Supporting Information

Cornelis et al. 10.1073/pnas.1115346109

SI Methods

Biological Samples. Hysterectomy or caesarian operations were performed on pregnant domestic dog (*Canis lupus familiaris*) females obtained from ENVA-INRA at day 30 of gestation, and placenta, uterus, and embryos were dissected and either stored in liquid nitrogen or processed for in situ hybridization. Total RNA was extracted from the frozen organs using the RNeasy RNA isolation kit (Qiagen), and genomic DNA was extracted by phenol–chloroform extraction. For other tissues from adult domestic dog (cortex, lung, muscle, testis, liver, spleen, kidney, intestine, and heart), total RNA was obtained from Zyagen. Pregnant domestic cat (*Felis catus*) females obtained from ENVA-INRA were similarly processed. Total RNA and genomic DNA were extracted from the frozen organs as above. Total RNA from adult domestic cat organs (cortex, lung, muscle, testis, liver, spleen, kidney, intestine, heart, and skin) were from Zyagen.

Placenta of a California sea lion (*Zalophus californianus*) was recovered after birth. A piece of skin and a blood sample were collected in the course of a benign surgical operation on a sea lion adult. Placenta and skin tissues were conserved in RNAlater (Qiagen) before extraction of total RNA using the RNeasy RNA isolation kit (Qiagen). Total RNA was extracted from the blood sample using the RNA Blood kit II (PaxGene).

Genomic DNA from red fox (*Vulpes vulpes*) was a gift from H. Gachot (Institut Pluridisciplinaire Hubert Curien, Strasbourg, France). Genomic DNA from long-tailed pangolin (*Manis longicaudata*) was extracted from tissue supplied by F. Catzeflis (Collection de Tissus de Mammifères de Montpellier, Institut des Sciences de l'Evolution, Montpellier, France). Genomic DNA from the Malayan pangolin (*Manis javanica*), the giant panda (*Ailuropoda melanoleuca*), the African palm civet (*Nandinia binotata*), the spotted linsang (*Prionodon pardicolor*), the striped skunk (*Mephitis mephitis*), the spectacled Andean bear (*Tremarctos ornatus*), and the walrus (*Odobenus rosmarus*) was extracted from tissues supplied by G. Veron (Paris, France). Genomic DNA from the bat-eared fox (*Otocyon megalotis*), the spotted hyena (*Crocuta crocuta*), the Oriental small-clawed otter (*Aonyx cinerea*), the meerkat (*Suricata suricatta*), the mountain lion or cougar (*Felis puma concolor*), the red panda (*Ailurus fulgens*), the Asian bearcat (*Arctictis binturong*), the fossa (*Cryptoprocta ferox*), the raccoon (*Procyon lotor*), the tiger (*Panthera tigris*), the African lion (*Panthera leo*), the South African white lion (*Panthera leo krugeri*), the California sea lion (*Zalophus californianus*), the serval (*Leptailurus serval*), the Persian leopard (*Panthera pardus saxicolor*), and the clouded leopard (*Neofelis nebulosa*) was purified from blood samples collected by B. Mulot and R. Potier (Saint Aignan, France) using the DNA Blood Kit II (PaxGene).

Dog and Cat Envelope Protein Coding Sequences Coordinates. Coding sequence locations were (chromosome number, strand orientation, and coding sequence coordinates indicated): *canis-env1* (chr3, +, 86665505–86666926), *canis-env2* (chr35, +, 27290446–27291789), and *canis-env3* (chrX, +, 53591484–53593214) for the dog and *felis-env1* (chrB1, –, 200328503–200329924), *felis-env2* (chrB2, –, 2944663–2946015), *felis-env3* (chrB1, +, 129108152–129109339), *felis-env4* (chrE1, +, 5984282–5985712 and chrX, +, 36710233–36711663), *felis-env5* (chrD3, +, 9476077–9477303), *felis-env6* (chrC2, –, 3766634–3767776), *felis-env7* (Un_ACBE01158631, –, 656–1948), and *felis-env8* (chrA2, +, 560053–562068 and Un_ACBE01496605, –, 1558–3546) for the cat.

