Spatial familial networks to infer demographic structure of wild populations

Samantha McFarlane[1], Micheline Manseau[1], and Paul Wilson[1,1]

[1]Trent University

September 2, 2020

Abstract

1. In many social species, reproductive success varies between individuals within a population, resulting in socially structured populations. Social network analyses of familial relationships may provide insights on how fitness influences population-level demographic patterns. These methods have however rarely been applied to genetically-derived pedigree data from wild populations. 2. Here we use social networks to reconstruct parent-offspring relationships and create a familial network from polygamous boreal woodland caribou (Rangifer tarandus caribou) in Saskatchewan, Canada, to inform recovery efforts. We collected samples from 933 individuals at 15 variable microsatellite loci along with caribou-specific primers for sex identification. Using social network metrics, we assess the contribution of individual caribou to the population with several centrality metrics and then determine which metrics are best suited to inform on the population demographic structure. We look at the centrality of individuals from eighteen different local areas, along with the entire population. 3. We found substantial differences in centrality of individuals in different local areas, that in turn contributed differently to the full network, highlighting the importance of analyzing social networks at different scales. The full network revealed that boreal caribou in Saskatchewan form a complex, interconnected social network with strong familial ties, as the removal of edges with high betweenness did not result in distinct subgroups. Alpha, betweenness, and eccentricity centrality were the most informative metrics to characterize the population demographic structure and for spatially identifying areas of highest fitness levels and social cohesion across the range. 4. Synthesis and applications: Our results demonstrate the value of different network metrics in assessing genetically-derived familial networks. The spatial application of the familial networks identified areas of higher fitness levels and social cohesion across the range in support of population monitoring and recovery efforts.

Keywords

boreal caribou; dispersal; network theory; pedigree reconstruction; *Rangifer tarandus* ; familial network; social network analysis; social structure.

Introduction

The influence of genetic relatedness on social interactions of conspecifics has been an ongoing question in conservation of wild populations for a long time (Wilson, 1975). On the one hand, pedigree reconstructions have provided insights into mating patterns, individual fitness, and social and genetic structure of populations (Harcourt, Kingston, Cameron, Waas, & Hindell, 2006; Lucena-Perez et al., 2018; vonHoldt et al., 2008). On the other hand, analyses of social connectivity and stability of individuals have produced valuable information on the viability of wild populations, especially at current selective environments due to human-induced environmental changes (Snijders, Blumstein, Stanley, & Franks, 2017). Yet social network

analyses and genetically-derived pedigree reconstruction have been used as two separate methodological frameworks to assist conservation of wild populations. The combination of these methods may highlight the interconnectedness between individuals, differences in reproductive success, and ultimately inform on the demographic structure of a population.

Reconstructing a reasonably complete and accurate familial network from pedigree data is especially relevant for endangered species, providing information on mating patterns and reproductive success (Lucena-Perez et al., 2018; Manlik et al., 2016). However, collecting reliable parentage information for cryptic and elusive species is difficult or directly unfeasible; pedigree information obtained through direct field observations are often limited to females, and may consistently overlook cryptic mating (Coltman et al., 1999; Gottelli, Wang, Bashir, & Durant, 2007). Molecular markers, such as microsatellites, have been used to infer parentage and familial relationships in wild populations (Pemberton, 2008) and assess individual heterogeneity in survival and reproduction (Bolnick et al., 2011; Hamel, Gaillard, Festa-Bianchet, & Côté, 2009; Kendall, Fox, Fujiwara, & Nogeire, 2011). Such heterogeneity can be the result of a number of common processes, such as persistent social rank (e.g. Holst et al., 2002; Stockley & Bro-Jørgensen, 2011), unequal allocation during parental care (e.g. Manser & Avey, 2000; Johnstone, 2004), fine-scale spatial habitat heterogeneity (Bollinger & Gavin, 2004; Franklin, Anderson, Gutiérrez, & Burnham, 2000; Manolis, Andersen, & Cuthbert, 2002), and genetics (Meyers & Bull, 2002; Nussey, 2005).

Social networks have been used to investigate complex webs of interconnections between individuals, providing an array of measurements of individual sociality and the extent to which an individual is connected to other individuals (Borgatti, Mehra, Brass, & Labianca, 2009; Wasserman & Faust, 1994). Several network-based measures are commonly used in social network analysis to quantify indirect connections between individuals in a network (Table 1). Although some network-based centrality measures may overlap, each measure captures a distinct aspect of the social network of a population; individuals with high scores for one measure may not necessarily have a high score in other measures (Brent, 2015; Sueur, Jacobs, Amblard, Petit, & King, 2011). Highly directly connected individuals may not necessarily be highly indirectly connected; individuals with the same degree (same number of social ties) may have different betweenness, depending if that individual's partners are from the same subgroup (low betweenness) or from different subgroups (high betweenness). In sperm whale (*Physeter macrocephalus*) association networks, centrality varied between and within individuals, with one sperm whale having the highest scores for strength and eigenvector centrality, but the lowest clustering coefficient (Lusseau, Whitehead, & Gero, 2008). In a captive chimpanzee (*Pan troglodytes*) population, the individuals with the largest degree (number of grooming partners) were not the individuals with the greatest betweenness or clustering coefficient (Kanngiesser, Sueur, Riedl, Grossmann, & Call, 2011). The extent to which an individual is directly and indirectly connected to the network has considerable quantifiable differences. Strongly directly connected individuals with weak indirect connections may have great influence over their immediate social partners, but have minimal influence over the rest of the population, whereas individuals that are weakly directly connected but with strong indirect connections may be the single tie that links otherwise unconnected network sections together, exerting considerable influence over the population (Brent, 2015).

