consistent with findings from other animal populations (Saatoglu *et al.* 2025). This is an important finding because group-living animals, including hyenas, often show mating bias toward members of their own group (Ellis *et al.* 2022), increasing the risk of inbreeding. Dispersal is therefore widely considered a mechanism to reduce inbreeding in kin-biased societies (Perrin & Mazalov 2000; Pusey 1987). Our results support this hypothesis because the high proportion of immigrant ancestry in the population might have resulted from biased mate choice toward males that immigrate from outside the population. This could facilitate the spread of beneficial alleles introduced through immigration or introduce greater genetic variation to the population, even though we found no direct link between immigrant ancestry and genetic diversity. Regardless, our findings suggest there is unlikely to be selection against immigrants in the population, under which we would have expected lower fitness for immigrant ancestry (Hendry 2004). Although immigration from outside the Crater is rare, having immigrant ancestry is extremely common in the population, and it is likely that it plays a

disproportionate role in shaping the quantitative genetic architecture of fitness related traits in the population (Reid & Arcese 2020).

Most wild populations are subdivided to some extent, resulting in population genetic structure (Bohonak 1999). In structured populations, genotypes are distributed non-randomly across heterogeneous environments, where both the availability of mates and ecological selection pressures vary spatially (Barton & Clark 1990). As such, population structure may significantly affect evolutionary dynamics (Allen *et al*. 2017; Frean *et al*. 2013; Nowak *et al*. 2010). Yet, foundational evolutionary models, such as the fundamental theorem of natural selection, typically assume panmixia when estimating selection on a trait or the rate of evolution, ignoring the effects of structure (Fisher 1930; Robertson & Lewontin 1968; Walsh & Lynch 2018). Our findings show that social structure can generate genetic differentiation and modulate evolutionary potential. Recent work reported that adaptive evolution occurs at biologically meaningful rates in many wild populations, including in the hyenas studied here (Bonnet *et al*. 2022). But, mispredictions are common in studies of microevolution in wild animals (Merilä *et al*. 2001; Pujol *et al*. 2018). One often overlooked reason for this may be that studies typically do not account for how social structure generates heterogeneous selective landscapes. Our findings demonstrate that social structure can generate genetic differentiation and create variable evolutionary potential across a single population. While the factors driving different rates of adaptive evolution between clans remain uncertain, the Lemala clan provides an illustrative case: it has the lowest estimated rate of evolution, the lowest male-to-female sex ratio, the highest proportion of philopatric males, and the most stable social hierarchy. It also shows strong reproductive skew among females, with a few individuals contributing disproportionately to the next generation (unpublished). These socio-demographic features likely reduce genetic variance for fitness by limiting mate pools and increasing relatedness, highlighting how fine-scale variation in social structure can influence evolutionary potential.

In conclusion, our study provides a comprehensive empirical examination of how social structure impacts evolution in the wild, revealing that social behaviours and structures in a population can profoundly shape microevolutionary processes. Interestingly, our results suggest broadly similar inferences from either LRS and lifespan, despite them reflecting different combinations of the core fitness components of survival and fecundity. This suggests that the adaptive benefits of immigration and rates of adaptive evolution per clan are not being driven by one or other fitness component, but more likely the combined effect of both. We cannot yet quantify how population structure or clan-level variation in the rate of adaptive evolution correspond to realised genetic evolution, which traits may be under selection, or if these are the same across all clans. Nonetheless, our findings illustrate that social behaviours and group structure are key mediators of gene flow and evolutionary potential, and should be more fully integrated into evolutionary theory and empirical study design.

Methods

Study population and data collection

We used data collected as part of a long-term individual-based study of a population of spotted hyenas occupying the floor of the Ngorongoro Crater in Tanzania (approx. 250 km^2, $3^{\circ}11^{'}$ S, $35^{\circ}34^{'}$ E). The population has been continuously monitored as part of a long-term individual-based study since 1996. All hyenas in the population are individually identified by their spot pattern and colouration of their coat, as well as other uniquely identifiable features (e.g., scars and ear notches, described in (Davidian *et al.* 2016; Höner *et al.* 2007)). Demographic, life-history and behavioural data are collected via almost daily surveys, during which observations are made from a vehicle that hyenas are well habituated to from birth (Davidian *et al.* 2021; Höner *et al.* 2007). DNA is isolated from tissue, faecal, skin or hair samples which are collected opportunistically.

Births and deaths are rarely observed in the wild, but are estimated using a combination of resighting data during daily surveys, behavioural observations and morphology (explained in (Davidian *et al.* 2016; Dheer *et al.* 2022)). Social ranks of all individuals are determined via established rules of inheritance for hyenas and verified via behavioural interaction data (Davidian *et al.* 2021; Vullioud *et al.* 2019). Males were considered to have chosen a breeding clan when they displayed a sexual interest in the females of either their natal clan (a philopatric male) or another clan (a disperser) and engaged in social interactions with members of that clan for at least 3 months (Davidian *et al.* 2016, 2021). Parentage is identified via a combination of observations of nursing and genetic parentage assignment, which is reconstructed using nine polymorphic microsatellite loci in CERVUS 3.0 (Höner *et al.* 2010; Kalinowski *et al.* 2007). In this study, we used records collected between April 1996 and February 2025. The full pedigree used in this study contained 3239 individuals sampled across a maximum depth of nine generations.

Field data collection and transport was permitted by the Vice President's office of the United Republic of Tanzania (Ref. No. BA 78/130/01/42), the Tanzania Wildlife Research Institute (TAWIRI), the Tanzania Commission for Science and Technology (COSTECH; Permit No. 2021-380-NA-1990), and the Ngorongoro Conservation Area Authority (NCAA). All study procedures were performed in compliance with the ethical regulations of these institutions and the Internal Committee for Ethics and Animal Welfare of the Leibniz Institute for Zoo and Wildlife Research Berlin (No. 2020-06-02).

Genotyping-by-sequencing

A detailed description of the methodological approach we used to generate the genotype-by-sequencing dataset is provided in (Arantes *et al.* 2025). Briefly, library construction followed the 3RADseq protocol (Bayona-Vásquez *et al.* 2019) using the restriction enzymes *EcoRI*, *XbaI*, and *NheI* and incorporating a modification to remove PCR duplicates (Hoffberg *et al.* 2016), which

enabled sequencing approximately 23500 loci (380 - 460 bp). Barcoded samples were pooled prior to fragment size selection (range to 480-640 bp) using the Blue Pippin (Sage Science) with a 1.5% cassette. The product was split into two aliquots and each was subjected to a single-cycle PCR to incorporate the iTru5-8N primer, followed by an indexing PCR. A low coverage sequencing run (approximately 2,000 reads per individual) was conducted to screen samples for endogenous DNA content and, based on the proportion of reads assigned to each individual, new volumes of the digestion/ligation product were calculated and re-pooled for a second library preparation to ensure equal representation across individuals (described in detail in (Arantes *et al.* 2025)). Final libraries were sequenced on an Illumina NovaSeq S4 platform using 150 bp paired-end reads targeting ≥30× coverage per individual. Reads were then trimmed of their adapters using Cutadapt (Martin 2011) before being demultiplexed using Flexbar (Roehr *et al.* 2017). PCR duplicates were removed using a custom Python script. We then concatenated the two replicates of each library and merged forward and reverse reads using PEAR v0.9.11 (maximum assembled sequence length of 270 bp (Zhang *et al.* 2014)). Unmerged reads were trimmed to a maximum length of 130 bp and a minimum quality score of 30 using Trimmomatic (Bolger *et al.* 2014). An *in silico* digestion was performed to filter out undigested or chimeric sequences and we only retained reads with correct sequences at both ends (using a custom Python script).