Real-Time RT-PCR. Syncytin-Car1 mRNA expression was determined by real-time qRT-PCR. Reverse transcription was performed with 1 μg DNase-treated RNA as in ref. 1. Real-time qPCR was with 5 μL diluted (1:15) cDNA in a final volume of 25 μL using SYBR Green PCR Master Mix (Applied Biosystems). PCR was carried out using an ABI PRISM 7000 sequence detection system. Primers are listed in Table S1. The transcript levels were normalized relative to the amount of the Peptidylprolyl isomerase A (PPIA) gene mRNA encoding Cyclophilin A. Samples were assayed in duplicate. Reverse-transcribed cDNA from placenta was PCR-amplified using primers on each side of the putative splice donor and acceptor sequences (primers listed in Table S1).

In Situ Hybridization. Freshly collected dog and cat placentae (at midgestation, the same placentae were used for real-time qRT-PCR expression assays) were fixed in 4% paraformaldehyde at 4 °C and embedded in paraffin; serial sections (7 μm) were either stained with HES or used for in situ hybridization. For dog and cat genes, two PCR-amplified *syncytin-Car1* fragments of 321 and 369 bp or 294 and 299 bp, respectively, (primers listed in Table S1) were cloned into pGEM-T Easy (Promega) for in vitro synthesis of the antisense and sense riboprobes generated with SP6 RNA polymerase and digoxigenin 11-UTP (Roche Applied Science). Sections were processed, hybridized at 42 °C overnight with the pooled riboprobes, and incubated further at room temperature for 2 h with alkaline phosphatase-conjugated antidigoxigenin antibody Fab fragments (Roche Applied Science). Staining was revealed with nitro-blue tetrazolium (NBT) and 5-bromo-4-chloro-3'-indolyl-phosphate (BCIP) alkaline phosphatase substrates as indicated by the manufacturer (Roche Applied Science).

***Syncytin-Car1* Expression Vector and Infection Assay.** *Syncytin-Car1* fragments PCR-amplified from the genomic DNA of each Carnivora species were cloned into the phCMV-G expression vector (GenBank accession AJ318514). Syncytin-Car1–HIV-1 pseudotypes were produced by cotransfecting 10^6 293T cells with 0.8 μg p8.91 (HIV-1 gag-pol expression vector) (2), 1 μg pSIN-nlsLacZ (lentiviral vector [3] with GFP replaced by nlsLacZ), and 0.35 μg *syncytin-Car1* or murine leukemia virus (ecotropic and amphotropic as described in refs. 4 and 5) *env* expression vectors using the Fugene6 transfection kit (Roche). Supernatants were harvested 48 h after transfection, filtered, supplemented with Polybrene (8 μg/mL), transferred to target ($5–8 \times 10^4$ cells/well), and spinoculated at 1,200 × g for 2.5 h. X-Gal staining was performed 3 d postinfection. All cell lines are described in refs. 6 and 7, except A72 (ECACC 89050908) and DK (ECACC 93120836) cells and HuH7 (JCRB 0403) cells; they were grown in DMEM supplemented with 10% FCS (Invitrogen).

1. de Parseval N, Lazar V, Casella JF, Benit L, Heidmann T (2003) Survey of human genes of retroviral origin: Identification and transcriptome of the genes with coding capacity for complete envelope proteins. *J Virol* 77:10414–10422.
2. Zufferey R, Nagy D, Mandel RJ, Naldini L, Trono D (1997) Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat Biotechnol* 15:871–875.
3. Demaison C, et al. (2002) High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency [correction of imunodeficiency] virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter. *Hum Gene Ther* 13:803–813.
4. Dewannieux M, Collins MK (2008) Spontaneous heteromerization of gammaretrovirus envelope proteins: A possible novel mechanism of retrovirus restriction. *J Virol* 82:9789–9794.

5. Ribet D, Harper F, Esnault C, Pierron G, Heidmann T (2008) The GLN family of murine endogenous retroviruses contains an element competent for infectious viral particle formation. *J Virol* 82:4413–4419.
6. Dupressoir A, et al. (2005) Syncytin-A and syncytin-B, two fusogenic placenta-specific murine envelope genes of retroviral origin conserved in Muridae. *Proc Natl Acad Sci USA* 102:725–730.
7. Heidmann O, Vernochet C, Dupressoir A, Heidmann T (2009) Identification of an endogenous retroviral envelope gene with fusogenic activity and placenta-specific expression in the rabbit: A new "syncytin" in a third order of mammals. *Retrovirology* 6:107.

