Journal of Heredity, 2025, **116**, 548–557
https://doi.org/10.1093/jhered/esaf001
Advance access publication 29 January 2025
Genome Resources

American Genetic Association

OXFORD

Genome Resources

Chromosome-length genome assembly of the stone marten (*Martes foina*, Mustelidae): A new view on one of the cornerstones in carnivore cytogenetics

Andrey Tomarovsky[1,2,*], Ruqayya Khan[3], Olga Dudchenko[3,4], Azamat Totikov[1,2], Natalia A. Serdyukova[1], David Weisz[3], Nadejda V. Vorobieva[1], Tatiana Bulyonkova[5], Alexei V. Abramov[6], Wenhui Nie[7], Jinhuan Wang[7], Svetlana A. Romanenko[1], Anastasiya A. Proskuryakova[1], Nikolay Cherkasov[8], Malcolm A. Ferguson-Smith[9], Fengtang Yang[10], Elena Balanovskaya[11], M. Thomas P. Gilbert[12,13], Alexander S. Graphodatsky[1], Erez Lieberman Aiden[3,4], Roger Powell[14], Klaus-Peter Koepfli[15], Polina L. Perelman[1] and Sergei Kliver[12]

[1]Department of Diversity and Evolution of Genomes, Institute of Molecular and Cellular Biology SB RAS, Novosibirsk, Russia,
[2]Department of Natural Sciences, Novosibirsk State University, Novosibirsk, Russia,
[3]Department of Molecular and Human Genetics, The Center for Genome Architecture, Baylor College of Medicine, Houston, TX, United States,
[4]The Center for Theoretical Biological Physics, Rice University, Houston, TX, United States,
[5]Youth Laboratory of Molecular Genetics, Yugra State University, Khanty-Mansiysk, Russia,
[6]Zoological Institute, Russian Academy of Sciences, Saint Petersburg, Russia,
[7]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China,
[8]Centre for Computational Biology, Peter the Great Saint Petersburg Polytechnic University, St. Petersburg, Russia,
[9]Department of Veterinary Medicine, Cambridge Resource Centre for Comparative Genomics, University of Cambridge, Cambridge, United Kingdom,
[10]School of Life Sciences and Medicine, Shandong University of Technology, Zibo, China,
[11]Laboratory of Human Population Genetics, Research Centre for Medical Genetics, Moscow, Russia,
[12]Center for Evolutionary Hologenomics, The Globe Institute, The University of Copenhagen, Copenhagen, Denmark,
[13]University Museum, NTNU, Trondheim, Norway,
[14]North Carolina State University, Raleigh, NC, United States,
[15]Smithsonian-Mason School of Conservation, George Mason University, Front Royal, VA, United States

[*]Corresponding author: Institute of Molecular and Cellular Biology SB RAS, 8/2 Acad. Lavrentiev Avenue, Novosibirsk 630090, Russia. Email: andrey.tomarovsky@gmail.com

Corresponding Editor: William Murphy

Abstract

The stone marten (*Martes foina*) is an important species for cytogenetic studies in the order Carnivora. ZooFISH probes created from its chromosomes provided a strong and clean signal in chromosome painting experiments and were valuable for studying the evolution of carnivoran genome architecture. The research revealed that the stone marten chromosome set is similar to the presumed ancestral karyotype of the Carnivora, which added an additional value for the species. Using linked-read and Hi-C sequencing, we generated a chromosome-length genome assembly of a male stone marten (Gansu province, China) from a primary cell line. The stone marten assembly had a length of 2.42 Gbp, scaffold N50 of 144 Mbp, and a 96.2% BUSCO completeness score. We identified 19 chromosomal scaffolds (2n = 38) and assigned them chromosome ids based on chromosome painting data. Annotation identified 20,087 protein-coding gene models, of which 18,283 were assigned common names. Comparison of the stone marten assembly with the cat, dog, and human genomes revealed several small syntenic blocks absent on the published painting maps. Finally, we assessed the heterozygosity and its distribution over the chromosomes. The detected low heterozygosity level (0.4 hetSNPs/kbp) and the presence of long runs of homozygosity require further research and a new evaluation of the conservation status of the stone marten in China. Combined with available carnivoran genomes in large-scale synteny analysis, the stone marten genome will highlight new features and events in carnivoran evolution, hidden from cytogenetic approaches.

Key words: centromere, genome assembly, *Martes foina*, stone marten, whole genome alignment

Received November 29, 2024; Accepted January 20, 2025

Introduction

Martes foina (the stone, or beech marten) is a member of the large family Mustelidae with ~68 known species, and the subfamily Guloninae, which includes the martens, fisher, tayra, and wolverine (Wilson and Reeder 2005). The stone marten diverged from other marten species about 2.5 to 3.5 million years ago and then spread from Asia to the forests and mountain areas of Europe (Koepfli et al. 2008; Law et al. 2018). This small nocturnal omnivore is an active predator with a tendency to become synanthropic. Chromosome painting and heterochromatin analyses showed that *M. foina* has a 2n = 38 karyotype (Liu et al. 1995; Nie et al. 2002; Beklemisheva et al. 2023). Flow cytometric studies revealed that the *M. foina* karyotype had remarkably well-resolved peaks, which resulted in clean probes with a high representation of each chromosome that produced a strong fluorescent signal. Using the *M. foina* 20-probe set (18 autosomes, X and Y) in a single fluorescent in situ hybridization experiment, it was possible to map the whole chromosome set of another species (Nie et al. 2002). So far, a total of 18 species from 9 families of the Carnivora have been studied using *M. foina* probes (Table 1) (A. S. Graphodatsky et al. 2002; Nie et al. 2002, 2012; Perelman et al. 2008, 2012; Beklemisheva et al. 2016, 2020; A. Graphodatsky et al. 2020). By mapping the chromosome rearrangements (mostly interchromosomal) onto the Carnivora phylogeny, it was shown that *M. foina* has a conserved chromosome set representative of the ancestral Carnivora karyotype (Nie et al. 2012). These facts allow us to label the stone marten a cornerstone of carnivore cytogenetics.

The original fibroblast cell line used in all the above experiments was established and described by Dr. Wenhui Nie's team in 1995. Biomaterial was sampled from a male marten captured in Gansu Province, China (Liu et al. 1995). After two decades of painting studies, we used this cell line to assemble the stone marten genome. Using a combination of linked and Hi-C reads, we generated a chromosome-length assembly, which we present here. It will provide a new point of view on such important species as the stone marten and bridge cytogenetic and bioinformatic studies of genome evolution within the order Carnivora.

Materials and methods

Samples and DNA extraction

Ultra-high molecular weight DNA (uHMW DNA) for genome sequencing was extracted from a primary fibroblast cell line of a stone marten using a phenol-chloroform protocol described below. The cell line was provided by the large-scale research facilities "Cryobank of cell cultures" Institute of Molecular and Cellular Biology (IMCB), Siberian Branch of the Russian Academy of Sciences (Novosibirsk, Russia). All experiments with the sample were evaluated and approved by the IMCB Ethical Committee (permit #01/21 issued on 26 January 2021).

Primary fibroblasts were cultured and dissociated using 0.25% trypsin/EDTA solution and resuspended in culturing media with 10% Fetal Bovine Serum to stop enzymatic reaction. After centrifugation at 300 g for 5 min, the supernatant

Table 1. Carnivora species were studied using stone marten chromosome painting probes (Zoo FISH experiments).

