APPENDIX A: Cumulative frequency distributions

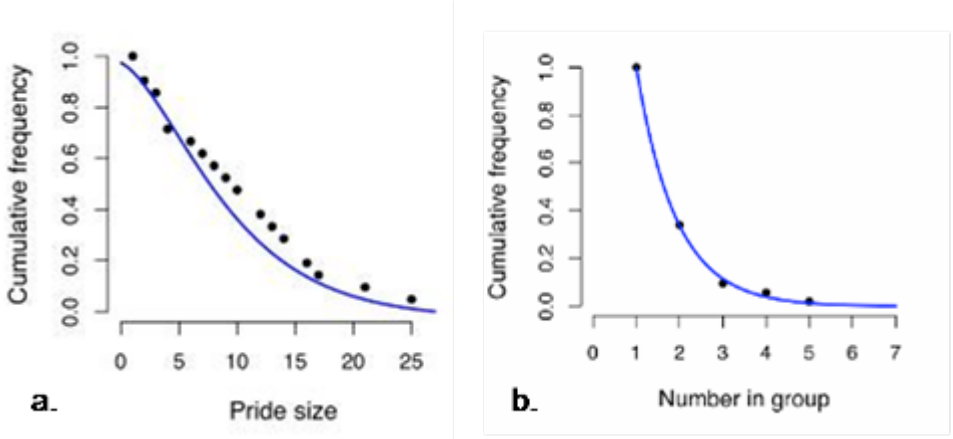


Figure A1. Cumulative frequency distributions of (a) pride size (X_P) and (b) nomadic group size (X_N). The black dots reflect the actual data while the blue lines are the best fit models (given in Table 1).

APPENDIX B: Focal-follow data

We used data from a study involving four-day continuous follows of individual lionesses (Packer et al. 1990; Scheel & Packer 1991) to compare contact rates between lions during the day and at night. We were unable to reject the null hypothesis that contact rates for individual lionesses differed between night and day (directional Wilcoxon Sign-Rank, $p = 0.109$). For each of the focal lionesses, we calculated the per hour (per lion) rate of contact with lions from other prides (mean = 0.0044 ± 0.0035). To compute a comparable quantity from the lion-sighting daytime data used in the model, we assumed that all pride lions over three months old had an equal probability of participating in any given contact with another pride. For each pride we calculated the following estimate of the per lion contact rate:

$$\frac{\text{average no. lions per contact}}{\text{no. of lions in pride}} \cdot \frac{\text{number of inter-pride contacts}}{\text{hour}}.$$

Averaging over all prides, we calculated an average per lion per hour rate of 0.0059 ± 0.00329, consistent with the focal-follow data.

APPENDIX C: Logistic Regression

In a dataset involving 25 prides and 2213 sightings, we observed 30 pride-pride interactions. For any pair of prides *A* and *B*, the binary response variable for our logistic regression analysis was whether or not a sighting of *A* includes an interaction with *B*. Each sighting of pride *A* thus yields whether or not it contacted each of the other 24 prides. We analyzed a logistic regression model given by:

$$\log\left[\frac{p}{1-p}\right] = \alpha + \beta_{\text{dist}} x_{\text{dist}} + \beta_{\text{net}} x_{\text{net}} + \beta_{\text{num}} x_{\text{num}} + \beta_{\text{split}} x_{\text{split}} \qquad (1)$$

where the β terms are logistic regression coefficients and α is the intercept. This relates the interaction probability for a pair of prides (p) with the factors listed above.

Figure A2. Contact rates decrease with distance and shared history. $M_{contact}$ gives the probability that a pride *A* will contact another pride *B* as a function of the network distance between the two prides and whether or not they recently split from each other. We use $M_{contact}$ to calculate these probabilities for every type of possible pride pair (d=1/recent split, d=1/no split, d=2, d=3, d=4, d=5, …). (Note that recent splits only occur for immediately adjacent pairs of prides.) We then use these probabilities as *contact weights* in our simulations. This graph shows the weights normalized so that the largest weight is one (for d=1, recent split pairs). For example, if a pride *A* is adjacent to pride *B* and recently split from pride *B* (d=1, split) and pride *A* has a network distance of two from another pride *C*, then *A* is 22 times more likely to contact *B* than *C*.

APPENDIX D: Neighbor models

We sought to develop a simple predictive model of pride proximity that determines whether two prides are neighbors based on the local distribution of pride centroids. Using the empirical territory data for 25 prides, we ran logistic regressions of the "neighborness" of each pair of prides (0=not neighbor, 1=neighbor) on each of the following quantities and chose the option that best predicted neighbor status. Let *a* and *b* denote the centroid locations of two distinct prides *A* and *B*.

