BMC
Evolutionary Biology

RESEARCH ARTICLE

Open Access

The dynamic proliferation of CanSINEs mirrors the complex evolution of Feliforms

Kathryn B Walters-Conte[1*], Diana LE Johnson[2], Warren E Johnson[3], Stephen J O'Brien[4] and Jill Pecon-Slattery[3]

Abstract

Background: Repetitive short interspersed elements (SINEs) are retrotransposons ubiquitous in mammalian genomes and are highly informative markers to identify species and phylogenetic associations. Of these, SINEs unique to the order Carnivora (CanSINEs) yield novel insights on genome evolution in domestic dogs and cats, but less is known about their role in related carnivores. In particular, genome-wide assessment of CanSINE evolution has yet to be completed across the Feliformia (cat-like) suborder of Carnivora. Within Feliformia, the cat family Felidae is composed of 37 species and numerous subspecies organized into eight monophyletic lineages that likely arose 10 million years ago. Using the Felidae family as a reference phylogeny, along with representative taxa from other families of Feliformia, the origin, proliferation and evolution of CanSINEs within the suborder were assessed.

Results: We identified 93 novel intergenic CanSINE loci in Feliformia. Sequence analyses separated Feliform CanSINEs into two subfamilies, each characterized by distinct RNA polymerase binding motifs and phylogenetic associations. Subfamily I CanSINEs arose early within Feliformia but are no longer under active proliferation. Subfamily II loci are more recent, exclusive to Felidae and show evidence for adaptation to extant RNA polymerase activity. Further, presence/absence distributions of CanSINE loci are largely congruent with taxonomic expectations within Feliformia and the less resolved nodes in the Felidae reference phylogeny present equally ambiguous CanSINE data. SINEs are thought to be nearly impervious to excision from the genome. However, we observed a nearly complete excision of a CanSINEs locus in puma (*Puma concolor*). In addition, we found that CanSINE proliferation in Felidae frequently targeted existing CanSINE loci for insertion sites, resulting in tandem arrays.

Conclusions: We demonstrate the existence of at least two SINE families within the Feliformia suborder, one of which is actively involved in insertional mutagenesis. We find SINEs are powerful markers of speciation and conclude that the few inconsistencies with expected patterns of speciation likely represent incomplete lineage sorting, species hybridization and SINE-mediated genome rearrangement.

Keywords: Incomplete lineage sorting, SINEs, Carnivora, Speciation, transposable elements, Adaptation, Feliformia, Felidae

Background

Repetitive short interspersed elements (SINEs) are ubiquitous eukaryotic retrotransposons. SINE sequences are approximately 70–700 base pairs (bp) averaging about 250 bp [1] with most organized into an RNA gene-derived region, a di-nucleotide repeat region and terminating in a poly A or poly A/T tail [2-4]. SINEs are "non-autonomous" such that amplification and integration is dependent on enzymes derived from the host genome and long

interspersed nuclear elements (LINEs) [5]. Proliferation is initiated via recognition of promoter boxes residing in the tRNA-related region of the genomic "master-copy" by host-derived RNA polymerase III and eventually results in novel retrotransposed copies [6]. SINEs constitute roughly 10% of the mammalian genome [1,7-10] and classification into family or subfamily designations is based on sequence variation and presence in specific evolutionary lineages [5,9,11-13,1,14].

Initially viewed as "junk" DNA without function, seminal studies in rodents [15,16] and primates [17-19] indicate a far more important role for SINEs in genome organization, gene evolution, and disease. For example, germ-line

* Correspondence: kwalt@american.edu
[1]Department of Biology, American University, 101 Hurst Hall 4440 Massachusetts Ave, Washington, DC 20016, USA
Full list of author information is available at the end of the article

insertions are correlated with non-homologous genome rearrangements, generation of novel coding sequences, alteration of regulatory elements and are linked with the origin and evolution of highly conserved non-coding elements in mammals [18,20-26]. Within somatic cells, *de novo* SINE integration can disrupt pathways involved with cell differentiation [27], modulate intracellular targeting of mRNAs [15] and potentially provide other cell-specific phenotypes [28].

Direct phenotypic variation is possible by altering gene expression via insertion into coding regions or interference from the internal RNA polymerase promoters in SINEs [29]. Analysis of the dog genome revealed SINE insertion polymorphisms resulting in anti-sense transcription that provide alternate splice site junctions [30]. For example, alterations of fur color [31], muscular disorders [32,33] and body size diversity [34,35] in Canidae are correlated with SINE insertions associated respectively with *SILV, PTPLA* and *IGF1*. In addition, SINE insertion into an exon of *STK38L* causes retinal degeneration [36] and an ancient SINE locus serves as an enhancer for fibroblast growth factor 8 (*Fgf8*) during mammalian brain formation [37].

SINEs are highly informative markers used in mammalian phylogenetic and population genetic studies of cetaceans [38], carnivores [39-42], primates [11,43,44], rodents [45,46], xenarthrans [47], marsupials [48,49] and the diverse assemblage of African species termed Afrotheria [50]. With few mechanisms for precise removal, SINE insertions are nearly homoplasy-free unidirectional markers and therefore informative in deciphering complex patterns of speciation [40,50-52]. In general, phylogenetic inferences rely upon presence and absence data of SINE loci among taxa. However, instances of parallel insertions unrelated to phylogenetic associations have been detected through sequence data [41,44,53,54] and indicate SINEs target specific sequence motifs during proliferation [1,55-57]. Incomplete lineage sorting of ancestral polymorphisms via ongoing hybridization or introgression among populations may cause contradictory findings in SINE-based phylogenetic reconstructions [58,59]. Consequently, accurate species trees are required to serve as a reference phylogeny for interpreting patterns of insertion and sequence divergence at SINE loci.

Here we utilized the well-resolved phylogeny of Felidae as a species tree to investigate the evolution a lesser known family of mammalian SINEs; those within the Order Carnivora, termed CanSINEs. The two suborders of Carnivora are Caniformia (dog-like) and Feliformia (cat-like). Caniformia is organized into Ursidae (bear), Canidae (domestic dogs, wolves, foxes, jackals, coyotes), Otariidae (eared seals), Odobenidae (walrus), Phocidae (earless seals), Mustelidae (badgers, weasels and otters), Mephitidae (skunks), Procyonidae (raccoons, coatis, kinkajous, olingos, ringtails and cacomistles), Ailuridae (red panda) [60,61]. Feliformia is composed of Felidae (cats), Viverridae (civets, genets, African linsang), Prionodontidae (Asiatic linsang), Eupleridae (Malagazy carnivores), Naniniidae (african palm civet), Herpestidae (mongooses), and Hyaenidae (hyenas) [60]. Initially discovered in multiple species of Caniformia [3,62,63], CanSINEs were presumed absent from Feliformia. This was revised upon further studies of the feline Y-chromosome [53,64] and through whole genome sequence analyses [8,65].
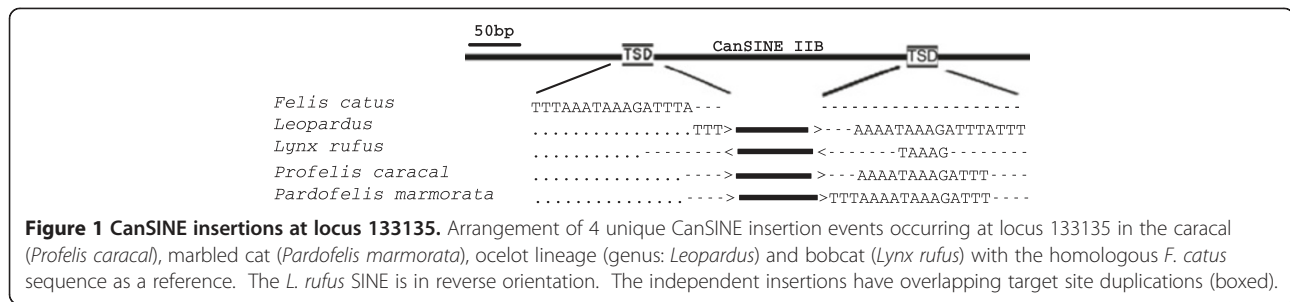
We used comparative methods to sequence CanSINEs within Feliformia with specific focus on the Felidae. Thirty-seven cat species augmented by representatives from related Feliform represent roughly 44 million years (MY) of divergence (see Additional file 1: Table S1) [40]. The extant cat species diverged into eight lineages in a nearly starburst pattern over 10 MY [66], and have largely maintained synteny in chromosome architecture [10]. Roughly 10-11% of a felid genome is comprised of SINEs [10]. We identified 93 new CanSINE loci, which were divided into quiescent and active subfamilies. In addition, we found empirical evidence of the effects of rapid speciation and imprecise SINE excision on phylogenetic consistency.

Results

We applied both *in silico* genome mining and PCR-based approaches to identify feliform CanSINE loci, which were then sequenced in 37 extant Felidae species and five additional representatives from Prionodontidae, Viverridae, Herpestidae, and Hyaenidae. First, direct *in silico* genome annotation of the domestic cat (*F. catus*), verified against the dog genome (*Canis familiaris*), identified 29 new CanSINE loci (see Additional file 2: Table S2A). Second, a SINE-to-SINE PCR method [67] isolated another 30 SINE-flanked genomic regions in exotic felids (see Additional file 2: Table S2B). Among the 59 total amplified regions, 21 (35%) included two or more independent insertions in Feliformia species. Together these represent 93 previously uncharacterized CanSINE loci (Additional file 3: Tables S3 and Additional file 4: Table S4).

CanSINE insertion hotspots

CanSINEs from different lineages targeted homologous loci during proliferation and retrotransposition within the genome. At least three inserts were found in 8 of the 21 multiple-insert loci (62%) in Felidae (See Additional file 5: Table S4). For example, inserts at locus 133135 occurred in unrelated *Lynx rufus*, *Profelis caracal* and *Pardofelis marmorata*, along with a synapomorphic insertion shared in the seven species of the ocelot lineage (Figure 1, see Additional file 5: Figure S1, see Additional file 4: Table S4). Each of these four CanSINEs was flanked by species-specific, overlapping target site duplication sequence (TSD). Independence of the four insertion

Figure 1 CanSINE insertions at locus 133135. Arrangement of 4 unique CanSINE insertion events occurring at locus 133135 in the caracal (*Profelis caracal*), marbled cat (*Pardofelis marmorata*), ocelot lineage (genus: *Leopardus*) and bobcat (*Lynx rufus*) with the homologous *F. catus* sequence as a reference. The *L. rufus* SINE is in reverse orientation. The independent insertions have overlapping target site duplications (boxed).

events is verified by multiple nucleotide indels in the microsatellite and poly A/T segments. Furthermore, the *L. rufus* SINE is in the reverse orientation. Similarly, locus 212075 contained six independent insertion events including: 1) a shared synapomorphy defining the bay cat lineage, 2) a shared synapomorphy of *P. caracal/P. aurata* and 3) autapomorphic insertions in *Felis nigripes*, *P. rubiginosus*, *P. bengalensis* and *P. planiceps*. In the latter case, insertions in *P. bengalensis* (n = 9) and *P. planiceps* (n = 7) were unfixed (Figure 2, see Additional file 6: Figure S2, see Additional file 4: Table S4,).
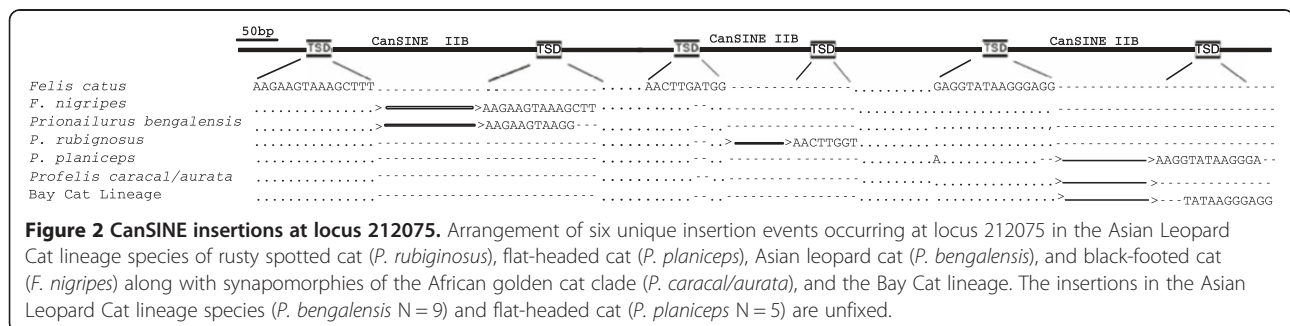
An examination of patterns of sequence divergence of both tRNA and genomic flanking regions suggests the insertions at 212075 occurred independently among species. In *F. nigripes* and *P. bengalensis* CanSINEs were flanked by different TSDs and the percent identity was 81.6% within the SINE regions compared to 96.4% in the regions flanking the SINE (Figure 2, see Additional file 6: Figure S2, see Additional file 2: Table S2). Similarly, *P. planiceps* and the bay cat lineage CanSINEs are flanked by different, but overlapping, TSDs and the percent identity was 82.8% within SINE regions compared to 97.2% in the 126 nucleotides flanking the SINE. While these sequence diversity estimates do not definitively preclude post insertion mutations, they are consistent with independent retrotransposition events by unique RNA templates.