Species	2n	Family	Original reference[a]
Martes flavigula	40	Mustelidae	(Nie et al. 2002)
Meles meles	44	Mustelidae	(A. S. Graphodatsky et al. 2002; Nie et al. 2002)
Melogale moschata	38	Mustelidae	(Nie et al. 2002)
Neogale vison	30	Mustelidae	(Hameister et al. 1997; A. S. Graphodatsky et al. 2000; Nie et al. 2002)
Procyon lotor	38	Procyonidae	(Perelman et al. 2008,Nie et al. 2012)
Ailurus fulgens	36	Ailuridae	(Nie et al. 2002; Tian et al. 2002)
Conepatus leuconotus	46	Mephitidae	(Perelman et al. 2008)
Mephitis macroura	50	Mephitidae	(Perelman et al. 2008)
Mephitis mephitis	50	Mephitidae	(Perelman et al. 2008)
Spilogale gracilis	60	Mephitidae	(Perelman et al. 2008)
Odobenus rosmarus	32	Odobenidae	(Beklemisheva et al. 2016)[b]
Arctocephalus forsteri	36	Otariidae	(Beklemisheva et al. 2020)[b]
Callorhinus ursinus	36	Otariidae	(Beklemisheva et al. 2020)[b]
Phocarctos hookeri	36	Otariidae	(Beklemisheva et al. 2020)[b]
Erignathus barbatus	34	Phocidae	(Beklemisheva et al. 2020)[b]
Pusa sibirica	32	Phocidae	(Beklemisheva et al. 2016)[b]
Felis catus	38	Felidae	(Rettenberger et al. 1995; Wienberg et al. 1997; Nash et al. 1998, 2001; A. S. Graphodatsky et al. 2000; Yang et al. 2000; Nie et al. 2002; Perelman et al. 2008)
Genetta pardina	52	Viverridae	(Perelman et al. 2012)

[a]Most of the painting maps were accessed via the *Atlas of Mammalian Chromosomes* (A. Graphodatsky et al. 2020). *Martes foina* painting probes were also mapped onto chromosomes of species from other orders: *Manis javanica* (Pholidota) (Yang et al. 2006); *Neophocaena phocainoides asiaeorientalis* (Artiodactyla) (Nie et al. 2012).
[b]The chromosomes of the pinniped species were mapped by a select set of stone marten painting probes.

was removed and the cell pellet was rinsed in Dulbecco's Phosphate-Buffered Saline (DPBS) and pelleted by 5 min centrifugation at 300 g. The supernatant was removed and lysis buffer (0.05 M Tris-HCl, pH 7.5, 0.066 M EDTA, 0.1 M NaCl) was added to the cells which were then lysed for 3 h at 56 °C. Next, we followed a 3-step phenol-chloroform DNA extraction protocol (phenol twice, phenol: chloroform 1:1 (v/v)) with 20 min incubation at each step with gentle rocking, followed by centrifugation at 1,500 g in 15 ml tubes with Dow Corning high vacuum grease. An equal volume of isopropanol was added to the aqueous phase (collected from above the vacuum grease layer), followed by slow precipitation over 2 h at room temperature, a 70% ethanol rinse, and then followed by gentle placement of the precipitated DNA into TE buffer (10 mM Tris-HCl, 1 mM disodium EDTA, pH 8.0).

Sequencing of the stone marten genome

The extracted ultra-high molecular weight genomic DNA was shipped for sequencing to MedGenome (California, USA). 10X Genomics linked-read libraries were prepared and purified with the Chromium Genome Reagent Kit version 2 and then run in a Chromium Controller instrument with a microfluidic Genome Chip according to the manufacturer's recommendations (10X Genomics, CA, USA). Fragment sizes of the input DNA and prepared libraries were quantitated using a 2100 Bioanalyzer (Agilent Technologies, CA, USA). Libraries were sequenced with 150 bp paired-end reads on an Illumina HiSeq X Ten instrument (Illumina, CA, USA). In situ Hi-C library preparations were performed from the primary fibroblast cell line according to the modified protocol (Rao et al. 2014). The constructed library was PCR amplified and paired-end sequenced with 150 bp reads on an Illumina HiSeq X Ten instrument.

De novo genome assembly

Genome assembly was performed in two stages. First, the draft assembly was generated from raw linked reads using Supernova v2 (N. I. Weisenfeld et al. 2017). Raw Hi-C data were aligned, filtered and deduplicated using Juicer and the 3D-DNA pipeline was used to scaffold the draft (Durand et al. 2016; Dudchenko et al. 2017). Finally, the assembly was manually corrected using Juicebox Assembly Tools (Dudchenko et al. 2018). The completeness of the final assembly was evaluated using BUSCO v.5.5.0 (Manni et al. 2021) and mammalia_odb10.2021-02-19 (9,226 orthologs) single-copy orthologs databases.

Raw data filtering and QC

10X Genomics barcodes were trimmed from the linked reads using EMA v.0.6.2 (Shajii et al. 2017). Next, Illumina adapters were removed from the reads using Cookiecutter (Starostina et al. 2015b). Then, the small remaining fragments of the adapters were trimmed, and a quality filtering was performed using Trimmomatic v.0.36 (Bolger et al. 2014) with the following parameters "ILLUMINACLIP:TruSeq2-PE.fa:2:30:10:1 SLIDINGWINDOW:8:20 MINLEN:50." Finally, we downsampled the reads to 22× coverage using the BBtools package (Bushnell 2018). The quality control was performed on both raw and filtered data using FastQC v.0.11.9 (Andrews 2010). As part of our quality control of the assembly, we estimated the genome size directly from the reads using a k-mer-based approach. 23-mers were counted from trimmed and filtered reads using Jellyfish v2.2.10 (Marçais and Kingsford 2011b), and the corresponding distribution was visualized using KrATER v2.5 (Kliver 2017), followed by estimation of the genome size by GenomeScope2 (Ranallo-Benavidez et al. 2020).

Read mapping and variant calling

Filtered reads were aligned to the final genome assembly using BWA v.0.7.11 (Li and Durbin 2009). Samtools v1.15.1 was used to sort the reads and mark duplicates (Li et al. 2009). Per base genome coverage was calculated using Mosdepth v.0.3.0 (Pedersen and Quinlan 2018). Obtained coverage values were used to identify the pseudoautosomal region (PAR) on the X chromosome. The algorithm used for PAR coordinate identification is described in detail by Totikov et al. (2021).

Variant calling and demographic reconstruction

Variant calling was performed using Bcftools v1.15.1 (Danecek et al. 2021) with the parameters: "-d 250 -q 30 -Q 30 --adjust-MQ 50 -a AD,INFO/AD,ADF,INFO/ADF,ADR,INFO/ADR,DP,SP,SCR,INFO/SCR" for Bcftools mpileup and "-m -O u -v -f GQ,GP" for Bcftools call (Danecek et al. 2021). We also used the PAR coordinates to correctly set the ploidy of X chromosome regions (diploid for PAR, haploid for the rest) for variant calling as our reference individual is male. Low-quality genetic variants ("QUAL < 20.0 || (FORMAT/SP > 60.0 || FORMAT/DP < 5.0 || FORMAT/GQ < 20.0)") were removed using the Bcftools filter. Additionally, genetic variants at sites with more than 250% or less than 33% of the median whole genome coverage were filtered out using Bcftools intersect. The filtered genetic variants were split into single nucleotide polymorphisms (SNPs) and insertions/deletions (indels) using Bcftools. Finally, the number of homozygous and heterozygous SNPs was calculated in 1 Mbp sliding windows with a step size of 100 kbp. Corresponding SNP densities were visualized on a heatmap using MACE v 1.1.32 commit 9e62d2c (Kliver 2024).

