mutations by SIAFE and directed evolution. Meanwhile, the amino acid sequence identity of evMBL8 with IMP-1 (25%) was almost double that of GlyII with IMP-1 (13%). As expected from the design strategy, more than 60% (49 amino acids) of mutations were concentrated in the catalytic and substrate-binding regions. In addition, many originally designed conserved residues in functional loops were changed during evolution. Nonetheless, the residues responsible for metal coordination (H66, H68, and H131 for metal 1; D70, C155, and H194 for metal 2) were not changed despite extensive mutagenesis. Replacement of any of these residues in evMBL8 with Ala resulted in loss of catalytic activity (*19*). Mutations C71H and C155D, which were attempts to restore GlyII-like metal coordination (H59 and D134) in evMBL8, also resulted in complete loss of catalytic activity (*19*). Thus, as evMBL8 evolved to have β-lactamase activity, it retained the designed metal coordination of IMP-1 that is essential for catalysis.

To further confirm the evolution of β-lactamase activity from the GlyII scaffold, evMBL8 was characterized in terms of in vivo biological activity, kinetic constants, and metal content. When *E. coli* cells expressing the evMBL8 were grown in liquid medium, distinct cell growth was observed up to cefotaxime concentrations of 1.0 μg/ml (Fig. 4A). A further increase in cefotaxime concentration (>2.0 μg/ml) caused a serious inhibition of cell growth due to cell lysis. On the other hand, cells expressing GlyII showed growth inhibition even at cefotaxime concentrations of 0.02 μg/ml. To study enzyme kinetics, evMBL8 was expressed as a fusion with maltose-binding protein (MBP) to improve stability (fig. S4). evMBL8 exhibited a typical saturation profile for cefotaxime, and k_{cat}^{app} and K_m^{app} were calculated to be 0.042 s^{-1} and 229 μM, respectively, from the reciprocal plot (Fig. 4B and Table 1). These values are lower by a factor of 150 and higher by a factor of 30 than the k_{cat} and K_m values of wild-type IMP-1, respectively, with the difference in K_m providing evidence against contamination by IMP-1 in conjunction with their relative activities for various substrates (table S2). Consequently, overall catalytic efficiency $(k_{cat}/K_m)^{app}$ of evMBL8 for cefotaxime was estimated to be 1.8×10^2 $M^{-1}s^{-1}$, which is about 3 to 4 orders of magnitude lower than that of wild-type IMP-1. Nonetheless, evMBL8 was active enough to provide *E. coli* cells a resistance against cefotaxime that is higher by a factor of 100 compared with cells having no evMBL8 (Fig. 4A). To check whether evMBL8 was still able to bind metal ions in its fused form with MBP, the total metal content of the enzyme was measured. Mutation H59C of GlyII was reported to cause a loss of zinc while maintaining iron content at a normal level, which suggests that H59C severely impaired the metal-binding site for zinc (*15*). However, evMBL8 with mutation H59C was found to bind both zinc (1.6 mol) and iron (0.4 mol) essential for hydrolytic reaction at levels comparable to GlyII (Table 1). It is likely that the disruption of metal coordination was reversed by other mutations during the directed evolution process.

Molecular modeling of evMBL8 with either an MBL family member (CcrA) or GlyII as the template gave an evMBL8 structure that is similar to that of the target IMP-1 (fig. S5A) but shows a distinct active site architecture compared with the template GlyII (fig. S5B). The model of the evMBL8-cefotaximne complex shows well-organized metal coordination (H66, H68, and H131 for metal 1; D70, C155, and H194 for metal 2) at the bottom of the active site (fig. S5C), which is consistent with the mutation study on putative metal ligands described above. Within the active site, loop 2 (T38-L48) and loop 6 (F156-P168) constitute two walls of the substrate-binding pocket, and loop 1 (T10-G22) forms the ceiling. Thus, we suggest that evMBL8 has acquired a new active site architecture with well-defined metal coordination and a substrate-binding pocket for new catalytic activity through the SIAFE and directed evolution process.

The design strategy presented here enabled the conversion of an enzyme in the metallohydrolase superfamily into a new family member with a different catalytic function, providing experimental support for the divergent evolution of mechanistically diverse family enzymes. We hope that the developed process can be extended to other scaffolds and create a larger variety of catalytic lineages performing diverse reactions and perhaps even reactions not found in nature.

References and Notes

1. S. J. Benkovic, S. Hammes-Schiffer, *Science* **301**, 1196 (2003).
2. F. H. Arnold, *Nature* **409**, 253 (2001).
3. A. Schmid *et al.*, *Nature* **409**, 258 (2001).
4. A. C. Joerger, S. Mayer, A. R. Fersht, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5694 (2003).
5. H. S. Park, K. H. Oh, H. S. Kim, *Methods Enzymol.* **388**, 187 (2004).
6. Y. H. Cheon *et al.*, *Biochemistry* **43**, 7413 (2004).
7. L. Tang *et al.*, *Biochemistry* **44**, 6609 (2005).
8. M. A. Dwyer, L. L. Looger, H. W. Hellinga, *Science* **304**, 1967 (2004).
9. S. Leopoldseder, J. Claren, C. Jurgens, R. Sterner, *J. Mol. Biol.* **337**, 871 (2004).
10. A. Aharoni *et al.*, *Nat. Genet.* **37**, 73 (2005).
11. H. Daiyasu, K. Osaka, Y. Ishino, H. Toh, *FEBS Lett.* **503**, 1 (2001).
12. A. D. Cameron, M. Ridderström, B. Olin, B. Mannervik, *Struct. Fold. Des.* **7**, 1067 (1999).
13. Z. Wang, W. Fast, A. M. Valentine, S. J. Benkovic, *Curr. Opin. Chem. Biol.* **3**, 614 (1999).
14. N. O. Concha *et al.*, *Biochemistry* **39**, 4288 (2000).
15. T. M. Zang *et al.*, *J. Biol. Chem.* **276**, 4788 (2001).
16. Z. Wang, W. Fast, S. J. Benkovic, *Biochemistry* **38**, 10013 (1999).
17. M. P. Yanchak, R. A. Taylor, M. W. Crowder, *Biochemistry* **39**, 11330 (2000).
18. S. D. Scrofani *et al.*, *Biochemistry* **38**, 14507 (1999).
19. Materials and methods are available as supporting material on *Science* Online.

20. We thank R. Sterner, M. Meyer, and F. Arnold for helpful comments, and J. H. Kim and S. C. Lee for technical assistance. Supported by National Research Laboratory program and Microbial Genomics and Application Center of Ministry of Science and Technology, Nano-Medical Project of Ministry of Health and Welfare, R & D program of Fusion Strategies for Advanced Technologies of Ministry of Commerce, Industry and Energy, and Brain Korea 21 of Ministry of Education and Human Resources Development, Korea.

