Host relatedness and landscape connectivity shape pathogen spread in the puma, a large secretive carnivore

Urban expansion can fundamentally alter wildlife movement and gene flow, but how urbanization alters pathogen spread is poorly understood. Here, we combine high resolution host and viral genomic data with landscape variables to examine the context of viral spread in puma (Puma concolor) from two contrasting regions: one bounded by the wildland urban interface (WUI) and one unbounded with minimal anthropogenic development (UB). We found landscape variables and host gene flow explained significant amounts of variation of feline immunodeficiency virus (FIV) spread in the WUI, but not in the unbounded region. The most important predictors of viral spread also differed; host spatial proximity, host relatedness, and mountain ranges played a role in FIV spread in the WUI, whereas roads might have facilitated viral spread in the unbounded region. Our research demonstrates how anthropogenic landscapes can alter pathogen spread, providing a more nuanced understanding of host-pathogen relationships to inform disease ecology in free-ranging species.
Understanding how pathogens spread through populations remains a fundamental challenge. The extent to which pathogen spread reflects movement patterns of their hosts is enigmatic but important for controlling disease\textsuperscript{12,13}. If pathogen spread mirrors host gene flow, host genetic structure/differentiation could be a valuable proxy for pathogen spread and be used as a basis to inform disease control\textsuperscript{3,4} (e.g., male vampire bat [Desmodus rotundus] genetics closely mirrors phylogenetic structure of rabies\textsuperscript{5}). A close relationship between host gene flow and pathogen spread may also be evidence for increased transmission between related conspecifics, and could affect evolutionary pressures on the pathogen, as closely related hosts may be more likely to have similar immune environments\textsuperscript{6}. In contrast, if host gene flow and pathogen spread are decoupled, fine-scale patterns of host movement, for example, may best predict spread and thus inform the strategy employed for disease control\textsuperscript{12,10,13}. The patterns of pathogen spread can also be influenced by characteristics of the pathogen itself, where host-specific and directly transmitted pathogens likely have the greatest concordance with host gene flow, relative to multi-host and environmentally transmitted pathogens. Given this range of scenarios, a better understanding of how host relatedness and environmental predictors drive these processes would improve our estimates of pathogen spread in heterogeneous landscapes.

Urbanization is one of the most destructive and large-scale of all anthropogenic landscape fragmentation processes, but how urbanization shapes pathogen spread in particular is still not well understood\textsuperscript{12,13}. As urban development fragments habitats and introduces barriers (the wildland–urban interface, WUI), it can cause reduced host gene flow between populations\textsuperscript{10,14}, altered animal behaviour (for example, animals becoming more nocturnal to avoid humans\textsuperscript{15,16}), and changes in feeding\textsuperscript{17} and movement\textsuperscript{18} patterns. If these anthropogenic impacts on host behaviour affect transmission dynamics, they may manifest in the demographics of pathogen populations\textsuperscript{19} (e.g., if transmission events are happening rapidly, the pathogen’s effective population size may be exponentially increasing\textsuperscript{19}). Comparing the factors that shape pathogen spread in populations that are affected by urbanization is often difficult due to a lack of high-resolution data (i.e., coupled host and pathogen genomic data for most individuals) or comparable populations (i.e., well-sampled populations impeded and unimpeded by urbanization). Quantifying how urbanization can affect host gene flow, and how this in turn impacts the transmission dynamics and spread of pathogens, can help address this important research gap.

Here, we determine how landscape variables (including those associated with urbanization) and host relatedness affect pathogen spread and transmission in puma (Puma concolor). Puma are useful indicators of the effects of urbanization on wildlife as they are sensitive to urban development\textsuperscript{17,20} but can persist in areas useful indicators of the effects of urbanization on wildlife as they spread and transmission in puma (\textit{P. concolor}) and host relatedness affect pathogen spread and transmission in puma (\textit{P. concolor}) and host relatedness affect pathogen spread and transmission in puma (\textit{P. concolor}).

