A Prelude to Conservation Genomics: First Chromosome-Level Genome Assembly of a Flying Squirrel (Pteromyini: *Pteromys volans*)

Authors:

Gerrit Wehrenberg[1] https://orcid.org/0000-0003-4828-3057,
Angelika Kiebler[1] https://orcid.org/0009-0001-2762-5057,
Carola Greve[2, 3] https://orcid.org/0000-0003-4993-1378,
Núria Beltrán-Sanz[4] https://orcid.org/0000-0001-8081-6234,
Alexander Ben Hamadou[2, 3] https://orcid.org/0009-0008-9826-902X,
René Meißner[5] https://orcid.org/0000-0003-2829-1607,
Sven Winter[5, 6, 7] https://orcid.org/0000-0002-1890-0977,
and Stefan Prost[1,8] https://orcid.org/0000-0002-6229-3596

Affiliations:

[1] Ecology and Genetics Research Unit, University of Oulu, P.O. Box 8000, 90014, Oulu, Finland

[2] Centre for Translational Biodiversity Genomics (LOEWE-TBG), Frankfurt am Main, Germany

[3] Senckenberg Forschungsinstitut und Naturmuseum, Senckenberganlage 25, 60325 Frankfurt am Main, Germany

[4] Centro di Ateneo Orto Botanico, University of Padua, Via Orto Botanico 15, 35123 Padua, Italy

[5] Research Institute of Wildlife Ecology, University of Veterinary Medicine, Vienna, Austria

[6] Faculty of Science and Technology, University of the Faroe Islands, Tórshavn, The Faroe Islands

[7] Senckenberg Biodiversity and Climate Research Centre, Frankfurt am Main, Germany

[8] Central Research Laboratories, Natural History Museum Vienna, Vienna, Austria

Correspondence to Gerrit Wehrenberg: gerrit.wehrenberg@oulu.fi

Keywords:

Pteromys volans, Siberian flying squirrel, Omni-C, PacBio HiFi, reference genome, Sciuridae, phylogeny

Abstract [English]

The Siberian flying squirrel (*Pteromys volans*) represents the only European Pteromyini species. Thus, it is biogeographically unique due to its specialised anatomy and biology as a volant rodent. As a result of habitat fragmentation and destruction, Siberian flying squirrels experience severe and ongoing population declines throughout most of their distribution. While considered *Least Concern* throughout their immense Eurasian distribution, this species is red-listed as *Vulnerable* and even *Critically Endangered* in parts of its range. More knowledge about the population structure and overall biology is needed to improve conservation efforts for this umbrella and flagship species of old-growth boreal forests. Here, we present the first chromosome-level genome assembly of any Pteromyini, represented by *P. volans* (Uoulu_pteVol_1.0). The final assembly has a total length of 2.85 Gbp in 19 chromosome-scale scaffolds with only minor differences in the chromosomal structure compared to other Sciuridae. All chromosome-scale scaffolds show indications for telomeres at both ends; the N50 value and busco as well as *k*-mer completeness scores are high with 157.39 Mbp and 97 – 99 %, respectively, indicating chromosome-level quality of the assembly. Based on whole-genome data from 17 rodent species, *P. volans* clusters according to known evolutionary relationships. Additionally, we present a new 16,511 bp long mitogenome unveiling differences from known conspecific mitogenomes. We propose the utility of the new reference genome for further research and development of conservation-applied genetic methods.

Introduction

The Siberian flying squirrel (*Pteromys volans* (LINNAEUS, 1758); Figure 1a) is the only native representative of its tribe (Pteromyini BRANDT, 1855) in Europe, with its characteristic patagium enabling a highly arboreal lifestyle, gliding between trees and largely avoiding the ground (Thorington et al., 2012; Koprowski et al., 2016). While being nocturnal, *P. volans* is active throughout the boreal seasons and therefore exposed to dramatic environmental and nutritional changes, adapting its activity and behaviour accordingly (Hokkanen et al., 1977; Törmälä et al., 1980; Hanski, 1998; Hanski et al., 2000). After the Eurasian red squirrel (*Sciurus vulgaris* LINNAEUS, 1758), *P. volans* has the second-largest continuous distribution range of any Sciuromorpha BRANDT, 1855 ranging from Northeast Europe throughout the Eurasian boreal forests until the East Asian Pacific coasts including Manchuria, the Korean Peninsula and the islands of Sakhalin and Hokkaidō southwards limited by arid regions of Central Asia (Thorington et al., 2012; Koprowski et al., 2016). The highest phylogenetic and (mitochondrial as well as nuclear) genetic variability can be found among the East Asian populations (Oshida et al., 2005; Lee et al., 2008; Ito Dos Santos et al., 2024), reflected in comprising all four described and currently accepted subspecies (Thorington et al., 2012; Holden-Musser et al., 2016; Hopkins, 2016; Koprowski et al., 2016). In large portions of its Eurasian range, including the last remaining European populations, it is represented by the nominate subspecies *P. v. volans* (Lee et al., 2008; Thorington et al., 2012; Koprowski et al., 2016; Ito Dos Santos et al., 2024).