1. The maximum value of cos(*cab*)*distance(*ac*)/distance(*ab*) where *c* is the centroid location of a third pride *C*.
2. Similar to (1) but using cotangent instead of cosine
3. Number of other prides *C* such that cos(*cab*)*distance(*ac*)/distance(*ab*) that exceeds a threshold, where *c* is the centroid location for *C*. Various thresholds were tried.
4. Number of prides within a *joint radius* of prides *A* and *B*. We define joint radius as the number of prides C such that distance(C,A) < distance(A,B) and distance(C,B) < distance(A,B)
5. Square root of number of prides within a joint radius of prides *A* and *B*
6. The number of prides in the union of the two intervening semicircles (Figure A3: gray area).

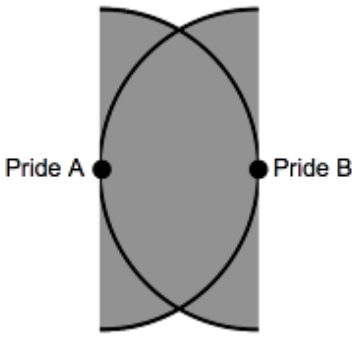


Figure A3. Model 6 is the number of prides in the union of the two overlapping gray semicircles.

We selected model 6 to the exclusion of the other inputs because it had the lowest AIC; we call it the "Neighboring prides model" (M_{neigh}).

This M_{neigh} model produced similar numbers of neighbors vs. non-neighbors; 47% of actual prides in Figure 1a and 53% of simulated prides in Figure 1b are classified as neighbors. More importantly, the M_{neigh} model produced distributions of number of neighboring prides (degree distributions) visually and statistically similar to the empirical data. To assess this, we simulated 30 full populations (each of N=180 prides); sampled 25 prides from a geographically restricted corner of the population (to mimic the number of prides within the geographic constraints of the Serengeti Lion Project study system (Craft et al. 2009)); then performed bootstrap Kolmogorov-Smirnov tests to compare the empirical degree distribution (based on 25 observed prides) and simulated degree distributions (Sekhon in press). We found that the simulated subsets resembled the empirical data: p-values ranged from 0.030 to 0.995 with significant differences at the 5% level for only 2 of the 30 simulated subsets, which is expected by chance; and $p = 0.533$ for the aggregated data (for examples, see Figure A4). However, the degree distributions for the simulated populations as a whole were often significantly different from the empirical degree distribution ($p < 0.05$ for 21 of the 30 simulated populations), confirming the claim in (Craft et al. 2009) that the SLP study area is on the periphery of the network of pride territories.