Region-specific viral demographic histories. We found not only distinct demographic histories in the viruses circulating in the WUI and UB regions, but also differing FIV\textsubscript{pco} subtypes. Bayesian time-scaled phylogenetic analysis of the FIV\textsubscript{pco} sequences revealed two co-circulating FIV\textsubscript{pco} subtypes: FIV\textsubscript{pco} CO, circulating among pumas in both regions (Fig. 1a) and FIV\textsubscript{pco} WY, which was only detected in the UB after having been previously detected in puma in Wyoming (Fig. 1b; see Fig. S1 for a maximum-likelihood tree that illustrates the broader phylogenetic context of these two subtypes across North America). Within FIV\textsubscript{pco} CO, we identified three clades (I, II and III, Fig. 1a) that had contrasting and landscape-specific demographic histories (Fig. S2). Clade I had been circulating predominantly in the WUI since \textbackslash~1995 (95% high posterior density interval (HPD): 1984–2003), and the effective population size of this clade has been gradually increasing through time (Fig. S2). Clade II showed a similar trajectory in population size (Fig. S2) and was found in puma from both regions, which is potentially indicative of long-distance dispersal of FIV\textsubscript{pco} (Fig. 1a). Clade III, in contrast, predominantly circulated in the UB, but had a much more distinctive demographic pattern (Fig. S2). We estimated that Clade III began circulating in the WUI in 2001 (95% HPD: 1992–2006) and arrived in the UB in 2006 (95% HPD: 2003–2008), afterwards going through a period of population growth which plateaued around 2012 (Fig. S3). In contrast, we found that FIV\textsubscript{pco} WY has likely circulated at low prevalence in the UB for over a hundred years (Fig. S3) and had a slower estimated evolutionary rate than FIV\textsubscript{pco} CO (Table S2).

Divergent patterns of viral and host relatedness across regions. Overall, despite regional fidelity, the FIV\textsubscript{pco} phylogeny did not evolv...
map closely onto the puma relatedness cladogram; yet there was some localized evidence for concordance between the two in the WUI region (grey boxes, Fig. 1a–c). For example, four related individuals in the WUI were infected with phylogenetically similar FIVpco (dark grey box, Fig. 1a–c) and were also captured in close spatial proximity to each other (dashed circle, Fig. 2d). In contrast, there was limited evidence of similar patterns in the UB as the most phylogenetically similar FIVpco isolates were sampled across unrelated individuals. In the UB, a significantly higher proportion of individuals in each ‘neighbourhood’ (i.e., pumas likely to have home-range overlap) were qPCR negative for FIVpco than puma in the WUI (Mann–Whitney U Test, \( p = 0.007 \), Fig. S4). In addition, there was qualitative evidence for spatial structuring with some FIVpco lineages being locally dominant (e.g., CO clade I and III in Fig. 2). However, overall FIVpco spread was a complex mixture of local and longer-distance jumps across both landscapes with uninfected individuals captured at the same time and within 500 m of infected individuals in both regions (Fig. 2). Similarly, individuals captured within months of each other at the same location were commonly infected with phylogenetically distinct FIVpco subtypes or clades (Fig. 2).

Predictors of FIV spread are region-dependent. We employed generalized dissimilarity models (GDM) and maximum likelihood of population-effects (MLPE) to test how host and landscape shaped two components (the overall phylogeographic pattern and lineage dispersal velocity) of spread in each region. Landscape variables for both techniques were formulated using a resistance/conductance approach. We calibrated our resistance/conductance costs based on expert opinion as well as optimizing the landscape variables using host genetic distance using the Resistance GA routine. In the UB, our optimization approach revealed that none of the landscape variables explained host gene flow more than the null model (i.e. models were >2 AICc units higher than the null or the model with no landscape variables included, Table S4). Subsequently, no host-genetic optimized surfaces were included in the UB model (see ‘Methods’). In contrast, our optimisation approach identified a stronger impact of spatial proximity on host gene flow in the WUI (Table S5). Canopy cover and urban land cover univariate models were also within 2 AICc units of the spatial proximity model, thus we combined these variables together to generate a multivariate, host genetics optimised resistance surface (hereafter called host-optimized resistance surface).
The host-optimized resistance surface was strongly correlated with interpolated host genetic resistance (Mantel $\rho = 0.76$, $p = 0.001$). We thus present models in the WUI region with host genetic resistance and host-optimized resistance included separately. In both regions, to tease apart the effect of host genetics and landscape on FIV spread, we also included the non-host genetics optimized landscape resistance/conductance surfaces (hereafter called landscape variables) in our models as well as spatial proximity and host variables.

Strikingly, landscape and host variables explained significant variation of FIVpco spread for the WUI only. In the WUI, our GDM models explained 20% of total model deviance ($p = 0.012$) but only 7% in the unbounded population ($p = 0.23$). Moreover, the most important variables that shaped spread in each case were different (Fig. 3a). To support these results, we compared our non-linear GDM models to MLPE. One advantage of the GDM method over MLPE and other methods is that it can capture non-linear associations between response and predictor matrices. However, model performance of MLPE has been more rigorously evaluated compared to GDM on landscape genetic datasets$^{38}$. Here, we highlight factors that explain the most deviance in our GDM models and are within two log units of the best performing MLPE models using BIC (Tables S6 and S7). In the UB, FIVpco spread was associated with roads (i.e., individuals more connected by roads had similar FIV isolates, Fig. 3b, Table S6). In contrast, the viral spread in the WUI was shaped by spatial proximity coupled with host relatedness and impervious surface (Fig. 3a, Table S7). As spatial proximity decreased, so did the FIVpco patristic distance between individuals; neighbouring individuals shared more phylogenetically similar FIVpco isolates (Fig. 3c). Similarly, as host relatedness decreased, so did FIVpco patristic distance (i.e., related individuals were more likely to share phylogenetically similar FIVpco isolates, Fig. 3e). A different measure of individual genetic distance (the Smouse measure$^{39}$) did not alter our results. We found a similar positive relationship between FIVpco patristic distance and impervious surface resistance (Fig. 3e). Furthermore, in complement to our relatedness measure, we also included host genetic resistance in our GDM models (see 'Methods' for details). Individuals in the WUI with low host genetic resistance values had more similar phylogenetically FIVpco isolates (Fig. 3f). However, this relationship plateaued with dissimilarity values of over 0.05 and was not significant ($p = 0.38$).

