Multiple types of genomic variation contribute to adaptive traits in the mustelid subfamily Guloninae

Running title: Genomic variation in mustelids

Lorena Derežanin[1*], Asta Blažytė[2], Pavel Dobrynin[3], David A. Duchêne[4], José Horacio Grau[5], Sungwon Jeon[2,#], Sergei Kliver[6], Klaus-Peter Koepfli[3,7,8], Dorina Meneghini[1], Michaela Preick[9], Andrey Tomarovsky[3,6,10], Azamat Totikov[3,6,10], Jörns Fickel[1,9], Daniel W. Förster[1]

1 Leibniz Institute for Zoo and Wildlife Research (IZW), Alfred Kowalke Straße 17, 10315 Berlin, Germany
2 Department of Biomedical Engineering, College of Information and Biotechnology, Ulsan National Institute of Science and Technology (UNIST), Ulsan, 44919, Republic of Korea
3 Computer Technologies Laboratory, ITMO University, 49 Kronverkskiy Pr., 197101 Saint Petersburg, Russia
4 Center for Evolutionary Hologenomics, The GLOBE Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Øster Farimagsgade 5, 1353 Copenhagen, Denmark
5 amedes Genetics, amedes Medizinische Dienstleistungen GmbH, Jägerstr. 61, 10117 Berlin, Germany
6 Institute of Molecular and Cellular Biology, SB RAS, 8/2 Acad. Lavrentiev Ave., Novosibirsk 630090, Russia
7 Smithsonian-Mason School of Conservation, 1500 Remount Road, Front Royal, VA 22630, USA
8 Smithsonian Conservation Biology Institute, Center for Species Survival, National Zoological Park, 1500 Remount Road, Front Royal, VA 22630, USA
9 Institute for Biochemistry and Biology, Faculty of Mathematics and Natural Sciences, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam OT Golm, Germany
10 Novosibirsk State University, 1 Pirogova str., Novosibirsk, 630090, Russia
 # current address: Clinomics Inc., Ulsan, 44919, Republic of Korea

*Corresponding author, *lorenaderezanin@gmail.com*

Abstract

Species of the mustelid subfamily Guloninae inhabit diverse habitats on multiple continents, and occupy a variety of ecological niches. They differ in feeding ecologies, reproductive strategies and morphological adaptations. To identify candidate loci associated with adaptations to their respective environments, we generated a *de novo* assembly of the tayra (*Eira barbara)*, the earliest diverging species in the subfamily, and compared this with the genomes available for the wolverine (*Gulo gulo*) and the sable (*Martes zibellina*). Our comparative genomic analyses included searching for signs of positive selection, examining changes in gene family sizes, as well as searching for species-specific structural variants (SVs). Among candidate loci associated with phenotypic traits, we observed many related to diet, body condition and reproduction. For example, for the tayra, which has an atypical gulonine reproductive strategy of aseasonal breeding, we observe species-specific changes in many pregnancy-related genes. For the wolverine, a circumpolar hypercarnivore that must cope with seasonal food scarcity, we observed many changes in genes associated with diet and body condition. All types of genomic variation examined contributed substantially to the identification of candidate loci. This strongly argues for consideration of variation other than single nucleotide polymorphisms in comparative genomics studies aiming to identify loci of adaptive significance.

Keywords

genomics, adaptation, structural variation, gene family evolution, positive selection, mustelids

Introduction

The Mustelidae are the most ecologically and taxonomically diverse family within the mammalian order Carnivora, representing a remarkable example of adaptive radiation among mammals

that is rich with recent speciation events (Koepfli et al., 2008; Liu et al., 2020). Closely related mustelid species often inhabit vastly different ecosystems, where they experience diverse environmental challenges and are thus exposed to different evolutionary pressures. This is particularly pronounced in the mustelid subfamily Guloninae, within which species occupy a variety of ecological niches, ranging from scansorial omnivores in the neotropics to terrestrial hypercarnivores in circumpolar regions. Members of the Guloninae display a range of behavioural and physiological adaptations associated with environment-specific resource availability, and consequently differ markedly in feeding ecology, reproductive strategy and morphology (Heldstab et al., 2018; Zhou et al., 2011). Here, we focus on tayra, wolverine and sable (Figure 1), for which genomic resources are now available.

The tayra (*Eira barbara)* is a predominantly diurnal, solitary species that inhabits tropical and subtropical forests of Central and South America, ranging from Mexico to northern Argentina (Wilson & Mittermeier, 2009). It is a scansorial, opportunistic omnivore, feeding on fruits, small mammals, birds, reptiles, invertebrates and carrion. Caching of unripe fruit for later consumption has been observed (Soley & Alvarado-Díaz, 2011). Unlike other gulonine species, which are characterized by seasonal breeding and embryonic diapause, the tayra is an aseasonal polyestrous breeder and does not exhibit delayed implantation (Proulx & Aubry, 2017), which may be due to the less prominent seasonality and fluctuation in food availability in neotropical habitats (Heldstab et al., 2018).

The largest terrestrial mustelid, the wolverine (*Gulo gulo)*, is a circumpolar species, inhabiting alpine and boreal zones across North America and Eurasia (Ekblom et al., 2018). The wolverine is an opportunistic predator and facultative scavenger, either feeding on carrion or actively hunting medium to large-sized mammals, such as roe deer, wild sheep and occasionally moose (Pasitschniak-Arts & Larivière, 1995). Morphological and behavioural adaptations such as dense fur, plantigrade locomotion facilitating movement through deep snow, and food caching, enable wolverines to survive in cold habitats with limited food resources (Copeland & Kucera, 1997). In addition, wolverines occupy large home

ranges, display territoriality, seasonal breeding and delayed implantation, traits indicating an adaptive response necessary for survival in scarce resource environments (Inman et al., 2012).

The sable (*Martes zibellina)* is distributed in the taiga and deciduous forests of north central and north eastern Eurasia. The sable is solitary and omnivorous, relying on hearing and olfaction to locate prey, even under a snow cover during winter months (Liu et al., 2020; Monakhov, 2011). Unlike wolverines, seasonal changes do not cause dramatic fluctuations in resource availability for sables as they are able to exploit a wider variety of food sources, and are adapted to tolerate short-term food scarcity (Mustonen et al., 2006). Their diet consists of small mammals, birds, nuts and berries, and in some instances food caching during the winter period has been reported (Monakhov, 2011). Similar to wolverines and many other species of mustelids, sables have a well-defined reproductive season and exhibit delayed blastocyst implantation (Proulx & Aubry, 2017).

To date, only a few studies have investigated adaptive variation in mustelids using comparative genomics (Abduriyim et al., 2019; Beichman et al., 2019; Liu et al., 2020; Miranda et al., 2021). Here, we generated a highly contiguous genome assembly of the tayra, an early diverging gulonine (Koepfli et al., 2008; Law et al., 2018), and compared it to previously published genomes of the wolverine and sable to identify the genetic basis underlying the adaptations to the diverse environments inhabited by these species.

In addition to identifying genes under positive selection, we investigated gene family evolution and structural variants (SVs), as these types of variants represent a significant source of intra- and interspecific genomic differentiation, affecting more nucleotides than single-nucleotide polymorphisms (Catanach et al., 2019). Gene copy number variation and large SVs can be associated with an adaptive response to new ecological circumstances (Rinker et al., 2019), and are thus an important source of genomic novelty to consider when studying adaptive divergence among species (Hecker et al., 2019). We focused on candidate loci linked to species-specific traits associated with response to environmental challenges, such as resource availability in the respective habitats of our study species.

Materials and Methods

Sequencing, Genome Assembly and Alignment

Whole blood from a captive male tayra was collected by the veterinary staff of the "Wildkatzenzentrum Felidae", Barnim (Germany), during a routine medical checkup. High molecular weight (HMW) genomic DNA extraction was performed using the Qiagen MagAttract HMW DNA Kit, following the manufacturer's protocol. We used 1ng of DNA and the Chromium Genome Reagents Kits Version 2 and the 10X Genomics Chromium Controller instrument with a microfluidic chip for library preparation. Sequencing was carried out on an Illumina NovaSeq 6000 (San Diego, CA, USA) with 300 cycles on an S1 lane.

We generated a *de novo* genome assembly using the 10x Genomics Supernova assembler v2.1.1 (Weisenfeld et al., 2017) with default parameters (assembly metrics given in SI Table S1). The assemblies of tayra, wolverine, sable and domestic ferret were assessed for gene completeness with BUSCO v4.1.2 using the mammalian lineage dataset mammalia_odb10 (Simão et al., 2015). To accurately identify repeat families, we used RepeatModeler v2 (Flynn et al., 2020) with the option "-LTRstruct" option, followed by Repeatmasker v4.1.2 (Smit, 2004) to identify and mask the modeled repeats in the tayra genome assembly.

For the purpose of demographic history and structural variation analyses, reads from all three Guloninae species were pre-processed (details given in Supplemental Information), and aligned to the domestic ferret (*Mustela putorius furo*) genome (MusPutFur1.0_HiC; Dudchenko et al., 2017, 2018; Peng et al., 2014) in local mode with Bowtie2 v2.3.5.1 (Langmead & Salzberg, 2012). Duplicated reads were removed with Picard Toolkit v2.23 (MarkDuplicates, Broad Institute 2019). Trimmed tayra reads were downsampled to ~38x with seqtk v1.3 (https://github.com/lh3/seqtk) prior to mapping to maintain uniformity among libraries and to avoid bias in variant calling.

Functional and biological roles of positively selected genes, loci affected by changes in gene family size, and structural variants, were explored using literature sources and online databases, including OrthoDB v10 (Kriventseva et al., 2019), Uniprot (The UniProt Consortium, 2017), and NCBI Entrez Gene (Maglott et al., 2011).

Demographic reconstruction

Analysis of demographic history for tayra and sable was performed with PSMC v0.6.5 (Li & Durbin, 2011) using the following parameters (repeated 100 times for bootstrapping): *psmc -N25 -t15 -r5 -b -p '4+25*2+4+6' -o round-${ARRAY_TASK_ID}.psmc ${name}.split.psmcfa.* Results for each genome were plotted with psmc_plot.pl, and the mutation rate was set to 1e-08 substitutions per site per generation (Cahill et al., 2016; Dobrynin et al., 2015). Generation times were set to 7.4 years for tayra, 5.7 for sable (Pacifici et al., 2013) and 6 years for wolverine (Ekblom et al., 2018).

Reference-based scaffolding

Using the domestic ferret genome as reference, we generated pseudochromosome assemblies for tayra, wolverine and sable, to visualize heterozygosity along chromosomes. Scaffolding was performed using RaGOO v1.1 (Alonge et al., 2019). The X chromosome in the domestic ferret assembly was identified via whole genome alignment to the domestic cat (*Felis catus*) Felis_catus_9.0 assembly (Buckley et al., 2020) and ZooFISH data available from the *Atlas of Mammalian Chromosomes* (Cavagna et al., 2000; O'Brien et al., 2020). Whole genome alignment was performed using LAST v971 (Frith & Kawaguchi, 2015).

Variant calling followed by quality filtration was performed using the BCFtools pipeline v1.10 (Poplin et al., 2018). Low-quality variants were removed (BCFtools filter, 'QUAL < 20.0 || (FORMAT/SP > 60.0 || FORMAT/DP < 5.0 || FORMAT/GQ < 20.0)'). In each sample, positions with coverage lower or higher than 50-250% of the whole genome median value were removed. Of the remaining positions only

those common to all samples were retained. Finally, SNPs with uncalled genotypes in any sample and variants with the same genotypes for all samples were removed. For visualization, heterozygous SNPs were counted in non-overlapping sliding windows of 1 Mbp (counts scaled to SNPs per kbp). Indels were not included due to the low quality of calls from short reads. SNP density plots were created using the MACE package (https://github.com/mahajrod/MACE).

Phylogenomic data preparation, analysis and dating

We performed sequence alignments and filtering of highly divergent alignment segments for 6020 coding genomic regions of single-copy orthologs shared across eight species of carnivores, using the software MACSE v2 (Ranwez et al., 2011). Alignments were further trimmed for highly divergent regions, excessive missing data, and substitution model adequacy (further details in SI). Gene trees were estimated from gene regions by first selecting the best substitution model from the GTR+F+Γ+I+R family (Kalyaanamoorthy et al., 2017), and calculating approximate likelihood-ratio test (aLRT) branch supports (Anisimova & Gascuel, 2006), as implemented in IQ-TREE v2 (Minh et al., 2020b).

Species tree estimates were performed using (i) concatenated sequence alignments for maximum likelihood inference using IQ-TREE v2, and (ii) gene trees for inference under the multi-species coalescent using the summary coalescent method in ASTRAL-III (Zhang et al., 2018). The maximum likelihood estimate of the species tree was accompanied by aLRT branch supports, while summary coalescent inference was accompanied by local posterior probabilities (Sayyari & Mirarab, 2016). The decisiveness of the data regarding the phylogenetic signals was examined using gene- and site-concordance factors, calculated in IQ-TREE v2 (Minh et al., 2020a).

Bayesian molecular dating analysis was performed using MCMCtree in PAML v4.8 (Yang, 2007), assuming the reconstructed species tree from ASTRAL-III and using the genomic regions with gene trees concordant with the species tree (see SI for further details). Regions were partitioned by codon positions, each modelled under individual GTR+Γ substitution models. We used an uncorrelated gamma

prior on rates across lineages and a birth-death prior for divergence times. Fossil calibrations are listed in the SI. The posterior distribution was sampled every $1x10^3$ MCMC steps over $1x10^7$ steps, after a burn-in phase of $1x10^6$ steps. We verified convergence to the stationary distribution by comparing the results from two independent runs, and confirming that the effective sample sizes for all parameters were above 1000 using the R package coda (Plummer et al., 2006).

Positive selection on single-copy orthologs

To investigate genes under positive selection, the CDS corresponding to 1:1 orthologs were aligned for the eight aforementioned carnivoran species. Multiple sequence alignments (MSAs) were constructed with PRANK v120716 (Löytynoja, 2014), and 17 MSAs were removed due to short alignment length. The CODEML module in the PAML v4.5 package was used to estimate the ratio of non-synonymous to synonymous substitutions, also called d_N/d_S or ω (Yang, 2007). We applied the one-ratio model to estimate the general selective pressure acting among all species, allowing only a single d_N/d_S ratio for all branches. A free-ratio model was also used to estimate the d_N/d_S ratio of each branch. Furthermore, the CODEML branch-site test for positive selection was performed on 6003 ortholog alignments for three separate foreground branches: *Eira barbara*, *Gulo gulo* and *Martes zibellina* (Zhang et al., 2005). Statistical significance was assessed using likelihood ratio tests (LRT) with a conservative 10% false discovery rate (FDR) criterion (Nielsen et al., 2005). Orthologs with a free-ratio > 2 in the branch model were considered for further analysis of signatures of positive selection.

To account for differences in genome assembly quality, we evaluated the alignments of selected orthologs based on the transitive consistency score (TCS), an extension to the T-Coffee scoring scheme used to determine the most accurate positions in MSAs (Chang et al., 2014). Additionally, alignments were visually inspected for potential low-scoring MSA portions.