To reconstruct the historical dynamics of effective population size (Ne) we used PSMC v0.6.5 (Li and Durbin 2011) with parameters "-N25 -t15 -r5 -p '4+25*2+4+6'." A consensus diploid sequence, required as input for PSMC, was generated using Samtools v0.1.19 and the vcfutils.pl script (Li et al. 2009) with the X chromosome excluded from the analysis. The coverage thresholds for minimum and maximum depths (parameters -d and -D in vcf2fq) were set to 33% and 250% of the median coverage of the genome, and variants outside this range were excluded. A generation time of 5 yr (Colella et al. 2021) and mutation rate of 2.2×10^{-9} (Kumar and Subramanian 2002; Colella et al. 2021) were used to scale the PSMC output.

Runs of homozygosity

Runs of homozygosity (ROH) search was based on the density distribution of heterozygous SNPs in sliding windows of 1 Mbp with a step size of 100 kbp. Filtering was based on the level of heterozygosity: windows were retained if heterozygosity did not exceed 0.05. For values between 0.05 and 0.1, windows were retained only if there were at least 5 consecutive windows. Windows with heterozygosity greater than 0.1 were discarded. Windows with low heterozygosity were merged into long regions if the distance between them was

less than half the window size. ROH coordinates were determined on all chromosomes except chrX. Visualization of ROHs was performed using MACE v 1.1.32.

Prediction and annotation of the protein-coding genes

We predicted gene models using the BRAKER pipeline v.3.0.8 (Brůna et al. 2024a), which involved a set of tools listed below. As input for the pipeline we used the public RNA-seq data generated from the sable, *Martes zibellina* (SRR13013010, SRR13013011, SRR8074161, SRR8074163, SRR8074165, SRR8074168, SRR8074169, SRR31089885, SRR31089886) and European pine marten, *Martes martes* (ERR11872609) (Liu et al. 2020; Xia et al. 2021; O'Brien et al. 2024; Tomarovsky et al. in prep), and the Mammalia_odb10 v2024-01-08 database and protein hints from Metazoa (OrthoDB v.11) generated (Kuznetsov et al. 2023) using orthodb-clades (https://github.com/tomasbruna/orthodb-clades). The GeneMark-ETP v1.02 (Brůna et al. 2024b) was used for the training and gene prediction based on both RNA-seq and protein data, while AUGUSTUS v.3.5.0 (Stanke et al. 2006, 2008) provided further gene predictions supported by extrinsic evidence. To assign gene names to the predicted models, we used eggNOG-mapper v.2.1.12 (Huerta-Cepas et al. 2019; Cantalapiedra et al. 2021) and the EggNOG database v5.0 (Mammalia subset).

Whole genome alignments and synteny analysis

Chromosome-length genome assemblies of domestic cat (*Felis catus*, Felis_catus_9.0), domestic dog (*Canis lupus familiaris*, CanFam3.1), and human (*Homo sapiens*, GRCh38.p13) were downloaded from NCBI's Genomes database (Lindblad-Toh et al. 2005; Schneider et al. 2017; Zhang and Schoenebeck 2020). For all three assemblies and the stone marten genome, we detected and softmasked tandem and interspersed repeats using RepeatMasker v4 (Smit et al. 2015), Tandem Repeat Finder (Benson 1999), and Windowmasker (Morgulis et al. 2006). Next, we performed a multiple whole genome alignment (mWGA) using the Progressive Cactus package (Armstrong et al. 2020). Raw synteny blocks were extracted from the mWGA using the halSynteny v2.2 (Krasheninnikova et al. 2020). Next, the raw blocks were filtered and visualized as described in Romanenko et al. (2023).

All the tools and databases used in the study are listed in Table 2.

Results and discussion

Assembly and annotation of the stone marten genome

We sequenced 564,762,322 linked reads (coverage 46.8×) and 297.4M Hi-C read pairs from the primary fibroblast cell line of a male stone marten. The generated assembly has a total length of 2.42 Gbp, a scaffold N50 of 144 Mbp (Supplementary Table ST4), and includes 19 chromosomal scaffolds (Fig. 1a) as expected from the known karyotype, 2n = 38 (Nie et al. 2002). The length of our assembly is slightly longer than that estimated from the reads (2.37 Gbp, Supplementary Fig. SF1). Assembly assessment using BUSCO (mammalia_odb10 database with 9226 BUSCOs) revealed 96.2% complete BUSCOs (95.6% single copy and 0.6% duplicated). Only 3.8% or 349 BUSCOs were fragmented or missing. These values are similar to those obtained for high-quality assemblies of other carnivoran species (Supplementary Table ST1). During the submission process one more genome assembly (mMarFoi2.1, GCA_964304585.1) of the stone marten was released to the scientific community, which was highlighted by one of the reviewers. Compared with our assembly it is slightly shorter (2.39 Gbp vs 2.42 Gbp), but its scaffold N50 is slightly higher (146 Mbp vs 144 Mbp). In terms of completeness, mMarFoi2.1 contains 8766 (95.1%) complete BUSCOs94.5% single-copy and 0.6% duplicated, which is lower than in our assembly (96.2%). The fraction of fragmented or missing BUSCOs (4.9% or 460 BUSCOs) also is higher compared to our assembly. mMarFoi2.1 assembly is based on the long but still very noisy Oxford Nanopore reads, corrected by the Illumina data (Martes Foina Genome Assembly mMarFoi2.1, n.d.). We suggest that imperfect correction of the sequencing errors is the main reason for the decreased BUSCO scores.

We identified and annotated 766 Mbp of repetitive sequences, corresponding to 31.7% of the assembly. Among these repeat sequences, most (18.89%) are long interspersed nuclear elements, followed by 4.26% of long terminal repeats, 2.70% of short interspersed nuclear elements and only 2.69% of DNA elements (Supplementary Table ST2). We used a coverage-based approach to localize the PAR. As expected, it is located on the end of the X chromosome (116.86 to 123.12 Mbp) and has the usual length of ~6 Mbp observed in Carnivora (Totikov et al. 2021; Yakupova et al. 2023). The prediction of the protein-coding genes resulted in 20,087 models with 30,708 transcripts (Supplementary Files SF1, SF2, SF3), and for 18,283 models we successfully assigned gene names (Supplementary Files SF4, SF5). The BUSCO-based evaluation of the predicted transcripts showed 97.8% complete BUSCOs, indicating high completeness of the annotation.

Synteny and centromeres

We aligned our assembly with the genomes of three species, previously used to create painting maps of the stone marten genome (Graphodatsky et al. 2020): the human, the domestic cat, and the domestic dog (Nie et al. 2002, 2012). Comparative chromosome painting (ZooFISH) has a resolution threshold of about 5 Mbp (Rens et al. 2006). Isolated homologous blocks of lower length have a high chance to be lost in ZooFISH experiments due to a weak signal, and sometimes even longer segments may be missed. We compared the synteny revealed from the whole genome alignment (Fig. 1c, b, and d) and the cat-stone marten painting map (Nie et al. 2002) and found no contradictions. However, we detected three blocks slightly longer than 5 Mbp, which are absent on the dog-stone marten map (Nie et al. 2012): CFAM 33 on MFOI 2, CFAM 27 on MFOI 9, and CFAM 26 on MFOI 12 (Fig. 1d, Supplementary File SF6, sheet "CFAM vs MFOI"). The human-stone marten painting map (Nie et al. 2012) lacks four such blocks, one of which (HSAP 1 on MFOI 12) nearly reaches 10 Mbp (Fig. 1b; Supplementary File SF6, sheet "HSAP vs MFOI").