We also found that there were region-specific impacts of landscape on FIVpco lineage dispersal velocity. Our analysis revealed that elevation tended to act as a conductance factor increasing the dispersal velocity of FIVpco lineages in the WUI, whereas none of the predictors we measured had any substantial effect on lineage dispersal velocity in the UB (positive $Q$ distribution and associated Bayes factor support $>3^{45}$; see Fig. S6 and Table S7 for a list of tested landscape factors). Furthermore, warmer minimum temperatures (as measured in the coolest month) tended to act as a resistance factor decreasing lineage dispersal velocity in the same region. Taken together, these results indicate that in the WUI, FIVpco tended to spread faster through colder, higher elevation areas less suitable for puma habitat. In the WUI, the areas of higher elevation tended to be away from the urban edge.
Discussion

Our multifaceted approach linking landscape, host genomic, and pathogen genomic data uncovered unique landscape-specific relationships with FIVpco spread, indicative of altered epidemiological dynamics associated with urban landscape structure. We found that spatial proximity was positively associated with FIVpco spread, but only in the WUI (Wildland Urban Interface). In the WUI, host relatedness, host-optimised resistance and impervious surface also had a minor association with FIVpco spread. These landscape factors mirrored the factors shaping host gene flow in the WUI26, providing support that host movement and viral spread were more intimately related in the WUI than in the UB (Unbounded) region. However, while there was some evidence for concordance between the FIVpco phylogeny and host cladogram in the WUI, they did not map precisely onto each other in either region (Fig. 1), indicating that transmission occurs outside of related individuals. There was little evidence of FIVpco/host concordance in the UB where host relatedness did not shape phylogeography and entirely different sets of predictors shaped host gene flow (spatial proximity and tree cover)26. These opposing patterns between host gene flow and viral spread could reflect regional differences in transmission. One potential scenario is that transmission between neighbouring related conspecifics may be more likely in the WUI due to altered puma movement, dispersal patterns22,23 and foraging behaviours17. The urban development that impacts the WUI is linear (i.e., where the Great Plains meet the Front Range of the Rocky Mountains). This linear development restricts juvenile dispersal23 and could lead to more opportunities for transmission among related conspecifics (i.e., individuals establish home ranges close to their parents). Our previous work supports this hypothesis as we demonstrated that family units were more clustered in space, where puma genetic distance per kilometre and sub-structure was greater in the WUI compared to the UB26. There has not been an extensive evaluation of relatedness in neighbouring home-range females, but female matrilines (groups of maternally-related females) are known to occur11,42. Furthermore, we found evidence that infection was more clustered in space in the WUI compared to the unbounded one, supporting the idea that spatial proximity increases transmission risk between related individuals in the WUI. This shift in transmission risk may reduce evolutionary pressure on the virus due to factors such as similarity in immune profile9.

