Transmission of one predicts another: Apathogenic proxies for transmission dynamics of a fatal virus

Marie L.J. Gilbertson[1†], Nicholas M. Fountain-Jones[2*], Jennifer L. Malmberg[3,4*], Roderick B. Gagne[3,5], Justin S. Lee[3], Simona Kraberger[6], Sarah Kechejian[3], Raegan Petch[3], Elliott Chiu[3], Dave Onorato[7], Mark W. Cunningham[8], Kevin R. Crooks[9], W. Chris Funk[10], Scott Carver[2], Sue VandeWoude[3], Kimberly VanderWaal[1], Meggan E. Craft[1,11]

[†]Corresponding author: jone1354@umn.edu

*These authors contributed equally

[1]Department of Veterinary Population Medicine, University of Minnesota, St Paul, MN 55108.

[2]School of Natural Sciences, University of Tasmania, Hobart Australia 7001.

[3]Department of Microbiology, Immunology, and Pathology, Colorado State University, Fort Collins, CO 80523

[4]Department of Veterinary Sciences, University of Wyoming, Laramie, Wyoming 82071.

[5]Wildlife Futures Program, Department of Pathobiology, University of Pennsylvania, Philadelphia, PA 19104

[6]The Biodesign Center for Fundamental and Applied Microbiomics, Arizona State University, Tempe, Arizona, AZ 85287, USA

[7]Fish and Wildlife Research Institute, Florida Fish and Wildlife Conservation Commission, Naples, FL 34114.

[8]Fish and Wildlife Research Institute, Florida Fish and Wildlife Conservation Commission, Gainesville, FL 32601.

[9]Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, CO 80523.

[10]Department of Biology, Graduate Degree Program in Ecology, Colorado State University, Fort Collins, CO 80523.

[11] Department of Ecology, Evolution and Behavior, University of Minnesota, St Paul, MN 55108.

Abstract

Identifying drivers of transmission prior to an epidemic—especially of an emerging pathogen—is a formidable challenge for proactive disease management efforts. To overcome this gap, we tested a novel approach hypothesizing that an apathogenic virus could elucidate drivers of transmission processes, and thereby predict transmission dynamics of an analogously transmitted virulent pathogen. We evaluated this hypothesis in a model system, the Florida panther (*Puma concolor coryi*), using apathogenic feline immunodeficiency virus (FIV) to predict transmission dynamics for another retrovirus, pathogenic feline leukemia virus (FeLV). We derived a transmission network using FIV genome sequences, and used exponential random graph models to determine drivers structuring this network. We used the identified drivers to predict transmission pathways among panthers; simulated FeLV transmission using these pathways and three alternate modeling approaches; and compared predictions against empirical data collected during a historical FeLV outbreak in panthers. FIV transmission was primarily driven by panther age class and distances between panther home range centroids. Prospective FIV-based predictions of FeLV transmission dynamics performed at least as well as simpler, often retrospective approaches, with evidence that FIV-based predictions could capture the spatial structuring of the observed FeLV outbreak. Our finding that an apathogenic agent can predict transmission of an analogously transmitted pathogen is an innovative approach that warrants testing in other host-pathogen systems to determine generalizability. Use of such apathogenic agents holds promise for improving predictions of pathogen transmission in novel host populations, and can thereby revolutionize proactive pathogen management in human and animal systems.

Significance Statement

Predicting infectious disease transmission dynamics is fraught with assumptions which limit our ability to proactively develop targeted control strategies. We show that transmission of non-disease causing (apathogenic) agents provides invaluable insight into drivers of transmission prior to outbreaks of more serious diseases. Integrating genomic and network approaches, we tested an apathogenic virus as a proxy for predicting transmission dynamics of a deadly virus in the Florida panther. We found that apathogenic virus-based predictions of pathogen transmission dynamics performed at least as well as simpler transmission models, and offered the advantage of prospectively identifying the underlying management-relevant drivers of transmission. Our innovative approach offers an opportunity to proactively design disease control strategies in at-risk animal and human populations.

Keywords

transmission tree; exponential random graph model; network modeling; disease model; Florida panther

Introduction

Infectious disease outbreaks can have profound impacts on conservation, food security, and global health and economics. Mathematical models have proven a vital tool for understanding transmission dynamics of pathogens (1, 2), but struggle to predict the dynamics of novel or emerging agents (3, 4). This is at least partially due to the challenges associated with characterizing contacts relevant to transmission processes. Common modeling approaches that assume all hosts interact and transmit infections to the same degree ignore key drivers of transmission. Such drivers can include specific transmission-relevant behaviors including grooming or fighting in animals (5), concurrent sexual partnerships in humans (6), or assortative mixing (7), and result in flawed epidemic predictions (8, 9). Further, identifying drivers of transmission and consequent control strategies for any given pathogen is typically done reactively or retrospectively in an effort to stop or prevent further outbreaks or spatial spread (e.g., 10–12). These constraints limit the ability to perform prospective disease management planning tailored to a given target population, increasing the risk of potentially catastrophic pathogen outbreaks, as observed in humans (13–15), domestic animals (16, 17), and species of conservation concern including Ethiopian wolves (18), African lions (19), black-footed ferrets (20), and Florida panthers (21).

A handful of studies have evaluated whether common infectious agents present in the healthy animal microbiome or virome can indicate contacts between individuals that may translate to interactions promoting pathogen transmission. Such an approach circumvents some of the uncertainties associated with more traditional approaches to contact detection (8). In these cases, genetic evidence from the transmissible agent itself is used to define between-individual interactions for which contact was sufficient for transmission to occur. Results of such studies show mixed success (22–27). For example, members of the same household have been found to share microbiota (28, 29), but disentangling social mechanisms of this sharing is complicated by shared diets, environments, behaviors, etc. (30). In animals,

studies of *Escherichia coli* in Verreaux's sifaka and giraffe have found strain sharing relationships to be tied to social interactions (22, 23), but the same was not found in a similar study of elephants (26).

These studies have, however, revealed ideal characteristics of non-disease inducing infectious agents (hereafter, *apathogenic agents*) for use as markers of transmission-relevant interactions. Such apathogenic agents should have rapid mutation rates to facilitate discernment of transmission relationships between individuals over time (31, 32). Furthermore, these agents should be relatively common and well-sampled in a target population, have a well-characterized mode of transmission, and feature high strain alpha-diversity (local diversity) and high strain turnover (32, 33). RNA viruses align well with these characteristics (34, 35) such that apathogenic RNA viruses could act as "proxies" of specific modes of transmission (i.e., direct transmission) and indicate which drivers underlie transmission processes. Such drivers, including but not limited to host demographics, relatedness, specific behaviors, or space use, would subsequently allow prediction of transmission dynamics of pathogenic agents with the same mode of transmission (32).

Here, we test the feasibility of this approach using a naturally occurring host-pathogen system to test if an apathogenic RNA virus can act as a proxy for direct transmission processes and subsequently predict transmission of a pathogenic RNA virus. Florida panthers (*Puma concolor coryi*) are an endangered subspecies of puma found only in southern Florida and have been extensively studied and monitored for almost four decades. We have documented that this population is infected by several feline retroviruses relevant to our study questions. Feline immunodeficiency virus (FIVpco; hereafter, FIV) occurs in approximately 50% of the population (36), and does not appear to cause significant clinical disease. FIV is transmitted by close contact (i.e., fighting and biting), generally has a rapid mutation rate (intra-individual evolution rate of 0.00129 substitutions/site/year; 37), and, as a chronic retrovirus infection, can be persistently detected after the time of infection. Panthers are also affected by feline leukemia

virus (FeLV), also a retrovirus, which spills over into their population following exposure of panthers to infected domestic cats (38). Once spillover occurs, FeLV is transmitted between panthers by close contact and results in progressive, regressive, or abortive infection states (21). Progressive cases are infectious and result in mortality; regressive infections are not expected to be infectious—though this is unclear in panthers—and recover (21, 39, 40). Abortive cases clear infection and are not themselves infectious (40). A high mortality FeLV outbreak was documented among panthers in 2002-2004 and has been characterized in several studies (21, 38, 41).