Gene family evolution

To investigate changes in gene family sizes, we constructed a matrix containing 7838 orthologs present as either complete 'single-copy', complete 'duplicated', or 'missing', identified using the BUSCO genome assembly completeness assessment of all eight carnivoran genomes. Orthologs were retained if they were detected in at least four species (including *Felis catus* as an outgroup) to obtain meaningful likelihood scores for the global birth and death (λ) parameter.

We applied a probabilistic global birth and death rate model of CAFE v4.2.1. (Han et al., 2013) to analyse gene gains ('birth') and losses ('death') accounting for phylogenetic history. First, we estimated the error distribution in our dataset, as genome assembly and annotation errors can result in biased estimates of the average rate of change (λ), potentially leading to an overestimation of λ. Following the error distribution modelling, we ran the CAFE analysis guided by the ultrametric tree estimated earlier, calculating a single λ parameter for the whole species tree. The CAFE results were summarized (SI Table 4A) with the python script *cafetutorial_report_analysis.py* (https://github.com/hahnlab/CAFE).

Structural variation

We applied an ensemble approach for structural variant (SV) calling, encompassing three SV callers: Manta v1.6.0 (Chen et al., 2016), Whamg v1.7.0 (Kronenberg et al., 2015) and Lumpy v0.2.13 (Layer et al., 2014). SV calls originating from reads mapping in low complexity regions and withpoor mapping quality were removed from all three call sets. We retained Manta calls with paired-read (PR) and split-read (SR) support of PR>=3 and SR>=3, respectively. To reduce the number of false positive calls, the Whamg call set was filtered for potential translocation events, as Whamg flags but does not specifically call translocations. We further removed calls with a low number of reads supporting the variant (PR, SR) from the Whamg ($A < 10$) and the Lumpy call set ($SU < 10$). All SV call sets were filtered based on genotype quality ($GQ \geq 30$). Whamg and Lumpy SV call sets were genotyped with

Svtyper v0.7.1 (Chiang et al., 2015) prior to filtering. Only scaffolds assigned to chromosomes were included in further analyses. Survivor v1.0.7 (Jeffares et al., 2017) was used to merge and compare SV call sets within and among samples. The union set of SV calls among the three gulonine species containing species-specific and shared variants was annotated, using Liftoff v1.5.1 (Shumate & Salzberg, 2020) for preparation of reference genome annotation, and Ensembl Variant Effect Predictor v101.0 (McLaren et al., 2016) for identifying variants affecting protein-coding genes. Gene ontology analysis was performed with Shiny GO (Ge et al., 2020) with an FDR < 0.05 for each SV type (excluding inversions) overlapping multiple protein-coding genes (> 5 genes).

Results

Genome Assembly

We generated a highly contiguous reference genome assembly for the tayra (*Eira barbara*). Extracted genomic DNA had an average molecule size of 50.75 kb and was sequenced to ~76-fold coverage (SI Table S1). The final assembly showed a total length of ~2.44 Gb (excluding scaffolds shorter than 5 kb), with a contig N50 of 290 kb, scaffold N50 of 42.1 Mb, and identity in 95% of all positions in an alignment with domestic ferret genome (SI Figure S1). The tayra assembly has higher contiguity than the Illumina-only based assemblies of both wolverine and sable, but it is more fragmented than the chromosome-length domestic ferret (*Mustela putorius furo)* assembly (Table 1, SI Figure S2A) that we used as reference genome for some analyses. The haploid tayra genome of ~2.4 Gb is contained in 162 scaffolds (>100 kb) with 40 scaffolds having a length above 50 Mb (SI Figure S2A).

The tayra assembly has high gene completeness as assessed with BUSCO v4.1.2 using 9226 conserved mammalian orthologs. In total, 8540 (92.5%) complete Benchmarking Universal Single-Copy Orthologs (BUSCOs) were identified, encompassing 8492 (92.0%) of complete and single-copy, and 48 (0.5%) complete and duplicated orthologs. Additionally, 104 (1.1%) orthologs were fragmented and 582

(6.4%) were missing. As measured by this metric, the tayra genome has higher gene completeness than the published genomes of wolverine, sable or domestic ferret (SI Figure S2B).

Repetitive elements

The repeat landscape of the tayra assembly contains ~0.85 Gb of repetitive elements (SI Table S2). L1 type LINE elements are the most abundant, constituting 23% of the tayra genome. L1 elements also show signs of recent proliferation in comparison to DNA transposons and LTR retroelements (SI Figure S3). Endogenous retroviruses constitute 3.8% of the tayra genome and can be classified as Gammaretroviruses and Betaretroviruses.

The overall repeat landscape of the tayra genome assembly is comparable to other carnivore genomes (Liu et al., 2020; Peng et al., 2018). It is similar to that of the sable genome, differing mostly in the number of L1 LINE elements, which have been recently proliferating and accumulating within the tayra genome more than in other Guloninae genomes. The diversity of endogenous retroviruses is similar to other mustelids. Although endogenous delta-retroviruses have been described from a broad range of mammal genomes, including several smaller carnivores like mongoose (family Herpestidae) and the fossa (*Cryptoprocta ferox*) (Hron et al., 2019), no delta-retroviruses were found in the genome of tayra.

Demographic reconstruction

Reconstruction of historical demography for tayra and sable using the Pairwise Sequentially Markovian Coalescent model (PSMC; Li & Durbin, 2011) showed similar effective population sizes (Ne) during the last one million years of the Pleistocene, but then start to differ around twenty thousand years ago, when the Ne of the sable increases relative to that of the tayra (SI Figure S4). Historical demography for the wolverine was reconstructed from Ekblom et al. 2018, and shows the lowest Ne among the three gulonine species.

Nucleotide diversity

The tayra, sable and wolverine assemblies were generated using different approaches and differ significantly in contiguity (Table 1). To address this issue, we generated pseudochromosome assemblies for each species using the chromosome-length assembly of the domestic ferret as reference. The domestic ferret has more chromosomes than the other species (2n=40 vs. 2n=38), and the same number (20) of pseudochromosomes (Lewin et al., 2019) were obtained after scaffolding in each case. For each assembly, we identified the X chromosome (labeled as *ps_chrX*) and arranged pseudoautosomes (labeled as *ps_aut1* - *ps_aut19*) according to the length of the original scaffolds in the domestic ferret reference. This allowed us to verify the sex of the animals using a coverage-based approach (SI Figure S5), which confirmed morphological sexing for the tayra (male) and wolverine (female) individuals. While the sable individual is referred to as a male (Liu et al., 2020), our analysis suggests it is a female (further details given in SI).

We counted heterozygous SNPs in 1 Mbp stacking windows for all three species and scaled it to SNPs per kbp (Figure 2). Median values for tayra, sable and wolverine were 1.89, 1.44 and 0.28 SNPs per kbp, respectively, the latter being in agreement with previous findings (Ekblom et al., 2018). All scaffolds of $\geq$ 1 Mbp in pseudochromosome assemblies were taken into account. Exclusion of ps_chrX resulted in slight increases of medians to 1.93, 1.47 and 0.29, respectively (SI Table S3, SI Figure S6). Regardless of whether the ps_chrX was included or excluded, the genome-wide diversities among the three species were significantly different (p-values $<<$ 0.001, Mann-Whitney test).

Phylogenomics and molecular dating

We reconstructed the phylogenetic relationships of the following carnivoran species: domestic cat, domestic dog, walrus, northern elephant seal, domestic ferret, wolverine, sable, and tayra. Phylogenomic analyses using concatenation and summary coalescent methods led to an identical resolution of the relationships among mustelid taxa. Within Guloninae, the wolverine and sable were

placed as sisters, to the exclusion of the tayra (Figure 3). Branch supports were maximal across all branches using both approximate likelihood-ratio test (aLRT) and local posterior probabilities. Similarly, concordance factors (CF) for genes and sites (gCF, sCF) were high across branches and consistently more than twice as high as the values of the discordance factors (gDF, sDF). The lowest concordance factors were those in support of the resolution of *Gulo* and *Martes* as sisters (gCF = 64.52, sCF = 54.38). However, the discordance factors were less than half these values (gDF < 14, sCF < 24), suggesting substantial decisiveness across genes and sites for this resolution.

Molecular dating was performed using 992 single-copy orthologous gene regions (comprising 0.53 million sites). Divergence time estimates across mustelids were largely in agreement with previous findings (Koepfli et al., 2008; Law et al., 2018; Li et al., 2014; Sato et al., 2012), placing the split between *Mustela* and Guloninae at 11.2 million years ago (Mya) (highest posterior density interval (HPDI) between 13.1 and 9.5 Mya), and the split between *Eira* and the *Gulo-Martes* group at 7.5 Mya (HPDI between 9 and 6.1 Mya). The split between *Gulo* and *Martes* was dated at 5.9 Mya (HPDI between 7.4 and 4.7 Mya).

Positive selection on single-copy orthologs

In the three Guloninae species, we found sites under positive selection (5 > dN/dS ratio > 1, (Barnett et al., 2020) in 55 single-copy orthologs that were highly significant (free-ratio > 2). Of these 55 positively selected genes (PSGs), 15 were observed in tayra, 22 in wolverine, and 18 in sable (Figure 4A-B). Gene names, descriptions and functions are given in SI Table S4.

Among the 15 PSGs we detected in tayra, five are associated with reproduction (*NSMCE1*, *ETV2*, *SPATA25*, *MUC15*, and *PIH1D2*) with functions involving spermatogenesis, placenta and embryo development, and blood vessel morphogenesis. Among the remaining ten PSGs, *HSPB6* is involved in vasodilation and muscle contraction, *DERA* is associated with environmental stressors, including exposure to toxins, and uricase (*UOX*) is a liver enzyme involved in purine catabolism and regulation of

fructose metabolism. Three PSGs (*IP6K3*, *MAGIX* and *FAM149B1*) are found to be associated with the nervous system, synapse formation and structural plasticity, as well as motor skills and coordination. Three further PSGs (*DUSP19*, *TNLG2B*, and *LRRC46*) are related to the immune system and *HEMK1* regulates methylation processes.

We detected 22 PSGs in wolverine, including six genes associated with energy production and conversion. Among them, *ATP6V0B*, *KMO*, and *SLC16A4* are primarily involved in insulin level regulation, and the metabolism of carbohydrates and fatty acids. Three PSGs (*OIP5*, *ZADH2* and *MTPAP*) are specifically associated with adipose tissue formation and intramuscular fat deposition. Additionally, we found three PSGs (*NBR1*, *TMEM38B*, *PPP1R18*) involved in selective autophagy as response to nutrient deprivation along with bone mass and density regulation, and resorption. We also detected PSGs (*DAB1*, *OPA1* and *CTNS*) linked to cognition, brain development and vision. Several PSGs (*BNIPL*, *IL18BP*, *CRNN*) were associated with the immune system; while three others (*ANAPC7*, *RNF212B*, *IZUMO3*) are involved in reproduction processes and *USB1* and *CLCN4* have a role in basal cell cycle processes. For the remaining two, *CEP95* and *FAM185A*, it was not possible to associate a specific phenotypic trait.

Among the 18 PSGs detected in the sable, three (*PRRT2*, *ATL2*, *SELENOI*) are associated with locomotion, and coordination, and *USP53* is associated with sensory perception and nervous system. Two PSGs, *VEGFC*, *RASA1*, are associated with blood vessel formation, three (*TTC4*, *ZBP1*, *CD247*) with the immune system and three (*IQUB*, *UBQLNL*, *MEIKIN*) with reproduction. Several PSGs (*EEF2KMT*, *DEUP1*, *ECD*, *IQCK*) are associated with cell cycle processes, and *ZC2HC1C* and *CCDC17* could not be associated with a particular biological process.

Gene family expansions and contractions

Adaptive divergence between species may also be caused by changes in gene family sizes that occur during genome evolution and are due to gains (expansions) or losses (contractions) of genes or

groups of genes (Olson, 1999; Tigano et al., 2020). All species analysed displayed more gene family contractions than expansions, with the wolverine having the highest contraction rate. This is most likely an artefact resulting from the fragmented genome assembly of this species (SI Figure S7). All identified expansions were in the form of gene duplications, with one putative triplication detected in tayra (SI Table S5B).

Tayra and sable had similar numbers of gene family expansions and contractions (SI Figure S7): 34 expansions and 169 contractions in tayra, and 33 expansions and 162 contractions in sable. The less contiguous wolverine genome contained 7 expansions and 649 contractions (SI Tables S5A, S5B). Due to the stochastic nature of gene losses and the potential inflation of estimates resulting from different genome assembly contiguities, we here focus on gains of gene copies.

Expanded gene families in the tayra genome are associated with reproduction, metabolism, nervous and immune system, and cell cycle, among others (Figure 4C-D; gene names, descriptions and functions are given in SI Table S5B). Of the three reproduction-related genes, *SLC38A2* regulates supply of nutrients for fetal growth through the placenta during the peri-implantation period. The second, *HSD17B10*, is associated with regulation of pregnancy-sustaining steroid hormones, and *RBP2* is involved in retinol binding and vitamin A metabolism, necessary for oogenesis and embryogenesis, as well as vision. Four genes (*ATP6V1D*, *DBX2*, *SLC38A1*, *MAPKAPK5*) are associated with cerebral cortex development, synapse formation, visual perception, and learning processes. *ANKRD13A* is also associated with vision, more specifically with lens fiber generation and vitamin A metabolism. *N6AMT1* is involved in modulation of arsenic-induced toxicity. The olfactory receptor gene *TAAR5* is involved in behavioral responses in mammals, and was duplicated in both tayra and sable. We detected one putative triplication of *FKBP3*, a gene associated with immunoregulation, predominantly of T-cell proliferation. *PRR11* could not be associated with a particular biological process.

In the wolverine, two duplicated genes are related to the nervous system: *GFRA4* is implicated in motor neuron development and *KCNS1* in regulating mechanical and thermal pain sensitivity. *MTM1* is

associated with positive regulation of skeletal muscle tissue growth and *MON1B* is implicated in immune response to viral infection.

In the sable, expansions involve gene families associated with the nervous system, specifically with sensory perception and locomotion, with angiogenesis, hair follicle development, and the immune system. *DUSP8* is involved in neuronal development, vision, olfaction, and spatial memory formation. *FBXL3* is associated with regulation of the circadian clock and *TIMM10* with hearing. Two genes, *CDC42* and *TCHHL1*, are implicated in hair-follicle development. Additionally, *CDC42*, a gene coding for a cell division control protein, is also involved in angiogenesis and hematopoiesis, alongside *TNFRSF12A* and *LXN*.

Structural variation

Structural variants (SVs) modify the structure of chromosomes and can affect gene synteny, repertoire, copy number and/or composition (e.g. gain or loss of exons), create linkage-blocks, and modify gene expression (Chiang et al., 2017; Mérot et al., 2020), leading to complex variation in phenotypes and genetic diseases (Weischenfeldt et al., 2013). We investigated four types of SVs (deletions, duplications, insertions, inversions) in the three Guloninae.