Roads, common but mostly unpaved in this UB region, were a modest predictor of spread in the UB. Radiotelemetry has shown that puma often move using unpaved roads43 and rapid viral evolution potentially allowed us to detect this modest effect. There was a smaller impact of roads on host gene flow26 which supports the idea that FIV phylogenetics may capture more contemporary movement patterns impossible to detect using host genetics alone10,11,44. When host gene flow and viral spread are decoupled, as was the case in the UB, the rapid accumulation of viral mutations may conversely obscure historical trends in connectivity45. This could explain why we detected no effect of tree cover on FIVpco spread even though puma are known to have a preference for tree cover to disperse and hunt46,47. The altered epidemiological history of the clade, dominant in the UB compared to all other detected clades, may reflect or be a consequence of relatively unrestricted spread in the UB (Fig. S3, FIVpco CO clade I). We postulate that recent arrival of FIVpco CO clade I in the UB and signature of rapid expansion19 across the Uncompahgre Plateau (Fig. S2) may only have been possible in a region where viral spread itself was unbounded. Further work is needed to assess the temporal dynamics of FIVpco in the UB. The high elevation Uncompahgre Plateau (averaging 2900 m a.s.l., Fig. 2) did not shape viral spread, even though puma are known to have a preference for not dispersing across cold, high altitude divides46. In contrast, we found that higher altitude areas increased dispersal velocity in the WUI. As most human activity in the Front Range (Fig. 2) occurs in lower altitude areas, it is plausible that increased viral velocity in higher altitude areas is a product of the host’s avoidance of the human ‘super predator’17,48.
FIV transmission events have been shown to be more likely to occur further from the urban edge in bobcat populations\textsuperscript{26} and the same may apply for pumas in the WUI region here. Velocity may also be faster through higher elevations in the Front Range as puma are more likely to rapidly move through unsuitable habitat. Unsuitable habitat has also been demonstrated to increase the velocity of rabies lineages in dogs\textsuperscript{49}. We acknowledge that in this study we did not sample all pumas in the system nor between the two sampled regions, and that FIV\textsubscript{pco} could not always be sequenced. This could mean that we missed, for example, some FIV\textsubscript{pco} lineages that could have altered our inference about FIV\textsubscript{pco} patterns. Nonetheless, this did not compromise our ability to gain complementary insights into the drivers of host connectivity by combining high-resolution host and pathogen genomic data, which would have been impossible to detect with either host or pathogen data alone.

Our findings have pathogen and host management implications, as we demonstrated that spatial proximity and host relatedness alterations in landscapes and host relatedness alter pathogen transmission will be increasingly important. We emphasise that impervious surface (further from the urban areas) was less constrained by impervious surface (further from the urban areas) may also reduce spread. Our work provides a valuable case study of how landscape context and host relatedness can be important in disease management plans. As urban landscapes continue to expand, improving our understanding of how heterogeneous landscapes and host relatedness alter pathogen transmission will be increasingly important.

**Methods**

**Samples**
Puma blood and tissue samples were collected from 103 individuals (48 males, 55 females) between 2005 and 2014 in the UB and 110 individuals (43 males, 54 females, 12 undetermined) between 2003 and 2015 from the WUI as part of monitoring efforts by Colorado Parks and Wildlife in the Rocky Mountain Range of Colorado, USA\textsuperscript{23,55}. This sampling effort is likely to represent a large proportion of the resident puma present in both regions during the sampling period\textsuperscript{23}. See Table S1 for further details on the samples.

**FIV detection and sequencing.** Total DNA was extracted from 50 µl whole blood samples using the QIAGEN DNeasy Blood & Tissue extraction kit (Qiagen, Valencia, CA) with an extended incubation period of two hours or from 200 µl whole blood samples using a phenol–chloroform extraction as per Pietro et al\textsuperscript{51}. Isolated DNA was quantified using a QuBit 2.0 fluorometer (Thermo-FisherScientific). DNA from individual puma were screened for the presence of FIV\textsubscript{pco} provirus using a specific qPCR assay as described by Lee et al\textsuperscript{82}. Using these data, we compared prevalence from both regions using a two-sample test for equality of proportions.

Full ORF and pol gene regions were isolated from those samples identified as qPCR positive using a nested PCR protocol. See the Supplementary Note for the sequencing protocol and Table S8 for primer details. While we also sequenced the env gene, our assessment of the temporal signal (see next section) indicated many discrepancies in the data regarding the use of a molecular clock (strict or relaxed) to analyse these data. Resulting genetic sequences with chromatograms were checked, assembled, trimmed and aligned using the MUSCLE algorithm\textsuperscript{176} using Geneious. Our pol and ORF4A datasets (GenBank accessions: MN563193–MN563239) were compiled together with those sequences available in the public database Genbank, previously isolated from across the USA\textsuperscript{54} (Fig S1). Recombination was detected using RDP software V4\textsuperscript{85}; parameters were set at default with linear topology. Events were determined as true if supported by three or more methods with p values <10\textsuperscript{−3} combined with phylogenetic support. Recombination-free datasets were used for all downstream phylogenetic analyses.

**Virul phyllogenetics.** To examine the broad placement of the FIV\textsubscript{pco} isolates sampled during this study, a maximum-likelihood tree was constructed for the FIV\textsubscript{pco} dataset comprised of all isolates recovered in the USA using PhyML\textsuperscript{86} with the TN93 + G + I model and aLRT branch support; branches with <50 support were collapsed using TreeGraph2\textsuperscript{37}. We used Temp\textsuperscript{82} to assess data quality control of our generated data through root-to-tip regression and observed largely varying deviations from the regression line for many of the sequenced env genes. These deviations precluded the use of strict or relaxed molecular clock models to analyse the data, and we hence decided to move forward with the pol and ORF4A sequence data. We used BEAST 1.10.9\textsuperscript{178} to perform MCMC and compare tree topology with assed datasets (GenBank accession: MN563193–MN563239) were compiled together with those sequences available in the public database Genbank, previously isolated from across the USA\textsuperscript{54} (Fig S1). Recombination was detected using RDP software V4\textsuperscript{85}; parameters were set at default with linear topology. Events were determined as true if supported by three or more methods with p values <10\textsuperscript{−3} combined with phylogenetic support. Recombination-free datasets were used for all downstream phylogenetic analyses.