Reads were mapped to the *C. crocuta* genome (accession number GCA_008692635.1) using Bowtie2 with default parameters and the options "-no-mixed" and "-no-discordant". Mitogenome and sex scaffolds identified with RADsex (Feron *et al.* 2021) were excluded using Samtools (Danecek *et al.* 2021). SNP calling was performed using Stacks reference-based pipeline v2.61 (Catchen *et al.* 2013; Rochette *et al.* 2019) and loci were retained if genotyped in ≥60% of individuals. Individuals with >2.5M reads were subsampled to normalize coverage (~60×). In total, 1187 spotted hyenas were genotyped at a total of 69816 SNP markers. For all downstream analyses, we further filtered genotypes to include genotype calls that had a read depth of at least 10 and no more than 110 before filtering the SNPs to remove SNP loci that had a minor allele frequency of less than 1% and a genotype missing rate of more than 30%. We also removed sequences for six individuals that had a genotyping rate of less than 50%. The final dataset used in all downstream analyses therefore consisted of 1181 individuals and 27219 unlinked SNPs.

Population structure

To characterise genetic substructure and differentiation in our study population, we ran two analyses: genomic principal component analysis (PCA) and isolation-by-distance. Genomic PCA was used to describe patterns in genetic variation through the population, and was conducted with all SNP loci using the *ADEGENET* package in R (Jombart & Ahmed 2011), retaining all of the principal components as per (Jombart *et al.* 2010). Isolation-by-distance was estimated via the correlation between geographic distance and genetic distance. Geographic

448 distance between pairs of individuals was estimated as the distance in metres between the
449 centroids of individuals' home ranges, which were calculated as 95% minimum convex polygons
450 (MCPs) using GPS coordinates collected when the individuals were sighted (Pebesma & Bivand
451 2023). Centroids of MCPs were calculated using geometric unary operations via the *sf* R package
452 (Pebesma 2018; Pebesma & Bivand 2023). Individuals with less than five sightings were removed
453 to ensure precision in estimation of home ranges (and therefore centroids). Genetic distance
454 between individuals was calculated using Nei's genetic distance (Nei 1972) using the *StAMPP* R
455 package (Pembleton *et al.* 2013).
456
457 *Impact of non-random dispersal between clans on genetic differentiation*
458
459 Genetic differentiation between clans was calculated as pairwise F_{ST} using the *StAMPP* R
460 package (Pembleton *et al.* 2013). Individuals were assigned to their natal clan, but analyses were
461 also conducted with individuals assigned to the clan in which they lived as adults. The majority
462 of males (>90%) dispersed to a breeding clan once and remained there for their entire adult life.
463 However, a small number of males had multiple dispersal events throughout their lifetime,
464 where they may disperse to a new clan later in life after spending some time living and
465 reproducing in the clan to which they dispersed first. For these males, their adult clan was
466 identified as the clan in which they spent the majority of their adult life. We calculated the
467 probability of dispersal between clans as the proportion of males that dispersed from one clan
468 to another, relative to the total number of dispersing males from the first clan. In this way, we
469 generated an asymmetric matrix that described the probability of dispersal between clans and
470 where the diagonal element of the matrix describes the proportion of adult males that are
471 philopatric and reproduce in their natal clan. All individuals and all clans were retained when
472 calculating F_{ST}, including individuals born to non-Crater clans that dispersed into a Crater clan.
473 Geographic distance between clans was calculated in metres between the centroid of clans'
474 territories. Clans and territories have remained very stable across time: the territory boundaries
475 have remained largely the same across time and there have been no instances of clan
476 replacement or formation during the study period (Figure 1A (Höner *et al.* 2005; Kruuk 1972)).
477 Clan territories were estimated as utilisation distributions (UD) via the kernel density methods
478 in the *adehabitatHR* R package (Calenge 2006) using sightings of resident individuals of each
479 clan. UDs for each clan were estimated with a smoothing parameter of 440 metres, which was
480 identified as the maximum value used when optimised per clan and when assuming a bivariate
481 normal distribution of geographic coordinates. We decided to use a single smoothing parameter
482 for all clans, rather than different parameters for each clan via the *href* option, to standardise the
483 estimation of clan territories across the population. Centroids per clan territory and geographic
484 distance between all centroids were then identified as described above.
485
486 To test whether pairwise F_{ST} between clans was greater than expected given the rate of dispersal
487 in the population, we ran a null model to generate F_{ST} that would be expected if males dispersed
488 randomly among clans. This model randomly permuted clan membership between males,

randomly distributing them through the population and simulating random dispersal. This permutation model was run a total of 1000 times, and in each one, we calculated pairwise F_{ST} between clans to generate a range of F_{ST} between clans expected under random male dispersal. We then identified whether observed pairwise F_{ST} between clans were greater than expected if the observed F_{ST} was greater than the 95th percentile of the distribution of the F_{ST} in the null model. This model was run in R using custom scripts and the *StAMPP* R package to calculate F_{ST} for each iteration of the permutation model.

To test if either the probability of dispersal or geographic distance between clans predicted pairwise F_{ST}, we ran a multiple regression quadratic assignment procedure (MRQAP) with the matrix of F_{ST} as the response matrix and the matrices of probability of dispersal and geographic distance as the predictor matrices. MRQAP is an extension of the mantel test (a statistical procedure commonly used to identify correlations between matrices) which allows the user to concurrently test for the effects of multiple predictor matrices. We ran MRQAP using the *asnipe* R package (Farine 2013) and used the double semi-partialling method (DSP) which permutes (N = 1000 randomisations) the matrix of residuals from the ordinary least regression of the dependent matrix on the independent matrices to estimate error and calculate the effects. These analyses were conducted using the 8 Crater clans. Non-Crater clans were discarded because their estimated rate of dispersal and home range are likely to be inaccurate due to lower monitoring effort.

Effect of group ancestry on fitness traits and estimated rate of evolution in clans

We used two measures of lifetime fitness: lifetime reproductive success (LRS), defined as the total number of offspring each individual had, and lifespan, defined as the age at which an individual died in years. LRS was calculated for each individual using the pedigree described above. Lifespan was estimated using estimated birth and death dates, as described above. We selected to use both LRS and lifespan in order to approximate the fitness components of survival and fecundity as best we could. More specifically, we expect that LRS should putatively reflect both survival and fecundity, whereas lifespan reflects survival alone. We note, however, that LRS and lifespan have a fairly high positive correlation ($r = 0.78$). Nevertheless, we decided to analyse both LRS and lifespan in an attempt to explore whether inferences about genetic variance in fitness, including differences between clans, may differ depending on the fitness component analysed. To ensure that we had observations across the complete lifespans of individuals, we only included individuals who were born after the beginning of the study period (1996) and were known to have died by February 2025. To ensure measures of fitness were accurate for each cohort in the dataset, we also removed individuals born after 2015 because 90% of individual hyenas die before 10 years of age (last observation date in analyses = February 2025). Analyses described below therefore included N =1635 Crater-born individuals for which mothers and social ranks at birth were known. Most individuals were of known sex; however, sex was unknown for 204 individuals, normally because they died before reaching an age where sex could be