Here, we infer population demographic structure by assessing different node-based metrics of centrality obtained from a familial pedigree network. First, we use microsatellite data to identify parent-offspring relationships and construct a spatial familial network from all relationships (familial pedigree) of boreal caribou in Saskatchewan, Canada. Then we create a spatial familial network to identify local area networks with varying distributions of centrality metrics, determining whether high centrality metrics and edge-to-node ratios at the fine scale correspond to high centrality in the full network. We also assess the community structure and cohesiveness within the full network using edge removal to identify boundaries that run between subgroups, with a particular focus on parts of the range presenting different levels of anthropogenic disturbance. Our findings allow us to discuss how different metrics of network centrality can be used to spatially identify areas of highest fitness levels and social cohesion across the landscape in support of population monitoring and recovery efforts.

Materials and Methods

Boreal caribou are part of the Boreal Caribou designatable unit (COSEWIC, 2011), listed as Threatened under the federal *Species at Risk Act* (Environment Canada, 2012) and as Vulnerable in Saskatchewan (SKCDC, 2020). In response to the listing, the Government of Saskatchewan initiated a comprehensive monitoring program along with range planning efforts with the goal of achieving a self-sustaining boreal caribou population (Johnson et al., 2020; Saskatchewan Ministry of Environment, 2013). More information on Saskatchewan's boreal caribou ecozones can be found in Appendix 3.

Fecal pellet collection and genetic analysis

We used samples from across the boreal caribou range in Saskatchewan, Canada, collected during winters of 2013-2019 (Figure S3.1; Table 2). This dataset was assembled primarily from systematic non-invasive fecal pellet surveys where aerial transects were systematically flown using a fixed-wing aircraft to locate caribou cratering locations (sites where caribou paw to uncover terrestrial lichens). Additional samples (90) from the northern part of the Saskatchewan Boreal Shield were obtained from blood blots or vials collected from individual boreal caribou handled during radio-collaring (McLoughlin et al., 2016; Priadka et al., 2018). All samples were kept frozen at -20°C until DNA extraction was performed.

In order to generate individual-specific genetic profiles and familial pedigree networks, DNA samples were amplified at 15 variable microsatellite loci (BM848, BM888, Map2C, Bishop et al. (1994); FCB193, Buchanan & Crawford (1993); NVHRT16, Røed & Midthjell (1998); OHEQ, Jones, Levine, & Banks (2000); RT1, RT5, RT6, RT7, RT9, RT13, RT24, RT27, RT30, Wilson, Strobeck, Wu, & Coffin (1997)) along with caribou-specific Zfx/Zfy primers for sex identification. DNA was extracted by removing the mucosal layer of cells coating the fecal pellets and followed the extraction protocol outlined in Ball et al. (2007). Microsatellite alleles were scored with the program GeneMarker® (SoftGenetics, State College, PA) and followed a protocol documented in Flasko et al. (2017). Unique individuals were identified using the program ALLELEMATCH (Galpern, Manseau, Hettinga, Smith, & Wilson, 2012). We retained samples that amplified at [?]5 loci and re-amplified apparent unique genetic profiles represented by a single sample using two independent scorers to confirm unique individual identities (Hettinga et al., 2012). An error rate per locus was calculated using these re-amplification results.

Defining familial relationships between individuals

We identified familial relationships of boreal caribou in the study area by reconstructing parent-offspring relationships using COLONY v2.0.6.5 (Jones & Wang, 2010). We calculated population allele frequencies using GenAlEx v6.5 (Peakall & Smouse, 2012). Input parameters were set to allow for female and male polygynous mating systems without inbreeding avoidance, and the probability of mothers or fathers being present in the sampled data set was set to 50% in the absence of other prior information. All sampled females were set as possible mothers, and all sampled males were set as possible fathers. COLONY infers the parental genotypes for each individual; inferred parents are genotypes that are not included in the candidate parent samples, either through that individual's genotype not being captured during sampling, or that parent is no longer living, resulting in a family network with more individuals than were sampled. Finally, individual fitness was calculated with the number of offspring each individual produced.

Modeling the social structure of the population

Identifying parts of the network that are highly connected and those individuals that are less connected to the network can help define the local and global structure of the familial network. We used the r package *CINNA* (Ashtiani, Mirzaie, & Jafari, 2018) to calculate individual node-based measures of network centrality. Nodes represent individuals and edges represent parent-offspring relationships, with directionality from parent to offspring. We calculated five direct and indirect node-based measures of centrality for each individual to

quantify distinct aspects of centrality: alpha, betweenness, closeness, degree, and eccentricity centrality (Table 1). We calculated correlation coefficients between measures to only select statistically independent aspects of centrality. We used principal component analysis (PCA) to collapse variance among any dependent centrality measures, as suggested by Brent (2015), and to identify the most the most important centrality types based on our network structure. We used the r package *FactoMineR* (Le, Josse, & Husson, 2008) to run the PCA, and package *factoextra* (Kassambara & Mundt, 2020) to visualize PCA results.

Network analysis

As boreal caribou mating systems are polygamous, with individuals having multiple mating partners, a dense and complicated network is created; visually analyzing the aspatial network along with the node-based measures of network centrality allows for easier identification of patterns and trends within the network. We used Cytoscape v3.7.2 (Shannon et al., 2003) for the non-spatial analyses of the local and full familial networks. We created the familial network from the reconstructed parent-offspring relationships identified by COLONY. As each individual has their parents identified by COLONY, as well their offspring, a network can be created from the multigenerational relationships among individuals.