**Host genomic data.** We genotyped 130 pumas (76 individuals from the UB and 54 individuals from the WUI) using a ddRADseq approach (see ref.\textsuperscript{26} for sequence and bioinformatics details). From these genomic data, we quantified individual relatedness using the inverse genetic distance. For the FIV\textsubscript{pco} WY, as there were only 6 sequences, we applied a different approach as there was not enough data to obtain stable results for these complicated evolutionary models. In this case, we assumed a strict clock and constant population size and set the root prior to a uniform distribution (0, 300). We also checked if our expectation that this subtype had been circulating for no more than 300 years. For both subtypes, duplicate MCMC chains were run for 200 million generations, with trees and parameters sampled every 20,000 steps. We used the program Tracer version 1.7\textsuperscript{186} to examine ESS values (with parameter estimates accepted if the ESS was >200) and obtain HPD intervals for estimated parameters.

To identify the hidden population structure in our time-scaled phylogenies, we applied the ‘tree structure’ R package\textsuperscript{69} using the default values. This analytical routine compares discrepancies between observed and idealised genealogies to identify clades under differing epidemiological or demographic processes\textsuperscript{87}. As we did not have enough sequences to find meaningful structure in FIV\textsubscript{pco} WY, we limited this analysis to FIV\textsubscript{pco} CO. Any clades identified as significantly departing from the idealised genealogy were further investigated using the R package ‘phyloDyn’\textsuperscript{70} to estimate effective population size through time. This nonparametric Bayesian approach uses integrated nested Laplace approximation (INLA) to estimate effective population size efficiently, while accounting for preferential sampling by modelling the sampling times as a Poisson process (see ref.\textsuperscript{71}).

**Landscape data.** We collected Geographic Information Systems (GIS)-based landscape data that we hypothesized would be important for puma relatedness in Colorado (see Table S7 for more detail on GIS data sources and ecological justifications). This included people (e.g., roads, buildings, etc.), land cover (forested, open-natural, and human developed as sub-rasters), percentage tree canopy cover, vegetation density, rivers/
streams, roads, minimum temperature of the coldest month, annual precipitation, topographic roughness, and elevation. Resistance surfaces were created from this landscape data using the Inverse Distance and Resistant and RasterGIS v. 10.1. We defined resistance/conductance cost values based on expert opinion as well as the ‘Resistance GA’78 routine in R. Resistance GA optimizes resistance/conductance surfaces based on the host pairwise genetic distances. We used this approach to test which landscape variables best approximated host gene flow and to formulate optimized resistance surfaces for use in viral spread analyses. To run the Resistance GA routine, we used an AIC objective function and tested all possible transformations (i.e. Ricker and Monomolecular transformations). As none of our landscape variables outperformed the null model in the UB region based on AICc, there was no need to further estimate the overall multivariate resistance surface shaping the host as more complex models get increasingly penalized in AICc scores (W. Peterman pers. comm). For the WUI region, canopy cover and urban landscape cover were within 2 AICc units of the best univariate model (spatial proximity), thus we included both features to create a multivariate optimised surface. We included resistance surfaces defined by expert opinion as well as host genetic optimized surfaces to help distinguish the effects of host genetics and landscape on FIVpco spread. Mantel tests were used to test for screen for collinearity between surfaces and for spatial autocorrelation prior to model construction.

We calculated the proportion of infected to infected individuals within a 5 km buffer (estimated average distance among individuals) of each FIVpco positive individual using the “summarize within” tool (also in ArcGIS v. 10.1). Estimates from each population were compared using a Mann–Whitney U Test.

**Statistics and reproducibility**

**Impact of host relatedness and landscape resistance on viral spread.** We used generalized dissimilarity modelling (GDM)35 to quantify if host relatedness and landscape resistance surfaces defined on FIVpco CO. GDM is a flexible non-linear regression approach that fits monotonic I-spline functions to pairwise matrix data35 to describe the rate and magnitude of, in this case, FIVpco phylogenetic change. We first calculated FIVpco patristic distance (using the maximum clade credibility tree) and converted this distance metric into a dissimilarity measure: 

$$d_{\text{GDM}} = \left( \frac{d}{\text{max}d} \right)^2$$

where $d$ is pairwise distance. The host Dps matrix and all of the landscape resistance matrices were converted to dissimilarities the same way. Specifically, GDM uses generalized linear models (GLMs) to model FIVpco patristic distance in the form of:

