Syst. Biol. XX(XX):XX–XX, 2025

https://doi.org/10.1093/sysbio/syaf021
Advance Access Publication March 25, 2025

Massive Inter-species Introgression Overwhelms Phylogenomic Relationships Among Jaguar, Lion, and Leopard

Sarah H. D. Santos[1,2], Henrique V. Figueiró[1,3], Tomas Flouri[4], Emiliano Ramalho[5,6], Laury Cullen Jr.[7], Ziheng Yang[4], William J. Murphy[8], and Eduardo Eizirik[1,6,*]

[1]*Laboratory of Genomics and Molecular Biology, School of Health and Life Sciences, PUCRS, Av. Ipiranga, 6681, prédio 12, sala 134, Porto Alegre, RS 90619-900, Brazil*
[2]*Department of Biology, University of Western Ontario, 1151 Richmond Street, London, Ontario N6A 3K7, Canada*
[3]*Vale Institute of Technology, Environmental Genomics Group, R. Boaventura da Silva 955, Nazaré, Belém, PA 66055-090, Brazil*
[4]*Department of Genetics, Evolution and Environment, University College London, Darwin Building, Gower Street, London WC1E 6BT, UK*
[5]*Instituto de Desenvolvimento Sustentável Mamirauá, Estrada do Bexiga 2584, Bairro Fonte Boa, Tefé, AM 69470-000, Brazil*
[6]*Instituto Pró-Carnívoros, Av. Horácio Netto, 1030, Chácaras Interlagos, Atibaia, SP 12945-010, Brazil*
[7]*Instituto de Pesquisas Ecológicas, Rod. Dom Pedro I, km 47, Nazaré Paulista, SP 12960-000, Brazil*
[8]*Department of Veterinary Integrative Biosciences, Texas A&M University, 402 Raymond Stotzer Pkwy building 2, College Station, TX 77843, USA*
[*]*Correspondence to be sent to: Eduardo Eizirik. Escola de Ciências da Saúde e da Vida, PUCRS, Av. Ipiranga, 6681, prédio 12, sala 134, Porto Alegre, RS 90619-900, Brazil. E-mail: eduardo.eizirik@pucrs.br.*

Received 16 July 2023; reviews returned 27 February 2025; accepted 22 March 2025
Associate Editor: Robert Thomson

Abstract.—Phylogenomic analyses of closely related species allow important glimpses into their evolutionary history. Although recent studies have demonstrated that inter-species hybridization has occurred in several groups, incorporating this process in phylogenetic reconstruction remains challenging. Specifically, the most predominant topology across the genome is often assumed to reflect the speciation tree, but rampant hybridization might overwhelm the genomes, causing that assumption to be violated. The notoriously challenging phylogeny of the 5 extant *Panthera* species (specifically jaguar [*P. onca*], lion [*P. leo*], and leopard [*P. pardus*]) is an interesting system to address this problem. Here we employed a *Panthera*-wide whole-genome-sequence data set incorporating 3 jaguar genomes and 2 representatives of lions and leopards to dissect the relationships among these 3 species. Maximum-likelihood trees reconstructed from non-overlapping genomic fragments of 4 different sizes strongly supported the monophyly of all 3 species. The most frequent topology (76–95%) united lion + leopard as a sister species (topology 1), followed by lion + jaguar (topology 2: 4–8%) and leopard + jaguar (topology 3: 0–6%). Topology 1 was dominant across the genome, especially in high-recombination regions. Topologies 2 and 3 were enriched in low-recombination segments, likely reflecting the species tree in the face of hybridization. Divergence times between sister species of each topology, corrected for local-recombination-rate effects, indicated that the lion-leopard divergence was significantly younger than the alternatives, likely driven by post-speciation admixture. Introgression analyses detected pervasive hybridization between lions and leopards, regardless of the assumed species tree. This inference was strongly supported by multispecies-coalescence-with-introgression analyses, which rejected topology 1 (lion+leopard) or any model without introgression. Interestingly, topologies 2 (lion+jaguar) and 3 (jaguar+leopard) with extensive lion-leopard introgression were unidentifiable, highlighting the complexity of this phylogenetic problem. Our results suggest that the dominant genome-wide tree topology is not the true species tree but rather a consequence of overwhelming post-speciation admixture between lion and leopard. [keyword: Evolution; hybridization; Felidae; Panthera; speciation.]

The availability of genome-wide data sets has led to new perspectives on the evolutionary history of present-day species. Studies on whole genomes have increasingly shown that species divergences are often tangled with admixture processes (e.g., Arnold 2015; Ackerman et al. 2019). Post-speciation hybridization resulting in introgression has been found to be a significant source of genome-wide phylogenetic (genealogical) discordance, in addition to the well-known stochastic process of incomplete lineage sorting (ILS) (e.g., Fontaine et al. 2015, Li et al. 2016a, 2019; Payseur et al. 2016; Figueiró et al. 2017; Kumar et al. 2017; Edelman et al. 2019; Martin et al. 2019). Admixture events may have occurred during or shortly after speciation, making it difficult to distinguish from ILS (Degnan and Rosenberg 2009; Payseur et al. 2016).

Mallet et al. (2016) have previously discussed the challenge to distinguish between a model of complete isolation and a model in which gene flow is so rampant that the genome is dominated by the incorrect gene tree induced by introgression. Various methods exist to evaluate these patterns (e.g., Payseur et al. 2016; Flouri et al. 2018; Hibbins et al. 2021), and the advancement and greater affordability of whole-genome sequencing have provided ample data for inferring gene flow. These novel analyses allow an improved dissection of the relationships among different species, but so far there have been few in-depth assessments of systems in which admixture may overwhelm the speciation signal.

Theoretical analyses suggest that even a small amount of gene flow can dramatically impact the genome-wide distribution of gene trees so that the dominating gene

tree may differ from the species tree (Long and Kubatko 2018; Jiao et al. 2020). Although several studies have investigated this problem using a single individual per species (e.g., Cahill et al. 2013; Kumar et al. 2017; Edelman et al. 2019), some have incorporated multiple individuals per taxon (e.g., Fontaine et al. 2015; de Manuel et al. 2016; Vianna et al. 2020; Santos et al. 2021; Bursell et al. 2022). Incorporating multiple individuals per species is particularly important to inferring gene flow (Jiao et al. 2021), as well as assessing species-level monophyly and the replicability of topological discordance patterns.

The phylogeny of the cat family (Felidae) has been extensively studied, revealing a complex evolutionary history that includes rapid radiation among extant lineages beginning ~10-15 million years ago (Ma) (Johnson et al. 2006, Li et al. 2016a, 2019). While some studies have mainly targeted family-wide phylogenetic problems (e.g., Johnson et al. 2006, Li et al. 2016a, 2019), others have focused on within-genus issues (e.g., Trigo et al. 2013; Figueiró et al. 2017). The *Panthera* genus and its sister-group, *Neofelis* spp., represent the most ancestral divergence within the extant Felidae. There are 5 living *Panthera* species, including jaguar (*P. onca*), leopard (*P. pardus*), lion (*P. leo*), tiger (*P. tigris*), and snow leopard (*P. uncia*) (Sunquist and Sunquist 2002), which have diverged from a common ancestor ca. 3.4–5.8 Ma (Johnson et al. 2006; Li et al. 2016a; Figueiró et al. 2017).

Numerous studies have focused on the phylogeny of this genus, revealing complex patterns of genealogical discordance, with distinct gene-tree topologies inferred using morphological and molecular data sets. Davis et al. (2010) used a Y-chromosome intron data set to infer the most likely species tree to be (((lion, leopard), jaguar), (tiger, snow leopard)), grouping the lion, and leopard as sister species. More recent analyses revealed new levels of complexity, including extensive historical hybridization among *Panthera* species, distinct genealogical patterns on the X chromosome and autosomes, with cytonuclear discordance (Li et al. 2016a; Figueiró et al. 2017). The first study using whole-genome sequences from all 5 species (Figueiró et al. 2017) demonstrated that the sister relationship between lion and leopard was the most common in the genome. However, its frequency varied along chromosomes, being lower in pericentromeric regions, particularly in a "recombination coldspot" located on the X chromosome, yet much higher near the ends of chromosomes. The 2 alternative topologies (lion + jaguar and leopard + jaguar) were enriched on the X chromosome, in which they tended to show a younger age for all nodes (including the 1 defining sister species) relative to the prevalent tree. These results led the authors to infer that the lion + leopard topology most likely reflected the sequence of speciation events (i.e., the species tree). At the same time, the other 2 were induced by a mixture of ILS and post-speciation admixture coupled with introgression, possibly driven by selection (Figueiró et al. 2017).

A subsequent study with whole genomes of 27 different cat species (including all 5 *Panthera*) expanded the assessment of genealogical discordance in the family (Li et al. 2019). Those analyses indicated a complex relationship among ILS, post-speciation gene flow, and recombination rate, generating discrepant patterns in different genomic regions, likely compounded by localized effects of natural selection. The authors concluded that genomic regions with low recombination rates are less prone to the impact of introgression and are more likely to retain the signal of the original branching order (or the true species tree). Given these previous findings, the *Panthera* genus emerges as an interesting system to employ novel phylogenomic approaches to dissect the processes leading to genome-wide patterns of genealogical discordance.

To address this, the present study specifically dissected the relationships among these 3 species using whole-genome sequences of an expanded sample of individuals. The usage of 2 or 3 individuals per species allowed the testing of genome-wide species-level monophyly and an assessment of the robustness of introgression analyses to changes in the species representatives. We characterized genome-wide patterns of genealogical discordance in detail and assessed their relationships to recombination rate. We also evaluated the variation across the genome in the divergence time between sister species defined by competing topologies, employing a strategy that corrects for local recombination-rate effects. Finally, we employed the recent implementation of the multispecies-coalescence-with-introgression (MSci) model in the bayesian phylogenetics and phylogeography program (BPP; Flouri et al. 2020) to statistically test competing models incorporating the 3 possible topologies and different admixture histories. These analyses allowed us to better distinguish the signatures of speciation, ILS, and introgression, and revealed that the evolutionary history of these species has been even more complex than previously envisaged.

Materials and Methods

Whole-Genome Sequencing

To assess the effect of including more than 1 individual per species in analyses focusing on genealogical discordance, we sequenced 2 jaguar genomes (jaguar 2, sampled in the Amazon region [NCBI accession no. SRR22298147]; jaguar 3, sampled in the Atlantic Forest [SRR22298151]; both have also been analyzed in a parallel study with a different focus [Lorenzana et al. 2022]). Using DNA extracts from a previous study focused on whole-exome sequencing (Figueiró 2016), we sequenced each genome on the Illumina HiSeqX platform with 300–350 bp insert libraries and paired-end 150-bp reads. Each genome was sequenced in half a lane, achieving ~22× and 19× coverage for jaguar 2 and 3, respectively.

Data Set Construction

The 2 novel genomes were analyzed jointly with our previously published jaguar genome (SRR4444360), derived from an individual sampled in the Pantanal region (referred to as "jaguar 1"). Therefore, our 3 jaguar genomes represent populations from 3 distinct South American biomes separated by up to 3,000 km, thus providing a good representation of extant genetic diversity and coalescent depth for this species. Our data set also included the following genomes: 1 domestic cat (*Felis catus* v.9.0: GCA_000181335.4), 2 leopards (SRR5382748; SRR5382749; SRR5382750; Kim et al. 2016), 2 lions (SRR836361; SRR836370), 1 tiger (SRR836306) and 1 snow leopard (SRR836372). For each genome, we assessed sequence quality using FastQC (v.0.11.7) (Andrews 2010) and trimmed sequences with Trimmomatic (v.0.36) (Bolger et al. 2014), excluding sites with Phred-score < 20 and length < 50 bp. Thus, our full data set comprised 10 complete genomes representing all 5 extant *Panthera* species (with 2–3 individuals per species in the case of jaguar, lion, and leopard) and the domestic cat to be used as an outgroup.

We mapped each genome with BWA-MEM (v.0.17.15) (Li and Durbin 2009) onto the domestic cat v.9 reference genome (Buckley et al. 2020) using default parameters. The resulting files for each genome were sorted based on genomic coordinate order and indexed using Samtools (v.1.5) (Li et al. 2009), followed by indel miscalling filtering with the Genome Analysis Toolkit (v.3.2-2) (McKenna et al. 2010). We obtained the genome depth of these filtered files with BEDTools (*genomecov*; v.2.17.0) (Quinlan and Hall 2010) (Supplementary Fig. S1) and generated a consensus for each genome with ANGSD (v.0.921) (Korneliussen et al. 2014), considering quality scores (*minMapQ: 30*; *minQ: 20*; *setMinDepth: 5*). In the case of heterozygote sites, the base represented in the consensus was randomly chosen using ANGSD so that the resulting genomes were pseudohaploids (*doFasta: 2*; *doCounts: 1*). We masked the repetitive regions of each genome based on the domestic cat coordinates using BEDTools (*maskfasta*).

