









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

Nicholas M. Fountain-Jones <sup>1,2✉</sup>, Simona Kraberger<sup>3</sup>, Roderick B. Gagne <sup>3</sup>, Daryl R. Trumbo<sup>4</sup>, Patricia E. Salerno<sup>4,5</sup>, W. Chris Funk <sup>4</sup>, Kevin Crooks<sup>6</sup>, Roman Biek <sup>7</sup>, Mathew Alldredge<sup>8</sup>, Ken Logan<sup>9</sup>, Guy Baele<sup>10</sup>, Simon Dellicour <sup>10,11</sup>, Holly B. Ernest<sup>12</sup>, Sue VandeWoude <sup>3</sup>, Scott Carver <sup>2</sup> & Meggan E. Craft <sup>1</sup>

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.

<sup>1</sup> Department of Veterinary Population Medicine, University of Minnesota, St Paul, MN 55108, USA. <sup>2</sup> School of Natural Sciences, University of Tasmania, Hobart, Australia 7001. <sup>3</sup> Department of Microbiology, Immunology, and Pathology, Colorado State University, Fort Collins, CO 80523, USA. <sup>4</sup> Department of Biology, Graduate Degree Program in Ecology, Colorado State University, Fort Collins, CO 80523, USA. <sup>5</sup> Universidad Regional Amazónica IKIAM, Km 7 vía Muyuna, Tena, Ecuador. <sup>6</sup> Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, CO 80523, USA. <sup>7</sup> Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow G12 8QQ, UK. <sup>8</sup> Colorado Parks and Wildlife, Fort Collins, CO 80526, USA. <sup>9</sup> Colorado Parks and Wildlife, Montrose, CO 81401, USA. <sup>10</sup> Department of Microbiology, Immunology and Transplantation, Rega Institute, KU Leuven, Herestraat 49, 3000 Leuven, Belgium. <sup>11</sup> Spatial Epidemiology Lab (SpELL), Université Libre de Bruxelles, CP160/12 50, av. FD Roosevelt, 1050 Bruxelles, Belgium. <sup>12</sup> Wildlife Genomics and Disease Ecology Lab, Department of Veterinary Sciences, University of Wyoming, Laramie, WY 82070, USA. ✉email: [Nick.FountainJones@utas.edu.au](mailto:Nick.FountainJones@utas.edu.au)

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<sup>1,2</sup>. 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<sup>3–7</sup> (e.g., male vampire bat [*Desmodus rotundus*] genetics closely mirrors phylogenetic structure of rabies<sup>8</sup>). 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<sup>9</sup>. 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<sup>2,10,11</sup>. 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<sup>12,13</sup>. As urban development fragments habitats and introduces barriers (the wildland–urban interface, WUI), it can cause reduced host gene flow between populations<sup>10,14</sup>, altered animal behaviour (for example, animals becoming more nocturnal to avoid humans<sup>15,16</sup>), and changes in feeding<sup>17</sup> and movement<sup>18</sup> patterns. If these anthropogenic impacts on host behaviour affect transmission dynamics, they may manifest in the demographics of pathogen populations<sup>19</sup> (e.g., if transmission events are happening rapidly, the pathogen's effective population size may be exponentially increasing<sup>19</sup>). 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<sup>17,20</sup> but can persist in areas impacted by urbanization provided sufficient landscape connectivity (e.g., ref. 21). As puma foraging, movement, and other behaviours are altered by urban development<sup>17,22,23</sup>, this species offers a valuable case study for how pathogen spread can be effected by urbanization—subject matter that is increasingly important for wildlife conservation and management<sup>24,25</sup>. We utilize data we collected from 217 pumas sampled from two geographically distinct regions (~500 km apart): one region bounded by the wildland–urban interface (hereafter the WUI) and the other in a more wild and rural setting relatively unbounded by anthropogenic development (hereafter UB). Our previous work found limited gene flow between pumas from these two regions but similar levels of genetic diversity<sup>26</sup>. From individuals in both regions, we collated high-resolution host genomic and spatial data alongside puma feline immunodeficiency virus (FIV<sub>pco</sub>) sampled from the same individuals. FIV<sub>pco</sub> is a rapidly