We identified the highest number of species-specific SVs in sable (22979), followed by tayra (8907), and wolverine (264) (Figure 5A). The most abundant SVs detected in all three species are deletions (> 50bp), ranging from 183 species-specific deletions in wolverine to 21713 in sable. Duplications were the least frequent SV type among the three species (SI Figure S8A). For all three species, the majority of SVs are located in intergenic regions (> 80%), with a smaller portion found in genic regions, completely or partially overlapping protein-coding genes (UTRs, exons, introns). According to Variant Effect Predictor (VEP) classification, SVs impacting genic regions are classified either as high-impact variants or modifiers (McLaren et al., 2016) with putative consequences on gene transcription ranging from transcript truncation to transcript ablation or amplification. The highest

number of species-specific genic SVs was detected in tayra, with 330 (3.70% of species-specific SVs), followed by 156 (0.68%) in sable and 53 (20.08%) in wolverine (SI Figure S8B). Other than the well documented impact of inversions on intra and interspecific gene flow (Porubsky et al., 2020; Wellenreuther & Bernatchez, 2018), determining the impact of inversions overlapping large sets of genes is still challenging, as the largest effect is likely to be restricted to genes near SV breakpoints. Therefore, we restricted our examination of gene function to loci affected by deletions, duplications, and insertions (Figure 5B; gene names, descriptions and functions are given in SI Table S6).

In the tayra genome, we observed 14 duplications spanning a combined length of 2.92 Mb, putatively affecting 24 protein-coding genes. Duplicated genes and gene blocks are associated with reproduction, olfaction, metabolism and energy conversion. This included *RNASEH2B*, a gene involved in *in utero* embryo development, and two genes involved in spermatogenesis, *DIAPH3* and *PCNX1*, with the latter an example of a complex SV involving heterozygous duplication and deletion of an exon (SV ~2kb in length). We detected 212 deletions in the tayra genome in relation to the domestic ferret reference, comprising a total length of 2.08 Mb, and affecting 247 genes, which are associated with reproduction, metabolism/energy conversion, nervous system, and cell cycle processes, among other functional categories (SI Table S6). Genes involved in placenta development and *in utero* embryogenesis include *HSF1*, *RSPO2*, and *DNMT3A*. Additionally, we detected *NLRP1* and *NLRP8*, both associated with preimplantation development, and highly expressed in oocytes. One short insertion was observed in *LIX1L*, a gene associated with anatomical structure morphogenesis.

In the wolverine genome, no duplications overlapped genic regions. However, 47 deletions spanning a combined length of 229 kb are putatively associated with transcript truncation or ablation in 48 genes. The majority of affected genes are associated with metabolism/energy conversion, development and basic cell cycle processes. This includes *GLUD1*, a gene involved in amino-acid induced insulin secretion, also found to be affected by a shorter deletion in sable, and *NSDHL*, a gene regulating cholesterol biosynthesis. Additionally, we detected deletions affecting *PARVA*, a gene associated with

angiogenesis and smooth muscle cell chemotaxis and *DNAJC7*, involved in positive regulation of ATPase activity and regulation of cellular response to heat. We also detected one insertion in a gene of unknown function.

In the sable genome, we detected eleven duplications spanning a combined length of 324 kb, overlapping 16 genes associated with sensory perception, development, cell cycle, and the immune system. The 130 detected deletions (combined length of 408 kb) overlap 125 protein-coding genes associated with reproduction, immune system, development, metabolism, sensory perception, and cell cycle. Deletions were identified in two genes involved in keratinocyte differentiation, *PPHLN1*, also affected in wolverine, and *IVL*, associated with hair follicle development. Additionally, two short insertions were found in *NCOA4*, and *YIPF5*, genes associated with mitochondrial iron homeostasis and protein transport, respectively.

Discussion

Here, we present a highly contiguous genome for the tayra (*Eira barbara*). Contiguity of the assembly and its gene completeness are similar to or higher than those of other carnivoran species using the same sequencing approach (Armstrong et al., 2019, Etherington et al., 2020, Kim et al., 2020), confirming the utility of linked-reads for assembly of mammal genomes.

Phylogenomic relationships among the mustelids resulted in a tree topology and divergence time estimates in agreement with previous studies using fewer loci (Koepfli et al., 2008; 2018; Li et al., 2014; Sato et al., 2012). We estimated the split between *Mustela* and Guloninae occurred 11.2 million years ago (Mya) (HPDI 13.1 - 9.5 Mya), followed by the split between *Eira* and the *Gulo-Martes* group 7.5 Mya (HPDI 9 - 6.1 Mya), and the split between *Gulo* and *Martes* at 5.9 Mya (HPDI 7.4 - 4.7 Mya).

Contrary to the findings by (Weissensteiner et al., 2020) in corvids, we did not observe a positive relationship between variation at the nucleotide level (heterozygous SNPs) and variation at the structural level (heterozygous SVs) within the gulonine species. The tayra displayed the highest nucleotide diversity

(1.89 SNPs per kbp), but only the second highest amount of heterozygous SVs (2543, 23.6% of the total SVs, SI Figure S9). The sable had the second highest nucleotide diversity (1.44 SNPs per kbp), but the highest number of heterozygous SVs (14823, 59.5% of total). The wolverine displayed the lowest variation for both (0.28 SNPs per kbp, and 153 or 43.1% heterozygous SVs in total). It is known that SV calling using short-read data can miss a large number of SVs (Ebert et al. 2021). The fact that we did not detect a positive correlation between variation at the nucleotide and structural level, as would be expected if diversity of SNPs and SVs are correlated with population size, may result from our SV analysis relying on short-read data only. Weissensteiner et al. (2020), who did report a positive correlation between SNP and SV diversity, performed long-read based SV typing.

Assessment of variation among genome assemblies of closely related species also strongly relies on the contiguity and completeness of analysed assemblies (Gurevich et al., 2013; Totikov et al., 2021). This needs to be accounted for when examining variation among discontiguous genome assemblies. Here, the low contiguity of the wolverine assembly has likely impacted the number of PSGs and gene family expansions/contractions detected. Multiple, short insert size libraries sequenced to low coverage have likely resulted in decreased SV detectability. However, low diversity in the Scandinavian wolverine population is expected given previous results (Ekblom et al. 2018) and the species' historical demography.

Adaptive genomic variation

Among positively selected genes, gene family expansions and coding regions impacted by SVs, we found numerous candidate loci related to species-specific traits in Guloninae.

The three gulonine species differed with respect to which reproduction-related processes were associated with candidate genes. In the aseasonally breeding tayra, genes detected in three analyses (PSGs, gene family expansions, SVs) were frequently pregnancy-related, while in the two seasonal breeders, sable and wolverine, genes were involved in a greater diversity of reproduction-related

processes, primarily in gametogenesis (Figure 6A). This may reflect the transition from seasonal to aseasonal breeding in the tayra lineage, the only extant aseasonal breeder among the Guloninae.

Shifts in dietary preferences have been linked to positive selection in single genes (Kosiol et al., 2008) and to copy number variation in metabolism-related gene families in mammals (Hecker et al., 2019; Rinker et al., 2019). Among Guloninae, the dietary gradient from omnivory in the neotropics to (hyper)carnivory in the northern Palearctic is reflected in the proportion of metabolism-related candidate genes observed (Figure 6B). The highest proportion of genes involved in carbohydrate metabolism ("omnivorous diet") were detected in tayra, followed by sable, and none detected in wolverine. However, the highest number of genes involved in maintaining a stable body condition through shifts in energy sources (autophagy, fat storage and bone resorption) was observed for the wolverine, potentially reflecting this species' adaptive response to unfavourable environmental conditions in its circumpolar habitat.

We note that in two analyses (PSGs and gene family evolution), we only considered variation in single-copy orthologs, not the entire gene repertoire of these species. Thus, our results are likely only an incomplete reflection of the genes involved in these traits.

Seasonal breeders in the north palaearctic: wolverine & sable

Obligate embryonic diapause or the delayed implantation of the blastocyst is a widespread reproductive strategy among mustelids and other carnivorans. The majority of mustelids in the Northern Hemisphere are seasonal breeders, which includes wolverines and sables that undergo a prolonged period of a delayed implantation lasting several months (Mead, 1981; Svishcheva & Kashtanov, 2011). There is a strong selection for short birthing periods at higher latitudes (Heldstab et al., 2018), because it is one of the most energetically demanding periods, and its timing in relation to prey abundance has a critical impact on reproductive fitness (Inman et al., 2012).

Candidate genes potentially involved in seasonal breeding in wolverine and sable were associated with pregnancy, gametogenesis and circadian rhythm (Figure 6A). These mustelids occupy large home ranges, display territoriality and low population density, leading to infrequent encounters between males and females (Inman et al., 2012; Kashtanov et al., 2015). Thus, a mechanism that facilitates induced ovulation during the encounter would be an adaptive advantage (Larivière & Ferguson, 2003). We detected signals of positive selection in *ANAPC7* in wolverine, a gene involved in progesterone-mediated oocyte maturation and release from cell arrest prior to fertilization (Papin et al., 2004; Reis et al., 2006), that may have a role in increasing progesterone secretion and renewed embryonic development, as observed in skunks and mink (Mead, 1989).

Delayed implantation serves as an important adaptation to ensure maximal survival chances of offspring depending on the environmental conditions (Heldstab et al., 2018). Previous studies in mink (Lopes et al., 2003) showed that increased levels of vascular endothelial growth factors (VEGF) and their receptors correlate with the implantation process. *VEGFC*, primarily associated with angiogenesis and regulation of permeability of blood vessels during embryogenesis, was positively selected in sable, suggesting its possible involvement in embryo implantation regulation in this species.

Changes in testicular activity and spermatogenesis strongly correlate with season in mink (Blottner et al., 2006), and lynx (Jewgenow et al., 2006). We observed spermatogenesis-related genes affected by different types of variation in the three gulonine species. In the wolverine, positively selected genes involved in spermatogenesis included *IZUMO3* and *RNF212B*. Testes-expressed *IZUMO3* is essential for gamete fusion during fertilization (Ellerman et al., 2009) and positive selection of *IZUMO3* has been linked to species-specific adaptations rather than sexual selection (Grayson & Civetta, 2012) *RNF212B* is critical for successful crossing over in gametes, and knockout of *RNF212* causes sterility in mice (Reynolds et al., 2013). In sable, candidate genes involved in spermatogenesis were *UBQLNL* (Marín, 2014) and *SEPT12*, with the latter also being duplicated. Furthermore, *MEIKIN* and *CDK2*, both involved in meiosis, show signals of positive selection and rapid evolution through gene family

expansion, respectively. *MEIKIN* has been found to be consistently rapidly evolving in other eutherian mammals (Pontremoli et al., 2021).

Seasonal breeding in many mammals is largely under photoperiod regulation, suggesting that the circadian system and pineal melatonin secretion play an important role in this reproductive strategy. Maintaining elevated melatonin in mink and spotted skunks has been shown to prolong embryonic diapause (Murphy & Fenelon, 2020). *FBXL3*, associated with maintenance of circadian clock oscillation in mammals (Fenelon et al., 2017; Shi et al., 2013) is duplicated in sable, indicating its possible relation to diapause regulation.

Aseasonal breeder in the neotropic: tayra

In tropical regions, reproductive seasonality is less pronounced as environmental conditions are relatively stable throughout the year (McNutt et al., 2019). Tayras are aseasonal breeders with multiple estrous cycles per year (Proulx & Aubry, 2017) and do not exhibit embryonic diapause (Poglayen-Neuwall et al., 1989). The tayra represents the most basal taxon of the Guloninae, and is an exception regarding its reproductive strategy.

Genes that may be involved in the aseasonal breeding of tayra include *ETV2*, *MUC15*, *SLC38A2*, *HSD17B10*, *RBP2* and *RNASEH2B*. *ETV2* and *MUC15* are involved in placental and embryo development, and regulation of implantation (Shyu et al., 2007; Singh et al., 2019). The *MUC15* protein has been associated with an anti-adhesive function, and prevention of premature blastocyst attachment in rats (Denker, 2000; Poon et al., 2014). Increased expression of the neutral amino acid transporter *SLC38A2* has been detected in the endometrium of cattle during the preimplantation period (Forde et al., 2014), and in mice placentas during late gestation, maintaining fetal growth when maternal growth is restricted by undernutrition (Coan et al., 2010). *SLC38A2* is a duplicated and rapidly evolving gene in tayra, suggesting a potential role in supporting multiple pregnancies during the year. Duplication of *HSD17B10* might indicate a gene dosage modulation in steroid hormone-regulated pregnancy in tayra.

This gene is involved in mitochondrial fatty acid metabolism, cell growth and stress response (Liu et al., 2020). It is highly expressed in fetal and maternal liver, influencing the maintenance of human pregnancy and providing protection against excitotoxicity (Hill et al., 2011). Besides being associated with vision, retinol binding protein *RBP2* has a significant role in regulating retinoids during oogenesis and embryogenesis, and positively impacts oocyte maturation in mice, cattle, pigs and sheep, (Brown et al., 2003; Harney et al., 1993). It was found to be duplicated in tayra, as well as a partial duplication detected in *RNASEH2B*. Mutations in this gene have been shown to cause interferonopathies, leading to impaired fetal brain development (Crow et al., 2015; Yockey & Iwasaki, 2018). In tayra, duplication of this gene potentially increases gene expression and may assist with maintenance of embryo development. Furthermore, five genes involved in placental development, implantation and embryogenesis (*HSF1*, *RSPO2*, *NLRP1*, *NLRP8* and *DNMT3A*) have been affected by partial deletions in tayra, suggesting further changes in dosage of pregnancy-related genes.

Resource availability in the northern Palaearctic: wolverine & sable

Overcoming demanding periods and staying active throughout the year is challenging for non-hibernating northern palearctic species and requires specific mechanisms to cope with seasonal food scarcity. Wolverines have evolved several adaptations to efficiently exploit scarce food resources and overcome constraints of the low-productivity environment (Inman et al., 2012)**.** A critical metric of body condition is the amount of stored fat, representing an important energy source between feeding opportunities and providing thermal insulation in cold environments. Undernutrition detrimentally affects reproduction, and may lead to decreased survival rates (Brown & Lasiewski, 1972; Schulte-Hostedde et al., 2005). Three PSGs detected in the wolverine are involved in adipose tissue formation, *MTPAP*, *OIP5 and ZADH2* (Han et al., 2012; Inoue et al., 2014; Yu et al., 2013), and may be associated with selective fatty acid mobilization stimulated by fasting periods (Inman et al., 2012; Krebs et al., 2004), as observed in mink (Nieminen et al., 2006), and raccoon dog (Mustonen et al., 2007).

One of the responses to prolonged periods of nutrient deprivation and extreme environmental conditions is suppressed bone resorption and formation (Lennox & Goodship, 2008; McGee-Lawrence et al., 2015). Beside regulating survival of blastocysts during delayed implantation (Lee et al., 2011; Lim & Song, 2014), it has been demonstrated that autophagy plays a critical role in maintenance of bone homeostasis (DeSelm et al., 2011; Montaseri et al., 2020). It is thus of note that genes involved in bone mass regulation, resorption (*PPP1R18*, *TMEM38B*) and autophagy (*NBR1*) are under positive selection in wolverines. We also detected a duplication of the muscle growth regulating gene *MTM1*. While a lack of *MTM1* will lead to muscle hypotrophy through unbalanced autophagy in humans and mice (Al-Qusairi et al., 2013), a gene duplication may facilitate muscle growth.