530 determined. Rather than remove these individuals, we randomly assigned a sex to them as
531 removing them would have systematically biased our distribution of fitness estimates as they
532 were mostly juveniles. Retaining them also allowed us to maximise the statistical power in the
533 models described below. The estimated effect of sex on either relative LRS or lifespan from
534 models that randomly assigned 204 individuals a sex (as described) was very similar to models
535 that were fitted to a dataset without these 204 individuals (LRS: without N = 204 individuals
536 β_{SEX}^{MALE} = -0.814 (-1.134 - -0.547), random assignment of sex β_{SEX}^{MALE} = -0.814 (-1.114 - -0.519);
537 Lifespan: without N = 204 individuals β_{SEX}^{MALE} = -0.228 (-0.372 - -0.068), random assignment of
538 sex β_{SEX}^{MALE} = -0.248 (-0.406 - -0.104)). As such, we determined that the random assignment of
539 sex did not have a significant impact on the estimated effect of sex, but removing them would
540 have both selectively removed individuals with low fitness and appreciably reduced our
541 statistical power. We therefore present results of models fitted to data that randomly assigned
542 the sex to these 204 individuals.
543
544 We estimated V_A in fitness using "genetic group animal models" (Aase *et al.* 2022; Muff *et al.*
545 2019; Quaas 1988; Wolak & Reid 2017). Animal models describe a type of linear mixed effects
546 model which fits relatedness information as a covariance matrix to estimate additive genetic
547 variance, that is, the variance of the additive genetic values, for the response variable (Kruuk
548 2004; Wilson *et al.* 2010). Genetic group animal models further allow the distribution of the
549 genetic values differ according to genetic substructures (so-called "genetic groups") in the
550 population. In our case, each genetic group corresponded to one of the clans, in addition to a
551 group lumping together non-Crater (immigrant) individuals. We fitted the response variables of
552 LRS and lifespan relative to the global population mean with a Gaussian distribution rather than
553 with alternative options on the raw distribution (e.g., zero-inflated Poisson for LRS) to aid the
554 fitting of these fairly complex models. We do not think that any estimated differences between
555 genetic groups (clans) in genetic variance of fitness was influenced by mean-variance
556 relationships that may remain after relativising fitness to the global mean because the
557 correlation between mean and genetic variance per clan was quite low for both measures of
558 fitness (r^2 (V_A^{LRS} and mean LRS) = -0.42, r^2 ($V_A^{Lifespan}$ and mean lifespan) = 0.27) (see also Table 2,
559 Table 3 and Figure 4).
560
561 For each response variable we fitted two models. The first model estimated a single V_A for the
562 whole population and the second estimated heterogeneous V_A specific to social clans. We fitted
563 fixed effects of individuals inbreeding coefficient (F, estimated from the pedigree), sex, size of
564 the clan at birth, social rank at birth and the proportion of genetic ancestry (genetic group
565 proportion) from each clan (Quaas 1988; Wolak & Reid 2017)). Ancestry proportions were
566 calculated using the full pedigree using the *nadiv* R package (Wolak 2012) and following the
567 methodology set out in (Muff *et al.* 2019; Wolak & Reid 2017). As random effects, we fitted
568 maternal identity, cohort (i.e., year of birth), and birth clan. We included a birth clan random
569 effect in our models to model the effect that the developmental environment is known to have
570 on both juvenile and adult fitness in hyenas (Gicquel *et al.* 2022; Höner *et al.* 2010). Whilst we

acknowledge that this will not account for the environment experienced by most males that survive to adulthood, we believe that this random effect accounts for the environment thought to be most influential in hyena development. The models that estimated homogeneous variances were therefore:

$$w_i = \mu + \beta_1 F_i + \beta_2 sex_i + \beta_3 socialrank_i + \beta_4 clansize_i + \beta_5 (sex_i \times socialrank_i) + \beta_6 qA_i + \beta_7 qE_i + \beta_8 qF_i + \beta_9 qL_i + \beta_{10} qM_i + \beta_{11} qN_i + \beta_{12} qT_i + \beta_{13} qS_i + m_k + c_l + y_z + a_i + \varepsilon_i \quad \text{(Equation 1)}$$

where w_i is the estimate of fitness (w) for individual i, μ is the population mean, $\beta_{1,..,13}$ are the fixed effects explained above, where *q* denotes the expected ancestry proportion for individual *i* from each Crater clan (A = Airstrip, F = Forest, T = Triangle, E = Engitati, L = Lemala, S = Shamba, M = Munge, N = Ngoitokitok), m is the maternal effect for mother *k*, c is the clan effect for clan *l*, y is the cohort effect for cohort z, a is the genetic breeding value (i.e., the effect of individual *i's* genome relative to μ), where breeding values are distributed as $(a_1,....,a_n) \sim N(0,V_A A)$ with additive genetic variance V_A and genetic relatedness matrix A (estimated via the population pedigree), and ε is the residual term. Given that ancestry proportions always sum to one, one genetic group is used as a reference level (as one would when fitting a categorical fixed effect), in this case the proportion of genetic ancestry from non-Crater individuals (Muff *et al*. 2019; Quaas 1988). Note that while the individuals contained in the dataset used in these analyses did not include individuals born outside the Crater, it is still possible that individuals will have genetic ancestry from outside the Crater and we used the full pedigree to determine these proportions.

In standard animal models, the underlying assumption is that all individuals come from the same genetic population, and the breeding values (a_i) therefore encode individuals' deviation from the mean of that population (with a mean of zero). The genetic group model in equation 1 relaxes the assumption of genetic homogeneity and incorporates genetic substructure caused by social grouping by allowing the clans to differ in mean breeding value. The coefficients β_6 through β_{13} can be viewed as mean breeding values for the respective clans, while additive genetic variances are still treated as homogeneous across clans. To allow the model to account for potentially heterogeneous additive genetic variances of different clans, we fitted a second model such that it splits individuals' genetic breeding values (a_i) into group specific contributions. To do this, we followed methods explained in detail in (Muff *et al*. 2019). Briefly, we constructed a model which estimates partial genetic breeding values (a_{ij}) for each group (clan), where a_{ij} represents the contribution to the breeding value of individual *i* from group *j*. These partial genetic breeding values are estimated for all groups (j=1,...,r) and are distributed as $(a_{1j},....,a_{nj}) \sim N(0, V_{Aj} A_j)$, where V_{Aj} is the additive genetic variance in group *j* and A_j is a group-specific relatedness matrix. Group specific relatedness matrices were specified as per (Muff et al., 2019) and the method can be found in the available R code, but the main principle is that relatedness between individuals is scaled according to their ancestry proportions in such a way that the group-specific relatedness

matrices estimate pairwise relatedness at genetic variants inherited from specific groups. Whilst methods exist to scale relatedness information based on genomic estimates of realised genetic ancestry from different groups (Aase *et al*. 2022), we opted to use a pedigree-based approach as per (Muff *et al*. 2019) to maximise the sample size available for the models (i.e., the full hyena pedigree contains 3239 individuals whereas the genetic dataset has 1181 individuals). To account for potential environmental differences that could result in between-clan differences in phenotypic variance in fitness, we also fitted heterogeneous residual variances between clans, where individuals were assigned to their natal clan.

All models were fitted in a Bayesian framework using *MCMCglmm* package in R (Hadfield 2010). We used the default weakly informative parameter expanded priors set to $F_{1,1}$ distribution setting the scale to 1000 for random effects, the default Inverse-Wishart prior for residual variance and non-informative priors for the fixed effects. Models were run for 18000 iterations, with a 3000 burn-in period and a thinning interval of 10 which was sufficient in all cases to achieve an effective sample size of at least 1000 for all parameters and for there to be low autocorrelation. Model convergence was assessed by ensuring effective sample sizes were at least 1000, visual inspection of trace plots to ensure sufficient sampling and low.

Acknowledgements

This work was supported by a European Research Council grant to L.E.B.K. (#101020503) and L.E.B.K. was funded by The Royal Society. We are grateful to the Tanzanian authorities (TAWIRI, COSTECH and NCAA) for their enduring support to the Ngorongoro Hyena Project. We thank Camila Mazzoni for contributions in the lab to the generation of the genetic dataset analysed in this paper. We also thank Alexandre Courtiol and members of the Data Zoo Gang for development and maintenance of the hyenaR package essential to data extraction. We also thank Philemon Naman, Arjun Dheer and Marta Mosna for field data collection, and Stephan Karl, Kerstin Wilhelm and Dagmar Thierer for assistance in the lab. We thank Bettina Wachter for her contribution to data collection and management in the early years of the Ngorongoro Hyena Project. We also thank Josephine Pemberton for useful feedback on early phases of analyses presented here.

Author contributions

Conceptualisation: KS, OH, ED; Data curation: OH, ED, LA; Resources: OH, ED, LA; Investigation: KS, OH, PN, LA, ED; Formal analysis: KS, LA; Methodology: KS, LA, JP, KA; Software: KS, LA; Validation: KS, OH, ED, JP, KA; Visualisation: KS; Supervision: LK; Project administration: KS, LK; Funding acquisition: LK, OH; Writing – original draft: KS; Writing – review and editing: all

References

Aase, K., Jensen, H. & Muff, S. (2022). Genomic estimation of quantitative genetic parameters in wild admixed populations. *Methods in Ecology and Evolution*, 13, 1014–1026.

Alexander, R.D. (1974). The evolution of social behavior. *Annual review of ecology and systematics*, 5, 325–383.

Allen, B., Lippner, G., Chen, Y.-T., Fotouhi, B., Momeni, N., Yau, S.-T., *et al.* (2017). Evolutionary dynamics on any population structure. *Nature*, 544, 227–230.