Further analysis of the synteny results allowed us to assign chromosome numbers to the chromosomal scaffolds of the stone marten assembly (Supplementary Table ST3) and even approximately localize centromeres (Supplementary Table ST4; Fig. 1b, c, and d). The most useful resource was

Table 2. Tools used for assembly, annotation, and analysis of the stone marten genome.

Stage of analysis	Software/Database	Version	Reference
Raw data filtering and QC	FastQC	v.0.11.9	(Andrews 2010)
	EMA	0.6.2	(Shajii et al. 2017)
	Trimmomatic	v.0.36	(Bolger et al. 2014)
	Cookiecutter	1.0.0	(Starostina et al. 2015a)
	BBtools	38.96	(Bushnell 2014)
	KrATER	2.5	https://github.com/mahajrod/krater
	Jellyfish	2.2.10	(Marçais and Kingsford 2011a)
	GenomeScope2	2.0	(Ranallo-Benavidez et al. 2020)
Genome assembly	Supernova	v2	(Weisenfeld et al. 2017)
	Juicer	v.2019	(Durand et al. 2016)
	3D-DNA	v.2019	(Dudchenko et al. 2017)
	Juicebox Assembly Tools	v.2019	(Dudchenko et al. 2018)
Genome assembly QC	BUSCO	5.5.0	(Manni et al. 2021)
	OrthoDB[a]	odb10	(Kriventseva et al. 2019)
Repeat detection and masking	RepeatMasker	4.0.7	(Smit et al. 2013)
	Dfam[a]	3.7	(Storer et al. 2021)
	Windowmasker	2.9	(Morgulis et al. 2006)
	TRF	4.0.9	(Benson 1999)
	Bedtools	2.29	(Quinlan and Hall 2010)
Read mapping and variant calling	BWA	v.0.7.11	(Li and Durbin 2009)
	Mosdepth	v.0.3.0	(Pedersen and Quinlan 2018)
	Samtools	1.15.1	(Li et al. 2009)
	Bcftools	1.15.1	(Danecek et al. 2021)
WGA	LAST	981	(Frith and Kawaguchi 2015)
	ProgressiveCactus	1.0	(Armstrong et al. 2020)
	halSynteny	2.2	(Krasheninnikova et al. 2020)
Gene prediction	BRAKER	v3.0.8	(Brůna et al. 2024)
	AUGUSTUS	v.3.5.0	(Stanke et al. 2006, 2008)
	eggNOG-mapper	v2.1.12	(Cantalapiedra et al. 2021)
	eggNOG (Mammalia)[a]	v5.0	(Huerta-Cepas et al. 2019)
Visualization	MAVR	0.113	https://github.com/mahajrod/mavr
	MACE	1.1.32	https://github.com/mahajrod/mace

[a]Databases.

the dog-stone marten painting map due to the high number of the rearrangements and just as high a number in the chromosomes in the dog genome. As a result, nearly all the centromeres are located between synteny blocks homologous to different dog chromosomes and are easy to detect (Fig. 1D). We note that our coordinates of the centromeres are approximate as centromeres themselves are absent in our assembly. It will be possible to resolve centromeres and their position in the future by improving the assembly with ultralong Nanopore reads (Cheng et al. 2024).

Heterozygosity, RoHs, and demographic history

In total, we found 976,869 heterozygous SNPs, which corresponds to the mean value of 0.4 hetSNPs/kbp. Compared to other mustelids and two iconic felid species (cheetah and clouded leopard), the heterozygosity of our individual is on the lower side (Fig. 2c). The distribution of heterozygosity on chromosomes is uneven (Fig. 2a), with long ROH on several chromosomes (Fig. 2b). Chromosome MFOI 7 contains the longest RoH of 114.4 Mbp, but smaller ROHs are present on other chromosomes (MFOI 2, 3, 4, 10, 13, and others) as well. Loss of heterozygosity (LoH) is known to occur in cancer cells and various human cell lines (C.-Y. Li et al. 1992; Happle 1999; Närvä et al. 2010; Ryland et al. 2015). Even whole chromosomal arms were reported to become homozygous in human embryonic stem cell lines (Närvä et al. 2010). There are several known mechanisms of LoH (Tischfield 1997; Happle 1999), but only one of them, a mitotic recombination (MR), can result in long RoHs. However, the location of the most RoHs, including the two longest ones on the MFOI 7 and MFOI 10, are within but not at the end of chromosomes (Fig. 2b), making such a scenario unlikely. The homozygous middle and heterozygous ends of a chromosome can arise only from a double MR. Given the low frequency ($<10^{-4}$) of MR even under treatment stimulating recombination (Turner et al. 2003), such an event has a low probability in the earlier passages of a primary cell line. A small number of the possible divisions during the lifetime of a primary somatic cell line (fibroblasts in our case) will not produce enough cells for fixation of the RoH in the cell line.