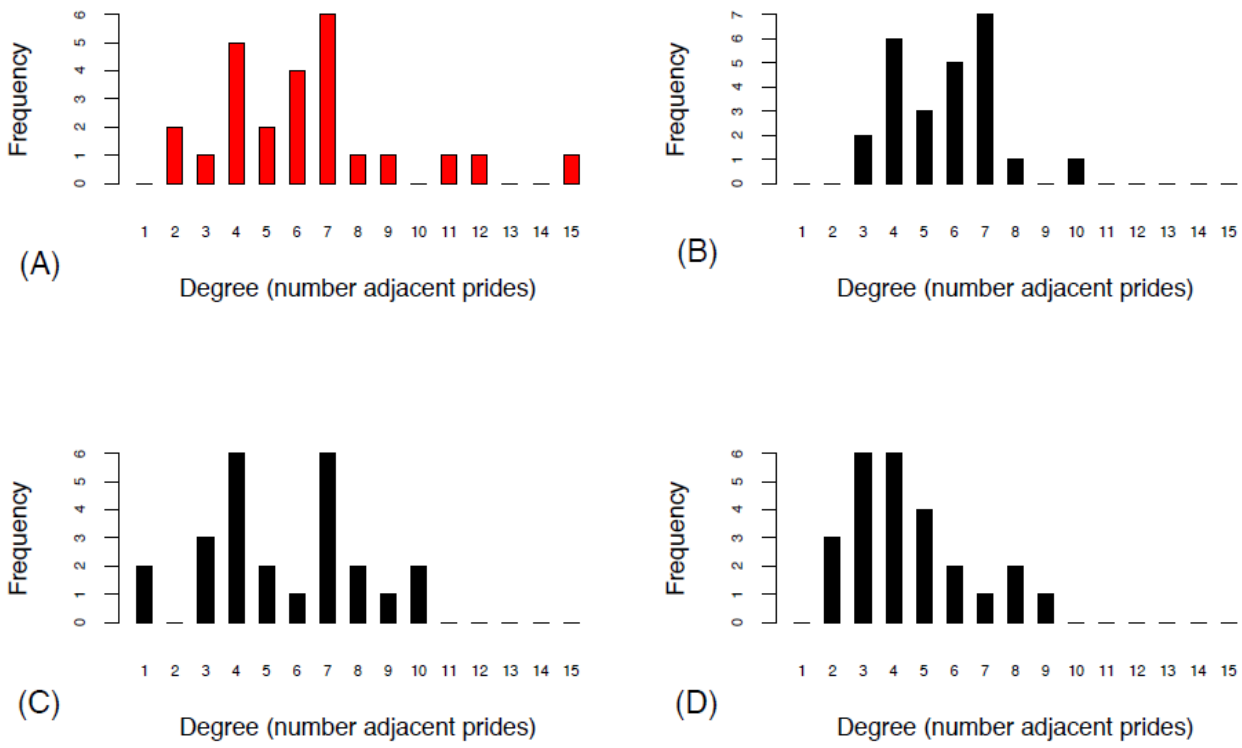


Figure A4. Comparison of empirical degree distribution to three simulated degree distributions. (A) Empirical degree distribution based on 25 prides in SLP study area. (B), (C) Simulated subsets that resemble empirical degree distribution (bootstrap Kolmogorov-Smirnov, $p = 0.87$ and $p = 0.91$, respectively). (D) One of the two simulated subsets (of 30) that differed significantly from the empirical distribution (bootstrap Kolmogorov-Smirnov, $p = 0.04$).

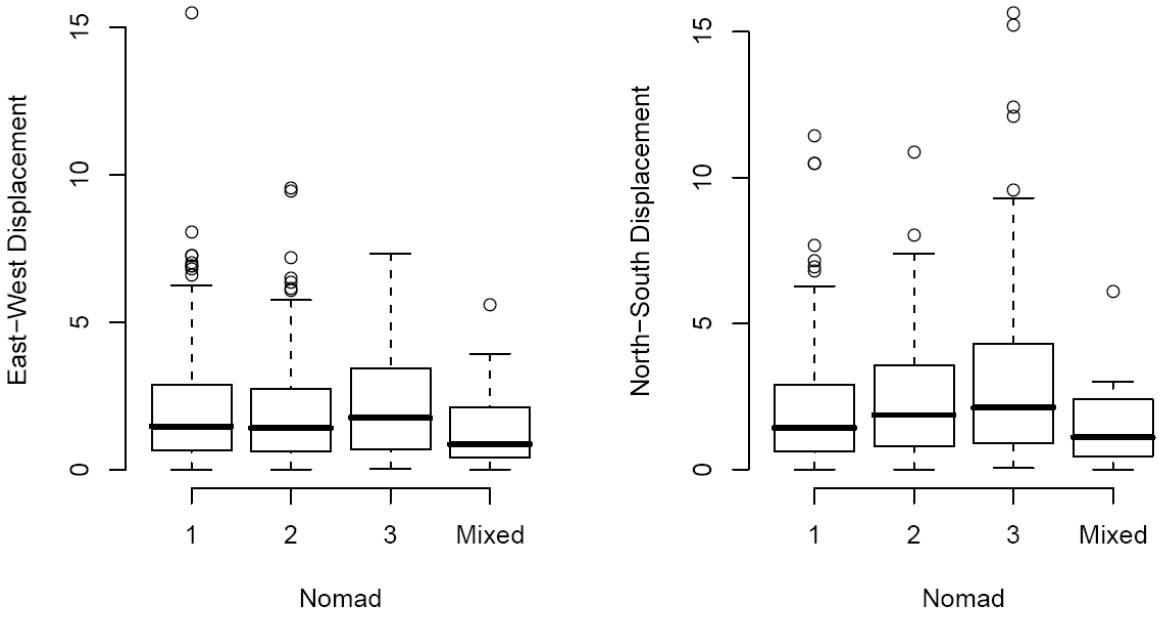
APPENDIX E: Nomad movement model

Variance Gamma (VG) processes are a type of Lévy random walk and can be efficiently simulated by taking the difference of two gamma random variables with scale proportional to time interval of the displacement. Specifically, the observed nomads approximately moved according to the following pair of equations