This well-observed historical outbreak is a key advantage of this naturally-occurring system, allowing us to compare our predictions against empirical observations. The objectives of this study were therefore: (1) to determine which factors shape FIV transmission in Florida panthers, and (2) test if these factors can predict transmission dynamics of analogously-transmitted FeLV in panthers. Success of this approach in our model system would pave the way for testing similar apathogenic agents in other host-pathogen systems, thereby improving our ability to predict transmission dynamics of novel agents in human and animal populations.

Methods

Dataset assembly

We assembled an extensive dataset covering almost 40 years of Florida panther research. Ongoing panther management has documented age and sex of monitored panthers. In addition, a subset of the population is monitored using very high frequency (VHF) telemetry collars, with relocations determined via aircraft typically three times per week. Previous panther research has generated a microsatellite dataset for monitored panthers (42), and a dataset of 60 full FIV genomes (proviral DNA sequenced within a tiled amplicon framework in: 43). In addition, the historical FeLV outbreak in panthers was well documented (21, 38), providing key

observations regarding FeLV dynamics in free-ranging panthers. To augment these observations, we also used an FeLV database which documents FeLV status (positive and negative) for 31 panthers as determined by qPCR from 2002-04.

FIV transmission inference

To determine drivers of FIV transmission, we first generated a "who transmitted to whom" transmission network using 60 panther FIV genomes collected from 1988 to 2011 (see Figure 1 for workflow across all analyses). We used the program Phyloscanner (44), which maximizes the information gleaned from next generation sequencing reads to infer transmission relationships. Phyloscanner operates in a two step process, first inferring within- and between-host phylogenies in windows along the FIV genome, and then analysing those phylogenies to produce transmission trees or networks. For step one, we used 150bp windows, allowing 25bp overlap between windows. To test sensitivity to this choice, we separately ran a full Phyloscanner analysis with 150bp windows, but without overlap between windows (supplementary methods). For step two, we held k, which penalizes within-host diversity, equal to 0. We used a patristic distance threshold of 0.05 and allowed missing and more complex transmission relationships. Because we had uneven read depth across FIV genomes, we downsampled to a maximum of 200 reads per host. The output of the full Phyloscanner analysis was a single transmission network (hereafter, *main FIV network*), but see supplementary methods for details regarding analysis of the sensitivity of our results to variations in and summary across multiple transmission networks (resulting in two *summary FIV networks*).

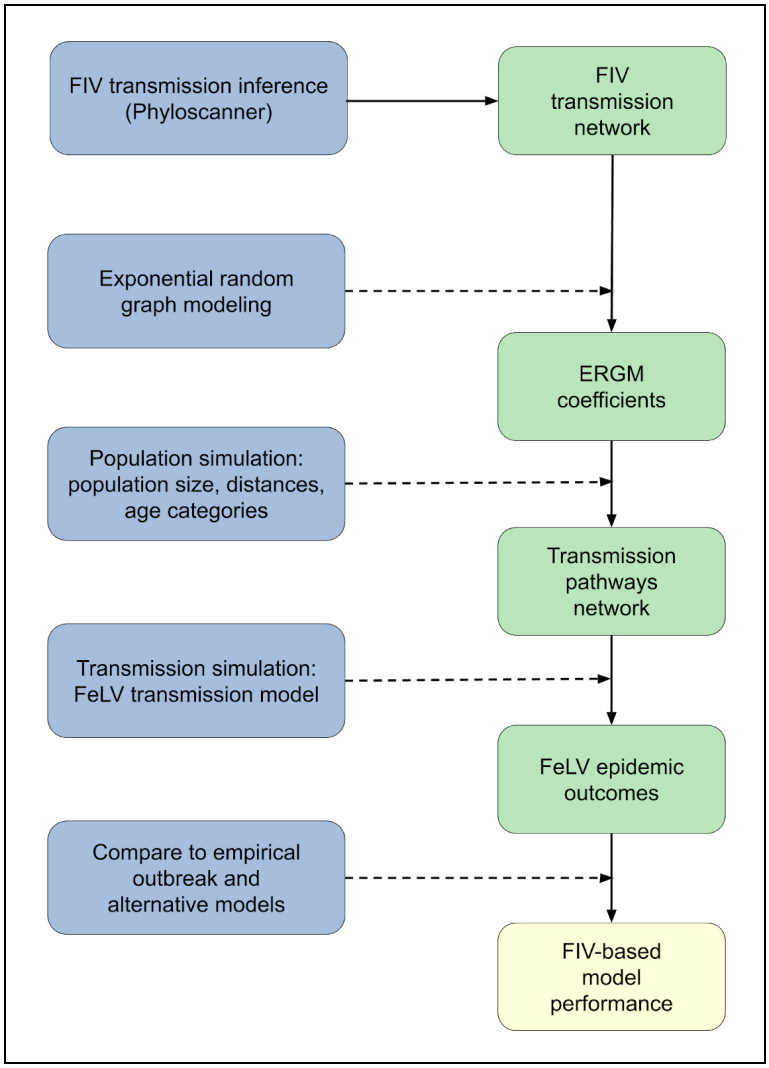

Figure 1: *Conceptual workflow across all analysis steps. Processes are shown on the left in blue; specific outcomes are shown on the right in green; the final analysis outcome is in yellow at the bottom right. Solid lines show direct flows or outcomes. Dashed lines show processes acting on or in concert with prior outcomes: for example, exponential random graph modeling was performed using the FIV transmission network, and the combination of the two produced the ERGM coefficients outcome.*

Statistical analysis of FIV transmission networks

We performed statistical analysis of unweighted, binary FIV transmission networks using exponential random graph models (ERGMs), which account for non-independence in network structure (45). The network structural terms we considered included an intercept-like edges term (45), geometrically weighted edgewise shared partner distribution (*gwesp*; representation of network triangles), alternating k-stars (*altkstar*; representation of star structures), and 2-paths (2 step paths from *i* to *k* via *j*; 46).

The dyad-independent variables included panther sex (both as a node factor and node mixing variable; see supplementary methods for terminology and additional variable details). We assessed panther age as a categorical variable (both as a node factor and node mixing variable), with subadults classified as individuals between the ages of 6 months and two years, and adults classified as individuals over two years of age. We included pairwise genetic relatedness from panther microsatellite data as an edge covariate. Spatial variables included a node-matching variable for the location of panthers' minimum convex polygon (MCP) home range centroid or capture location (hereafter *centroid*; see supplementary methods) north versus south of the major I-75 freeway. In addition, we included a node covariate term for the distance from the centroid to the nearest urban area (in km; USA Urban Areas layer, ArcGIS; 47). Pairwise geographic distances between panthers were calculated using distances between centroids (in km), and log-transformed for use as an edge covariate. Lastly, we included a spatial overlap edge covariate based on the pairwise utilization distribution overlap indices of 95% home range kernels (48), using the adehabitat package in R (49).

Because ERGMs are prone to degeneracy with increasing complexity (45), we first performed forward selection for network structural variables, followed by forward selection of dyad-independent variables, while controlling for network structure. Model selection was based on AIC and goodness of fit, and MCMC diagnostics were assessed for the final model (supplementary methods).

Panther population simulations

To test if predictors of FIV transmission identified in the ERGM analysis can predict FeLV transmission, we next simulated FeLV transmission through networks which were based on these FIV predictors among simulated populations representing panthers during the well-characterized FeLV outbreak (2002-2004; 21). Hereafter, a *full-simulation* includes both simulation of the panther population and its network of likely transmission pathways during the historical outbreak period and simulation of FeLV transmission within that population. Below, we describe the process for a single simulation, but these procedures were repeated for each full simulation.