Supporting Online Material

www.sciencemag.org/cgi/content/full/311/5760/535/DC1
Materials and Methods
SOM Text
Figs. S1 to S6
Tables S1 and S2
References

16 August 2005; accepted 7 December 2005
10.1126/science.1118953

A Virus Reveals Population Structure and Recent Demographic History of Its Carnivore Host

Roman Biek,[1]*† Alexei J. Drummond,[3] Mary Poss[1,2]

Directly transmitted parasites often provide substantial information about the temporal and spatial characteristics of host-to-host contact. Here, we demonstrate that a fast-evolving virus (feline immunodeficiency virus, FIV) can reveal details of the contemporary population structure and recent demographic history of its natural wildlife host (*Puma concolor*) that were not apparent from host genetic data and would be impossible to obtain by other means. We suggest that rapidly evolving pathogens may provide a complementary tool for studying population dynamics of their hosts in "shallow" time.

The genetic population structure of human pathogens often reflects known patterns of human migration (*1–4*). Moreover, the rapid evolutionary rate of many viral parasites means that this information can manifest itself over months or years (*5*). With HIV, for example, the origin and demographic history of epidemics have been determined retrospectively

from genetic data (*6*–*8*). Especially because of their high temporal resolution, viruses hold great potential as genetic tags of their hosts. Yet, this exciting potential remains largely unexplored for the study of wildlife populations where such an approach could be used to address questions in numerous disciplines, ranging from epidemiology to conservation.

Many wild feline species are natural hosts to species-specific types of feline immunodeficiency virus (FIV), analogous to simian immunodeficiency viruses (SIVs) in wild primates (*9*, *10*). Cougars (*Puma concolor*), a large New World feline species, are commonly infected by their own type of FIV (FIV_{Pco}) (*11*, *12*), apparently without suffering major fitness consequences (*13*). Transmission, which requires direct contact between individuals, occurs vertically and horizontally (*14*) and causes persistent infections. Exchange of virus with other cat species in the wild either does not occur or must be extremely rare (*10*). Importantly, FIV_{Pco} measurably evolves within years (*14*) and thus has the potential to be informative about virus and host dynamics over extremely short periods.

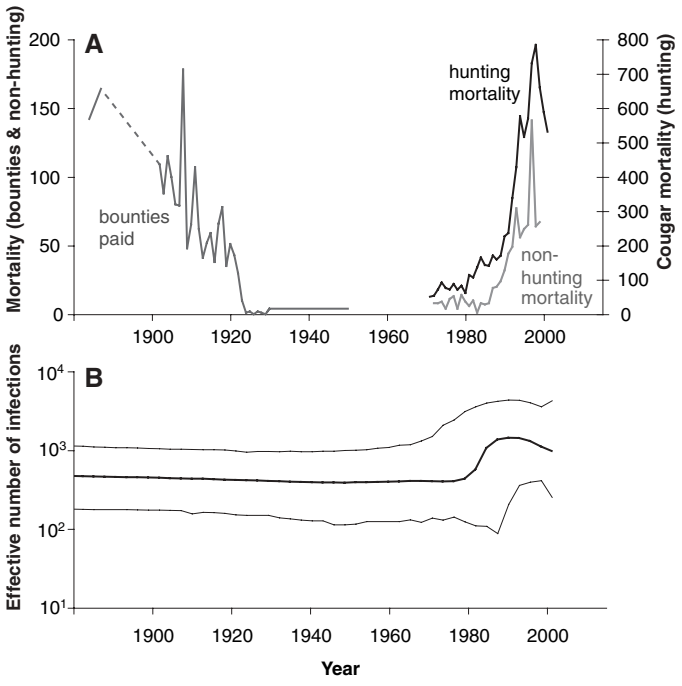


Fig. 1. Demographic trends for cougars and FIV_{Pco}, 1880 to 2001. (**A**) Cougar population indices for Montana. Dashed line indicates years with incomplete records. Nonhunting mortality represents cougars that died primarily as a result of animal control and road mortality. Numbers are taken from Riley (*17*) and from public records by Montana Fish Wildlife and Parks. (**B**) Median effective number of FIV_{Pco} infections (thin lines indicate 95% highest posterior density interval) as estimated from virus sequence data in program BEAST (*24*) using a Bayesian skyline plot technique (*22*). Estimates for years before 1880 remained essentially unchanged and are omitted.

As in most parts of North America, cougar populations in the northern Rocky Mountains underwent drastic declines during the early 20th century as a consequence of human persecution and depleted prey populations (*15*, *16*). In Montana, for example, the number of cougar carcasses annually submitted for bounty payments plummeted between 1908 and 1929 from more than 170 to zero and stayed extremely low for decades thereafter despite the bounty incentive remaining in place (*17*) (Fig. 1A). Populations in neighboring states and provinces apparently experienced a similar fate (*16*). More recently, cougar populations throughout the region have shown signs of recovery following legislation that limits the take of both cougars and their ungulate prey (*15*) (Fig. 1A). This strong demographic signal of decline followed by rapid growth offered a unique opportunity to test whether the dynamic history of cougars in Montana in the recent past could be discerned from the genetic structure of its obligate viral parasite.

We collected samples from 352 cougars across western Montana, as well as adjacent areas [Wyoming, British Columbia, and Alberta (*18*); table S1, Fig. 2A), and conducted genetic analysis of 11 host microsatellite loci using different individual-based methods [i.e., without relying on predefined populations (*18*)]. Despite samples being collected over distances exceeding 1000 km, results revealed little population structure. Although a southern and northern group could be distinguished, both overlapped substantially, indicating genetic admixture (fig. S2). This result was confirmed by a weak but significant isolation-by-distance pattern of cougar relatedness (*18*). Low genetic differentiation has also been noted among cougar populations in neighboring areas (*19*, *20*). Evidently, low population densities in the recent past, possibly associated with reduced movement, have not resulted in pronounced genetic structure for cougars in the northern Rocky Mountains.

More than one-quarter of the cougars we sampled were infected with FIV_{Pco} (98 or 28%; Fig. 2A, table S1), consistent with prevalence levels found in previous studies (*11*, *14*). Phylogenetic analysis based on combined sequence data (~1400 base pairs) from two viral genes identified two major virus clades with >20% divergence. Within clades, a total of eight FIV_{Pco} lineages (L1 to L8) could be distinguished that were up to 5% divergent and had high bootstrap and/or posterior support (Fig. 2B). Seven cougars yielded sequences of more than one lineage, indicative of coinfection or potentially circulating recombinants (*18*).