$$\begin{aligned} x(t+h) &= x(t) + \delta_{x,R}(h) - \delta_{x,L}(h) \\ y(t+h) &= y(t) + \delta_{x,U}(h) - \delta_{x,D}(h) \end{aligned} \qquad (2)$$

where $x(t)$ and $y(t)$ are the horizontal and vertical positions of a nomad at time t; $\delta_{x,R}(h)$ and $\delta_{x,L}(h)$ are the right and left displacements in an interval h and are random variables from identical gamma distributions, Γ_x; $\delta_{x,U}(h)$ and $\delta_{x,D}(h)$ are the upwards and downwards displacements in an interval h and are random variables from identical gamma distributions, $\Gamma_y(\mathrm{M}_{\mathrm{nomad}})$.

APPENDIX F: Comparisons of nomad movements



(A)

(B)

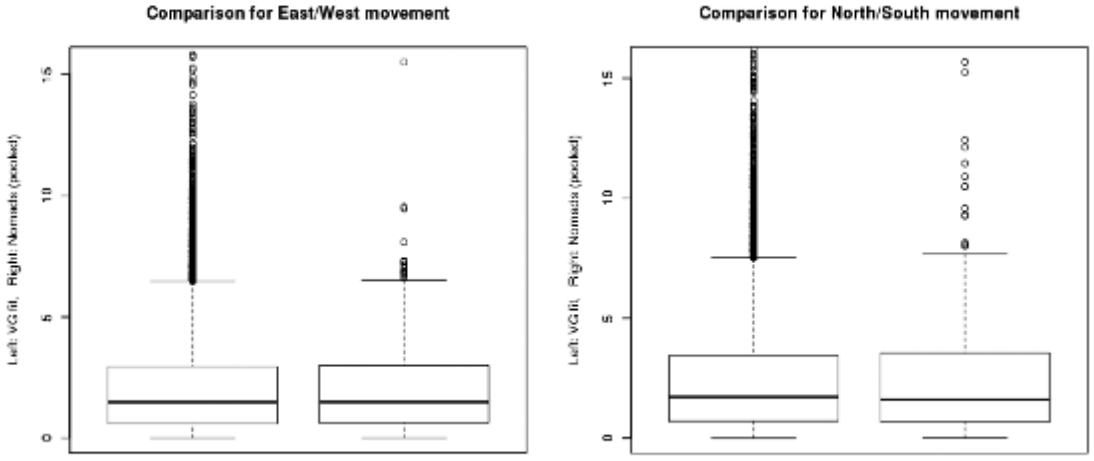


Figure A5. (a) Daily displacements in kilometers for three GPS-collared nomads ("Nomads 1, 2, and 3") and pooled daily displacements for 7 VHF-collared nomads ("Mixed"). Median daily displacements for Nomads 1, 2, and 3 are not significantly different from each other (Kruskal-Wallis $p = 0.1078$), nor are they significantly different from the median displacement for the seven VHF-

collared nomads pooled data (Kruskal-Wallis East/West $p = 0.8876$, North/South $p = 0.1735$). There were 198, 109, and 93 non-overlapping daily displacements for nomads 1, 2 and 3, respectively, and 23 for the VHF-collared nomads. (B) Boxplot comparison of daily displacements (km) for East/West movement and North/South movements for a model nomad simulated under the variance gamma model M_{nomad} (left) vs. the pooled displacements of the three GPS-collared nomads (right).

APPENDIX G: Contact group sizes

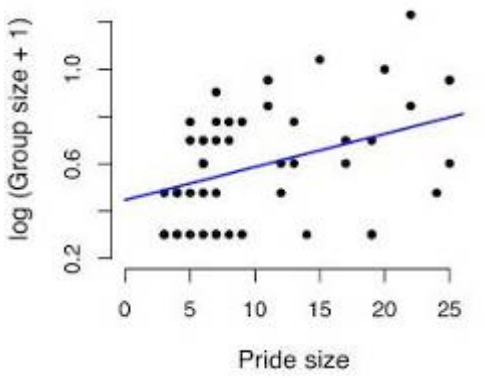


Figure A6. Total pride size vs. group size participating in an inter-pride interaction (G). The black dots reflect empirical data whereas the blue line is the fit of the regression (Regression R^2= 0.138, p = 0.003).