The global genome-wide consensus of each individual (already aligned to the domestic cat reference) was divided into non-overlapping genomic fragments (GFs) of different sizes (50 kb, 100 kb, 1 Mb, and 5 Mb, based on the domestic cat coordinates) with BEDTools. We first created the intervals with BEDTools *makewindows* using the domestic cat genome. The resulting interval files were used to generate the fasta files for each window size using BEDTools *getfasta*. In addition, to minimize potential biases induced by recombination in divergence time estimates (see below), we also created a data set comprising 10 kb GFs spaced from each other by 190 kb, thus decreasing intra-fragment recombination and increasing inter-fragment independence. The files (1 per individual) obtained per window for each GF size were merged into a single alignment with a custom Python script (Figueiró 2019).

Phylogenomic Analyses With Individual GFs

For the initial set of phylogenomic analyses, individual GFs of each size (10 kb, 50 kb, 100 kb, 1 Mb, and 5 Mb) of jaguars 1 and 2, 2 lions, 2 leopards, 1 tiger, 1 snow leopard, and the domestic cat outgroup were used to create input alignments. We used these data sets to reconstruct the maximum-likelihood (ML) phylogenetic tree for each GF with RAxML-PTHREADS (v.8.2.10) (Stamatakis 2014). We selected the algorithm that performs the best-scoring search and rapid bootstrap analysis simultaneously (option *-f a*). The substitution model employed was the general time reversible with gamma correction for rate variation among sites (GTR-GAMMA). We assessed the nodal support with 100 bootstrap replicates. For the main phylogenomic analyses of the 50 kb, 100 kb, 1 Mb, and 5 Mb GFs, we only included segments whose ML reconstruction supported 1 of the 3 focal topologies with bootstrap values > 90% since the goal was to consider segments whose local genealogy had high confidence. We also explored the use of all GFs (with no bootstrap filtering) to assess the impact of this selection criterion on phylogenomic analyses. We used the bootstrap-filtered GFs to assess species-level monophyly and estimate nucleotide diversity/divergence. All other downstream analyses were performed with all the GFs obtained for each size (i.e., with no bootstrap filter), including inferring the "species tree" with ASTRAL-II, the divergence time analyses performed with BASEML, the MSci analyses performed with BPP, and the introgression analyses with D_{FOIL} (see below for details).

We initially tested the species-level monophyly of jaguars, lions, and leopards using the same data set (with 2 individuals from each species) for each GF size. We repeated the same procedure with a data set that included the 3 jaguar genomes to assess whether this impacted the species-level support. We calculated the mean diversity per nucleotide site (π) using the genomes from the pairs (or trios) of individuals sampled for lion, leopard, and jaguar using a Python egglib package (De Mita and Siol 2012). This was estimated from 100 kb GFs filtered for bootstrap values > 90%. To obtain π per site and GF, we estimated the total number of sites showing diversity with the egglib.stats.ComputeStats function, and divided that by the total number of sites used in each GF.

Subsequent analyses focused on the relationship among jaguars, leopards, and lions. Based on their genome-wide frequencies (see Results), the 3 supported topologies are referred to as topology 1 ((lion, leopard), jaguar), topology 2 ((jaguar, lion), leopard), and topology 3 ((jaguar, leopard), lion) (Fig. 1; Table 1; Supplementary Fig. S2a–c). We surveyed the genome-wide frequencies of each topology for the various GF sizes and characterized their spatial distribution along each chromosome with the recombination rate estimated for the domestic cat v.8 reference genome (Li et al. 2016b).

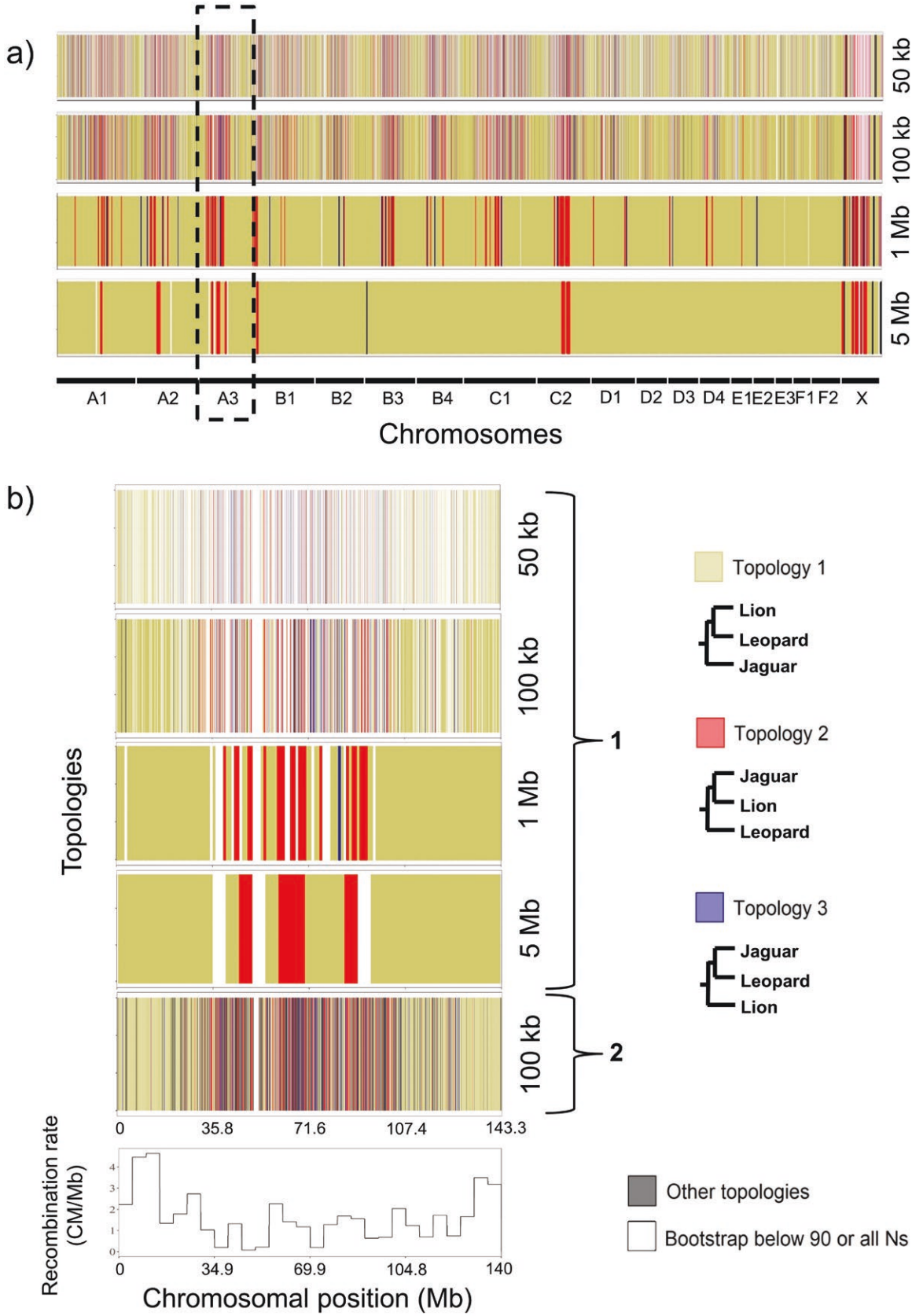

FIGURE 1. Variation in the phylogenetic tree topology for jaguar, lion, and leopard inferred using a) GFs of different sizes (50 kb, 100 kb, 1 Mb, and 5 Mb). The chromosome is indicated at the bottom. b) Detailed views of chromosome A3 either using a bootstrap cutoff (> 90) or without bootstrap filtering (for the 100-kb windows). The recombination rate is shown at the bottom. The topologies are as follows: 1: ((lion,leopard),jaguar); 2: ((lion,jaguar),leopard); and 3: ((jaguar,leopard),lion).

Table 1. Chromosome-level summary of the frequency of each topology, enrichment in low-recombination regions, relative age (proportional time depth) of sister pair, and *P*-value for testing the age difference (see Supplementary Figs. S13–S31 for details)

Chromosome	Percent frequency of topology[1]			T2/T3 in LowRec regions[2]	Mean normalized (relative) age of sister species node[3]			*P*-values for age difference[4]		
	Top1	Top2	Top3		Top1	Top2	Top3	Top1-2	Top1-3	Top2-3
All chrs	**30.76**	10.95	10.69	Yes	**68.18**	76.08	78.55	1E-52	4E-82	3E-04
Autosomes	**31.19**	10.75	10.71	Yes	**67.90**	75.67	78.24	2E-48	7E-79	2E-04
A1	**28.68**	12.04	11.87	Yes	**67.76**	73.18	78.36	1E-03	3E-09	2E-02
A2	**26.90**	12.02	10.36	Yes	**69.03**	75.22	79.11	3E-03	8E-06	0.1
A3	**27.41**	11.22	10.65	Yes	**69.51**	73.92	78.08	0.06	7E-05	0.1
B1	**26.58**	10.86	12.71	Yes	**65.73**	72.21	76.22	2E-03	4E-08	0.06
B2	**32.85**	10.25	9.20	Yes	**66.72**	74.70	79.23	8E-05	1E-07	0.1
B3	**33.51**	11.67	10.45	Yes	**69.51**	77.85	77.98	2E-05	8E-05	0.8
B4	**32.58**	10.39	9.83	Yes	**66.64**	75.62	75.94	2E-04	9E-05	0.7
C1	**30.37**	10.97	11.88	Yes	**68.34**	75.87	79.82	2E-05	6E-11	4E-02
C2	**26.51**	12.94	12.44	Yes	**69.39**	75.64	79.02	1E-03	4E-06	0.2
D1	**31.88**	10.10	9.23	Yes	**66.80**	80.51	80.64	9E-09	6E-08	0.6
D2	**38.18**	13.18	9.55	Yes	**67.98**	78.46	76.42	6E-05	2E-04	0.8
D3	**34.61**	10.19	9.34	Yes	**69.02**	78.44	79.83	7E-05	2E-05	0.6
D4	**29.51**	10.19	10.62	Yes	**67.42**	75.34	77.67	2E-03	1E-04	0.4
E1	**46.43**	10.39	6.82	Yes	**68.17**	78.82	76.69	4E-04	1E-02	0.8
E2	**37.74**	8.39	10.00	Yes	**70.84**	78.91	80.65	8E-03	0.4	0.6
E3	**37.02**	10.58	12.02	Yes	**68.40**	73.66	77.25	5E-02	2E-03	0.7
F1	**36.42**	6.94	12.43	Yes	**65.46**	74.94	75.23	7E-03	3E-04	0.6
F2	**32.78**	8.13	7.89	Yes	**66.10**	80.83	78.97	2E-06	2E-05	0.4
X	**22.80**	14.66	10.26	Yes	**75.46**	81.65	84.65	1E-0	2E-04	0.3

[1]Frequency (in percentage) of each topology in 100 kb GFs for each chromosome. Percentages are expressed relative to the total number of GFs in that partition, including those that did not achieve 90% bootstrap support for one of the focal topologies or that harbored other trees. The most prevalent topology is highlighted in bold.

[2]Enrichment of GFs supporting topologies 2 or 3 in low-recombination-rate regions.

[3]Normalized (relative) age is depicted as the percent depth of the internal node representing the sister species in each topology (lion-leopard in topology 1; lion-jaguar in topology 2; jaguar-leopard in topology 3) relative to the depth estimated for the preceding (trio) node. This analysis was performed with 10 kb GFs spaced by 190 kb. The topology yielding each chromosome's most recent (i.e., lowest percentage) age is highlighted.

[4]Significance assessment (Mann–Whitney test) for pairwise comparisons of the relative age of the sister-pair node for the different topologies (e.g., Top 1–2 for comparing the lion-leopard node in Topology 1 with the lion-jaguar node in Topology 2).