Evolutionary assessment of feliform CanSINEs

Based on alignment and phylogenetic reconstruction of conserved tRNA regions, we identified major CanSINE

lineages defined by distinct motifs, which we have designated as subfamilies and subtypes (Figures 3 and 4). Subfamily I members share a diagnostic 'TCCTGAT' motif at position 36 within the 5' tRNA-related region. Additional variants within the tRNA-related region of 'CA' or 'GT' at position 116 and 'GGGA' or AAGA' at position 138 were diagnostic for subtypes IA and IB respectively (Figure 3). Loci in subfamily II share a 'GG CTCGG' motif at position 118 within the tRNA region and subtypes IIA and IIB are delineated by an insertion/deletion ('T') at position 51 within the 5' tRNA-related region (Figure 3). Notably, there is an A > T polymorphism at position 70 within the RNA polymerase B box that segregates nearly perfectly with subfamilies I and II and a T > G polymorphism in the RNA polymerase A box that is specific to subtype IIB (Figure 3). In addition, published SINE voucher sequences annotated from *F. catus* clustered within the two subtypes of subfamily II [i.e. SINEC_Fc1 grouped with subtype IIB and SINEC_Fc2 grouped with subtype IIA (Figure 4)].

Phylogenetic differences between the two subfamilies are concordant with ancestral versus recent nodes within the Feliformia species tree. The most ancestral CanSINE lineage is subfamily I, composed of loci conserved among the 6 families in Feliformia and thus likely arose in a common ancestor ~50 MYA (Figure 5). Within subfamily II, four of the 10 subtype IIA CanSINEs arose ~35 MYA in a common ancestor of Felidae and Priondontidae, in the progenitor of Felidae or early in the initial Felidae radiation, while the remaining six are scattered among more recent lineages (Figure 4). In contrast, all 49 of subtype
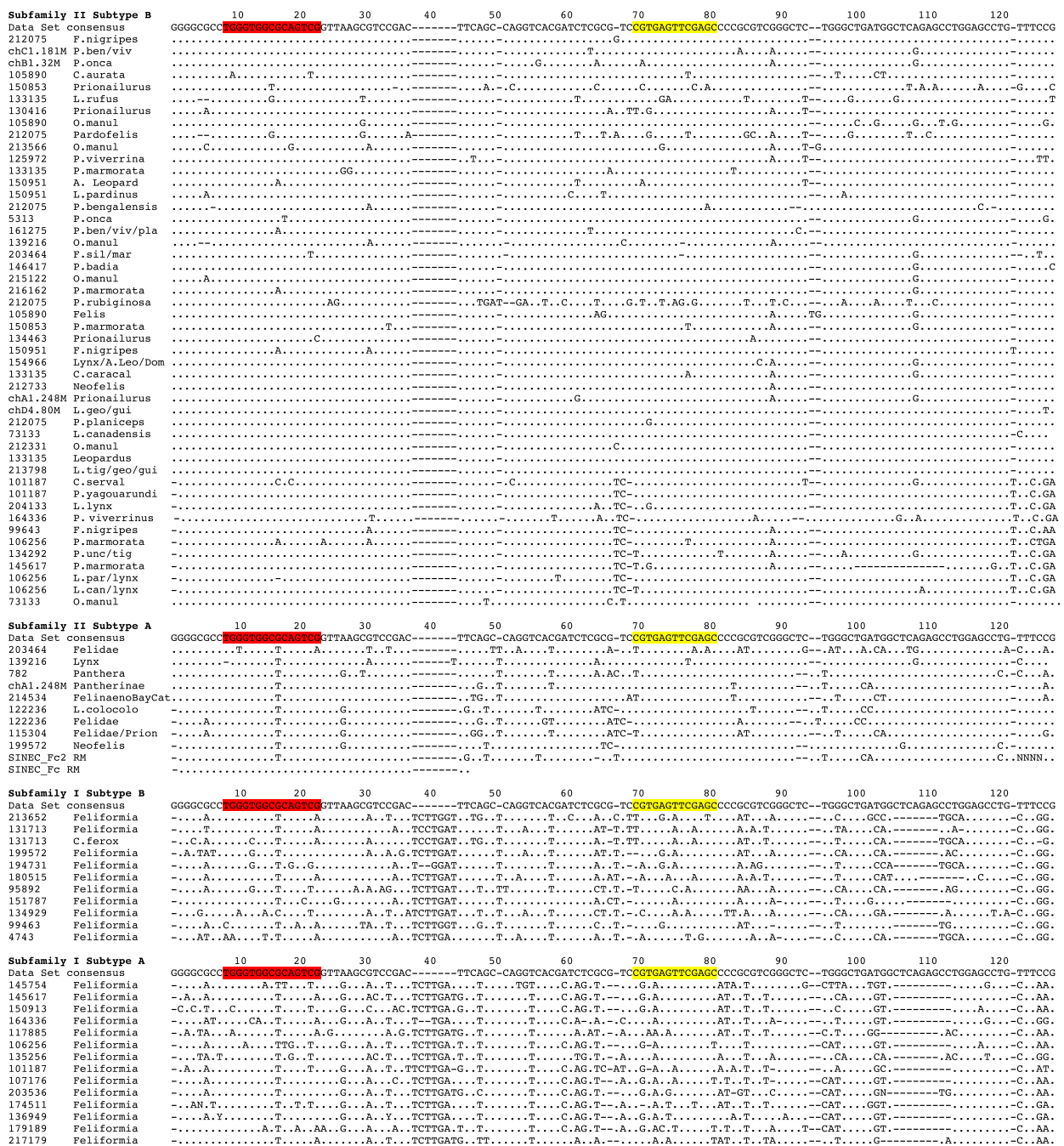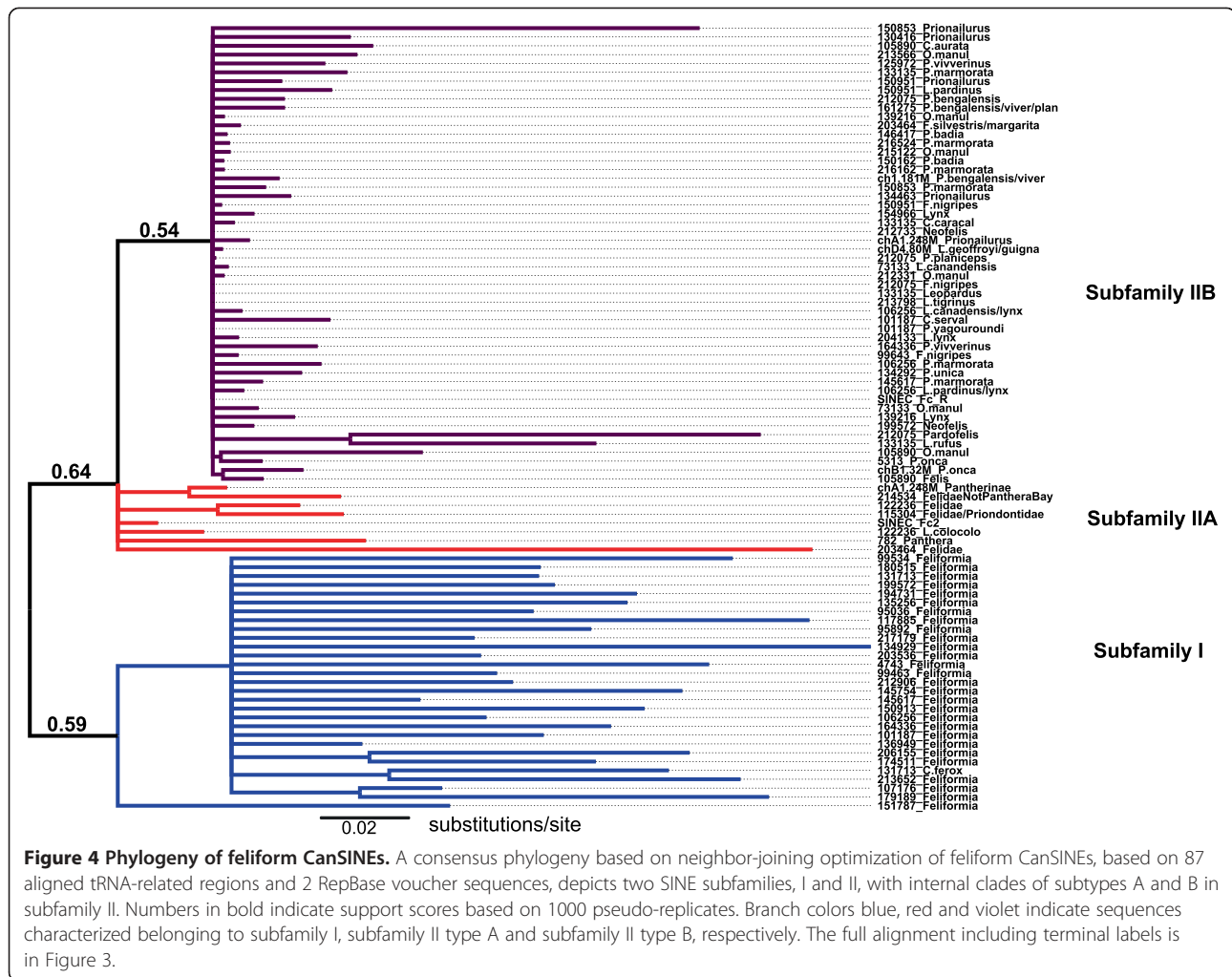

Figure 2 CanSINE insertions at locus 212075. Arrangement of six unique insertion events occurring at locus 212075 in the Asian Leopard Cat lineage species of rusty spotted cat (*P. rubiginosus*), flat-headed cat (*P. planiceps*), Asian leopard cat (*P. bengalensis*), and black-footed cat (*F. nigripes*) along with synapomorphies of the African golden cat clade (*P. caracal/aurata*), and the Bay Cat lineage. The insertions in the Asian Leopard Cat lineage species (*P. bengalensis* N = 9) and flat-headed cat (*P. planiceps* N = 5) are unfixed.

Figure 3 Alignment of tRNA-related SINE subregions found among Feliformia. Diagnostic indels are attributed to SINE subfamily and subtype distinctions. The RNA polymerase III boxes are indicated in red (A) and yellow (B).

IIB insertions are localized to individual felid clades or are unique to a single species and thus likely arose within the last 5 million years (Figure 5).