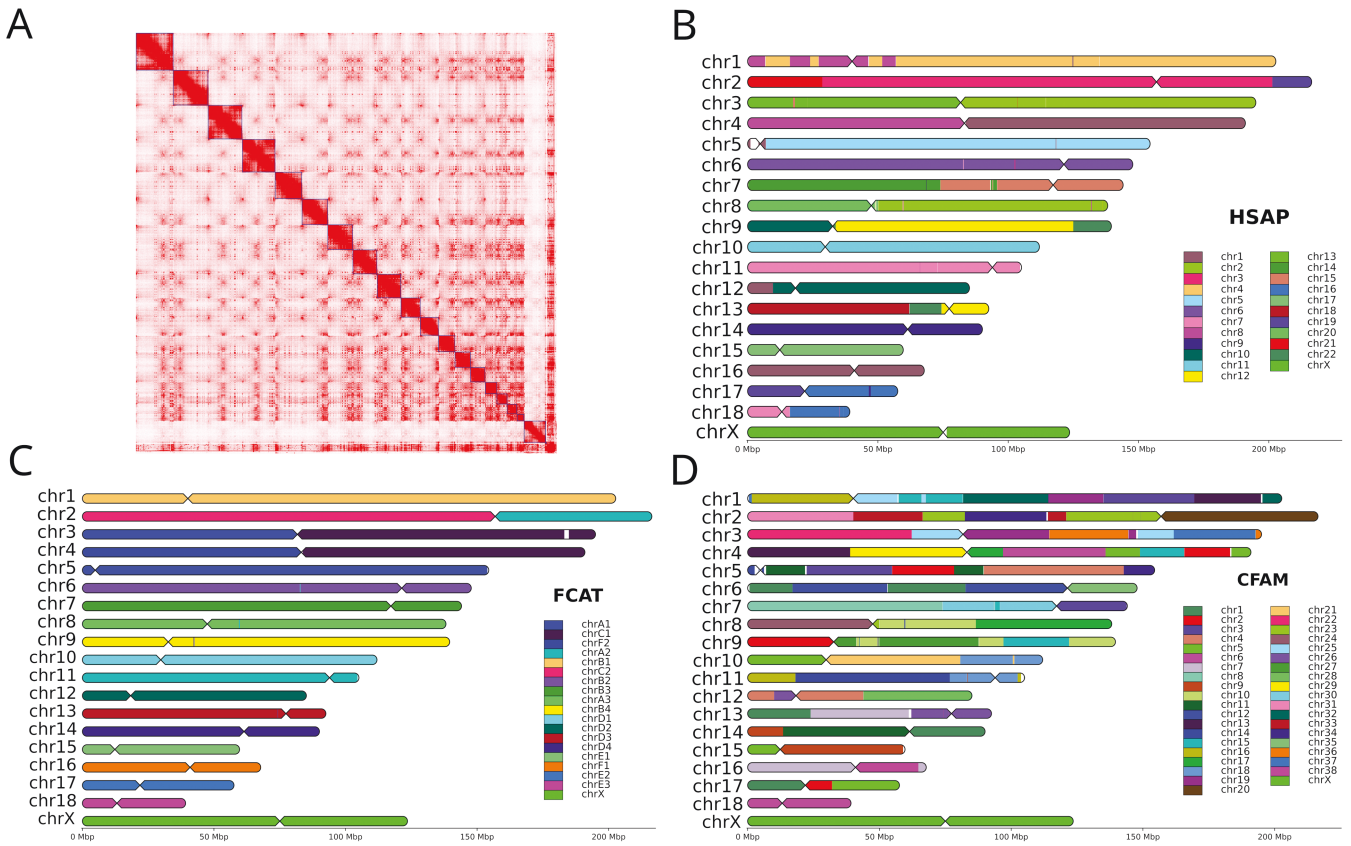

Fig. 1. Hi-C map of the stone marten genome assembly a) and synteny to the human (b, *Homo sapiens*, HSAP), domestic cat (c, *Felis catus*, FCAT), and domestic dog (d, *Canis familiaris*, CFAM) chromosomes. On figure a) autosomes (sorted by lengths) are shown first with X Chromosome following them.

Thus, the presence of the long RoHs most likely is not related to cell culturing. Long RoHs are known to be a consequence of recent inbreeding (Marras et al. 2015; Mastrangelo et al. 2016). We analyzed only a single individual, and our results do not describe diversity of a population. However, it is sufficient to raise questions about the conservation and genetic diversity status of the stone marten at least in Gansu province in China.

The trajectory of the demographic history (Fig. 2d) has two peaks of a similar height (at ~1.08 to 1.3 mya and ~230 to 280 kya) separated by a gap with ~30% decrease. Our analysis shows that the last peak was followed by a severe bottleneck that continued up to modern times, which reduced the effective population size tenfold. Such a recent decrease was previously detected for the American marten, *Martes americana* and the Pacific marten, *Martes caurina* (Colella et al. 2021).

Conclusions

The roots of this stone marten genome assembly are in the 1990s and the story of it doesn't end here. Even in this short report we connected this stone marten genome assembly with previous studies and used the benefits of additional cytogenetic data to provide novel insights about our assembly. The stone marten as well as some other species with recently assembled genomes was involved in the reconstruction of the carnivoran ancestral karyotype. Now it is possible to repeat the same analysis with the same set of species, but on the new level by using whole genome sequences instead of chromosome painting maps. The chromosome painting experiments highlighted mostly interchromosomal translocations, but only few inversions were detected because of methodological limitations. Whole genome alignments and consequent synteny analysis can overcome this gap. In addition to the generation of the assembly, we performed an initial and simple analysis of genome-wide diversity. Although our results are based on a single individual collected in the 1990s, the observed low heterozygosity and multiple long RoHs provide a justification for future conservation and population genomic studies of the stone marten not only in Gansu Province (the origin of the reference individual) but across the whole Eurasian range of the species. Our genome can be used as a reference for such research.

Our work is a reminder of the importance of long-term cryobanks (preserving living cell lines) for genomic studies (Mooney et al. 2023). The last decade has been the era of genomics, when not only genomes of new species are generated, but when the old assemblies are replaced by better ones or undergo continuous upgrades (Whibley et al. 2021). Current quality standards imply that the data for the improvement should be generated from the original sample individual. By using a cell line as a DNA source for sequencing, researchers can keep this door open.

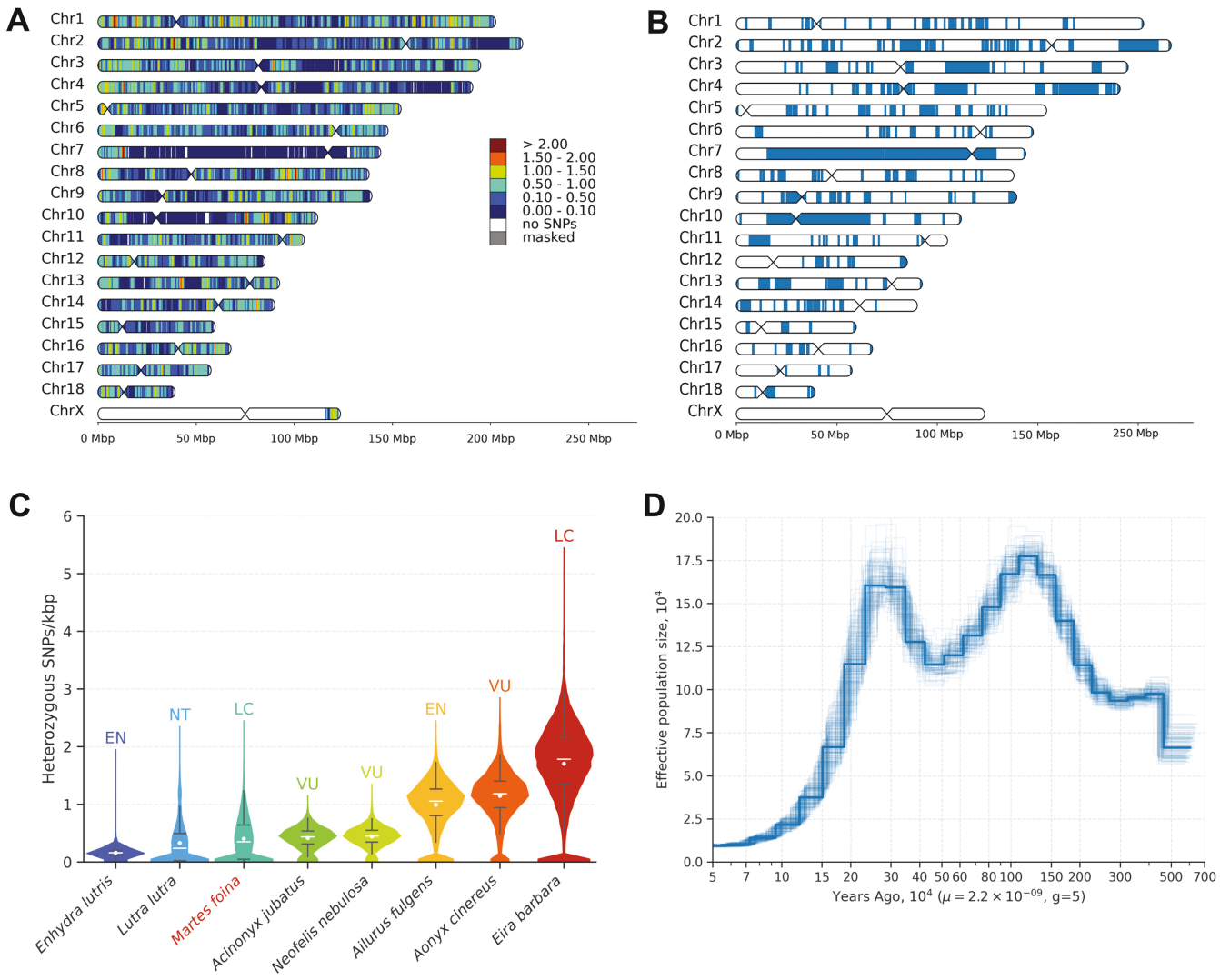

Fig. 2. Heterozygosity and demographic history reconstructed from the reference genome of the stone marten individual. a) density of the heterozygous SNPs (hetSNPs/kbp); b) RoHs; c) heterozygosity of the stone marten (red label) and selected carnivores of different conservation status (EN—endangered, NT—not threatened, LC—least concern, VU—vulnerable); d) PSMC demographic history of the stone marten with 100 bootstrap replicates.