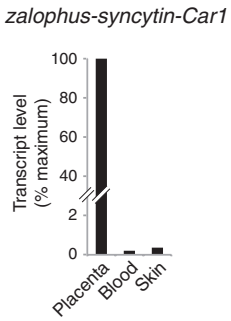


Fig. S1. Real-time quantitative RT-PCR analysis of the *syncytin-Car1* gene from the sea lion (*Z. californianus*) shows placenta-specific expression. Same experimental conditions as in Fig. 4 using the *canis-env1* primers; transcript levels are expressed as percent of maximum with an ordinate scale change to see the low values from nonplacental organs, and they were normalized relative to the amount of RNA from a control gene [peptidyliprolyl isomerase A (PPIA) using the *canis-PPIA* primers] (*Methods*). A piece of placental tissue was collected a few hours postpartum inside the sea lion enclosure (at the Zooparc de Beauval, Saint Aignan, France) and immediately transferred into RNAlater for RNA extraction; a blood sample was collected from a sea lion adult into a PAXgene Blood RNA Tube (PreAnalytiX), and RNA was extracted using the PAXgene Blood RNA kit (PreAnalytiX). A piece of skin was collected in the course of a benign surgical operation on a sea lion adult and processed as the placental tissue.

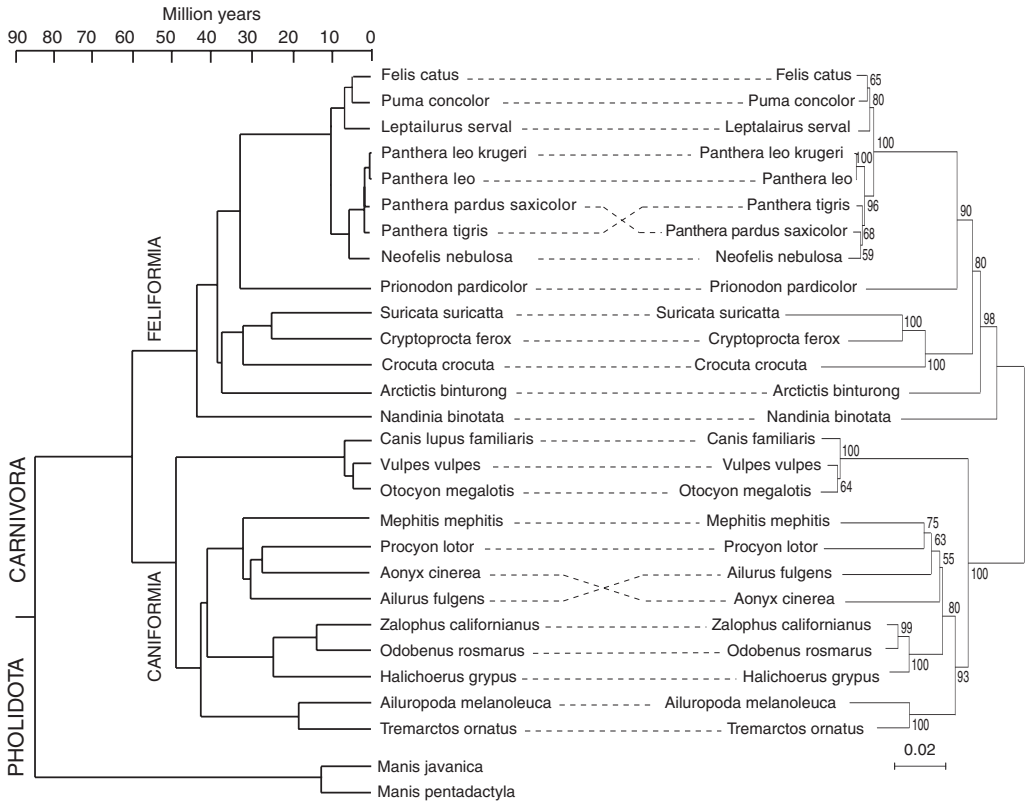
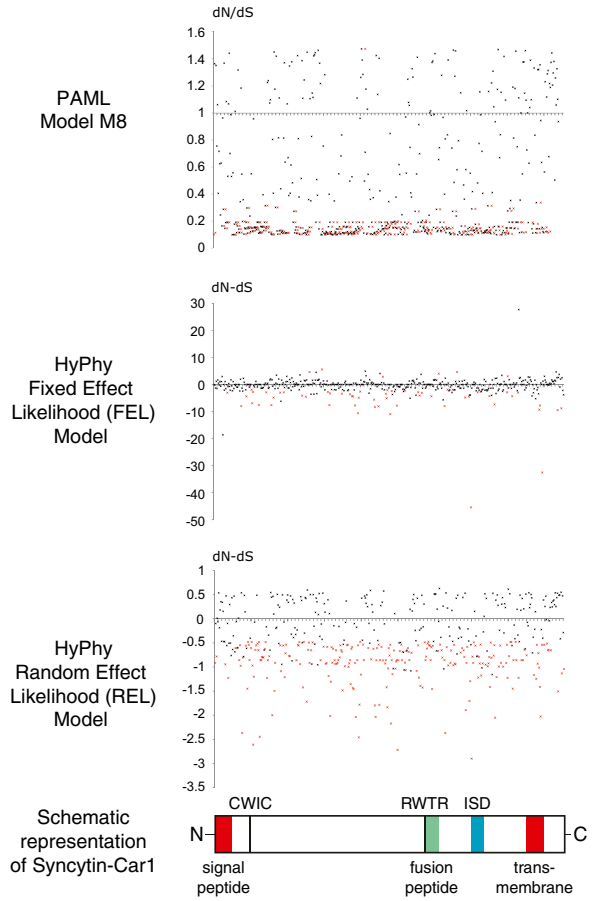
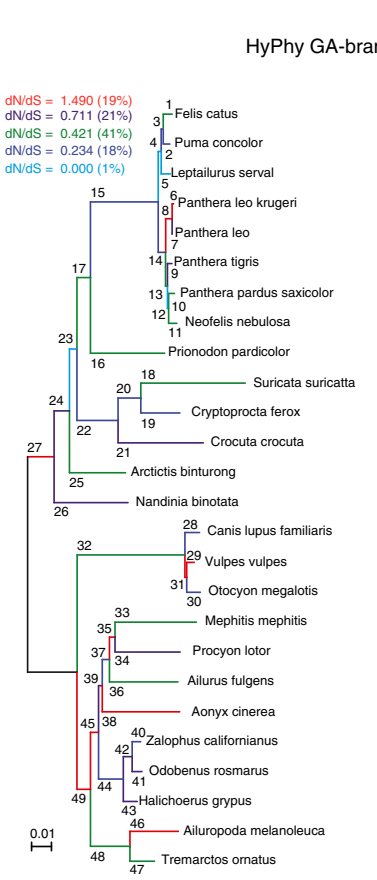


Fig. S2. Comparison between the Carnivora phylogenetic tree and the Syncytin-Car1–based phylogenetic tree. The phylogenetic trees are from Figs. 9 and 10, and they display strong congruence. Note that the limited cases where divergence can be observed correspond to nodes with low bootstrap values.