If CanSINE proliferation and subsequent sequence divergence is correlated with evolutionary time, then more ancient inserts will have greater nucleotide variation then those of recent origin. The more ancestral subfamily I is three times more diverse (0.298 substitutions/site) than subfamily II (0.090 substitutions/site) (Table 1). The most variable CanSINE was subtype IA (0.271 substitutions/site) and the least variable was subtype IIB (0.068 substitutions/site) (Table 1). These results suggest that measures of average sequence divergence observed in CanSINE lineages, calibrated by the feliform phylogeny, are estimates of time since periods of active proliferation.

Figure 4 Phylogeny of feliform CanSINEs. A consensus phylogeny based on neighbor-joining optimization of feliform CanSINEs, based on 87 aligned tRNA-related regions and 2 RepBase voucher sequences, depicts two SINE subfamilies, I and II, with internal clades of subtypes A and B in subfamily II. Numbers in bold indicate support scores based on 1000 pseudo-replicates. Branch colors blue, red and violet indicate sequences characterized belonging to subfamily I, subfamily II type A and subfamily II type B, respectively. The full alignment including terminal labels is in Figure 3.
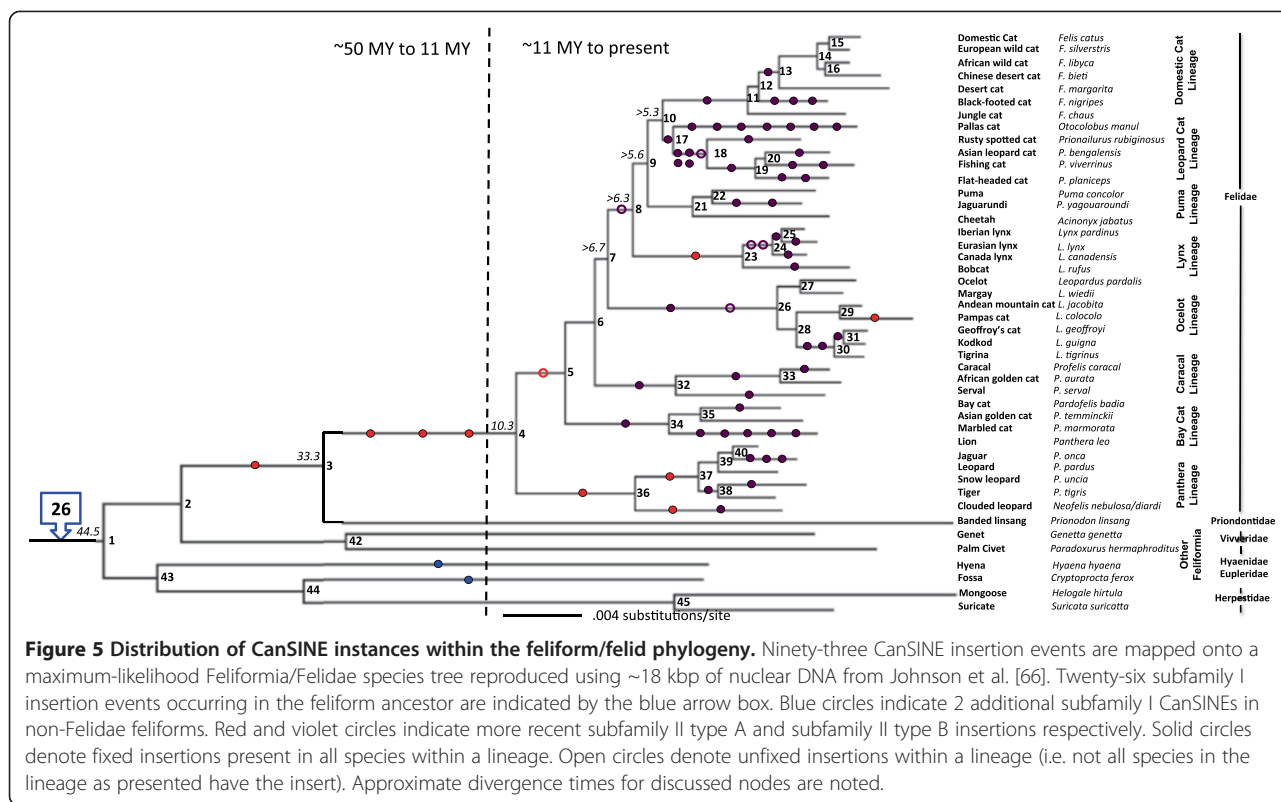
CanSINE evolution in the cat family Felidae

The phylogenetic fidelity of the 93 CanSINE loci varied among the hierarchical nodes within Feliformia. The 33 species-specific loci were distributed among the eight major Felidae lineages. An additional 26 Can-SINEs supported the monophyly of Feliformia (Figure 5: node 1). Another three insertions supported the monophyly of Felidae and one CanSINE locus supported the sister group relationship between Felidae and Priondontidae (Figure 5: nodes 3 and 4). Two unique insertions were found in non-Felidae representatives of *H. hyaena* and *C. ferox.* Twenty CanSINE loci were diagnostic for internal clades within Felidae while seven of the eight major felid lineages had diagnostic loci (Figure 5: nodes 36, 34, 32, 26, 23, 17 and 11 respectively). Intra-lineage markers included loci defining clades within the panthera, Asian leopard cat, caracal, ocelot and felis lineages (Figure 5: nodes 13, 18, 19, 30, 31, 33, 37 and 38).

Discordant phylogenetic inferences correlate with polymorphic loci

The 93 CanSINE loci presented here were mapped to a phylogeny based on multiple optimality criterion described by Johnson et al. (Figure 5) [66]. However, alternate branch topologies and phylogenetic ambiguities are indicated by six of the 93 CanSINE loci. In the lynx lineage, an orthologous insertion at locus 106256 was homozygous in all *L. pardinus* individuals while a second independently derived insertion, occurring 315 bp downstream of the first, was homozygous in *L. canadensis.* These orthologous insertions are polymorphic within *L. lynx*, with individuals either homozygous for one of the insertions or heterozygous, containing one copy of each (Figure 6A, see Additional file 7: Figure S3). No correlation was observed between the geographic origin of *L. lynx* individuals and CanSINE profile (see Additional file 8: Table S5). Further ambiguity of the *Lynx* genus topology was indicated by the presence of CanSINE locus 134463 in

Figure 5 Distribution of CanSINE instances within the feliform/felid phylogeny. Ninety-three CanSINE insertion events are mapped onto a maximum-likelihood Feliformia/Felidae species tree reproduced using ~18 kbp of nuclear DNA from Johnson et al. [66]. Twenty-six subfamily I insertion events occurring in the feliform ancestor are indicated by the blue arrow box. Blue circles indicate 2 additional subfamily I CanSINEs in non-Felidae feliforms. Red and violet circles indicate more recent subfamily II type A and subfamily II type B insertions respectively. Solid circles denote fixed insertions present in all species within a lineage. Open circles denote unfixed insertions within a lineage (i.e. not all species in the lineage as presented have the insert). Approximate divergence times for discussed nodes are noted.

all *L. canadensis* and *L. lynx* individuals with absence from *L. pardinus* (Figure 6a, See Additional file 9: Figure S4).

Likewise, CanSINEs did not always map to expected species associations within the South American ocelot lineage. CanSINE locus 133135 is fixed in *Leopardus pardalis, L. jacobita, L. tigrina, L. guigna* and *L. geoffroyi*, yet absent in *L. wiedii* and *L. colocolo* (Figure 6B, see Additional file 10: Figure S5). In the Asian leopard cat lineage, locus 161275 is polymorphic in *P. rubiginosus* and *P. bengalensis,* fixed in *P. viverrinus* and absent from *P. planiceps* (Figure 6c, see Additional file 11: Figure S6).

Table 1 Mean genetic distance (substitutions/site) of feliform SINE subfamilies and subtypes with standard deviation

SINE type	Genetic distance	SD
All SINEs	0.205	0.133
Subfamily I	0.298	0.083
Subfamily II	0.090	0.068
Subtype IA	0.271	0.079
Subtype IB	0.189	0.045
Subtype IIA	0.137	0.066
Subtype IIB	0.068	0.052

Using an unpaired T-test, differences in the mean genetic distances between each subfamily and subtypes within each family were significant with p <0.0001.

To account for the possibility that the CanSINE profiles described above are the result of recent hybridization events between closely related species, mitochondrial profiles at the *NADH5* gene were obtained from all individuals representing the lynx, ocelot and Asian leopard cat lineages. We found most mtDNA haplotypes to be consistent with species designation and the previously proposed phylogenetic relationships among the Felidae species. The exception was *P. rubiginosus*, which had differing *NADH5* haplotypes identical to those found among *P. bengalensis* (see Additional file 11: Figure S6).

Two CanSINE loci that may be mapped to the "backbone" of the Felidae tree were also inconsistent with prior estimations of the initial Felidae radiation. Locus 154966 is present in all species of the domestic cat, Asian leopard cat and lynx lineages, and absent in the puma, ocelot, caracal, bay cat and panthera lineages (Figure 5: nodes 4–10, see Additional file 12: Figure S7). Locus 214534 is present in all Felidae species except those of the caracal and panthera lineages (Figure 5: nodes 4–10, see Additional file 13: Figure S8).

Evidence for SINE Excision
CanSINE locus 174511 is present in all feliform taxa with one exception. In *Puma concolor* locus 174511 includes an 18 bp reverse-oriented SINE fragment rather than a full-length SINE and mapped to 63 bp of upstream
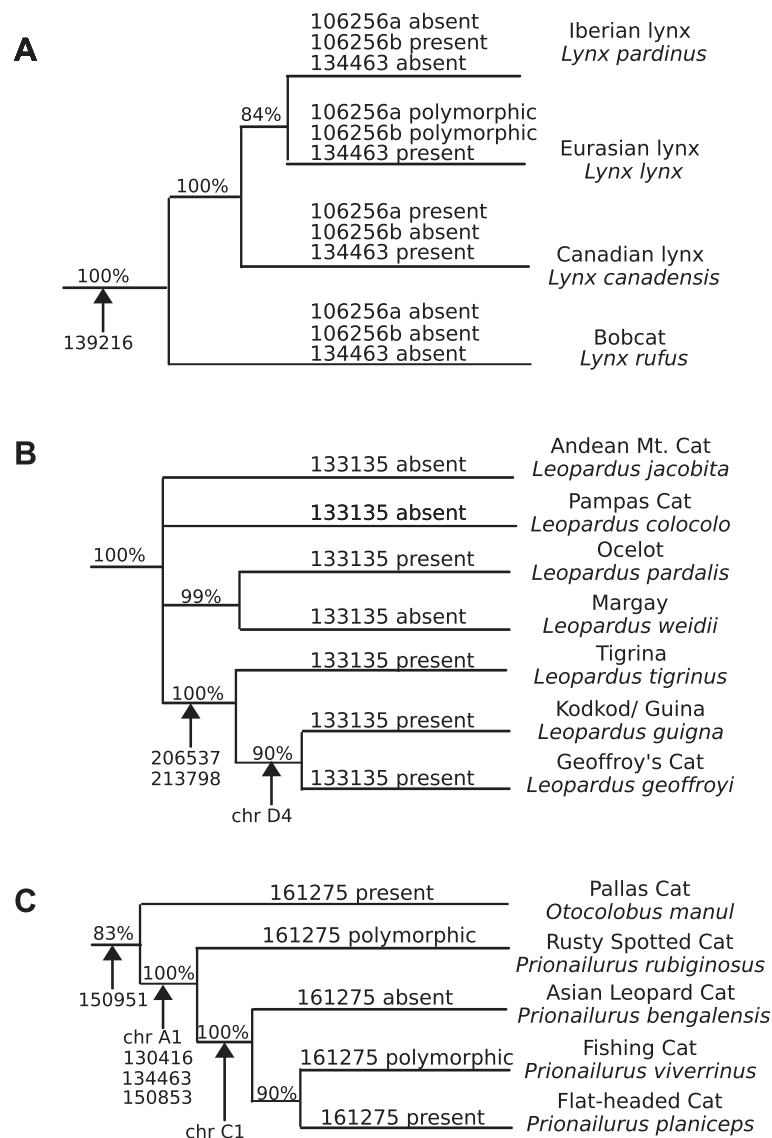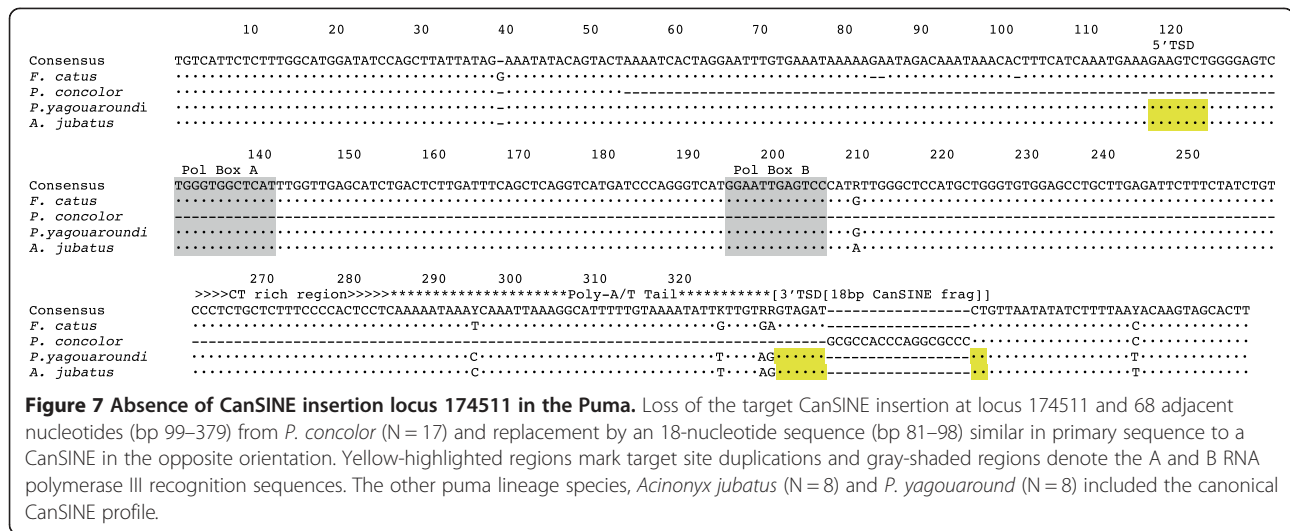
A