To detect community structure and assess network cohesiveness within the full network, we used the Girvan-Newman algorithm to look for boundaries that run between social groups to find natural divisions within the network by removing edges with the highest betweenness scores (Girvan & Newman, 2002; Lusseau & Newman, 2004; Newman & Girvan, 2004). We used an edge betweenness centrality measure (Freeman, 1977) calculated in the NetworkAnalyzer (Assenov, Ramırez, Schelhorn, Lengauer, & Albrecht, 2007) plugin for Cytoscape. Edge betweenness quantifies how often an edge is crossed when moving between any pair of individuals in the network; bottlenecks are identified in edges that have higher betweenness, as these edges are passed the most often when connecting individuals. Edges were systematically removed until social groups can be identified.

Spatial application of network analysis

We examined how local areas presenting high and low edge-to-node ratios (Box 1) contributed to the full network by comparing centrality metrics across local areas within the network. The local areas were of management interest, had a comparable number of individuals and similar geographic sizes. We plotted the spatial locations of all sampled individuals and parent-offspring relationships in ArcGIS (ESRI Inc., 2018) to spatially identify local areas. From these, we selected areas with the highest and lowest ratio of edges (parent-offspring relationships) to nodes (individuals) to compare local area networks within the larger spatial familial network. We examined the centrality metrics for all sampled individuals within each local area network, as well as for their first neighbours (individuals one degree away from individuals in these areas - as inferred parents do not have spatial locations, this captures inferred individuals) and compared each local area network.
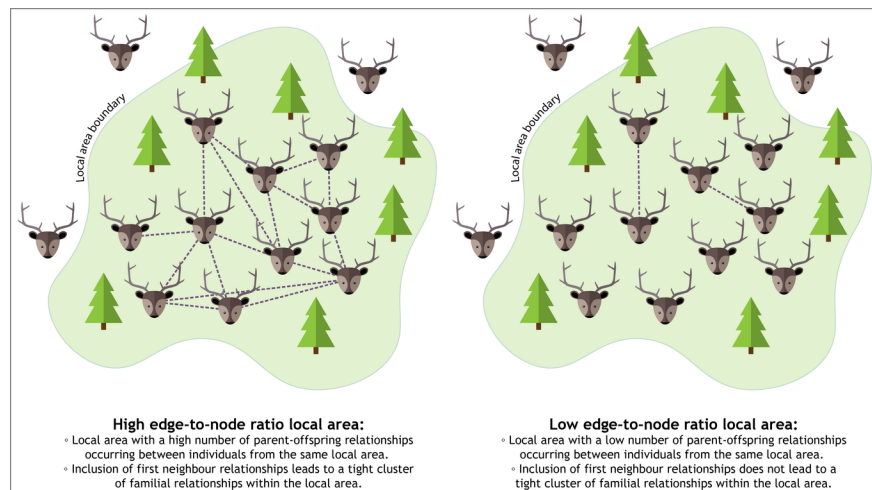
Box 1: Edge-to-node ratio definition for local areas.

Results

A total of 2,198 samples were collected (2,099 fecal and 99 blood blot). 1,970 were successfully scored (average success rate of 91.4%), and 933 unique individuals were identified (Table 2). Overall, the average dropout rate was 0.0028% and the average false allele rate was 0.011%. Pedigree reconstruction inferred 310 females and 319 males, for a total familial network of 1,562 individuals. 355 inferred and sampled males and 360 inferred and sampled females were identified as parents. 1,487 (95.2%) individuals were linked in one network, with the remaining 75 individuals linked in five smaller clusters (Figure S2.1). We used the 1,487 individuals identified in the primary network for calculating node-based metrics of centrality.

Spatial network analysis

Local area networks

We identified 18 local area networks in order to determine the cohesiveness and centrality of individuals. The local areas with the lowest edge-to-node ratios were all located in the northern Boreal Shield, with the high edge-to-node ratio areas found further south in the western Boreal Plains and southern Boreal Shield (Figure 1). We found significant differences between the distribution of centrality metrics between high and low edge-to-node ratio local areas (Table S2.2, Figure 2). The largest edge-to-node ratio was Canoe Lake in the western Boreal Plains (ratio of 15; Table S2.1, Figure S2.3). We identified three other local areas with similarly high edge-to-node ratios (Figure S2.4, Figure S2.5, Figure S2.6, Table S2.1). The smallest edge-to-node ratio (Central SK Shield) had zero parent-offspring relationships (Table S2.1; Figure S2.7). We identified two other local areas with similarly low edge-to-node ratios, with very few parent-offspring relationships occurring within these local areas (Figure S2.8-Figure S2.9, Table S2.1), indicating that Boreal Shield individuals are not presenting the same proximity to related individuals as observed in the Boreal Plains. Overall, edge-to-node ratios correlated positively to closeness (Figure S2.2A), alpha (Figure S2.2C), betweenness (Figure S2.2D), and degree centrality (Figure S2.2E). However, edge-to-node ratios decreased with eccentricity centrality (Figure S2.2B).

When bringing in the first neighbours of all individuals within a local area, the high edge-to-node ratio areas formed a tighter cluster of individuals than in the low edge-to-node ratio areas. Including first neighbours in the area with the highest edge-to-node ratio (Canoe Lake) increased the ratio to 1.14 and connected 73.6%

of individuals into one cluster (Figure S2.3). A large proportion of each high edge-to-node ratio local area became connected into one or two large clusters with the inclusion of first neighbours (Figure S2.4, Figure S2.5, Figure S2.6). In comparison, including first neighbours in the lowest edge-to-node ratio local area (Central SK Shield) increased the ratio to 0.86, but did not connect many individuals into one cluster (only 12.8% of individuals; Figure S2.7). The large variation in local area edge-to-node ratios show that individuals within the high edge-to-node ratio areas in the Boreal Plains and southern Boreal Shield are forming tighter family groups (high number of parent-offspring relationships within each local area).