While its conservation status is globally considered as *Least Concern* by the International Union for Conservation of Nature and Natural Resources (IUCN), populations experience declining trends throughout most of their distribution (Hokkanen et al., 1982; Shar et al., 2016; Zagorodniuk, 2022). Lately, this species has been listed as *Near Threatened* throughout its European distribution (Amori, 2024). Nationally, they are red-listed as *Vulnerable* in Finland (Liukko et al., 2019) and as *Critically Endangered* in Estonia (EELIS, 2025). In fact, the species went extinct in several European countries over the last centuries. Recently, it disappeared in Latvia and Lithuania (Pilāts, 2010, 2023), while a minor population was rediscovered in Northeast Belarus (Abramchuk, 2021). They are legally protected

within the European Union (EU) through the Habitats Directive (*Council Directive 92/43/EEC on the Conservation of natural habitats and of wild fauna and flora*) under *Annexes II* and *IV* (European Commission, 2021). On the contrary, it is not continuously protected throughout most of its distribution in Russia (Kurhinen et al., 2011), which represents the majority of its range. Given its relatively large population size, Finland holds special conservation responsibilities within the EU. Regardless of its protection status, its populations are still decreasing mostly due to habitat loss and fragmentation through extensive forestry (Lampila et al., 2009). Additionally, recent changes in Finnish conservation and forestry legislations made the execution of protection of this native rodent ineffective (Santangeli et al., 2013; Wistbacka et al., 2018). Even though Northeast Europe is much less populated by humans compared to the rest of the subcontinent, it is nevertheless highly exploited and thus causes the loss of old-growth mixed forests. The foliage of conifers and cavities in large tree trunks provide cover all year round, while the flying squirrel largely feeds on leaves and catkins of birch and aspen during the warmer season (Hanski et al., 2000; Thorington et al., 2012; Koprowski et al., 2016; Selonen and Wistbacka, 2016). Such dependency on old-growth forests (Mönkkönen et al., 1997; Hanski, 1998) and its charismatic, ecological and phylogenetic uniqueness in Europe makes *P. volans* also an umbrella and flagship species for habitat conservation of intact boreal forests (Hurme et al., 2008; Selonen and Mäkeläinen, 2017). Despite the current conservation focus, the ecology and intraspecific phylogeography of the Siberian flying squirrel is widely unknown.

High-contiguity genome assemblies serve as invaluable references for uncovering such unknown aspects of a species' biology and advancing applied conservation genetics (Brandies et al., 2019; Paez et al., 2022). Those assemblies not only enhance our understanding of population structures but also enable the identification of informative genetic markers, such as single-nucleotide polymorphisms (SNPs) and microsatellites. With their high contiguity, these reference genomes lay the foundation for developing genetic tools with broad applications in conservation efforts as needed for the Siberian flying squirrel.

While a draft genome assembly of *Glaucomys volans* (LINNAEUS, 1758), native to eastern and southern North America (Koprowski et al., 2016), was published (Wolf et al., 2022), our study provides the first chromosome-level assembly of any Pteromyini.

Only such high-contiguity assemblies enable comprehensive analysis of chromosomal structures, including improved SNP localisation, which facilitates linkage disequilibrium mapping and the study of functional diversity in genes, as well as the identification of runs of homozygosity and repeat regions. Here, we present the first high-contiguity, chromosome-level *de novo* genome assembly of *Pteromys volans*.

Material and Methods

Sample, DNA extraction and sequencing library preparations

High molecular weight DNA was extracted from muscle tissue taken from a whole-frozen specimen. The specimen died of natural causes and was collected post-mortem at 62.917950°N, 27.631830°E (WGS84) near Kuopio, Finland at the 26th March 2020. It was morphologically assigned as a male of the nominate subspecies of *Pteromys volans volans*. The specimen is stored at -20°C in the collection of the Zoological Museum of the University of Oulu (ZMUO; Collection Management System Kotka/FinBIF: http://id.zmuo.oulu.fi/OV.36069; GBIF: https://www.gbif.org/occurrence/4977508301 (Finbif, 2024)).

We used the *NEB Monarch® HMW DNA Extraction Kit for Cells & Blood* (NEB, USA) for the DNA extraction following the manufacturers protocol. Next, we measured DNA concentration and length utilising the *Qubit dsDNA BR Assay* kit at a Qubit Fluorometer 2.0 (Thermo Fisher Scientific), gel electrophoresis and the *Agilent High Sensitivity DNA* Kit for the *Agilent 2100 Bioanalyzer* (Agilent Technologies, USA). Two *SMRTbell* libraries were prepared according to the instructions of the *SMRTbell Express Prep Kit* v3.0 and sequenced on a *PacBio Revio* system (PacBio, Menlo Park, USA). Next, we generated a proximity ligation sequencing library from the same individual using the *Dovetail Omni-C* Kit (Cantata Bio, USA) following the manufacturers protocol, which was sequenced by *Novogene* in Munich on a *NovaSeq X* platform (Illumina, Inc., San Diego, CA, USA).

Raw data processing, assembly and scaffolding

First, we extracted the HiFi reads from the BAM file and converted them to the FASTQ format using *BamTools* v2.5.2 (Barnett et al., 2011) with the *-tag "rq":">=0.99"*

option and *BEDtools* v2.31.1 (Quinlan and Hall, 2010) with the *bamToFastq* option. We then assembled baseline contig-level genome assemblies using the HiFi reads and (1) the *Hifiasm* v0.19.6-r597 assembler (Cheng et al., 2021) and (2) the *Flye* v2.9.4 assembler (Kolmogorov et al., 2019). Based on the *BUSCO* v5.4.7 (Manni et al., 2021) scores and assembly continuity statistics (calculated using *abyss-fac*, (Simpson et al., 2009)) we selected the *Hifiasm* assembly for scaffolding using the Omni-C data. We ran *BUSCO* v5.4.7 in *euk_genome_met* mode and with the gene predictor *metaeuk* (Levy Karin et al., 2020) and the lineage dataset *mammalia_odb10* (which included 24 genomes of 9,226 buscos).