We first based the simulated population size on the range of empirical estimates from 2002-2004 (50). Additional characteristics of the simulated population included those identified as significant variables in the ERGM analysis, which were: age category and pairwise geographic distances between panther home range centroids. We randomly assigned age categories to the simulated population based on the proportion of adults versus subadults. Age proportions were based on age distributions in the western United States (51) which qualitatively align with the historically elevated mean age of the Florida panther population (52). Pairwise geographic distances for the simulated population were generated by randomly assigning simulated home range centroids based on the distribution of observed centroids on the landscape (supplementary methods).

We then used ERGM coefficients to generate network edges in the simulated panther population representing potential transmission pathways between panthers. The FIV transmission network spanned 15 years of observations and represents a subset of the actual contact network, as it includes only those interactions that resulted in successful transmission (53). We therefore had a high degree of uncertainty regarding the appropriate network density for our simulations. To manage this uncertainty, we constrained density in our network simulations across a range of parameter space (net_dens, Table 1).

Table 1: Network and transmission simulation parameters

Parameter	Definition	Range	Reference
Pop_size	Population size	80-120	(50)
Adult_prop	Proportion adults versus subadults	0.82-0.99	(51)
Net_dens	Simulated network density	0.05-0.15	*NA*
β	Probability of transmission from progressives, given effective contact	0.17-0.29	(54)
C	Constant multiplier for probability of transmission from regressives, given effective contact	0, 0.1, 0.5, 1	*NA*
ω	Weekly probability of contact	0.1-0.4	(55)
μ	Weekly probability of death from progressive infection	1/18, 1/26*	(21)
K	Constant multiplier for weekly probability of recovery from regressive infection	0.5, 1	*NA*
ν	Weekly probability of territory repopulation ("respawn rate")	1/12-1/4	*NA*
τ	Weekly probability of vaccination	0.5-1	*NA*
ve	Probability of vaccine efficacy	0.4-1	(21)
P	Proportion randomly assigned to progressive, regressive	0.25	(21)

*Note: Parameter gives parameter symbols or abbreviations; definition gives the description for each parameter. Range shows the continuous range or discrete values sampled from in simulations, with references giving literature supporting ranges or values. *We tested a lower death rate (prolonged duration of infection) due to the low number of observed panther cases and the generally longer infection duration in domestic cats (39).*

Simulation of FeLV transmission on FIV-based networks

The next step in each full simulation was to model FeLV transmission on the network generated from FIV predictors of transmission. FeLV transmission was based on a chain binomial process on the simulated networks, following a modified SIR compartmental model (Figure 2). Simulations were initiated with one randomly selected infectious individual and

proceeded in weekly time steps. Transmission simulations lasted until no infectious individuals remained or until 2.5 years, whichever came first.

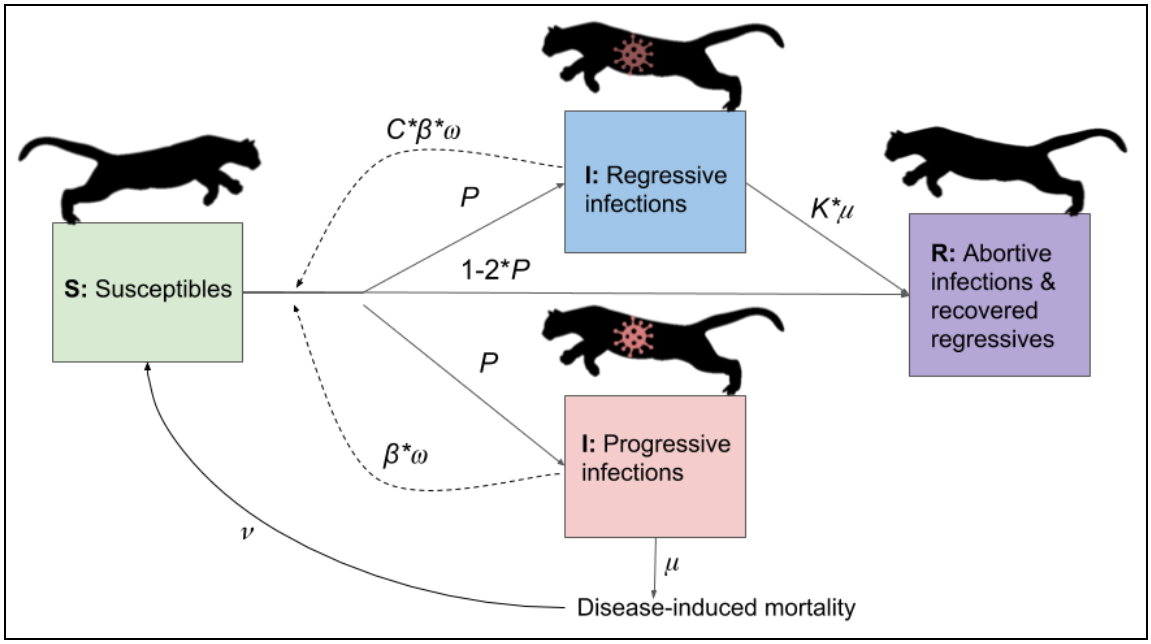

Figure 2: *Diagram of flows of individuals between compartments in transmission model. Virus icons indicate infectious states, with the regressive infection icon darkened to represent reduced or uncertain infectiousness of this class. Note: a vaccination process was also included in the transmission model, but is not shown for simplicity. With vaccination, susceptibles could be vaccinated, and vaccinated individuals subsequently infected analogously to susceptibles, but with an additional probability of (1-*ve*). See Table 1 for definitions of parameters.*

Transmission (Figure 2) was dependent on the following: (1) existence of an edge between two individuals, (2) the dyad in question involving a susceptible and infectious individual, and (3) a random binomial draw based on the probability of transmission given contact (*β*, Table 1). In addition, *Puma concolor* generally have low expected weekly contact rates (55); we therefore included an additional weekly contact probability, represented as a random binomial draw for contact in a given week (ω, Table 1).

Upon successful transmission, infectious individuals were randomly assigned to one of three outcomes of FeLV infection (21). *Progressive* infections (probability *P*, Table 1) are infectious, develop clinical disease, and die due to infection at a mortality rate, μ. *Regressive* infections (also probability *P*) recover at a rate based on a constant, *K*, multiplied by the mortality rate of progressives (Table 1). Anecdotal evidence suggests regressive individuals are not infectious (21), but given ongoing uncertainty, we allowed regressives to be infectious by multiplying the probability of transmission for progressives (β) by a constant, *C* (Table 1). *Abortive* infections (probability 1-2*P*) are never infectious, clearing infection and joining the recovered class.

A vaccination process was included in simulations as panthers were vaccinated against FeLV during the historical FeLV outbreak starting in 2003. Vaccination occurred at a rate, τ, and applied to the whole population as wildlife managers are unlikely to know if a panther is susceptible at the time of capture or darting. However, only susceptible individuals transitioned to the vaccinated class (i.e. vaccination failed in non-susceptibles). Because panthers were vaccinated in the empirical outbreak with a domestic cat vaccine with unknown efficacy in panthers, we allowed vaccinated individuals to become infected in transmission simulations by including a binomial probability for vaccine failure (1-vaccine efficacy, *ve*, Table 1).

The panther population size remained roughly static through the course of the FeLV outbreak (50, 56). We therefore elected not to include background mortality, but did include infection-induced mortality. To maintain a consistent population size, we therefore included a birth/recruitment process. Because FIV-based simulated networks drew edges based on population characteristics, we treated births as a “respawning” process, in which territories vacated due to mortality were reoccupied by a new susceptible at rate, ν. This approach allowed us to maintain the ERGM-based network structure and is biologically reasonable, as vacated panther territories are unlikely to remain unoccupied for long.