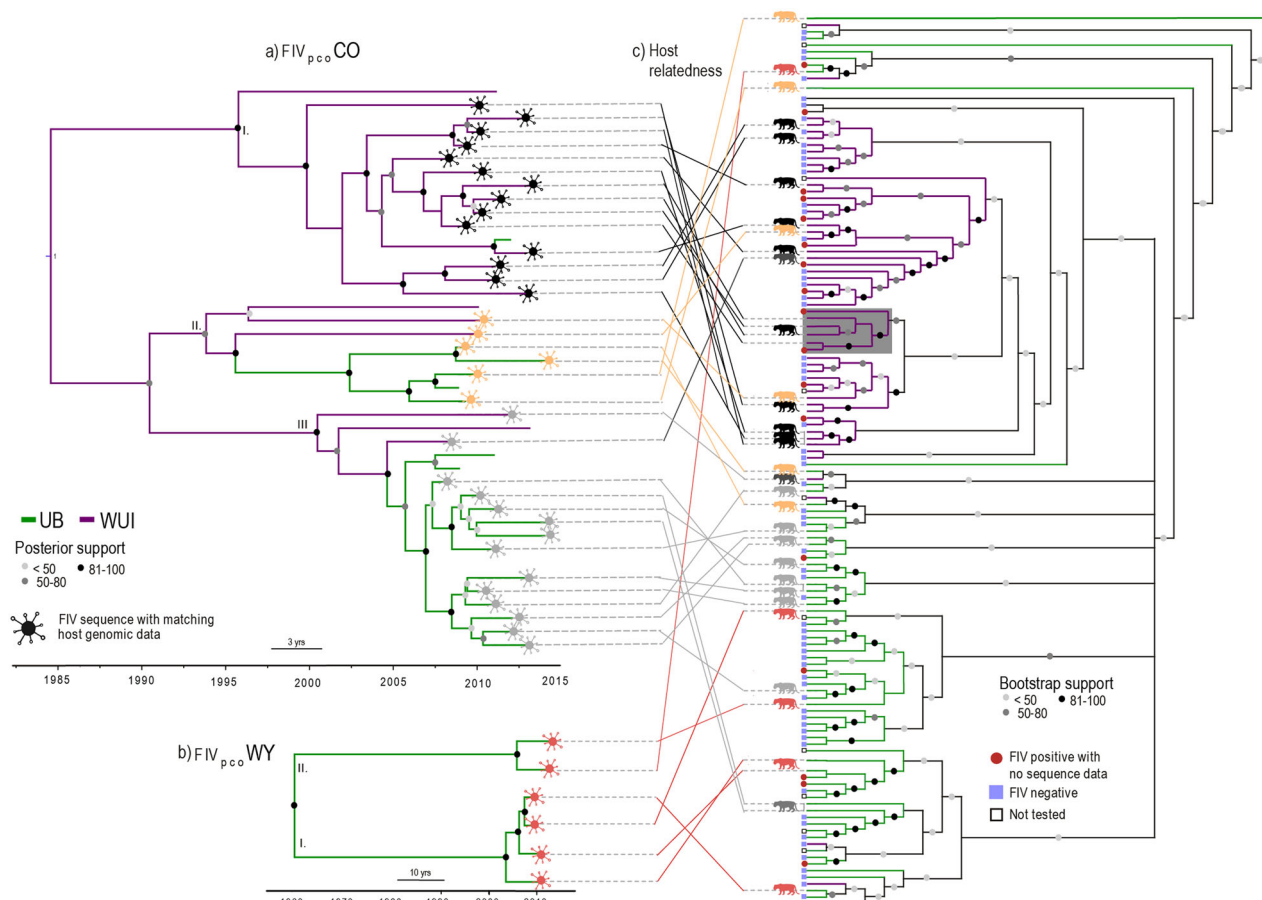
evolving retrovirus<sup>27</sup> endemic to puma populations, and is thought to be predominantly transmitted horizontally via aggressive encounters<sup>28</sup>, although vertical transmission has been documented by phylogenetic analyses<sup>29</sup>. Because FIV<sub>pco</sub> is essentially apathogenic in puma<sup>30,31</sup>, it is an ideal model pathogen to understand transmission dynamics in wild systems without potential confounding effects of disease on behaviour and demography<sup>29,32</sup>. We examine what factors impact FIV<sub>pco</sub> spread using a novel pipeline synthesizing phylodynamic, phylogeographic and landscape genetic techniques (an ecophylogenetic approach<sup>33</sup>). We employ this pipeline to test for (1) differences in FIV<sub>pco</sub> demographic histories and transmission dynamics across regions, (2) concordant patterns of host relatedness, viral phylogenetics and spatial distance, and (3) the relative roles of host relatedness and landscape predictors, such as urban development, in shaping the pattern of spread of the virus. We hypothesized that as anthropogenic factors impact puma movement (e.g., ref. 34) and gene flow<sup>26</sup> at the WUI, that transmission opportunities would be restricted and spatial proximity and host relatedness would be more important in shaping spread in this region.

## Results

Of the 217 individuals we tested, we found that FIV<sub>pco</sub> prevalence was higher in the WUI than UB puma (58% vs 41%,  $p = 0.04$  (two-sample test for equality of proportions), see Table S1 for site and population characteristics). We sequenced FIV<sub>pco</sub> from a total of 46 animals representing most of the infected animals in both populations over a 10-year period. For 43 of the pumas, we obtained both FIV<sub>pco</sub> sequences (the conserved *pol*, *ORF4* and *env* genes representing 36% of the FIV genome) and the corresponding puma genomic data (consisting of a dataset of 12,444 neutral single nucleotide polymorphisms [SNPs] per individual<sup>26</sup>).

**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<sub>pco</sub> subtypes. Bayesian time-scaled phylogenetic analysis of the FIV<sub>pco</sub> sequences revealed two co-circulating FIV<sub>pco</sub> subtypes: FIV<sub>pco</sub> CO, circulating among pumas in both regions (Fig. 1a) and FIV<sub>pco</sub> 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<sub>pco</sub> 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 ~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<sub>pco</sub> (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<sub>pco</sub> 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<sub>pco</sub> CO (Table S2).