In sable, fatty acids are uniformly mobilized from anatomically distinct fat deposits (Nieminen & Mustonen, 2007), and we observed duplications of *ASB6*, *SLC25A10*, *RNF186* and *BORCS6*, which have been implicated in effective regulation of fat storage and response to nutrient availability (Mizuarai et al., 2005; Okamoto et al., 2020; Schweitzer et al., 2015; Wilcox et al., 2004). Moreover, partial deletions of *APOD*, *PDHB*, *LDLR*, and *CERS5*, associated with lipid metabolic processes (Carmo et al., 2009; Gosejacob et al., 2016; Serão et al., 2011; Tavori et al., 2015), indicate a possible association with adaptive fasting and energy conservation in this species.

We observed partial deletions in *DNAJC7*, a gene involved in thermoregulation (Sonna et al., 2002) in both sable and wolverine, but not in tayra, indicating a potential role in gulonines inhabiting colder environments. Furthermore, we detected partial deletions in *GLUD1* in sable and wolverine. *GLUD1* regulates insulin secretion (Fahien & Macdonald, 2011) and might play a role in mobilizing lipids, causing states of reversible insulin resistance as an adaptive response during prolonged periods of nutrient deprivation, as has been observed in some carnivorans (Martinez & Ortiz, 2017; Viscarra et al., 2013).

In sable and tayra, we observed a partial deletion of *PPP1CC*, associated with glycogen metabolism (Li et al., 2021), tentatively facilitating a more 'omnivorous diet' in these species compared to hypercarnivory in wolverines (Lopes-Marques et al., 2020; Sharma et al., 2018).

There are notable differences in hair follicle structure and fur density between palearctic and neotropic species (Kitchener et al., 2018), and we observed two gene copy gains in sable, *TCHHL1*, involved in hair morphogenesis (Wu et al., 2011), and *CDC42*, required for differentiation of hair follicle progenitor cells (Wu et al., 2006).

Resource availability in the Neotropics: tayra

Tayras exploit diverse food sources and experience relatively stable resource availability all year round, as (sub)tropical habitats have higher primary productivity compared to habitats at higher latitudes (Heldstab et al., 2018; Zhou et al., 2011). We found candidate genes associated with regulation of fructose metabolism, consistent with dietary preferences of tayra involving a variety of fruit and honey (Kratzer et al., 2014). The uricase (*UOX*) gene, which is involved in regulation of purine metabolism and conversion of fructose to fat (Johnson & Andrews, 2010) was under positive selection in tayra. Old World monkeys show low levels of uricase enzymatic activity, and uricase is pseudogenized in humans, great apes, and New World monkeys due to silencing mutations (Oda et al., 2002).

High rates of lineage-specific variation in gene family size, especially those families involved in immune response or detoxification of xenobiotic molecules (Thomas, 2007), are likely associated with environmental changes during speciation (Lynch & Conery, 2000; Zhang, 2003). We found a duplication of *N6AMT1*, a member of a gene family implicated in environmental stress response. *N6AMT1* is associated with conversion of the arsenic metabolite monomethylarsonous acid to the less toxic dimethylarsonic acid (Ren et al., 2011). Chronic arsenic exposure through water and soil, has been monitored in the neotropics, as it represents a severe public health issue, leading to various diseases (Zhang et al., 2015). This duplication may represent an adaptation of tayra to this xenobiotic compound.

Finally, we also found candidate genes associated with vision and cognition, which may be associated with the tayra's primarily diurnal foraging activity. Tayras supposedly detect prey primarily by smell, as their eyesight has been described as being relatively poor (Galef et al., 1976; Wilson & Mittermeier, 2009). This is contradictory to the observed behaviour of caching of unripe but mature stages of both native and non-native fruit by tayras (Soley & Alvarado-Díaz, 2011). We detected variation in several genes implicated with lens fiber formation and retinal vascularization, including gene expansions of *ANKRD13A* (Avellino et al., 2013) and *RBP2* (D'Ambrosio et al., 2011), suggesting that tayras might not in fact exhibit poor eyesight. The caching behavior mentioned above indicates remarkable sensory perception (Presley, 2000), and recognition of different stages of fruit development, likely evolved as a response to avoid competition with other species during seasons when fruit becomes ripe (Soley & Alvarado-Díaz, 2011). We observed genes related to memory and learning such as *SLC38A1* (Qureshi et al., 2019) and *CRBN* (Higgins et al., 2010) that may have been involved in the evolution of caching behavior in tayra.

Mustelids are a remarkable example of adaptive radiation, and we show how, together with positively selected loci, changes in gene family size and structural variants have shaped genomes in this diverse taxonomic group. We demonstrated that the latter two sources of variation contribute significantly to the number of loci potentially involved in adaptive genomic evolution. These variant types are often not considered in comparative genomic studies of non-model species, even though they impact more nucleotides than SNPs.

Acknowledgments

We thank Dr. Rafael from the Felidae Wildkatzen- und Artenschutzzentrum Barnim for kindly providing the tayra sample, and Michael Hofreiter from the Adaptive Genomics group (University of Potsdam) for assistance in generating the 10x Genomics linked-read library.

David Duchêne was funded by a Carlsbergfondet postdoctoral fellowship (grant number CF18-0223). Sergei Kliver, Andrey Tomarovsky and Azamat Totikov were funded by the Russian Foundation for Basic Research (grant number 20-04-00808). Azamat Totikov and Andrey Tomarovsky were additionally funded by JetBrains Research.

References

Abduriyim, S., Nishita, Y., Abramov, A. V., Solovyev, V. A., Saveljev, A. P., Kosintsev, P. A., Kryukov, A. P., Raichev, E., Väinölä, R., Kaneko, Y., & Masuda, R. (2019). Variation in pancreatic amylase gene copy number among Eurasian badgers (Carnivora, Mustelidae, Meles) and its relationship to diet. *Journal of Zoology*, *308*(1), 28–36.

Alonge, M., Soyk, S., Ramakrishnan, S., Wang, X., Goodwin, S., Sedlazeck, F. J., Lippman, Z. B., & Schatz, M. C. (2019). RaGOO: fast and accurate reference-guided scaffolding of draft genomes. *Genome Biology*, *20*(1), 224.

Al-Qusairi, L., Prokic, I., Amoasii, L., Kretz, C., Messaddeq, N., Mandel, J.-L., & Laporte, J. (2013). Lack of myotubularin (MTM1) leads to muscle hypotrophy through unbalanced regulation of the autophagy and ubiquitin-proteasome pathways. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology*, *27*(8), 3384–3394.

Anisimova, M., & Gascuel, O. (2006). Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Systematic Biology*, *55*(4), 539–552.

Armstrong, E. E., Taylor, R. W., Prost, S., Blinston, P., van der Meer, E., Madzikanda, H., Mufute, O., Mandisodza-Chikerema, R., Stuelpnagel, J., Sillero-Zubiri, C., & Petrov, D. (2019). Cost-effective assembly of the African wild dog (Lycaon pictus) genome using linked reads. *GigaScience*, *8*(2). https://doi.org/10.1093/gigascience/giy124

Avellino, R., Carrella, S., Pirozzi, M., Risolino, M., Salierno, F. G., Franco, P., Stoppelli, P., Verde, P., Banfi, S., & Conte, I. (2013). miR-204 targeting of Ankrd13A controls both mesenchymal neural crest and lens cell migration. *PloS One*, *8*(4), e61099.

Barnett, R., Westbury, M. V., Sandoval-Velasco, M., Vieira, F. G., Jeon, S., Zazula, G., Martin, M. D., Ho, S. Y. W., Mather, N., Gopalakrishnan, S., Ramos-Madrigal, J., de Manuel, M., Zepeda-Mendoza, M. L., Antunes, A., Baez, A. C., De Cahsan, B., Larson, G., O'Brien, S. J., Eizirik, E., … Gilbert, M. T. P. (2020). Genomic Adaptations and Evolutionary History of the Extinct Scimitar-Toothed Cat, Homotherium latidens. *Current Biology: CB*, *30*(24), 5018–5025.e5.

Beichman, A. C., Koepfli, K.-P., Li, G., Murphy, W., Dobrynin, P., Kilver, S., Tinker, M. T., Murray, M. J., Johnson, J., Lindblad-Toh, K., Karlsson, E. K., Lohmueller, K. E., & Wayne, R. K. (2019). Aquatic adaptation and depleted diversity: a deep dive into the genomes of the sea otter and giant otter. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/msz101

Blottner, S., Schön, J., & Jewgenow, K. (2006). Seasonally activated spermatogenesis is correlated with increased testicular production of testosterone and epidermal growth factor in mink (Mustela vison). *Theriogenology*, *66*(6-7), 1593–1598.

Broad Institute. (2019). Picard Toolkit. GitHub Repository. https://broadinstitute.github.io/picard/

Brown, J. A., Eberhardt, D. M., Schrick, F. N., Roberts, M. P., & Godkin, J. D. (2003). Expression of retinol-binding protein and cellular retinol-binding protein in the bovine ovary. *Molecular*

Reproduction and Development, *64*(3), 261–269.
Brown, J. H., & Lasiewski, R. C. (1972). Metabolism of weasels: The cost of being long and thin. *Ecology*, *53*(5), 939–943.
Buckley, R. M., Davis, B. W., Brashear, W. A., Farias, F. H. G., Kuroki, K., Graves, T., Hillier, L. W., Kremitzki, M., Li, G., Middleton, R. P., Minx, P., Tomlinson, C., Lyons, L. A., Murphy, W. J., & 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 Genetics*, *16*(10), e1008926.
Cahill, J. A., Soares, A. E. R., Green, R. E., & Shapiro, B. (2016). Inferring species divergence times using pairwise sequential Markovian coalescent modelling and low-coverage genomic data. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *371*(1699). https://doi.org/10.1098/rstb.2015.0138
Carmo, S. D., Fournier, D., Mounier, C., & Rassart, E. (2009). Human apolipoprotein D overexpression in transgenic mice induces insulin resistance and alters lipid metabolism. *American Journal of Physiology-Endocrinology and Metabolism*, *296*(4), E802–E811.
Catanach, A., Crowhurst, R., Deng, C., David, C., Bernatchez, L., & Wellenreuther, M. (2019). The genomic pool of standing structural variation outnumbers single nucleotide polymorphism by threefold in the marine teleost Chrysophrys auratus. *Molecular Ecology*, *28*(6), 1210–1223.
Cavagna, P., Menotti, A., & Stanyon, R. (2000). Genomic homology of the domestic ferret with cats and humans. *Mammalian Genome: Official Journal of the International Mammalian Genome Society*, *11*(10), 866–870.
Chang, J.-M., Di Tommaso, P., & Notredame, C. (2014). TCS: a new multiple sequence alignment reliability measure to estimate alignment accuracy and improve phylogenetic tree reconstruction. *Molecular Biology and Evolution*, *31*(6), 1625–1637.
Chen, X., Schulz-Trieglaff, O., Shaw, R., Barnes, B., Schlesinger, F., Källberg, M., Cox, A. J., Kruglyak, S., & Saunders, C. T. (2016). Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* , *32*(8), 1220–1222.
Chiang, C., Layer, R. M., Faust, G. G., Lindberg, M. R., Rose, D. B., Garrison, E. P., Marth, G. T., Quinlan, A. R., & Hall, I. M. (2015). SpeedSeq: ultra-fast personal genome analysis and interpretation. *Nature Methods*, *12*(10), 966–968.
Chiang, C., Scott, A. J., Davis, J. R., Tsang, E. K., Li, X., Kim, Y., Hadzic, T., Damani, F. N., Ganel, L., GTEx Consortium, Montgomery, S. B., Battle, A., Conrad, D. F., & Hall, I. M. (2017). The impact of structural variation on human gene expression. *Nature Genetics*, *49*(5), 692–699.
Coan, P. M., Vaughan, O. R., Sekita, Y., Finn, S. L., Burton, G. J., Constancia, M., & Fowden, A. L. (2010). Adaptations in placental phenotype support fetal growth during undernutrition of pregnant mice. *The Journal of Physiology*, *588*(Pt 3), 527–538.
Copeland, J. P., & Kucera, T. E. (1997). Wolverine (Gulo gulo). In *Mesocarnivores of Northern California: Biology, Management, and Survey Techniques*, Workshop Manual. Harris, John E., and Chester V. Ogan., Eds. Humboldt State Univ., Arcata, CA. The Wildlife Society, California North Coast Chapter, Arcata, CA. 127 p
Crow, Y. J., Chase, D. S., Lowenstein Schmidt, J., Szynkiewicz, M., Forte, G. M. A., Gornall, H. L., Oojageer, A., Anderson, B., Pizzino, A., Helman, G., Abdel-Hamid, M. S., Abdel-Salam, G. M., Ackroyd, S., Aeby, A., Agosta, G., Albin, C., Allon-Shalev, S., Arellano, M., Ariaudo, G., … Rice, G. I. (2015). Characterization of human disease phenotypes associated with mutations in TREX1, RNASEH2A, RNASEH2B, RNASEH2C, SAMHD1, ADAR, and IFIH1. *American Journal of Medical Genetics. Part A*, *167A*(2), 296–312.
D'Ambrosio, D. N., Clugston, R. D., & Blaner, W. S. (2011). Vitamin A metabolism: an update. *Nutrients*, *3*(1), 63–103.
Denker, H. W. (2000). Structural dynamics and function of early embryonic coats. *Cells, Tissues, Organs*,