We used Dovetail Genomics's library QC pipeline (https://omni-c.readthedocs.io/en/latest/index.html) to assess the quality of the Omni-C library. For the Omni-C scaffolding, we applied the Vertebrate Genome Project (VGP) Hi-C scaffolding pipeline (Rhie et al., 2021). In brief, we first filtered the Omni-C reads using the filter_five_end.pl *perl* script (https://github.com/ignacio3437/HiC_mapping_pipeline/tree/master). Next, the reads were merged and PCR duplicates were removed using *Sambamba* v1.0.1 (Tarasov et al., 2015). We used *YaHS* v1.2 (Zhou et al., 2023) for the Omni-C scaffolding and displayed the contact map using *JuicerTools* v1.22.01 (https://github.com/aidenlab/juicertools). Next, we carried out gapclosing using *TGS-GapCloser* v2.0.0 (Xu et al., 2020). We then assessed the scaffolded assembly quality using *BUSCO* v5.4.7 and abyss-fac. We then performed another round of Omni-C scaffolding using the VGP Hi-C scaffolding pipeline and gapclosing using *TGS-GapCloser* v2.0.0, followed by another assembly quality assessment (as described above). Finally, we visualised the generated Omni-C contact map and manually curated the scaffolding using *Juicebox* v2.17.00 (Dudchenko et al., 2018).

Quality assessment, Repeat and Gene Annotation

The final assembly was assessed with *QUAST* v5.2.0 (Gurevich et al., 2013), *compleasm* v0.2.6, which was shown to be more accurate than *BUSCO* (Huang and Li, 2023), and *BUSCO* v5.4.7, to be able to compare with previous runs. The latter two each utilising the lineage dataset *mammalia_odb10*. In addition, we also investigated the *k*-mer completeness (%; k = 21), the consensus quality value (QV; Phred scale), and the *k*-mer multiplicity using *Meryl* v1.4.1 and *Merqury* v1.3 (Rhie et

al., 2020). The *k*-mer multiplicity was also esteimented with *GenomeScope* v2.0 (Ranallo-Benavidez et al., 2020). We further estimated the genome-level and scaffold-level coverage using *Qualimap* v2.3 (Okonechnikov et al., 2016) and assessed potential contamination using *BlobTools* v1.1 (Laetsch et al., 2017; Laetsch and Blaxter, 2017) and visualised in *BlobPlots* based on protein sequences found in the database 'Swiss-Prot' (release: 04-2024; The UniProt Consortium, 2019) by *diamond* v2.1.6 (Buchfink et al., 2021). To do so, we mapped the HiFi reads onto the Omni-C scaffolded assembly using *minimap2* v2.24 (Li, 2018, 2021). We further identified mitochondrial contigs in the assembly by blasting our assembled mitochondrial genome (outlined in the next section) against the final genome assembly using *NCBI BLAST* v2.16.0 (Boratyn et al., 2013) and removed them from the assembly. The SRY gene (XM_047537220.1) extracted from an eastern grey squirrel genome assembly (*Sciurus carolinensis* GMELIN, 1788; GCA_902686445.2 (Mead et al., 2020a)) was aligned to the novel genome assembly using blastn (*NCBI BLAST* v2.16.0) to determine its genomic location.

To annotate repeats in the genome assembly, we first used *RepeatModeler* v2.0.5 (Flynn et al., 2020) with the search engine *rmblast* v2.14.1+ and *RepeatMasker* v4.1.6 (Smit et al., 2023) with the *Repbase* 'rodentia' database (Jurka et al., 2005) and the additional options of: *-nolow* to not hard-mask low complexity DNA or simple repeats and *-gccalc* to calculate the GC content of each contig/scaffold individually. Next, we applied *RepeatMasker* v4.1.6 to the masked genome but with the options: *-xsmall* to change the masking mode from hard to soft-masked, *-noint* to only hard-masks low complex/simple repeats, excluding intronic repeats, and *-gccalc.*

Genes in the masked assembly were predicted based on homology with Gene Model Mapper *GeMoMa* v.1.9 (Keilwagen et al., 2016, 2018) using the following publicly available assemblies at NCBI GenBank: human (*Homo sapiens* LINNAEUS, 1758; GCA_000001405.29), house mouse (*Mus musculus* LINNAEUS, 1758; GCA_000001635.9), eastern grey squirrel (*Sciurus carolinensis*; GCA_902686445.2 (Mead et al., 2020a)), thirteen-lined ground squirrel (*Ictidomys tridecemlineatus* (MITCHILL, 1821); GCA_016881025.1 (Fu et al., 2021)), Arctic ground squirrel (*Urocitellus parryii* RICHARDSON, 1825; GCA_003426925.1), Alpine marmot (*Marmota marmota marmota* (LINNAEUS, 1758); GCA_001458135.2), and groundhog or woodchuck (*Marmota monax* (LINNAEUS, 1758); GCA_021218885.2 (Clarke and

Bader, 2024)). The busco completeness of the transcript annotation was assessed with *BUSCO* v5.4.7 utilising the lineage dataset *mammalia_odb10*. Functional annotation of the predicted proteins by *GeMoMa* was conducted by using *diamond* (*blastp* mode) v2.1.6 (Buchfink et al., 2015) with an e-value significance cutoff of ≤ 10^{-6} against the 'Swiss-Prot' database (release: 04-2024; The UniProt Consortium, 2019). Furthermore, we annotated gene ontology (GO) terms, domains, and motifs using *InterProScan* v5.73-104.0 (Quevillon et al., 2005; Jones et al., 2014). The gene annotation results were combined utilising modified functional annotation scripts from *MAKER* v3.01.03 (Cantarel et al., 2008; https://github.com/schellt/maker-functional). Gene annotation files are available on Dryad (https://doi.org/10.5061/dryad.3xsj3txth).

Telomere regions were identified using *seqtk* with the option *telo* (https://github.com/lh3/seqtk), and *quarTeT* v1.2.5 (Lin et al., 2023) with the options *-m 10* to *100*.

Lastly, we used the *circlize R* package v0.4.16 (Gu et al., 2014) to perform a six-track circos plot for the largest 19 scaffolds. As an input, we used (A) the telomeres length generated using -m 10, (B) the gene content clustered in 500 kbp windows, (C; D) the repetitive elements (simple and complex repeats) were grouped according to their type (the same TE type found sequentially) and plotted according to their total relative length, and (E) GC content (%; compare Figure 2a).

