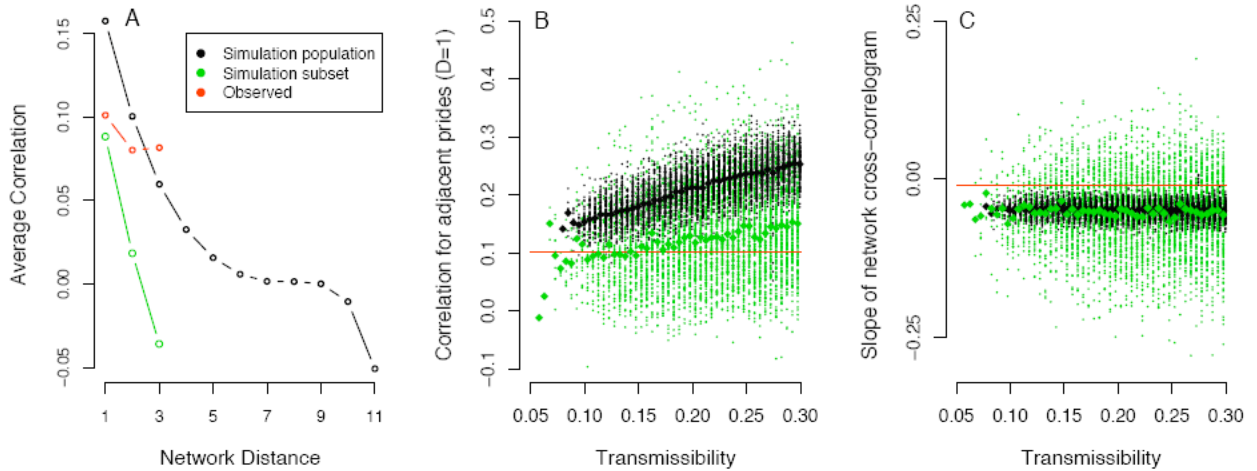


**Table S1.** Geographic and network location versus the probability that a pride is infected during an epidemic. Table S1 gives the results of the multivariate logistic regression analysis for distance to edge, degree, and closeness centrality vs. the probability that a pride is infected during an epidemic. Grey rows highlight significant factors.

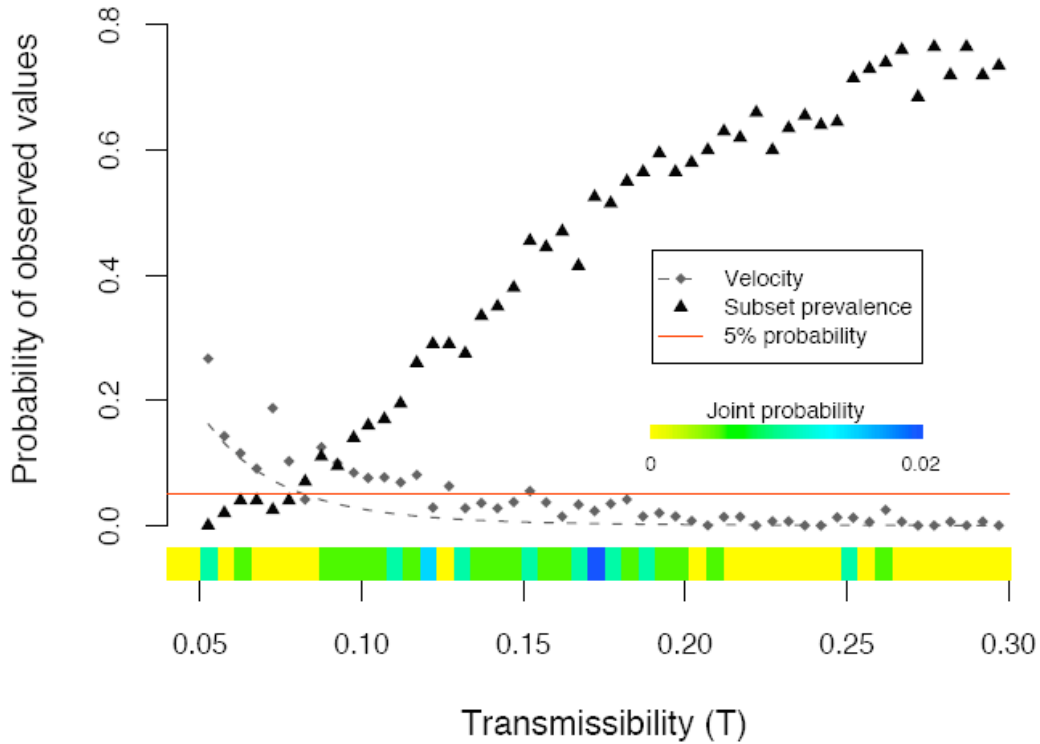
<b>Source</b>	<b>df</b>	<b>Likelihood-ratio chi-square</b>	<b><i>P</i>-value</b>
Distance to edge ( <i>DE</i> )	1	0.09326	0.7601
Degree ( <i>Deg</i> )	1	118.5862	<.0001
Closeness Centrality ( <i>CC</i> )	1	28.5187	<.0001