APPENDIX H: SEIR model

Lions frequently contact all other lions in their group, so we assume that any given pride, nomadic group, or resident-male coalition moves through the four disease classes as a unit, as would occur with a novel disease for which there is no immunity in the population. A group is considered exposed when its first member becomes infected; the group transitions stochastically from exposed to infectious at a rate of once per ε days and from infectious to recovered at a rate of once per ι days. We investigated the empirical pattern of within-pride deaths in order to inform the infectious period distribution, but because the data was too sparse (yet consistent with an exponential tail) we relied on convention and accepted the exponential distribution.

In addition, the expected number of transmissions is the same whether or not we consider individual lions progressing through the infectious state. Specifically, consider pride A which becomes infected at t_0. Scenario 1: All lions in the pride recover at the same time. Scenario 2: All lions recover at a random time, with each recovery time chosen independently from the same distribution (i.e., individual infectious periods are independent and identically distributed random variables). The distribution of recovery rates in both cases is exponential with parameter μ. Other variables described in the two scenarios include m_A (mean group size from prides contacted by A), T (transmissibility per lion-to-lion contact), C_P (rate of pride-pride contacts), $\langle G \rangle$ (size of a group of lions from pride A involved in a contact), and N_A (replicate from a distribution of pride size X_P. Derived variables are:

- Expected number of pride contacts before time t = $C_P t$
- Probability that a given lion takes part in a pride contact = $\langle G \rangle / N_A$
- Probability that a pride still infectious at time t recovers = $\mu e^{-\mu(t-t_0)}$
- Probability that a lion still infectious at time t recovers = $\mu e^{-\mu(t-t_0)}$
- Expected number of transmissions per group contact = $m_A \langle G \rangle T$

- Expected number of transmissions by a given lion = $m_A T$

The expected number of transmissions in scenario 1 is:

$$\begin{aligned}
&\int_{t=t_0}^{\infty} \Pr(\text{infectious pride recovers at t}) \\
&\times (\text{Expected number of pride contacts prior to t}) \\
&\times (\text{Expected number of transmissions per group contact})\, dt \\
&= C_p m_A \langle G \rangle T \int_{t=t_0}^{\infty} \mu t e^{-\mu(t-t_0)} \\
&= C_p m_A \langle G \rangle T / \mu
\end{aligned}$$

The expected number of transmissions in scenario 2 is:

$$\begin{aligned}
&\sum_{i=1}^{N_A} \int_{t=t_0}^{\infty} \Pr(\text{infectious lion i recovers at t}) \\
&\times (\text{Expected number of pride contacts prior to t}) \\
&\times (\text{Probability that lion i took part in contact}) \\
&\times (\text{Expected number of transmissions by lion i per group contact})\, dt \\
&= N_A C_p m_A T (\langle G \rangle / N_A) \int_{t=t_0}^{\infty} \mu t e^{-\mu(t-t_0)} dt \\
&= C_p m_A \langle G \rangle T / \mu
\end{aligned}$$

where the expected number of transmissions in both scenarios is the same. It should be noted that sequential infection among pride members and longer latent and infectious periods would likely slow the spread of disease, but would not change the final number infected.

When an infected group (*A*) contacts a susceptible group (*B*), the probability of disease transmission is a function of the number of individuals involved in the interaction and a per-contact transmissibility parameter (*T*), given by

$$\tau_{AB} = \sum_{j,k} p_j q_k \left(1 - (1-T)^{j \cdot k}\right) \qquad (3)$$

where p_j and q_k are the probabilities that the subgroup sizes from *A* and *B* are *j* and *k*, respectively. This assumes that every lion in one subgroup encounters every lion in the other subgroup; this is the

network analog to density-dependence. We believe that this is a realistic assumption because contacts can be of long duration and generally involve small groups (averaging 1.51 nomads and 3.65 pride lions per contact group). When a susceptible coalition of territorial males resides with an infected pride, the coalition is immediately infected; when an infected coalition of territorial males switches to a susceptible pride, it immediately infects the second pride (Craft et al. 2009).

REFERENCES

Craft, M. E., Volz, E., Packer, C. & Meyers, L. A. 2009 Distinguishing epidemic waves from disease spillover in a wildlife population. *Proceedings of the Royal Society B: Biological Sciences* **276**, 1777-1785.

Packer, C., Scheel, D. & Pusey, A. E. 1990 Why Lions Form Groups: Food is Not Enough. *The American Naturalist* **136**, 1-19.

Scheel, D. & Packer, C. 1991 Group hunting behaviour of lions: a search for cooperation. *Animal Behaviour* **41**, 697-709.

Sekhon, J. S. in press Multivariate and propensity score matching software with automated balance optimization. *Journal of Statistical Software*.