$$-\ln(d_{ij}) = \beta_0 + \sum_{p=\text{predictors}} \beta_p x_{ij}$$

where $i$ and $j$ are individual puma, $d_{ij}$ is the intercept, $p$ is the number of covariates and $x_{ij}$ have I-spline transformed versions of the predictors (see refs. 35, 58; for further details). We used a backward elimination model selection approach and permutation tests ($n = 99$) to test for significance.56,77 The model with the highest deviance (±2% deviance explained) with the smallest number of predictors was reported. We performed GDM using the same predictor sets as in Trumbel et al.26 for each region but in our reduced dataset (i.e., with individuals with FIVpco data), temperature and elevation were strongly correlated with space (Mantel $r = 0.93$). We tested different values of $K$ (see below) and treated each as resistance or conductance surfaces yet this made no difference to the GDM models (we present results from resistance surfaces only with $K = 100$). Analysis of FIVpco WY was not possible given the small sample size.

We compared our non-linear GDM models to maximum likelihood of population-estimate (MLPE) models48 to test the validity of our results. Instead of fitting non-linear splines, the MLPE approach uses linear mixed models to fit variables. We compared univariate MLPE models using Bayesian information criterion (BIC) model selection.

**Impact of environmental factors on viral dispersal velocity.** The analysis of the impact of environmental factors and host genetic differentiation on the dispersal velocity of viral lineages was performed using R functions of the package ‘seraphim’98 (see refs. 91, 92 for a similar workflow). In this analysis, each environmental factor, as well as the interpolated host genetic distance surface, was described by a raster that describes its spatial heterogeneity and that was used to compute an environmental distance for each branch in the phylogeny using two different path models: (i) the least-cost path model, which uses a least-cost algorithm to determine the route taken between the starting and ending point83, and (ii) the Circuitscape path model. Here, we investigated the impact of the environmental rasters listed in Table S5 as well as the resistance raster generated from host genetic distance interpolation. We generated distinct land cover rasters from the original categorical land cover raster (resolution = 0.5 arcmin) by creating lower resolution rasters (2 arcmin) whose cell values equaled the number of occurrences of each land cover category within the 2 arcmin cells.41 For each considered environmental factor, several distinct rasters were also generated by transforming original raster cell values with the following formula: 

$$v_{ij} = 1 + k(f_{ij}/v_{ij})$$

where $v_{ij}$ and $v_{ij}$ are the transformed and original raster cell values, and $v_{ij}$ is the minimum raster cell value recorded in the raster. The rescaling parameter $k$ here allowed the definition and testing of different strengths of raster cell conductance or resistance, relative to the conductance/resistance of a cell with a minimum value set to 1. For each environmental factor, we tested three different values for $k$ (i.e., 10, 100, and 1000). Finally, all these rasters were tested as potential conductance factors (i.e., factors facilitating movement) and as possible resistance factors (i.e., factors impeding movement). The statistic Q was used to estimate the correlations between phylogenetic branch duration and environmental distances, Q is defined as the difference between two coefficients of determination ($R^2$): (i) $R^2$ obtained when branch durations are regressed against environmental distances computed on the environmental raster, and (ii) $R^2$ obtained when branch durations are regressed against environmental distances computed on a null raster, i.e., an environmental raster with a value of 1 assigned to all the cells. For positive distributions of estimated Q values (i.e., with at least 90% of positive values), statistical support was then evaluated against a null distribution generated by a randomization procedure and formalized as an approximated Bayes factor (BF) support 84. To account for the uncertainty related to the Bayesian inference, this analysis was based on 1000 trees sampled from the post-burn-in posterior distribution inferred using the continuous phylogeographic model. We performed two distinct analyses, one per region, gathering all phylogenetic branches occurring on each study area.

**Ethics statement.** Puma samples were collected as part of ongoing studies by Colorado Parks and Wildlife (CPW) between 2006 and 2014. We handled all pumas in accordance with approved CPW Animal Care and Use Committee (ACUC) capture and handling protocols (ACUC file 408-2004, ACUC protocol 03-2007 ACUC 16-2008). Samples were provided to Colorado State University for diagnostic evaluation. Colorado State University and Colorado Parks and Wildlife (CPW) Institutional Animal Care and Use Committees reviewed and approved this work prior to initiation (CSU IACUC protocol 05-061A).

**Data availability**

DNA sequences—GenBank accession numbers MN563193–MN563239. The sequence alignment file used to create the phylogenies in this paper as well as the data to reproduce the generalized dissimilarity models and phylogeographic models is available on GitHub: https://github.com/nflj1380/ColoradoPumaFIVproject.

**Code availability**

All code is available on GitHub: https://github.com/nflj1380/ColoradoPumaFIVproject.