166(2), 180–207.
DeSelm, C. J., Miller, B. C., Zou, W., Beatty, W. L., van Meel, E., Takahata, Y., Klumperman, J., Tooze, S. A., Teitelbaum, S. L., & Virgin, H. W. (2011). Autophagy proteins regulate the secretory component of osteoclastic bone resorption. *Developmental Cell*, *21*(5), 966–974.
Dobrynin, P., Liu, S., Tamazian, G., Xiong, Z., Yurchenko, A. A., Krasheninnikova, K., Kliver, S., Schmidt-Küntzel, A., Koepfli, K.-P., Johnson, W., Kuderna, L. F. K., García-Pérez, R., Manuel, M. de, Godinez, R., Komissarov, A., Makunin, A., Brukhin, V., Qiu, W., Zhou, L., … O'Brien, S. J. (2015). Genomic legacy of the African cheetah, Acinonyx jubatus. *Genome Biology*, *16*, 277.
Dudchenko, O., Batra, S. S., Omer, A. D., Nyquist, S. K., Hoeger, M., Durand, N. C., Shamim, M. S., Machol, I., Lander, E. S., Aiden, A. P., & Aiden, E. L. (2017). De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. *Science*, *356*(6333), 92–95.
Dudchenko, O., Shamim, M. S., Batra, S. S., Durand, N. C., Musial, N. T., Mostofa, R., Pham, M., St Hilaire, B. G., Yao, W., Stamenova, E., Hoeger, M., Nyquist, S. K., Korchina, V., Pletch, K., Flanagan, J. P., Tomaszewicz, A., McAloose, D., Estrada, C. P., Novak, B. J., … Aiden, E. L. (2018). *The Juicebox Assembly Tools module facilitates de novo assembly of mammalian genomes with chromosome-length scaffolds for under $1000*. In *bioRxiv (*p. 254797) https://doi.org/10.1101/254797
Ebert, P., Audano, P. A., Zhu, Q., Rodriguez-Martin, B., Porubsky, D., Bonder, M. J., Sulovari, A., Ebler, J., Zhou, W., Serra Mari, R., Yilmaz, F., Zhao, X., Hsieh, P., Lee, J., Kumar, S., Lin, J., Rausch, T., Chen, Y., Ren, J., … Eichler, E. E. (2021). Haplotype-resolved diverse human genomes and integrated analysis of structural variation. *Science*, *372*(6537).
Ekblom, R., Brechlin, B., Persson, J., Smeds, L., Johansson, M., Magnusson, J., Flagstad, Ø., & Ellegren, H. (2018). Genome sequencing and conservation genomics in the Scandinavian wolverine population. *Conservation Biology: The Journal of the Society for Conservation Biology*, *32*(6), 1301–1312.
Ellerman, D. A., Pei, J., Gupta, S., Snell, W. J., Myles, D., & Primakoff, P. (2009). Izumo is part of a multiprotein family whose members form large complexes on mammalian sperm. *Molecular Reproduction and Development*, *76*(12), 1188–1199.
Etherington, G. J., Heavens, D., Baker, D., Lister, A., McNelly, R., Garcia, G., Clavijo, B., Macaulay, I., Haerty, W., & Di Palma, F. (2020). Sequencing smart: De novo sequencing and assembly approaches for a non-model mammal. *GigaScience*, *9*(5). https://doi.org/10.1093/gigascience/giaa045
Fahien, L. A., & Macdonald, M. J. (2011). The complex mechanism of glutamate dehydrogenase in insulin secretion. *Diabetes*, *60*(10), 2450–2454.
Fenelon, J. C., Shaw, G., Frankenberg, S. R., Murphy, B. D., & Renfree, M. B. (2017). Embryo arrest and reactivation: potential candidates controlling embryonic diapause in the tammar wallaby and mink†. *Biology of Reproduction*, *96*(4), 877–894.
Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., & Smit, A. F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(17), 9451–9457.
Forde, N., Simintiras, C. A., Sturmey, R., Mamo, S., Kelly, A. K., Spencer, T. E., Bazer, F. W., & Lonergan, P. (2014). Amino acids in the uterine luminal fluid reflects the temporal changes in transporter expression in the endometrium and conceptus during early pregnancy in cattle. *PloS One*, *9*(6), e100010.
Frith, M. C., & Kawaguchi, R. (2015). Split-alignment of genomes finds orthologies more accurately. *Genome Biology*, *16*, 106.
Galef, B. G., Mittermeier, R. A., & Bailey, R. C. (1976). Predation by the Tayra (Eira barbara). *Journal of Mammalogy*, *57*(4), 760–761.
Ge, S. X., Jung, D., & Yao, R. (2020). ShinyGO: a graphical gene-set enrichment tool for animals and

plants. *Bioinformatics* , *36*(8), 2628–2629.
Gosejacob, D., Jäger, P. S., Vom Dorp, K., Frejno, M., Carstensen, A. C., Köhnke, M., Degen, J., Dörmann, P., & Hoch, M. (2016). Ceramide Synthase 5 Is Essential to Maintain C16:0-Ceramide Pools and Contributes to the Development of Diet-induced Obesity. *The Journal of Biological Chemistry*, *291*(13), 6989–7003.
Grayson, P., & Civetta, A. (2012). Positive Selection and the Evolution of izumo Genes in Mammals. *International Journal of Evolutionary Biology*, *2012*, 958164.
Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* , *29*(8), 1072–1075.
Han, M. V., Thomas, G. W. C., Lugo-Martinez, J., & Hahn, M. W. (2013). Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Molecular Biology and Evolution*, *30*(8), 1987–1997.
Han, X., Jiang, T., Yu, L., Zeng, C., Fan, B., & Liu, B. (2012). Molecular characterization of the porcine MTPAP gene associated with meat quality traits: chromosome localization, expression distribution, and transcriptional regulation. *Molecular and Cellular Biochemistry*, *364*(1-2), 173–180.
Harney, J. P., Ott, T. L., Geisert, R. D., & Bazer, F. W. (1993). Retinol-binding protein gene expression in cyclic and pregnant endometrium of pigs, sheep, and cattle. *Biology of Reproduction*, *49*(5), 1066–1073.
Hecker, N., Sharma, V., & Hiller, M. (2019). Convergent gene losses illuminate metabolic and physiological changes in herbivores and carnivores. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(8), 3036–3041.
Heldstab, S. A., Müller, D. W. H., Graber, S. M., Bingaman Lackey, L., Rensch, E., Hatt, J.-M., Zerbe, P., & Clauss, M. (2018). Geographical Origin, Delayed Implantation, and Induced Ovulation Explain Reproductive Seasonality in the Carnivora. *Journal of Biological Rhythms*, *33*(4), 402–419.
Higgins, J. J., Tal, A. L., Sun, X., Hauck, S. C. R., Hao, J., Kosofosky, B. E., & Rajadhyaksha, A. M. (2010). Temporal and spatial mouse brain expression of cereblon, an ionic channel regulator involved in human intelligence. *Journal of Neurogenetics*, *24*(1), 18–26.
Hill, M., Pařízek, A., Kancheva, R., & Jirásek, J. E. (2011). Reduced progesterone metabolites in human late pregnancy. *Physiological Research / Academia Scientiarum Bohemoslovaca*, *60*(2), 225–241.
Hron, T., Elleder, D., & Gifford, R. J. (2019). Deltaretroviruses have circulated since at least the Paleogene and infected a broad range of mammalian species. *Retrovirology*, *16*(1), 33.
Inman, R. M., Magoun, A. J., Persson, J., & Mattisson, J. (2012). The wolverine's niche: linking reproductive chronology, caching, competition, and climate. *Journal of Mammalogy*, *93*(3), 634–644.
Inoue, K., Maeda, N., Mori, T., Sekimoto, R., Tsushima, Y., Matsuda, K., Yamaoka, M., Suganami, T., Nishizawa, H., Ogawa, Y., Funahashi, T., & Shimomura, I. (2014). Possible involvement of Opa-interacting protein 5 in adipose proliferation and obesity. *PloS One*, *9*(2), e87661.
Jeffares, D. C., Jolly, C., Hoti, M., Speed, D., Shaw, L., Rallis, C., Balloux, F., Dessimoz, C., Bähler, J., & Sedlazeck, F. J. (2017). Transient structural variations have strong effects on quantitative traits and reproductive isolation in fission yeast. *Nature Communications*, *8*, 14061.
Jewgenow, K., Goeritz, F., Neubauer, K., Fickel, J., & Naidenko, S. V. (2006). Characterization of reproductive activity in captive male Eurasian lynx (Lynx lynx). *European Journal of Wildlife Research*, *52*(1), 34–38.
Johnson, R. J., & Andrews, P. (2010). Fructose, uricase, and the Back-to-Africa hypothesis. *Evolutionary Anthropology*, *19*(6), 250–257.
Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589.
Kashtanov, S. N., Svischeva, G. R., Pishchulina, S. L., Lazebny, O. E., Meshchersky, I. G., Simakin, L.

V., & Rozhnov, V. V. (2015). Geographical structure of the sable (Martes zibellina L.) gene pool on the basis of microsatellite loci analysis. *Russian Journal of Genetics*, *51*(1), 69–79.

Kim, B.-M., Lee, Y. J., Kim, J.-H., Kim, J.-H., Kang, S., Jo, E., Lee, S. J., Lee, J. H., Chi, Y. M., & Park, H. (2020). The Genome Assembly and Annotation of the Southern Elephant Seal Mirounga leonina. *Genes*, *11*(2). https://doi.org/10.3390/genes11020160

Kitchener, A. C., Meloro, C., & Williams, T. M. (2018). Form and function of the musteloids. In *Biology and Conservation of Musteloids*. Oxford University Press.

Koepfli, K.-P., Deere, K. A., Slater, G. J., Begg, C., Begg, K., Grassman, L., Lucherini, M., Veron, G., & Wayne, R. K. (2008). Multigene phylogeny of the Mustelidae: resolving relationships, tempo and biogeographic history of a mammalian adaptive radiation. *BMC Biology*, *6*, 10.

Kosiol, C., Vinar, T., da Fonseca, R. R., Hubisz, M. J., Bustamante, C. D., Nielsen, R., & Siepel, A. (2008). Patterns of positive selection in six Mammalian genomes. *PLoS Genetics*, *4*(8), e1000144.

Kratzer, J. T., Lanaspa, M. A., Murphy, M. N., Cicerchi, C., Graves, C. L., Tipton, P. A., Ortlund, E. A., Johnson, R. J., & Gaucher, E. A. (2014). Evolutionary history and metabolic insights of ancient mammalian uricases. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(10), 3763–3768.

Krebs, J., Lofroth, E., Copeland, J., Banci, V., Cooley, D., Golden, H., Magoun, A., Mulders, R., & Shults, B. (2004). Synthesis of survival rates and causes of mortality in north American wolverines. *The Journal of Wildlife Management*, *68*(3), 493–502.

Kriventseva, E. V., Kuznetsov, D., Tegenfeldt, F., Manni, M., Dias, R., Simão, F. A., & Zdobnov, E. M. (2019). OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Research*, *47*(D1), D807–D811.

Kronenberg, Z. N., Osborne, E. J., Cone, K. R., Kennedy, B. J., Domyan, E. T., Shapiro, M. D., Elde, N. C., & Yandell, M. (2015). Wham: Identifying Structural Variants of Biological Consequence. *PLoS Computational Biology*, *11*(12), e1004572.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359.

Larivière, S., & Ferguson, S. H. (2003). Evolution of Induced Ovulation in North American Carnivores. *Journal of Mammalogy*, *84*(3), 937–947.

Law, C. J., Slater, G. J., & Mehta, R. S. (2018). Lineage Diversity and Size Disparity in Musteloidea: Testing Patterns of Adaptive Radiation Using Molecular and Fossil-Based Methods. *Systematic Biology*, *67*(1), 127–144.

Layer, R. M., Chiang, C., Quinlan, A. R., & Hall, I. M. (2014). LUMPY: a probabilistic framework for structural variant discovery. *Genome Biology*, *15*(6), R84.

Lee, J.-E., Oh, H.-A., Song, H., Jun, J. H., Roh, C.-R., Xie, H., Dey, S. K., & Lim, H. J. (2011). Autophagy regulates embryonic survival during delayed implantation. *Endocrinology*, *152*(5), 2067–2075.

Lennox, A. R., & Goodship, A. E. (2008). Polar bears (Ursus maritimus), the most evolutionary advanced hibernators, avoid significant bone loss during hibernation. *Comparative Biochemistry and Physiology. Part A, Molecular & Integrative Physiology*, *149*(2), 203–208.

Lewin, H. A., Graves, J. A. M., Ryder, O. A., Graphodatsky, A. S., & O'Brien, S. J. (2019). Precision nomenclature for the new genomics. *GigaScience*, *8*(8). https://doi.org/10.1093/gigascience/giz086

Li, B., Wolsan, M., Wu, D., Zhang, W., Xu, Y., & Zeng, Z. (2014). Mitochondrial genomes reveal the pattern and timing of marten (Martes), wolverine (Gulo), and fisher (Pekania) diversification. *Molecular Phylogenetics and Evolution*, *80*, 156–164.

Li, H., & Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. *Nature*, *475*(7357), 493–496.

Li, X., Ye, Y., Wang, B., & Zhao, S. (2021). miR-140-5p Aggravates Insulin Resistance via Directly

Targeting GYS1 and PPP1CC in Insulin-Resistant HepG2 Cells. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy*, *14*, 2515–2524.

Lim, H. J., & Song, H. (2014). Evolving tales of autophagy in early reproductive events. *The International Journal of Developmental Biology*, *58*(2-4), 183–187.

Liu, G., Zhao, C., Xu, D., Zhang, H., Monakhov, V., Shang, S., Gao, X., Sha, W., Ma, J., Zhang, W., Tang, X., Li, B., Hua, Y., Cao, X., Liu, Z., & Zhang, H. (2020). First Draft Genome of the Sable, Martes zibellina. *Genome Biology and Evolution*, *12*(3), 59–65.

Liu, L., Chen, S., Yu, M., Ge, C., Ren, M., Liu, B., Yang, X., Christian, T. W., Jr, Hou, Y.-M., Zou, J., Zhu, W.-G., & Luo, J. (2020). Deacetylation of HSD17B10 by SIRT3 regulates cell growth and cell resistance under oxidative and starvation stresses. *Cell Death & Disease*, *11*(7), 563.

Lopes, F. L., Desmarais, J., Gevry, N. Y., Ledoux, S., & Murphy, B. D. (2003). Expression of vascular endothelial growth factor isoforms and receptors Flt-1 and KDR during the peri-implantation period in the mink, Mustela vison. *Biology of Reproduction*, *68*(5), 1926–1933.

Lopes-Marques, M., Serrano, C., Cardoso, A. R., Salazar, R., Seixas, S., Amorim, A., Azevedo, L., & Prata, M. J. (2020). GBA3: a polymorphic pseudogene in humans that experienced repeated gene loss during mammalian evolution. *Scientific Reports*, *10*(1), 11565.

Löytynoja, A. (2014). Phylogeny-aware alignment with PRANK. *Methods in Molecular Biology* , *1079*, 155–170.

Lynch, M., & Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes [Review of *The evolutionary fate and consequences of duplicate genes*]. *Science*, *290*(5494), 1151–1155.

Maglott, D., Ostell, J., Pruitt, K. D., & Tatusova, T. (2011). Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Research*, *39*(Database issue), D52–D57.

Marín, I. (2014). The ubiquilin gene family: evolutionary patterns and functional insights. *BMC Evolutionary Biology*, *14*, 63.

Martinez, B., & Ortiz, R. M. (2017). Thyroid Hormone Regulation and Insulin Resistance: Insights From Animals Naturally Adapted to Fasting. *Physiology* . https://doi.org/10.1152/physiol.00018.2016

McGee-Lawrence, M., Buckendahl, P., Carpenter, C., Henriksen, K., Vaughan, M., & Donahue, S. (2015). Suppressed bone remodeling in black bears conserves energy and bone mass during hibernation. *The Journal of Experimental Biology*, *218*(Pt 13), 2067–2074.

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., Flicek, P., & Cunningham, F. (2016). The Ensembl Variant Effect Predictor. *Genome Biology*, *17*(1), 122.

McNutt, J. W., Groom, R., & Woodroffe, R. (2019). Ambient temperature provides an adaptive explanation for seasonal reproduction in a tropical mammal. *Journal of Zoology* , *309*(3), 153–160.

Mead, R. A. (1981). Delayed implantation in mustelids, with special emphasis on the spotted skunk. *Journal of Reproduction and Fertility. Supplement*, *29*, 11–24.