FeLV infection is uncommon in puma as a species, resulting in a high degree of uncertainty regarding differences in within-individual infection dynamics in panthers (e.g., epidemiological parameters such as infectious period). Given these uncertainties, we performed all simulations across a range of parameter space. To more efficiently cover this parameter space, we generated parameter sets using a Latin hypercube design (LHS), using the lhs package in R (57). We generated 150 parameter sets, conducting 50 full simulations per parameter set.

Comparison of simulation predictions to observed FeLV outbreak

We compared FIV-based simulation predictions to the observed FeLV outbreak as well as predictions from three simpler types of models: random networks, home range overlap-based networks, and a well-mixed model. All models used the parameterizations from our LHS parameter sets, as relevant.

For our random networks model, we generated Erdős-Rényi random networks, with the simulated network densities from our LHS parameter sets (Table 1). Overlap-based networks were generated using the degree distributions of panther home range overlap networks from 2002-2004 and simulated annealing with the R package statnet (58, 59; supplementary methods). These overlap-based networks were not spatially explicit, as they were based only on the degree distributions from real spatial overlap networks. For both random and overlap-based networks, FeLV transmission was simulated as in the FIV-based simulations. The well-mixed model was a Gillespie algorithm (stochastic, continuous time compartmental model), with rate functions aligning with the chain binomial FeLV transmission probabilities (supplementary methods).

Target ranges for predicted outcomes were based on observed FeLV dynamics (21), with ranges to account for uncertainty in observations and population size in this cryptic species (supplementary methods). The primary outcomes of interest were (1) duration of outbreak:

78-117 weeks, (2) total number of progressive infections: 5-20, and (3) presence of spatial clustering (see below). While our primary focus was progressive infections, we also included an expectation that at least 5 individuals were abortive infections. Empirically, these individuals were the most numerous, but as they were not clinically ill, abortive infections were less likely to be detected in normal panther management; we therefore did not include an upper bound for this target.

Using our database of qPCR results for FeLV in panthers (positive and negative tests), we performed a local spatial clustering analysis of FeLV cases and controls using SaTScan (50% maximum, circular window; 60), and a global cluster analysis with Cuzick and Edward's test in the R package *smacpod* (1, 3, 5, 7, 9, and 11 nearest neighbors; 999 iterations; 61, 62). These analyses found evidence of local (weak) and global clustering (at 3, 5, and 7 neighborhood levels) among progressive and regressive cases (see results, supplementary methods and results). In simulations, we therefore included spatial clustering of progressive and regressive cases as a target outcome.

For the duration of outbreaks and total number of progressives, we calculated median values of both outcomes for each parameter set (i.e., 50 simulations) within each model type. If a parameter set's medians were within the target ranges for both of these outcomes, it was considered *feasible*. To quantify differences in model prediction performance, we fit a binomial generalized linear mixed model (GLMM), assuming a logistic regression with "feasible" (vs "unfeasible") as the outcome, model type as a predictor variable, and a random intercept for LHS parameter set.

To determine if simulated results demonstrated spatial clustering, we performed SaTScan spatial cluster analysis (50% maximum, circular window) and Cuzick and Edward's tests (at 3, 5, and 7 nearest neighbors) on simulation results. Only FIV-based simulations were spatially explicit, and we performed spatial analyses only on those parameter sets that were classified as feasible. To determine if any detected clustering in FIV simulations was simply

based on our respawning protocol, we also performed both spatial analyses with feasible overlap-based simulation results. For these, we assigned the same geographic locations to nodes in the overlap-based networks from the corresponding FIV-based networks (i.e., matching simulation number from matching parameter set).

Results

FIV transmission network analysis

The main FIV network included 19 nodes (individuals) with 42 edges (representing potential transmission events; network density = 0.25) after removing 9 edges that were between individuals known not to be alive at the same time (Figure S1). ERGM results for the main FIV network identified geometrically weighted edgewise shared partner distribution (*gwesp*) and alternating k-stars (*altkstar*) as key structural variables, and age category (as a node-level factor) and log transformed pairwise geographic distance as key dyad-independent variables (Table 2). Though altkstar was not statistically significant, inclusion of this variable contributed to improved AIC and goodness of fit outcomes. Adults were more likely to be involved in transmission events (but see discussion of sample size limitations) and inferred transmission events were more likely between individuals which were geographically closer to each other. The fitted model showed reasonable goodness of fit (Figure S2). ERGM results were largely consistent across two replicate analyses with alternative transmission networks formed by summarizing across four single Phyloscanner outputs (see supplementary results for further details).

Table 2: Main FIV transmission network exponential random graph model results

Variable	**Estimate**	**SE**	***p*-value**
Edges (intercept)	-2.56	1.33	0.055

gwesp	0.98	0.26	<0.001
altkstar	-0.70	0.96	0.47
Age (Adult)	0.93	0.44	0.03
Log pairwise distance	-0.45	0.21	0.03

Note: "gwesp" is geometrically weighted edgewise shared partner distribution and "altkstar" is alternating k-stars. Age classes were subadult and adult; pairwise distances were between home range centroids. Only those variables from the final model are shown. Estimates shown are untransformed; SE represents standard error; p*-values less than 0.05 were considered statistically significant.*

FeLV simulations

SaTScan analysis of observed FeLV status found weak evidence of local spatial clustering (two clusters detected, but not statistically significant with p=0.165 and 0.997, respectively; Figure S4). Cuzick and Edward's tests found evidence of global clustering at 3, 5, and 7 nearest neighbor levels (test statistic T_k where k is number of nearest neighbors considered: T_3 = 20, p = 0.049; T_5 = 32, p = 0.028; T_7 = 43, p = 0.023). Both sets of spatial analysis results were then compared against FeLV predictions from FIV and overlap-based models.

About 9% of parameter sets across all model types were classified as feasible (Figures 3, S5-S6). The GLMM for model type performance (i.e., FIV-based, random, overlap-based, or well-mixed) did not find statistically significant differences between odds of generating feasible simulation outcomes, though the FIV-based model had the highest odds of feasibility (exponentiated estimate = 1.55, though p = 0.30; Table S2). Feasible parameter sets from both the FIV-based and overlap-based models produced some evidence of local and global spatial clustering of simulated FeLV cases (Figures 4, S7). However, the FIV-based model was better

able to capture the size and strength (observed/expected FeLV cases) of predicted local clusters (Figure 4) and was moderately better at capturing global spatial patterns (Figure S7).

In order to determine if certain transmission parameters were particularly important for feasible performance, we performed *post hoc* random forest analyses using the R package *randomForest* (63, 64) for each of the four model types (see supplementary results). While random forests typically showed poor balanced accuracy and area under the curve (AUC) results, the parameter shaping transmission from regressively infected individuals (C), showed support for weak to moderate transmission from regressives (i.e., C = 0.1 or 0.5; Figure S10).