We also inferred the "species tree" from our genome-wide data set using the multispecies coalescent approach implemented in ASTRAL-II (Mirarab and Warnow 2015). The phylogenetic trees inferred with RAxML for 100 kb GFs (using the data set with 2 jaguars) were used to estimate the "species tree" for each chromosome and the complete genome-wide data set (with and without the X chromosome). All ASTRAL-II parameters were left at default values. Notably, ASTRAL and RAxML analyses assume no inter-species gene flow. Additionally, we estimated a species network with the maximum-likelihood approach in PhyloNet (*InferNetwork_ML*; Than et al. 2008 ; Wen et al. 2018). All gene trees generated with 100 kb GFs were used (N = 24,201). We estimated networks with a maximum of 1 reticulation across 50 runs (*-x 50*). For each run, we selected the 5 best networks and optimized their inheritance probabilities and branch lengths (*-po*). The analyses included a taxa map that combined 2 individuals of each focal species (a: 2 lions; b: 2 leopards; c: 2 jaguars). We visualized these results with Dendroscope v.3 (Huson and Scornavacca 2012) and included the inheritance probabilities in the networks. Finally, we calculated information criterion tests (akaike information criterion [AIC] and bayesian information criterion [BIC]) as proposed by Yu et al. (2014) to determine the best number of reticulations.

Divergence Time Estimates for Individual GFs

To characterize divergence patterns across the genomes, we initially estimated the mean divergence per site (d_{xy}) between sister species (lions/leopards in topology 1; lions/jaguars in topology 2; leopards/jaguars in topology 3). This analysis was based on 100-kb GFs that remained after filtering for bootstrap values > 90% for 1 of the 3 focal topologies and employed a Python egglib package. To obtain d_{xy} per site and GF, we estimated the total number of sites showing diversity with the egglib.stats.ComputeStats function, and divided that by the total number of sites used in each GF. To statistically compare d_{xy} estimates between alternative topologies, we first used a Kolmogorov–Smirnov test to assess whether the data were normally distributed. We then applied the non-parametric Mann–Whitney test, given their non-normal distribution. Both steps were conducted with IBM SPSS Statistics v.27.

To estimate the depth of divergence between the sister species defined by each topology (see Supplementary Fig. S2a–c) and its variation across the genome, we used the data set comprising 10-kb GFs, spaced by 190 kb, to decrease intra-fragment recombination and increase independence among sampled segments. Only GFs whose ML tree was topology 1, 2, or 3 were considered (i.e., GFs yielding other topologies were discarded). We

used BASEML (Yang 2007) to estimate the divergence depth at the node defining the innermost sister species pair in each of the topologies, as well as the depth of the preceding (trio) node, which provided an internal control for substitution rate and other effects of the local-recombination rate. This correction is warranted since low-recombining regions may lead to slower substitution rates, resulting in younger estimated depths if a local calibration is not employed. For each 10-kb GF, we estimated the divergence times assuming a global clock (clock = 1) and the JC69 model of nucleotide substitution, given the high similarity among the sequences. All other parameters were left at default values.

We then calculated the ratio of the sister-pair depth to that of the respective trio node to compute a locally calibrated, "normalized" (relative) age of the sister-pair node in each GF. Using this "normalized age" approach, one can compare the depths of local trees across the genome to assess signatures of introgression. The expectation is that local trees induced by introgression will have a younger relative depth of the sister-pair node compared to species-tree compatible topologies. We plotted the divergences at both nodes and their ratio across the genome to assess spatial patterns and their relationship with the recombination landscape. Finally, we compared the relative ages of the sister-pairs between topologies (1 vs. 2; 1 vs. 3; 2 vs. 3) across the genome and within each chromosome, using an alpha value of 0.05. As in the analyses described above, we first tested whether the data were normally distributed with the Kolmogorov–Smirnov test. When the distribution was normal, we employed a One-way analysis of variance (ANOVA); otherwise, we applied non-parametric tests. For the latter, we compared all 3 topologies simultaneously with the Kruskal–Wallis test and performed pairwise comparisons with the Mann–Whitney test. All tests were conducted with IBM SPSS Statistics v.27.

Introgression Analysis With D_{FOIL}

We used the software D_{FOIL}, which extends the *D*-statistics approach, to detect introgression signals on 4 taxa (Pease and Hahn 2015). This program assumes a symmetric 5-taxon phylogeny (Supplementary Fig. S2d), the fifth being the outgroup, which made it possible to compare multiple individuals from our focal trio. The test considers the presence and frequency of ancestral and derived alleles in different taxon sequences, allowing differentiation between introgression and ILS (Pease and Hahn 2015).

We tested all possible combinations using the different individuals of lions, leopards, and jaguars (assuming they represent genetically distinct populations). We employed the tiger as an outgroup (Supplementary Fig. S3) with non-overlapping 100 kb GFs for each data set. Data sets containing 2 jaguar individuals P_1 and P_2 were named data sets 1, 2, and 3 for the 3 possible combinations of individuals. The same strategy was used to label data sets with the 2 leopards as P_1 and P_2 (data sets 4, 5, and 6) and the 2 lions as P_1 and P_2 (data sets 7, 8, and 9). Moreover, within each data set, we also tested all combinations of individuals from the other taxa (P_3 and P_4), resulting in 24 different sub-data sets that were separately assessed with D_{FOIL}. To be conservative in the survey of GFs bearing signatures of introgression, we merged the results from all sub-data sets within each data set (e.g., 1.1–1.4 for data set 1) and only considered GFs that exhibited significant scores ($P < 0.01$) for all of them. All other parameters were left at default.

MSci Analyses

We used BPP (Flouri et al. 2018, 2020) to fit Multispecies Coalescent (MSC) and MSci models to the 3 competing topologies with different scenarios of post-speciation admixture (including no admixture in the MSC model). We created a data set with 2 individuals per species in our focal trio and one individual for each remaining species in this analysis. We sampled a short coding segment of at least 500 bp (mean 1,107.83 bp) from every other 100 kb GF so that there was at least 100 kb space between consecutive samples, resulting in a total of 2,455 loci. We compiled 2 independent data sets to assess consistency, using odd- and even-numbered GFs. The alternative data set consisted of 2,407 loci with a mean length of 1091.27 bp. Since our data set comprised pseudohaploid genomes (see above), that is, alternate alleles from heterozygous sites were randomly chosen, we used the diploid option set to 0 in BPP to consider this feature in the models. BPP implements the full-likelihood method and accommodates phylogenetic uncertainties and errors in the gene trees (Rannala et al. 2020).

For all settings, we used the JC substitution model, and inverse-gamma priors were assigned to parameters θ (population size) and τ_0 (the root age), with $\theta \sim$ IG(3, 0.003) and $\tau_0 \sim$ IG(3, 0.024). The uniform prior U(0, 1) was used for parameter φ (introgression probability). We used burn-ins of 8×10^4 iterations and generated 8×10^4 samples, sampling every 10 iterations. We conducted 4 independent runs for each scenario to assess convergence and consistency among runs. For the MSC (no admixture) models, all settings remained the same as in the case of admixture, except that the chains were shorter with burn-in set to 4×10^4 and 4×10^4 generated samples, again sampling every 10 iterations.

To compare different models with and without admixture, we calculated their marginal likelihoods (Bayes factors) with thermodynamic integration using Gaussian quadrature (Lartillot and Philippe 2006; Rannala and Yang 2017) with 32 quadrature points. We then used the best-fit models to calculate the ages of the divergence episodes among extant *Panthera* lineages by using the ratio of the estimated τ (tau) values for each node in the topology to τ_0 (the root age) and calibrating the root age (divergence between *Panthera* and the domestic cat) at 11.5 Ma (Li et al. 2016a).

Results

Phylogenetic Analyses

Data set features and tests of species-level monophyly.—The genomes were mapped against the domestic cat reference assembly and presented a depth of coverage profile suitable for all the analyses performed in this study (Supplementary Fig. S1). The final alignment matrix spanned an average of 52.4% of each reconstructed genome (after repeat masking). Our analyses recovered well-supported monophyly for each species across our genome-wide alignment when considering only GFs in which the bootstrap was above 90% (Supplementary Table S1). Analyses that included all 3 jaguars yielded 99.9% of GFs supporting species-level monophyly for all 3 focal species. Within jaguars, individuals 2 and 3 were the most frequent grouping (45%), followed by jaguars 1 and 2 (28.2%) and 1 and 3 (26.7%).

The observed monophyly at the species level agreed with lower intraspecific genetic diversity than sister species divergence (Supplementary Fig. S4). Jaguars presented the highest level of intraspecific diversity (given the current set of individuals), either for the data set comprising only jaguars 1 and 2 (Supplementary Fig. 4a) or jaguars 1, 2, and 3 (Supplementary Fig. S4b).

Topological variation across the genome.—The choice of genome fragment size influenced the inferred pattern of topological variation. In most cases, we observed all 3 topologies; however, in some chromosomes and GF sizes, topologies 2 and 3 were rare or not even retrieved (Fig. 1 and Supplementary Table S2). GFs spanning 50 kb each (n = 17,411) encompassed 35.4% of the global alignment and retrieved topologies 1, 2, and 3 with frequencies of 76.1%, 8.4%, and 6.4%, respectively, at the 90% bootstrap threshold (Fig. 1 and Supplementary Table S2). The 100-kb GFs (n = 13,432) represented 54.6% of the alignment (Fig. 1 and Supplementary Table S2) and showed an increasing dominance of topology 1 (84.2% vs. 6.9% and 4.5% for topologies 2 and 3, respectively). When longer GFs were employed, topology 1 was even more predominant: the frequencies of the 3 topologies were 92.5%, 5.7%, and 1% for 1-Mb GFs (which represented 92.6% of the alignment), and 95.3%, 3.9%, and 0% for 5-Mb GFs (which represented 99.2% of the alignment). Other topologies were also observed at lower frequencies (see Fig. 1), for example, with the leopard at a basal position in *Panthera* (see Figueiró et al. [2017]). They were not explored further, given the focus of this study. Since the overall spatial patterns were congruent among the different GF sizes, we observed a tradeoff between phylogenetic resolution (i.e., larger GFs contained more informative sites and could thus have their dominant topology resolved more conclusively) and sensitivity to genealogical discordance (i.e., smaller GFs provided finer-scale resolution of the phylogenetic landscape). We selected 100-kb GFs for some in-depth analyses. This included the introgression analyses performed with D_{FOIL} (see "Introgression Analyses with D_{FOIL}") and an assessment of topological frequency across the genome. Topology 1 was dominant on every chromosome (Table 1), even chromosome X, on which this pattern was less pronounced (Fig. 1).

Interestingly, topology 1 was highly predominant in distal chromosomal regions, congruent with areas of higher recombination rates (Fig. 1). In contrast, topologies 2 and 3 were more frequent in central portions of the chromosomes, and their local enrichment matched segments with lower-recombination rates (Fig. 1 and Supplementary Figs. S5–S12). Although the spatial patterns varied among chromosomes, the relationship with recombination rates was consistent. For example, chromosome E3 has mostly high recombination rates throughout its length and a very strong predominance of topology 1 (Fig. 1). Such trends could be detected with all different GF sizes but could be discerned more clearly with the finer-scale assessment allowed by the 50-kb and 100-kb segments.

The "species tree" approach implemented in ASTRAL-II retrieved topology 1 as the species tree for all scenarios. This included the joint analysis of all GFs across the genome, all GFs for autosomes only, and each chromosome separately, including the X chromosome. Furthermore, when applying no or one reticulation in PhyloNet, we also retrieved topology 1 as the "species tree" (Supplementary Fig. S37).

Divergence Estimates Based on Individual GFs

In addition to the observation that intraspecific diversity was consistently lower than sister species divergence, fitting the strong support for genome-wide species-level monophyly (see above), there was a clear pattern of increased diversity and divergence in telomeric regions of all chromosomes (Supplementary Fig. S4e–g). Moreover, we observed that GFs harboring topology 1, enriched in high-recombination regions, exhibited significantly higher divergence between sister species than those harboring topologies 2 and 3. These results indicated strong recombination-driven effects on divergence estimates, highlighting the importance of controlling for them when estimating relative node ages among the competing topologies.

We estimated node ages with 10-kb GFs spaced from each other by 190 kb to minimize intra-fragment recombination and increase independence among adjacent samples. We initially observed that GFs containing topology 1 presented larger sister species sequence distances than those harboring topologies 2 and 3, with significant differences in Chromosome X (Supplementary Figs. S13–S31 and Supplementary Tables S3 and S4). However, when we calculated "normalized" relative ages by dividing the sister-pair divergence by that of the preceding (trio) node, compensating for local recombination effects, we observed a consistent pattern across all chromosomes (Supplementary Figs. S13–S31 and Supplementary Table S3) in which topology 1 exhibited a significantly younger relative date for its sister-pair (lion-leopard) than topologies 2 or 3 (Table 1).