Received: 12 November 2019; Accepted: 25 November 2020; Published online: 04 January 2021

**References**

1. Daversa, D. R., Fenton, A., Dell, A. L., Garner, T. W. J. & Manica, A. Infections on the move: how transient phases of host movement influence disease spread. Proc. R. Soc. B Biol. Sci. 284, 20171807 (2017).

2. Martinez-Guilmo, E., Blanchet, S., McCoy, K. D. & Loot, G. Host dispersal as the driver of parasite genetic structure: a paradigm lost? Ecol. Lett. 19, 336–347 (2016).

3. Biek, R. & Real, L. A. The landscape genetics of infectious disease emergence and spread. Mol. Ecol. 19, 3515–3531 (2010).

4. Kozakiewicz, C. P. et al. Pathogens in space: Advancing understanding of pathogen dynamics and disease ecology through landscape genetics. Evol. Appl. https://doi.org/10.1111/eva.12678 (2018).

5. Brüniche-Olsen, A., Burridge, C. P., Austin, J. J. & Jones, M. E. Disease induced changes in gene flow patterns among Tasmanian devil populations. Biol. Conserv. 165, 69–78 (2013).

6. Kyle, C. J. et al. Spatial patterns of neutral and functional genetic variations reveal patterns of local adaptation in raccoon (Procyon lotor) populations exposed to raccoon rabies. Mol. Ecol. 23, 2287–2298 (2014).

7. Schwabl, P. et al. Prediction and prevention of parasitic diseases using a landscape genomics framework. Trends Parasitol. 33, 264–275 (2017).

8. Streicker, D. G. et al. Host-pathogen evolutionary signatures reveal dynamics and future invasions of vampire bat rabies. Proc. Natl. Acad. Sci. USA 113, 10926–10931 (2016).

9. Gisberts, E. F. et al. Low level of HIV-1 evolution after transmission from mother to child. Sci. Rep. 4, 6460–6465 (2014).

10. Lee, J. S. et al. Gene flow and pathogen transmission among bobcats (Lynx rufus) in a fragmented urban landscape. Mol. Ecol. 21, 1617–1631 (2012).

11. Fountain-Jones, N. M. et al. Urban landscapes can change virus gene flow and evolution in a fragmentation-sensitive carnivore. Mol. Ecol. 26, 6487–6497 (2018).

12. Brearley, G. et al. Wildlife disease prevalence in human-modified landscapes. Biol. Rev. Camb. Philos. Soc. 88, 427–442 (2013).
13. Mcdonald, R. I., Kareiva, P. & Forman, R. T. T. The implications of current and future urbanization for global protected areas and biodiversity conservation. Biol. Conserv. 141, 1695–1703 (2008).

14. Riley, S. P. D. et al. A southern California freeway is a physical and social barrier to gene flow in carnivores. Mol. Ecol. 15, 1733–1741 (2006).

15. Gaynor, K. M., Hojnecwski, C. E., Carter, N. H. & Brashares, J. S. The influence of human disturbance on wildlife nocturnality. Science 360, 1232–1235 (2018).

16. Rützler, S. P. et al. Effects of urbanization and habitat fragmentation on bobcats and coyotes in southern California. Conserv. Biol. 17, 566–576 (2003).

17. Smith, J. A. et al. Fear of the human ‘super predator’ reduces feeding time in large carnivores. Proc. Biol. Sci. 284, 20170433 (2017).

18. Tracey, J. A., Bevins, S. N., VandeWoude, S. & Crooks, K. R. An agent-based movement model to assess the impact of landscape fragmentation on disease transmission. Ecol. Appl. 5, 119 (2014).

19. Volz, E. M., Koelle, K. & Bedford, T. Viral phylodynamics. PLoS Comput. Biol. 9, e1002947 (2013).

20. Ordeñana, M. A. et al. Effects of urbanization on carnivore species distribution and richness. J. Mammal. 91, 1322–1331 (2010).

21. Crooks, K. R. Relative sensitivities of mammalian carnivores to habitat fragmentation. Conserv. Biol. 16, 488–502 (2002).

22. Blecha, K. A., Boone, R. B. & Allredge, M. W. Hunger mediates apex predator’s risk avoidance response in wildland-urban interface. J. Anim. Ecol. 87, 609–622 (2018).

23. Lewis, J. S. et al. The effects of urbanization on population density, occupancy, and detection probability of wild felids. Ecol. Appl. 25, 1880–1895 (2015).

24. Bradley, C. A. & Altizer, S. Urbanization and the ecology of wild diseases. Trends Ecol. Evol. 22, 95–102 (2007).

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

26. Trumbo, D. et al. Urbanization impacts apex predator gene flow but not genetic diversity across an urban-rural divide. Mol. Ecol. 28, 4926–4940 (2019).

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

28. Brown, E. W., Yuhki, N., Packer, C. & O’Brien, S. J. A lion lentivirus related to feline immunodeficiency virus: epidemiologic and phylogenetic aspects. J. Virol. 68, 5953–5968 (1994).