**Table S1.** Logistic regression of population structure on epidemiological risk.

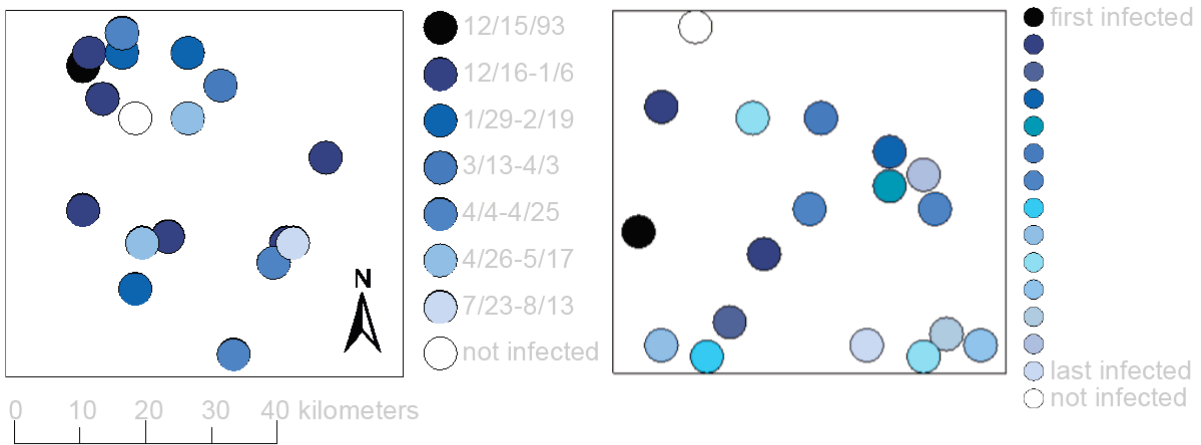
**Video S1.** Online animation of one simulated epidemic (open video link). Red “x” is infected (latent or infectious), blue is recovered, and one frame corresponds to one day ( $T = 0.10$ ). The larger green circles (lower left) comprise the subset. The final outbreak sizes were 17 of 18 prides in the subset and 169 of 180 prides overall; and the epidemic lasted 12.5 months.



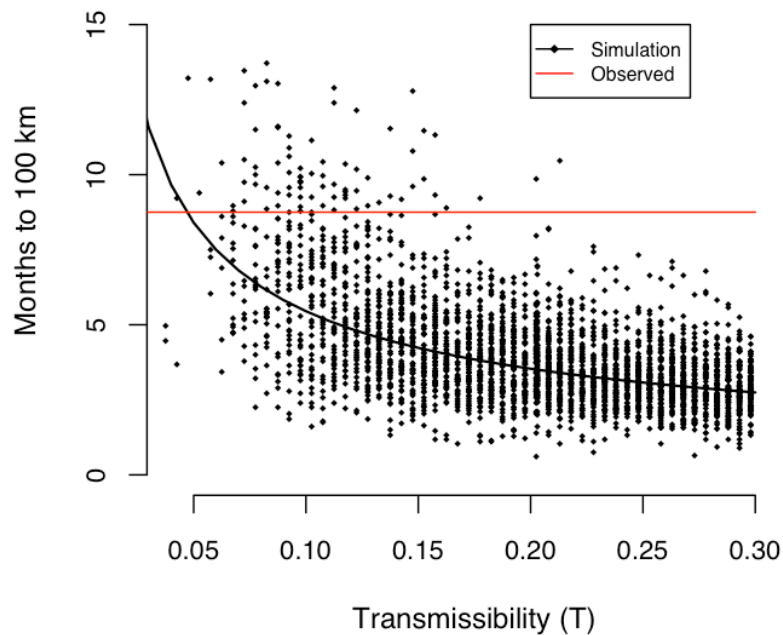
**Figure S1.** Spatial spread of CDV. (A) Network correlograms for simulated and observed epidemics. In simulated epidemics (with  $T = 0.1725$ ), the average correlation in the timing of infectious periods between randomly chosen prides decreases with increasing network distance. Correlations between adjacent prides were lower in both the observed 1994 CDV outbreak (red) and simulated subsets (green). (B) Average correlation in infectious period for all directly adjacent prides in the subset (green) and population (black). Small points show averages from individual simulations and large points show overall means at each transmissibility. The red line is the estimated correlation from the 1994 outbreak. (C) Slope of the network correlograms for the subset (green) and population (black). Small points show slopes from individual simulations and large points show mean slope across all simulations. Red line is the estimated slope from the 1994 outbreak.