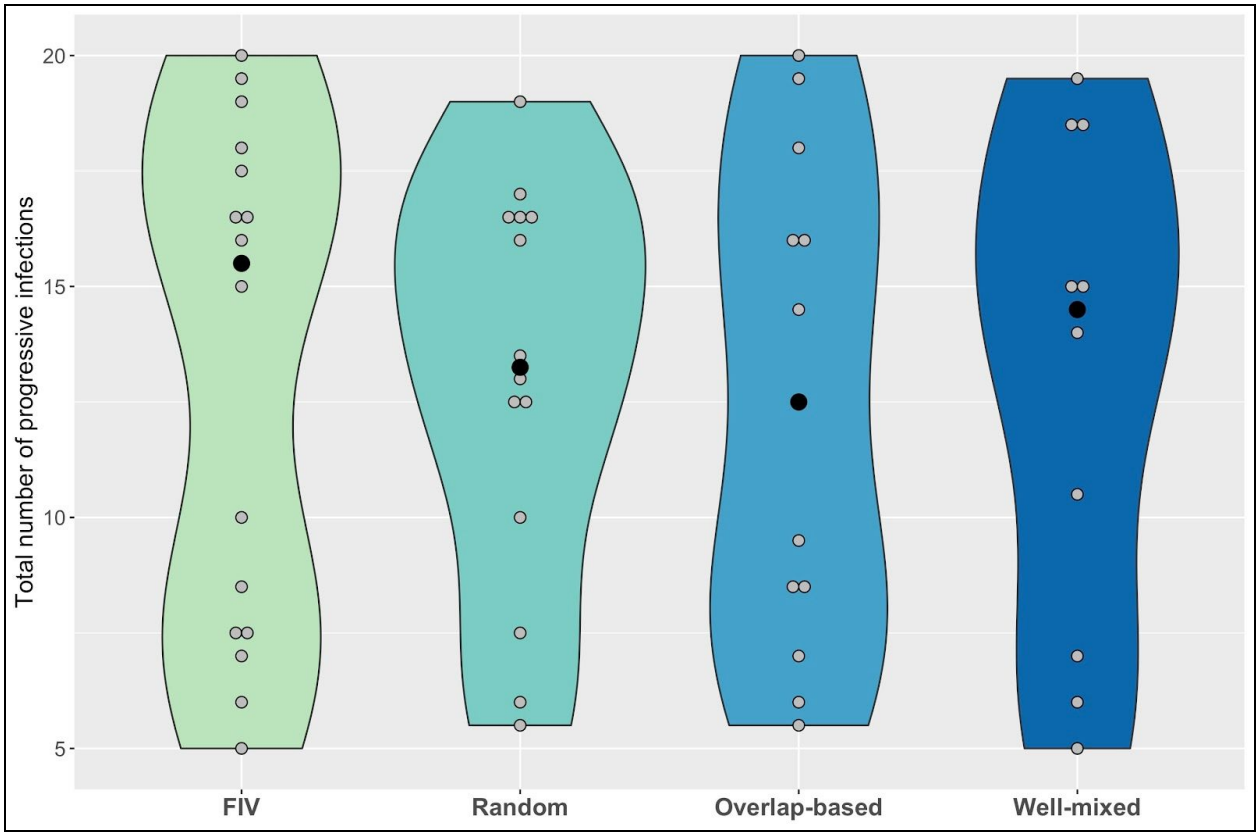

Figure 3: *FIV-based networks perform at least as well as other models in predicting number of progressive infections, as seen in violin plots of median total number of progressive infections from parameter sets classified as "feasible." To be feasible, medians needed to fall between 5 and 20 progressive infections, while also having median epidemic duration between 78-117*

weeks, and a median of at least 5 abortive infections. Gray points show median values from each feasible parameter set; black points are the median value within each violin plot. Model types are given on the x-axis.

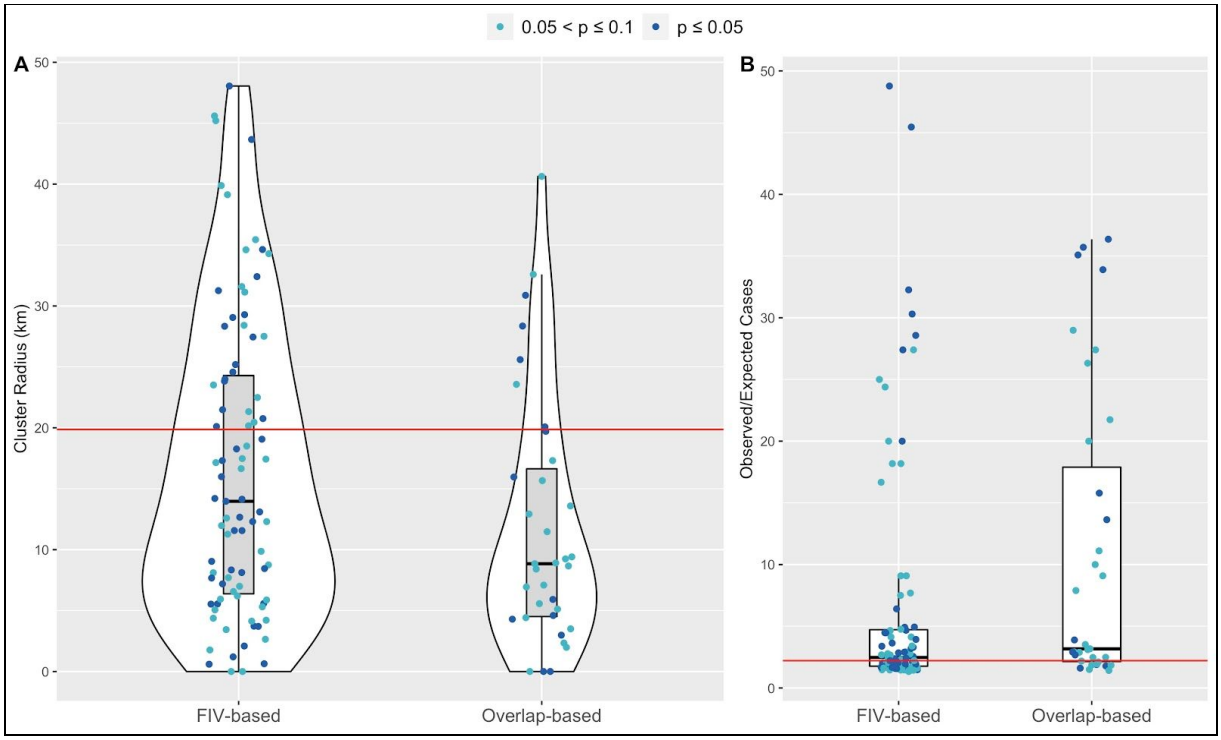

Figure 4: *SaTScan cluster analysis for feasible FIV-based and overlap-based network simulations show stronger agreement between empirical observations (red horizontal lines) and FIV-based predictions for (A) predicted FeLV cluster size and (B) Observed/Expected FeLV cases associated with the top detected cluster. Shown are feasible simulation results in which at least one cluster was detected with* p*-value less than or equal to 0.1; further, only the results from the top cluster are shown.*

Discussion

In this study we develop a new approach whereby we leveraged knowledge of transmission dynamics of a common apathogenic agent to prospectively predict dynamics of an uncommon and virulent pathogen. Our approach was distinctly different from simpler models we tested, as the apathogenic (FIV)-based approach could be used to prospectively identify predictors of transmission and develop disease control plans prior to an outbreak of a virulent pathogen (FeLV), while other approaches either make broad assumptions about transmission-relevant contacts (e.g., homogeneous mixing), or rely on retrospective or reactive modeling. We found that FIV transmission in panthers is primarily driven by distance between home range centroids and age class, and that our prospective FIV-based approach predicted FeLV transmission dynamics at least as well as simpler or more reactive approaches. While we do not propose that this apathogenic agent approach could accurately predict exactly when, where, and to whom transmission might occur, our results support the role of apathogenic agents as novel tools for prospectively identifying relevant drivers of transmission and consequently improving proactive disease management.

Pairwise geographic distances and panther age class predict FIV transmission

Combining genomic and network approaches, we determined that pairwise geographic distances and age category structure FIV transmission in the Florida panther. Because FIV is a persistent infection, we would expect cumulative risk of transmission to increase over an individual's lifetime and adults would consequently be involved in more transmission events. The low number of subadult individuals in our dataset, however, means that this finding must be interpreted with caution.