Consistent with cougar host genetics, FIV lineages fell into a northern (L1 to L4) and a southern group (L5 to L8) (Fig. 2A). However, in contrast to host genetics, spatial structure of viral populations was much more pronounced. All lineages exhibited more or less contiguous ranges, and these ranges were often spatially exclusive (Fig. 2A) such as for northern (L1 to L4) versus southern lineages (L5 to L8). Furthermore, individual lineages were restricted to smaller geographic areas. For example, L1 dominated the central part of our study area, whereas L3, L4, L5, and L8 were found only in small areas at the periphery. Lineage L2 was exceptional in that it co-occurred with all other lineages even though it was found only in 14% of infected cougars. All seven individuals yielding virus sequences from more than one lineage were found where ranges of these lineages overlapped (Fig. 2A).

The lack of spatial admixture of most virus lineages was clearly at odds with the microsatellite results because high levels of cougar movement should tend to quickly randomize virus distributions. Therefore, we hypothesized that the current virus pattern may be a direct consequence of previously low cougar population sizes and thus a rather recent, and potentially transient, phenomenon. To determine how long ago the most recent common ancestor (MRCA) of each of the FIV lineages had existed, we estimated evolutionary rates for the two FIV_{Pco} genes, *pol* and *env*, on the basis of temporally spaced sampling of cougars. Both genes were found to evolve at a rate of about 4×10^{-4} per site per year [*pol*: 2.9 (95% highest posterior density interval, HPD: 1.2 to 4.8) $\times 10^{-4}$; *env*: 4.0 (HPD: 1.6 to 6.3) $\times 10^{-4}$], slightly lower than but generally consistent with previous estimates from a different cougar data set (*14*).

On the basis of the estimated rate, diversification of all eight viral lineages had occurred within the last 20 to 80 years (Fig. 3). Because this interval coincides with the period of low cougar densities (Fig. 1A), it seems likely that the FIV_{Pco} lineages currently found are survivors of a demographic bottleneck that must have resulted in frequent local extinction and severe spatial restriction of lineages. Moreover, the time elapsed since the MRCA was posi-

[1]Wildlife Biology Program, [2]Division of Biological Sciences, University of Montana, Missoula, MT 59812, USA. [3]Department of Computer Science, University of Auckland, Auckland, New Zealand.

*Present address: Department of Biology, Emory University, Atlanta, GA 30322, USA.

†To whom correspondence should be addressed. E-mail: rbiek@emory.edu

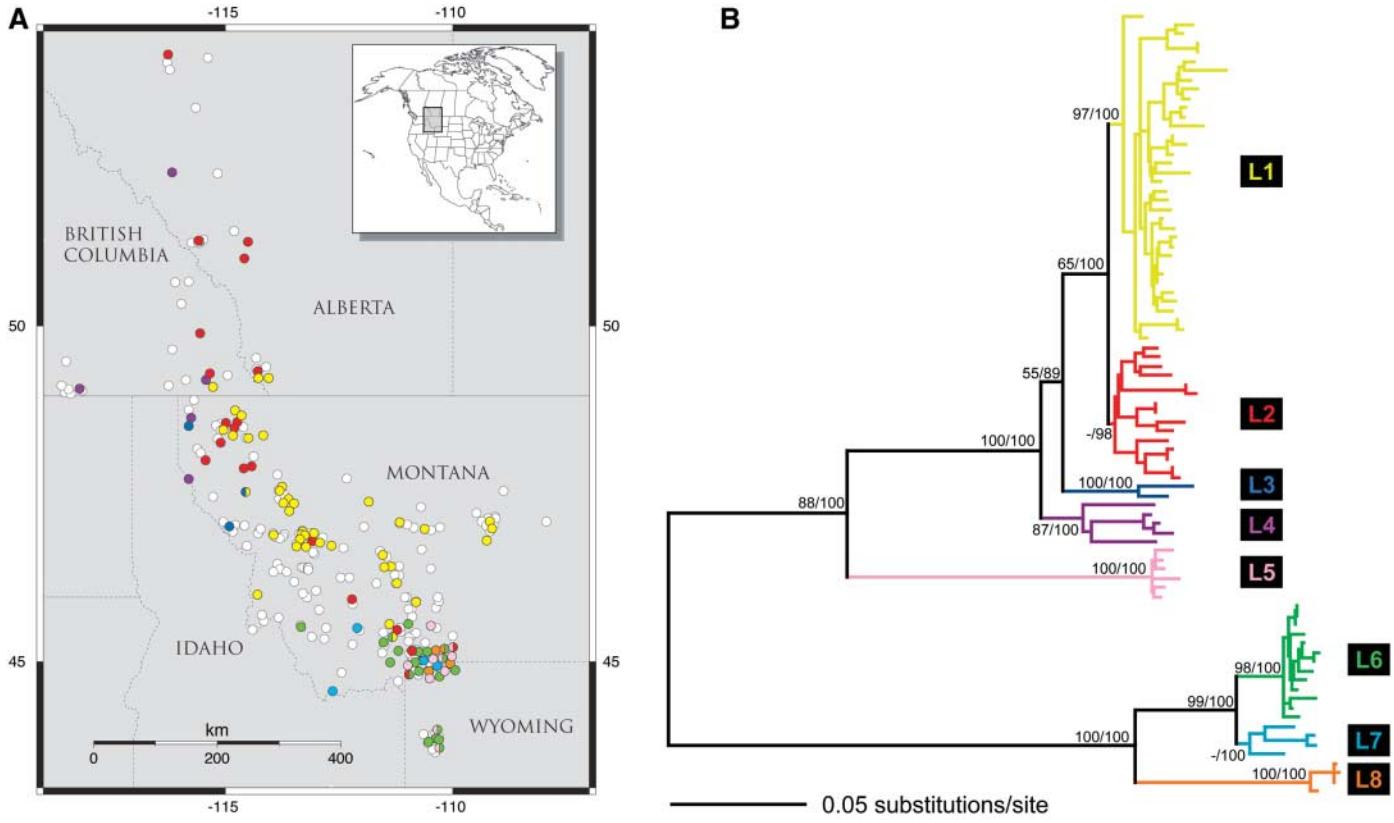


Fig. 2. FIV_{Pco} lineages and their spatial distribution. (**A**) Origin of FIV_{Pco}-positive [colors as in (B)] and -negative samples (white circles). The map was generated using Online Map Creator (*25*). The area covered by samples in Montana roughly reflects distribution of cougar habitat (forest). (**B**) Phylogeny of FIV_{Pco} in the northern Rocky Mountains constructed from concatenated *pol* and *env* sequences (n = 85) and designation of eight viral lineages based on >5% divergence. Labeled nodes represent bootstrap support based on 1000 neighbor-joining trees generated from maximum-likelihood distances estimated from the original likelihood model and posterior proportions from the Bayesian analysis. Values for within-lineage nodes are omitted for clarity.

tively correlated with the distribution of virus lineages, in that samples of more recently diverged lineages exhibited smaller maximum geographic distances (P = 0.012) (Fig. 3A). This suggests that, concurrent with cougar population recovery, surviving lineages have consistently expanded their range. This scenario of spatial expansion was supported by a separate analysis in which the geographic distance between the sampled viruses and their inferred ancestors was regressed against the corresponding time separating the sample from the origin of its lineage. This regression of geographic distances against times revealed an estimated mean viral diffusion rate of 3.7 km per year (Fig. 3B). Interestingly, the recovery of cougar populations in Alberta is thought to have started earlier than it did further south, possibly during the 1930s (*21*). This may explain the particularly wide distribution of lineage L2, which is the dominant type in Alberta today and which may have benefited, in terms of spread, from the earlier expansion of cougars in that area (Fig. 2A).