Mead, R. A. (1989). The Physiology and Evolution of Delayed Implantation in Carnivores. In J. L. Gittleman (Ed.), *Carnivore Behavior, Ecology, and Evolution* (pp. 437–464). Springer US.

Mérot, C., Oomen, R. A., Tigano, A., & Wellenreuther, M. (2020). A Roadmap for Understanding the Evolutionary Significance of Structural Genomic Variation. *Trends in Ecology & Evolution*. https://doi.org/10.1016/j.tree.2020.03.002

Minh, B. Q., Hahn, M. W., & Lanfear, R. (2020). New Methods to Calculate Concordance Factors for Phylogenomic Datasets. *Molecular Biology and Evolution*, *37*(9), 2727–2733.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, *37*(5), 1530–1534.

Miranda, I., Giska, I., Farelo, L., Pimenta, J., Zimova, M., Bryk, J., Dalén, L., Mills, L. S., Zub, K., & Melo-Ferreira, J. (2021). Museomics dissects the genetic basis for adaptive seasonal colouration in the least weasel. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/msab177

Mizuarai, S., Miki, S., Araki, H., Takahashi, K., & Kotani, H. (2005). Identification of dicarboxylate

carrier Slc25a10 as malate transporter in de novo fatty acid synthesis. *The Journal of Biological Chemistry*, *280*(37), 32434–32441.
Monakhov, V. G. (2011). Martes zibellina (Carnivora: Mustelidae). *Mammalian Species*, *43*(876), 75–86.
Montaseri, A., Giampietri, C., Rossi, M., Riccioli, A., Del Fattore, A., & Filippini, A. (2020). The Role of Autophagy in Osteoclast Differentiation and Bone Resorption Function. *Biomolecules*, *10*(10). https://doi.org/10.3390/biom10101398
Murphy, B. D., & Fenelon, J. C. (2020). Models of embryonic diapause in Carnivora. *Bioscientifica Proceedings*. https://doi.org/10.1530/biosciprocs.10.006
Mustonen, A.-M., Käkelä, R., Käkelä, A., Pyykönen, T., Aho, J., & Nieminen, P. (2007). Lipid metabolism in the adipose tissues of a carnivore, the raccoon dog, during prolonged fasting. *Experimental Biology and Medicine* , *232*(1), 58–69.
Mustonen, A.-M., Puukka, M., Saarela, S., Paakkonen, T., Aho, J., & Nieminen, P. (2006). Adaptations to fasting in a terrestrial mustelid, the sable (Martes zibellina). *Comparative Biochemistry and Physiology. Part A, Molecular & Integrative Physiology*, *144*(4), 444–450.
Nielsen, R., Bustamante, C., Clark, A. G., Glanowski, S., Sackton, T. B., Hubisz, M. J., Fledel-Alon, A., Tanenbaum, D. M., Civello, D., White, T. J., J Sninsky, J., Adams, M. D., & Cargill, M. (2005). A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biology*, *3*(6), e170.
Nieminen, P., Käkelä, R., Pyykönen, T., & Mustonen, A.-M. (2006). Selective fatty acid mobilization in the American mink (Mustela vison) during food deprivation. *Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology*, *145*(1), 81–93.
Nieminen, P., & Mustonen, A.-M. (2007). Uniform fatty acid mobilization from anatomically distinct fat depots in the sable (Martes zibellina). *Lipids*, *42*(7), 659–669.
O'Brien, S. J., Graphodatsky, A. S., & Perelman, P. L. (2020). *Atlas of Mammalian Chromosomes*. John Wiley & Sons.
Oda, M., Satta, Y., Takenaka, O., & Takahata, N. (2002). Loss of urate oxidase activity in hominoids and its evolutionary implications. *Molecular Biology and Evolution*, *19*(5), 640–653.
Okamoto, T., Imaizumi, K., & Kaneko, M. (2020). The Role of Tissue-Specific Ubiquitin Ligases, RNF183, RNF186, RNF182 and RNF152, in Disease and Biological Function. *International Journal of Molecular Sciences*, *21*(11). https://doi.org/10.3390/ijms21113921
Olson, M. V. (1999). When less is more: gene loss as an engine of evolutionary change. *American Journal of Human Genetics*, *64*(1), 18–23.
Pacifici, M., Santini, L., Di Marco, M., Baisero, D., Francucci, L., Grottolo Marasini, G., Visconti, P., & Rondinini, C. (2013). Generation length for mammals. *Nature Conservation*, *5*, 89–94.
Papin, C., Rouget, C., Lorca, T., Castro, A., & Mandart, E. (2004). XCdh1 is involved in progesterone-induced oocyte maturation. *Developmental Biology*, *272*(1), 66–75.
Pasitschniak-Arts, M., & Larivière, S. (1995). Gulo gulo. *Mammalian Species*, *499*, 1–10.
Peng, C., Niu, L., Deng, J., Yu, J., Zhang, X., Zhou, C., Xing, J., & Li, J. (2018). Can-SINE dynamics in the giant panda and three other Caniformia genomes. *Mobile DNA*, *9*, 32.
Peng, X., Alföldi, J., Gori, K., Eisfeld, A. J., Tyler, S. R., Tisoncik-Go, J., Brawand, D., Law, G. L., Skunca, N., Hatta, M., Gasper, D. J., Kelly, S. M., Chang, J., Thomas, M. J., Johnson, J., Berlin, A. M., Lara, M., Russell, P., Swofford, R., … Katze, M. G. (2014). The draft genome sequence of the ferret (Mustela putorius furo) facilitates study of human respiratory disease. *Nature Biotechnology*, *32*(12), 1250–1255.
Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: convergence diagnosis and output analysis for MCMC. *R News*, *6*(1), 7–11.
Poglayen-Neuwall, I., Durrant, B. S., Swansen, M. L., Williams, R. C., & Barnes, R. A. (1989). Estrous cycle of the tayra,Eira barbara. *Zoo Biology*, *8*(2), 171–177.
Pontremoli, C., Forni, D., Pozzoli, U., Clerici, M., Cagliani, R., & Sironi, M. (2021). Kinetochore

proteins and microtubule-destabilizing factors are fast evolving in eutherian mammals. *Molecular Ecology*, *30*(6), 1505–1515.

Poon, C. E., Lecce, L., Day, M. L., & Murphy, C. R. (2014). Mucin 15 is lost but mucin 13 remains in uterine luminal epithelial cells and the blastocyst at the time of implantation in the rat. *Reproduction, Fertility, and Development*, *26*(3), 421–431.

Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T. J., Carneiro, M. O., Van der Auwera, G. A., Kling, D. E., Gauthier, L. D., Levy-Moonshine, A., Roazen, D., Shakir, K., Thibault, J., Chandran, S., Whelan, C., Lek, M., Gabriel, S., Daly, M. J., Neale, B., MacArthur, D. G., & Banks, E. (2018). Scaling accurate genetic variant discovery to tens of thousands of samples. In *bioRxiv* (p. 201178). https://doi.org/10.1101/201178

Porubsky, D., Sanders, A. D., Höps, W., Hsieh, P., Sulovari, A., Li, R., Mercuri, L., Sorensen, M., Murali, S. C., Gordon, D., Cantsilieris, S., Pollen, A. A., Ventura, M., Antonacci, F., Marschall, T., Korbel, J. O., & Eichler, E. E. (2020). Recurrent inversion toggling and great ape genome evolution. *Nature Genetics*, *52*(8), 849–858.

Presley, S. J. (2000). Eira barbara. *Mammalian Species*, *2000*(636), 1–6.

Proulx, G., & Aubry, K. B. (2017). *The Martes complex: A monophyletic clade that shares many life-history traits and conservation challenges*. In *The Martes Complex in the 21st Century: ecology and conservation*, Proulx, Gilbert, Eds. Mammal Research Institute, Polish Academy of Sciences, Białowieża, Poland: 3-24.

Qureshi, T., Sørensen, C., Berghuis, P., Jensen, V., Dobszay, M. B., Farkas, T., Dalen, K. T., Guo, C., Hassel, B., Utheim, T. P., Hvalby, Ø., Hafting, T., Harkany, T., Fyhn, M., & Chaudhry, F. A. (2019). The Glutamine Transporter Slc38a1 Regulates GABAergic Neurotransmission and Synaptic Plasticity. *Cerebral Cortex* , *29*(12), 5166–5179.

Ranwez, V., Harispe, S., Delsuc, F., & Douzery, E. J. P. (2011). MACSE: Multiple Alignment of Coding SEquences accounting for frameshifts and stop codons. *PloS One*, *6*(9), e22594.

Reis, A., Chang, H.-Y., Levasseur, M., & Jones, K. T. (2006). APCcdh1 activity in mouse oocytes prevents entry into the first meiotic division. *Nature Cell Biology*, *8*(5), 539–540.

Ren, X., Aleshin, M., Jo, W. J., Dills, R., Kalman, D. A., Vulpe, C. D., Smith, M. T., & Zhang, L. (2011). Involvement of N-6 adenine-specific DNA methyltransferase 1 (N6AMT1) in arsenic biomethylation and its role in arsenic-induced toxicity. *Environmental Health Perspectives*, *119*(6), 771–777.

Reynolds, A., Qiao, H., Yang, Y., Chen, J. K., Jackson, N., Biswas, K., Holloway, J. K., Baudat, F., de Massy, B., Wang, J., Höög, C., Cohen, P. E., & Hunter, N. (2013). RNF212 is a dosage-sensitive regulator of crossing-over during mammalian meiosis. *Nature Genetics*, *45*(3), 269–278.

Rinker, D. C., Specian, N. K., Zhao, S., & Gibbons, J. G. (2019). Polar bear evolution is marked by rapid changes in gene copy number in response to dietary shift. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(27), 13446–13451.

Sato, J. J., Wolsan, M., Prevosti, F. J., D'Elía, G., Begg, C., Begg, K., Hosoda, T., Campbell, K. L., & Suzuki, H. (2012). Evolutionary and biogeographic history of weasel-like carnivorans (Musteloidea). *Molecular Phylogenetics and Evolution*, *63*(3), 745–757.

Sayyari, E., & Mirarab, S. (2016). Fast Coalescent-Based Computation of Local Branch Support from Quartet Frequencies. *Molecular Biology and Evolution*, *33*(7), 1654–1668.

Schulte-Hostedde, A. I., Zinner, B., Millar, J. S., & Hickling, G. J. (2005). Restitution of mass–size residuals: Validating body condition indices. *Ecology*, *86*(1), 155–163.

Schweitzer, L. D., Comb, W. C., Bar-Peled, L., & Sabatini, D. M. (2015). Disruption of the Rag-Ragulator Complex by c17orf59 Inhibits mTORC1. *Cell Reports*, *12*(9), 1445–1455.

Serão, N. V. L., Veroneze, R., Ribeiro, A. M. F., Verardo, L. L., Braccini Neto, J., Gasparino, E., Campos, C. F., Lopes, P. S., & Guimarães, S. E. F. (2011). Candidate gene expression and intramuscular fat content in pigs. *Journal of Animal Breeding and Genetics = Zeitschrift Fur*

Tierzuchtung Und Zuchtungsbiologie, *128*(1), 28–34.
Sharma, V., Hecker, N., Roscito, J. G., Foerster, L., Langer, B. E., & Hiller, M. (2018). A genomics approach reveals insights into the importance of gene losses for mammalian adaptations. *Nature Communications*, *9*(1), 1215.
Shi, G., Xing, L., Liu, Z., Qu, Z., Wu, X., Dong, Z., Wang, X., Gao, X., Huang, M., Yan, J., Yang, L., Liu, Y., Ptácek, L. J., & Xu, Y. (2013). Dual roles of FBXL3 in the mammalian circadian feedback loops are important for period determination and robustness of the clock. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(12), 4750–4755.
Shumate, A., & Salzberg, S. L. (2020). *Liftoff: an accurate gene annotation mapping tool* (p. 2020.06.24.169680). https://doi.org/10.1101/2020.06.24.169680
Shyu, M.-K., Lin, M.-C., Shih, J.-C., Lee, C.-N., Huang, J., Liao, C.-H., Huang, I.-F., Chen, H.-Y., Huang, M.-C., & Hsieh, F.-J. (2007). Mucin 15 is expressed in human placenta and suppresses invasion of trophoblast-like cells in vitro. *Human Reproduction* , *22*(10), 2723–2732.
Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* , *31*(19), 3210–3212.
Singh, B. N., Gong, W., Das, S., Theisen, J. W. M., Sierra-Pagan, J. E., Yannopoulos, D., Skie, E., Shah, P., Garry, M. G., & Garry, D. J. (2019). Etv2 transcriptionally regulates Yes1 and promotes cell proliferation during embryogenesis. *Scientific Reports*, *9*(1), 9736.
Smit, A. F. A. (2004). Repeat-Masker Open-3.0. *Http://www. Repeatmasker. Org*. https://ci.nii.ac.jp/naid/10029514778/
Soley, F. G., & Alvarado-Díaz, I. (2011). Prospective thinking in a mustelid? Eira barbara (Carnivora) cache unripe fruits to consume them once ripened. *Die Naturwissenschaften*, *98*(8), 693–698.
Sonna, L. A., Fujita, J., Gaffin, S. L., & Lilly, C. M. (2002). Invited review: Effects of heat and cold stress on mammalian gene expression. *Journal of Applied Physiology*, *92*(4), 1725–1742.
Svishcheva, G. R., & Kashtanov, S. N. (2011). Reproductive strategy of the sable (Martes zibellina Linnaeus, 1758): An analysis of litter size inheritance in farm-raised populations. *Russian Journal of Genetics: Applied Research*, *1*(3), 221–225.
Tavori, H., Rashid, S., & Fazio, S. (2015). On the function and homeostasis of PCSK9: reciprocal interaction with LDLR and additional lipid effects. *Atherosclerosis*, *238*(2), 264–270.
The UniProt Consortium. (2017). UniProt: the universal protein knowledgebase. *Nucleic Acids Research*, *45*(D1), D158–D169.
Thomas, J. H. (2007). Rapid birth-death evolution specific to xenobiotic cytochrome P450 genes in vertebrates. *PLoS Genetics*, *3*(5), e67.
Tigano, A., Colella, J. P., & MacManes, M. D. (2020). Comparative and population genomics approaches reveal the basis of adaptation to deserts in a small rodent. *Molecular Ecology*, *29*(7), 1300–1314.
Totikov, A., Tomarovsky, A., Prokopov, D., Yakupova, A., Bulyonkova, T., Derezanin, L., Rasskazov, D., Wolfsberger, W. W., Koepfli, K.-P., Oleksyk, T. K., & Kliver, S. (2021). Chromosome-Level Genome Assemblies Expand Capabilities of Genomics for Conservation Biology. *Genes*, *12*(9), 1336.
Viscarra, J. A., Rodriguez, R., Vazquez-Medina, J. P., Lee, A., Tift, M. S., Tavoni, S. K., Crocker, D. E., & Ortiz, R. M. (2013). Insulin and GLP-1 infusions demonstrate the onset of adipose-specific insulin resistance in a large fasting mammal: potential glucogenic role for GLP-1. *Physiological Reports*, *1*(2), e00023.
Weischenfeldt, J., Symmons, O., Spitz, F., & Korbel, J. O. (2013). Phenotypic impact of genomic structural variation: insights from and for human disease. *Nature Reviews. Genetics*, *14*(2), 125–138.
Weisenfeld, N. I., Kumar, V., Shah, P., Church, D. M., & Jaffe, D. B. (2017). Direct determination of diploid genome sequences. *Genome Research*, *27*(5), 757–767.
Weissensteiner, M. H., Bunikis, I., Catalán, A., Francoijs, K.-J., Knief, U., Heim, W., Peona, V., Pophaly,

S. D., Sedlazeck, F. J., Suh, A., Warmuth, V. M., & Wolf, J. B. W. (2020). Discovery and population genomics of structural variation in a songbird genus. *Nature Communications*, *11*(1), 3403.