Panthers are wide-ranging animals but maintain home ranges, and this appears to translate to increased transmission between individuals that are close geographically. This finding is further supported by the tendency for FIV phylogenies to show distinct geographic clustering (65, 66), but is in contrast to other infectious agents of puma. An additional feline

retrovirus, feline foamy virus (FFV), does not show distinct geographic clustering but is commonly transmitted between domestic cats and puma (67). A prior study of several pathogens in puma across the United States rarely identified spatial autocorrelation in pathogen exposures, but notably found that FIV infection status approached statistical significance specifically in Florida panthers (68). The wide-ranging nature of puma appears to limit geographic clustering of many infectious agents, with FIV a notable exception to this pattern. Multi-host agents such as FFV are presumably more able to escape geographical limitations and/or may be transmitted prior to dispersal and thereby lack spatial structuring. More generally, the importance of distance rather than panther relatedness in structuring transmission may support the resource dispersion or land tenure hypotheses as drivers of spatial and social structuring in panthers, rather than kinship. The high inbreeding among panthers (52) may limit our power for identifying a relationship between relatedness and FIV transmission, but support for resource dispersion or land tenure would be in agreement with findings in other puma systems (69) and even other territorial carnivores (70).

Surprisingly, sex was not a significant predictor of FIV transmission. FIV force of infection is generally higher in male panthers, likely due to their increased fighting behaviors (71). However, studies in other felid species have found mixed importance of sex for FIV transmission, ranging from little to no importance (puma in the western United States: 72; bobcat: 73) to importance only among certain FIV subtypes (African lions: 74). Our results add to this body of research to suggest that the relationship between host sex and FIV transmission is more complex than can be explained by sex-specific behaviors or susceptibility alone.

An FIV-based model captures FeLV transmission dynamics

In our study, a network model based on principles of FIV transmission produced FeLV outbreak predictions consistent with the observed FeLV outbreak. The FIV-based approach performed at least as well as simpler models, per our GLMM analysis, with evidence that FIV

better predicted the observed spatial dynamics for FeLV transmission. A key difference between the FIV-based approach and other spatially-explicit methods is that FIV allowed us to determine the importance of spatial dynamics prospectively and then translate to predictions of FeLV transmission, rather than relying on retrospective FeLV spatial analyses. Furthermore, while more complex potential drivers of transmission (e.g., host relatedness or assortative mixing by age or sex) were not found to be important for FIV transmission, these may yet be key factors structuring transmission in other systems. Simpler model types like random networks or metapopulation models may struggle to make transmission predictions that incorporate these factors as drivers of transmission-relevant contact. The predictive power we observed here using an apathogenic virus could thus significantly shape proactive epidemic management strategies for pathogens such as FeLV.

FIV and FeLV have different epidemiologies (e.g., progressive versus regressive infections, duration of infectiousness); despite this, FIV-based predictors of transmission were able to capture dynamics of FeLV transmission. Here, FIV determined the key drivers of close, direct contact transmission in panthers, fundamentally acting as a proxy for this type of contact. FeLV simulations were then able to independently account for differences in epidemiology to produce appropriate predictions for a different but analogously transmitted virus. Key components of this success are likely that (1) FIV is a largely species-specific virus with transmission pathways closely matching intra-species transmission of FeLV, and (2) both FIV and FeLV, perhaps unusually for infectious agents of puma, display spatial clustering of infection. If, for example, FIV also exhibited strong vertical or environmental transmission, we would no longer expect the predictive success for FeLV we observed here.

Domestic cats are the reservoir host for FeLV (38) and panthers are subject to repeated spillover events from domestic cats (41). Determining the predictors and frequency of these spillover events would be less feasible with FIVpco, which adapted to the puma host (75–77). Rather, targeted investigation of spillover dynamics could draw on studies of other apathogenic

viral candidates that are frequently transmitted from domestic cats to puma, such as FFV (67). Using such agents to identify drivers of spillover events could be key for better understanding the dynamics of "pathogen release from reservoir hosts" (4), which is of profound relevance across wildlife, domestic animal, and human systems.

While few parameter sets were classified as feasible across all model types, this appears to be predominantly the result of the wide range of parameter space explored through our LHS sampling design. This limitation was fundamentally due to uncertainties in FeLV transmission parameters, and is representative of the uncertainties experienced in predicting transmission of emerging or understudied pathogens. Key here were the interacting uncertainties regarding infectiousness of regressives, the number of introductions of FeLV to the panther population, and the duration of FeLV infection in panthers. All three of these dynamics can have significant impacts on the duration of a simulated epidemic, allowing an epidemic to continue to "stutter" along at low levels (78), much as was observed in the empirical FeLV outbreak. Our *post hoc* random forest analysis provided some evidence of weak transmission from regressive individuals, but this finding would need to be validated with additional research, as it is in stark contrast to FeLV dynamics in domestic cats. Reducing uncertainties in these three key dynamics would significantly narrow the range of our predictions, and even assist in ongoing management efforts for FeLV in endangered panthers.

Furthermore, the effect of transmission parameter uncertainties underscores the importance of linking laboratory and model-based research to generate more accurate transmission forecasts (79). Experimental research could help to narrow the range of parameter space for FeLV—or other emerging pathogens—to produce more consistent and accurate model predictions. This necessity is all the more apparent during the current COVID-19 pandemic, in which mathematical models have benefited from rapid laboratory and epidemiological research to reduce uncertainty in model parameters.

Limitations and future directions

This study found evidence for the utility of an apathogenic agent to predict transmission of a related pathogenic agent, but this approach must now be tested in additional host-pathogen systems. The mixed results when using commensal agents to identify close social relationships in other systems (22–27) highlights that some host-apathogenic agent combinations will work better than others for determining drivers of transmission. More research is therefore necessary to determine which apathogenic agents may be most suitable as markers of transmission, and how divergent an apathogenic agent may be from a pathogen of interest while still predicting transmission dynamics.

The suite of tools for inferring transmission networks from infectious agent genomes is rapidly expanding (31, 80). In this study, we used the program Phyloscanner as it maximized the information from our next generation sequencing viral data. However, our FIV sequences were generated within a tiled amplicon framework (43, 81), which biases intrahost diversity and likely limits viral haplotypes (82). Phyloscanner was originally designed to analyze RNA from virions and not proviral DNA, as we have done here. We have attempted to mitigate the effects of these limitations by analyzing several different Phyloscanner outputs to confirm consistency in our results, and by using only binary networks to avoid putting undue emphasis on transmission network edge probabilities, as these are likely highly uncertain. Further, our primary conclusions from the transmission networks—that age and pairwise distance are important for transmission—are biologically plausible and supported by other literature, as discussed above. Nevertheless, future work should evaluate additional or alternative transmission network inference platforms.

In addition, ERGMs assume the presence of the “full network” and it is as yet unclear how missing data may affect transmission inferences (45). ERGMs are also prone to degeneracy with increased complexity and do not easily capture uncertainty in transmission events, as most weighted network ERGM (or generalized ERGM) approaches have been

tailored for count data (e.g., 83). ERGMs may therefore not be the ideal solution for identifying drivers of transmission networks in all systems. Alternatives may include advancing dyad-based modeling strategies (84), which may more easily manage weighted networks and instances of missing data.

Our FIV-based approach required extensive field sampling, and many disciplines from viral genomics through simulation modeling. However, with increasing availability of virome data and even field-based sequencing technology, our approach may become more accessible with time. Further, the predictive benefits seen here, while needing further testing and validation, could become a key strategy for proactive pathogen management in species of conservation concern, populations of high economic value (e.g. production animals), or populations at high risk of spillover, all of which may most benefit from rapid, efficient epidemic responses.

Conclusions

Here, we integrated genomic and network approaches to identify drivers of FIV transmission in the Florida panther. This apathogenic agent acted as a marker of close, direct contact transmission, and was subsequently successful in predicting the observed transmission dynamics of the related pathogen, FeLV. Further testing of apathogenic agents as markers of transmission and their ability to predict transmission of related pathogenic agents is needed, but holds great promise for revolutionizing proactive epidemic management across host-pathogen systems.

Acknowledgements

Thanks to M. Michalska-Smith, K. Worsley-Tonks, J. Mistrick, and S. N. Hart for key feedback. This research was supported by the National Science Foundation (DEB-1413925, 1654609, and 2030509). MLJG was supported by the Office of the Director, National Institutes of Health (NIH T32OD010993), the University of Minnesota Informatics Institute MnDRIVE program, and the

Van Sloun Foundation. JLM was supported by the ACVP/STP Coalition for Veterinary Pathology Fellows and the Linda Munson Fellowship for Wildlife Pathology Research. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Puma icon in Figure 2 by Freepik at Flaticon.com.