Arantes, L.S., Caccavo, J.A., Sullivan, J.K., Sparmann, S., Mbedi, S., Höner, O.P., *et al.* (2025). Scaling-up RADseq methods for large datasets of non-invasive samples: Lessons for library construction and data preprocessing. *Molecular Ecology Resources*, 25.

Bailey, L.D., Höner, O.P., Davidian, E., Dheer, A., Radchuk, V., Walter, L.F., *et al.* (2024). Effects of environmental change on population growth: monitoring time-varying carrying capacity in free-ranging spotted hyenas. *bioRxiv*, 2024.04.11.589105.

Barton, N. & Clark, A. (1990). Population Structure and Processes in Evolution. In: *Population Biology: Ecological and Evolutionary Viewpoints* (eds. Wöhrmann, K. & Jain, S.K.). Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 115–173.

Bayona-Vásquez, N.J., Glenn, T.C., Kieran, T.J., Pierson, T.W., Hoffberg, S.L., Scott, P.A., *et al.* (2019). Adapterama III: Quadruple-indexed, double/triple-enzyme RADseq libraries (2RAD/3RAD). *PeerJ*, 7, e7724.

Bohonak, A.J. (1999). Dispersal, Gene Flow, and Population Structure. *The Quarterly Review of Biology*, 74, 21–45.

Bolger, A.M., Lohse, M. & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30, 2114–2120.

Bonnet, T., Morrissey, M.B., de Villemereuil, P., Alberts, S.C., Arcese, P., Bailey, L.D., *et al.* (2022). Genetic variance in fitness indicates rapid contemporary adaptive evolution in wild animals. *Science*, 376, 1012–1016.

Bowler, D.E. & Benton, T.G. (2005). Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biological Reviews*, 80, 205–225.

Calenge, C. (2006). The package “adehabitat” for the R software: a tool for the analysis of space and habitat use by animals. *Ecological modelling*, 197, 516–519.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A. & Cresko, W.A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, 22, 3124–3140.

Clobert, J., Le Galliard, J.-F., Cote, J., Meylan, S. & Massot, M. (2009). Informed dispersal, heterogeneity in animal dispersal syndromes and the dynamics of spatially structured populations. *Ecology Letters*, 12, 197–209.

Clutton-Brock, T. (2021). Social evolution in mammals. *Science*, 373, eabc9699.

Clutton-Brock, T.H. (2002). Breeding together: Kin selection and mutualism in cooperative vertebrates. *Science*, 296, 69–72.

Clutton-Brock, T.H. & Lukas, D. (2012). The evolution of social philopatry and dispersal in female mammals. *Molecular Ecology*, 21, 472–492.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., *et al.* (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, 10, giab008.

Davidian, E., Courtiol, A., Wachter, B., Hofer, H. & Höner, O.P. (2016). Why do some males choose to breed at home when most other males disperse? *Science Advances*, 2, e1501236.

Davidian, E. & Höner, O.P. (2022). Kinship and similarity drive coordination of breeding-group choice in male spotted hyenas. *Biology Letters*, 18, 20220402.
Davidian, E., Wachter, B., Heckmann, I., Dehnhard, M., Hofer, H. & Höner, O.P. (2021). The interplay between social rank, physiological constraints and investment in courtship in male spotted hyenas. *Functional Ecology*, 35, 635–649.
Dheer, A., Davidian, E., Courtiol, A., Bailey, L.D., Wauters, J., Naman, P., *et al.* (2022). Diurnal pastoralism does not reduce juvenile recruitment nor elevate allostatic load in spotted hyenas. *Journal of Animal Ecology*, 91, 2289–2300.
East, M.L. & Hofer, H. (2001). Male spotted hyenas (Crocuta crocuta) queue for status in social groups dominated by females. *Behavioral Ecology*, 12, 558–568.
Ellis, S., Johnstone, R.A., Cant, M.A., Franks, D.W., Weiss, M.N., Alberts, S.C., *et al.* (2022). Patterns and consequences of age-linked change in local relatedness in animal societies. *Nature Ecology & Evolution*, 6, 1766–1776.
Engh, A.L., Funk, S.M., Horn, R.C.V., Scribner, K.T., Bruford, M.W., Libants, S., *et al.* (2002). Reproductive skew among males in a female-dominated mammalian society. *Behavioral Ecology*, 13, 193–200.
Estandía, A., Recalde, N.M., Slate, J. & Sheldon, B.C. (2025). The genetic consequences of dispersal and immigration in a wild great tit population. *bioRxiv*, 2025.06.26.660565.
Farine, D.R. (2013). Animal social network inference and permutations for ecologists in R using asnipe. *Methods in Ecology and Evolution*, 4, 1187–1194.
Feron, R., Pan, Q., Wen, M., Imarazene, B., Jouanno, E., Anderson, J., *et al.* (2021). RADSex: A computational workflow to study sex determination using restriction site-associated DNA sequencing data. *Molecular Ecology Resources*, 21, 1715–1731.
Fisher, R.A. (1930). *The genetical theory of natural selection.* The genetical theory of natural selection. Clarendon Press, Oxford, England.
Frank, L.G. (1986). Social organization of the spotted hyaena Crocuta crocuta. II. Dominance and reproduction. *Animal Behaviour*, 34, 1510–1527.
Frean, M., Rainey, P.B. & Traulsen, A. (2013). The effect of population structure on the rate of evolution. *Proceedings of the Royal Society B: Biological Sciences*, 280, 20130211.
Gardner, M.G., Bull, C.M., Cooper, S.J.B. & Duffield, G.A. (2001). Genetic evidence for a family structure in stable social aggregations of the Australian lizard Egernia stokesii. *Molecular Ecology*, 10, 175–183.
Gicquel, M., East, M.L., Hofer, H. & Benhaiem, S. (2022). Early-life adversity predicts performance and fitness in a wild social carnivore. *Journal of Animal Ecology*, 91, 2074–2086.
Hadfield, J.D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software*, 33, 1–22.
Hamilton, W.D. (1964). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, 7, 1–16.
He, P., Maldonado-Chaparro, A.A. & Farine, D.R. (2019). The role of habitat configuration in shaping social structure: a gap in studies of animal social complexity. *Behavioral Ecology and Sociobiology*, 73, 9.
Hendry, A.P. (2004). Selection against migrants contributes to the rapid evolution of ecologically dependent reproductive isolation. *Evolutionary Ecology Research*, 6, 1219–1236.
Hofer, H. & East, M.L. (2003). Behavioral processes and costs of co-existence in female spotted hyenas: a life history perspective. *Evolutionary Ecology*, 17, 315–331.

Hoffberg, S.L., Kieran, T.J., Catchen, J.M., Devault, A., Faircloth, B.C., Mauricio, R., *et al.* (2016). RADcap: sequence capture of dual-digest RADseq libraries with identifiable duplicates and reduced missing data. *Molecular Ecology Resources*, 16, 1264–1278.

Holekamp, K.E., Smale, L. & Szykman, M. (1996). Rank and reproduction in the female spotted hyaena. *Reproduction*, 108, 229–237.

Holekamp, K.E., Smith, J.E., Strelioff, C.C., Van Horn, R.C. & Watts, H.E. (2012). Society, demography and genetic structure in the spotted hyena. *Molecular Ecology*, 21, 613–632.

Höner, O.P., Wachter, B., East, M.L., Runyoro, V.A. & Hofer, H. (2005). The effect of prey abundance and foraging tactics on the population dynamics of a social, territorial carnivore, the spotted hyena. *Oikos*, 108, 544–554.

Höner, O.P., Wachter, B., East, M.L., Streich, W.J., Wilhelm, K., Burke, T., *et al.* (2007). Female mate-choice drives the evolution of male-biased dispersal in a social mammal. *Nature*, 448, 798–801.

Höner, O.P., Wachter, B., Hofer, H., Wilhelm, K., Thierer, D., Trillmich, F., *et al.* (2010). The fitness of dispersing spotted hyaena sons is influenced by maternal social status. *Nature Communications*, 1, 60.

Jensen, H., Moe, R., Hagen, I.J., Holand, A.M., Kekkonen, J., Tufto, J., *et al.* (2013). Genetic variation and structure of house sparrow populations: is there an island effect? *Molecular Ecology*, 22, 1792–1805.