Mitochondrial genome assembly

We used *MitoHifi* v3.2.2 (Uliano-Silva et al., 2023) to reconstruct and the *Flye* v2.9.4 polisher (Kolmogorov et al., 2019) to polish the mitochondrial genome.

Protein-coding genes, rRNAs and tRNAs were annotated and visualised with *Chlorobox* online software tools (https://chlorobox.mpimp-golm.mpg.de/), *GeSeq* (Tillich et al., 2017), and *OGDraw* (Greiner et al., 2019). For *GeSeq*, a circular mitochondrial sequence source was selected, keeping only the best annotation. Annotation via "*BLAT search*" (standard settings) with NCBI GenBank reference MW722951.1 (*P. volans*) and third-party references from five Sciuridae FISCHER DE WALDHEIM, 1817 individuals: southern flying squirrel or assapan (*Glaucomys volans*; NC_050026.1); particolored flying squirrel (*Hylopetes alboniger* (HODGSON, 1836); NC_031847.1), red and white giant flying squirrel (*Petaurista alborufus* (MILNE-

EDWARDS, 1870); NC_023922.1), *P. volans* (NC_019612.1 (Ryu et al., 2013)), and Eurasian red squirrel (*Sciurus vulgaris*; NC_002369.1 (Reyes et al., 2000)). GenBank annotation was manually double-checked, and redundant replicates were removed (keeping the version with best alignment scores). For *OGDraw*, a mitochondrial sequence source was selected to visualise the generated GenBank file.

Synteny and phylogenetic analyses

Next, we assessed genome synteny to the chromosome-level reference genomes for the Eurasian red squirrel (*Sciurus vulgaris*; GenBank: GCA_902686455.2 (Mead et al., 2020b)), the eastern grey squirrel (*Sciurus carolinensis*; GenBank: GCA_902686445.2 (Mead et al., 2020a)), and the European ground squirrel or European souslik (*Spermophilus citellus* LINNAEUS, 1766; GenBank: GCA_964194105.1) using JupiterPlot v1.1 (https://github.com/JustinChu/JupiterPlot) using options: *ng = 101-110*, *m = 10000000*.

We then generated a phylogenetic tree for all available Sciuridae genomes and two Muridae species as an outgroup on the NCBI GenBank: house mouse (*Mus musculus*; GCA_000001635.9), black rat (*Rattus rattus* (LINNAEUS, 1758); GCF_011064425.1 (Rane et al., 2020)), southern flying squirrel or assapan (*Glaucomys volans*; GCA_020662805.1 (Wolf et al., 2022)), Eurasian red squirrel (*Sciurus vulgaris*; GCA_902686455.2 (Mead et al., 2020b), eastern grey squirrel (*Sciurus carolinensis*; GCA_902686445.2 (Mead et al., 2020a), eastern fox squirrel (*Sciurus niger* LINNAEUS, 1758; GCA_020740815.1 (Kang et al., 2021)), unstriped ground squirrel (*Xerus rutilus* (CRETZSCHMAR, 1828); GCA_028644305.1), Cape ground squirrel (*Geosciurus inauris* (ZIMMERMAN, 1780); GCA_004024805.1), Siberian chipmunk (*Eutamias sibiricus* (LAXMANN, 1769); GCA_025594165.1 (Li et al., 2022)), Daurian ground squirrel (*Spermophilus dauricus* BRANDT, 1843; GCA_002406435.1), thirteen-lined ground squirrel (*Ictidomys tridecemlineatus*; GCA_016881025.1 (Fu et al., 2021)), Arctic ground squirrel (*Urocitellus parryii*; GCA_003426925.1), Himalayan marmot (*Marmota himalayana* (HODGSON, 1841); GCA_005280165.1 (Bai et al., 2019)), Alpine marmot (*Marmota marmota marmota*; GCA_001458135.2), Vancouver Island marmot (*Marmota vancouverensis* (SWARTH, 1911); GCA_005458795.1), and yellow-bellied marmot or rock chuck (*Marmota flaviventer* (AUDUBON & BACHMAN, 1841); GCA_003676075.3) using 3,269 busco genes. To do so, we first obtained fasta files for each busco gene using

BUSCO_phylogenomics (https://github.com/jamiemcg/BUSCO_phylogenomics) with default settings, and we aligned them individually using *Mafft* v7.505 (Katoh and Standley, 2013). Next, we reconstructed phylogenetic trees for each gene separately with *iqtree* v2.3.4 (Minh et al., 2020) utilising a GTR+I+G substitution model. Lastly, we combined the gene trees into a species tree using a multi-species coalescent model using *ASTRAL-III* v5.7.8 (Zhang et al., 2018). Phylogenetic discordance was investigated and visualised using *DiscoVista* (Sayyari et al., 2018).

Results & Discussion

The 2.8 Gbp large genome of *Pteromys volans* is organised in the 19 largest scaffolds (Figure 1b), representing the expected gonosome and 18 autosomes (2*n* = 38; Oshida and Yoshida, 1996; Oshida et al., 2000; Table 1). Mapping the PacBio HiFi reads back to the genome assembly indicates a mean coverage of 22.6×. Scaffold CM117769.1 shows roughly half the mean genome coverage (12.1×; Supplementary Figure S10, Supplementary Table S3) and minimal interchromosomal contacts (Figure 1b), indicating weak or no interaction with other chromosomes. This along with the chromosome assignment by NCBI GenBank and a blastn-based search of the SRY gene against the novel genome assembly identities this scaffold as the heteromorph gonosomes, which confirms the morphological sex determination of the sequenced specimen. The total sequencing depth was not enough to assemble the Y chromosome. Two runs of PacBio generated in total 64.7 Gbp in filtered HiFi reads that were used for contig genome assembly (run 1: 63.9 Gbp, 7,143,058 reads; run 2: 836.4 Mbp, 90,539 reads). We were able to close two assembly gaps with *TGS-GapCloser*. *Compleasm* scores (99.9% complete, 0.78% duplicated, 0% fragmented and only 0.03% missing genes; Table 2) and *BUSCO* scores (97% complete, 2% duplication, 0.6% fragmented and only 2.4% missing genes; Table 2), a *k*-mer completeness of 97.04% (Consensus QV = 42.7294, corresponding to an error rate of $5.33 \times 10^{-}$□), the *k*-mer multiplicity spectrum (Supplementary Figure S9), and assembly statistics (scaffold-level N50 = 157 Mbp, contig-level N50 = 60Mbp, scaffold-level L90 = 18, contig-level L90 = 73, etc.; Table 1) indicate a high contiguity. The *k*-mer multiplicity corresponds closely to the expected genome coverage of ~22× (Supplementary Figure S9). *GenomeScope* showed a *k*-mer spectrum of an individual with extremely low heterozygosity,