branch number	estimated dN/dS	standard deviation	probability (dN/dS>1)
1	0.42	0.10	7.1E-03
2	0.23	0.06	7.1E-04
3	0.25	0.11	4.3E-03
4	0.00	0.02	3.0E-07
5	0.03	0.14	4.1E-03
6	11.25	313.25	9.3E-01
7	0.53	0.44	1.4E-01
8	1.39	0.28	8.9E-01
9	0.68	0.17	3.1E-02
10	0.41	0.10	4.1E-03
11	0.45	0.12	3.7E-03
12	0.02	0.09	2.0E-03
13	0.42	0.16	1.8E-02
14	0.22	0.04	1.3E-04
15	0.26	0.06	0.0E+00
16	0.41	0.02	1.5E-05
17	0.41	0.07	8.8E-04
18	0.41	0.02	2.2E-06
19	0.28	0.07	1.4E-05
20	0.27	0.07	2.6E-04
21	0.69	0.03	6.5E-04
22	0.22	0.03	0.0E+00
23	0.01	0.05	2.7E-04
24	0.55	0.13	3.4E-04
25	0.70	0.13	2.6E-02
26	0.69	0.04	8.0E-04
27	**1.47**	**0.10**	**9.8E-01**
28	0.30	0.09	3.1E-04
29	1.37	0.31	8.8E-01
30	0.28	0.08	3.4E-04
31	7.56	247.30	9.4E-01
32	0.41	0.02	0.0E+00
33	0.41	0.02	7.2E-06
34	0.69	0.06	6.9E-03
35	**1.50**	**2.70**	**1.0E+00**
36	0.40	0.03	3.1E-06
37	0.27	0.15	7.5E-03
38	**1.48**	**0.06**	**1.0E+00**
39	0.65	0.10	2.7E-03
40	0.23	0.05	2.3E-04
41	0.71	0.16	4.1E-02
42	0.72	0.19	5.6E-02
43	0.69	0.08	6.0E-03
44	0.27	0.07	2.9E-04
45	**1.45**	**0.19**	**9.6E-01**
46	**1.49**	**0.05**	**1.0E+00**
47	0.41	0.03	1.2E-04
48	0.41	0.03	1.1E-04
49	**1.49**	**0.08**	**1.0E+00**

Fig. S3. Analysis of site- and branch-specific selections within *syncytin-Car1*. (*A*) Site-specific analysis of selection on *syncytin-Car1* gene codons. Three independent analyses were performed using PAML (M8 model) or HyPhy (random effect likelihood [REL] and fixed effect likelihood [FEL]) packages. HyPhy package was run on the webserver www.datamonkey.org. The relevant indexes for selective pressure are provided for each codon. A schematic representation of Syncytin-Car1 protein domains is also given, with the same conventions as in Fig. 3. At variance with the PAML model (M8), where a synonymous mutation rate (dS) is estimated for the entire sequence, HyPhy models (FEL and REL) independently estimate nonsynonymous mutation rate (dN) and dS for each codon, allowing the dS value to be null. Consequently, for the HyPhy models, dN − dS values are represented instead of dN/dS. Significant values (probability ≥ 0.95) are represented in red, whereas nonsignificant values are represented in black. All three models predict only a few codons—if any—to be under positive selection (i.e., dN/dS > 1 or dN − dS > 0), with only one site (M8), six sites (FEL), or zero sites (REL) being detected with weakly positive values. Moreover, no specific domain emerges, and weakly positive codons are rather equally distributed among the entire sequence. (*B*) Branch-specific analysis of selection along the *syncytin-Car1*–based phylogenetic tree. Analysis was performed using the GA-branch method from the HyPhy Package on the webserver www.datamonkey.org, and the selected model was the one with the best Akaike Information Criterion. Such analysis identifies five classes of branches: four classes under strong purifying selection (dN/dS values < 1) and one under weak positive selection (dN/dS = 1.49). (*Left*) The neighbor-joining tree of *syncytin-Car1* (same as in Fig. 10) is represented together with the branch numbers, and the color code for each class is indicated on the top. (*Right*) Estimated branch-specific dN/dS values are indicated ± SD as well as the probability for each branch of being under positive selection. Positively selected branches with probabilities ≥0.95 are highlighted in bold and include two terminal branches (*Ailuropoda melanoleuca* and *Aonyx cinerea*) and four internal branches. However, the corresponding dN/dS values remain low (≤1.5).