Jombart, T. & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27, 3070–3071.

Jombart, T., Devillard, S. & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, 11, 94.

Kalinowski, S.T., Taper, M.L. & Marshall, T.C. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular ecology*, 16, 1099–1106.

Kruuk, H. (1966). Clan-system and feeding habits of spotted hyaenas (Crocuta crocuta Erxleben). *Nature*, 209, 1257–1258.

Kruuk, H. (1972). The spotted hyena: a study of predation and social behavior. *(No Title)*.

Kruuk, L.E.B. (2004). Estimating genetic parameters in natural populations using the “animal model.” *Proceedings of the Royal Society of London, Series B.*, 359, 873–890.

Kurvers, R.H.J.M., Krause, J., Croft, D.P., Wilson, A.D.M. & Wolf, M. (2014). The evolutionary and ecological consequences of animal social networks: emerging issues. *Trends in Ecology & Evolution*, 29, 326–335.

Li, X.-Y. & Kokko, H. (2019). Sex-biased dispersal: a review of the theory. *Biological Reviews*, 94, 721–736.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal*, 17, 10–12.

McPeek, M.A. & Holt, R.D. (1992). The Evolution of Dispersal in Spatially and Temporally Varying Environments. *The American Naturalist*, 140, 1010–1027.

Merilä, J., Sheldon, B.C. & Kruuk, L.E.B. (2001). Explaining stasis: microevolutionary studies in natural populations. *Genetica*, 112, 199–222.

Montiglio, P.-O., Ferrari, C. & Réale, D. (2013). Social niche specialization under constraints: personality, social interactions and environmental heterogeneity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368, 20120343.

Morinay, J., Woodward, B.K., Russell, A.F., Sharp, S.P. & Hatchwell, B.J. (2025). Ecological and demographic drivers of kin-directed cooperation in a social bird: Insights from a long-term study. *Journal of Animal Ecology*, 94, 485–500.

Muff, S., Niskanen, A.K., Saatoglu, D., Keller, L.F. & Jensen, H. (2019). Animal models with group-specific additive genetic variances: extending genetic group models. *Genetics Selection Evolution*, 51, 7.

Nei, M. (1972). Genetic Distance between Populations. *The American Naturalist*, 106, 283–292.

Nowak, M.A., Tarnita, C.E. & Antal, T. (2010). Evolutionary dynamics in structured populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365, 19–30.

Parreira, B.R. & Chikhi, L. (2015). On some genetic consequences of social structure, mating systems, dispersal, and sampling. *Proceedings of the National Academy of Sciences*, 112, E3318–E3326.

Pebesma, E. (2018). Simple features for R: standardized support for spatial vector data.

Pebesma, E. & Bivand, R. (2023). *Spatial data science: With applications in R*. Chapman and Hall/CRC.

Pembleton, L.W., Cogan, N.O.I. & Forster, J.W. (2013). StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, 13, 946–952.

Peniston, J.H., Backus, G.A., Baskett, M.L., Fletcher, R.J. & Holt, R.D. (2024). Ecological and evolutionary consequences of temporal variation in dispersal. *Ecography*, 2024, e06699.

Pereira, A.S., De Moor, D., Casanova, C. & Brent, L.J.N. (2023). Kinship composition in mammals. *Royal Society Open Science*, 10, 230486.

Perrin, N. & Mazalov, V. (2000). Local competition, inbreeding, and the evolution of sex-biased dispersal. *The American Naturalist*, 155, 116–127.

Pujol, B., Blanchet, S., Charmantier, A., Danchin, E., Facon, B., Marrot, P., *et al.* (2018). The Missing Response to Selection in the Wild. *Trends in Ecology & Evolution*, 33, 337–346.

Pusey, A.E. (1987). Sex-biased dispersal and inbreeding avoidance in birds and mammals. *Trends in Ecology & Evolution*, 2, 295–299.

Quaas, R.L. (1988). Additive Genetic Model with Groups and Relationships. *Journal of Dairy Science*, 71, 1338–1345.

Räsänen, K. & Hendry, A.P. (2008). Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification. *Ecology Letters*, 11, 624–636.

Reid, J.M. & Arcese, P. (2020). Recent immigrants alter the quantitative genetic architecture of paternity in song sparrows. *Evolution Letters*, 4, 124–136.

Robertson, A. & Lewontin, R. (1968). Population biology and evolution. *The spectrum of genetic variation. Syracuse Univ. Press, New York*, 5–16.

Rochette, N.C., Rivera-Colón, A.G. & Catchen, J.M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, 28, 4737–4754.

Roehr, J.T., Dieterich, C. & Reinert, K. (2017). Flexbar 3.0 – SIMD and multicore parallelization. *Bioinformatics*, 33, 2941–2942.

Saatoglu, D., Niskanen, A.K., Froy, H., Ranke, P.S., Goedert, D., Reid, J., *et al.* (2025). Metapopulation-level analyses reveal positive fitness consequences of immigration in a small bird. *Journal of Animal Ecology*, 94, 1180–1192.

Snyder-Mackler, N., Burger, J.R., Gaydosh, L., Belsky, D.W., Noppert, G.A., Campos, F.A., *et al.* (2020). Social determinants of health and survival in humans and other animals. *Science*, 368, eaax9553.

Székely, T., Moore, A.J. & Komdeur, J. (2010). *Social behaviour: genes, ecology and evolution*. Cambridge University Press.

Vullioud, C., Davidian, E., Wachter, B., Rousset, F., Courtiol, A. & Höner, O.P. (2019). Social support drives female dominance in the spotted hyaena. *Nature Ecology & Evolution*, 3, 71–76.

Walsh, B. & Lynch, M. (2018). *Evolution and selection of quantitative traits*. Oxford University Press.

Waples, R.S. (2010). Spatial-temporal stratifications in natural populations and how they affect understanding and estimation of effective population size. *Molecular Ecology Resources*, 10, 785–796.

Watts, H.E., Scribner, K.T., Garcia, H.A. & Holekamp, K.E. (2011). Genetic diversity and structure in two spotted hyena populations reflects social organization and male dispersal. *Journal of Zoology*, 285, 281–291.

Wemmer, C. (1973). Kruuk, H. The Spotted Hyena, a study of predation and social behavior. Univ. Chicago Press, xvi + 335 pp., 1972. *Journal of Mammalogy*, 54, 553–554.

Wilson, A.J., Réale, D., Clements, M.N., Morrissey, M.M., Postma, E., Walling, C.A., *et al.* (2010). An ecologist's guide to the animal model. *Journal of Animal Ecology*, 79, 13–26.

Wolak, M.E. (2012). nadiv: an R package to create relatedness matrices for estimating non-additive genetic variances in animal models. *Methods in Ecology and Evolution*, 3, 792–796.

Wolak, M.E. & Reid, J.M. (2017). Accounting for genetic differences among unknown parents in microevolutionary studies: how to include genetic groups in quantitative genetic animal models. *Journal of Animal Ecology*, 86, 7–20.

Yearsley, J.M., Viard, F. & Broquet, T. (2013). THE EFFECT OF COLLECTIVE DISPERSAL ON THE GENETIC STRUCTURE OF A SUBDIVIDED POPULATION. *Evolution*, 67, 1649–1659.

Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*, 30, 614–620.