If the population dynamics of cougars had affected the current spatial distribution of FIV_{Pco}, it should also have had a discernable effect on virus demography. We estimated changes in the effective number of infections through time from the distribution of coalescent events in the genealogy using a Bayesian skyline plot (*22*). Results support an exponential increase in the absolute number of infections after 1980, i.e., during the demographic recovery of cougars in Montana (Fig. 1B). The maximum number of infections, as well as indexes of cougar abundance, was reached during the late 1990s (Fig. 1, A and B). Although this result fits expectations in terms of correlated demographic trends for both cougars and virus, evidence of a decline in infections before 1950 was weak (Fig. 1B). We suspect that this inability to detect a signature of decline during the early 20th century was due to the scarcity of coalescent events (and hence data points) in our phylogeny for this time period.

In principle, an increase in FIV_{Pco} infections could reflect elevated prevalence rather than changes in host population size; however, we believe that this alone is unlikely to explain the observed trend. For example, data collected from cougars in and around Yellowstone National Park between 1990 and 2004 show that the proportion of infected individuals remained relatively stable (~20%) throughout this period (*13*). It is important to stress, however, that our results only allow us to infer a simultaneous increase in virus and cougar host qualitatively, because reliable estimates of cougar population sizes do not exist and the proportion of infected hosts through time ultimately remains unknown.

In conclusion, we have shown that FIV_{Pco} genetics are highly informative about the recent history of its cougar host, whereas the same information was either not apparent or could not have been inferred from host genetic data. Given that lineages appear to be spatially expanding, we predict that the fine-scale spatial structure currently characterizing FIV_{Pco} will likely disappear in the near future, assuming cougar population sizes and movement remain sufficiently high. Maintaining functional populations and natural levels of gene flow remain important priorities for the conservation of wildlife populations. Monitoring the distribution and diversity of FIV_{Pco} in cougars may therefore represent a future tool for wildlife managers to ensure that these goals are met.

More generally, rapid evolutionary rates and host specificity are traits not limited to FIV_{Pco} but are typical of many parasites. Although our knowledge of parasite diversity in nondomestic species is still limited, the approaches described here should be applicable to a wide range of species and systems (*23*), provided that the ecology of interaction between parasite and host is sufficiently well understood. Natural ecosystems and populations are currently undergoing changes

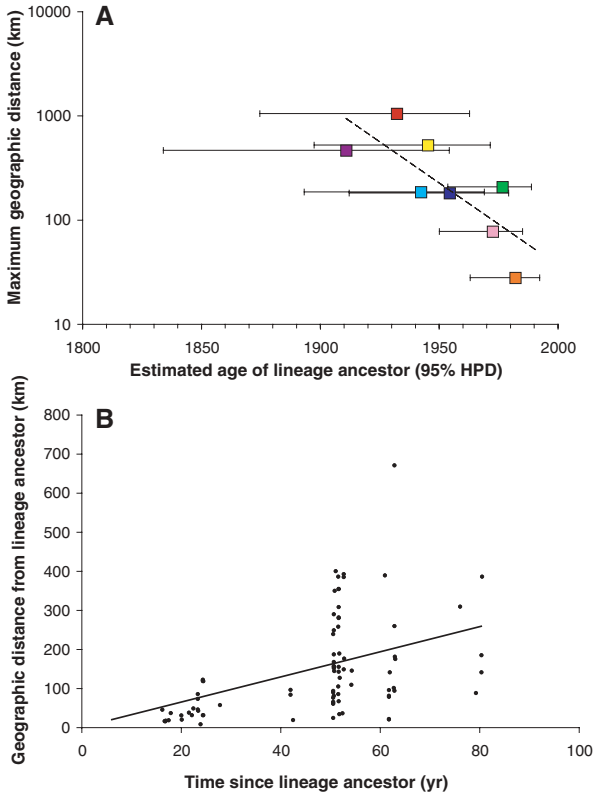


Fig. 3. Space-time relationships of virus distributions. (**A**) Median age of common ancestor (95% highest posterior density interval; HPD) for eight FIV_{Pco} lineages is correlated with their spatial range (expressed as maximum pairwise geographic distance) [$R^2 = 0.61$, $P = 0.023$; log(distance) = 0.0182(years before present) + 1.4847]. (**B**) Spatial diffusion of viruses over time. Temporal divergence for each sequence was calculated as number of years elapsed between the median age of lineage ancestor [as in (A)] and sampling date. Similarly, spatial divergence was calculated relative to the putative location of the respective lineage ancestor determined through ancestral reconstruction in COMPARE (*26*). Regression parameter estimates were obtained by comparative analysis, also using COMPARE (*26*), and excluded zero for slope ($\beta_1 = 3.74$; confidence interval: 1.04 to 6.44) but not intercept ($\beta_0 = -43.41$; confidence interval: −188.67 to 101.85); therefore, the regression line shown is based on a model without intercept. See supporting online material for details.

at unprecedented rates owing to human activities. Molecular markers with an equally short temporal resolution will be of increasing value to researchers trying to understand and mitigate these changes.