**Figure S2.** Sensitivity analysis of model based on 200 replicate simulations at each of 50 transmissibility values. For each simulation, we randomly drew all parameter values from the ranges given in the third column of Table 1. This figure was then calculated using the same methods described for Figure 5. The qualitative and quantitative agreement between the two figures show that the basic conclusion of the paper – that lions probably did not sustain the 1994 CDV epidemic themselves – is robust to uncertainties in the parameters.



**Figure S3.** Spatio-temporal progression of CDV in both the observed study area and a model subset. Disease moves through prides in (A) the observed study area during the 1994 outbreak (the timing of a pride’s infection corresponds to the first date that an infected or seropositive lion from the pride was observed) and (B) a simulated epidemic with  $T = 0.1275$ . The units of time are weeks. The black circle shows the first pride infected and color changes from dark blue to light blue as the epidemic progresses. Empty circles indicate uninfected prides. The rest of the ecosystem would extend from the left and top of each region as in Figure 1A and 1B.



**Figure S4.** Epidemic velocity. Each point represents the time until the disease reached 100 km from the first infected pride for a single simulated epidemic starting at a randomly chosen pride in the subset. The black line shows the least squares linear regression on log-log transformed values. The red line shows the estimated velocity for the observed 1994 outbreak.

**Text S1. Statistical Methods**

*Centrality analysis.* For any given pride, distance to edge is calculated as the shortest Euclidean distance from its centroid to the ecosystem boundary; degree is simply the number of prides with adjacent territories; and closeness centrality is the reciprocal of the sum of shortest paths to all other prides in the population. We calculate shortest paths using the function `networkx.path.all_pairs_shortest_path_length` in the `networkx` software package

<http://networkx.lanl.gov>. The centrality analysis presented in Figure 2 is based on 1400

epidemic simulations at  $T = 0.10$ , each based on a unique randomly generated lion population. For each centrality metric, we (1) calculated the average within the subset and the average overall for every simulation and conducted a Wilcoxon signed rank test on the data, and (2) binned prides into five equal-sized bins and calculated both the fraction of prides infected across all simulations and the fraction of all outbreaks originating at a pride within the bin that ultimately infected at least 50% of prides. Finally, we performed a logistic regression using the three centrality metrics as predictor variables (distance to edge, degree, and closeness centrality) and the infection state of the pride (infected or not during an epidemic) as the response variable (Table S1).

*Network correlograms.* The network correlograms show the average correlation in infectious periods at each discrete network distance class. Each pride has a series of binary-valued disease states  $a_1, a_2, \dots, a_L \in \{0, 1\}$ , where zero and one correspond to uninfected and infected days, respectively, and  $L$  is the length of the epidemic in days. For each simulation, let  $r_{xy}$  denote the Pearson product-moment correlation coefficient between these times series for every pair of prides  $x$  and  $y$ . The average correlation coefficient for network distance  $d$  in a given simulation is given by

$$\langle r_d \rangle = \frac{1}{|\Delta_d|} \sum_{(x,y) \in \Delta_d} r_{xy}$$

where  $\Delta_d$  is the set of all pairs of prides that have minimum path length of  $d$  in the territory network and  $|\Delta_d|$  is the number of pairs in that set. These calculations were adapted from the *R* library *ncf* (Bjornstad & Falck 2001).

The epidemiological data from the 1994 CDV outbreak are incomplete. For each of the 17 infected prides in the study area, the onset of infection was based on the date of first observed death or first confirmed seropositive individual in that pride, whichever occurred first (Supplementary Figure S3). We discarded the date of onset for three prides that only provided serological evidence of infection, which revealed only that infection had occurred at an indefinite time during the 1994 outbreak (for dates of infection categorized by serology or death, see Craft et al 2008). For each of the remaining prides, we stochastically reconstructed the disease state time series assuming that (1) pride infectious periods are random variables distributed exponentially with a mean of two weeks (as in our SEIR model) and (2) the date at which disease was first observed is selected uniformly from the infectious period of the pride. The network correlogram based on observed 1994 data (Figure S1A) gives averages over 1000 time series reconstructions. The network correlograms for the simulations were calculated using complete time-series data. The full population analysis includes all 180 prides, whereas the subset analysis is based on a randomly drawn sample of 15 prides (to replicate the incompleteness of the empirical data).

### **References for Supporting Information**

Bjornstad O. N. and Falck W. 2001 Nonparametric spatial covariance functions: Estimation and testing. *Environ ecol stat* **8**, 53-70.