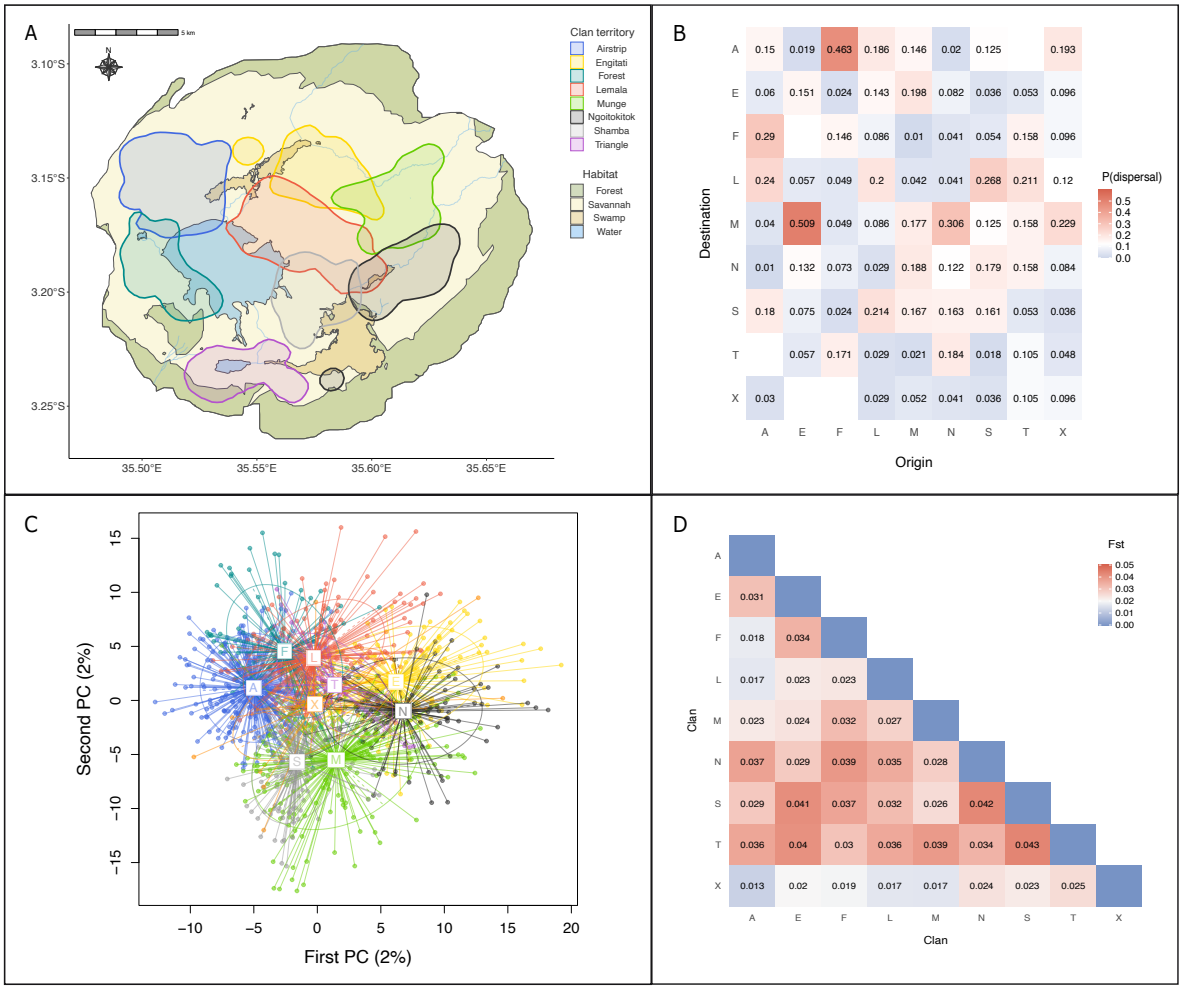

Figure 1. Summary plots describing the distribution of individuals and genetic variation across the population of spotted hyenas in Ngorongoro Crater, Tanzania. (A) Map of the Ngorongoro Crater and each clan's territory (see methods) (B) Heatmap showing the probability of male dispersal between all clans, where the diagonal element of the matrix describes the proportion of males that remain philopatric to each clan. (C) Summary of the first two axes of genomic PCA analysis describing the distribution of genetic variation. Each point is an individual and is coloured by the clan they were born into, and ellipses group individuals born into the same clan. (D) Heatmap of pairwise Fst's between clans. Clan names: A = Airstrip, F = Forest, T = Triangle, E = Engitati, L = Lemala, S = Shamba, M = Munge, N = Ngoitokitok, X = non-Crater (i.e., immigrant).

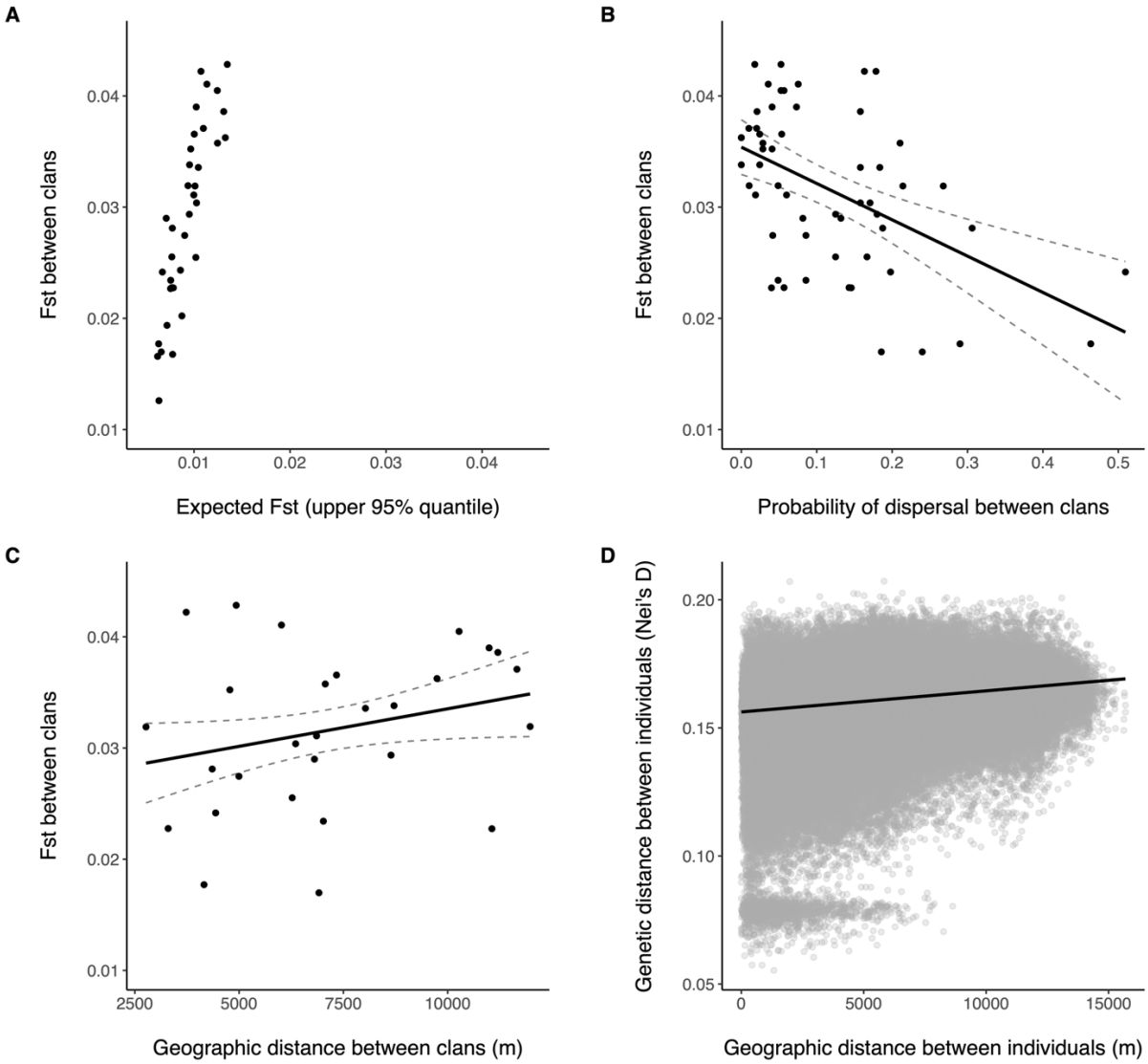

Figure 2. Results from analyses used to test for a relationship between genetic differentiation, male dispersal and geographic distance between clans in spotted hyenas in Ngorongoro Crater. (A) Relationship between pairwise F_{ST} between clans expected when male dispersal between clans is random (x axis) and observed F_{ST} between all clans. Upper bound of that range of expected F_{ST} under random dispersal is shown under the expectation that if observed F_{ST} was equal or lower to this expected value, it could have occurred by chance, (B) relationship between probability of dispersal between clans and genetic differentiation between clans, which is measured as F_{ST} (solid line shows the mean predicted relationship with standard errors as dashed lines), (C) relationship between geographic distance (in metres) between the centroid of each clan's territory and genetic differentiation between clans, which is measured as F_{ST} (solid line shows the mean predicted relationship with standard errors as dashed lines), (D) Isolation-by-distance. Each point is an individual and genetic distance on the y axis is measured as Nei's genetic distance between individuals. Geographic distance on the x axis is measured as linear distance in metres between the centroids of individuals home ranges. Note that diagonals of all

matrices were removed prior to running analyses as we were interested in patterns occurring *between* clans rather than *within* clans.