Full network

Individuals from high edge-to-node ratio local areas were located more centrally within the full family network and clustered with other individuals from the same local area. Individuals from low edge-to-node ratio local areas were dispersed throughout the network and primarily found on the outer edges of the network (Figure 3). Although all local areas were of similar geographic size (Figure 1), individuals from low edge-to-node ratio local areas were not closely connected to each other in the network. Individuals from these local areas were not found within a few edges of other individuals from the same local area, indicating that individuals encountered in each low edge-to-node ratio local area are from different familial lines, or are dispersers that were sampled in that local area (Figure 3); as the edges in the familial network represent parent-offspring relationships, these individuals are likely not highly related to one another, and do not form a cohesive social group. In contrast, individuals from high edge-to-node local areas were highly connected to one another within the full network, indicating they are closely related, with a high density of familial ties (parent-offspring relationships).

Removal of edges with high betweenness did not alter the overall network structure (Figure S2.10). Most edges within the network had low betweenness centrality (score of 1 – 81.5% of edges; Table 3). Only 2.97% of edges were removed after sequentially removing edges with the highest edge betweenness score until only edges with an edge betweenness > 4 remained (Table 3), with no clear delineation of groups within the main network (Figure S2.10), but the removal of these edges led to individuals becoming disconnected from the main network. The high edge-to-node ratio local areas of Canoe Lake, Peter Pond Lake, and SK2West remained central and clustered within the edge removal network (Figure S2.10). Individuals from Trade Lake maintained a high level of clustering, but became separated from the main network, forming a separate subgroup (Figure S2.10). Removal of high betweenness edges did not result on individuals from low edge-to-node ratio areas becoming separate subgroups; individuals remained dispersed throughout the network (Figure S2.10).

Discussion

Studies have applied social network analysis to quantify and explore the social structure of populations across numerous taxa (Brent, 2015; Brent, Lehmann, & Ramos-Fernández, 2011; Sueur et al., 2011), but to our knowledge this is the first to combine genetically-derived pedigree data with social network analysis to infer social structure of wild populations. Social network analyses are powerful and flexible methods for investigating the complex networks of interconnections between individuals within and between populations, providing numerous measures of individual sociality (Brent, 2015; Wasserman & Faust, 1994). With a large interconnected network of 1,562 nodes (individuals) and 1,866 edges (parent-offspring relationships) between individuals, it can be difficult to identify significant differences within the network. By bringing the familial network into a spatial framework and incorporating aspatial node-based centrality metrics, we were able to identify different levels of sociality within the network, with some local areas composed of stable family groups and others that are less sedentary. Comparing local area networks of management interest allowed us to identify areas of higher and lower fitness and connectivity in the overall boreal caribou familial network.

There are numerous network centrality metrics available, and many are often correlated (Brent, 2015; Freeman, 1977; Newman, 2003; Sueur et al., 2011; Wasserman & Faust, 1994; Wey & Blumstein, 2012). Indirect

measures of centrality are commonly used in social network analysis to identify central individuals within a network (Brent, 2015; Wey, Blumstein, Shen, & Jordán, 2008), but there has been little research done to determine which of these metrics are suitable for measuring centrality in pedigree-based familial networks. As familial networks are not built on associations between individuals or constructed through direct observation, measures of network centrality apply differently, and not all metrics may be suitable for familial-based networks. By identifying local areas within the network, we were able to gain a better understanding of which areas contributed most to the familial network. We found significant differences in centrality measures between local areas in the full familial network, and these variations in individual centrality would have remained hidden if only the full familial network was examined. We used five centrality metrics in our social network analysis of familial networks (Figure 2), and found that alpha, betweenness, and eccentricity centrality were the most informative measures of individual centrality (Figure S1.1). Degree centrality in familial networks represents the parents of an individual (in-degree) and the offspring of an individual (out-degree), giving a direct measure of an individual's reproductive output and fitness levels. It is important to note, however, that inferred individuals in the pedigree will always have an in-degree of 0, as it is not possible to infer the parents of inferred individuals. Alpha centrality is an important metric for familial networks, as it indicates those individuals who are connected to individuals who themselves are highly connected, giving an indication of individual status, even if that individual does not have a lot of direct connections (offspring). Reproductive output can be highly asymmetrical, with the number of offspring varying between individuals (McFarlane et al., 2018), and alpha centrality can indicate if that individual is part of a high-fitness family if they are connected to highly connected individuals. We found that local areas with high edge-to-node ratios had a wider distribution of alpha and degree centrality, indicating that more higher fitness individuals are found in these local areas than in low edge-to-node local areas (Figure 2C), and are better connected to other well-connected individuals. Three of the four high edge-to-node ratio local areas we identified are located in the western part of Saskatchewan's Boreal Plains, which has the highest levels of both anthropogenic and fire disturbance in the Boreal Plains (Figure S3.2), and the tight family groups we observed in these areas may be a result of decreased dispersal propensity due to high levels of fragmentation between local areas.

Betweenness centrality is another important metric for network analysis, as it captures the interconnectedness of subgroups; individuals with high betweenness interact with individuals who do not interact with one another, therefore making betweenness important for maintaining group cohesion, and connecting disparate parts of the network (Brent, 2015). Our familial network was not comprised of subgroups, as most individuals (94.2%) had a betweenness centrality of 0, and 95.2% of all sampled individuals formed one large familial network. Even after the removal of edges with the highest edge betweenness, the overall network structure did not change, with most individuals still connected in one main network, with no clear subgroups (Figure S2.10). Our study species displays a polygamous mating system, with individuals potentially having multiple partners, producing a complex network of parent-offspring relationships and full- and half-siblings, with high interconnectedness among individuals across the network (Figure S2.1). Our highly interconnected network with no evidence of subgroups and low average betweenness centrality is the result of the polygamous mating system and high dispersal ability.