Data Availability

Full R code for simulations is available on GitHub (https://github.com/mjones029/FIV_FeLV_Transmission) and upon acceptance will be archived at Zenodo.

Works Cited

1. R. M. Anderson, R. M. May, *Infectious Diseases of Humans: Dynamics and Control* (OUP Oxford, 1991).

2. J. Antonovics, Transmission dynamics: critical questions and challenges. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372** (2017).

3. C. J. E. Metcalf, J. Lessler, Opportunities and challenges in modeling emerging infectious diseases. *Science* **357**, 149–152 (2017).

4. R. K. Plowright, *et al.*, Pathways to zoonotic spillover. *Nat. Rev. Microbiol.* **15**, 502–510 (2017).

5. J. A. Drewe, Who infects whom? Social networks and tuberculosis transmission in wild meerkats. *Proc. Biol. Sci.* **277**, 633–642 (2010).

6. M. Morris, M. Kretzschmar, Concurrent partnerships and transmission dynamics in networks. *Soc. Networks* **17**, 299–318 (1995).

7. S. Cauchemez, *et al.*, Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 2825–2830 (2011).

8. M. E. Craft, D. Caillaud, Network models: an underutilized tool in wildlife epidemiology? *Interdiscip. Perspect. Infect. Dis.* **2011**, 676949 (2011).

9. M. J. Keeling, K. T. D. Eames, Networks and epidemic models. *J. R. Soc. Interface* **2**,

295–307 (2005).

10. M. J. Keeling, M. E. J. Woolhouse, R. M. May, G. Davies, B. T. Grenfell, Modelling vaccination strategies against foot-and-mouth disease. *Nature* **421**, 136–142 (2003).

11. D. L. Smith, B. Lucey, L. A. Waller, J. E. Childs, L. A. Real, Predicting the spatial dynamics of rabies epidemics on heterogeneous landscapes. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 3668–3672 (2002).

12. J. O. Lloyd-Smith, *et al.*, Epidemic dynamics at the human-animal interface. *Science* **326**, 1362–1367 (2009).

13. A. P. Dobson, *et al.*, Ecology and economics for pandemic prevention. *Science* **369**, 379–381 (2020).

14. J. Pike, T. Bogich, S. Elwood, D. C. Finnoff, P. Daszak, Economic optimization of a global strategy to address the pandemic threat. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 18519–18523 (2014).

15. M. R. Keogh-Brown, R. D. Smith, The economic impact of SARS: How does the reality match the predictions? *Health Policy* **88**, 110–120 (2008).

16. T. J. D. Knight-Jones, J. Rushton, The economic impacts of foot and mouth disease--What are they, how big are they and where do they occur? *Prev. Vet. Med.* **112**, 161–173 (2013).

17. A. Blake, M. T. Sinclair, G. Sugiyarto, Quantifying the Impact of Foot and Mouth Disease on Tourism and the UK Economy. *Tourism Econ.* **9**, 449–465 (2003).

18. C. Sillero-Zubiri, A. A. King, D. W. Macdonald, Rabies and mortality in Ethiopian wolves (*Canis simensis*). *J. Wildl. Dis.* **32**, 80–86 (1996).

19. M. E. Roelke-Parker, *et al.*, A canine distemper virus epidemic in Serengeti lions (*Panthera leo*). *Nature* **379**, 441–445 (1996).

20. E. S. Williams, E. T. Thorne, M. J. Appel, D. W. Belitsky, Canine distemper in black-footed ferrets (*Mustela nigripes*) from Wyoming. *J. Wildl. Dis.* **24**, 385–398 (1988).

21. M. W. Cunningham, *et al.*, Epizootiology and management of feline leukemia virus in the Florida puma. *J. Wildl. Dis.* **44**, 537–552 (2008).

22. K. L. VanderWaal, E. R. Atwill, L. A. Isbell, B. McCowan, Linking social and pathogen transmission networks using microbial genetics in giraffe (*Giraffa camelopardalis*). *J. Anim. Ecol.* **83**, 406–414 (2014).

23. A. Springer, A. Mellmann, C. Fichtel, P. M. Kappeler, Social structure and *Escherichia coli* sharing in a group-living wild primate, Verreaux's sifaka. *BMC Ecol.* **16**, 6 (2016).

24. C. M. Bull, S. S. Godfrey, D. M. Gordon, Social networks and the spread of Salmonella in a sleepy lizard population. *Mol. Ecol.* **21**, 4386–4392 (2012).

25. A. Blasse, *et al.*, Mother-offspring transmission and age-dependent accumulation of simian foamy virus in wild chimpanzees. *J. Virol.* **87**, 5193–5204 (2013).

26. P. I. Chiyo, *et al.*, The influence of social structure, habitat, and host traits on the transmission of *Escherichia coli* in wild elephants. *PLoS One* **9**, e93408 (2014).

27. M. D. J. Blyton, S. C. Banks, R. Peakall, D. M. Gordon, High temporal variability in commensal *Escherichia coli* strain communities of a herbivorous marsupial. *Environ. Microbiol.* **15**, 2162–2172 (2013).

28. S. Lax, *et al.*, Longitudinal analysis of microbial interaction between humans and the indoor environment. *Science* **345**, 1048–1052 (2014).

29. S. J. Song, *et al.*, Cohabiting family members share microbiota with one another and with their dogs. *Elife* **2**, e00458 (2013).

30. E. A. Archie, J. Tung, Social behavior and the microbiome. *Current Opinion in Behavioral Sciences* **6**, 28–34 (2015).

31. M. D. Hall, M. E. J. Woolhouse, A. Rambaut, Using genomics data to reconstruct transmission trees during disease outbreaks. *Rev. Sci. Tech.* **35**, 287–296 (2016).

32. M. L. J. Gilbertson, N. M. Fountain-Jones, M. E. Craft, Incorporating genomic methods into contact networks to reveal new insights into animal behaviour and infectious disease dynamics. *Behaviour* (2018).

33. M. D. J. Blyton, S. C. Banks, R. Peakall, D. B. Lindenmayer, D. M. Gordon, Not all types of host contacts are equal when it comes to *E. coli* transmission. *Ecol. Lett.* **17**, 970–978 (2014).

34. B. T. Grenfell, *et al.*, Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* **303**, 327–332 (2004).

35. E. A. Archie, G. Luikart, V. O. Ezenwa, Infecting epidemiology with genetics: a new frontier in disease ecology. *Trends Ecol. Evol.* **24**, 21–30 (2009).

36. S. Carver, *et al.*, Pathogen exposure varies widely among sympatric populations of wild and domestic felids across the United States. *Ecol. Appl.* **26**, 367–381 (2016).

37. E. Krakoff, R. B. Gagne, S. VandeWoude, S. Carver, Variation in Intra-individual Lentiviral Evolution Rates: a Systematic Review of Human, Nonhuman Primate, and Felid Species. *J. Virol.* **93** (2019).

38. M. A. Brown, *et al.*, Genetic characterization of feline leukemia virus from Florida panthers. *Emerg. Infect. Dis.* **14**, 252–259 (2008).

39. K. Hartmann, Clinical aspects of feline retroviruses: a review. *Viruses* **4**, 2684–2710 (2012).

40. C. E. Greene, *Infectious diseases of the dog and cat*, 4th ed.. (Elsevier/Saunders, 2012).

41. E. S. Chiu, *et al.*, Multiple Introductions of Domestic Cat Feline Leukemia Virus in Endangered Florida Panthers. *Emerg. Infect. Dis.* **25**, 92–101 (2019).