estimated as 0.2%, consistent with a previous study reporting very low heterozygosities for the Finnish population of *P. volans* from which the sequenced individual originates (Ito Dos Santos et al., 2024). The largest 19 scaffolds cover 92% of the assembly and showed at least 10 copies of the telomere repeat TTAGGG on either end. More than 100 copies were found on both ends for nine and at one end for seven additional scaffolds. Only three scaffolds did not show telomeres with more than 100 copies of TTAGGG (Figure 2a, Supplementary Figure S6). The 19 largest scaffolds reached *compleasm* scores of 99.89% complete, 0.78% duplicated, 0.07% fragmented and 0.04% missing genes, only one complete gene less than the full assembly, and *BUSCO* scores equal to the whole assembly, indicating a high completeness of the potential chromosomes. The GC-content with 40.77% lies slightly below eutherian levels of 40.9 – 41.8% (Duret and Galtier, 2009). The homology-based gene prediction with *GeMoMa* identified a total of 27,383 genes and 54,238 transcripts with a *BUSCO* completeness of 91.5% (Supplementary Table S2). More than 97.26% (52,752) of those unique protein query sequences were annotated to sequences from the 'Swiss-Prot' database with the *Diamond blastp*-option. The functional annotation with *InterProScan* identified 53,845 unique protein sequences, while 73.3% (39,487) could be assigned to at least one Gene Ontology (GO) term. *Blobtools* identified 17 scaffolds belonging to viruses of the phylum Artverviricota (Supplementary Figure S8), which were removed from the assembly. In total, 970,478,186 bp (34.07%; Table 3) were masked as repeat elements.

The novel genome assembly phylogenetically clusters with the genomic data of the other published members of Sciuridae as expected based on other studies (Zelditch et al., 2015; Menéndez et al., 2021; Figure 3).

Members of the sister tribe Sciurini LINNAEUS, 1758 showed only minor synteny differences in their chromosomal architecture (Figure 1c). Noteworthy is the apomorph chromosomal fission of scaffold 1 shown in both *Sciurus* species investigated here (Figure 1c). A fused scaffold 1 seems to be a plesiomorph trait kept in *P. volans* if compared to a member of Xerinae OSBORN, 1910 as a clade outside of Sciurinae FISCHER DE WALDHEIM, 1817 (compare to the phylogenetic tree in Figure 3). More distant taxa, such as the European ground squirrel, show a higher degree of chromosomal rearrangements (Figure 1c).

As mentioned above, *Blobtools* identified scaffolds belonging to the viral family Retroviridae. Due to this phylogenetic assignment, those sequences might originate from proviral DNA rather than pararetroviral virions. But we cannot infer if the viruses were active or rather in a proviral phase. Since, the cause for its natural death is unknown and another adult male individual was found with the sequenced specimen, it could shed light on a potential pathogenic relationship of *P. volans* with retroviruses.

The mitochondrial genome has a length of 16,511 bp, including 13 protein-coding genes (CytB, ND1-6, COX1-3, ATP6, and ATP8), two ribosomal RNA genes (12S rRNA and 16S rRNA), and 22 tRNA genes (Figure 2b). This new mitogenome, from the westernmost distribution, contains two fewer base pairs than that of an individual from the Korean Peninsula (Ryu et al., 2013), one of the easternmost ranges, mostly considered to belong to the subspecies *Pteromys volans buechneri* SATUNIN, 1903 (Thorington et al., 2012; Koprowski et al., 2016). Previous genetic studies showed a consistent differentiation between the East Asian mainland populations and the northern Eurasian populations (Oshida et al., 2005; Koh et al., 2008; Ito Dos Santos et al., 2024). However, a more comprehensive sampling and its genomic analysis is required to further resolve the intraspecific phylogeography of the Siberian flying squirrel (Ito Dos Santos et al., 2024), which could be facilitated by the newly presented whole-genome assembly.

For the first time, we provide a chromosome-level genome assembly of a flying squirrel species, namely *Pteromys v. volans* inhabiting extremely fluctuating environments, which will help to understand many aspects of the species' biology, such as phylogeny, ecological, and nutritional genomics (Ungerer et al., 2008). The sequenced specimen is a representative of the species' westernmost population, presumably only connected by a narrow corridor to the metapopulation (Kurhinen et al., 2011) with a current conservation focus (Flying Squirrel LIFE; LIFE17 NAT/FI/000469). A previously shown low standing genetic variation within this western population (Ito Dos Santos et al., 2024) is of considerable importance for its adaptive capacity and inferred conservation strategies. Thus, necessary in-depth investigations on inbreeding through runs of homozygosity (ROH) or the genetic health looking at mutational load can only be accomplished with whole-genome assemblies, as provided here. It will also constitute the reference to identify informative genetic markers for conservation applications such as population

monitoring. Therefore, this novel genome assembly has the potential to not only shed light on processes of natural selection and adaptation interesting beyond the species context but will provide the basis for applied conservation-relevant studies on the population genetics of the species itself.