Topologies 2 and 3 did not differ significantly in most chromosomes (Table 1). Furthermore, we observed a consistent visual association of the younger dates for the lion-leopard relationship with regions of high recombination when plotting the relative ages along the chromosomes (Supplementary Figs. S13–S31).

Introgression Analyses With D_{FOIL}

We then assessed genome-wide patterns of D-statistics to investigate the relationships between the observed genealogical discordance and inferences of introgression. The D_{FOIL} results indicated that introgression was highly prevalent throughout the genomes, regardless of the topology assumed as the species tree (Fig. 2). The number of 100-kb GFs exhibiting some significant signatures of introgression ranged from ~3,200 GFs when assuming topology 1 to > 12,000 GFs when assuming topology 3.

Our multiple data sets assessed the effects of varying the individuals representing each species, which yielded interesting insights. For example, when assuming topology 1 and using all 3 combinations of jaguar genomes as P_1 and P_2, along with all combinations of lion and leopard as P_3 and P_4, we observed consistent patterns of introgression, with ca. 2-fold more GFs with signatures of admixture between jaguars and lions than between jaguars and leopards (Fig. 2a). There was variation in the detection of introgression in some GFs depending on the individuals used, for example, each data set uniquely detected ~100 GFs with lion-jaguar introgression, which did not yield significant results in the other data sets with topology 1 (6.0–6.5%). However, there were > 780 GFs with signatures of lion-jaguar introgression detected by 2 different data sets (13.8–16.0%) and > 1,100 GFs detected by all 3 (62.6–64.2%). Using a conservative assessment that only considered significant signatures of introgression identified with all other data sets (and all their respective sub-data sets), we still observed > 150 GFs with some signal of introgression between lion and jaguar or leopard and jaguar.

Consistent signals of introgression across data sets (with varying individuals) were also observed when we assumed topologies 2 or 3 as the species tree (Fig. 2b,c and Supplementary Fig. S3). In these cases, we used the 2 leopards or the 2 lions as P_1 and P_2, respectively, allowing for the detection of admixture affecting their common ancestral branch (Fig. 2). Although there were some differences in the detection of introgression when varying the individuals (i.e., some GFs bore significant signatures with only 1 or 2 data sets), there was a substantial amount of congruence with most of the admixture signals being captured by all 3 data sets (85.6–86.3%). The admixture results were striking with both topologies, even when only considering GFs detected by all data sets. In both cases, there were > 8,500 GFs with significant signatures of lion-leopard admixture and others that implicated the jaguar in hybridization processes.

The genome-wide spatial distribution of introgressed GFs also presented intriguing patterns. When assuming topology 1 (Fig. 2a), the blocks containing significant signatures of introgression were primarily present in the central portion of chromosomes, correlating with the low recombination regions enriched for topologies 2 and 3 (Supplementary Figs. S32–36). Conversely, when assuming topologies 2 or 3, the signatures of introgression were more concentrated in distal portions of the chromosomes, correlating with the higher recombination regions enriched for topology 1 (Supplementary Figs. S32–36).

Multispecies Coalescent Analysis With and Without Introgression (MSC and MSci Models)

As an assessment of the processes underlying the observed genealogical discordance, we used BPP to compare the 3 competing topologies with different scenarios of post-speciation admixture (Fig. 3). The 2 best-fitting models were topologies 2 and 3 with bidirectional introgression between lion and leopard (log marginal likelihoods of −4,063,179.4 and −4,063,175.3, respectively). These 2 models are unidentifiable, so their marginal likelihood difference reflects sampling errors in the Bayes factor calculation. We note that the models assuming no introgression (MSC models) provided the worst fit to the data. All models assuming topology 1 (even accounting for admixture) were also poorly supported.

We used the posterior means of divergence times (t) at each node, calibrated by fixing the root age (*Panthera* vs. domestic cat divergence) at 11.5 Ma, to estimate the timing of separation among extant *Panthera* lineages (Fig. 4) for both best-fitting models. Identical dates were obtained with both models (see below): the base of crown *Panthera* was estimated at 5.2 Ma, closely followed (at 4.98 Ma) by the base of the lion-leopard-jaguar trio; tiger and snow leopard diverged from each other at 3.55 Ma. Within the lion-leopard-jaguar trio, both models supported the initial separation of a lineage identified as y in Fig. 4, followed by the split between another lineage (x) and the jaguar 3.87 Ma. There was then intense admixture between lineages × and y, giving rise to the present-day genomes of lion and leopard. The estimated introgression parameters between these lineages were as follows: genomic contribution from × into the leopard (parameter α in Fig. 4) was 0.76, with the 95% highest posterior density (HPD) CI estimated at 0.71–0.82; genomic contribution from y into the lion (parameter β in Fig. 4) was 0.18, with an HPD CI of 0.12–0.22.

These 2 best-fitting models (Figs. 3 and 4) are unidentifiable with our data set because they predict the same statistical distributions of the data. Consider tracing back the history of a sample of present-day sequences from the leopard. In topology 2, when they reach t_x, they follow branch x (or branch t-x) with a probability a of 76% and coalesce at the rate $2/q_x$ before joining jaguar at 3.87 Ma (t_t). In topology 3, the leopard sequences

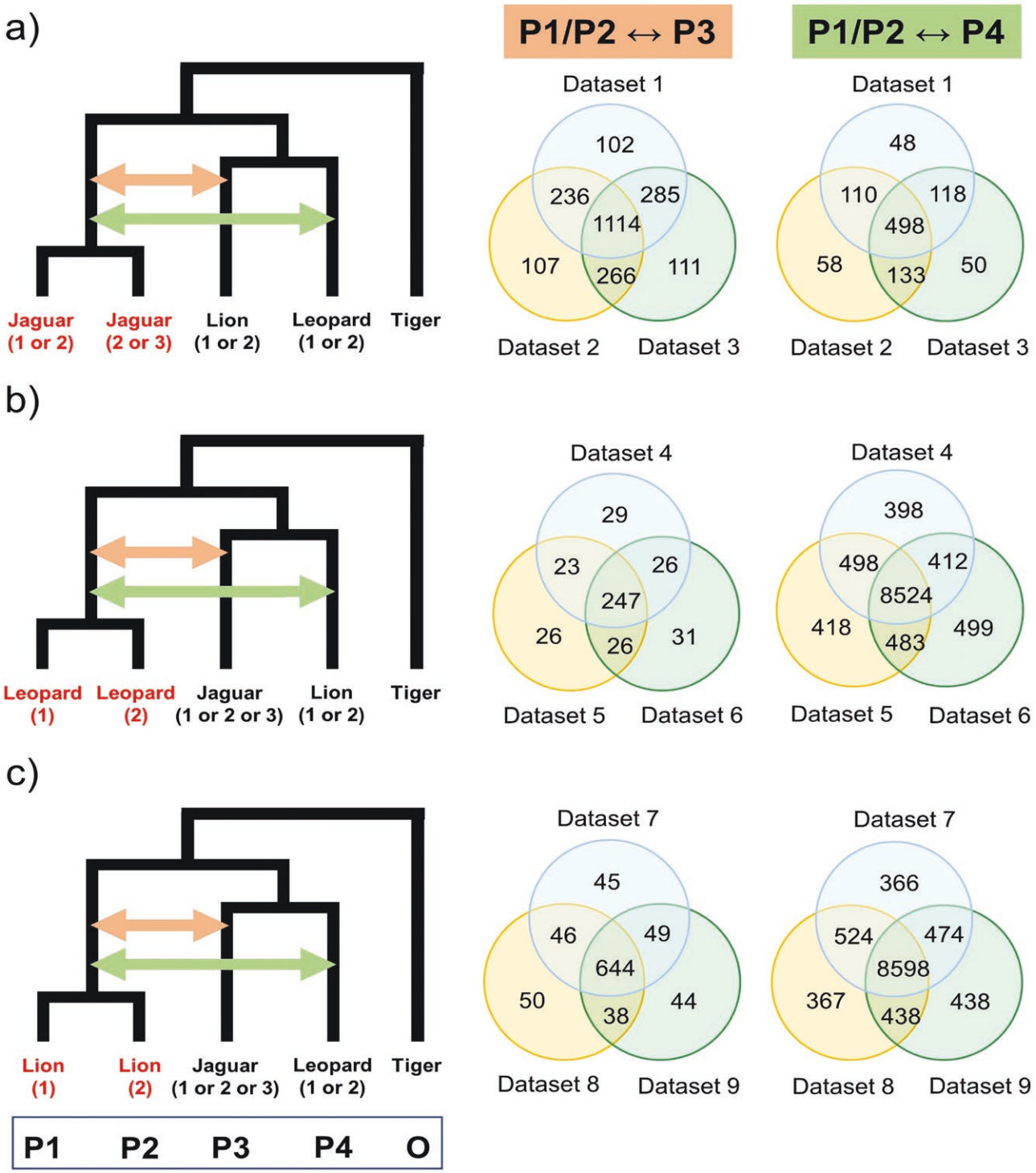

FIGURE 2. Results from the D_{FOIL} analyses for all possible combinations of individuals. The following were used as P_1 and P_2: (a) all 3 jaguars (Data sets 1–3) and assuming topology 1; (b) the 2 leopards (Data sets 4–6) and assuming topology 2; and (c) the 2 lions (Data sets 7–9) and assuming topology 3. Numbers in parentheses (e.g., Jaguar 1 or 2) represent the different individuals of each species. All combinations of individuals used as P_3 and P_4 were also assessed and considered sub-data sets within each data set (see Supplementary Fig. 3 for the exact composition of all sub-data sets). Only GFs with significant evidence of introgression for all sub-data sets within a data set were considered. Trees on the left depict the topological structure of each data set, with colored arrows indicating the branch combinations that were assessed for introgression. Venn diagrams on the right illustrate the number of 100 kb GFs with significant signatures of introgression for each data set and each branch combination (labeled at the top).

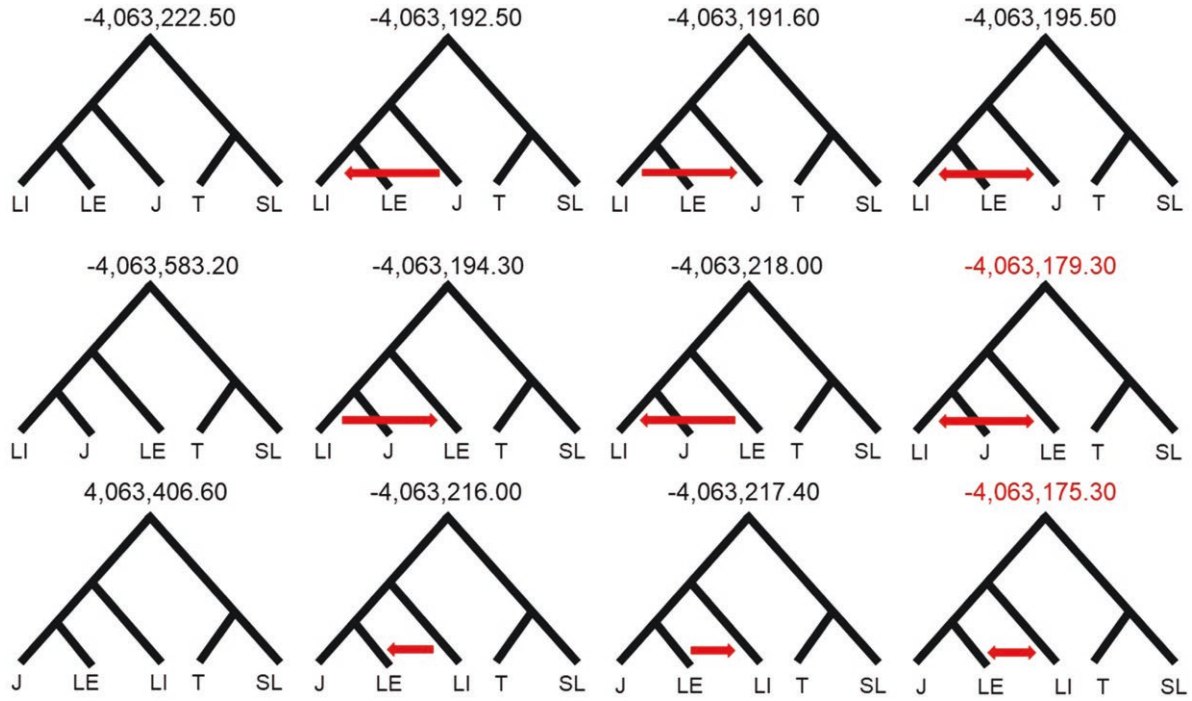

FIGURE 3. Log marginal likelihoods for different models of species divergence (topology) and post-speciation admixture calculated with BPP using thermodynamic integration with Gaussian quadrate (32 quadrature points). Results are shown for topology 1 (top), topology 2 (middle), and topology 3 (bottom), with no admixture (MSC model) on the left, unidirectional admixture combinations in the 2 central columns (the arrow indicates the direction), and bidirectional admixture on the right. Species are coded as LI: lion; LE: leopard; J: jaguar; T: tiger; SL: snow leopard. The respective log marginal likelihood is indicated above each model (best-fit models are shown in red font [gray font when seen in black-and-white]).