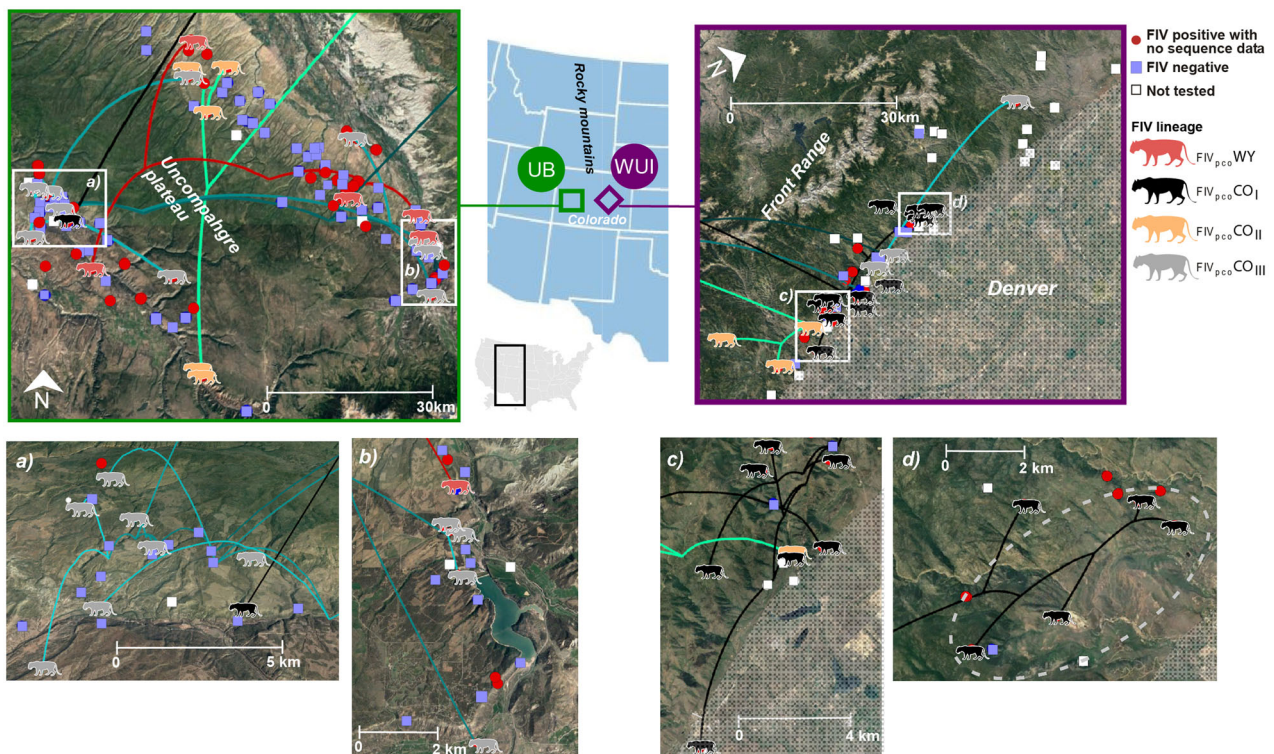
**Divergent patterns of viral and host relatedness across regions.** Overall, despite regional fidelity, the FIV<sub>pco</sub> phylogeny did not



**Fig. 1 Tanglegram revealing how FIV<sub>pco</sub> phylogenetic relationships overall mapped imprecisely onto the puma relatedness cladogram in our study regions.** **a** Bayesian time-scaled phylogenetic tree for FIV<sub>pco</sub> subtype CO found in both the wildland-urban interface (WUI) and unbounded (UB) regions. **b** Bayesian time-scaled phylogenetic tree for FIV<sub>pco</sub> subtype WY which was only found in the UB. I–III represent the different clades identified using tree structure analysis<sup>69</sup>. Virus branch colours are based on population assignment posterior values from our FIV<sub>pco</sub> subtype CO discrete trait analysis. **c** Host relatedness cladogram constructed using singular value decomposition (SVD) quartets<sup>37</sup> based on over 12,000 SNPs from 130 individual puma across both study areas. The grey shaded box encompasses related individuals with phylogenetically similar FIV<sub>pco</sub> isolates. Virus symbols (from panels **a** and **b**) are coloured based on viral lineage membership. This colour matches the puma infected with that isolate (puma silhouettes **c**) and the lines connecting each isolate to each host in the tanglegram. Tips without virus symbols indicate that there was no matching host genomic data for this FIV<sub>pco</sub> isolate. Branch colours indicate which region each individual puma and matching virus was sampled from (WUI or UB).

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 FIV<sub>pco</sub> (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 FIV<sub>pco</sub> 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 FIV<sub>pco</sub> than puma in the WUI (Mann–Whitney U Test,  $p = 0.007$ , Fig. S4). In addition, in both regions, there was qualitative evidence for spatial structuring with some FIV<sub>pco</sub> lineages being locally dominant (e.g., CO clade I and III in Fig. 2). However, overall FIV<sub>pco</sub> 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 FIV<sub>pco</sub> subtypes or clades (Fig. 2).