106256a absent
106256b present
134463 absent — Iberian lynx *Lynx pardinus*

84%
106256a polymorphic
106256b polymorphic
134463 present — Eurasian lynx *Lynx lynx*

100%
106256a present
106256b absent
134463 present — Canadian lynx *Lynx canadensis*

100%
139216

106256a absent
106256b absent
134463 absent — Bobcat *Lynx rufus*

B

133135 absent — Andean Mt. Cat *Leopardus jacobita*

133135 absent — Pampas Cat *Leopardus colocolo*

100%

99%
133135 present — Ocelot *Leopardus pardalis*

133135 absent — Margay *Leopardus weidii*

100%
133135 present — Tigrina *Leopardus tigrinus*

206537
213798

90%
133135 present — Kodkod/ Guina *Leopardus guigna*

133135 present — Geoffroy's Cat *Leopardus geoffroyi*

chr D4

C

83%
161275 present — Pallas Cat *Otocolobus manul*

161275 polymorphic — Rusty Spotted Cat *Prionailurus rubiginosus*

150951

100%

chr A1
130416
134463
150853

100%
161275 absent — Asian Leopard Cat *Prionailurus bengalensis*

90%
161275 polymorphic — Fishing Cat *Prionailurus viverrinus*

161275 present — Flat-headed Cat *Prionailurus planiceps*

chr C1

Figure 6 CanSINE insertion sites incongruent with prior phylogenic analyses. The model topologies shown are based on maximum-likelihood reconstruction using 18,853 bp of nuclear DNA, with bootstrap scores noted to the left of each node and divergence time estimates between nodes in gray (Johnson et al 2006) [66]. **A)** An insertion at locus 106256 is present in *L. canadensis* (N=22), polymorphic in *L. lynx* (N=23) and absent in *L. pardinus* (N=8), while another insertion near the same site is absent from *L. canadensis*, polymorphic in *L. lynx* and present in *L. pardinus*. A third insertion at locus 134463 is present in all *L. canadensis* and *L. lynx* and absent in all *L. pardinus*. **B)** An insertion at locus 133135 has a paraphyletic distribution among the ocelot lineage species; *Leopardus pardalis* (N=10), *L. jacobita* (N=2), *L. tigrina* (N=9), *L. guigna* (N=3) and *L. geoffroyi* (N=11), *L. wiedii* (margay, N=6) and *L. colocolo* (pampas cat, N=3). The placement of *L. jacobita* and *L. colocolo* with respect to the remaining *Leopardus* species has yet to be determined with statistical confidence and thus is depicted here as a polytomy. **C)** An insertion at locus 161275 is present in all *P. viverrinus* (N=7), absent in *P. planiceps* (N=5) and polymorphic among *P. bengalensis* (N=9) and *P. rubiginosus* (N=2). However, 4 other SINE insertion sites support the monophyly of *Prionailurus* and one insertion supports the monophyly of *P. bengalensis*, *P. viverrinus* and *P. planiceps*.

sequence (Figure 7). By contrast, the full-length SINE locus 174511 is fixed in the two puma lineage sister species, *Acinonyx jubatus* and *P. yagouaroundi*.

Discussion

Genomic characterization of 93 novel CanSINEs in Feliformia clarifies, amends and extends existing hypotheses on SINE evolution and strongly supports the

phylogenetic fidelity of these retrotransposons. In using the well-supported phylogeny of the cat family, Felidae, as a reference species tree, we provide empirical evidence for long speculated, but rarely observed, processes such as co-evolution of SINE families with the host genome, targeted insertion during CanSINE proliferation, lineage sorting of ancestral polymorphisms among closely related species, and instances of SINE excision from the genome.

Figure 7 Absence of CanSINE insertion locus 174511 in the Puma. Loss of the target CanSINE insertion at locus 174511 and 68 adjacent nucleotides (bp 99–379) from *P. concolor* (N = 17) and replacement by an 18-nucleotide sequence (bp 81–98) similar in primary sequence to a CanSINE in the opposite orientation. Yellow-highlighted regions mark target site duplications and gray-shaded regions denote the A and B RNA polymerase III recognition sequences. The other puma lineage species, *Acinonyx jubatus* (N = 8) and *P. yagouaround* (N = 8) included the canonical CanSINE profile.

CanSINE integration targets homologous loci

The discovery of 93 CanSINE loci includes a high frequency of multiple insertions within orthologous intergenic regions. Some loci serve as apparent "hotspots" of CanSINE activity within the Felidae. For example, CanSINE locus 133135 displays four independent insertion events defined by different, yet overlapping TSDs. Likewise, locus 212075 supports six independent insertions, three of which occur in a single genus, *Prionailurus*. Similar patterns of CanSINE integration have been observed in the Caniformia suborder wherein amplification of five putative *C. familiaris* CanSINE loci revealed eight additional insertions in related species [40], and amplification of 13 intronic segments amongst caniforms revealed 26 independent insertion events [58]. Possible explanations for the likelihood of additional CanSINEs co-occurring at orthologous loci involve signature motifs associated with the L1 long interspersed element (LINE) derived endonuclease [1]. In primates, integration of SINEs (*Alu* repeats) is facilitated by the motif TTAAAA $(N)_{0-8}$TYTNR [68]. A similar mechanism is hypothesized in whole genome assessments that found over 20% of *C. familiaris* CanSINE integration sites include a TTAAAA motif [1,69]. Likewise, CanSINE integration sites within Feliformia share similar AT-rich motifs (Tables S3 and S4) indicating target site preferences [53].

Dynamic evolution of Feliform CanSINE lineages

Beyond the initial reporting of voucher sequences (see Repbase database www.girinst.org/repbase) within the domestic cat whole genome sequence [8], little is known of CanSINE evolution in Feliformia [70]. Until now, even the most current SINE resource (SINEbase, http://sines.eimb.ru) includes only one feliform specific voucher [1]. Here, phylogenetic analysis of the conserved tRNA-like regions of the 93 newly described CanSINEs reveal two distinct subfamily lineages defined by time of emergence within Felidae and further differentiated into subtypes marked by specific sequence motifs and adaptive reverse transcriptase promoter sites (Figures 3 and 4).

Subfamily I likely originated roughly 45–60 MYA when the Carnivore order first split to form two major lineages of Feliformia and Caniformia [60]. Subfamily II arose during the emergence of sister families of Priono-dontidae and Felidae, with continuous diversification until present-day Felidae. The relatively smaller mean genetic distance among subfamily II CanSINEs compared to subfamily I (0.090 substitutions/site versus 0.298 substitutions/site) is consistent with subfamily II insertions being acquired more recently from either a single master copy or multiple yet similar master copies, whereas subfamily I insertions are derived from a now quiescent set of master copies and have since accumulated substitutions.

Historic and ongoing patterns of CanSINE proliferation can be inferred by both position with the feliform phylogeny and the extent of sequence divergence among loci. The more basal subfamily I is comprised of subtypes IA and IB, which each arose in a common ancestor to Feliformia (Figure 5). Significant genetic distance estimates for subtype IA and IB (Table 1) imply that each lineage may have originated from different master copies and that subtype IA may have proliferated before IB. Within subfamily II, subtype IIA master copy or copies may have had an ancient origin, inserting into a common ancestor of Felidae and Priondontidae (Figure 5). However, this subtype apparently remains a source of extant species-specific insertions as indicated by species-specific presence in *L. colocolo* and *N. neofelis* (Figure 5). Subtype IIB CanSINEs are more recent, not as genetically diverse as IIA loci, and are the source of most phylogenetically informative sites in extant Felidae (Table 1, Figure 5).

While the genetic distances among each CanSINE subfamily provide substantial evidence for a progressive evolution of CanSINEs from the Feliform ancestor to present, the phylogenetic support scores for subfamily I, subtype IIA and subtype IIB remain relatively low, 50-65%. In addition, subfamily I cannot be resolved into subtypes based on consensus of 1000 minimum evolution replicates (Figure 4). A possible explanation for this lack of resolution could be the existence of multiple master copies that can concurrently convey insertional mutagenesis, leading to the paraphyletic pattern observed in the CanSINE phylogeny. This mechanism, also known as the 'sprout' model, has been proposed for human *Alu*s and allows for secondary master copies to provide a minor portion of a subfamily's members [71].

In addition, CanSINE subfamilies have distinctive polymorphisms in the pol A and pol B RNA polymerase III binding sites that may indicate adaptive evolution. As non-autonomous transposable elements, changes in host polymerase specificity can cause SINE quiescence or adaptation [72,73]. Here, the A > T mutation in polymerase box B of the recent subfamily II and the T > G transversion in polymerase box A of subfamily IIB is not observed in the more ancestral subfamily I and thus could be evidence of functional adaptation driving ongoing subfamily II proliferation (Figure 3). However, RNA polymerase III A and B boxes are known to contain degenerate sites [74] and evidence of adaptive evolution during speciation, as opposed to accumulation of random mutations, awaits further sequence analyses of RNA polymerases in Felidae.

Deciphering CanSINE proliferation against a backdrop of rapid speciation

SINEs are generally viewed as ideal markers of genetic divergence and phylogenetic reconstruction [52,75,76]. However, inconsistencies between SINE-based results and other molecular data may occur and are tangible evidence of complex speciation events, revealing dynamic evolutionary histories. SINEs can provide an advantage over SNP-based molecular phylogenetic analyses, wherein determining inconsistency due to homoplasy (i.e. parallelism or multiple-hits) versus hemiplasy (i.e. lineage-sorting) is ambiguous [77]. Here, the Felidae reference species tree as a framework for SINE evolution is robust [66] while the few alternate topologies [42,66,78-80], provide an opportunity to test the accuracy CanSINEs as cladistic markers during rapid speciation.