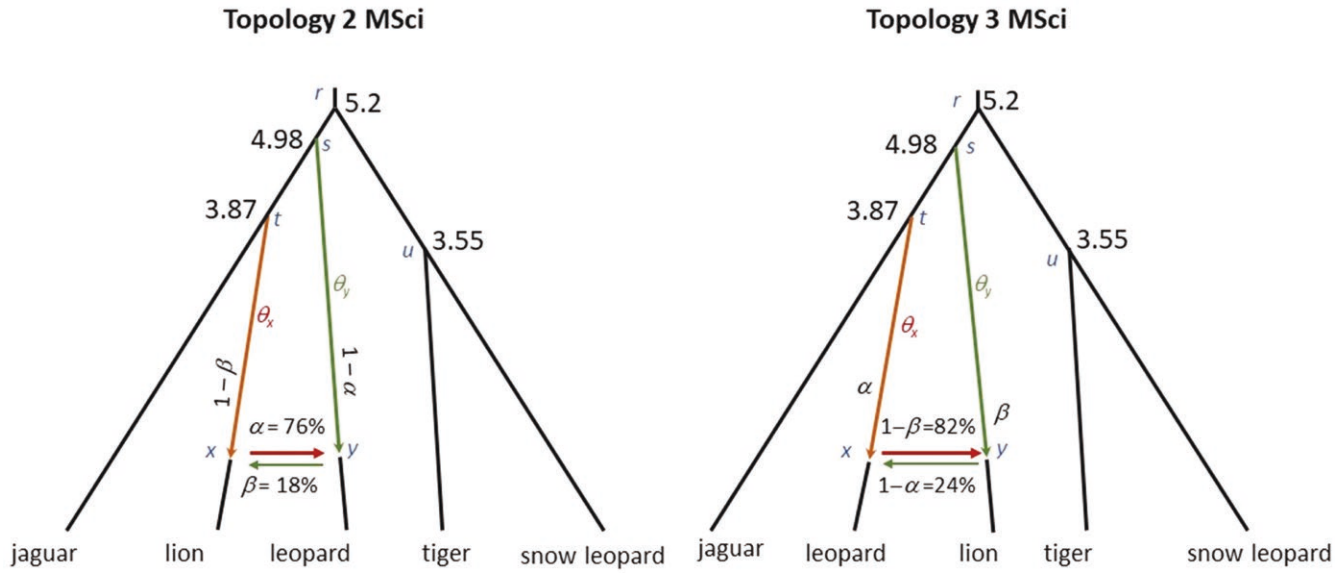

FIGURE 4. Unidentifiability of the 2 best-fitting MSci models based on topologies 2 and 3. There was strong evidence for bidirectional introgression between lion and leopard in both models (the genomic proportion arising from introgression in each direction is indicated by the colored arrows connecting these lineages in each tree). Numbers placed on nodes are divergence times (in a million years ago [Ma]) estimated in the respective model (compensating for subsequent episodes of introgression) for each topology; branch lengths are scaled to time.

again follow branch × at $t_x = t_y$ with a probability a of 76% and coalesce at the rate $2/q_x$ before joining the jaguar lineage at 3.87 Ma (t_t). In both models, the leopard sequences follow branch y with a probability of 24% (1–a) at t_x and coalesce at the rate $2/q_y$ before reaching node s at 4.98 Ma (t_s). The 2 models also make the same

predictions for a sample of lion sequences or any combinations of lion and leopard sequences. Both models predict that today's lions trace 82% of their genes back to lineage *tx* and 18% to lineage *sy*, while today's leopards trace 76% of their genes back to lineage *tx* and 24% to lineage *sy*. In other words, lineage *tx* accounts for nearly 80% of the genomes of lions and leopards, while lineage *sy* accounts for ca. 20%.

Discussion

Identifying the True Speciation History in the Face of Massive Admixture

Distinguishing between a speciation model with complete isolation from one in which inter-species gene flow dominated the genome may be a major methodological challenge. Our results demonstrate that the notorious difficulty in resolving the *Panthera* phylogeny has been driven by this problem, but that reliable inference can be achieved with the use of multiple, complementary phylogenomic approaches that incorporate hybridization. We were able to refute topology 1 (which is massively prevalent in the genome) as the species tree, while discovering that the 2 alternative topologies are unidentifiable due to the complex history of admixture in this clade. Below, we discuss our findings in detail, along with their implications for the development of best practices in phylogenomic inference in groups with widespread historical admixture.

Panthera Phylogenomic Relationships

The whole-genome data set analyzed here allowed an in-depth assessment of the complex evolutionary relationships among lions, leopards, and jaguars. The different approaches complemented and corroborated one another, providing a better understanding of the evolution of *Panthera*. Initially, we tested species-level monophyly, which was strongly supported for the 3 focal species across their genomes and might compensate for potential biases derived from using pseudohaploid genomes in downstream analyses. Similar findings of species-level monophyly have been observed for groups such as whales, canids, and wild sheep (Árnason et al. 2018; Gopalakrishnan et al. 2018; Santos et al. 2021; Upadhyay et al. 2021), suggesting that this may be a broad pattern for mammals.

We then assessed inter-species relationships by partitioning the alignments into GFs of different lengths. Topology 1 was more common than topologies 2 or 3 for all GF sizes (Table 1). Smaller GF sizes revealed more detailed patterns of genealogical discordance but often showed less support for the local phylogenetic reconstruction. Conversely, larger GFs contained more phylogenetic information, but their averaging effect swamped the information on regional discordance. Similar patterns were observed when employing 100-kb GFs for chromosome A3 (e.g., Fig. 1b) or 10-kb GFs for all chromosomes without the 90% bootstrap threshold.

We also observed that topology 1 was consistently recovered as the "species tree" when employing the multispecies coalescent method with ASTRAL-II, which was undoubtedly a consequence of it being overwhelmingly prevalent in the genome (Table 1). ASTRAL-II builds a species tree based on the conflicting gene trees due to ILS but does not account for other evolutionary processes, such as admixture (Mallo and Posada 2016). Similarly, when using no reticulation (or even allowing one reticulation) in PhyloNet or the Multispecies Coalescent model without introgression in BPP, topology 1 was also favored (Fig. 3).

Comparing the positions representing the species tree or alternative topologies to recombination rates across the genomes can confirm or provide evidence that further investigation is needed. Recombination rates appear to influence the distribution of topologies (Elworth et al. 2019; Edelman et al. 2019; Li et al. 2019; Hennelly et al. 2021; Pilot 2021). Higher recombination rates typically retain introgression blocks more easily than low-recombining regions, where the true speciation history is observed. The relationship with recombination rates was better demonstrated when employing GF sizes of 50 and 100 kb (Fig. 1 and Supplementary Figs. S5–S12): topologies 2 and 3 were more often present in lower-recombination regions compared to topology 1, as illustrated by chromosome A3 (Fig. 1b). A fine-scaled approach may allow an in-depth dissection of the history of particular chromosomal segments. These initial steps demonstrated the complex relationship among lions, leopards, and jaguars.

Divergence Time Estimates

Given the observation that the less-frequent topologies (2 and 3) were enriched in genomic regions with lower-recombination rates, where we expect the species tree to be preferentially retained, we further explored this problem by estimating divergence times based on each focal tree. The expectation was that topologies derived from introgression would have proportionally younger ages for the internal (sister species) node relative to alternative topologies (species tree or ILS-induced tree). The initial (uncorrected) result indicated that topology 1 exhibited the oldest divergence times for the sister species pair, followed by topologies 3 and 2 (as also observed by Figueiró et al. 2017 and Li et al. 2019). Although this is a consistent result and would support the view that topology 1 is the species tree, its spatial distribution is at odds with this interpretation. It is improbable that repeated introgressions would occur and be retained consistently in lower-recombination regions, as these are usually less permeable to introgression (Li et al. 2019; Hennelly et al. 2021).

To further investigate this issue, we estimated "normalized" ages for the sister species node, considering the possibility that low-recombination regions have downward biases in substitution rates in *Panthera*.

These biases may not be sufficiently corrected with the domestic cat calibration, thus leading divergence times to be underestimated in those segments. Using this correction, we observed that topology 1 was consistently (and in most cases significantly) younger than topologies 2 and 3 (Table 1), which fits the expectations that it is derived from introgression and is consistent with its prevalence in high-recombination regions of the genome (Nachman and Payseur 2012; Li et al. 2019; Hennelly et al. 2021). Interestingly, we observed that GFs bearing topologies 2 and 3 had a similar recombination profile and sister species "normalized" age. The X chromosome specifically demonstrates such a pattern (Table 1) and is known to often preserve the species tree, as proposed for *Anopheles* mosquitoes (Fontaine et al. 2015; Thawornwattana et al. 2018), cat species in general (Li et al. 2019), and wolves (Hennelly et al. 2021).

Direct Inferences of Inter-species Introgression

The next step included evaluating the 3 main topologies and inferring introgression patterns between these species using different combinations of individuals in D_{FOIL}. Previous studies have used multiple individuals of the same species to assess introgression events, which increases the robustness of the analysis (e.g., Guo et al. 2019, 2022; Scott 2019; Santos et al. 2021); Upadhyay et al. 2021). We observed less congruence when using topology 1 (ca. 64%) than topologies 2 and 3 (ca. 86%). Such a result might indicate that topology 1 cannot retrieve the correct introgression patterns when varying the individuals used. We then focused on GFs with consistent results throughout the different data sets, thus obtaining more conservative estimates of introgression.

When we assumed topology 1 (i.e., jaguars as P_1 and P_2), there was an introgression signal between jaguars and lions, as well as jaguars and leopards, enriched in the central portion of chromosomes, with lower-recombination rates (Fig. 3a and Supplementary Figs. S32–S36). This spatial pattern did not fit the recombination-based expectation, since these regions are usually less permeable to introgression. When placing lions or leopards as P_1 and P_2 (i.e., assuming topologies 2 or 3), we observed extensive introgression signals between lions and leopards (Fig. 3b,c) enriched in high-recombination segments (Supplementary Figs. S32–S36), which fits the recombination-based expectation. Interestingly, when the X chromosome was assessed in detail, there was little signature of introgression in the recombination desert highlighted in our previous studies (Figueiró et al. 2017; Li et al. 2019). Since this was the case regardless of the assumed species tree, this region possibly had insufficient signals for D_{FOIL} to detect significant signatures of introgression. It harbors less variation due to its lower recombination and substitution rates (Li et al. 2019). These analyses support the interpretation that topology 1, despite being vastly prevalent across the genome, probably results from post-speciation introgression between lions and leopards, implying a large-scale genomic replacement due to secondary admixture.

The same inference emerged from the BPP analyses, which revealed a much more detailed (and complex) picture of the evolutionary history of these *Panthera* species. Most implementations of full-likelihood models accounting for gene flow are too expensive computationally to apply to data sets with more than 100 or 200 loci (Hey et al. 2018; Wen and Nakleh 2018; Zhang et al. 2018). BPP allowed us to leverage the power of whole-genome data and to perform marginal likelihood calculations to directly compare the different models of lion-leopard-jaguar relationships (Fig. 3). This showed that, even though topology 1 is overwhelmingly present in the genome, our estimates unequivocally supported topologies 2 or 3 with extensive lion-leopard introgression and refuted topology 1 as the species tree. Previously, Mallet et al. (2016) have pondered if one could distinguish between models of complete isolation from those where extreme gene flow dominated the genomes. We suggest that the answer to this question is "yes," if one leverages the information contained in genome-wide patterns of genealogical discordance and the use of methods that consider/incorporate introgression. By using such approaches, we were able to refute topology 1 as the species tree and conclude that it has been induced by extensive hybridization between the ancestors of lions and leopards.