References and Notes

1. T. Wirth, A. Meyer, M. Achtman, *Mol. Ecol.* **14**, 3289 (2005).
2. D. Falush *et al.*, *Science* **299**, 1582 (2003).
3. E. C. Holmes, *Mol. Ecol.* **13**, 745 (2004).
4. C. Sugimoto, T. Kitamura, J. Guo, *Proc. Natl. Acad. Sci. U.S.A* **94**, 9191 (1997).
5. A. J. Drummond, O. G. Pybus, A. Rambaut, R. Forsberg, A. G. Rodrigo, *Trends Ecol. Evol.* **18**, 481 (2003).
6. V. V. Lukashov, J. Goudsmit, *J. Mol. Evol.* **54**, 680 (2002).
7. P. Lemey *et al.*, *Proc. Natl. Acad. Sci. U.S.A* **100**, 6588 (2003).
8. B. Korber *et al.*, *Science* **288**, 1789 (2000).
9. R. A. Olmsted *et al.*, *J. Virol.* **66**, 6008 (1992).
10. J. L. Troyer *et al.*, *J. Virol.* **79**, 8282 (2005).
11. M. A. Carpenter *et al.*, *J. Virol.* **70**, 6682 (1996).
12. R. J. Langley *et al.*, *Virology* **202**, 853 (1994).
13. R. Biek, thesis, University of Montana (2004).
14. R. Biek *et al.*, *J. Virol.* **77**, 9578 (2003).
15. K. A. Logan, L. L. Sweanor, *Desert Puma: Evolutionary Ecology and Conservation of an Enduring Carnivore* (Island Press, Washington, DC, 2001).
16. R. M. Nowak, "The cougar in the United States and Canada" (U.S. Department of Interior Fish and Wildlife Service and New York Zoological Society, 1976).
17. S. J. Riley, thesis, Cornell University (1998).
18. Materials and methods are available as supporting material on *Science* Online.
19. C. R. J. Anderson, F. G. Lindzey, D. B. McDonald, *J. Mammal.* **85**, 1207 (2004).
20. E. A. Sinclair *et al.*, *Anim. Conserv.* **4**, 257 (2001).
21. M. G. Jalkotzy, "Management plan for cougar in Alberta" (Alberta Forestry, Land and Wildlife, Fish and Wildlife Division, 1992).
22. A. J. Drummond, A. Rambaut, B. Shapiro, O. G. Pybus, *Mol. Biol. Evol.* **22**, 1185 (2005).
23. N. K. Whiteman, P. G. Parker, *Anim. Conserv.* **8**, 175 (2005).
24. A. J. Drummond, A. Rambaut, BEAST, version 1.2; available at http://evolve.zoo.ox.ac.uk/beast/ (2003).
25. Online Map Creator available at www.aquarius.geomar.de/omc/.
26. E. P. Martins, Computer programs for the statistical analysis of comparative data. Distributed by the author at http://compare.bio.indiana.edu/ (2004).

27. For help collecting cougar samples, we thank T. K. Ruth, K. Murphy, R. DeSimone, R. Grey, C. Chetkiewicz. T. Shury, A. Kortello, H. Schwantje, M. Hall, Montana Fish Wildlife and Parks, Montana State Houndsmen Association, and Alberta Tree Hound Association. See supporting online material for more extensive acknowledgements. Research was funded by the Wilburforce Foundation, the Yellowstone to Yukon Conservation Initiative, and the NSF. We thank S. Painter, K. Pilgrim, K. Samuel, and N. Akamine for lab assistance and L. A. Real and M. K. Schwartz for comments. Sequences have been submitted to GenBank under accession numbers DQ106985 to DQ107160.

Supporting Online Material

www.sciencemag.org/cgi/content/full/311/5760/538/DC1
Materials and Methods
SOM Text
Figs. S1 and S2
Table S1
References

14 October 2005; accepted 21 December 2005
10.1126/science.1121360

Supporting Online Material for

A Virus Reveals Population Structure and Recent Demographic History of Its Carnivore Host

Roman Biek,* Alexei J. Drummond, Mary Poss

*To whom correspondence should be addressed. E-mail: rbiek@emory.edu

Published 27 January 2006, *Science* **311**, 538 (2006)
DOI: 10.1126/science.1121360

This PDF file includes:

Supporting Online Material

Materials and Methods

Cougar sampling and DNA extraction

We extracted DNA from blood or tissue samples (lymph nodes or salivary glands) from 352 cougars, caught for other research purposes or killed by hunters (Supporting Table S1). Fresh blood samples were taken from live animals and small quantities of blood on Isocode filter paper (Schleicher and Schuell, Keene, NH), lymph nodes, or salivary glands were collected from cougar carcasses. Precise sampling location was recorded for all live-caught and most hunter-killed individuals; for the remaining cases the center point of the hunting district of origin was taken as the sampling location. DNA extraction from blood or tissue was done according to standard protocols (e.g. *S1*). For extraction from Isocode filter paper, the entire bloodstained area was cut out with a razorblade (discarded after use), cut into smaller pieces, and transferred to a 1.5 ml Eppendorf tube. After immersing the pieces completely in water, the tube was put at 95°C for 1h, pulse-vortexing 30 times after 30 and 60min. Paper and elution were transferred to Costar filters and spun for 3min at 14,000rpm. The filtrate was brought to 100µM TRIS by adding 1M stock and treated with a Qiagen DNA extraction kit, with amounts adjusted to the volume of the filtrate. Extracted DNA was eluted in 70µl of 100µM TRIS.

Microsatellite amplification

After excluding kittens and other known relatives from the data set, 273 cougars were genotyped at eleven microsatellite loci: Fca30, Fca35, Fca57, Fca77, Fca90, Fca96, Fca132, Fca176, Fca391, Fca559 (*S2*), and Lc109 (*S3*). Three combinations of two

primer sets (Fca90/Fca391; Fca35/Fca132, Fca77/Fca559) were set up as multiplex PCR reactions, the remaining loci were amplified individually. PCR reactions contained up to 200 ng genomic DNA, 10 mM Tris–HCl (pH8.3), 2.0 mM MgCl2, 20 mM of each dNTP, 0.2 mM of each primer (10nM for labeled primer) and 0.8 U Taq DNA polymerase and water for a total volume of 10 ml. PCR conditions were: 3 min at 94ºC (5 min for Lc109), 10 cycles of 94ºC for 15 sec, 55ºC for 15 sec (58ºC for Fca30 and Fca176), 72ºC for 30 sec, followed by 20 cycles of 89ºC for 15 sec, 55ºC for 15 sec (58ºC for Fca30 and Fca176), 72ºC for 30 sec, and a final extension at 72ºC for 10 min. Negative and positive controls in the form of water and DNA from previously genotyped individuals were included each time. PCR products were visualized on a LICOR DNA analyzer (Lincoln, Nebraska, USA) and scored relative to known genotypes.

We used GenePop 3.1 (*S4*) to test for linkage disequilibrium (LD) and deviations from Hardy-Weinberg equilibrium (HWE), expressed as F_{IS} (*S5*), using the method of Guo and Thompson (*S6*). After Bonferroni correction, none of the loci deviated significantly from HWE and no pair of loci showed LD (Results not shown).

Generation of FIV_{Pco} sequence data

All PCR reactions were set up in a room that was kept free of plasmids and PCR products. No more than four samples were handled at any given time and, after initial genotype determination, genetically distinct samples (e.g. different lineages) were chosen wherever possible for such sets of four, so that potential cross-contamination would become readily apparent.