**Predictors of FIV spread are region-dependent.** We employed generalized dissimilarity models<sup>35</sup> (GDM) and maximum likelihood of population-effects<sup>36</sup> (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<sup>37</sup>. 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<sup>35</sup>. In the UB, our optimization approach revealed that none of the landscape variables (Table S3) 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



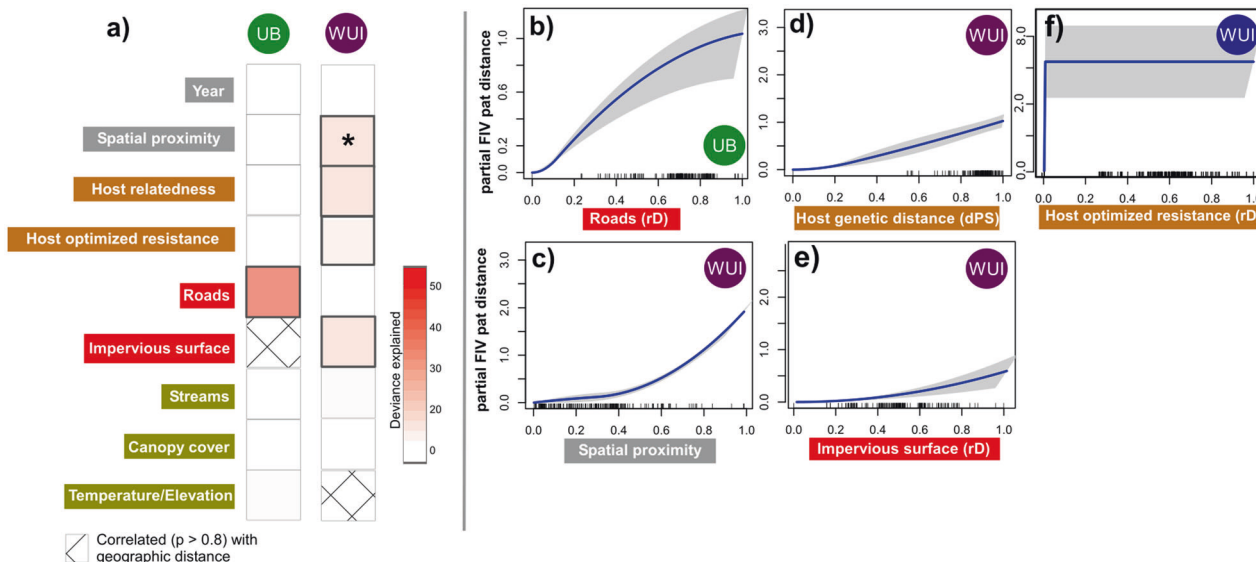
**Fig. 2 Spatially projected FIV phylogenies showing the configuration of FIV<sub>pco</sub> spread in the unbounded (UB) and wildland-urban interface (WUI) regions.** Panels **a-d** represent enlarged sections of each area. Tree tips represent the capture locations. Lines indicate FIV<sub>pco</sub> estimated branch locations, with colours of each line as well as the coloured puma symbols reflecting FIV<sub>pco</sub> subtype or lineage membership based on tree structure analysis<sup>69</sup>. The dashed circle indicates the group of individuals in the WUI in which the viral phylogeny mapped onto the host cladogram (grey boxes in Fig. 1). Grey shading shows the extent of the Denver metropolitan areas. Branches coming from outside the study area in the top panels are the branches connecting each region. See Supplementary Data 1-4 for .kml files to recreate this map. FIV<sub>pco</sub> or host genomic data could not be tested in some individuals (white boxes).

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-optimised 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 FIV<sub>pco</sub> 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<sup>38</sup>. 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, FIV<sub>pco</sub> 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 FIV<sub>pco</sub> patristic distance between individuals; neighbouring individuals

shared more phylogenetically similar FIV<sub>pco</sub> isolates (Fig. 3c). Similarly, as host relatedness decreased, so did FIV<sub>pco</sub> patristic distance (i.e., related individuals were more likely to share phylogenetically similar FIV<sub>pco</sub> isolates, Fig. 3e). A different measure of individual genetic distance (the Smouse measure<sup>39</sup>) did not alter our results. We found a similar positive relationship between FIV<sub>pco</sub> 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 FIV<sub>pco</sub> 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 FIV<sub>pco</sub> lineage dispersal velocity. Our analysis revealed that elevation tended to act as a conductance factor increasing the dispersal velocity of FIV<sub>pco</sub> 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  $>3^{40}$ ; 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, FIV<sub>pco</sub> 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.



**Fig. 3 Results from the GDM analyses of FIV<sub>pco</sub> spread.** **a** Heat map showing the deviance explained by each predictor in the unbounded (UB) and wildland-urban interface (WUI) \**p* value 0.01–0.05. Predictor labels in grey boxes = spatiotemporal predictors, predictors in orange = host genetic measures, predictors in red = associated with anthropogenic development, predictors in olive green = natural landscape features. The right panel shows partial effect plots (all other variables held at the mean) of the GDM-fitted l-splines showing the relationship between partial FIV<sub>pco</sub> patristic distance for the variables selected in the UB final model (**b**, see ‘Methods’) and WUI (**c–f**) in order of importance. Pat distance = patristic distance. Lines across the x-axis represent a rug plot showing the distribution of the data and the grey surrounding the blue line (the fit) indicates 95% confidence intervals. rD: Resistance distance. dps: Proportion of shared alleles. See Tables S6/S7 for corresponding MLPE results. See Fig. S5 for the model with host genetic resistance is substituted for optimized host resistance.