The high eccentricity centrality and low closeness centrality informs on the social structure and the presence of animals dispersing longer distances in the Boreal Shield. The Boreal Shield is less fragmented than the Boreal Plains, with significantly less anthropogenic disturbance (Figure S3.2; Table S3.1). Very few parent-offspring relationships occurred within or between the northern Boreal Shield local areas (Figure 1). This suggests that individuals in the Boreal Shield are not central to the familial network and have lower individual fitness, not reproducing many offspring that survive until fall (low degree centrality). Individuals in low edge-to-node local areas are not from the same familial lines and are not highly related to any other individuals in the network. The removal of high betweenness edges led to some individuals becoming disconnected from the full network, but these disconnected individuals were not from one local area, instead located throughout both ecozones, again highlighting the interconnectedness of the familial network.

In most animal network studies, nodes represent observed individuals, with relationships between pairs of individuals (dyads) defined by an association index (the time the pair of individuals spent together), with

edges representing observed relationships, forming an interaction network (Morrison, 2016; Whitehead & Dufault, 1999). For many species, it is not possible or feasible to directly observe rare and elusive species, and therefore association information cannot be obtained. Pedigree reconstruction can give direct information about dyads between closely related individuals (parent-offspring and full-siblings), with these relationships forming the basis of the familial network. In comparison to association networks, in familial networks, only the sampled individuals are known or observed, and the edges between individuals and the unsampled individuals (parents) are inferred by the data analysis (Morrison, 2016). Reconstructing a familial network from genetically-derived pedigree data gives valuable information about the number of mating partners, the number of offspring, and the structure of the reproductive network of a population (McFarlane et al., 2018; Pemberton, 2008). Pedigrees represent historical and evolutionary connections between generations; these relationships have long been recognized as reticulating but are instead commonly presented as simplified trees instead of networks, where reticulations caused by inbreeding are absent (Morrison, 2016). Pedigrees represent a network of relationships, and therefore reconstructed pedigrees inherently contain information that can be used to construct a network. With a wide spectrum of mating systems present in species (Clutton-Brock, 1989), almost all species have pedigree networks, with multiple partners and/or offspring attributed to each individual, therefore creating a complex network of familial relationships (Morrison, 2016). Although boreal caribou have a skewed reproductive rate, with varying levels of individual fitness (McFarlane et al., 2018), our network does not appear to be vulnerable to sudden population crashes resulting from changes in social structure or social isolation and inbreeding. Due to the polygamous mating system and long-range dispersal ability, the boreal caribou network is highly interconnected, and removal of edges with high betweenness did not change the overall network structure or lead to distinct clusters or social groups, although family groups can be identified within the network, with varied levels of dispersal and fitness levels among family groups.

Social networks are powerful methods to assist in wildlife conservation (Snijders et al., 2017), but most wild populations cannot be directly observed, and social association networks cannot be constructed. By constructing a familial network based on genetically-derived parent-offspring relationships, we calculated informative metrics to draw a much finer picture of their individual fitness levels, pattern of demographic structure, and relative contribution of local areas to the larger population. The spatial application of the familial network allowed us to identify areas of higher fitness levels and social cohesion across the range in support of population monitoring and recovery efforts.

Author's contributions

SM implemented the analysis with contributions from MM. PW coordinated DNA extraction and analysis. SM wrote the first draft of the manuscript with help from MM. All authors contributed to subsequent drafts and gave final approval for publication and have no conflicts of interest to declare.

Acknowledgements

Data availability

Data deposited via the Dryad Digital Repository (*link*).

References

Ashtiani, M., Mirzaie, M., & Jafari, M. (2018). CINNA: An r/CRAN package to decipher central informative nodes in network analysis.*Bioinformatics* , *35* (8), 1436–1437. doi: 10.1093/bioinformatics/bty819

Assenov, Y., Ramırez, F., Schelhorn, S.-E., Lengauer, T., & Albrecht, M. (2007). Computing topological parameters of biological networks. *Bioinformatics* ,*24* (2), 282–284. doi: 10.1093/bioinformatics/btm554

Ball, M. C., Pither, R., Manseau, M., Clark, J., Petersen, S. D., Kingston, S., . . . Wilson, P. (2007). Characterization of target nuclear DNA from faeces reduces technical issues associated with the assumptions of low-quality and quantity template. *Conservation Genetics* , *8* (3), 577–586. doi: 10.1007/s10592-006-9193-y

Bishop, M. D., Kappes, S. M., Keele, J. W., Stone, R. T., Sunden, S. L., Hawkins, G. A., . . . Yoo, J. (1994). A genetic linkage map for cattle. *Genetics* ,*136* (2), 619–639.