Evolution of modern Felidae is marked by a nearly star-burst pattern of speciation from a common ancestor approximately 10 MYA [60]. As such, CanSINE analyses presented here reveal limitations to correct phylogenetic interpretations even at higher-order nodes within the topology. For example, the insertion at locus 154966 suggests that the lynx lineage (Figure 5: node 8) is more recently derived than the puma lineage, which is consistent with prior minimum evolution, maximum parsimony and Bayesian analysis, yet inconsistent with maximum likelihood reconstructions [66]. Similarly, the insertion at locus 214534 suggests a more basal position of the caracal lineage within Felidae rather than the bay cat lineage, while previous phylogenetic reconstructions place the bay cat lineage at a more basal position than the caracal lineage (Figure 5: node 6), with statistical support from 50-100% depending on the optimality criterion [66]. The insertion patterns at loci 154966 and 214534 can be attributed to the nearly simultaneous divergences of the lynx and puma lineages ~7 MYA and the bay cat and caracal lineages ~9 MYA [66], resulting in "ancient" incomplete lineage sorting, a phenomenon previously observed in SINE profiles of cichlid species that diverged during a similar span of time, ~5–10 MY [81].

Similarly, rapid evolution has resulted in mosaic SINE profiles that reflect complex intra-lineage speciation patterns. In the ocelot lineage, *L. jacobita* and *L. colocolo* diverged within 20,000 years from the stem lineage (Figure 5: nodes 26–27, Figure 6B), and *L. tigrinus*, *L. guigna* and *L. geoffroyi* all arose within a brief 20,000-year interval (Figure 5: nodes 30–31, Figure 6B) [66,82]. In the lynx lineage, 40,000 years separates the *L. canadensis*, *L. lynx* and *L. pardinus* species complex (Figure 5: nodes 24–25, Figure 6A). Likewise, in the Asian leopard cat lineage *P. bengalensis*, *P. viverrinus* and *P. planiceps* diverged within a 40,000-year interval (Figure 5: nodes 19–20, Figure 6C) [66,83,84]. In addition, documented instances of ongoing hybridization between species in the wild further complicate phylogenetic analyses and taxonomy [82,84,85].

These instances of rapid speciation in Felidae are correlated with incomplete lineage sorting of ancestral polymorphisms among CanSINE loci. In the lynx lineage, maximum likelihood phylogeny derived from concatenated segments of nuclear DNA indicate *L. lynx* and *L. pardinus* are sister taxa [66,83], contrary to recent Bayesian reconstructions including mitochondrial DNA [66,78] that support a more basal position of *L. pardinus* with respect to *L. lynx* and *L. canadensis*. CanSINE distributions described here reflect the nearly simultaneous and successive speciation of the lynx, a process observed repeatedly amongst mammalian lineages [59,86]. In this instance, rapid divergence resulted an ancestral polymorphism at locus 106256 becoming fixed for presence or absence in *L. canadensis* and *L. pardinus* while remaining polymorphic in *L. lynx*. In contrast, a fixed insertion at locus 134463 supports a sister taxa relationship between *L. canadensis* and *L. lynx* (Figure 6A). Additional evidence, possibly from upcoming whole-genome efforts, should reveal a more comprehensive view of lynx phylogeny [87].

Previous analyses also failed to fully resolve the phylogenetic position of *L. jacobita* and *L. colocolo* within the ocelot lineage. Depending on the molecular data types examined and the optimality criterion employed, these two species have been placed as sister taxa or as belonging to other clades within the ocelot lineage [66,82,88]. Hence, whether the presence CanSINE locus 133135 in *L. jacobita* is due to incomplete lineage sorting of a CanSINE that was present in the *Leopardus* ancestor or due to a closer evolutionary relationship between *L. jacobita* and the *L. tigrinus*, *L. geoffroyi* and *L. guigna* clade, rather than *L. colocolo*, cannot be determined (Figure 6b). Intraspecific single nucleotide polymorphisms (SNPs) present in the *L. pardinus* 133135 locus indicate the insertion was present during the genesis of this species and not inherited more recently through hybridization (see Additional file 10: Figure S5).

In some instances CanSINEs reflect ongoing and ancestral episodes of hybridization in Felidae. For example an orthologous insert at locus 161275 in *P. rubiginosus*, *P. bengalensis,* and *P. viverrinus* to the exclusion of *P. planiceps* is incongruent with prior strongly supported species associations and is in direct conflict with a fixed insertion site at chromosome C1 diagnostic of the *P. bengalensis*/*P. planiceps*/*P. viverrinus* clade [66,84,89] (Figure 5: nodes 18–19, Figure 6c). Notably, the two heterozygous *P. rubiginosus* CanSINE sequences differ yet are each identical to CanSINE 161275 copies in *P. bengalensis.* This in conjunction with the *P. rubiginosus NADH5* haplotype, indicates hybridization with *P. bengalensis* after the initial radiation of *Prionailurus* (see Additional file 11: Figure S6).

Further, *P. bengalensis* serves as a model of an ongoing SINE fixation process. *P. bengalensis* is divided into two putative subspecies that diverged ~2.5 MYA: a 'northern' population on the Asian mainland and a 'southern' population on the Malay Peninsula, [84,89]. The four individuals examined from the northern population are polymorphic at locus 161275, compared with four southern homozygous individuals. Albeit a small sample size, the data suggest that the populations differ in CanSINE fixation at locus 161275 and is perhaps linked with ongoing genetic drift.

Overall, our findings suggest that rapid speciation results in mosaic genomes with conflicting phylogenetic signals [43,86]. In such instances a polytomy or split network, which recognizes shared alleles between paraphyletic groups, may be a more accurate depiction of evolutionary history. As with large scale genome sequences, CanSINE data did not unequivocally resolve the Felidae into a series of bifurcating lineages, a pattern seen even in the reconstruction of basal mammalian lineages [59,90].

SINE locus loss
Although rarely observed, perfect or near-perfect SINE excision can occur via inter or intra chromosomal recombination between insertions of the same SINE family or between flanking TSDs [9,23,21]. The excision of locus 174511 in *P. concolor,* marked by an inverted 18 bp segment, is consistent with a mechanism of non-homologous recombination. Alternatively, simple repeats that surround the insertion site may have formed a loop structure that was omitted during DNA replication (Figure 7) leading to excision. Similar evidence of SINE removal occurs in other vertebrate lineages, such as in the squamate *Darevskia* subspecies [76] and primates [23,21].

Conclusions
The availability of whole genome sequences has dramatically increased our understanding of mammalian non-coding DNAs. By employing comparative genomics methods to identify SINE loci in domestic and exotic feliforms, two feliform-specific CanSINE subfamilies were defined based on sequence structure and taxonomic distribution. Identification of a currently active SINE subfamily with Felidae will provide opportunities to test hypotheses about the role of CanSINEs in somatic functional diversity. Patterns of insertion also support species designations, affirming CanSINEs as systematic markers and confirming complex evolutionary processes including incomplete lineage sorting following rapid species divergence, hybridization and SINE mediated genome rearrangement.

Methods
CanSINE distribution was assessed in one or more individuals representing each of the extant Felidae species including four subspecies of the domestic cat complex, *F. silvestris*. We also examined representative samples from five additional Feliformia families, Prionodontidae, Hyaenidae, Herpestidae, Eupleridae and Vivveridae. Taxa are listed in Additional file 1: Table S1. Commercial genomic DNA from *F. catus* was purchased from EMBD Biosciences Product No: 69235. Genomic DNA for the remaining taxa was extracted from blood and/or tissue samples using the Qiagen DNeasy Blood & Tissue Kit. All tissue samples for the Laboratory of Genomic Diversity were collected in full compliance with specific Federal Fish and Wildlife permits from the Conservation of International Trade in Endangered Species of Wild flora and Fauna: Endangered and Threatened Species, Captive Bred issued to the National Cancer Institute (NCI)-National Institutes of Health (NIH) (S.J.O. principal officer) by the U.S. Fish and Wildlife Services of the Department of the Interior.

Genome mining
From a list of 322 felid SINEs identified during the initial *F. catus* whole genome annotation [8], select loci were retrieved from the March 2006 genome assembly on the

UCSC genome browser (http://genome.ucsc.edu) and matched to corresponding cat chromosome locations using a *F. catus* genome browser, GARField (formerly at http://lgd.abcc.ncifcrf.gov/cgi-bin/gbrowse/cat/) [91]. Within the context of this study, each region is named for the UCSC genome browser scaffold from which the reference sequence was obtained (see Additional file 2: Table S2). Sixty regions containing feliform CanSINEs found in the *F. catus* whole genome sequence with homologous flanking sequence in *C. familiaris* were selected for amplification in all extant felids and five feliform outgroup taxa. Forward and reverse PCR primers were designed within 300 bp of the putative SINE insertion sites.

Direct PCR, sequencing and cloning

Approximately 20 ng of extracted genomic DNA was used in each PCR reaction. All reactions consist of 0.1U of AmpliTaq DNA polymerase, 0.75 μM forward and reverse primer, 2.5 mM $MgCl_2$, 0.2 mM of each deoxynucleotide triphosphate and the appropriate amount of 10X AmpliTaq Buffer II and water for a 20 μl reaction. Touchdown PCR conditions were 5 min at 94°C, 10 cycles of 30 sec at 94°C, 30 sec at 63°C* and 60 sec at 72°C, with a decrease in the annealing temperature at a rate of 0.5°C per cycle, followed by 30 cycles of 30 sec at 95°C, 30 sec at 58°C** and 60 sec at 72°C, then a final elongation step of 7 min at 72°C. **Final annealing temperatures varied from 50-64°C depending on the primer set. *Initial annealing temperatures were set to 5°C warmer than the final annealing temperature. To confirm amplification and assess the sizes of DNA fragments, 5 μl of PCR product was fractionated by gel electrophoresis in a 1.0% agarose gel containing ethidium bromide. Prior to cloning or sequencing, 20 μl of PCR product was purified using the ExoSAP protocol with 0.72 μl shrimp alkaline phosphatase (SAP) and 1.44 μl exonuclease I (ExoI) (Amersham Pharmacia, Piscataway, NJ).

Cycle sequencing reactions consisted of 0.25U BigDye® Terminator v3.1 Ready Reaction Mix, 0.075 μM primer, 5 μl of sequencing buffer (Applied Biosystems), 1.5 μl of purified PCR product and enough water for a 10 μl reaction. Cycle sequencing was performed under the following conditions: 94°C for 10 sec, 52°C for 5 sec, and 72°C for 2 min for 45 cycles. Products from cycle sequencing reactions were run on an ABI 3730 DNA Analyzer. Sequence results were visualized and edited with Sequencher v4.8 (GeneCodes).

Multiple gel electrophoresis bands or illegible preliminary sequencing traces were resolved by cloning PCR amplification products with the TOPO TA Cloning Kit (Invitrogen) followed by purification with the Qiagen GeneClean Kit according to manufacturer's instructions. Cycle sequencing of the purified fragments was performed using 0.25U BigDye® Terminator v3.1 Ready Reaction

Mix, 1 μl of forward or reverse M13 primer provided in the TOPO TA Cloning Kit, 5 μl of sequencing buffer (Applied Biosystems), 2.5 μl of purified PCR product and enough water for a 10 μl reaction. Cycle sequencing was performed under the following conditions: 94°C for 10 sec, 52°C for 5 sec, and 72°C for 4 min for 45 cycles.