Wellenreuther, M., & Bernatchez, L. (2018). Eco-Evolutionary Genomics of Chromosomal Inversions. *Trends in Ecology & Evolution*, *33*(6), 427–440.

Wilcox, A., Katsanakis, K. D., Bheda, F., & Pillay, T. S. (2004). Asb6, an adipocyte-specific ankyrin and SOCS box protein, interacts with APS to enable recruitment of elongins B and C to the insulin receptor signaling complex. *The Journal of Biological Chemistry*, *279*(37), 38881–38888.

Wilson, D. E., & Mittermeier, R. A. (Eds.). (2009). Family Mustelidae. In *Handbook of the mammals of the world, vol. 1 Carnivores* (Vol. 1, pp. 627–637). Lynx Editions, Barcelona.

Wu, X., Quondamatteo, F., Lefever, T., Czuchra, A., Meyer, H., Chrostek, A., Paus, R., Langbein, L., & Brakebusch, C. (2006). Cdc42 controls progenitor cell differentiation and beta-catenin turnover in skin. *Genes & Development*, *20*(5), 571–585.

Wu, Z., Latendorf, T., Meyer-Hoffert, U., & Schröder, J.-M. (2011). Identification of trichohyalin-like 1, an s100 fused-type protein selectively expressed in hair follicles. *The Journal of Investigative Dermatology*, *131*(8), 1761–1763.

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, *24*(8), 1586–1591.

Yockey, L. J., & Iwasaki, A. (2018). Interferons and Proinflammatory Cytokines in Pregnancy and Fetal Development. *Immunity*, *49*(3), 397–412.

Yu, Y.-H., Chang, Y.-C., Su, T.-H., Nong, J.-Y., Li, C.-C., & Chuang, L.-M. (2013). Prostaglandin reductase-3 negatively modulates adipogenesis through regulation of PPARγ activity. *Journal of Lipid Research*, *54*(9), 2391–2399.

Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, *19*(Suppl 6), 153.

Zhang, H., Ge, Y., He, P., Chen, X., Carina, A., Qiu, Y., Aga, D. S., & Ren, X. (2015). Interactive Effects of N6AMT1 and As3MT in Arsenic Biomethylation. *Toxicological Sciences: An Official Journal of the Society of Toxicology*, *146*(2), 354–362.

Zhang, J. (2003). Evolution by gene duplication: an update. *Trends in Ecology & Evolution*, *18*(6), 292–298.

Zhang, J., Nielsen, R., & Yang, Z. (2005). Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Molecular Biology and Evolution*, *22*(12), 2472–2479.

Zhou, Y.-B., Newman, C., Xu, W.-T., Buesching, C. D., Zalewski, A., Kaneko, Y., Macdonald, D. W., & Xie, Z.-Q. (2011). Biogeographical variation in the diet of Holarctic martens (genus Martes, Mammalia: Carnivora: Mustelidae): adaptive foraging in generalists. *Journal of Biogeography*, *38*(1), 137–147.

Data Accessibility and Benefit-Sharing Statement

The tayra primary and alternate genome sequences have been deposited in GenBank under the accession JAHRIG000000000-JAHRIH000000000. All data generated or analysed during this study are included in this published article and its supplemental information files. Bioinformatic scripts will be

publicly available in a Github repository specifically designated to this study, with a DOI assigned to it. Raw sequencing data will be deposited and available for download from the NCBI SRA repository.

Author contributions

LD conceived the study, conducted genome assembly, gene family evolution and structural variation analysis, interpreted the data, and wrote the manuscript. AB and SJ conducted positive selection analysis, and interpreted the data. PD carried out demographic history reconstruction and data interpretation. DAD designed and performed phylogenomic analysis and molecular dating, and wrote the manuscript. JHG conducted repetitive elements analysis and interpreted the data. SK, AT and AT performed reference-based scaffolding, alignment to pseudochromosome assemblies, sex verification and nucleotide diversity assessment, interpreted the data and worte the manuscript. KPK advised comparative genomic analyses, and edited the manuscript. DM supervised and performed gene family evolution analysis, and advised other comparative genomic analyses. MP conducted whole genome sequencing and sequencing data interpretation. JF and DWF conceived the study, interpreted the data, wrote and edited the manuscript. All authors have reviewed and approved the manuscript.

Tables and Figures

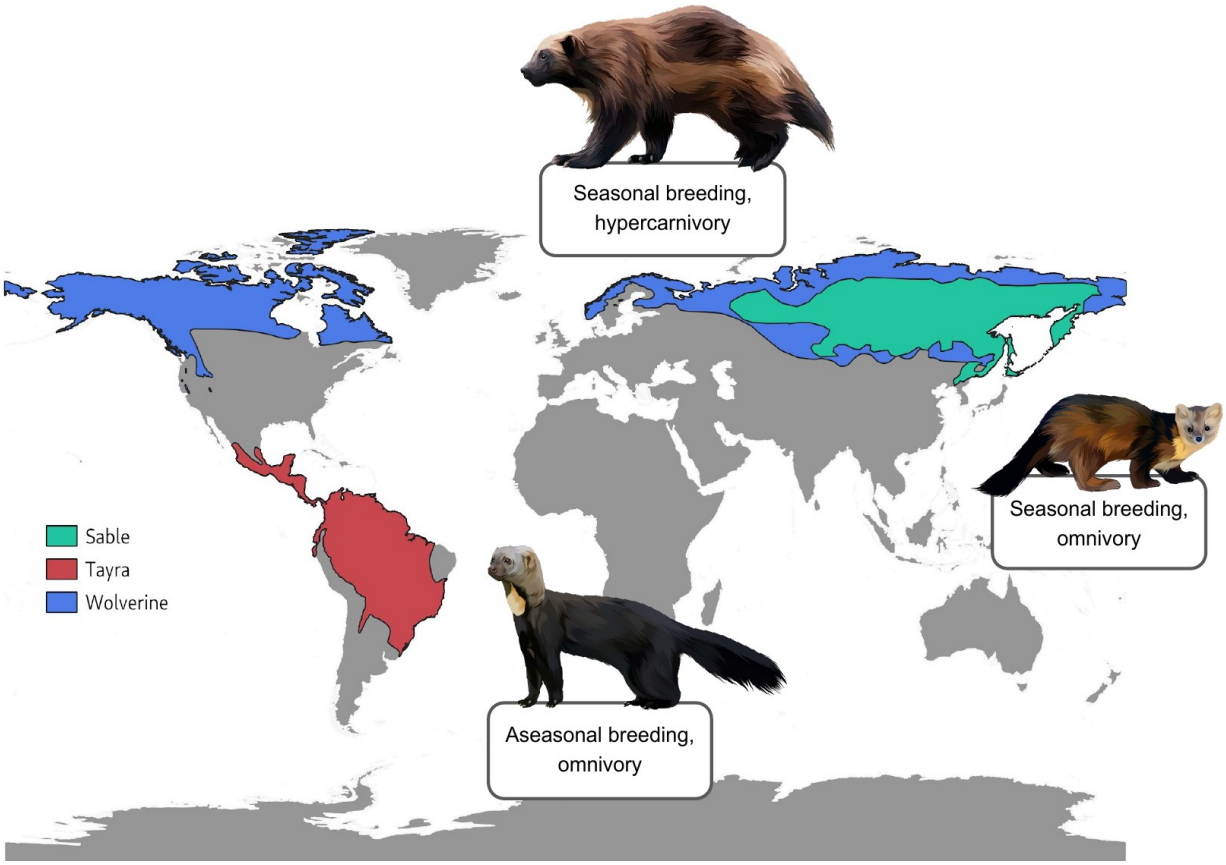

Figure 1. Distribution and species-specific traits of the tayra (*Eira barbara*), wolverine (*Gulo gulo*), and sable (*Martes zibellina*). Vector graphics of species are created based on royalty-free images (source: Shutterstock).

Table 1. Comparison of genome assembly metrics among four mustelid species.

Species	Tayra (*E. barbara*)	Domestic ferret (*M. putorius furo*)	Sable (*M. zibellina*)	Wolverine (*G. gulo*)
Assembly Accession/ reference	JAHRIG000000000 (This study)	(Dudchenko et al., 2017, 2018; X. Peng et al., 2014)	GCA_012583365.1 (G. Liu et al., 2020)	GCA_900006375.2 (Ekblom et al., 2018)
Sequencing/assembly approach	Illumina + 10xGenomics/ Supernova	Illumina/ ALLPATHS-LG + Hi-C scaffolding	Illumina/ SOAPdenovo2	Illumina/ SOAPdenovo
Raw coverage	75.6x	162x	114.5x	76x
Contig N50	289.9 kb	44.7 kb	41.7 kb	3.6 kb
Scaffold N50	42.0 Mb	145.3 Mb	5.2 Mb	0.2 Mb
Number of scaffolds	14579	7428	15814	47417
Total genome length	2.44 Gb	2.40 Gb	2.42 Gb	2.42 Gb

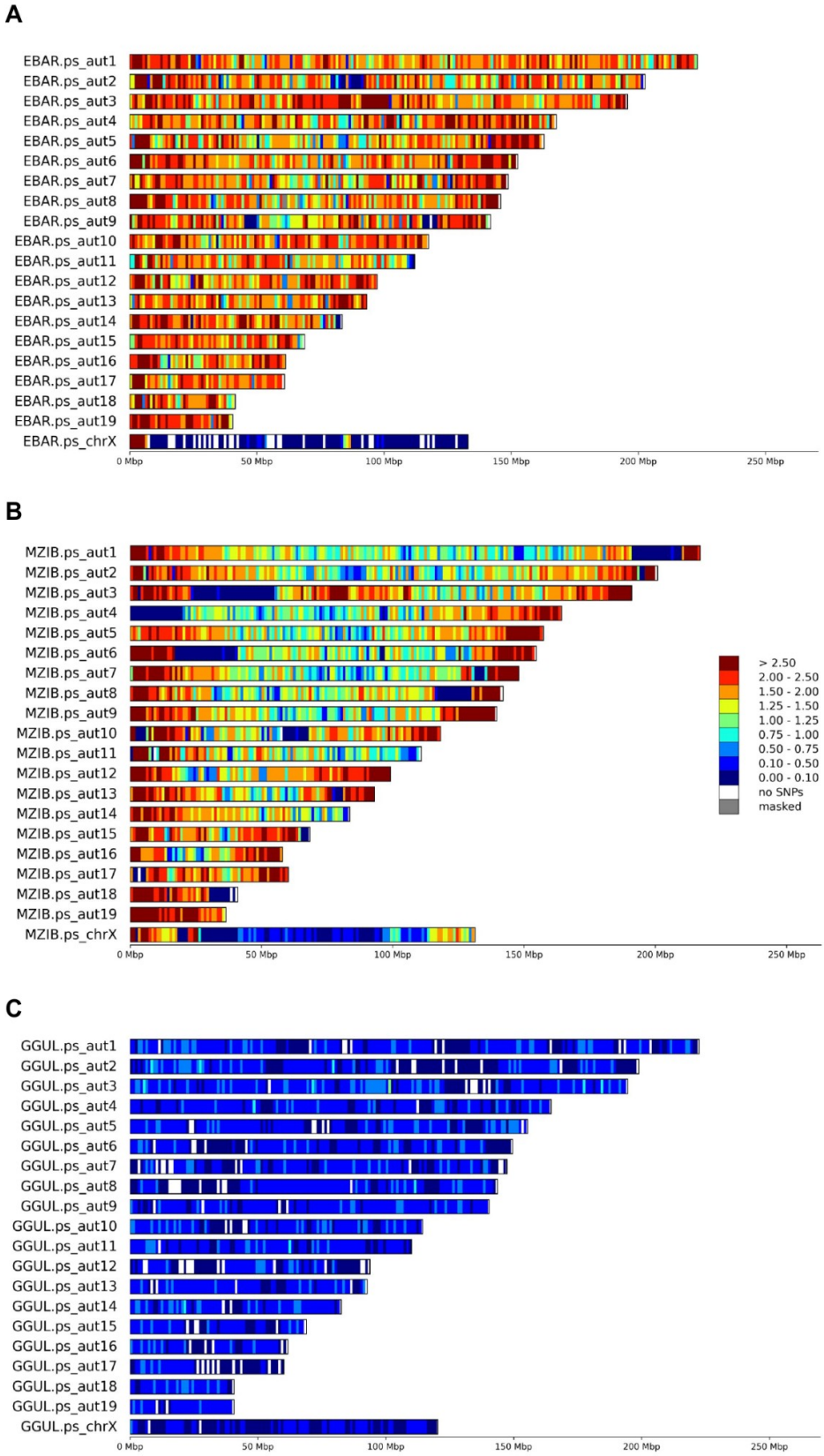

Figure 2. Heterozygosity density among pseudochromosomes for (A) tayra, (B) sable and (C) wolverine. Heterozygous SNPs were counted in stacking windows of 1 Mbp and scaled to SNPs per kbp. Tayra is a male individual and thus heterozygous SNP density is underestimated (due to only one X chromosome), while sable and wolverine are females and therefore likely representative of true SNP density.