References

Abramchuk, A.V., 2021. The Siberian flying squirrel (Pteromys volans) in Belarus: distribution, abundance, threats, and conservation. Theriol. Ukr. 2021, 69–79. https://doi.org/10.15407/TU2208

Amori, G., 2024. Pteromys volans (Europe assessment). The IUCN Red List of Threatened Species 2024: e.T18702A221782992. https://doi.org/10.2305/IUCN.UK.2024-2.RLTS.T18702A221782992.en

Bai, L., Liu, B., Ji, C., Zhao, S., Liu, S., Wang, R., Wang, W., Yao, P., Li, X., Fu, X., Yu, H., Liu, M., Han, F., Guan, N., Liu, H., Liu, D., Tao, Y., Wang, Z., Yan, S., Florant, G., Butcher, M.T., Zhang, J., Zheng, H., Fan, J., Liu, E., 2019. Hypoxic and Cold Adaptation Insights from the Himalayan Marmot Genome. iScience 11, 505–507. https://doi.org/10.1016/j.isci.2019.01.019

Barnett, D.W., Garrison, E.K., Quinlan, A.R., Strömberg, M.P., Marth, G.T., 2011. BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics 27, 1691–1692. https://doi.org/10.1093/bioinformatics/btr174

Boratyn, G.M., Camacho, C., Cooper, P.S., Coulouris, G., Fong, A., Ma, N., Madden, T.L., Matten, W.T., McGinnis, S.D., Merezhuk, Y., Raytselis, Y., Sayers, E.W., Tao, T., Ye, J., Zaretskaya, I., 2013. BLAST: a more efficient report with usability improvements. Nucleic Acids Research 41, W29–W33. https://doi.org/10.1093/nar/gkt282

Brandies, P., Peel, E., Hogg, C.J., Belov, K., 2019. The Value of Reference Genomes in the Conservation of Threatened Species. Genes 10, 846. https://doi.org/10.3390/genes10110846

Buchfink, B., Reuter, K., Drost, H.-G., 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods 18, 366–368. https://doi.org/10.1038/s41592-021-01101-x

Buchfink, B., Xie, C., Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. Nat Methods 12, 59–60. https://doi.org/10.1038/nmeth.3176

Cantarel, B.L., Korf, I., Robb, S.M.C., Parra, G., Ross, E., Moore, B., Holt, C., Sánchez Alvarado, A., Yandell, M., 2008. MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res. 18, 188–196. https://doi.org/10.1101/gr.6743907

Cheng, H., Concepcion, G.T., Feng, X., Zhang, H., Li, H., 2021. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nat Methods 18, 170–175. https://doi.org/10.1038/s41592-020-01056-5

Clarke, Z., Bader, G., 2024. Healthy woodchuck genome with viral sequences appended used for single-cell RNA-seq analysis. https://doi.org/10.5281/ZENODO.10855128

Dudchenko, O., Shamim, M.S., Batra, S.S., Durand, N.C., Musial, N.T., Mostofa, R., Pham, M., Glenn St Hilaire, B., Yao, W., Stamenova, E., Hoeger, M., Nyquist, S.K., Korchina, V., Pletch, K., Flanagan, J.P., Tomaszewicz, A., McAloose, D., Pérez Estrada, C., Novak, B.J., Omer, A.D., Aiden, E.L., 2018. The Juicebox

Assembly Tools module facilitates *de novo* assembly of mammalian genomes with chromosome-length scaffolds for under $1000. https://doi.org/10.1101/254797

Duret, L., Galtier, N., 2009. Biased Gene Conversion and the Evolution of Mammalian Genomic Landscapes. Annu. Rev. Genom. Hum. Genet. 10, 285–311. https://doi.org/10.1146/annurev-genom-082908-150001

EELIS, 2025. EELIS (Eesti looduse infosüsteem). Keskkonnaagentuur. Pteromys volans (lendorav). https://www.eelis.ee/lnim/-646374363.

European Commission, 2021. Commission notice Guidance document on the strict protection of animal species of Community interest under the Habitats Directive.

Finbif, 2024. ZMUO Common vertebrate collection. https://doi.org/10.15468/U5MR6A

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. Proc. Natl. Acad. Sci. U.S.A. 117, 9451–9457. https://doi.org/10.1073/pnas.1921046117

Fu, R., Gillen, A.E., Grabek, K.R., Riemondy, K.A., Epperson, L.E., Bustamante, C.D., Hesselberth, J.R., Martin, S.L., 2021. Dynamic RNA Regulation in the Brain Underlies Physiological Plasticity in a Hibernating Mammal. Front. Physiol. 11, 624677. https://doi.org/10.3389/fphys.2020.624677

Greiner, S., Lehwark, P., Bock, R., 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Research 47, W59–W64. https://doi.org/10.1093/nar/gkz238

Gu, Z., Gu, L., Eils, R., Schlesner, M., Brors, B., 2014. *circlize* implements and enhances circular visualization in R. Bioinformatics 30, 2811–2812. https://doi.org/10.1093/bioinformatics/btu393

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29, 1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Hanski, I.K., 1998. Home ranges and habitat use in the declining flying squirrel *Pteromys volans* in managed forests. Wildlife Biology 4, 33–46. https://doi.org/10.2981/wlb.1998.013

Hanski, I.K., Stevens, P.C., Ihalempiä, P., Selonen, V., 2000. Home-Range Size, Movements, and Nest-Site Use in the Siberian Flying Squirrel, *Pteromys volans*. Journal of Mammalogy 81, 798–809. https://doi.org/10.1644/1545-1542(2000)081<0798:HRSMAN>2.3.CO;2

Hokkanen, H., Törmälä, T., Vuorinen, H., 1982. Decline of the flying squirrel Pteromys volans l. populations in Finland. Biological Conservation 23, 273–284. https://doi.org/10.1016/0006-3207(82)90083-0

Hokkanen, H., Törmälä, T., Vuorinen, H., 1977. Seasonal changes in the circadian activity of Pteromys volans L. in central Finland. Annales Zoologici Fennici 14, 94–97.