Scanning via SINE-to-SINE PCR and Cloning

A second SINE discovery method was adapted from a SINE-to-SINE amplification protocol [67] to allow identification of novel SINE loci in exotic Felidae species. Similar methods have been applied to illuminate human *Alu* loci [92,93]. Primers were developed that anneal to diagnostic motifs within the tRNA-related region of feliform CanSINEs: primer 1 (ATCAGACTCTTGATT TCAGCTCA), primer 2 (AGCTCAGGTCATGATCCCA GG), primer 3 (TCCGACTTCAGCCAGGTC), primer 4 (TGATGGCTCGGAGCCT) and primer 5 (TCCGACTT CGGCTCAGGTC). Single primer PCR was performed on approximately 20 ng of extracted genomic DNA from eight species representing the major Felidae lineages: *Neofelis nebulosa, Panthera onca, Pardofelis marmorata, Pardofelis badia, Leopardus guigna, Leopardus rufus, Octocolobus manul* and *Prionailurus viverrinus*. Reactions consisted of 0.1U of AmpliTaq DNA polymerase, 1.5 μM primer, 2.5 mM $MgCl_2$, 0.2 mM of each deoxynucleotide triphosphate and the appropriate amount of 10X AmpliTaq Buffer II and water in a 20 μl total volume. PCR conditions were 5 min at 94°C, 40 cycles of 30 sec at 94°C, 30 sec at 54°C and 90 sec at 72°C followed by a final elongation step of 5 min at 72°C.

SINE-to-SINE amplifications resulted in a collection of DNA fragments flanked by head-to-head oriented CanSINE segments. To confirm amplification and assess the size range of DNA fragments, 5 μl of PCR product was fractionated by gel electrophoresis in a 1.0% agarose gel containing ethidium bromide. Prior to cloning, 15 μl of PCR product was purified using the ExoSAP protocol with 0.72 μl shrimp alkaline phosphatase (SAP) and 1.44 μl exonuclease I (ExoI) (Amersham Pharmacia, Piscataway, NJ). Isolation of SINE flanked fragments was completed using the TOPO TA Cloning kit (Invitrogen). Twelve to 24 clones from each query species were purified and sequenced following the protocol described in the previous section.

Identification of novel informative SINE loci

Sequenced DNA fragments consisted of genomic sequence from the host species flanked at either end by the tRNA-related region of a feliform specific CanSINE insertion. After masking for low complexity repeats using RepeatMasker [70], the segments were aligned to the December 2008 10X *F. catus* whole genome sequence with the BLAST algorithm. When possible, the resulting

homologous *F. catus* regions were extended 200 bp on either end, imported into Sequencher and aligned in appropriate contigs. Two screening strategies were then employed depending on insertion presence or absence status in *F. catus*. If SINEs identified in exotic species were absent in *F. catus*, primers were built around the putative insertion sites and all Felidae species were then amplified by direct PCR. Alternatively, if a SINE was initially identified a non-panthera lineage species and *F. catus*, primers were built around the putative insertion site and direct PCR was performed on a Pantherinae species. If the insertion is present in Pantherinae, then the insertion must have occurred in the ancestor of all Felidae. However, if the insertion is absent from Pantherinae, the insertion event must have occurred during the subsequent Felidae radiation. The site was then assessed by direct PCR and sequencing in all Felidae species as described in the previous section. After confirmation of amplification by gel electrophoresis, PCR products were purified and sequenced.

Determining SINE presence or absence

A specific SINE insertion site is delimited by the exact sequence of the 6–20 base pair target site duplication (TSD). If a SINE is present the amplification product will include; the forward primer sequence, 5' genomic sequence, one copy of the TSD, the SINE element, the second copy (duplicate) of the TSD, 3' genomic sequence and the reverse primer sequence. If a SINE is absent, the amplification product will include; the primer sequences plus 5' and 3' genomic sequence bracketing one copy of the TSD sequence (canonical genomic DNA). Note that the absence of any PCR product signifies amplification failure and does not imply that the SINE is absent from the homologous region. Thus, criteria for successful amplification loci are 1) PCR products from *F. catus* include the target SINE insertion and therefore are about 200–400 base pairs larger than the amplification products of *C. familiaris* that lack the target SINE insertion, 2) the sequence of the TSD can be determined by examining sequence traces of *F. catus* 3) PCR products yielded sufficiently legible sequences such that SINE presence or absence at the TSD can be ascertained in at least 80% of the sample taxa.

Evolutionary analysis of SINE subfamilies

Representing 87 full-length SINE insertion loci, 5' tRNA regions were aligned using the MAFFT algorithm implemented in the Geneious software package version 5 [94-96]. Phylogenetic analyses were performed using minimum evolution, maximum parsimony and maximum likelihood methods. The Tamura-Nei plus gamma (TrN + G) model was selected as the optimal nucleotide substitution model for likelihood analyses using Modeltest with the AIC criterion [97,98]. Minimum evolution was

implemented in Geneious [96] using the neighbor-joining algorithm, maximum parsimony was implemented using PAUP [99] and maximum likelihood was implemented in GARLI through the Lattice Project Grid computing system using the general time reversible model (nearest option to TrN) and a gamma distribution to account for among-site rate variation [100,101]. Bootstrap support values for all three analyses were obtained from 1000 repetitions. Genetic distances were obtained from the distance matrix calculated for the minimum evolution phylogeny. The mean rate of substitution for the tRNA-derived regions from each SINE subfamily and subtype as well as for all SINEs examined here were calculated by averaging the quotients: D/T where D is the genetic distance between each SINE pair and T minimum age of the most recent common ancestor of the lineages in which the pair of SINEs occur [60,66]. Tests for significance between substitution rates were calculated using the unpaired T-test, with significance at $p < 0.05$.

Availability of supporting data

DNA sequences are catalogued in GenBank. Accession numbers are indicated in Additional tables 3 and 4. *Note, sequences under 200 base pairs cannot be catalogued but are available from the corresponding author. The phylogenetic data set supporting the results of this article is available in the TreeBase repository, at http://purl.org/phylo/treebase/phylows/study/TB2:S15822 [102].

Additional files

Additional file 1: Table S1. List of species used in study.

Additional file 2: Table S2. PCR Primers used to amplify CanSINE containing regions. A) PCR primers flanking 29 feliform specific CanSINE insertions identified initial *F. catus* genome annotation (Pontius et al. [92]). B) PCR primers for 30 genomic loci containing informative CanSINE loci. Each primer pair is designated by the corresponding UCSC genome browser scaffold number (if known) and chromosome coordinates (if known).

Additional file 3: Table S3. Genomic regions containing single CanSINE insertion events among feliforms. Target site duplications, distribution among taxa and corresponding GenBank accession numbers are indicated.

Additional file 4: Table S4. Genomic regions containing multiple CanSINE insertion events among feliforms. Target site duplications, distribution among taxa and corresponding GenBank accession numbers are indicated.

Additional file 5: Figure S1. Four unique CanSINE insertion events at locus 133135. Alignment of four CanSINE insertion events occurring at locus 133135 in the caracal (*Profelis caracal*), marbled cat (*Pardofelis marmorata*), ocelot lineage (genus: *Leopardus*) and bobcat (*Lynx rufus*). The homologous region without CanSINE from *F. catus* sequence in included as a reference. The *L. rufus* CanSINE is in reverse orientation, RC (reverse complement). Yellow-highlighted regions mark target site duplications and gray-shaded regions denote the A and B RNA polymerase III recognition sequences.

Additional file 6: Figure S2. Six unique insertion events occurring at locus 212075. Alignment of six insertion events occurring at locus 212075 in the asian leopard cat lineage species of rusty spotted cat (*P. rubiginosus*), flat-headed cat (*P. planiceps*), Asian leopard cat

(*P. bengalensis*); black-footed cat (*F. nigripes*); along with synapomorphies of the African golden cat clade (*P. caracal/aurata*), and the bay cat lineage. The homologous region without CanSINE from *F. catus* sequence in included as a reference. The insertions in the Asian leopard cat (*P. bengalensis*) and flat-headed cat (*P. planiceps*) are unfixed. Yellow-highlighted regions mark target site duplications and gray-shaded regions denote the A and B RNA polymerase III recognition sequences.

Additional file 7: Figure S3. Alignment of Lynx lineage individuals at locus 106256. Alignment of Lynx lineage individuals at locus 106256 reveals a conserved Can-SINE in all species, a Can-SINE immediately adjacent in *L. pardinus* and some *L. lynx* individuals, and a different Can-SINE insertion that begins 260 bp 3' of the conserved SINE in all *L. canadensis* and some *L. lynx* individuals. Each insertion has unique TSDs, as well as distinct SNPs in the tRNA related region. Yellow-highlighted regions mark target site duplications and gray-shaded regions denote the A and B RNA polymerase III recognition sequences.

Additional file 8: Table S5. Distribution of locus 106265 CanSINEs among *Lynx pardinus*, *L. lynx*, and *L. canadensis* individuals. Ningxia, Qinghai, Yunnan are located in China.

Additional file 9: Figure S4. Alignment of Lynx lineage individuals at locus 134463. Alignment of Lynx lineage individuals at locus 134463 reveals, a Can-SINE insert in *L. canadensis* (N = 7) and *L. lynx* (N = 3), yet absent from all 5 *L. pardinus* (N = 5). Yellow-highlighted regions mark target site duplications and gray-shaded regions denote the A and B RNA polymerase III recognition sequences.

Additional file 10: Figure S5. Alignment of ocelot (*Leopardus*) lineage individuals at locus 133135. Alignment of ocelot (*Leopardus*) lineage individuals at locus 133135 reveals a Can-SINE insert in *L. pardalis*(N = 9) *L. tigrina* (N = 8), *O. geoffroyi* (N = 8) and *O. guigna* (N = 1) and *L.jacobita* (N = 1) individuals, yet absent from *L. colocolo* (N = 3) and *L. wiedii* (N = 2). Yellow-highlighted regions mark target site duplications and gray-shaded regions denote the A and B RNA polymerase III recognition sequences.

Additional file 11: Figure S6. Alignment of DNA sequences from locus 161275. Alignment of DNA sequences from locus 161275 SINE amongst Asian leopard cat lineage species, with NADH5 haplotypes for each individual indicated in parentheses. Yellow-highlighted regions mark target site duplications and gray-shaded regions denote the A and B RNA polymersase III recognition sequences. The 2 rusty-spotted cat (*P. rubiginosa*) individuals have SINE sequences that differ from each other and are found amongst the asian leopard cat (*P. bengalensis*). Thus it cannot be determined whether the rusty-spotted cats acquired their SINEs through incomplete lineage sorting of ancestral polymorphisms or through hybridization with the Asian Leopard cat.

Additional file 12: Figure S7. Alignment of a CanSINE insertion at locus 154966. Proliferation of the CanSINE at locus 154966 occurred during the initial Felidae radiation. Sequence from one species represents each lineage. The insertion is present in the domestic cat, asian leopard cat and lynx lineages. Yellow-highlighted regions mark target site duplications and gray-shaded regions denote the A and B RNA polymerase III recognition sequences.

Additional file 13: Figure S8. Alignment of a CanSINE insertion at locus 214534. Proliferation of the CanSINE at locus 214534 occurred during the initial Felidae radiation. Sequence from one species represents each lineage. The insertion is present in all lineages except the Panthera lineage. Within the caracal lineage, the insertion is unfixed within the african golden cat species but absent in multiple caracal and serval individuals. Yellow-highlighted regions mark target site duplications and gray-shaded regions denote the A and B RNA polymerase III recognition sequences.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
KWC conceived of the study, carried out the laboratory molecular genetic work, participated in the phylogenetic/statistical analyses and drafted the manuscript. DJ participated in the design and coordination of the study and critically revised the manuscript. WJ participated in the phylogenetic analyses and critically revised the manuscript. SO participated in the design of the study, critically revised the manuscript and provided final approval for publication. JPS conceived of the study, participated in its design and coordination, participated in the phylogenetic/statistical analyses and critically revised the manuscript. All authors read and approved the final manuscript.

Author details
[1]Department of Biology, American University, 101 Hurst Hall 4440 Massachusetts Ave, Washington, DC 20016, USA. [2]Department of Biological Sciences, The George Washington University, 2036 G St, Washington, DC 20009, USA. [3]Smithsonian Conservation Biology Institute, National Zoological Park, Front Royal, VA 22630, USA. [4]Dobzhansky Center for Genome Bioinformatics, St. Petersburg State University, 41 A, Sredniy Avenue St., Petersburg 199034, Russia.