Initial detection of FIV was based on amplification of the polymerase (*pol*) gene, using primers Pcp8R (5' GAG TTG CCC AAT TAA TCT TTC CTA 3') and Pcp11F (5' ACA CGA CCC CGA GTT GAG GC 3') followed by Pcp5F (5' AGA TTA GAA AAA GAA GGG AAA GTA G 3') and Pcp6R (5'AGT TCA TAC CCC ATC CAT TTA 3'). Samples were set up as 3-8 replicates in 50ml reactions that contained 0.2 mM dNTPs, 2mM MgCl, 1mM of each primer, 1.25 U Taq and on average 500-1000 ng DNA (smaller amounts from filter paper samples and up to 2500 ng for tissues other than blood). Second rounds used 40ml reactions with reagents as 1st round except for 1.5mM MgCl and 1.5 ml 1st round template. Samples that were negative after two rounds were also tested in a third round (reagents as in 2nd round) using primers 2581F (5' AAA TCA GGA AAA TGG AGA A 3') and 3012R (5' CTG TGG TGG GGA TTT GAA ACT 3'). If amplification was successful at 3rd round, the original sample was revisited and the protocol repeated using additional replicates. Samples were only considered positive if FIV detection was repeatable.

For verified positive individuals, amplification of *env* was conducted with primers PcEn4R (5' TGT TTC AAA TCC CCA CCA 3') and PcEn5F (5' AGA TTA GAA AAA GAA GGG AAA GTA G 3') followed by PcEn2R (5' CGT GGT GCC AGT GGT TGC TC 3') and PcEn3F (5' TGG AAT AGG ACA AAT GAG ACA GA 3') using the above described protocol. If amplification was unsuccessful after two rounds, primers 7406F (5' CCG TGG GGT GGA AGT AGA T 3') and 8115R (5' GTG CCA GTG GTT GCT CCT ATC A 3') were used in a third round. PCR regimes were the same for both genes. For the first round conditions were: 3 min at 94°C, 15 cycles of 94°C for 30 s, 44°C for 30 s, and 71°C for 90 s, followed by 30 cycles of 94°C for 30 s, 52°C for 30 s, and 71°C for

90s and a final extension for 5 min at 71°C. Second run conditions were: 3 min at 94°C, 45 cycles (30 cycles for 3rd round) of 94°C for 30 s, 50°C for 40 s, and 71°C for 50 s, and a final extension for 5 min at 72°C.

PCR products (*pol*: 399-538 bp, *env*: 678-885bp) were either cloned using PCR4 vector (Invitrogen, Inc., San Diego, Calif.) or used for direct sequencing. These two sequence-generating approaches were deemed equivalent since previous work had shown that for both *pol* and *env*, viral diversity within hosts based on cloned sequences was low (usually <1%)(*S7*). Further, we found in this study that sequences obtained with both approaches for six individuals were 99.6-100.0% identical, corroborating that sequences generated in either way well characterize the virus infecting an individual.

Preliminary examination revealed that sequences clustered in distinct lineages and that for all but seven individuals, phylogenetic clustering of *pol* and *env* sequences from the same infected individual was consistent in respect to these groups. For these seven cougars, including one previously reported to be co-infected (*S8*), sequences fell into different lineages. The presence of two lineages was either found for the same gene (n=4), suggestive of co-infection, or for different genes (n=3), potentially the result of recombination. Because we could not distinguish conclusively between these two possibilities and to avoid potential complications arising in the subsequent analysis, all sequences from these individuals were removed from the data set. Geographic distribution patterns of virus lineages remained largely unchanged by the omission of these viruses (see Fig. 2a in the main text). Although we cannot rule out recombination in our final data, we believe that this is negligible because (i) we would have likely noticed

recombination among lineages and (ii) recombination within lineages would have had very little effect given low (<5%) within lineage divergence.

For six individuals only *pol,* but not *env,* could be amplified. Five of the six cases were samples from filter paper but involved various lineages suggesting that insufficient DNA quality for amplification of longer fragments, rather than primer mismatch, was generally responsible for a failure to amplify *env*.

Phylogenetic analysis

A single sequence for either gene was included in the data set for each individual (randomly chosen where more than one sequence had been obtained). Excluding the individuals for which only *pol* could be amplified, data sets from both viral genes were then combined (n=85). A partition homogeneity test (*S9*) with 1000 replicates in PAUP* v4.0b10 (*S10*) did not reject the null hypothesis of homogeneity between the two genes (p=0.146), thus the concatenated data (1423 bp) were used in all further analyses.

An appropriate model of molecular evolution (GTR+ I + G) was found (*S11*) based on Akaike's Information Criterion for small sample sizes (AIC_C) and implemented in a heuristic Maximum Likelihood (ML) search in PAUP*. An FIV sequence from a Vancouver Island cougar served as an outgroup (*S7, S8*). Node support was evaluated based on 1000 bootstrap Neighbor Joining trees constructed from ML distances under the original substitution model. From the ML tree, we defined eight viral lineages based on high bootstrap support (>85%) and genetic divergence of up to 5%.

Sequences have been submitted to GenBank under accession no. DQ106985-DQ107160.

Molecular clock analysis

To estimate substitution rates, divergence times of lineages, and demographic trends for FIV_{Pco}, we analyzed the complete combined sequence data in BEAST v1.2 (*S12*). This program is aimed at molecular clock analyses based on dated sequences and uses Markov Chain Monte Carlo (MCMC) integration to estimate parameters across the entire space of possible trees, whereby trees are weighted by their posterior probability (*S13*). Three independent runs with 20 million generations were performed; subsequently the first 2 million from each run were removed as burn-in. Samples from the Markov Chain were taken every 4000 steps and the combined samples from all three runs were used to estimate parameters. All runs used a HKY+I+G model for codon position one and two and an HKY+G model for position three. Also, separate transition/ transversion ratios were estimated for each codon position in each of the two genes.

Although the Bayesian skyline plot technically estimates effective population size, we interpret our estimate as the effective number of infections (*S14*) since changes in virus population size at the host population level (rather than the level of the infected individual) will be most affected by the number of infection events, each of which represents a bottleneck event in which very few viruses are transmitted (*15*).

Comparative phylogenetic analysis

We used a comparative method termed phylogenetic generalized least squares (PGLS; see *S16*), implemented in COMPARE 4.6 (*S17*) to test for evidence of viral spatial diffusion. All calculations were based on the maximum *a posteriori* (MAP) tree, i.e. the tree associated with the highest posterior likelihood, identified in the previous Bayesian

analysis in BEAST (see above). As the first step, latitude and longitude associated with each sequence were treated as 'traits', in order to infer the ancestral coordinates of each lineage ancestor in the phylogeny using the 'PGLS-ancestor' option in COMPARE. Next, we calculated the geographic distance of each sampled virus from its lineage ancestor and regressed it against the time that had elapsed since that ancestor (based on median age estimated in BEAST) using the 'PGSL-relationship' option in COMPARE. This method takes into account that 'distance' and 'time' are potentially not independent variables due to shared ancestry of viral sequences and adjusts the resulting regression coefficient and parameter estimates accordingly. The program does not allow constraint of the regression model to a zero intercept, which clearly would have been most appropriate in our case. However, confidence intervals for the intercept obtained in PGLS analysis included zero (CI: -188.67 – 101.85), suggesting that a zero intercept model would provide an adequate fit.