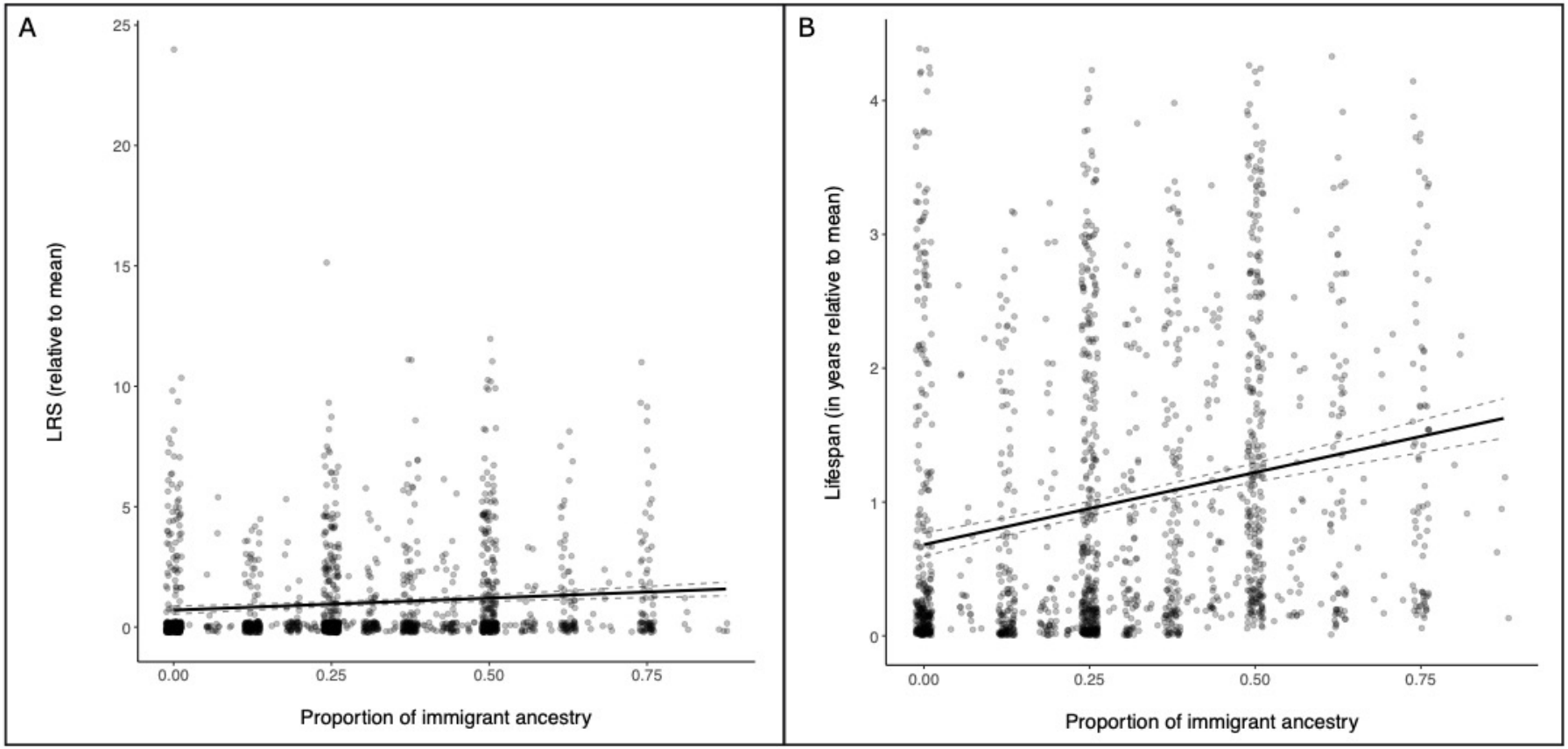

Figure 3. The relationship between the proportion of ancestry an individual has from immigrants (i.e., ancestors that immigrated into the Crater from outside the Crater) and fitness, measured as lifetime reproductive success (LRS, A) and lifespan (B). Both fitness measures are plotted relative to the population mean (mean LRS = 1.74, mean lifespan = 4.15). Points have been jittered across the x axis, the solid black line demonstrates the predicted effect of immigrant ancestry and dotted lines demonstrate the 95% confidence intervals on that prediction.

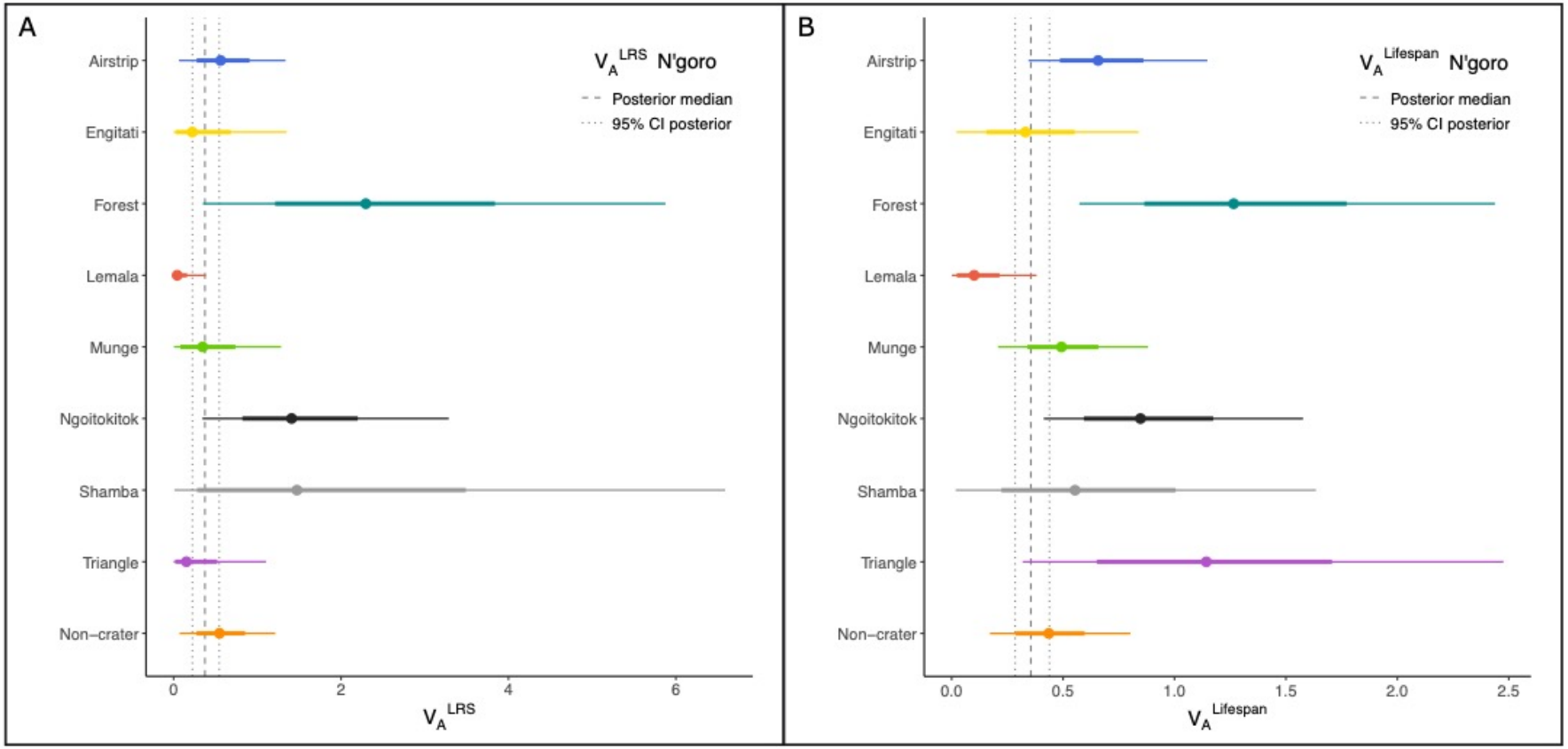

Figure 4. Estimates of additive genetic variance (V_A) for two measures of fitness (lifetime reproductive success, LRS, and lifespan) in each clan of spotted hyenas in Ngorongoro Crater. Estimates of V_A are on relative fitness (i.e., relative to the global population mean). Points show the median estimate of V_A, thick bars show the 75% credible intervals of the full posterior distribution and thin lines show 95% credible intervals of the full posterior. Vertical lines intersect estimated V_A for the population that was derived from a model that does not estimate V_A for each clan independently. Dashed lines show the median of the posterior of those estimates and dotted lines show the 95% credible intervals of the posterior.