References
1. Vassetzky NS, Kramerov DA: **SINEBase: a database and tool for SINE analysis.** *Nucleic Acids Res* 2013, **41**(Database issue):D83–89.
2. Ohshima K, Okada N: **Generality of the tRNA origin of short interspersed repetitive elements (SINEs): characterization of three different tRNA-derived retroposons in the octopus.** *J Mol Biol* 1994, **243**(1):25–37.
3. Coltman DW, Wright JM: **Can SINEs: a family of tRNA-derived retroposons specific to the superfamily Canoidea.** *Nucl Acids Res* 1994, **22**(14):2726–2730.
4. Churakov G, Smit A, Brosius J, Schmitz J: **A novel abundant family of retroposed elements (DAS-SINEs) in the nine-banded armadillo (*Dasypus novemcinctus*).** *Mol Biol Evol* 2005, **22**:886–893.
5. Kajikawa M, Okada N: **LINEs mobilize SINEs in the eel through a shared 3' sequence.** *Cell* 2002, **111**(3):433–444.
6. Cordaux R, Batzer M: **The impact of retrotransposons on human genome evolution.** *Nat Rev Genet* 2009, **10**:691–703.
7. Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, Clamp M, Chang JL, Kulbokas EJ, Zody MC 3rd, Mauceli E, Xie X, Breen M, Wayne RK, Ostrander EA, Ponting CP, Galibert F, Smith DR, DeJong PJ, Kirkness E, Alvarez P, Biagi T, Brockman W, Butler J, Chin CW, Cook A, Cuff J, Daly MJ, DeCaprio D, Gnerre S, *et al*: **Genome sequence, comparative analysis and haplotype structure of the domestic dog.** *Nature* 2005, **438**(7069):803–819.
8. Pontius J, Mullikin J, Smith D, Lindblad-Toh K, Gnerre S, Clamp M, Chang J, Stephens R, Neelam B, Volfovsky N, Schaffer A, Agarwala R, Narfstrom K, Murphy W, Giger U, Roca A, Antunes A, Menotti-Raymond M, Yuhki N, Pecon-Slattery J, Johnson W: **Initial sequence and comparative analysis of the cat genome.** *Genome Res* 2007, **17**(11):1675–1689.
9. Jurka J, Kapitonov V, Kohany O, Jurka MV: **Repetitive sequences in complex genomes: structure and evolution.** *Annu Rev Genomics Hum Genet* 2007, **8**:241–259.
10. Cho YS, Hu L, Hou H, Lee H, Xu J, Kwon S, Oh S, Kim H-M, Jho S, Kim S, Shin Y-A, Kim BC, Kim H, Kim C-u, Luo S-J, Johnson WE, Koepfli K-P, Schmidt-K√É¬¬ntzel A, Turner JA, Marker L, Harper C, Miller SM, Jacobs W, Bertola LD, Kim TH, Lee S, Zhou Q, Jung H-J, Xu X, Gadhvi P, *et al*: **The tiger genome and comparative analysis with lion and snow leopard genomes.** *Nat Commun* 2013, **4**:2433. doi: 10.1038/ncomms3433.

11. Liu G, Alkan C, Jiang L, Zhao S, Eichler EE: **Comparative analysis of *Alu* repeats in primate genomes.** *Genome Res* 2009, **19**:876–885.

12. Salem AH, Ray DA, Batzer MA: **Identity by descent and DNA sequence variation of human SINE and LINE elements.** *Cytogenet Genome Res* 2005, **108**:63–72.

13. Zhao F, Qi J, Schuster SC: **Tracking the past: interspersed repeats in an extinct Afrotherian mammal, *Mammuthus primigenius*.** *Genome Res* 2009, **19**:1384–1392.

14. Wang J, Wang A, Han Z, Zhang Z, Li F, Li X: **Characterization of three novel SINE families with unusual features in Helicoverpa armigera.** *PLOS One* 2012, **7**(2):e31355.

15. Buckley PT, Lee MT, Sui J-Y, Miyashiro KY, Bell TJ, Fisher SA, Kim J, Eberwine J: **Cytoplasmic intron sequence-retaining transcripts can be dedritically targeted via ID element retrotransposons.** *Neuron* 2011, **69**:877–884.

16. Zhou Y, Zheng JB, Gu X, Li W, Saunders GF: **A novel Pax-6 binding site in rodent B1 repetitive elements: coevolution between developmental regulation and repeated elements?** *Gene* 2000, **245**(2):319–328.

17. Hasler J, Strub K: **Survey and summary- Alu elements as regulators of gene expression.** *Nucl Acids Res* 2006, **34**:5491–5497.

18. Krull M, Brosius J, Schmitz J: **Alu-SINE exonization: en route to protein-coding function.** *Mol Biol Evol* 2005, **22**(8):1702–1711.

19. Singer SS, Mannel DN, Hehlgans T, Brosius J, Schmitz J: **From "junk" to gene: curriculum vitae of a primate receptor isoform gene.** *J Mol Biol* 2004, **341**(4):883–886.

20. Batzer MA, Deininger P: **Alu repeats and human genomic diversity.** *Nat Rev Genet* 2002, **3**:370–379.

21. Callinan P, Wang J, Herke S, Garber R, Liang P, Batzer M: **Alu retrotransposition-mediated deletion.** *J Mol Biol* 2005, **348**(4):791–800.

22. Deininger P, Batzer M: **Alu repeats and human disease.** *Mol Genet Metab* 1999, **67**(3):183–193.

23. van de Lagemaat LN, Gagnier L, Medstrand P, Mager DL: **Genomic deletions and precise removal of transposable elements mediated by short identical DNA segments in primates.** *Genome Res* 2005, **15**:1243–1249.

24. Krull M, Petrusma M, Makalowski W, Brosius J, Schmitz J: **Functional persistence of exonized mammalian-wide interspersed repeat elements (MIRs).** *Genome Res* 2007, **17**(8):1139–1145.

25. Lin L, Shen S, Tye A, Cai J, Jiang P, Davidson B, Xing Y: **Diverse splicing patterns of exonized Alu elements in human tissues.** *PLoS Genet* 2008, **4**:e10000225.

26. Nishihara H, Smit AFA, Okada N: **Functional noncoding sequences derived from SINEs in the mammalian genome.** *Genome Res* 2006, **16**(7):864–874.

27. Lunyak VV, Prefontaine GG, Nunez E, Cramer T, Ju B-G, Ohgi KA, Hutt K, Roy A, Garcia-Diaz A, Zhu X, Yung Y, Montoliu L, Glass CK, Rosenfeld MG: **Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis.** *Science* 2007, **317**(5835):248–251.

28. Goldman A, Capoano CA, Gonzalez-Lopez E, Geisinger A: **Identifier (ID) elements are not preferentially located to brain-specific genes: High ID element representation in other tissue-specific- and housekeeping genes of the rat.** *Gene* 2014, **533**(1):72–77.

29. Ponicsan SL, Kugel JF, Goodrich JA: **Genomic gems: SINE RNAs regulate mRNA production.** *Curr Opin Genet Dev* 2010, **20**(2):149–155.

30. Wang W, Kirkness EF: **Short interspersed elements (SINEs) are a major source of canine genomic diversity.** *Genome Res* 2005, **15**:1798–1808.

31. Clark LA, Wahl JM, Rees CA, Murphy KE: **Retrotransposon insertion in SILV is responsible for merle patterning of the domestic dog.** *PNAS* 2006, **103**(5):1376–1381.

32. Hu L-J, Laporte J, Kioschis P, Heyberger S, Kretz C, Poustka A, Mandel J-L, Dahl N: **X-linked myotubular myopathy: refinement of the gene to a 280-kb region with new and highly informative microsatellite markers.** *Hum Genet* 1996, **98**(2):178–181.

33. Laporte J, Hu LJ, Kretz C, Mandel JL, Kioschis P, Coy JF, Klauck SM, Poustka A, Dahl N: **A gene mutated in X-linked myotubular myopathy defines a new putative tyrosine phosphatase family conserved in yeast.** *Nat Genet* 1996, **13**(2):175–182.

34. Parker HG, VonHoldt B, Quignon P, Margulies E, Shao S, Mosher D, Spady T, Elkahloun A, Michele C, Jones PG, Maslen CL, Acland GM, Sutter N, Kuroki K, Bustamante C, Wayne R, Ostrander EA: **An expressed *Fgf4* retrogene is associated with breed-defining chondrodyplasia in domestic dogs.** *Science* 2009, **325**(5943):995–998.

35. Sutter N, Bustamante C, Chase K, Gray M, Zhao K, Lan Z, Padhukasahasram B, Karlins E, Davis S, Jones PG, Quignon P, JOhnson GS, Parker HG, Fretwell N, Mosher D, Lawler DF, Satyaraj E, Nordborg M, Lark KG, Wayne RK, Ostrander E: **A single IGF1 allele is a major determinant of small size in dogs.** *Science* 2007, **316**(5821):112–115.

36. Goldstein O, Kukekova AV, Aguirre GD, Acland GM: **Exonic SINE insertion in STK38L causes canine early retinal degeneration (erd).** *Genomics* 2010, **96**(6):362–368.

37. Nakanishi A, Kobayashi N, Suzuki-Hirano A, Nishihara H, Sasaki T, Hirakawa M, Sumiyama K, Shimogori T, Okada N: **A SINE-derived element constitutes a unique modular enhancer for mammalian diencephalic Fgf8.** *PLoS One* 2012, **7**(8):e43785.

38. Nikaido M, Matsuno F, Hamilton H, Brownell RL Jr, Cao Y, Ding W, Zuoyan Z, Shedlock AM, Fordyce RE, Hasegawa M, Okada N: **Retroposon analysis of major cetacean lineages: The monophyly of toothed whales and the paraphyly of river dolphins.** *PNAS* 2001, **98**(13):7384–7389.

39. Lopez-Giraldez F, Andres O, Domingo-Roura X, Bosch M: **Analyses of carnivore microsatellites and their intimate association with tRNA-derived SINEs.** *BMC Genomics* 2006, **7**:269.

40. Schröder C, Bleidorn C, Hartmann S, Tiedemann R: **Occurrence of Can-SINEs and intron sequence evolution supports robust phylogeny of pinniped carnivores and their terrestrial relatives.** *Gene* 2009, **448**(2):221–226.

41. Yu L, Zhang Y-p: **Evolutionary implications of multiple SINE insertions in an intronic region from diverse mammals.** *Mamm Genome* 2005, **16**(9):651–660.

42. Yu L, Zhang Y-p: **Phylogenetic studies of pantherine cats (Felidae) based on multiple genes, with novel application of nuclear β-fibrinogen intron 7 to carnivores.** *Mol Phylogenet Evol* 2005, **35**(2):483–495.

43. Meyer TJ, McLain AT, Oldenburg JM, Faulk C, Bourgeois MG, Conlin EM, Mootnick AR, de Jong PJ, Roos C, Carbone L, Batzer MA: **An Alu-based phylogeny of gibbons (hylobatidae).** *Molecular biology and evolution* 2012, **29**(11):3441–3450.

44. McLain AT, Meyer TJ, Faulk C, Herke SW, Oldenburg JM, Bourgeois MG, Abshire CF, Roos C, Batzer MA: **An alu-based phylogeny of lemurs (infraorder: Lemuriformes).** *PLOS One* 2012, **7**(8):e44035.

45. Farwick A, Jordan U, Fuellen G, Huchon D, Catzeflis F, Brosius J, Schmitz J: **Automated scanning for phylogenetically informative transposed elements in rodents.** *Syst Biol* 2006, **55**(6):936–948.

46. Steppan S, Adkins R, Anderson J: **Phylogeny and divergence-date estimates of rapid radiations in muroid rodents based on multiple nuclear genes.** *Syst Biol* 2004, **53**(4):533–553.