Supplementary material

Supplementary material can be found at http://www.jhered.oxfordjournals.org/.

Acknowledgments

DNA Zoo acknowledges support from Illumina, IBM, and Pawsey Supercomputing Center.

Funding

S.K. and M.T.P.G. were funded by the Carlsbergfondet Research Infrastructure Grant CF22-0680 and the Danish National Research Foundation award DNRF143. F.Y. was supported by a China National Natural Science Foundation grant No. 32370689. N.C. was funded by the Ministry of Science and Higher Education of the Russian Federation as a part of the World-class Research Center program: Advanced Digital Technologies (contract No. 075-15-2022-311). The reported study was funded in by RSF grant 22-24-01076 and supported in part through computational resources of HPC facilities at the collaborative center «Bioinformatics» ICG SB RAS.

Conflict of interest statement

None declared.

Author contributions

Andrey Tomarovsky (Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing—original draft, Writing—review & editing), Ruqayya Khan (Formal analysis, Investigation, Methodology), Olga Dudchenko (Data curation, Formal analysis, Methodology, Software, Validation, Writing—review & editing), Azamat Totikov (Formal analysis, Methodology,

Writing—review & editing), Natalia Serdyukova (Formal analysis, Methodology), David Weisz (Data curation, Formal analysis, Methodology, Software), Nadezhda Vorobieva (Funding acquisition, Project administration), Tatiana Bulyonkova (Writing—original draft, Writing—review & editing), Alexei Abramov (Project administration, Writing—review & editing), Wenhui Nie (Resources, Writing—review & editing), Jinhuan Wang (Methodology, Resources), Svetlana Romanenko (Methodology, Resources, Writing—review & editing), Anastasiya Proskuryakova (Methodology, Resources, Writing—review & editing), Nikolay Cherkasov (Project administration, Resources, Software), Malcolm A. Ferguson-Smith (Resources, Writing—review & editing), Fengtang Yang (Methodology, Resources), Elena Balanovskaya (Methodology, Resources), Tom Gilbert (Funding acquisition, Resources), Alexander Graphodatsky (Funding acquisition, Methodology, Project administration, Resources, Supervision), Erez Aiden (Funding acquisition, Project administration, Resources, Supervision), Roger Powell (Funding acquisition, Project administration, Resources), Klaus-Peter Koepfli (Methodology, Project administration, Resources, Validation, Writing—original draft, Writing—review & editing), Polina Perelman (Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Validation, Visualization, Writing—original draft, Writing—review & editing), and Sergei Kliver (Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Writing—original draft, Writing—review & editing)

Data availability

Martes foina reads are available at NCBI SRA under SRR22412409 (linked reads) and SRR16086879 (Hi-C) accessions. The stone marten genome assembly is available from the DNAzoo website (https://www.dnazoo.org/assemblies/martes_foina) or NCBI under accession GCA_040938555.1.

References

Andrews S. FastQC: a quality control tool for high throughput sequence data. Cambridge, United Kingdom: Babraham Bioinformatics, Babraham Institute; 2010.

Armstrong J, Hickey G, Diekhans M, Fiddes IT, Novak AM, Deran A, Fang Q, Xie D, Feng S, Stiller J, et al. Progressive Cactus is a multiple-genome aligner for the thousand-genome era. Nature. 2020:587:246–251. doi:10.1038/s41586-020-2871-y

Beklemisheva VR, Lemskaya NA, Prokopov DY, Perelman PL, Romanenko SA, Proskuryakova AA, Serdyukova NA, Utkin YA, Nie W, Ferguson-Smith MA, et al. Maps of constitutive-heterochromatin distribution for four Martes species (Mustelidae, Carnivora, Mammalia) show the formative role of Macrosatellite repeats in interspecific variation of chromosome structure. Genes. 2023:14:Article 2. doi:10.3390/genes14020489

Beklemisheva VR, Perelman PL, Lemskaya NA, Kulemzina AI, Proskuryakova AA, Burkanov VN, Graphodatsky AS. The ancestral carnivore karyotype as substantiated by comparative chromosome painting of three pinnipeds, the walrus, the Steller sea lion and the Baikal seal (Pinnipedia, Carnivora). PLoS One. 2016:11:e0147647. doi:10.1371/journal.pone.0147647

Beklemisheva VR, Perelman PL, Lemskaya NA, Proskuryakova AA, Serdyukova NA, Burkanov VN, Gorshunov MB, Ryder O, Thompson M, Lento G, et al. Karyotype evolution in 10 pinniped species: variability of heterochromatin versus high conservatism of Euchromatin as revealed by comparative molecular cytogenetics. Genes. 2020:11:1485. doi:10.3390/genes11121485

Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999:27:573–580. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC148217/

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014:30:2114–2120. doi:10.1093/bioinformatics/btu170

Brůna T, Gabriel L, Hoff KJ. Navigating eukaryotic genome annotation pipelines: a route map to BRAKER, Galba, and TSEBRA (No. arXiv:2403.19416). arXiv. 2024a. doi:10.48550/arXiv.2403.19416

Brůna T, Lomsadze A, Borodovsky M. GeneMark-ETP significantly improves the accuracy of automatic annotation of large eukaryotic genomes. Genome Res. 2024b:34:757–768. doi:10.1101/gr.278373.123

Bushnell B. BBMap: a fast, accurate, splice-aware aligner; 2014. [accessed 29 Jan 2025]. https://escholarship.org/uc/item/1h3515gn.

Bushnell B. BBTools: a suite of fast, multithreaded bioinformatics tools designed for analysis of DNA and RNA sequence data. Joint Genome Institute; 2018.

Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol Biol Evol. 2021:38:5825–5829. doi:10.1093/molbev/msab293

Cheng H, Asri M, Lucas J, Koren S, Li H. Scalable telomere-to-telomere assembly for diploid and polyploid genomes with double graph. Nat Methods. 2024:21:967–970. doi:10.1038/s41592-024-02269-8

Colella JP, Lan T, Talbot SL, Lindqvist C, Cook JA. Whole-genome resequencing reveals persistence of forest-associated mammals in Late Pleistocene refugia along North America's North Pacific Coast. J Biogeogr. 2021:48:1153–1169. doi:10.1111/jbi.14068

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. Twelve years of SAMtools and BCFtools. GigaScience. 2021:10:giab008. doi:10.1093/gigascience/giab008

Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, Shamim MS, Machol I, Lander ES, Aiden AP, et al. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. Science. 2017:356:92–95. doi:10.1126/science.aal3327

Dudchenko O, Shamim MS, Batra SS, Durand NC, Musial NT, Mostofa R, Pham M, Glenn St Hilaire B, Yao W, Stamenova E, et al. The juicebox assembly tools module facilitates de novo assembly of mammalian genomes with chromosome-length scaffolds for under $1000 [Preprint]. Genomics. 2018. doi:10.1101/254797

Durand NC, Shamim MS, Machol I, Rao SSP, Huntley MH, Lander ES, Aiden EL. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. Cell Systems. 2016:3:95–98. doi:10.1016/j.cels.2016.07.002

Frith MC, Kawaguchi R. Split-alignment of genomes finds orthologies more accurately. Genome Biol. 2015:16:1–17.

Graphodatsky A, Perelman P, O'Brien SJ. Atlas of mammalian chromosomes. John Wiley & Sons, Incorporated; 2020.