Table 1. Genetic diversity of each clan of spotted hyenas in Ngorongoro Crater, Tanzania. Standard deviations for all estimates given in parentheses, showing the variation across individuals in all estimates.

Clan	Observed heterozygosity	Expected heterozygosity	F_{IS}
Airstrip	0.262 (0.172)	0.256 (0.164)	-0.016 (0.095)
Engitati	0.259 (0.181)	0.251 (0.17)	-0.022 (0.111)
Forest	0.262 (0.180)	0.254 (0.168)	-0.024 (0.114)
Lemala	0.261 (0.176)	0.254 (0.166)	-0.020 (0.101)
Munge	0.260 (0.176)	0.254 (0.167)	-0.012 (0.108)
Ngoitokitok	0.260 (0.184)	0.250 (0.171)	-0.026 (0.12)
Shamba	0.261 (0.182)	0.251 (0.169)	-0.027 (0.116)
Triangle	0.264 (0.192)	0.251 (0.173)	-0.040 (0.146)
Non-crater	0.262 (0.168)	0.266 (0.159)	0.017 (0.172)

Observed heterozygosity (He) estimated as the proportion of heterozygous loci per individual, averaged across the population. Expected heterozygosity estimated as He * (2 * N_{ID} / (2 * N_{ID} – 1)), where N_{ID} is the number of individuals in that clan. F_{IS} is the inbreeding coefficient and is estimated as 1 – (Heterozygosity / Expected heterozygosity). All estimates calculated in the *dartR* R package (Gruber & Georges, 2018).

Table 2. Parameter estimates for fixed and random effects estimated from a genetic group animal model used to estimate additive genetic variance for two measures of fitness (lifetime reproductive success, LRS, and lifespan). Both fitness measures were modelled relative to the global population mean and were fit with Gaussian errors. All parameter estimates shown are posterior medians with 95% credible intervals shown in parentheses.

Parameter	LRS	Lifespan
Intercept	3.295 (2.58 - 4.042)	2.329 (1.852 - 2.755)
Sex_{Males}	-0.808 (-1.13 - -0.509)	-0.237 (-0.382 - -0.093)
Social rank	-0.022 (-0.031 - -0.012)	-0.006 (-0.01 - 0)
F	-5.101 (-10.171 - 0.254)	-2.875 (-5.399 - -0.218)
Clan size	-0.02 (-0.03 - -0.011)	-0.01 (-0.015 - -0.005)
$GG^{Airstrip}$	-1.024 (-1.805 - -0.241)	-1.244 (-1.766 - -0.824)
$GG^{Engitati}$	-1.629 (-2.538 - -0.686)	-1.929 (-2.46 - -1.34)
GG^{Forest}	-1.527 (-2.627 - -0.334)	-1.71 (-2.42 - -0.998)
GG^{Lemala}	-1.419 (-2.26 - -0.562)	-1.819 (-2.249 - -1.319)
GG^{Munge}	-0.852 (-1.577 - -0.058)	-1.455 (-1.979 - -1.028)
$GG^{Ngoitokitok}$	-1.359 (-2.301 - -0.374)	-1.606 (-2.182 - -1.028)
$GG^{Triangle}$	-1.479 (-2.841 - -0.156)	-2.096 (-2.759 - -1.359)
GG^{Shamba}	-1.611 (-2.596 - -0.517)	-1.298 (-1.876 - -0.723)
Sex_M:Social rank	0.018 (0.006 - 0.027)	0.004 (-0.001 - 0.009)
V_A	0.379 (0.226 - 0.546)	0.360 (0.285 - 0.443)
V_{Mother}	0.15 (0.033 - 0.269)	0.011 (0 - 0.034)
V_{Cohort}	0.134 (0.035 - 0.321)	0.155 (0.062 - 0.298)
V_{Clan}	0.159 (0.003 - 0.544)	0.02 (0 - 0.073)
V_R	3.208 (2.976 - 3.442)	0.629 (0.564 - 0.696)

Social rank describes the position in the social hierarchy, where 1 is the highest rank and descends to N = clan size. F = individuals inbreeding coefficient estimated from the pedigree. GG = genetic group, which estimates the proportion of ancestry each individual has from each clan. Immigrant ancestry was fit as the reference level, so all parameter estimates show the effect of ancestry to each clan relative to immigrant ancestry. V_A is the additive genetic effects variance, V_{Mother} is the maternal effect, V_{Cohort} is the variance associated with between birth-year effects V_{Clan} is the variance between clans, and V_R is the residual variance.

Table 3. Variance estimates for random effects from genetic group animal models used to estimate additive genetic variance for two measures of fitness (lifetime reproductive success, LRS, and lifespan) in each clan of spotted hyenas in Ngorongoro Crater. Both measures of fitness were modelled relative to the global population mean, and were fit with Gaussian errors. Results shown for set of models with clan-specific V_A estimates. Estimates shown are posterior medians with 95% credible intervals shown in subscript parentheses.

Parameter	LRS	Lifespan
$V_A^{Airstrip}$	0.57 (0.096 - 1.187)	0.66 (0.372 - 0.992)
$V_A^{Engitati}$	0.372 (0.003 - 1.113)	0.358 (0.048 - 0.75)
V_A^{Forest}	2.46 (0.524 - 5.091)	1.273 (0.652 - 2.117)
V_A^{Lemala}	0.087 (0 - 0.328)	0.12 (0.002 - 0.313)
V_A^{Munge}	0.409 (0.009 - 1.044)	0.484 (0.221 - 0.812)
$V_A^{Ngoitokitok}$	1.5 (0.48 - 2.849)	0.88 (0.462 - 1.421)
V_A^{Shamba}	1.88 (0.027 - 5.479)	0.668 (0.083 - 1.54)
$V_A^{Triangle}$	0.279 (0.001 - 0.942)	1.179 (0.446 - 2.161)
$V_A^{Non-Crater}$	0.585 (0.127 - 1.16)	0.434 (0.212 - 0.718)
V_{Mother}	0.074 (0.003 - 0.171)	0.011 (0 - 0.035)
V_{Cohort}	0.228 (0.071 - 0.471)	0.179 (0.071 - 0.338)
V_{Clan}	0.084 (0 - 0.326)	0.028 (0 - 0.1)
$V_R^{Airstrip}$	4.515 (3.859 - 5.276)	0.672 (0.537 - 0.832)
$V_R^{Engitati}$	2.152 (1.765 - 2.615)	0.526 (0.398 - 0.673)
V_R^{Forest}	6.23 (4.852 - 7.927)	0.443 (0.299 - 0.61)
V_R^{Lemala}	2.203 (1.881 - 2.552)	0.678 (0.568 - 0.808)
V_R^{Munge}	3.537 (3.039 - 4.098)	0.667 (0.535 - 0.805)
$V_R^{Ngoitokitok}$	2.706 (2.148 - 3.352)	0.627 (0.482 - 0.802)
V_R^{Shamba}	2.836 (2.234 - 3.513)	0.518 (0.375 - 0.679)
$V_R^{Triangle}$	1.436 (0.981 - 2.003)	0.804 (0.45 - 1.262)

V_A is the additive genetic effects variance, V_{Mother} is the maternal effect, V_{Cohort} is the variance associated with between birth-year effects V_{Clan} is the variance between clans, and V_R is the residual variance.