Supporting Text

Test for genetic population structure in cougars using individual-based methods

Without using information about the origin of individuals, microsatellite data were examined for evidence of population subdivision using a Bayesian approach implemented in program Structure (*S18*). To determine the number of distinguishable population clusters (K) in the data, 5 independent runs were conducted for K =1-10. Runs assumed admixture and correlated allele frequencies and used a Markov Chain Monte Carlo and a burn-in period of 500,000. Plotting averages across runs we found that Pr(data | K) reached a plateau for values of K between 2 and 8, with a peak at K=5 (Fig. S1). Using

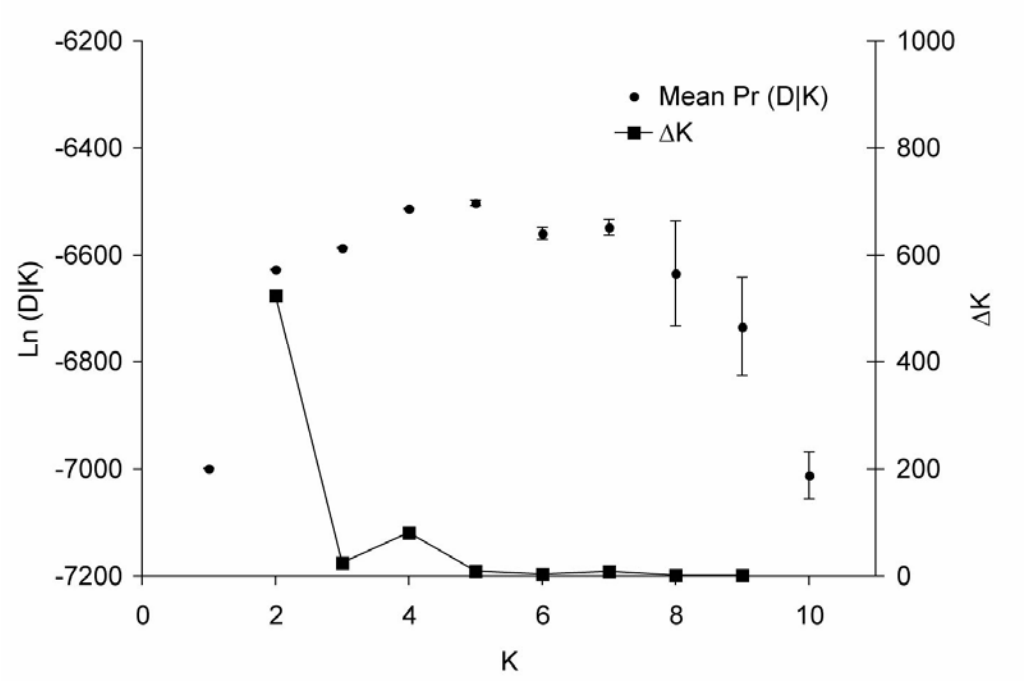
simulated data, Evanno et al. (*S19*) recently showed that the highest Pr (data | K) does not reliably identify the true number of clusters but that the uppermost hierarchical level of structure can usually be found based on the second order rate of change for K (ΔK), which in our case showed an unambiguous maximum at K=2 (Fig. S1).

The distribution of individuals assigned to each of the two clusters revealed that the two identified clusters in fact represent the end points of a genetic continuum (Fig. S2a). Only cougars in the northwest and extreme south and east of our study area could be consistently assigned to one particular genetic cluster, whereas individuals in a broad transition zone in between were assigned to either one in roughly equal proportions. Thus, the pattern is consistent with isolation by distance but otherwise does not suggest major disruptions in the genetic continuity of cougars throughout our study area.

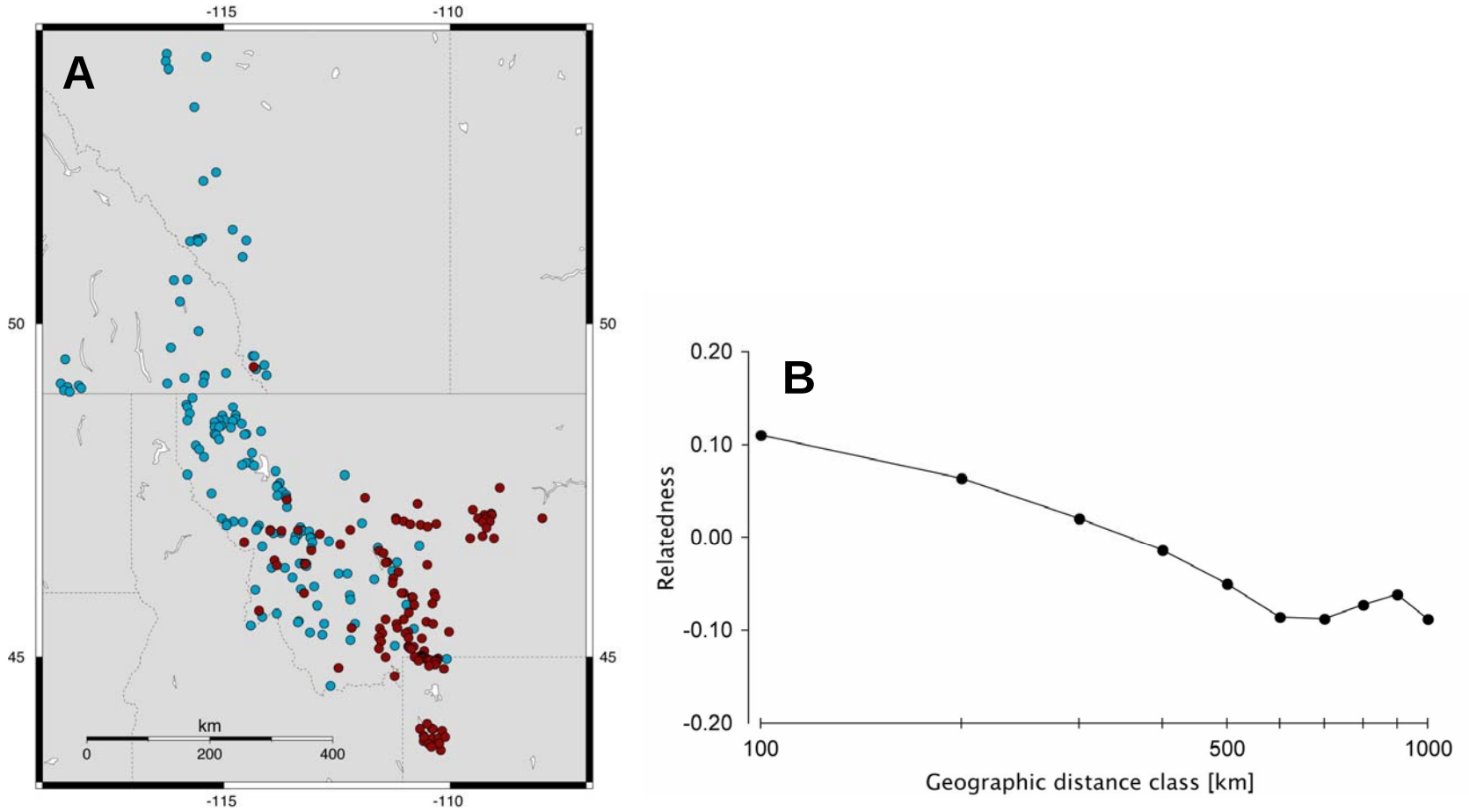
This conclusion was further substantiated by examining the relationship between pairwise relatedness of individuals (*S20*) and the geographic distance (broken down into distance classes) between them. The former was calculated in RELATEDNESS 5.0.8 (available at http://www.gsoftnet.us/GSoft.html) and statistical significance of the relationship was assessed in program SPAGEDI (*S21*) using 100 permutations. Results show a weak but significant decline ($R=0.92$; $P<0.001$) in relatedness with increasing spatial distance up to distances of 600 km, after which relatedness becomes uncorrelated (Fig. S2b). This result indicates isolation by distance and thus supports the notion of a genetically contiguous cougar population.