Interestingly, the models assuming topologies 2 and 3 were unidentifiable, that is, they provided an equal fit to the data (Fig. 3). This is because they reconstruct the same evolutionary history for this set of *Panthera* species, involving ancestral lineages that hybridized extensively, subsequently giving rise to modern-day lions and leopards (Fig. 4). Such a complex evolutionary history, including pervasive hybridization and subsequent species formation, poses great challenges to phylogenetic reconstruction and divergence dating, which should be performed using approaches that incorporate historical gene flow among lineages. In this context, the divergence dates shown in Fig. 4, which take into account this complex history of inter-species admixture, likely represent a more realistic assessment of the timing of *Panthera* divergences than previously reported estimates. More broadly, these results illustrate the power of genome-scale data sets and complete likelihood methods to address such complex evolutionary problems.

Historical Biogeography

Biogeographic considerations are relevant while discussing our findings and alternative scenarios for the inferred speciation and admixture events. Previous studies have suggested that these species diverged either in Asia (Fig. 5a; e.g., Johnson et al. 2006; O'Brien and Johnson 2007; Mazak 2010; Tseng et al. 2014) or in Africa (Fig. 5b–d; e.g., Hemmer et al. 2010; Werdelin et al. 2010). An Asian origin of all 3 species is the most plausible scenario, given the available evidence. This scenario would imply waves of dispersal of the ancestors of lions and leopards to Africa and Europe and

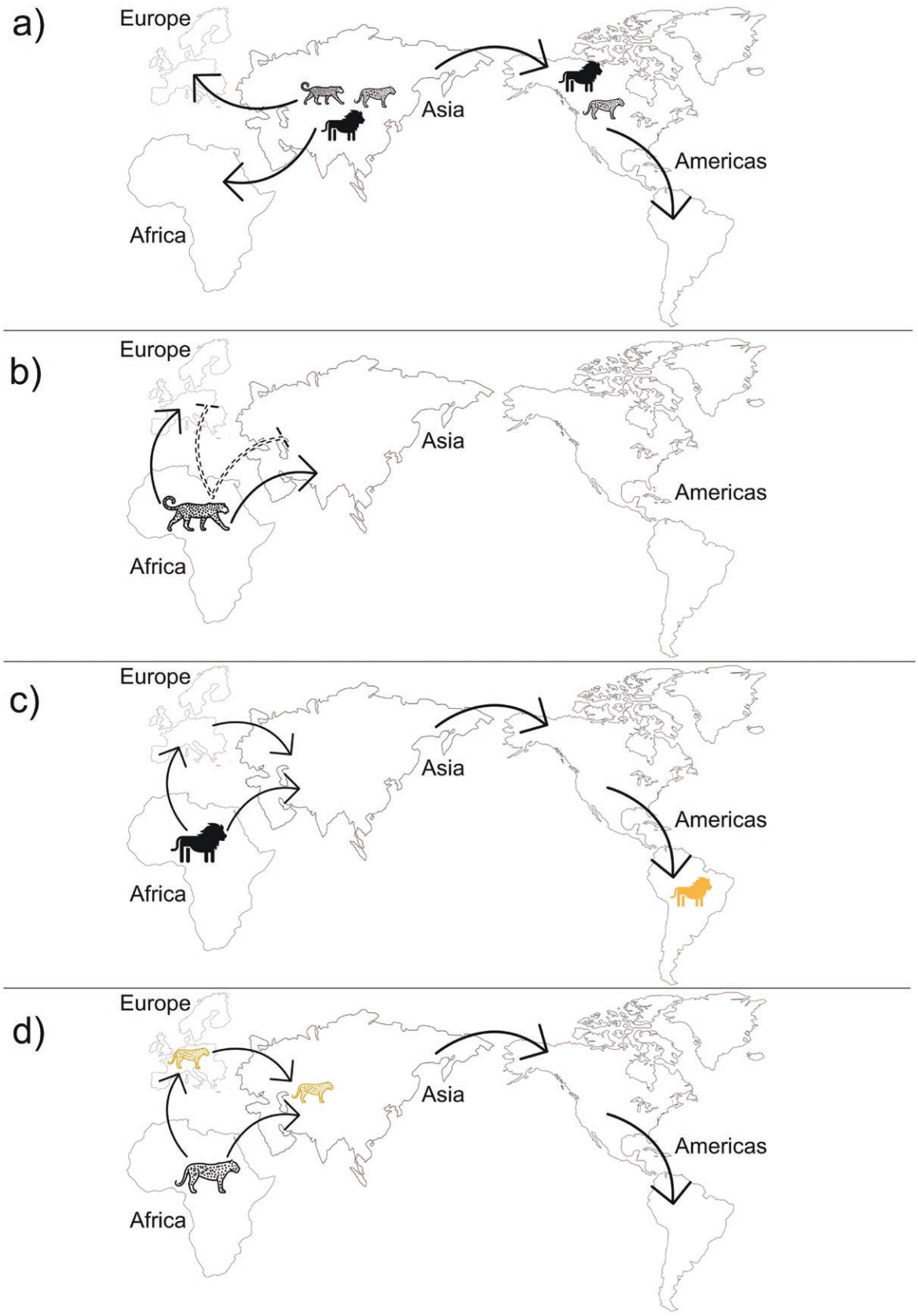

FIGURE 5. Possible origin and dispersal scenarios for the focal species based on the fossil record. The first scenario includes all *Panthera* species originating from central/northern Asia, with the ancestors migrating to other regions afterwards (Johnson et al. 2006; O'Brien and Johnson 2007; Mazak 2010; Tseng et al. 2014). The second scenario includes an African origin and respective expansion for (b) leopards

of those of lions and jaguars to the Americas (Fig. 5a; Johnson et al. 2006; O'Brien and Johnson 2007; Mazak 2010; Tseng et al. 2014).

Regardless of the assumed biogeographic scenario, our genomic data strongly indicate sufficient sympatry between the ancestors of these lineages to allow extensive admixture among them. The ancestors of modern lions and leopards hybridized extensively, perhaps repeatedly and on multiple occasions, leading to the overwhelming signal observed here (Figs. 1–4). The reduced introgression signal between leopards and jaguars may have been influenced by their lack of geographical overlap after divergence and dispersal (Uphyrkina et al. 2001; Jacobson et al. 2016). Moreover, the less frequent but detectable introgression signal between them might have been passed on by lions (i.e., intermediary species admixture), given their broad distribution. This hypothetical scenario is similar to the one postulated for the widely distributed brown bears, potentially serving as a conduit for genetic exchange between the ancestors of polar or American black bears and Asian black bears, whose distributions do not currently overlap (Kumar et al. 2017).

Fossils of *Panthera* are rare, fragmented, and one or more ghost ancestors could be missing (Werdelin et al. 2010; Tseng et al. 2014), complicating the interpretation of their origin and subsequent dispersal. In addition, several fossils are difficult to reliably place phylogenetically, given the conservation of skull morphology and likely instances of evolutionary convergence in this group. Therefore, additional paleontological studies and continued attempts to obtain genomic data from extinct *Panthera* lineages are needed, especially for controversial lineage. Bears, for example, have blocks of introgression between living and extinct species (Barlow et al. 2018; Cahill et al. 2018), a scenario that is very likely for *Panthera*, especially given the genomic evidence for "fossil" topologies that imply such admixture events (e.g., trees placing leopard as the most basal divergence in the genus [Figueiró et al. 2017]). In this sense, even though elements in the fossil record fit the scenario of lions or leopards and jaguars being sister species, several aspects of this complex history should be further explored. Hopefully, additional fossil-based and genomic analyses, including testing more complex evolutionary models, will help shed further light on these processes. Overall, we can conclude that *Panthera* is an increasingly interesting system for investigating the interplay among different evolutionary forces.

Broader Considerations: Dissecting Complex Evolutionary Histories

The advancement of whole-genome technologies and novel analytical methods can shed new light on the evolution of species across the Tree (or Web) of Life, which may be particularly relevant in the case of groups that have been recalcitrant to conclusive phylogenetic reconstruction. We propose here some recommendations that can be considered when addressing such problems.

A critical aspect is to employ whole-genome data, whenever possible comparing different fragment sizes, the spatial distribution of topologies, and its relationship with recombination rates. The relative frequency of topologies can be an important driver of global (and potentially erroneous) inferences, especially when employing methods that do not incorporate hybridization. For example, we showed here that topology 1, now refuted as the species tree, is extremely prevalent throughout the genome, which led it to be retrieved as the speciation topology by ASTRAL-II, PhyloNet, and BPP (under the MSC model). It is noteworthy that even PhyloNet, when allowed to incorporate one reticulation, still retrieved this topology as the most likely species tree.

The relative age of lineage divergence in alternative topologies is one of the criteria that can be used to distinguish the true species tree from other reconstructions. However, the consistent occurrence of alternative topologies in genomic regions with varying recombination rates can also impact local substitution rates and thus bias the inference of divergence times, which may mislead such a comparison. The local correction applied here was able to compensate for this effect, reconstructing a younger age for topology 1, which is consistent with the results from the other analyses.

Another aspect to be considered is that inferring introgression signals with D-statistics can be impacted by the assumed species tree. In addition, using multiple individuals per species is an interesting approach to assess the consistency of such signals. When we assumed topology 2 or 3 as the species tree and used different combinations of individuals, we observed more consistent results when compared to assuming topology 1 (which was revealed no to be the true species tree). The generality of this observation can be further investigated in additional systems with the use of a varying set of individuals per species.

Finally, we observed that the genealogical history of the focal trio of *Panthera* was so complex that only the full-likelihood approach implemented in BPP was able to

(Uphyrkina et al. 2001; Paijmans et al. 2018), (c) lions (Yamaguchi et al. 2004; Barnett et al. 2006; de Manuel et al. 2020), and (d) jaguars (Hemmer et al. 2010; Argant and Argant 2011; Ruiz-Ramoni et al. 2020). (b) Ancestors of the modern leopard migrated to Eurasia from Africa. A theory postulates that they went extinct at some point (dashed lines), except in Africa, from which they later migrated to Eurasia once more (black arrow; Uphyrkina et al. 2001). (c–d) Ancestors of lions and jaguars migrated to multiple regions, including Europe, Asia, and the Americas. Yellow species icons (lighter gray when visualized in black-and-white) are those whose fossil records present conflicting species designations. The *Panthera atrox* fossil (c), an extinct lion in South America, may actually be a jaguar (Christiansen and Harris 2009). The Eurasian *Panthera gombaszogensis* (d) might be more related to other *Panthera* species (Chatar et al. 2022) as opposed to an immediate jaguar relative. The icons used to represent a leopard (Anniken and Andreas) and a jaguar (Yu Luck) are from the Noun Project (CCBY3.0: https://thenounproject.com/).

completely clarify the effects of massive historical admixture on the genomic composition of present-day species. The results from this method were completely consistent with the ones we derived from the other approaches (i.e., refuting the most prevalent topology as the species tree), but further revealed a history in which none of the assumed trees can truly reflect the actual genealogy of this clade, given that massive admixture between 2 ancient lineages generated 2 present-day species. Taken together, our results illustrate how multiple analytical approaches can be employed and compared to understand complex genealogical histories, and may serve as a basis to design strategies to investigate other taxonomic groups that present challenging phylogenetic problems.

SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository: https://dx.doi.org/10.5061/dryad.gtht76hwr

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

Funding

We thank the Brazilian National Council for Scientific and Technological Development (CNPq) for financial support (grant numbers 487396/2012-0, 310803/2015-2 and 309068/2019-3). This study is a contribution of the National Institutes for Science and Technology (INCT) in Ecology, Evolution, and Biodiversity Conservation, supported by MCTIC/CNPq/Brazil (proc. 465610/2014-5) and FAPEG/Brazil (proc. 201810267000023). This study was supported by Biotechnology and Biological Sciences Research Council grants (BB/T003502/1 and BB/X007553/1) to T.F. and Z.Y., and a Natural Environment Research Council (NERC) grant (NSFDEB-NERC NE/X002071/1) to Z.Y.