29. Biek, R. et al. Epidemiology, genetic diversity, and evolution of enzootic feline immunodeficiency virus in a population of wild cougars. J. Virol. 77, 9578–9589 (2003).

30. Biek, R., Ruth, T. K., Murphy, K. M., Anderson, C. R. Jr. & Poss, M. Examining effects of persistent retroviral infection on predator transmission model to assess the impact of landscape fragmentation on disease flow in carnivores. PLoS Comput. Biol. 12, e1005276 (2016).

31. Reynolds, J. J. H. et al. Feline immunodeficiency virus in puma: estimation of force of infection reveals insights into transmission. Ecol. Evol. ece3558, https://doi.org/10.1002/ece3.5584 (2019).

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

33. Fountain-Jones, N. M. et al. Towards an eco-phylogenetic framework for infectious disease ecology. Biol. Rev. 93, 950–970 (2018).

34. Smith, J. A. et al. Fear of the human ‘super predator’ reduces feeding time in large carnivores. Proc. R. Soc. London B Biol. Sci. 284, 20170433 (2017).

35. Ferrier, S., Manion, L., Elith, J. & Richardson, K. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. Divers. Distrib. 13, 252–264 (2007).

36. Clarke, R. T., Rothey, R. & Raybould, A. F. Conidium limits for regression model comparison while accommodating phylogenetic uncertainty. Syst. Biol. 68, 1052–1061, https://doi.org/10.1093/sysbio/syz020 (2019).

37. Lemey, P., Bambaut, A., Welch, J. J. & Suchard, M. A. Phylogeny takes a relaxed random walk in continuous space and time. Mol. Biol. Evol. 27, 1877–1885 (2010).

38. Lefort, V., Longueville, J.-E. & Gascuel, O. SMS: Smart model selection in PhyML. Mol. Biol. Evol. 34, 2422–2424 (2017).

39. Gill, M. S. et al. Improving Bayesian population dynamics inference: a coalescent-based model for species with widespread distributions: temporal and spatial patterns of genetic diversity. Phil. Trans. R. Soc. B 373 (2018).

40. Baelde, G. et al. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. Mol. Biol. Evol. 29, 2157–2162 (2012).

41. Lartillot, N. & Philippe, H. Computing Bayes factors using thermodynamic integration. Syst. Biol. 55, 195–207 (2006).

42. Xie, W., Lewis, P. O., Fan, Y., Kuo, L. & Chen, M.-H. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. Syst. Biol. 60, 150–160 (2011).

43. Baele, G., Lemey, P. & Vansteelandt, S. Make the most of your samples: Bayes factor estimators for high-dimensional models of sequence evolution. BMC Bioinformatics. 14, 85 (2013).

44. Bambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901–904 (2018).

45. Volz, E. M. et al. Identification of hidden population structure in time-scaled phylogenies. Syst. Biol. 69, 884–896, https://doi.org/10.1093/sysbio/syy009 (2019).

46. Karcher, M. D., Palacios, J. A., Lan, S. & Minin, V. N. phylod: an R package for phylogenetic simulation and inference. Mol. Ecol. Resour. 17, 96–100 (2017).

47. Karcher, M. D., Palacios, J. A., Bedford, T., Suchard, M. A. & Minin, V. N. Quantifying and mitigating the effect of preferential sampling on phylophanydemic inference. PLoS Comput. Biol. 12, e1004789 (2016).

48. Stowcock, A. M. et al. High resolution of human evolutionary trees with polymorphic microsatellites. Nature 368, 455–457 (1994).

49. McRae, B. H. Isolation by resistance. Evolution 60, 1551–1561 (2006).
Author contributions
N.F.J. conducted the analysis and wrote the initial draft of the paper to which all authors contributed. K.L. and M.A. studied the puma in the field and provided the blood samples. S.K., D.T., P.S., E.G. and S.V. collected virus and host genetic data. S.D., R.B. and G.B. contributed to the biogeographic and phylogenetic analyses. M.C., S.V., C.F., H.E. and S.C. conceived of the project.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s42003-020-01548-2.

Correspondence and requests for materials should be addressed to N.M.F.-J.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements
This project was funded by a National Science Foundation Ecology of Infectious Diseases research program grants (DEB 1413925). M.C. was funded by the University of Minnesota’s Office of the Vice President for Research and Academic Health Center Seed Grant. G.B. acknowledges support from the Interne Fondsen KU Leuven/Internal Funds KU Leuven under grant agreement C14/18/094, and the Research Foundation—Flanders (“Fonds voor Wetenschappelijk Onderzoek—Vlaanderen,” G0E1420N). We thank Bill Peterman for his advice on the resistanceGA pipeline.