**Discussion**

Our multifaceted approach linking landscape, host genomic, and pathogen genomic data uncovered unique landscape-specific relationships with FIV<sub>pco</sub> spread, indicative of altered epidemiological dynamics associated with urban landscape structure. We found that spatial proximity was positively associated with FIV<sub>pco</sub> 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 FIV<sub>pco</sub> spread. These landscape factors mirrored the factors shaping host gene flow in the WUI<sup>26</sup>, 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 FIV<sub>pco</sub> 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 FIV<sub>pco</sub>/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<sup>26</sup>). 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 patterns<sup>22,23</sup> and foraging behaviours<sup>17</sup>. 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 dispersal<sup>23</sup> 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 UB<sup>26</sup>. There has not been an extensive evaluation of relatedness in neighbouring home-range females, but female matriline (groups of maternally-related females) are

known to occur<sup>41,42</sup>. 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 profile<sup>9</sup>.

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 roads<sup>43</sup> and rapid viral evolution potentially allowed us to detect this modest effect. There was a smaller impact of roads on host gene flow<sup>26</sup> which supports the idea that FIV phylogenetics may capture more contemporary movement patterns impossible to detect using host genetics alone<sup>10,11,44</sup>. 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 connectivity<sup>45</sup>. This could explain why we detected no effect of tree cover on FIV<sub>pco</sub> spread even though puma are known to have a preference for tree cover to disperse and hunt<sup>46,47</sup>. 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, FIV<sub>pco</sub> CO clade I). We postulate that recent arrival of FIV<sub>pco</sub> CO clade I in the UB and signature of rapid expansion<sup>19</sup> 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 FIV<sub>pco</sub> 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 divides<sup>46</sup>. 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’<sup>17,48</sup>.

FIV transmission events have been shown to be more likely to occur further from the urban edge in bobcat populations<sup>26</sup> 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<sup>49</sup>. We acknowledge that in this study we did not sample all pumas in the system nor between the two sampled regions, and that FIV<sub>pco</sub> could not always be sequenced. This could mean that we missed, for example, some FIV<sub>pco</sub> lineages that could have altered our inference about FIV<sub>pco</sub> 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 may be relevant predictors of pathogen spread in regions impacted by urban development. This may mean that the difficult task of disease control in a large apex predator (such as vaccinating against feline leukaemia virus in a puma population<sup>25</sup>) may be more tractable in bounded populations. Because juveniles set up home ranges near their parents, dispersal events important for pathogen jumps in the landscape are more constrained. Further, as host gene flow and viral spread were tightly linked in the WUI, targeting individuals where gene flow 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<sup>23,50</sup>. This sampling effort is likely to represent a large proportion of the resident puma present in both regions during the sampling period<sup>23</sup>. 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, Inc., 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.<sup>51</sup>. Isolated DNA was quantified using a Qubit 2.0 fluorometer (ThermoFisherScientific). DNA from individual puma were screened for the presence of FIV<sub>pco</sub> provirus using a specific qPCR assay as described by Lee et al.<sup>52</sup>. Using these data, we compared prevalence from both regions using a two-sample test for equality of proportions.

Full *ORFA* 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<sup>53</sup> using Geneious 7.0.6. Our *Pol* and *ORFA* datasets (GenBank accession: MN563193–MN563239) were compiled together with those sequences available in the public database Genbank, previously isolated from across the USA<sup>54</sup> (Fig. S1). Recombination was detected using RDP software V4<sup>55</sup>; parameters were set at default with linear topology. Events were determined as true if supported by three or more methods with *p* values <10e<sup>-3</sup> combined with phylogenetic support. Recombination-free datasets were used for all downstream phylogenetic analyses.

**Viral phylogenetics.** To examine the broad placement of the FIV<sub>pco</sub> isolates sampled during this study, a maximum-likelihood tree was constructed for the FIV<sub>pco</sub> dataset comprised of all isolates recovered in the USA using PhyML<sup>56</sup> with

the TN93 + G + I model and aLRT branch support; branches with <80 support were collapsed using TreeGraph2<sup>57</sup>.

We used TempEst<sup>58</sup> 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 *ORFA* sequence data. We used BEAST 1.10<sup>59</sup> with BEAGLE 3.1<sup>60</sup> to perform both discrete and continuous phylogeographic analyses<sup>61</sup> based on the concatenated *pol* and *ORFA* sequences from each subtype using an HKY substitution model (found most suitable for this smaller number of sequences using ‘smart model selection’ in PhyML<sup>62</sup>). For FIV<sub>pco</sub> CO, we performed Bayesian model selection on various model combinations comprising a strict molecular clock and an uncorrelated relaxed clock model with an underlying lognormal distribution, as well as three different coalescent models: a constant population size model, an exponential growth model and a non-parametric Bayesian skygrid model<sup>63</sup>. We also tested three relaxed random walk (RRW) models of continuous diffusion for the phylogeographic analyses and included each population (WUI and UB) as a discrete trait. All Bayesian model selection experiments were performed by (log) marginal likelihood estimation using path sampling and stepping-stone sampling<sup>64–67</sup> (see Table S9). Multiple replicates were run with different starting seeds to ensure convergence. Based on the outcome of the Bayesian model selection procedure, we presented the results obtained for the relaxed molecular clock models, the exponential population size coalescent model and the Cauchy RRW model. For FIV<sub>pco</sub> 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) reflecting 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<sup>68</sup> 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<sup>69</sup> using the default values. This analytical routine compares discrepancies between observed and idealised genealogies to identify clades under differing epidemiological or demographic processes<sup>69</sup>. As we did not have enough sequences to find meaningful structure in FIV<sub>pco</sub> WY, we limited this analysis to FIV<sub>pco</sub> CO. Any clades identified as significantly departing from the idealised genealogy were further investigated using the R package ‘phylodyn’<sup>70</sup> 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. 71).