Acknowledgements

This work would clearly not have been possible without the help and generous support of the following people and organizations: T. K. Ruth, K. Murphy, R. DeSimone, R. Grey, C. Chetkiewicz. T. Shury, A. Kortello, H. Schwantje, M. Hall, Montana Fish Wildlife and Parks (especially J. Firebaugh, R. Wiesner, K. Alt, K. Choates, N. Anderson, R. Rauscher, T. Stivers, C. Loecker), M. Dean and the Montana State Houndsmen Association, Murray Selzer and the Alberta Tree Hound Association, Alberta Professional Outfitter Society, Montana Outfitters and Guides Association, Guide and Outfitter Association British Columbia, the Wilburforce Foundation, the Yellowstone to Yukon Conservation Initiative, the National Science Foundation, S. Painter, K. Pilgrim, C. Engkjer, K. Samuel, N. Akamine, L.A. Real, M.K. Schwartz, F.W. Allendorf, L.S. Mills, D. Pletscher, and A. Rodrigo.



Supporting Figure S1. Results from program Structure (*S18*), indicating the number of genetically homogeneous clusters K in the cougar microsatellite data. Shown is the posterior probability of the data given a particular value of K ("Pr (D|K)") averaged over five runs (±SD). Also shown is the second order rate of change for K (ΔK), which has its maximum at K=2.



Supporting Figure S2. Analysis of cougar genetic structure based on eleven microsatellite loci. **(A)** Spatial distribution of cougars assigned to two clusters based on results from STRUCTURE. Map generated using Online Map Creator available at http://www.aquarius.geomar.de/omc/. **(B)** Decaying relatedness among individuals with increasing spatial distance (log scale).

Supporting Table S1. Cougar samples collected for analysis of host and retroviral DNA in the northern Rocky Mountains, 1990-2004. Samples represent lymph nodes, salivary glands and blood, the latter in the form dried blood of Isocode filter paper (hunter samples) or as fresh peripheral blood mononuclear cells (PBMC, for cougar population studies). Numbers shown reflect individuals, i.e. no repeat samples.

Sampling area	Sampling period	Source[a]	Number of samples/FIV positive			
			Blood	Lymph nodes	Salivary glands	Total
Alberta	2001-04	H, S[b,c]	26/9	-	-	26/9
British Columbia	1998-03	H	13/3	4/1	-	17/4
Montana	2001-04	H, S[d]	98/28	76/28	14/2	188/58
Yellowstone (Montana/Wyoming)	1990-04	S[e]	103/21	-	-	103/21
Teton Range (Wyoming)	2001-04	S[f]	18/6	-	-	18/6
All areas						**352/98**

[a] H= hunter samples; S=samples from cougar population studies, see footnote for investigator

[b] Todd Shury, Banff National Park

[c] Cheryl Chetkiewicz, University of Alberta

[d] Rich DeSimone, Montana Fish Wildlife and Parks

[e] Kerry Murphy & Toni K. Ruth, Hornocker Wildlife Institute/ Wildlife Conservation Society

[f] Rachel Grey, Beringia South Institute

References

1. S. D. Werman, M. S. Springer, R. J. Britten, in *Molecular Systematics* D. M. Hillis, C. Moritz, B. K. Mable, Eds. (Sinauer Associates, Inc., Sunderland, Massachusetts, 1996) pp. 169-204.
2. M. Menotti-Raymond *et al.*, *Genomics* 57, 9 (1999).
3. L. E. Carmichael, W. Clark, C. Strobeck, *Mol. Ecol.* 9, 2197 (2000).
4. M. Raymond, F. Rousset, *J. Hered.* 86, 248 (1995).
5. B. S. Weir, C. C. Cockerham, *Evolution* 38, 1358 (1984).
6. S. W. Guo, E. A. Thompson, *Biometrics* 48, 361 (1992).
7. R. Biek *et al.*, *J. Virol.* 77, 9578 (2003).
8. M. A. Carpenter *et al.*, *J. Virol.* 70, 6682 (1996).
9. J. S. Farris, M. Kallersjo, A. G. Kluge, C. Bult, *Cladistics* 10, 315 (1994).
10. D. L. Swofford, PAUP* 4.0b10 (Sinauer Associates, Sunderland, Massachusetts, 2003).
11. D. Posada, K. A. Crandall, *Bioinformatics* 14, 817 (1998).
12. A. J. Drummond, A. Rambaut, BEAST v2.1. Available from http://evolve.zoo.ox.ac.uk/beast/ (2003).
13. A. J. Drummond, G. K. Nicholls, A. G. Rodrigo, W. Solomon, *Genetics* 161, 1307 (2002).
14. P. Lemey *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* 100, 6588 (2003).
15. K. M. McGrath, *Vir. Res.* 76, 137 (2001).
16. E. P. Martins, T. F. Hansen, *Am. Nat.* 149, 646 (1997).

17. E. P. Martins. Computer programs for the statistical analysis of comparative data. Distributed by the author at http://compare.bio.indiana.edu/ (2004).

18. J. K. Pritchard, M. Stephens, P. Donnelly, *Genetics* 155, 945 (2000).

19. G. Evanno, S. Regnaut, J. Goudet, *Mol Ecol* 14, 2611 (2005).

20. D. C. Queller, K. F. Goodnight, *Evolution* 43, 258 (1989).

21. O. J. Hardy, X. Vekemans, *Mol. Ecol. Notes* 2, 618 (2002).