Bollinger, E. K., & Gavin, T. A. (2004). Responses of nesting bobolinks (*Dolichonyx oryzivorus*) to habitat edges. *The Auk* , *121* (3), 767. doi: 10.1642/0004-8038(2004)121[0767:ronbdo]2.0.co;2

Bolnick, D. I., Amarasekare, P., Araujo, M. S., Burger, R., Levine, J. M., Novak, M., . . . Vasseur, D. A. (2011). Why intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution* , *26* (4), 183–192. doi: 10.1016/j.tree.2011.01.009

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science* , *323* (5916), 892–895. doi: 10.1126/science.1165821

Brent, L. J. N. (2015). Friends of friends: Are indirect connections in social networks important to animal behaviour? *Animal Behaviour* , *103* , 211–222. doi: 10.1016/j.anbehav.2015.01.020

Brent, L. J. N., Lehmann, J., & Ramos-Fernández, G. (2011). Social network analysis in the study of nonhuman primates: A historical perspective. *American Journal of Primatology* , *73* (8), 720–730. doi: 10.1002/ajp.20949

Buchanan, F. C., & Crawford, A. M. (1993). Ovine microsatellites at the OarFCB11, OarFCB128, OarFCB193, OarFCB266 and OarFCB304 loci. *Animal Genetics* ,*24* , 145. doi: 10.1111/j.1365-2052.1993.tb00269.x

Clutton-Brock, T. H. (1989). Review lecture: Mammalian mating systems. *Proceedings of the Royal Society of London. B. Biological Sciences* , *236* (1285), 339–372. doi: 10.1098/rspb.1989.0027

Coltman, D. W., Bancroft, D. R., Robertson, A., Smith, J. A., Clutton-brock, T. H., & Pemberton, J. M. (1999). Male reproductive success in a promiscuous mammal: Behavioural estimates compared with genetic paternity. *Molecular Ecology* , *8* (7), 1199–1209. doi: 10.1046/j.1365-294x.1999.00683.x

COSEWIC. (2011).*Designatable units for caribou (Rangifer tarandus in Canada* (p. xi + 88 pp.).

Environment Canada. (2012).*Recovery strategy for the woodland caribou (Rangifer tarandus caribou), Boreal population, in Canada. Species at Risk Act Recovery Strategy Series* . Ottawa. xi + 138 pp.

ESRI Inc. (2018). *ArcGIS 10.6.1* .

Flasko, A., Manseau, M., Mastromonaco, G., Bradley, M., Neufeld, L., & Wilson, P. (2017). Fecal DNA, hormones, and pellet morphometrics as a noninvasive method to estimate age class: An application to wild

populations of central mountain and boreal woodland caribou (*Rangifer tarandus caribou*).*Canadian Journal of Zoology* , *95* (5), 311–321. doi: 10.1139/cjz-2016-0070

Franklin, A. B., Anderson, D. R., Gutiérrez, R. J., & Burnham, K. P. (2000). Climate, habitat quality, and fitness in Northern Spotted Owl populations in northwestern California. *Ecological Monographs* , *70* (4), 539–590. doi: 10.1890/0012-9615(2000)070[0539:chqafi]2.0.co;2

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry* ,*40* (1), 35. doi: 10.2307/3033543

Galpern, P., Manseau, M., Hettinga, P., Smith, K., & Wilson, P. (2012). Allelematch: An R package for identifying unique multilocus genotypes where genotyping error and missing data may be present. *Molecular Ecology Resources* ,*12* (4), 771–778. doi: 10.1111/j.1755-0998.2012.03137.x

Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks.*Proceedings of the National Academy of Sciences* , *99* (12), 7821–7826. doi: 10.1073/pnas.122653799

Gottelli, D., Wang, J., Bashir, S., & Durant, S. M. (2007). Genetic analysis reveals promiscuity among female cheetahs. *Proceedings of the Royal Society B: Biological Sciences* , *274* (1621), 1993–2001. doi: 10.1098/rspb.2007.0502

Hamel, S., Gaillard, J.-M., Festa-Bianchet, M., & Côté, S. D. (2009). Individual quality, early-life conditions, and reproductive success in contrasted populations of large herbivores. *Ecology* , *90* (7), 1981–1995. doi: 10.1890/08-0596.1

Harcourt, R. G., Kingston, J. J., Cameron, M. F., Waas, J. R., & Hindell, M. A. (2006). Paternity analysis shows experience, not age, enhances mating success in an aquatically mating pinniped, the weddell seal (leptonychotes weddellii).*Behavioral Ecology and Sociobiology* , *61* (4), 643–652. doi: 10.1007/s00265-006-0294-x

Hettinga, P. N., Arnason, A. N., Manseau, M., Cross, D., Whaley, K., & Wilson, P. J. (2012). Estimating size and trend of the North Interlake woodland caribou population using fecal-DNA and capture-recapture models. *Journal of Wildlife Management* , *76* (6), 1153–1164. doi: https://doi.org/10.1002/jwmg.380

Holst, D. von, Hutzelmeyer, H., Kaetzke, P., Khaschei, M., Rodel, H. G., & Schrutka, H. (2002). Social rank, fecundity and lifetime reproductive success in wild european rabbits (oryctolagus cuniculus). *Behavioral Ecology and Sociobiology* , *51* (3), 245–254. doi: 10.1007/s00265-001-0427-1

Johnson, C. A., Sutherland, G. D., Neave, E., Leblond, M., Kirby, P., Superbie, C., & McLoughlin, P. D. (2020). Science to inform policy: Linking population dynamics to habitat for a threatened species in Canada. *Journal of Applied Ecology* , *57* (7), 1314–1327. doi: 10.1111/1365-2664.13637

Johnstone, R. A. (2004). Begging and sibling competition: How should offspring respond to their rivals? *The American Naturalist* , *163* (3), 388–406. doi: 10.1086/375541

Jones, K. C., Levine, K. F., & Banks, J. D. (2000). DNA-based genetic markers in black-tailed and mule deer for forensic applications. *Calif. Fish Game* , *86* (2), 115–126.