47. Moller-Krull M, Delsuc F, Churakov G, Marker C, Superina M, Brosius J, Douzery E, Schmitz J: **Retroposed Elements and Their Flanking Regions Resolve the Evolutionary History of Xenarthran Mammals (Armadillos, Anteaters, and Sloths).** *Mol Biol Evol* 2007, **24**(11):2573–2482.

48. Gu W, Ray DA, Walker JA, Barnes EW, Gentles AJ, Samollow PB, Jurka J, Batzer MA, Pollock DD: **SINEs, evolution and genome structure in the opossum.** *Gene* 2007, **396**(1):46–58.

49. Munemasa M, Nikaido M, Nishihara H, Donnellan S, Austin CC, Okada N: **Newly discovered young CORE-SINEs in marsupial genomes.** *Gene* 2008, **407**(1–2):176–185.

50. Nishihara H, Satta Y, Nikaido M, Thewissen JGM, Stanhope M, Okada N: **Retrotransposon analysis of Afrotherian phylogeny.** *Mol Biol Evol* 2005, **9**:1823–1833.

51. Lum JK, Nikaido M, Shimamura M, Hidetoshi S, Shedlock AM, Okada N, Hasegawa M: **Consistency of SINE insertion topology and flanking sequence tree: quantifying relationships among Cetartiodactyls.** *Mol Biol Evol* 2000, **17**(10):1417–1424.

52. Ray DA, Xing J, Salem AH, Batzer MA: **SINEs of a nearly perfect character.** *Syst Biol* 2006, **55**(6):928–935.

53. Pecon-Slattery J, Wilkerson AJ, Murphy WJ, O'Brien SJ: **Phylogenetic assessment of introns and SINEs within the Y chromosome using the cat family Felidae as a species tree.** *Mol Biol Evol* 2004, **21**(12):2299–2309.

54. Cantrell MA, Filanoski BJ, Ingermann AR, Olsson K, DiLuglio N, Lister Z, Wichman HA: **An ancient retrovirus-like element contains hot spots for SINE Insertion.** *Genetics* 2001, **158**:769–777.

55. Jurka J: **Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons.** *PNAS* 1997, **94**:1872–1877.

56. Jurka J: **Repeats in genomic DNA: mining and meaning.** *Current Opinion in Structural Biology* 1998, **8**:333–337.

57. Jurka J, Klonowski P: **Integration of retroposable elements in mammals: selection of target sites.** *J Mol Evol* 1996, **43**:685–689.

58. Yu L, Luan P-T, Jin W, Ryder OA, Chemnick LG, Davis HA, Zhang Y-p: **Phylogenetic Utility of Nuclear Introns in Interfamilial Relationships of Caniformia (Order Carnivora).** *Syst Biol* 2011, **60**(2):175–187.

59. Churakov G, Kreigs J, Baertsch R, Zemann A, Brosius J, Schmitz J: **Mosaic retrotransposon insertion patterns in placental mammals.** *Genome Res* 2009, **19**:868–875.

60. Eizirik E, Murphy W, Koepfli K, Johnson W, Dragoo J, Wayne R, O'Brien SJ: **Pattern and timing of diversification of the mammalian order Carnivora inferred from multiple nuclear gene sequences.** *Mol Phylogenet Evol* 2010, **56**(1):49–63.

61. Wilson DE, Reeder DM: **Mammal Species of the World.** In *A Taxonomic and Geographic Reference*. 3rd edition. Baltimore, Maryland: Johns Hopkins University Press; 2005.

62. Lavrentieva MV, Rivkin MI, Shilov AG, Kobetz ML, Rogozin IB, Serov OL: **B2-like repetitive sequence from the X chromosome of the American mink (Mustela vison).** *Mamm Genome* 1991, **1**(3):165–170.

63. Minnick MF, Stillwell LC, Heineman JM, Stiegler GL: **A highly repetitive DNA sequence possibly unique to canids.** *Gene* 1992, **110**:235–238.

64. Pecon-Slattery J, Murphy WJ, O'Brien SJ: **Patterns of diversity among SINE elements isolated from three Y-chromosome genes in carnivores.** *Mol Biol Evol* 2000, **17**(5):825–829.

65. Vassetzky NS, Kramerov DA: **CAN—a pan-carnivore SINE family.** *Mamm Genome* 2002, **13**(1):50–57.

66. Johnson WE, Eizirik E, Pecon-Slattery J, Murphy WJ, Antunes A, Teeling E, O'Brien SJ: **The late Miocene radiation of modern Felidae: a genetic assessment.** *Science* 2006, **311**(5757):73–77.

67. Borodulina OR, Kramerov DA: **PCR-based approach to SINE isolation: Simple and complex SINEs.** *Gene* 2005, **349**:197–205.

68. Jurka J: *Evolutionary impact of human Alu repetitive elements Current Opinion in Genetics and Development* 2004, **14**(6):603–608.

69. Gentles A, Kohany O, Jurka J: **Evolutionary diversity and potential recombinogenic role of integration targets of non-LTR retrotransposons.** *Mol Biol Evol* 2005, **22**:1983–1991.

70. Smit A, Hubley R, Green P: **Repeat Masker Website and Server.** In: http://www.repeatmasker.org/. 1996-2010.

71. Cordaux R, Hedges DJ, Batzer M: **Retrotransposition of Alu elements: how many sources?** *Trends Genet* 2004, **20**(10):464–467.

72. Ohshima K, Okada N: **SINEs and LINEs: symbionts of eukaryotic genomes with a common tail.** *Cytogenet Genome Res* 2005, **110**(1–4):475–490.

73. Weiner AM: **SINEs and LINEs: the art of biting the hand that feeds you.** *Curr Opin Cell Biol* 2002, **14**(IS - 3):343–350.

74. Borodulina OR, Kramerov DA: **Transcripts synthesized by RNA polymerase III can be polyadenylated in an AAUAAA-dependent manner.** *Rna* 2008, **14**(9):1865–1873.

75. Ray DA: **SINEs of progress: Mobile element applications to molecular ecology.** *Mol Ecol* 2007, **16**:19–33.

76. Kosushkin SA, Grechko VV: **Molecular genetic relationships and some issues of systematics of rock lizards of the genus Darevskia (Squamata: Lacertidae) based on locus analysis of SINE-type repeats (Squam1).** *Russian Journal of Genetics* 2013, **49**(9):857–869.

77. Avise J, Robinson T: **Hemiplasy: a new term in the lexicon of phylogenetics.** *Syst Biol* 2008, **57**:503–507.

78. Agnarsson I, Kunter M, May-Collado L: **Dog, cats and kin: A molecular species-level phylogeny of Carnivora.** *Mol Phylogenet Evol* 2010, **54**(3):726–745.

79. Wesley-Hunt G, Flynn J: **Phylogeny of the carnivora: Basal relationships among the carnivoramorphans, and assessment of the position of the 'Miacoidea' relative to carnivora.** *Journal of Systematic Palaeontology* 2005, **3**(1):1–28.

80. Yu L, Li Q-w, Ryder O, Zhang Y-p: **Phylogenetic relationships within mammalian order Carnivora indicated by sequences of two nuclear DNA genes.** *Mol Phylogenet Evol* 2004, **33**:694–705.

81. Terai Y, Takahashi K, Nishida M, Sato T, Okada N: **Using SINEs to probe ancient explosive speciation: "Hidden" radiation of African cichlids?** *Mol Biol Evol* 2003, **20**(6):924–930.

82. Johnson WE, Pecon-Slattery J, Eizirik E, Kim J-H, Menotti-Raymond M, Bonacic C, Cambre R, Crawshaw P, Nunes A, Seuanez HN, Moreira MAM, Seymour KL, Simon F, Swanson W, O'Brien SJ: **Disparate phylogeographic patterns of molecular genetic variation in four closely related South American small cat species.** *Mol Ecol* 1999, **8**:S79–S94.

83. Johnson WE, Godoy JA, Palomares F, Delibes M, Fernandes M, Revilla E: O'Brien aSJ: **Phylogenetic and phylogeographic analysis of Iberian lynx populations.** *J Hered* 2004, **95**(1):19–28.

84. Luo S, Johnson WE, Martelli P, Antunes A, Smith JLD, O'Brien SJ: **Phylogenetic partitions of Asian felids reveal significant Indochinese-Sundaic transition.** *Mol Biol Evol* 2010. In Press.

85. Perelman PL, Graphodatsky AS, Serdukova NA, Nie W, Alkalaeva EZ, Fu B, Robinson TJ, Yang F: **Karyotypic conservatism in the suborder Feliformia (Order Carnivora).** *Cytogenet Genome Res* 2005, **108**:348–354.

86. Hallstrom B, Janke A: **Mammalian evolution may not be strictly bifurcating.** *Mol Biol Evol* 2010, **27**(12):2804–2816.

87. Scientists GKCo: **Genome 10K: A Proposal to Obtain Whole-Genome Sequence for 10 000 Vertebrate Species.** *J Hered* 2009, **100**(6):659–674.

88. Johnson WE, Culver M, Iriate JA, Eizirik E, Seymour KL, O'Brien S: **Tracking the evolution of the elusive Andean mountain cat (Oreailurus jacobita) from mitochondrial DNA.** *J Hered* 1998, **89**(3):227–232.

89. Luo S-J: **Comparative phylogeography of sympatric wild cats: Implications for biogeography and conservation in Asian biodiversity hotspots.** In *University of Minnesota- Doctoral Dissertation*. Minneapolis: Doctoral Dissertation-University of Minnesota; 2006.

90. Nishihara H, Maruyama S, Okada N: **Retrotransposon analysis and recent geological data suggest near-simultaneous divergence of the three superorders of mammals.** *PNAS* 2009, **106**:5235–5240.

91. Pontius JU, O'Brien SJ: **Genome annotation resource fields—GARFIELD: a genome browser for Felis catus.** *J Hered* 2007, **98**(5):386–389.

92. Mei L, Ding X, Tsang SY, Pun FW, Ng SK, Yang J, Zhao C, Li D, Wan W, Yu CH, Tan TC, Poon WS, Leung GK, Ng HK, Zhang L, Xue H: **AluScan: a method for genome-wide scanning of sequence and structure variations in the human genome.** *BMC Genomics* 2011, **12**:564.

93. Walker JA, Kilroy GE, Xing J, Shewale J, Sinha SK, Batzer MA: **Human DNA quantitation using Alu element-based polymerase chain reaction.** *Analytical Biochemistry* 2003, **315**:122–128.

94. Drummond A, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A: **Geneious v4.7.** 2009. http://www.geneious.com.

95. Katoh K, Asimenos G, Toh H: **Multiple alignment of DNA sequences and MAFFT.** *Methods Mol Biol* 2009, **537**:39–64.

96. Drummond A, Ashton B, Buxton S, Cheung M, Cooper A, Heled J, Kearse M, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A: **Geneious v5.1.** 2010. In: http://www.geneious.com.

97. Posada D, Buckley TR: **Model selection and averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests.** *Syst Biol* 2004, **53**:793–818.

98. Posada D, Crandall KA: **Selecting the best-fit model of nucleotide substitution.** *Syst Biol* 2001, **50**:580–601.

99. Swofford D: **PAUP*. Phylogenetic analysis using parsimony (*and other methods) version 4.** In *Sinauer Associates* 2003.

100. Bazinet A, Cummings M: **The lattice project: a grid research and production environment combining multiple grid computing models.** In *In: Distributed & Grid Computing- Science Made Transparent for Everyone Principles, Applications and Supporting Communities*. 2008:2–13.

101. Zwickl D: **Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under maximum likelihood criterion.** In *The University of Texas at Austin - Doctoral Dissertation*. Austin: Doctoral Dissertation-The University of Texas at Austin; 2006.

102. Walters-Conte KB, Johnson DLE, Johnson WE, O'Brien SJ, Pecon-Slattery J: **The Dynamic Proliferation of CanSINEs Mirrors the Complex Evolution of Feliforms.** *TreeBase* 2014, http://purl.org/phylo/treebase/phylows/study/TB2: S15822.