42. M. van de Kerk, D. P. Onorato, J. A. Hostetler, B. M. Bolker, M. K. Oli, Dynamics, persistence, and genetic management of the endangered florida panther population. *Wildlife Monogr.* **203**, 3–35 (2019).

43. J. L. Malmberg, *et al.*, Altered lentiviral infection dynamics follow genetic rescue of the Florida panther. *Proc. Biol. Sci.* **286**, 20191689 (2019).

44. C. Wymant, *et al.*, PHYLOSCANNER: Inferring Transmission from Within- and Between-Host Pathogen Genetic Diversity. *Mol. Biol. Evol.* **35**, 719–733 (2018).

45. M. J. Silk, D. N. Fisher, Understanding animal social structure: exponential random graph models in animal behaviour research. *Anim. Behav.* **132**, 137–146 (2017).

46. M. Morris, M. S. Handcock, D. R. Hunter, Specification of Exponential-Family Random Graph Models: Terms and Computational Aspects. *J. Stat. Softw.* **24**, 1548–7660 (2008).

47. Esri, National Atlas of the United States, United States Geological Survey, Department of Commerce, Census Bureau-Geography Division, USA Urban Areas (FeatureServer).

48. J. Fieberg, C. O. Kochanny, Lanham, Quantifying home-range overlap: the importance of the utilization distribution. *J. Wildl. Manage.* **69**, 1346–1359 (2005).

49. C. Calenge, The package adehabitat for the R software: tool for the analysis of space and habitat use by animals. *Ecological Modelling* **197**, 1035 (2006).

50. B. T. McClintock, D. P. Onorato, J. Martin, Endangered Florida panther population size determined from public reports of motor vehicle collision mortalities. *J. Appl. Ecol.* **52**, 893–901 (2015).

51. K. A. Logan, L. L. Sweanor, *Desert Puma: Evolutionary Ecology And Conservation Of An Enduring Carnivore* (Island Press, 2001).

52. W. E. Johnson, *et al.*, Genetic restoration of the Florida panther. *Science* **329**, 1641–1645 (2010).

53. M. E. Craft, Infectious disease transmission and contact networks in wildlife and livestock. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **370** (2015).

54. E. Fromont, M. Artois, M. Langlais, F. Courchamp, D. Pontier, Modelling the feline leukemia virus (FeLV) in natural populations of cats (*Felis catus*). *Theor. Popul. Biol.* **52**, 60–70 (1997).

55. L. M. Elbroch, H. Quigley, Social interactions in a solitary carnivore. *Curr. Zool.* **63**, 357–362 (2016).

56. R. T. McBride, R. T. McBride, R. M. McBride, C. E. McBride, Counting Pumas by Categorizing Physical Evidence. *sena* **7**, 381–400 (2008).

57. R. Carnell, lhs: Latin hypercube samples. *R package version 0.10, URL http://CRAN. R-project.org/package=lhs* (2012).

58. M. S. Handcock, D. R. Hunter, C. T. Butts, S. M. Goodreau, M. Morris, statnet: Software tools for the representation, visualization, analysis and simulation of network data. *J. Stat. Softw.* **24**, 1548 (2008).

59. J. J. H. Reynolds, B. T. Hirsch, S. D. Gehrt, M. E. Craft, Raccoon contact networks predict seasonal susceptibility to rabies outbreaks and limitations of vaccination. *J. Anim. Ecol.* **84**,

1720–1731 (2015).

60. M. Kulldorff, A spatial scan statistic. *Communications in Statistics - Theory and Methods* **26**, 1481–1496 (1997).

61. J. French, smacpod: Statistical Methods for the Analysis of Case-Control Point Data (2020).

62. J. Cuzick, R. Edwards, Spatial clustering for inhomogeneous populations. *J. R. Stat. Soc.* **52**, 73–96 (1990).

63. A. Liaw, M. Wiener, Classification and Regression by randomForest. *R News* **2**, 18–22 (2002).

64. L. A. White, S. VandeWoude, M. E. Craft, A mechanistic, stigmergy model of territory formation in solitary animals: Territorial behavior can dampen disease prevalence but increase persistence. *PLoS Comput. Biol.* **16**, e1007457 (2020).

65. J. S. Lee, *et al.*, Evolution of puma lentivirus in bobcats (*Lynx rufus*) and mountain lions (*Puma concolor*) in North America. *J. Virol.* **88**, 7727–7737 (2014).

66. S. P. Franklin, *et al.*, Frequent transmission of immunodeficiency viruses among bobcats and pumas. *J. Virol.* **81**, 10961–10969 (2007).

67. S. Kraberger, *et al.*, Frequent cross-species transmissions of foamy virus between domestic and wild felids. *Virus Evol* **6**, vez058 (2020).

68. M. L. J. Gilbertson, S. Carver, S. VandeWoude, Is pathogen exposure spatially autocorrelated? Patterns of pathogens in puma (*Puma concolor*) and bobcat (*Lynx rufus*) (2016).

69. L. M. Elbroch, P. E. Lendrum, H. Quigley, A. Caragiulo, Spatial overlap in a solitary carnivore: support for the land tenure, kinship or resource dispersion hypotheses? *J. Anim. Ecol.* **85**, 487–496 (2016).

70. E. E. Brandell, *et al.*, Group density, disease, and season shape territory size and overlap of social carnivores. *J. Anim. Ecol.* (2020) https:/doi.org/10.1111/1365-2656.13294.

71. J. J. H. Reynolds, *et al.*, Feline immunodeficiency virus in puma: Estimation of force of infection reveals insights into transmission. *Ecol. Evol.* **9**, 11010–11024 (2019).

72. N. M. Fountain-Jones, *et al.*, Host relatedness and landscape connectivity shape pathogen spread in a large secretive carnivore. *bioRxiv*, 816009 (2019).

73. N. M. Fountain-Jones, *et al.*, Urban landscapes can change virus gene flow and evolution in a fragmentation-sensitive carnivore. *Mol. Ecol.* **26**, 6487–6498 (2017).

74. N. M. Fountain-Jones, *et al.*, Linking social and spatial networks to viral community phylogenetics reveals subtype-specific transmission dynamics in African lions. *J. Anim. Ecol.* **86**, 1469–1482 (2017).

75. D. M. Lagana, *et al.*, Characterization of regionally associated feline immunodeficiency virus (FIV) in bobcats (*Lynx rufus*). *J. Wildl. Dis.* **49**, 718–722 (2013).

76. S. VandeWoude, C. Apetrei, Going wild: lessons from naturally occurring T-lymphotropic lentiviruses. *Clin. Microbiol. Rev.* **19**, 728–762 (2006).

77. S. VandeWoude, J. Troyer, M. Poss, Restrictions to cross-species transmission of lentiviral infection gleaned from studies of FIV. *Vet. Immunol. Immunopathol.* **134**, 25–32 (2010).

78. S. Blumberg, J. O. Lloyd-Smith, Inference of R(0) and transmission heterogeneity from the size distribution of stuttering chains. *PLoS Comput. Biol.* **9**, e1002993 (2013).

79. R. K. Plowright, S. H. Sokolow, M. E. Gorman, P. Daszak, J. E. Foley, Causal inference in disease ecology: investigating ecological drivers of disease emergence. *Front. Ecol. Environ.* **6**, 420–429 (2008).

80. S. M. Firestone, *et al.*, Reconstructing foot-and-mouth disease outbreaks: a methods comparison of transmission network models. *Sci. Rep.* **9**, 4809 (2019).

81. J. Quick, *et al.*, Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat. Protoc.* **12**, 1261–1276 (2017).

82. N. D. Grubaugh, *et al.*, An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biol.* **20**, 8 (2019).

83. P. N. Krivitsky, Exponential-family random graph models for valued networks. *Electron. J. Stat.* **6**, 1100–1128 (2012).

84. A. Yang, *et al.*, Effects of social structure and management on risk of disease establishment in wild pigs. *J. Anim. Ecol.* (2020) https:/doi.org/10.1111/1365-2656.13412.