**Host genomic data.** We genotyped 130 pumas (76 individuals from the UB and 54 individuals from the WUI) using a ddRADseq approach (see ref. 26 for sequence and bioinformatics details). From these genomic data, we quantified individual relatedness using the inverse proportion of shared alleles (Dps<sup>72</sup>) and used the resultant pairwise distance matrix in the downstream analyses. We also calculated the Smouse measure of individual genetic distance<sup>39</sup> to check if our results were sensitive to the distance measure used. We were unsuccessful in obtaining ddRAD data for seven individuals for which corresponding viral genomic data were available (see Fig. 1) and for these individuals we used the mean population relatedness value. Removing these individuals from the analyses did not qualitatively alter the results. Furthermore, we used the complete Dps dataset from all individuals and interpolated this distance across each landscape using a kriging approach. We converted this interpolated surface into a resistance raster in R and calculated resistance distances between each individual with FIV<sub>pco</sub> sequence data using a Circuitscape model, which uses circuit theory to accommodate uncertainty in the route taken<sup>73</sup>. See <https://github.com/nfj1380/ColoradoPumaFIVproject> for details.

We estimated a coalescent-based phylogeny with bootstrap support values using SVD quartets<sup>74</sup> as implemented in PAUP<sup>75</sup>. SVD quartets, which is currently considered the most robust and computationally efficient SNP-based phylogenetic estimation method<sup>37</sup> and uses a site-based approach (considering each SNP has an independent genealogy) to estimate combinations of four-taxon relationships and heuristically summarise the resulting trees into a species tree phylogeny<sup>37,74</sup>. We evaluated a maximum of 100,000 random quartets using the QFM quartet assembly method and the multispecies coalescent tree model for generating the topology, and performed 100 bootstrap replicates for assessing topological branch support for all terminal taxa.

**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 justification for each landscape variable). This included percent impervious surface (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 Reclassify and Raster Calculator tools in ArcGIS v. 10.1. We defined resistance/conductance cost values based on expert opinion as well the ‘Resistance GA’<sup>76</sup> routine in R. Resistance GA optimises 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 FIV<sub>pco</sub> spread. Mantel tests were used to screen for collinearity between surfaces and for spatial autocorrelation prior to model construction.

We calculated the proportion of uninfected to infected individuals within a 5 km buffer (estimated average distance between individuals<sup>77</sup>) of each FIV<sub>pco</sub> 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)<sup>35</sup> to quantify if host relatedness and landscape shaped FIV<sub>pco</sub> spread for FIV<sub>pco</sub> CO. GDM is a flexible non-linear regression approach that fits monotonic I-spline functions to pairwise matrix data<sup>35</sup> to describe the rate and magnitude of, in this case, FIV<sub>pco</sub> phylogenetic change. We first calculated FIV<sub>pco</sub> patristic distance (using the maximum clade credibility tree)

and converted this distance metric into a dissimilarity measure:  $ds_{FIVple} =$

$\left(\frac{d_i}{\max(d_{ij})}\right)^2$  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 FIV<sub>pco</sub> patristic distance in the form of:

$$-\ln(ds_x) = a_0 + \sum_{p=1}^n |f_p(x_{pi}) - f_p(x_{pj})|$$

where  $i$  and  $j$  are individual puma,  $a_0$  is the intercept,  $p$  is the number of covariates and  $f_p(x)$  have I-spline transformed versions of the predictors (see refs. 35,78 for further details). We used a backward elimination model selection approach and permutation tests ( $n = 99$ ) to test for significance<sup>35,79</sup>. The model with the highest deviance ( $\pm 2\%$  deviance explained) with the smallest number of predictors was reported. We performed GDM using the same predictor sets as in Trumbo et al.<sup>26</sup> for each region but in our reduced dataset (i.e., with individuals with FIV<sub>pco</sub> 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 FIV<sub>pco</sub> WY was not possible given the small sample size.

We compared our non-linear GDM models to maximum likelihood of population-estimate (MLPE) models<sup>36</sup> 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’<sup>80</sup> (see refs. 81,82 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 defines 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 points<sup>83</sup>, 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 equalled the number of occurrences of each land cover category within the 2 arcmin cells<sup>81</sup>. For each considered environmental factor, several distinct rasters were also generated by transforming original raster cell values with the following formula:  $v_t = 1 + k^*(v_o/v_{max})$ , where  $v_t$  and  $v_o$  are the transformed and original raster cell values, and  $v_{max}$  the maximum 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<sup>84</sup>. 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 #08-2004, 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/nfj1380/ColoradoPumaFIVproject>.