Data availability

The data used in this study are publicly available in the National Center for Biotechnology Information database, including the domestic cat v.9.0 assembly (GCA_000181335.4). The short-read sequences were obtained from the following accession numbers: (a) jaguars 1-3: PRJNA348348; (b) leopard 1: PRJNA380636; (c) leopard 2: PRJNA306701; (d) lions 1-2, snow leopard and tiger: PRJNA182708.

References

Ackerman R.R., Arnold M.L., Baiz M.D., Cahill J.A., Cortés-Ortiz L., Evans B.J., Grant B.R., Grant P.R., Hallgrimsson B., Humphreys R., Jolly C.J., Malukiewicz J., Percival C.J., Ritzman T., Roos C., Roseman C.C., Schroeder L., Smith F.H., Warren K., Wayne R., Zinner D. 2019. Hybridization in human evolution: insights from other organisms. Evol. Anthropol. 28(4):189-209. doi: 10.1002/evan.21787.

Andrews S. 2010. FASTQC: a quality control tool for high throughput sequence data. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Argant A., Argant J. 2011. The *Panthera gombaszogensis* story: the contribution of the Château breccia (Saône-et-Loire, Burgundy, France). Quaternaire 4:247-269. https://hal.science/hal-01778589.

Árnason Ú., Lammers F., Kumar V., Nilsson M.A., Janke A. 2018. Whole-genome sequencing of the blue whale and other rorquals finds signatures for introgressive gene flow. Sci. Adv. 4(4):eaap9873. doi: 10.1126/sciadv.aap9873.

Arnold M.L. 2015. Divergence with genetic exchange. Oxford (England): Oxford University Press.

Barlow A., Cahill J.A., Hartmann S., Theunert C., Xenikoudakis G., Fortes G.G., Paijmans J.L.A., Rabeder G., Frischauf C., Grandal-d'Anglade A., García-Vázquez A., Murtskhvaladze M., Saarma U., Anijalg P., Skrbinsek T., Bertorelle G., Gasparian B., Bar-Oz G., Pinhasi R., Slatkin M., Dalén L., Shapiro B., Hofreiter M. 2018. Partial genomic survival of cave bears in living brown bears. Nat. Ecol. Evol. 2(10):1563-1570. doi: 10.1038/s41559-018-0654-8.

Barnett R., Yamaguchi N., Barnes I., Cooper A. 2006. The origin, current diversity and future conservation of the modern lion (*Panthera leo*). Proc. Biol. Sci. 273(1598):2119-2125. doi: 10.1098/rspb.2006.3555.

Bolger A.M., Lohse M., Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114-2120. doi: 10.1093/bioinformatics/btu170.

Buckley R.M., Davis B.W., Brashear W.A., Farias F.H.G., Kuroki K., Graves T., Hillier LW, Kremitzki M., Li G., Middleton R.P., Minx P., Tomlinson C., Lyons L.A., Murphy WJ, Warren W.C. 2020 A new domestic cat genome assembly based on long sequence reads empowers feline genomic medicine and identifies a novel gene for dwarfism. PLoS Genet. 16(10):e1008926. doi: 10.1371/journal.pgen.1008926.

Bursell M.G., Dikow R.B., Figueiró H.V., Dudchenko O., Flanagan J.P., Aiden E.L., Goossens B., Nathan S.K., Johnson W.E., Koepfli K.P., Frandsen P.B. 2022. Whole genome analysis of clouded leopard species reveals an ancient divergence and distinct demographic histories. iScience 25(12):105647. doi: 10.1016/j.isci.2022.105647.

Cahill J.A., Green R.E., Fulton T.L., Stiller M., Jay F., Ovsyanikov N., Salamzade R., St John J., Stirling I., Slatkin M., Shapiro B. 2013. Genomic evidence for island population conversion resolves conflicting theories of polar bear evolution. PLoS Genet. 9(3):e1003345. doi: 10.1371/journal.pgen.1003345.

Cahill J.A., Heintzman P.D., Harris K., Teasdale M.D., Kapp J., Soares A.E.R., Stirling I., Bradley D., Edwards C.J., Graim K., Kisleika A.A., Malev A.V., Monaghan N., Green R.E., Shapiro B. 2018. Genomic evidence of widespread admixture from polar bears into brown bears during the last ice age. Mol. Biol. Evol. 35(5):1120-1129. doi: 10.1093/molbev/msy018.

Chatar N., Michaud M., Fischer V. 2022. Not a jaguar after all? Phylogenetic affinities and morphology of the Pleistocene felid *Panthera gombaszoegensis*. Pap. Palaeontol. 8(5):e1464. doi: 10.1002/spp2.1464.

Christiansen P., Harris J.M. 2009. Craniomandibular morphology and phylogenetic affinities of *Panthera atrox*: implications for the evolution and paleobiology of the lion lineage. J. Vertebr. Paleontol. 29(3):934-945. doi: 10.1671/039.029.0314.

Davis B.W., Li G., Murphy W.J. 2010. Supermatrix and species tree methods resolve phylogenetic relationships within the big cats, *Panthera* (Carnivora: Felidae). Mol. Phylogenet. Evol. 56(1):64-76. doi: 10.1016/j.ympev.2010.01.036.

de Manuel M., Barnett R., Sandoval-Velasco M., Yamaguchi N., Vieira F.G., Mendoza L.Z., Liu S., Guo C., Zheng J., Zazula G., Baryshnikov G., Eizirik E., Koepfli K.P., Johnson W.E., Antunes A., Sicheritz-Ponten T., Gopalakrishnan S., Larson G., Yang H., O'Brien S.J., Hansen A.J., Zhang G., Marques-Bonet T., Gilbert M.T.P. 2020. The evolutionary history of extinct and living lions. PNAS 117(20):10927-10934. doi: 10.1073/pnas.1919423117.

de Manuel M., Kuhlwilm M, Frandsen P., Sousa V.C., Desai T., Prado-Martinez J., Hernandez-Rodrigues J., Dupanloup I., Lao O., Hallast P., Schmidt J.M., Heredia-Genestar J.M., Benazzo A., Barbujani G., Peter B.M., Kuderna L.F., Casals F., Angedakin S., Boesch C., Kühl H., Vigilant L., Langergraber K., Novembre J., Gut M., Navarro A., Carlsen F., Andrés A.M., Siegismund H.R., Scally A., Excoffier L., Tyler-Smith C., Castellano S., Yue Y., Hvilsom C., Marques-Bonet T. 2016. Chimpanzee genomic diversity reveals ancient admixture with bonobos. Science 354(6311):477-481. doi: 10.1126/science.aag2602.

de Mita S., Siol M. 2012. EggLib: processing, analysis and simulation tools for population genetics and genomics. BMC Genet. 13:27. doi: 10.1186/1471-2156-13-27.

Degnan J.H., Rosenberg N.A. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. 24(6):332-340. doi: 10.1016/j.tree.2009.01.009.

Edelman N.B., Frandsen P.B., Miyagi M., Clavijo B., Davey J., Dikow R.B., Garcia-Accinelli G., Van Belleghem S.M., Patterson N., Neafsey D.E., Challis R., Kumar S., Moreira G.R.P., Salazar C., Chouteau M, Counterman B.A., Papa R., Blaxter M., Reed R.D., Dasmahapatra K.K., Kronforst M., Joron M., Jiggins C.D., McMillan W.O., Di Palma F., Blumberg A.J., Wakeley J., Jaffe D., Mallet J. 2019. Genomic architecture and introgression shape a butterfly radiation. Science 366(6465):594-599. doi: 10.1126/science.aaw2090.

Li G., Davis B., Eizirik E., Murphy W. 2016a. Phylogenomic evidence for ancient hybridization in the genomes of living cats (Felidae). Genome Res. 26(1):1-11. doi: 10.1101/gr.186668.114.

Elworth R.A.L., Ogilvie H.A., Zhu J., Nakhleh L. 2019. Advances in computational methods for phylogenetic networks in the presence of hybridization. In: Warnow T., editor. Bioinformatics and phylogenetics, computational biology. New York City (NY): Springer Cham. p. 317-360.

Figueiró H.V. 2016. Análises genômicas da onça-pintada (*Panthera onca*): caracterização do genoma completo e investigação de regiões sob seleção através de comparações interespecíficas e populacionais [Ph.D. dissertation]. Porto Alegre (BR): Pontificia Universidade Catolica do Rio Grande do Sul. p. 184. Available from: http://tede2.pucrs.br/tede2/handle/tede/7248.

Figueiró H.V. 2019. Introgression-penguins [Source code]. Available from: https://github.com/henriquevf/Introgression-penguins.

Figueiró H.V., Li G., Trindade F.J., Assis J., Pais F., Fernandes G., Santos S.H.D., Hughes G.M., Komissarov A., Antunes A., Trinca C.S., Rodrigues M.R., Linderoth T., Bi K., Silveira L., Azevedo F.C.C., Kantek D., Ramalho E., Brassaloti R.A., Villela P.M.S., Nunes A.L.V., Teixeira R.H.F., Morato R.G., Loska D., Saragüeta P., Gabaldón T., Teeling E.C., O'Brien S.J., Nielsen R., Coutinho L.L., Oliveira G., Murphy W.J., Eizirik E. 2017. Genome-wide signatures of complex introgression and adaptive evolution in the big cats. Sci. Adv. 3(7):e1700299. doi: 10.1126/sciadv.1700299.

Flouri T., Jiao X., Rannala B., Tang Z. 2018. Species tree inference with BPP using genomic sequences and the multispecies coalescent. Mol. Biol. Evol. 35(10):2585-2593. doi: 10.1093/molbev/msy147.

Flouri T., Jiao X., Rannala B., Yang Z. 2020. A bayesian implementation of the multispecies coalescent model with introgression for phylogenomic analysis. Mol. Biol. Evol. 37(4):1211-1223. doi: 10.1093/molbev/msz296.

Fontaine M.C., Pease J.B., Steele A., Waterhouse R.M., Neafsey D.E., Sharakhov I.V., Jiang X., Hall A.B., Catteruccia F., Kakani E., Mitchell S.N., Wu Y.C., Smith H.A., Love R.R., Lawniczak M.K., Slotman M.A., Emrich S.J., Hahn M.W., Besansky N.J. 2015. Extensive introgression in a malaria vector species complex revealed by phylogenomics. Science 347:1258524-1258524. doi: 10.1126/science.1258524.

Gopalakrishnan S., Sinding M.H.S., Ramos-Madrigal J., Niemann J., Castruita J.A.S., Vieira F.G., CarØe C., Montero M.M., Kuderna L., Serres A., González-Basallote V.M., Liu Y.H., Eang G.D., Marques-Bonet T., Mirarab S., Fernandes C., Gaubert P., Koepfli K.P., Budd J., Rueness E.K., Heide-JØrgensen M.P., Peterson B., Sicheritz-Ponten T., Bachmann L., Wiig Ø., Hansen A.J., Gilbert M.T.P. 2018. Interspecific gene flow shaped the evolution of the genus *Canis*. Curr. Biol. 28(21):3441-3449.e5. doi: 10.1016/j.cub.2018.08.041.

Guo B., Fang B., Shikano T., Momigliano P., Wang C., Kravchenko A., Merila J. 2019. A phylogenomic perspective on diversity, hybridization and evolutionary affinities in the stickleback genus *Pungitius*. Mol. Ecol. 28(17):4046-4064. doi: 10.1111/mec.15204.

Guo W., Sun D., Cao Y., Xiao L., Huang X., Ren W., Xu S., Yang G. 2022. Extensive interspecific gene flow shaped complex evolutionary history and underestimated species diversity in rapidly radiated dolphins. J. Mamm. Evol. 29:353-367. doi: 10.1007/s10914-021-09581-6.