Holden-Musser, M.E., Juškaitis, R., Musser, G., 2016. Family Dormice (Gliridae), in: Handbook of the Mammals of the World. Vol. 6. Lagomorphs and Rodents I. pp. 838–892.

Hopkins, S.S.B., 2016. Family Aplodontiidae (Mountain Beaver), in: Handbook of the Mammals of the World. Vol. 6. Lagomorphs and Rodents I. pp. 642–647.

Huang, N., Li, H., 2023. compleasm: a faster and more accurate reimplementation of BUSCO. Bioinformatics 39, btad595. https://doi.org/10.1093/bioinformatics/btad595

Hurme, E., Mönkkönen, M., Sippola, A.-L., Ylinen, H., Pentinsaari, M., 2008. Role of the Siberian flying squirrel as an umbrella species for biodiversity in northern boreal forests. Ecological Indicators 8, 246–255. https://doi.org/10.1016/j.ecolind.2007.02.001

Ito Dos Santos, F., Lilley, T., Laine, V.N., Prost, S., Kurhinen, J., Susi, H., Gashev, S., Saveljev, A., Bondarchuk, S., Babina, S., Shishikin, A., Tirski, D., Vasina, A., Karpin, V., Kekkonen, J.A., 2024. Population structure of the endangered Siberian flying squirrel (*Pteromys volans*) revealed by genomic and mitochondrial data. https://doi.org/10.1101/2024.09.13.612918

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R., Hunter, S., 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30, 1236–1240. https://doi.org/10.1093/bioinformatics/btu031

Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O., Walichiewicz, J., 2005. Repbase Update, a database of eukaryotic repetitive elements. Cytogenet Genome Res 110, 462–467. https://doi.org/10.1159/000084979

Kang, L., Michalak, P., Hallerman, E., Moncrief, N.D., 2021. A draft genome assembly for the eastern fox squirrel, *Sciurus niger*. G3 Genes|Genomes|Genetics 11, jkab315. https://doi.org/10.1093/g3journal/jkab315

Katoh, K., Standley, D.M., 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution 30, 772–780. https://doi.org/10.1093/molbev/mst010

Keilwagen, J., Hartung, F., Paulini, M., Twardziok, S.O., Grau, J., 2018. Combining RNA-seq data and homology-based gene prediction for plants, animals and fungi. BMC Bioinformatics 19, 189. https://doi.org/10.1186/s12859-018-2203-5

Keilwagen, J., Wenk, M., Erickson, J.L., Schattat, M.H., Grau, J., Hartung, F., 2016. Using intron position conservation for homology-based gene prediction. Nucleic Acids Res 44, e89–e89. https://doi.org/10.1093/nar/gkw092

Koh, H.-S., Jin, Y., Yang, B.-G., Lee, B.-K., Heo, S.-W., Jang, K.-H., 2008. Taxonomic Status of Siberian Flying Squirrel from Korea (Pteromys volans aluco Thomas 1907). Animal Systematics, Evolution and Diversity 24, 169–172. https://doi.org/10.5635/KJSZ.2008.24.2.169

Kokot, M., Długosz, M., Deorowicz, S., 2017. KMC 3: counting and manipulating *k* -mer statistics. Bioinformatics 33, 2759–2761. https://doi.org/10.1093/bioinformatics/btx304

Kolmogorov, M., Yuan, J., Lin, Y., Pevzner, P.A., 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37, 540–546. https://doi.org/10.1038/s41587-019-0072-8

Koprowski, J., Goldstein, E., Bennett, K., Pereira, C., 2016. Family Sciuridae (Tree, Flying and Ground Squirrels, Chipmunks, Prairie Dogs and Marmots), in: Handbook of the Mammals of the World. Vol. 6. Lagomorphs and Rodents I. pp. 648–837.

Kurhinen, J., Kulebyakina, E., Zadiraka, E., Mamontov, V., Muravskaya, E., Hanski, I.K., 2011. Distribution of the Siberian flying squirrel (*Pteromys volans* L.) in taiga isthmuses between Baltic and White Sea regions. Acta Zoologica Lituanica 21, 306–310. https://doi.org/10.2478/v10043-011-0037-8

Laetsch, D.R., Blaxter, M.L., 2017. BlobTools: Interrogation of genome assemblies. F1000Res 6, 1287. https://doi.org/10.12688/f1000research.12232.1

Laetsch, D.R., Koutsovoulos, G., Booth, T., Stajich, J., Sujai Kumar, 2017. DRL/blobtools: BlobTools v1.0.1. https://doi.org/10.5281/ZENODO.845347

Lampila, S., Wistbacka, R., Mäkelä, A., Orell, M., 2009. Survival and population growth rate of the threatened Siberian flying squirrel (*Pteromys volans*) in a fragmented forest landscape. Écoscience 16, 66–74. https://doi.org/10.2980/16-1-3184

Lee, Mu-Yeong, Park, S., Hong, Y., Kim, Y., Voloshina, I., Myslenkov, A., Saveljev, A.P., Choi, T., Piao, R., An, J., Lee, Mun-Han, Lee, H., Min, M., 2008. Mitochondrial genetic diversity and phylogenetic relationships of Siberian flying squirrel *(Pteromys volans)* populations. Animal Cells and Systems 12, 269–277. https://doi.org/10.1080/19768354.2008.9647182

Levy Karin, E., Mirdita, M., Söding, J., 2020. MetaEuk—sensitive, high-throughput gene discovery, and annotation for large-scale eukaryotic metagenomics. Microbiome 8, 48. https://doi.org/10.1186/s40168-020-00808-x

Li, H., 2021. New strategies to improve minimap2 alignment accuracy. Bioinformatics 37, 4572–4574. https://doi.org/10.1093/bioinformatics/btab705

Li, H., 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34, 3094–3100. https://doi.org/10.1093/bioinformatics/bty191