### Code availability

All code is available on GitHub: <https://github.com/nfj1380/ColoradoPumaFIVproject>.

Received: 12 November 2019; Accepted: 25 November 2020;

Published online: 04 January 2021

### References

- Daversa, D. R., Fenton, A., Dell, A. I., 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).
- Mazé-Guilmo, E., Blanchet, S., McCoy, K. D. & Looft, G. Host dispersal as the driver of parasite genetic structure: a paradigm lost? *Ecol. Lett.* **19**, 336–347 (2016).
- Biek, R. & Real, L. A. The landscape genetics of infectious disease emergence and spread. *Mol. Ecol.* **19**, 3515–3531 (2010).
- 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).
- 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).
- 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).
- Schwabl, P. et al. Prediction and prevention of parasitic diseases using a landscape genomics framework. *Trends Parasitol.* **33**, 264–275 (2017).
- 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).
- Gijsbers, E. F. et al. Low level of HIV-1 evolution after transmission from mother to child. *Sci. Rep.* **4**, 4650–4655 (2014).
- 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).
- Fountain-Jones, N. M. et al. Urban landscapes can change virus gene flow and evolution in a fragmentation-sensitive carnivore. *Mol. Ecol.* **26**, 6487–6498 (2017).
- Brearely, G. et al. Wildlife disease prevalence in human-modified landscapes. *Biol. Rev. Camb. Philos. Soc.* **88**, 427–442 (2013).

13. McDonald, R. L., 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., Hohnowski, C. E., Carter, N. H. & Brashares, J. S. The influence of human disturbance on wildlife nocturnality. *Science* **360**, 1232–1235 (2018).
16. Riley, S. P. D. 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. *Ecosphere* **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 wildlife diseases. *Trends Ecol. Evol.* **22**, 95–102 (2007).
25. Cunningham, M. W. et al. Epizootiology and management of feline leukemia virus in the Florida puma. *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 endemic 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 fitness and pathogen susceptibility in a natural feline host. *Can. J. Zool.* **84**, 365–373 (2006).
31. Reynolds, J. J. H. et al. Feline immunodeficiency virus in puma: estimation of force of infection reveals insights into transmission. *Ecol. Evol.* ece3.5584, <https://doi.org/10.1002/ece3.5584> (2019).
32. Fountain-Jones, N. M. 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).
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, G., 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., Rothery, P. & Raybould, A. F. Condence limits for regression relationships between distance matrices: estimating gene flow with distance. *J. Agric. Biol. Environ. Stat.* <https://doi.org/10.1198/108571102320> (2002).
37. Chou, J. et al. A comparative study of SVDquartets and other coalescent-based species tree estimation methods. *BMC Genomics* **16**, S2 (2015).
38. Shirk, A. J., Landguth, E. L. & Cushman, S. A. A comparison of regression methods for model selection in individual-based landscape genetic analysis. *Mol. Ecol. Resour.* **18**, 55–67 (2018).
39. Smouse, P. E. & Peakall, R. Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity (Edinb.)* **82**, 561–573 (1999).
40. Kass, R. E. & Raftery, A. E. Bayes factors. *J. Am. Stat. Assoc.* **90**, 773–795 (1995).
41. Logan, K. A. & Swenor, L. L. *Desert Puma: Evolutionary Ecology and Conservation of an Enduring Carnivore* (Island Press, 2001).
42. Biek, R. et al. Genetic consequences of sex-biased dispersal in a solitary carnivore: yellowstone cougars. *Biol. Lett.* **2**, 312–315 (2006).
43. Dickson, B. G., Jenness, J. S. & Beier, P. Influence of vegetation, topography, and roads on cougar movement in Southern California. *J. Wildl. Manag.* **69**, 264–276 (2005).
44. Kerr, T. J. et al. Viruses as indicators of contemporary host dispersal and phylogeography: an example of feline immunodeficiency virus (FIV Ple) in free-ranging African lion (*Panthera leo*). *J. Evol. Biol.* <https://doi.org/10.1111/jeb.13348> (2018).
45. Epps, C. W. & Keyghobadi, N. Landscape genetics in a changing world: disentangling historical and contemporary influences and inferring change. *Mol. Ecol.* **24**, 6021–6040 (2015).
46. Hornocker, M. G. & Negri, S. *Cougar: Ecology and Conservation* (University of Chicago Press, 2010).
47. Swenor, L. L., Logan, K. A. & Hornocker, M. G. Cougar dispersal patterns, metapopulation dynamics, and conservation. *Conserv. Biol.* **14**, 798–808 (2000).
48. Suraci, J. P., Clinchy, M., Zanette, L. Y. & Wilmers, C. C. Fear of humans as apex predators has landscape-scale impacts from mountain lions to mice. *Ecol. Lett.* e13344 <https://doi.org/10.1111/ele.13344> (2019).
49. Tian, H. et al. Transmission dynamics of re-emerging rabies in domestic dogs of rural China. *PLOS Pathog.* **14**, e1007392 (2018).
50. Carver, S. et al. Pathogen exposure varies widely among sympatric populations of wild and domestic felids across the United States. *Ecol. Appl.* **26**, 367–381 (2016).
51. Di Pietro, F., Ortenzi, F., Tilio, M., Concetti, F. & Napolioni, V. Genomic DNA extraction from whole blood stored from 15- to 30-years at  $-20^{\circ}\text{C}$  by rapid phenol–chloroform protocol: a useful tool for genetic epidemiology studies. *Mol. Cell. Probes* **25**, 44–48 (2011).
52. Lee, J. S. et al. Targeted enrichment for pathogen detection and characterization in three felid species. *J. Clin. Microbiol.* **55**, 1658–1670 (2017).
53. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
54. Lee, J. S. et al. Evolution of puma lentivirus in bobcats (*Lynx rufus*) and mountain lions (*Puma concolor*) in North America. *J. Virol.* **88**, 7727–7737 (2014).
55. Martin, D. P., Murrell, B., Golden, M., Khoosal, A. & Muhire, B. RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol.* **1**, vev003 (2015).
56. Guindon, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
57. Stöver, B. C. & Müller, K. F. TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinforma.* **11**, 7 (2010).
58. Rambaut, A., Lam, T. T., Max Carvalho, L. & Pybus, O. G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2**, vew007 (2016).
59. Suchard, M. A. et al. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4** (2018).
60. Ayres, D. L. et al. BEAGLE 3: Improved performance, scaling, and usability for a high-performance computing library for statistical phylogenetics. *Syst. Biol.* **68**, 1052–1061, <https://doi.org/10.1093/sysbio/syz2020> (2019).
61. Lemey, P., Rambaut, A., Welch, J. J. & Suchard, M. A. Phylogeography takes a relaxed random walk in continuous space and time. *Mol. Biol. Evol.* **27**, 1877–1885 (2010).
62. Lefort, V., Longueville, J.-E. & Gascuel, O. SMS: Smart model selection in PhyML. *Mol. Biol. Evol.* **34**, 2422–2424 (2017).
63. Gill, M. S. et al. Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. *Mol. Biol. Evol.* **30**, 713–724 (2013).
64. Baele, G. et al. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Mol. Biol. Evol.* **29**, 2157–2167 (2012).
65. Lartillot, N. & Philippe, H. Computing Bayes factors using thermodynamic integration. *Syst. Biol.* **55**, 195–207 (2006).
66. 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).
67. Baele, G., Lemey, P. & Vansteelandt, S. Make the most of your samples: Bayes factor estimators for high-dimensional models of sequence evolution. *BMC Bioinforma.* **14**, 85 (2013).
68. Rambaut, 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).
69. 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/syaa009> (2019).
70. Karcher, M. D., Palacios, J. A., Lan, S. & Minin, V. N. phylodyn: an R package for phylodynamic simulation and inference. *Mol. Ecol. Resour.* **17**, 96–100 (2017).
71. Karcher, M. D., Palacios, J. A., Bedford, T., Suchard, M. A. & Minin, V. N. Quantifying and mitigating the effect of preferential sampling on phylodynamic inference. *PLoS Comput. Biol.* **12**, e1004789 (2016).
72. Bowcock, A. M. et al. High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368**, 455–457 (1994).
73. McRae, B. H. Isolation by resistance. *Evolution* **60**, 1551–1561 (2006).