Graphodatsky AS, Yang F, O'Brien PCM, Serdukova N, Milne BS, Trifonov V, Ferguson-Smith MA. A comparative chromosome map of the arctic fox, red fox and dog defined by chromosome painting and high resolution G-banding. Chromosome Res. 2000:8:253–263. doi:10.1023/A:1009217400140

Graphodatsky AS, Yang F, Perelman PL, O'Brien PCM, Serdukova NA, Milne BS, Biltueva LS, Fu B, Vorobieva NV, Kawada S-I, et al. Comparative molecular cytogenetic studies in the order Carnivora: mapping chromosomal rearrangements onto the phylogenetic tree. Cytogenet Genome Res. 2002:96:137–145. doi:10.1159/000063032

Hameister H, Klett C, Bruch J, Dixkens C, Vogel W, Christensen K. Zoo-FISH analysis: the American mink Mustela vison) closely

resembles the cat karyotype. Chromosome Res. 1997:5:5–11. doi:10.1023/A:1018433200553
Happle R. Loss of heterozygosity in human skin. J Am Acad Dermatol. 1999:41:143–164. doi:10.1016/S0190-9622(99)70042-3
Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. 2019:47:D309–D314. doi:10.1093/nar/gky1085
Kliver S. KrATER: K-mer analysis tool easy to run [Computer software]; 2017. [accessed 2025 Jan 29]. https://github.com/mahajrod/KrATER (Original work published 2016)
Kliver S. MACE v1.1.32 [Python]; 2024. [accessed 2025 Jan 29]. https://github.com/mahajrod/MACE.
Koepfli K-P, Deere KA, Slater GJ, Begg C, Begg K, Grassman L, Lucherini M, Veron G, Wayne RK. Multigene phylogeny of the Mustelidae: resolving relationships, tempo and biogeographic history of a mammalian adaptive radiation. BMC Biol. 2008:6:10. doi:10.1186/1741-7007-6-10
Krasheninnikova K, Diekhans M, Armstrong J, Dievskii A, Paten B, O'Brien S. halSynteny: a fast, easy-to-use conserved synteny block construction method for multiple whole-genome alignments. GigaScience. 2020:9:giaa047. doi:10.1093/gigascience/giaa047
Kriventseva EV, Kuznetsov D, Tegenfeldt F, Manni M, Dias R, Simão FA, Zdobnov EM. OrthoDB v10: Sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. Nucleic Acids Res. 2019:47:D807–D811. doi:10.1093/nar/gky1053
Kumar S, Subramanian S. Mutation rates in mammalian genomes. Proc Natl Acad Sci USA. 2002:99:803–808. doi:10.1073/pnas.022629899
Kuznetsov D, Tegenfeldt F, Manni M, Seppey M, Berkeley M, Kriventseva EV, Zdobnov EM. OrthoDB v11: annotation of orthologs in the widest sampling of organismal diversity. Nucleic Acids Res. 2023:51:D445–D451. doi:10.1093/nar/gkac998
Law CJ, Slater GJ, Mehta RS. Lineage diversity and size disparity in Musteloidea: testing patterns of adaptive radiation using molecular and fossil-based methods. Syst Biol. 2018:67:127–144. doi:10.1093/sysbio/syx047
Li C-Y, Yandell DW, Little JB. Molecular mechanisms of spontaneous and induced loss of heterozygosity in human cells in vitro. Somat Cell Mol Genet. 1992:18:77–87. doi:10.1007/BF01233450
Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009:25:1754–1760. doi:10.1093/bioinformatics/btp324
Li H, Durbin R. Inference of human population history from individual whole-genome sequences. Nature. 2011:475:493–496. doi:10.1038/nature10231
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and SAMtools. Bioinformatics. 2009:25:2078–2079. doi:10.1093/bioinformatics/btp352
Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, Clamp M, Chang JL, Kulbokas EJ, Zody MC, et al. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature. 2005:438:Article 7069. doi:10.1038/nature04338
Liu G, Zhao C, Xu D, Zhang H, Monakhov V, Shang S, Gao X, Sha W, Ma J, Zhang W, et al. First draft genome of the sable, *Martes zibellina*. Genome Biol Evol. 2020:12:59–65. doi:10.1093/gbe/evaa029
Liu R, Nie W, Chen Y. Chromosome study of stone marten (*Martes foina intermedia*). Zool Res. 1995:16:275–279.
Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol Biol Evol. 2021:38:4647–4654. doi:10.1093/molbev/msab199
Marçais G, Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics. 2011a:27:764–770. doi:10.1093/bioinformatics/btr011
Marçais G, Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics. 2011b:27:764–770. doi:10.1093/bioinformatics/btr011
Marras G, Gaspa G, Sorbolini S, Dimauro C, Ajmone-Marsan P, Valentini A, Williams JL, Macciotta NPP. Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. Anim Genet. 2015:46:110–121. doi:10.1111/age.12259
Martes foina genome assembly mMarFoi2.1. NCBI; n.d. [accessed 2024 Dec 26]. https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_964304585.1/.
Mastrangelo S, Tolone M, Gerlando RD, Fontanesi L, Sardina MT, Portolano B. Genomic inbreeding estimation in small populations: evaluation of runs of homozygosity in three local dairy cattle breeds. Animal. 2016:10:746–754. doi:10.1017/S1751731115002943
Mooney A, Ryder OA, Houck ML, Staerk J, Conde DA, Buckley YM. Maximizing the potential for living cell banks to contribute to global conservation priorities. Zoo Biol. 2023:42:697–708. doi:10.1002/zoo.21787
Morgulis A, Gertz EM, Schäffer AA, Agarwala R. WindowMasker: Window-based masker for sequenced genomes. Bioinformatics (Oxford, England) 2006:22:134–141. doi:10.1093/bioinformatics/bti774
Närvä E, Autio R, Rahkonen N, Kong L, Harrison N, Kitsberg D, Borghese L, Itskovitz-Eldor J, Rasool O, Dvorak P, et al. High-resolution DNA analysis of human embryonic stem cell lines reveals culture-induced copy number changes and loss of heterozygosity. Nat Biotechnol. 2010:28:371–377. doi:10.1038/nbt.1615
Nash WG, Menninger JC, Wienberg J, Padilla-Nash HM, O'Brien SJ. The pattern of phylogenomic evolution of the Canidae. Cytogenet Cell Genet. 2001:95:210–224. doi:10.1159/000059348
Nash WG, Wienberg J, Ferguson-Smith MA, Menninger JC, O'Brien SJ. Comparative genomics: Tracking chromosome evolution in the family Ursidae using reciprocal chromosome painting. Cytogenet Cell Genet. 1998:83:182–192. doi:10.1159/000015176
Nie W, Wang J, O'Brien PCM, Fu B, Ying T, Ferguson-Smith MA, Yang F. The genome phylogeny of domestic cat, red panda and five mustelid species revealed by comparative chromosome painting and G-banding. Chromosome Res. 2002:10:209–222. doi:10.1023/A:1015292005631
Nie W, Wang J, Su W, Wang D, Tanomtong A, Perelman PL, Graphodatsky AS, Yang F. Chromosomal rearrangements and karyotype evolution in carnivores revealed by chromosome painting. Heredity. 2012:108:17–27. doi:10.1038/hdy.2011.107
O'Brien D, Januszczak I; Natural History Museum Genome Acquisition Lab. The genome sequence of the European pine marten, *Martes martes* (Linnaeus, 1758). Wellcome Open Res. 2024:9:325. doi:10.12688/wellcomeopenres.22458.1
Pedersen BS, Quinlan AR. Mosdepth: quick coverage calculation for genomes and exomes. Bioinformatics. 2018:34:867–868. doi:10.1093/bioinformatics/btx699
Perelman PL, Beklemisheva VR, Yudkin DV, Petrina TN, Rozhnov VV, Nie W, Graphodatsky AS. Comparative chromosome painting in Carnivora and Pholidota. Cytogenet Genome Res. 2012:137:174–193. doi:10.1159/000341389
Perelman PL, Graphodatsky AS, Dragoo JW, Serdyukova NA, Stone G, Cavagna P, Menotti A, Nie W, O'Brien PCM, Wang J, et al. Chromosome painting shows that skunks (Mephitidae, Carnivora) have highly rearranged karyotypes. Chromosome Res. 2008:16:1215–1231. doi:10.1007/s10577-008-1270-2
Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010:26:841–842. doi:10.1093/bioinformatics/btq033