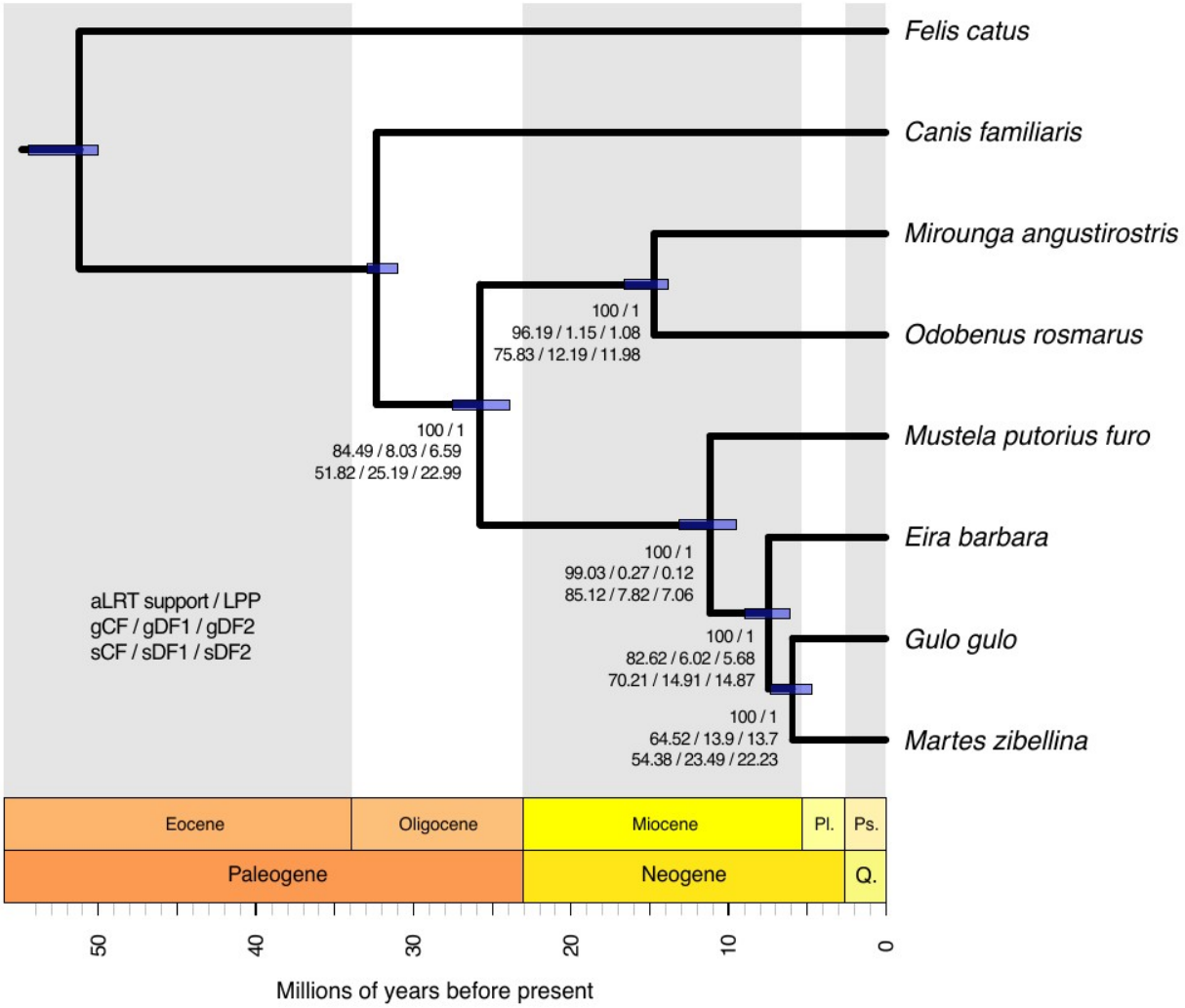

Figure 3. Phylogenetic tree and divergence times of Guloninae and five other carnivorans. The mean age of each node is shown, with 95% confidence intervals depicted as purple bars. The gene and site concordance (gCF, sCF) and discordance (gDF, sDF) factors are given, along with the branch support for both approximate likelihood-ratio test (aLRT) and local posterior probabilities (LPP).

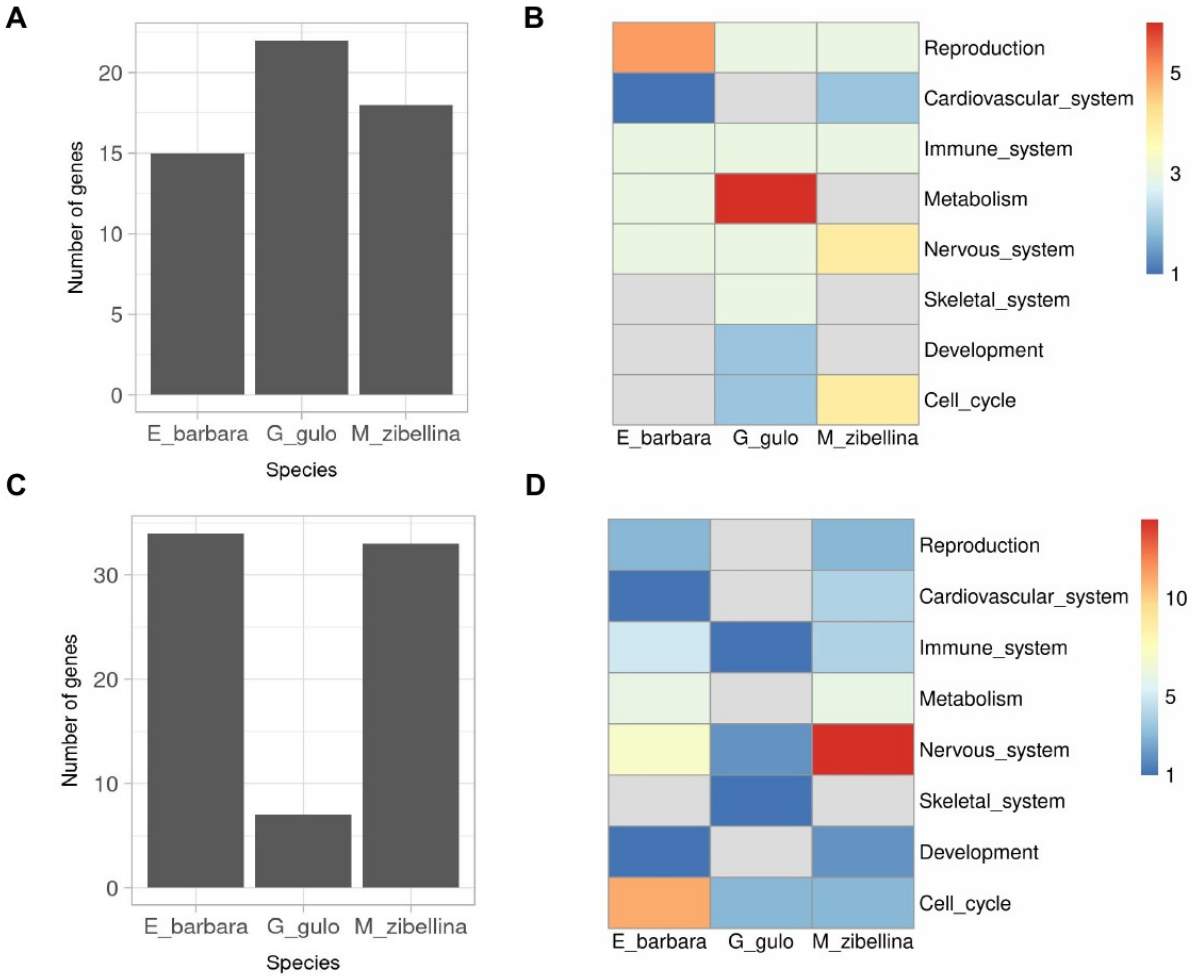

Figure 4. Number of candidate genes and their functional groups.
Genes identified from analyses of (A,B) positive selection on single genes (PSG), and (C,D) gene family expansions. Heatmap scale represents the number of genes. Heatmap cells in grey color indicate no observations for a given variable.

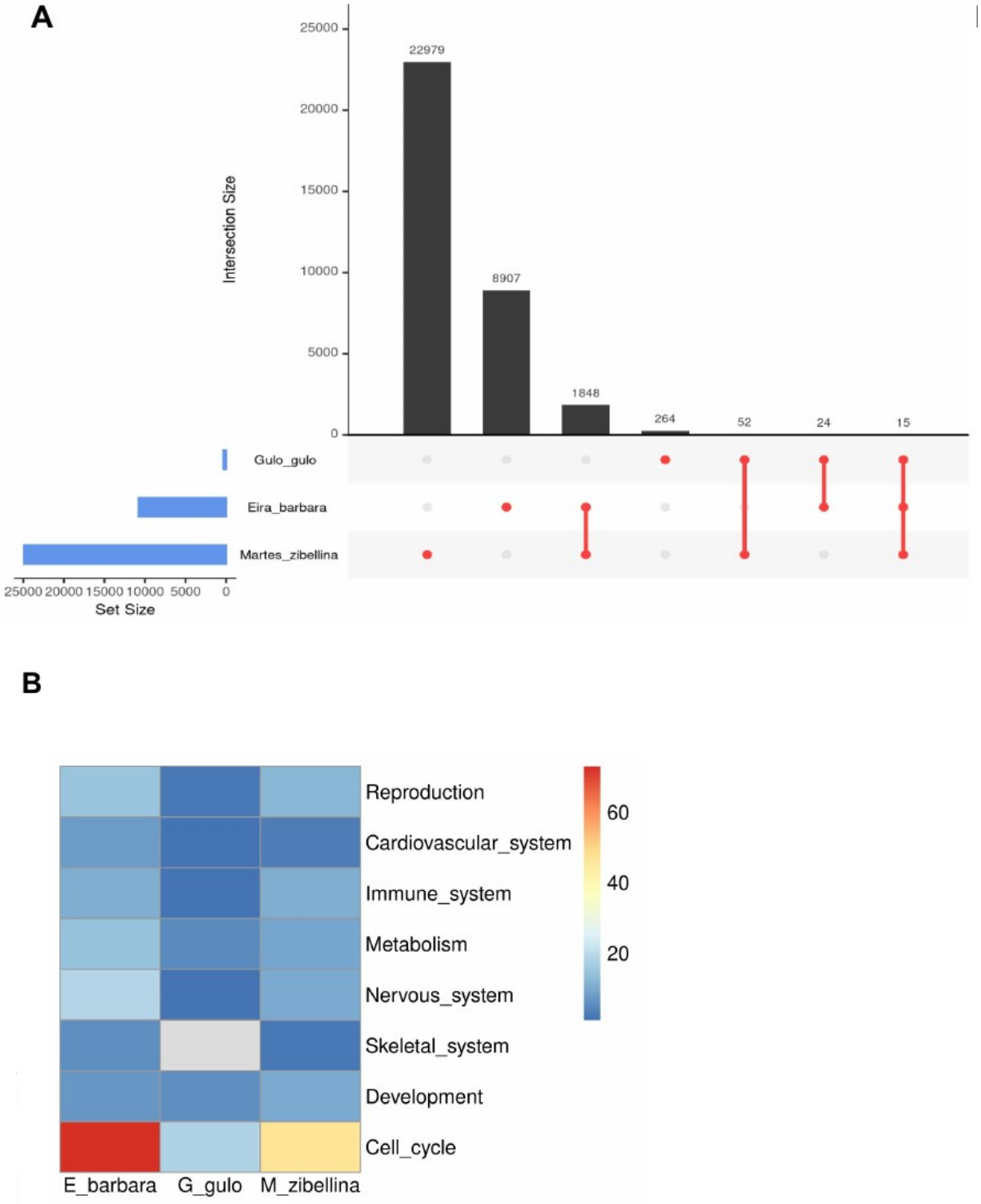

Figure 5. Structural variants detected in gulonine species.
A) Shared and species-specific structural variants detected in wolverine (*Gulo gulo*), tayra (*Eira barbara*), and sable (*Martes zibellina*), B) Functional groups of genes affected by species-specific structural variants in 3 gulonine species (SV types: DEL, DUP, INS). Heatmap scale represents the number of genes. Heatmap cells in grey color indicate no observations for a given variable.

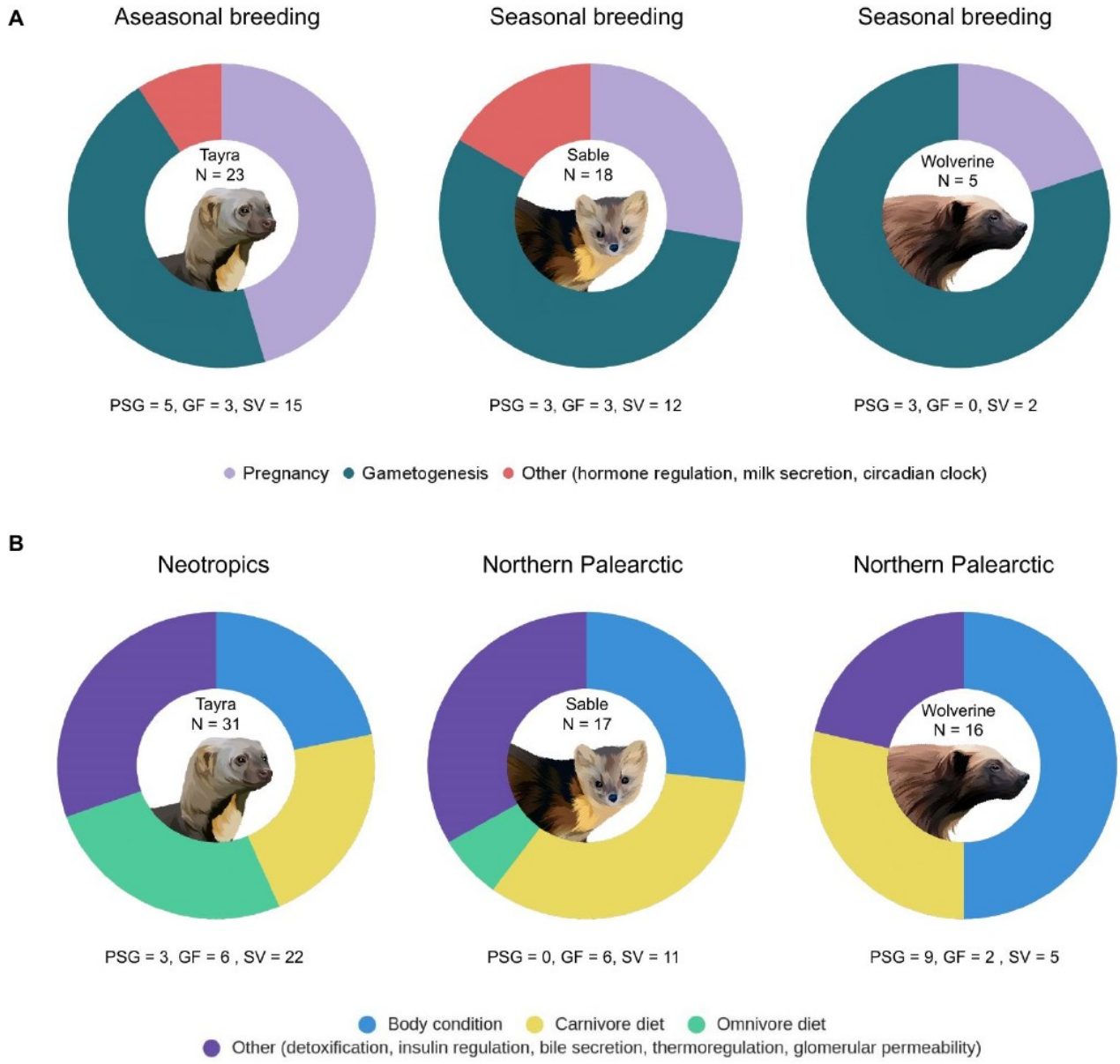

Figure 6. Summary of functional categories of (A) reproduction and (B) metabolism-related genes derived from analyses of positively selected genes (PSG), gene family expansion (GF) and structural variation (SV). N represents the total number of detected genes. Vector graphics of species are created based on royalty-free images (source: Shutterstock).