Hemmer H., Kahlke R.D., Vekua A.K. 2010. *Panthera onca georgica* ssp. nov. from the Early Pleistocene of Dmanisi (Republic of Georgia) and the phylogeography of jaguars (Mammalia, Carnivora, Felidae). Neues Jahrb. Geol. Paläontol., Abh. 257(1):115-127. doi: 10.1127/0077-7749/2010/0067

Hennelly LM, Habib B, Modi S, Rueness EK, Gaubert P, Sacks BN. 2021. Ancient divergence of Indian and Tibetan wolves revealed by recombination-aware phylogenomics. Mol. Ecol. 30(24):6687-6700. doi: 10.1111/mec.16127

Hey J., Chung Y., Sethuraman A., Lachance J., Tishkoff S., Sousa V.C., Wang Y. 2018. Phylogeny estimation by integration over isolation with migration models. Mol. Biol. Evol. 35(11):2805-2818. doi: 10.1093/molbev/msy162

Hibbins M.S., Hahn M.W., Turelli M. 2021. Phylogenomic approaches to detecting and characterizing introgression. Genetics 220(2). doi:10.1093/genetics/iyab173

Huson D.H., Scornavacca C. 2012. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. Syst. Biol. 61(6):1061-1067. doi:10.1093/sysbio/sys062

Jacobson A.P., Gerngross P., Lemeris Jr J.R., Schoonover R.F., Anco C., Breitenmoser-Würsten C., Durant S.M., Farhadinia M.S., Henschel P., Kamler J.F., Laguardia A., Rostro-Garcia S., Stein A.B., Dollar L. 2016. Leopard (*Panthera pardus*) status, distribution, and the research efforts across its range. PeerJ. 4:e1974. doi: 10.7717/peerj.1974.

Jiao X., Flouri T., Rannala B., Yang Z. 2020. The impact of cross-species gene flow on species tree estimation. Syst. Biol. 69(5):830-847. doi: 10.1093/sysbio/syaa001

Jiao X., Flouri T., Yang Z. 2021. Multispecies coalescent and its applications to infer species phylogenies and cross-species gene flow. Nat. Sci. Rev. 8. doi: 10.1093/nsr/nwab127.

Johnson W.E., Eizirik E., Pecon-Slattery J., Murphy W.J., Antunes A., Teeling E., O'Brien S.J. 2006. The late Miocene radiation of modern Felidae: a genetic assessment. Science 311(5757):73-77. doi: 10.1126/science.1122277.

Kim S., Cho Y.S., Kim H.M., Chung O., Kim H., Jho S., Seomun H., Kim J., Bang W.Y., Kim C., An J., Bae C.H., Bhak Y., Jeon S., Yoon H., Kim Y., Jun J., Lee H., Cho S., Uphyrkina O, Kostyria A., Goodrich J., Miquelle D., Roelke M., Lewis J., Yurchenko A., Bankevich A., Cho J., Lee S., Edwards J.S., Webber J.A., Cook J., Kim S., Lee H., Manica A., Lee I., O'Brien S.J., Bhak J., Yeo J.H. 2016. Comparison of carnivore, omnivore, and herbivore mammalian genomes with a new leopard assembly. Genome Biol. 17:1-12. https://doi.org/10.1186/s13059-016-1071-4.

Korneliussen T.S., Albrechtsen A., Nielsen R. 2014. ANGSD: analysis of next generation sequencing data. BMC Bioinform. 15:1-13. doi: 10.1186/s12859-014-0356-4.

Kumar V., Lammers F., Bidon T., Pfenninger M., Kolter L., Nilsson M.A., Janke A. 2017. The evolutionary history of bears is characterized by gene flow across species. Sci. Rep. 7:1-10. doi: 10.1038/srep46487.

Lartillot N., Philippe H. 2006. Computing Bayes factors using thermodynamic integration. Syst. Biol. 55(2):195-207. doi: 10.1080/10635150500433722.

Li H., Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler Transform. Bioinformatics. 25(14):1754-1760. doi: 10.1093/bioinformatics/btp324.

Li G, Figueiró H.V., Eizirik E., Murphy W.J. 2019. Recombination-aware phylogenomics reveals the structured genomic landscape of hybridizing cat species. Mol. Biol. Evol. 36(10):2111-2126. doi: 10.1093/molbev/msz139.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R., 1000 Genome Project Data

Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25(16):2078-2079. doi: 10.1093/bioinformatics/btp352.
Li G., Hillier L.W., Grahn R.A., Zimin A.V., David V.A., Menotti-Raymond M., Middleton R., Hannah S., Hendrickson S., Makunin A., O'Brien S.J., Minx P., Wilson R.K., Lyons L.A., Warren W.C., Murphy W.J. 2016b. A high-resolution SNP array-based linkage map anchors a new domestic cat draft genome assembly and provides detailed patterns of recombination. G3. 6(6):1607-1616. doi: 10.1534/g3.116.028746.
Long C., Kubatko L. 2018. The effect of gene flow on coalescent-based species-tree inference. Syst. Biol. 67(5):770-785. doi: 10.1093/sysbio/syy020.
Lorenzana G.P., Figueiró H.V., Kaelin C.B., Barsh G.S., Johnson J., Karlsson E., Morato R.G., Sana D.A., Cullen L., May J.A., Moraes E.A., Kantek D.L.Z., Silveira L., Murphy W.J., Ryder O.A., Eizirik E. 2022. Whole-genome sequences shed light on the demographic history and contemporary genetic erosion of free-ranging jaguar (*Panthera onca*) populations. J. Genet. Genom. 49(1):77-80. doi: 10.1016/j.jgg.2021.10.006.
Mallet J., Besansky N., Hahn M.W. 2016. How reticulated are species? BioEssays 38(2):140-149. doi: 10.1002/bies.201500149.
Mallo D., Posada D. 2016. Multilocus inference of species trees and DNA barcoding. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 371(1702):20150335. doi: 10.1098/rstb.2015.0335.
Martin S.H., Davey J.W., Salazar C., Jiggins C.D. 2019. Recombination rate variation shapes barriers to introgression across butterfly genomes. PLoS Biol. 17(2):e2006288. doi: 10.1371/journal.pbio.2006288.
Mazak, J.H. 2010. What is *Panthera palaeosinensis*? Mamm. Rev. 40(1), 90-102. doi: 10.1111/j.1365-2907.2009.00151.x.
Mckenna A., Hanna M., Banks E., Sivachenko A., Cibulskis K., Kernytsky A., Garimella K, Altshuler D., Gabriel S., Daly M., DePristo M.A. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20(9):1297-1303. doi: 10.1101/gr.107524.110.
Mirarab S., Warnow T. 2015. Astral-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics 31(12):i44–52. doi: 10.1093/bioinformatics/btv234.
Nachman M.W., Payseur B.A. 2012. Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 367(1587):409-421. doi: 10.1098/rstb.2011.0249.
O'Brien, S.J., Johnson W.E. 2007. The evolution of cats. Sci. Am. 297(1):68-75. doi: 10.1038/scientificamerican0707-68.
Paijmans, J.L., Barlow A., Förster D.W., Henneberger K., Meyer M., Nickel B., Nagel D., Worsøe Havmøller R., Baryshnikov F., Joger U., Rosendahl W., Hofreiter M. 2018. Historical biogeography of the leopard (*Panthera pardus*) and its extinct Eurasian populations. BMC Evol. Biol. 18(1):1-12. doi: 10.1186/s12862-018-1268-0.
Payseur B.A., Rieseberg L.H. 2016. A genomic perspective on hybridization and speciation. Mol. Ecol. 25(11):2337-2360. doi: 10.1111/mec.13557.
Pease J.B., Hahn M.W. 2015. Detection and polarization of introgression in a five-taxon phylogeny. Syst. Biol. 64(4):651-662. doi: 10.1093/sysbio/syv023.
Pilot M. 2021. Disentangling the admixed trails of grey wolf evolution. Mol. Ecol. 30(24):6509-6512. doi: 10.1111/mec.16261.
Quinlan A.R., Hall I.M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26(6):841-842. doi: 10.1093/bioinformatics/btq033.
Rannala B., Edwards S.V., Leaché A., Yang Z. 2020. The multi-species coalescent model and species tree inference. In: Scornavacca C., Delsuc F., Galtier N., editors. Phylogenetics in the genomic era. No commercial publisher. p. 3.3:1-3.3-20.
Rannala B., Yang Z. 2017. Efficient Bayesian species tree inference under the multispecies coalescent. Syst. Biol. 66(5):823-842. doi: 10.1093/sysbio/syw119.
Ruiz-Ramoni D., Montellano-Ballesteros M., Arroyo-Cabrales J., Caso A., Carvajal-Villarreal S. 2020. The large jaguar that lived in the past of México: a forgotten fossil. Therya. 11(1):33-40. doi: 10.12933/therya-20-821.
Santos S.H.D., Peery R.M., Miller J.M., Dao A., Lyu F.H., Li X., Li M.H., Coltman D.W. 2021. Ancient hybridization patterns between bighorn and thinhorn sheep. Mol. Ecol. 30(23):6273-6288. doi: 10.1111/mec.16136.
Scott P.A., Glenn T.C., Rissler L.J. 2019. Formation of a recent hybrid zone offers insight into the geographic puzzle and maintenance of species boundaries in musk turtles. Mol. Ecol. 28(4):761-771. doi: 10.1111/mec.14983.
Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9):1312-1313. doi: 10.1093/bioinformatics/btu033.
Sunquist M., Sunquist F. 2002. Wild cats of the world. Chicago (IL): University of Chicago Press.
Than C., Ruths D., Nakhleh L. 2008. PhyloNet: a software package for analyzing and reconstructing reticulate evolutionary relationships. BMC Bioinformatics. 9(1). doi:10.1186/1471-2105-9-322
Thawornwattana Y., Dalquen D., Yang Z. 2018. Coalescent analysis of phylogenomic data confidently resolves the species relationships in the *Anopheles gambiae* species complex. Mol. Biol. Evol. 35(10):2512-2527. doi: 10.1093/molbev/msy158.
Trigo T.C., Schneider A., Oliveira T.G., Lehugeur L.M., Silveira L., Freitas T.R., Eizirik E. 2013. Molecular data reveal complex hybridization and a cryptic species of neotropical wild cat. Curr. Biol. 23(14):2528-2533. doi: 10.1016/j.cub.2013.10.046.
Tseng Z.J., Wang X., Slater G.J., Takeuchi G.T., Li Q., Liu J., Xie G. 2014. Himalayan fossils of the oldest known pantherine establish ancient origin of big cats. Proc. Biol. Sci. 281(1774):20132686. doi: 10.1098/rspb.2013.2686.
Upadhyay M., Kunz E., Sandoval-Castellanos E., Hauser A., Krebs S., Graf A., Blum H., Dotsev A., Okhlopkov I., Shakhin A., Bagirov V., Brem G., Fries R., Zinovieva N., Medugorac I. 2021. Whole genome sequencing reveals a complex introgression history and the basis of adaptation to subarctic climate in wild sheep. Mol. Ecol. 30(24):6701-6717. doi: 10.1111/mec.16184.
Uphyrkina O., Johnson W.E., Quigley H., Miquelle D., Marker L., Bush M., O'Brien S.J. 2001. Phylogenetics, genome diversity and origin of modern leopard, *Panthera pardus*. Mol. Ecol. 10(11):2617-2633. doi: 10.1046/j.0962-1083.2001.01350.x.
Vianna J.A., Fernandes F.A., Frugone M.J., Figueiró H.V., Pertierra L.R., Noll D., Bi K., Wang-Claypool C.Y., Lowther A., Parker P., Le Bohec C., Bonadonna F., Wienecke B., Pistorius P., Steinfurth A., Burridge C.P., Dantas G.P.M., Poulin E., Simison W.B., Henderson J., Eizirik E., Nery M.F., Bowie R.C.K. 2020. Genome-wide analyses reveal drivers of penguin diversification. Proc. Natl. Acad. Sci. U S A. 117(36):22303-22310. doi: 10.1073/pnas.2006659117.
Wen D, Nakhleh L. 2018. Coestimating reticulate phylogenies and gene trees from multilocus sequence data. Syst. Biol. 67(3):439-457. doi: 10.1093/sysbio/syx085.
Werdelin L, Yamaguchi N, Johnson W.E., O'Brien S.J. 2010. Phylogeny and evolution of cats (Felidae). In: Macdonald D., Loverage A., editors. Biology and conservation of wild felids. Oxford (England): Oxford University Press. p. 59-82. doi: 10.1017/S0952836904005242.
Yamaguchi N, Cooper A, Werdelin L, Macdonald D.W. 2004. Evolution of the mane and group-living in the lion (*Panthera leo*): a review. J. Zool. 263(4):329-342.
Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24(8):1586-1591. doi: 10.1093/molbev/msm088.
Yu Y., Dong J., Liu K.J., Nakhleh L. 2014. Maximum likelihood inference of reticulate evolutionary histories. Proc. Natl. Acad. Sci 111(46):16448-16453. doi:10.1073/pnas.1407950111
Zhang C, Ogilvie H.A., Drummond A.J., Stadler T. 2018. Bayesian inference of species networks from multilocus sequence data. Mol. Biol. Evol. 35(2):504-517. doi: 10.1093/molbev/msx307.