Jones, O. R., & Wang, J. (2010). COLONY: A program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* , *10* (3), 551–555. doi: 10.1111/j.1755-0998.2009.02787.x

Kanngiesser, P., Sueur, C., Riedl, K., Grossmann, J., & Call, J. (2011). Grooming network cohesion and the role of individuals in a captive chimpanzee group.*American Journal of Primatology* , *73* (8), 758–767. doi: 10.1002/ajp.20914

Kassambara, A., & Mundt, F. (2020). *Factoextra: Extract and visualize the results of multivariate data analyses* . Retrieved from https://CRAN.R-project.org/package=factoextra

Kendall, B. E., Fox, G. A., Fujiwara, M., & Nogeire, T. M. (2011). Demographic heterogeneity, cohort selection, and population growth. *Ecology* , *92* (10), 1985–1993. doi: 10.1890/11-0079.1

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R Package for multivariate analysis.*Journal of Statistical Software* , *25* (1). doi: 10.18637/jss.v025.i01

Lucena-Perez, M., Soriano, L., López-Bao, J. V., Marmesat, E., Fernández, L., Palomares, F., & Godoy, J. (2018). Reproductive biology and genealogy in the endangered iberian lynx: Implications for conservation. *Mamm. Biol.* ,*89* (WOS:000426540000002), 7–13. doi: 10.1016/j.mambio.2017.11.006

Lusseau, D., & Newman, M. E. J. (2004). Identifying the role that animals play in their social networks. *Proceedings of the Royal Society of London. Series B: Biological Sciences* , *271* (suppl_6). doi: 10.1098/rsbl.2004.0225

Lusseau, D., Whitehead, H., & Gero, S. (2008). Incorporating uncertainty into the study of animal social networks. *Animal Behaviour* , *75* (5), 1809–1815. doi: 10.1016/j.anbehav.2007.10.029

Manlik, O., McDonald, J. A., Mann, J., Raudino, H. C., Bejder, L., Krützen, M., . . . Sherwin, W. B. (2016). The relative importance of reproduction and survival for the conservation of two dolphin populations. *Ecology and Evolution* ,*6* (11), 3496–3512. doi: 10.1002/ece3.2130

Manolis, J. C., Andersen, D. E., & Cuthbert, F. J. (2002). Edge effect on nesting success of ground nesting birds near regenerating clearcuts in a forest-dominated landscape. *The Auk* , *119* (4), 955. doi: 10.1642/0004-8038(2002)119[0955:eeonso]2.0.co;2

Manser, M. B., & Avey, G. (2000). The effect of pup vocalisations on food allocation in a cooperative mammal, the meerkat (*Suricata suricatta*).*Behavioral Ecology and Sociobiology* , *48* (6), 429–437. doi: 10.1007/s002650000248

McFarlane, S., Manseau, M., Flasko, A., Horn, R. L., Arnason, N., Neufeld, L., . . . Wilson, P. (2018). Genetic influences on male and female variance in reproductive success and implications for the recovery of severely endangered mountain caribou. *Global Ecology and Conservation* , *16* , e00451. doi: 10.1016/j.gecco.2018.e00451

McLoughlin, P. D., Stewart, K., Superbie, C., Perry, T., Tomchuk, P., Greuel, R., . . . Johnstone, J. F. (2016). *Population dynamics and critical habitat of woodland caribou in the saskatchewan boreal shield. Interim project report, 2013-2016.* (p. 162). Saskatoon: Department of Biology, University of Saskatchewan.

Meyers, L. A., & Bull, J. J. (2002). Fighting change with change: Adaptive variation in an uncertain world. *Trends in Ecology & Evolution* , *17* (12), 551–557. doi: 10.1016/s0169-5347(02)02633-2

Morrison, D. A. (2016). Genealogies: Pedigrees and phylogenies are reticulating networks not just divergent trees. *Evolutionary Biology* , *43* (4), 456–473. doi: 10.1007/s11692-016-9376-5

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review* ,*45* (2), 167–256. doi: 10.1137/s003614450342480

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks.*Physical Review E* , *69* (2). doi: 10.1103/physreve.69.026113

Nussey, D. H. (2005). Selection on heritable phenotypic plasticity in a wild bird population.*Science* , *310* (5746), 304–306. doi: 10.1126/science.1117004

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in excel. Population genetic software for teaching and researching - an update.*Bioinformatics* , *28* , 2537–2539. doi: 10.1093/bioinformatics/bts460

Pemberton, J. M. (2008). Wild pedigrees: The way forward. *Proceedings of the Royal Society B: Biological Sciences* , *275* (1635), 613–621. doi: 10.1098/rspb.2007.1531

Priadka, P., Manseau, M., Trottier, T., Hervieux, D., Galpern, P., McLoughlin, P. D., & Wilson, P. J. (2018). Partitioning drivers of spatial genetic variation for a continuously distributed population of boreal caribou: Implications for management unit delineation. *Ecology and Evolution* . doi: 10.1002/ece3.4682

Røed, K. H., & Midthjell, L. (1998). Microsatellites in reindeer, *Rangifer tarandus* , and their use in other cervids. *Mol. Ecol.* , *7* (12), 1773–1776. doi: 10.1046/j.1365-294x.1998.00514.x

Saskatchewan Ministry of Environment. (2013). *Conservation strategy for boreal woodland caribou (Rangifer tarandus caribou) in Saskatchewan* . Saskatchewan Ministry of Environment.