Ranallo-Benavidez TR, Jaron KS, Schatz MC. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nat Commun. 2020:11:Article 1. doi:10.1038/s41467-020-14998-3

Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, et al. A 3D map of the human genome at Kilobase resolution reveals principles of chromatin looping. Cell. 2014:159:1665–1680. doi:10.1016/j.cell.2014.11.021

Rens W, Fu B, O'Brien PCM, Ferguson-Smith M. Cross-species chromosome painting. Nat Protocols. 2006:1:783–790. doi:10.1038/nprot.2006.91

Rettenberger G, Klett C, Zechner U, Bruch J, Just W, Vogel W, Hameister H. ZOO-FISH analysis: cat and human karyotypes closely resemble the putative ancestral mammalian karyotype. Chromosome Res. 1995:3:479–486. doi:10.1007/BF00713962

Romanenko SA, Kliver SF, Serdyukova NA, Perelman PL, Trifonov VA, Seluanov A, Gorbunova V, Azpurua J, Pereira JC, Ferguson-Smith MA, et al. Integration of fluorescence in situ hybridization and chromosome-length genome assemblies revealed synteny map for guinea pig, naked mole-rat, and human. Sci Rep. 2023:13:Article 1. doi:10.1038/s41598-023-46595-x

Ryland GL, Doyle MA, Goode D, Boyle SE, Choong DYH, Rowley SM, Li J, Bowtell DD, Tothill RW, Campbell IG, et al; Australian Ovarian Cancer Study Group. Loss of heterozygosity: what is it good for? BMC Med Genomics. 2015:8:45. doi:10.1186/s12920-015-0123-z

Schneider VA, Graves-Lindsay T, Howe K, Bouk N, Chen H-C, Kitts PA, Murphy TD, Pruitt KD, Thibaud-Nissen F, Albracht D, et al. Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. Genome Res. 2017:27:849–864.

Shajii A, Numanagić I, Berger B. Latent variable model for aligning barcoded short-reads improves downstream analyses. bioRxiv. 2017:220236. doi:10.1101/220236

Smit A, Hubley R, Green P. 2015 RepeatMasker Open-4.0; 2013.

Smit A, Hubley R, Green P. RepeatMasker Open-4.0. 2013–2015; 2015.

Stanke M, Diekhans M, Baertsch R, Haussler D. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics. 2008:24:637–644. doi:10.1093/bioinformatics/btn013

Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. AUGUSTUS: Ab initio prediction of alternative transcripts. Nucleic Acids Res. 2006:34:W435–W439. doi:10.1093/nar/gkl200

Starostina E, Tamazian G, Dobrynin P, O'Brien S, Komissarov A. Cookiecutter: a tool for kmer-based read filtering and extraction. BioRxiv. 2015a:024679.

Starostina E, Tamazian G, Dobrynin P, O'Brien S, Komissarov A. Cookiecutter: a tool for kmer-based read filtering and extraction [Preprint]. Bioinformatics. 2015b. doi:10.1101/024679

Storer J, Hubley R, Rosen J, Wheeler TJ, Smit AF. The Dfam community resource of transposable element families, sequence models, and genome annotations. Mobile DNA. 2021:12:2. doi:10.1186/s13100-020-00230-y

Tian Y, Nie W-H, Wang J-H, Yang Y-F, Yang F-T. Comparative chromosome painting shows the red panda (*Ailurus fulgens)* has a highly conserved karyotype. Yi Chuan Xue Bao. 2002:29:124–127.

Tischfield JA. Loss of heterozygosity or: how I learned to stop worrying and love mitotic recombination. Am J Hum Genet. 1997:61:995–999. doi:10.1086/301617

Tomarovsky A, Totikov A, Beklemisheva V, Bulyonkova T, Perelman PL, Serdyukova N, Yakupova A, Abramov AV, Prokopov D, Krasheninnikova K. et al. Chromosome length genome assemblies of sable (*Martes zibellina*) and pine marten (*Martes martes*) combined with resequencing data elucidate hybridization between these species; in prep.

Totikov A, Tomarovsky A, Prokopov D, Yakupova A, Bulyonkova T, Derezanin L, Rasskazov D, Wolfsberger WW, Koepfli K-P, Oleksyk TK, et al. Chromosome-level genome assemblies expand capabilities of genomics for conservation biology. Genes. 2021:12:1336. doi:10.3390/genes12091336

Turner DR, Dreimanis M, Holt D, Firgaira FA, Morley AA. Mitotic recombination is an important mutational event following oxidative damage. Mutat Res. 2003:522:21–26. doi:10.1016/S0027-5107(02)00194-X

Weisenfeld NI, Kumar V, Shah P, Church DM, Jaffe DB. Direct determination of diploid genome sequences. Genome Res. 2017:27:757–767. doi:10.1101/gr.214874.116

Whibley A, Kelley JL, Narum SR. The changing face of genome assemblies: guidance on achieving high-quality reference genomes. Mol Ecol Resour. 2021:21:641–652. doi:10.1111/1755-0998.13312

Wienberg J, Stanyon R, Nash WG, O'Brien PCM, Yang F, O'Brien SJ, Ferguson-Smith MA. Conservation of human vs. Feline genome organization revealed by reciprocal chromosome painting. Cytogenet Cell Genet. 1997:77:211–217. doi:10.1159/000134579

Wilson DE, Reeder DM. Mammal species of the world: a taxonomic and geographic reference. Vol. 1. JHU Press; 2005.

Xia T, Zhang L, Sun G, Yang X, Zhao C, Zhang H. Insights into cold tolerance in sable (*Martes zibellina*) from the adaptive evolution of lipid metabolism. Mamm Biol. 2021:101:861–870. doi:10.1007/s42991-021-00135-0

Yakupova A, Tomarovsky A, Totikov A, Beklemisheva V, Logacheva M, Perelman PL, Komissarov A, Dobrynin P, Krasheninnikova K, Tamazian G, et al. Chromosome-length assembly of the Baikal seal (*Pusa sibirica*) genome reveals a historically large population prior to isolation in lake Baikal. Genes. 2023:14:619. doi:10.3390/genes14030619

Yang F, Graphodatsky AS, Li T, Fu B, Dobigny G, Wang J, Perelman PL, Serdukova NA, Su W, O'Brien PC, et al. Comparative genome maps of the pangolin, hedgehog, sloth, anteater and human revealed by cross-species chromosome painting: further insight into the ancestral karyotype and genome evolution of eutherian mammals. Chromosome Res. 2006:14:283–296. doi:10.1007/s10577-006-1045-6

Yang F, Milne BS, Schelling C, Dolf G, Schläpfer J, Switonski M, Ladon D, Pienkowska A, Bosma AA, Sargan DR, et al. Chromosome identification and assignment of DNA clones in the dog using a red fox and dog comparative map. Chromosome Res. 2000:8:93–100. doi:10.1023/A:1009251917072

Zhang W, Schoenebeck JJ. The ninth life of the cat reference genome, *Felis catus*. PLoS Genet. 2020:16:e1009045.