74. Chifman, J. & Kubatko, L. Quartet inference from SNP data under the coalescent model. *Bioinformatics* **30**, 3317–3324 (2014).
75. Swofford, D. L. *PAUP\* Phylogenetic Analysis Using Parsimony (\*and Other Methods)*. Version 4.0b10 (Sinauer Associates, 2002).
76. Peterman, W. E. ResistanceGA: an R package for the optimization of resistance surfaces using genetic algorithms. *Methods Ecol. Evol.* **9**, 1638–1647 (2018).
77. Pierce, B. M., Bleich, V. C. & Bowyer, R. T. Social organization of mountain lions: does a land-tenure system regulate population size? *Ecology* **81**, 1533–1543 (2000).
78. Fitzpatrick, M. C. & Keller, S. R. Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecol. Lett.* **18**, 1–16 (2015).
79. Fitzpatrick, M. C. et al. Forecasting the future of biodiversity: a test of single- and multi-species models for ants in North America. *Ecography* **34**, 836–847 (2011).
80. Dellicour, S., Rose, R., Faria, N. R., Lemey, P. & Pybus, O. G. SERAPHIM: studying environmental rasters and phylogenetically informed movements. *Bioinformatics* **32**, 3204–3206 (2016).
81. Dellicour, S. et al. Explaining the geographic spread of emerging epidemics: a framework for comparing viral phylogenies and environmental landscape data. *BMC Bioinforma.* **17**, 82–94 (2016).
82. Laenen, L. et al. Spatio-temporal analysis of Nova virus, a divergent haantavirus circulating in the European mole in Belgium. *Mol. Ecol.* **25**, 5994–6008 (2016).
83. Dijkstra, E. W. A note on two problems in connexion with graphs. *Numer. Math.* **1**, 269–271 (1959).
84. Dellicour, S. et al. Using viral gene sequences to compare and explain the heterogeneous spatial dynamics of virus epidemics. *Mol. Biol. Evol.* **34**, 2563–2571 (2017).

## 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.

## 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.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021