Saskatchewan Ministry of Environment. (2019). *Range plan for woodland caribou in Saskatchewan: Boreal plain ecozone, SK2 Central Caribou Administrative Unit* . Saskatchewan Ministry of Environment.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., . . . Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research* ,*13* (11), 2498–2504. doi: 10.1101/gr.1239303

SKCDC. (2020). Species at risk in Saskatchewan. Retrieved from http://biodiversity.sk.ca/SppList/verts.pdf

Snijders, L., Blumstein, D. T., Stanley, C. R., & Franks, D. W. (2017). Animal social network theory can help wildlife conservation. *Trends in Ecology & Evolution* , *32* (8), 567–577. doi: 10.1016/j.tree.2017.05.005

Stockley, P., & Bro-Jørgensen, J. (2011). Female competition and its evolutionary consequences in mammals. *Biological Reviews* , *86* (2), 341–366. doi: 10.1111/j.1469-185x.2010.00149.x

Sueur, C., Jacobs, A., Amblard, F., Petit, O., & King, A. J. (2011). How can social network analysis improve the study of primate behavior? *American Journal of Primatology* , *73* (8), 703–719. doi: 10.1002/ajp.20915

vonHoldt, B. M., Stahler, D. R., Smith, D. W., Earl, D. A., Pollinger, J. P., & Wayne, R. K. (2008). The genealogy and genetic viability of reintroduced yellowstone grey wolves. *Mol. Ecol.* , *17* (1), 252–274. doi: 10.1111/j.1365-294X.2007.03468.x

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* . Cambridge, U.K.: Cambridge University Press.

Wey, T., Blumstein, D. T., Shen, W., & Jordán, F. (2008). Social network analysis of animal behaviour: A promising tool for the study of sociality. *Animal Behaviour* ,*75* (2), 333–344. doi: 10.1016/j.anbehav.2007.06.020

Wey, T. W., & Blumstein, D. T. (2012). Social attributes and associated performance measures in marmots: Bigger male bullies and weakly affiliating females have higher annual reproductive success. *Behavioral Ecology and Sociobiology* ,*66* (7), 1075–1085. doi: 10.1007/s00265-012-1358-8

Whitehead, H., & Dufault, S. (1999). Techniques for analyzing vertebrate social structure using identified individuals: Review and recommendations. In *Advances in the study of behavior* (pp. 33–74). Elsevier. doi: 10.1016/s0065-3454(08)60215-6

Wilson, E. O. (1975).*Sociobiology: The new synthesis* . Harvard University Press.

Wilson, G. A., Strobeck, C., Wu, L., & Coffin, J. W. (1997). Characterization of microsatellite loci in caribou *Rangifer tarandus* , and their use in other artiodactyls.*Molecular Ecology* , *6* (7), 697–699. doi: 10.1046/j.1365-294X.1997.00237.x

Supporting Information

Additional supporting information may be found online in the Supporting Information section.

Tables

Table 1: Node-based measures of connectivity.

Metric

Type

Definition

Alpha centrality

Indirect

Alpha centrality of all vertices. A generalization of eigenvector centrality to directed graphs (Bonanich 2001).

Betweenness centrality

Indirect

Reflects the amount of control that a node exerts over the interactions of other nodes in the network.

Closeness centrality

Indirect

A measure of how fast information spreads from a given node to other reachable nodes in the network. The Latora closeness centrality is used in networks with disconnected components (Latora and Marchiori 2001).

Degree centrality

Direct

The number of edges connected to a node.

Eccentricity centrality

Indirect

The maximum non-infinite length of a shortest path between n and another node in the network.

Table 2: Sampling data

Sampling Area

Survey Years

Sample Type

Number of Samples Collected

Number of Samples Successfully Scored

Number of Unique Genotypes

Genotyping Success (%)

Dropouts (%)

False Alleles (%)

Flin Flon

13

2014

Fecal

336

320

104

95.2

0.0077

0.032

La Ronge

2013 & 2015

Fecal

497

403

162

81.1

0.0032

0.0097

SK Boreal Plains West

2016

Fecal

242

233

122

96.3

0

0

Patterson Lake

2018

Fecal

21

19

9

90.5

0.0089

0.036

SK2Central

2017

Fecal

452

371

150

82.1

0

0

SK Shield

2014

Fecal

99

98

98

99

0

0

SKShield

2019

Blood

551

526

288

95.5

0

0

Total

-

-

2,198

1,970

933

-

-

-

Table 3: Edge betweenness scores for each edge in the full familial network.

Edge betweenness

Count

20

1

18

1

12

5

9

2

7

5

6

9

5

30

4

26

3

201

2

50

1

1454

Table 4: Correlation coefficients between node-based measures of network connectivity.

Degree Centrality

Eccentricity Centrality

Betweenness Centrality

Closeness Centrality

Alpha Centrality

-0.216

-0.124

0.152

0.208

Degree Centrality

-0.118

0.371

0.284

Eccentricity Centrality

-0.11

-0.544

Betweenness Centrality

0.234

Table 5: The amount of variation retained by each principal component.

eigenvalue

variance.percent

cumulative.variance.percent

Dim.1

1.8768792

37.537584

37.53758

Dim.2

1.2519155

25.038310

62.57589

Dim.3

0.9720742

19.441484

82.01738

Dim.4

0.5130135

10.260270

92.27765

Dim.5

0.3861176

7.722353

100.00000

Hosted file

Social_Network_Manuscript_FINAL_08-12-2020_Figure1.docx available at https://authorea.com/users/309851/articles/478205-spatial-familial-networks-to-infer-demographic-structure-of-wild-populations

Hosted file

Social_Network_Manuscript_FINAL_08-12-2020_Figure2.docx available at https://authorea.com/users/309851/articles/478205-spatial-familial-networks-to-infer-demographic-structure-of-wild-populations

Hosted file

Social_Network_Manuscript_FINAL_08-12-2020_Figure3.docx available at https://authorea.com/users/309851/articles/478205-spatial-familial-networks-to-infer-demographic-structure-of-wild-populations