Table S1. List of primers

Primer names	Primer sequences
RT-qPCR	
Canis-env1-F	5′-TATGGGACGAATTTATGTGGCT
Canis-env1-R	5′-CCAATCAGTAGCAAACCAGTCA
Canis-env2-F	5′-GCTTGCCCTCCATTTATCCTAT
Canis-env2-R	5′-CCCCCACCTAGTTCTGAGTCC
Canis-env3-F	5′-AAGGCAGCGGTACCTCAC
Canis-env3-R	5′-AGTTGGTGGATGACCCTT
Canis-PPIA-F	5′-GTCAACCCCACCGTGTTCTT
Canis-PPIA-R	5′-CTGCTGTCTTTGGAACTTTGT
Felis-env1-F	5′-GGCACAGACTGGCATTACTT
Felis-env1-R	5′-ATTAGCGTTGGCATATTCATT
Felis-env2-F	5′-CGAGGATATGACGTGTGTGA
Felis-env2-R	5′-GAAGTTAGGTTGGTCGCCC
Felis-env3-F	5′-CTGGAAGCCTACCAAGGT
Felis-env3-R	5′-AGTGGCTAAGTAAACCCTCC
Felis-env4-F	5′-GTCCTGCTGACCACGCCCAC
Felis-env4-R	5′-GTTCTGCCGTCAGTAGATCC
Felis-env5-F	5′-GACGGCTAATAGAAGATG
Felis-env5-R	5′-GTTGAGTCCGAACCATAA
Felis-env6-F	5′-ACAATCCCTGGACCCACT
Felis-env6-R	5′-AAATAGGTGTACGCCATGTTA
Felis-env7-F	5′-CCGACCCACAACATAAACTG
Felis-env7-R	5′-TACTCTGGGCCATAACTCAA
Felis-env8-F	5′-GCATTCTGACATACAGGCAC
felis-env8-R	5′-GTTTGAGTCTCTCCCTCAGC
Felis-PPIA-F	5′-GTCAACCCCATCGTGTTTTT
Felis-PPIA-R	5′-CTGCTGTCTTGGGAACTTTGTC
Splice site determination	
Syncytin-Car1-splice-F	5′-ATACATCTCGAGACCCGTATCATCTTGGCGACCACAAA
Syncytin-Car1-splice-R	5′-ATACATACGCGTTCTTCTGTAGCTTCTTCAGGC
In situ hybridization probe synthesis	
Canis-syncytin-Car1-ISH-F1	5′-TATGTGGCTTACTCCTACC
Canis-syncytin-Car1-ISH-R1	5′-TCTTTACGCAATTTCCCTT
Canis-syncytin-Car1-ISH-F2	5′-TTAAATTCTGAAGTTGCATTG
Canis-syncytin-Car1-ISH-R2	5′-TTGTTCTGATAATTTTTCTGA
Felis-syncytin-Car1-ISH-F1	5′-CTTGGTGGGTTTCTCCTTT
Felis-syncytin-Car1-ISH-R1	5′-AGCCATCATCAGTATATCTCA
Felis-syncytin-Car1-ISH-F2	5′-CAAACTATACCACAAAAGCCC
Felis-syncytin-Car1-ISH-R2	5′-GCCCATATACCTCCACTAGAC
Amplification of genomic *syncytin-Car1* in Carnivora species	
Forward primer: syncytin-Car1-genomic-F	5′-ATACATCTCGAGACCCGTAAAGGACCAATAAGTAAGGC
Reverse primer for all species but *Ailurus fulgens* and *Mephitis mephitis*: syncytin-Car1-genomic-R1	5′-ATACATACGCGTTGAAGTATAGTTTGAAATCAGGA
Reverse primer for *A. fulgens* and *M. mephitis*: syncytin-Car1-genomic-R2	5′-ATACATACGCGTCAGCTTTTTGGTGGAGTCTT
Search for *syncytin-Car1* in Pholidota species	
Syncytin-Car1-ORF-F	5′-ATACATCTCGAGACCCGTAAAGGACCAATAAGTAAGGC
Syncytin-Car1-ORF-R	5′-ATACATACGCGTTCTTCTGTAGCTTCTTCAGGC
Syncytin-Car1-locus-F1	5′-TAATCACTCAGGTGGCAT
Syncytin-Car1-locus-F2	5′-ACAGTCTTTCACATCAAAGGA
Syncytin-Car1-locus-F3	5′-AGAAGGAAAAACAGGAATTT
Syncytin-Car1-locus-R1	5′-ATACATACGCGTTGAAGTATAGTTTGAAATCAGGA
Syncytin-Car1-locus-R2	5′-ATACATACGCGTCAGCTTTTTGGTGGAGTCTT
Syncytin-Car1-locus-R3	5′-ATACATACGCGTTTCAACAATGTTATTTCTTCTGA
Syncytin-Car1-internal-F1	5′-GAATTTATGTGGCTTACTCCTAC
Syncytin-Car1-internal-F2	5′-TATGTGGCTTACTCCTACCATAG
Syncytin-Car1-internal-R	5′-CCAGATTATGTCTTCTATTCCCA