Li, R., Zhang, M., Cha, M., Xiang, J., Yi, X., 2022. Chromosome-level genome assembly of the Siberian chipmunk (Tamias sibiricus). Sci Data 9, 783. https://doi.org/10.1038/s41597-022-01910-5

Lin, Y., Ye, C., Li, X., Chen, Q., Wu, Y., Zhang, F., Pan, R., Zhang, S., Chen, S., Wang, X., Cao, S., Wang, Y., Yue, Y., Liu, Y., Yue, J., 2023. quarTeT: a telomere-to-telomere toolkit for gap-free genome assembly and centromeric repeat identification. Horticulture Research 10, uhad127. https://doi.org/10.1093/hr/uhad127

Liukko, U.-M., Henttonen, H., Kauhala, K., Kojola, I., Kyheröinen, E.-M., Pitkänen, J., 2019. Mammals, in: The 2019 Red List of Finnish Species. Ministry of the Environment & Finnish Environment Institute. Helsinki.

Manni, M., Berkeley, M.R., Seppey, M., Simão, F.A., Zdobnov, E.M., 2021. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. Molecular Biology and Evolution 38, 4647–4654. https://doi.org/10.1093/molbev/msab199

Mead, D., Fingland, K., Cripps, R., Portela Miguez, R., Smith, M., Corton, C., Oliver, K., Skelton, J., Betteridge, E., Dolucan, J., Dudchenko, O., Omer, A.D., Weisz, D., Lieberman Aiden, E., Fedrigo, O., Mountcastle, J., Jarvis, E., McCarthy, S.A., Sims, Y., Torrance, J., Tracey, A., Howe, K., Challis, R., Durbin, R., Blaxter, M., 2020b. The genome sequence of the Eurasian red squirrel, Sciurus vulgaris Linnaeus 1758. Wellcome Open Res 5, 18. https://doi.org/10.12688/wellcomeopenres.15679.1

Mead, D., Fingland, K., Cripps, R., Portela Miguez, R., Smith, M., Corton, C., Oliver, K., Skelton, J., Betteridge, E., Doulcan, J., Quail, M.A., McCarthy, S.A., Howe, K., Sims, Y., Torrance, J., Tracey, A., Challis, R., Durbin, R., Blaxter, M., 2020a. The genome sequence of the eastern grey squirrel, Sciurus carolinensis Gmelin, 1788. Wellcome Open Res 5, 27. https://doi.org/10.12688/wellcomeopenres.15721.1

Menéndez, I., Gómez Cano, A.R., Cantalapiedra, J.L., Peláez-Campomanes, P., Álvarez-Sierra, M.Á., Hernández Fernández, M., 2021. A multi-layered approach to the diversification of squirrels. Mammal Review 51, 66–81. https://doi.org/10.1111/mam.12215

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, 1530–1534. https://doi.org/10.1093/molbev/msaa015

Mönkkönen, M., Reunanen, P., Nikula, A., Inkeröinen, J., Forsman, J., 1997. Landscape characteristics associated with the occurrence of the flying squirrel *Pteromys volans* in old-growth forests of northern Finland. Ecography 20, 634–642. https://doi.org/10.1111/j.1600-0587.1997.tb00432.x

Okonechnikov, K., Conesa, A., García-Alcalde, F., 2016. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics 32, 292–294. https://doi.org/10.1093/bioinformatics/btv566

Oshida, T., Abramov, A., Yanagawa, H., Masuda, R., 2005. Phylogeography of the Russian flying squirrel (*Pteromys volans*): implication of refugia theory in arboreal small mammal of Eurasia. Molecular Ecology 14, 1191–1196. https://doi.org/10.1111/j.1365-294X.2005.02475.x

Oshida, T., Yanagawa, H., Tsuda, M., Indue, S., Yoshida, M.C., 2000. Comparisons of the banded karyotypes between the smallJapanese flying squirrel,Pteromys momonga and the Russian flying squirrel, P. volans(Rodentia, Sciuridae). Caryologia 53, 133–140. https://doi.org/10.1080/00087114.2000.10589188

Oshida, T., Yoshida, M.C., 1996. Banded karyotypes and the localization of ribosomal RNA genes of Eurasian flying squirrel, *Pteromys volans orii* (Mammalia, Rodentia). Caryologia 49, 219–225. https://doi.org/10.1080/00087114.1996.10797366

Paez, S., Kraus, R.H.S., Shapiro, B., Gilbert, M.T.P., Jarvis, E.D., Vertebrate Genomes Project Conservation Group, Al-Ajli, F.O., Ceballos, G., Crawford, A.J., Fedrigo, O., Johnson, R.N., Johnson, W.E., Marques-Bonet, T., Morin, P.A., Mueller, R.C., Ryder, O.A., Teeling, E.C., Venkatesh, B., 2022. Reference genomes for conservation. Science 377, 364–366. https://doi.org/10.1126/science.abm8127

Pilāts, V., 2023. Abstracts of the 81st Scientific Conference of the University of Latvia: January – February 2023. EEB 21, 33–44. https://doi.org/10.22364/eeb.21.05

Pilāts, V., 2010. Results of flying squirrel survey in Latvia., Wildlife census methods: reliability and application. Workshop proceedings. International Scientific Practical Workshop. Kaunas.

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., Lopez, R., 2005. InterProScan: protein domains identifier. Nucleic Acids Research 33, W116–W120. https://doi.org/10.1093/nar/gki442

Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. https://doi.org/10.1093/bioinformatics/btq033

Ranallo-Benavidez, T.R., Jaron, K.S., Schatz, M.C., 2020. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nat Commun 11, 1432. https://doi.org/10.1038/s41467-020-14998-3

Rane, R., Pichlmueller, F., Lee, R., Edwards, O., Gemmell, N., Tompkins, D., 2020. Rattus rattus chromosomal assembly (New zealand strain). https://doi.org/10.25919/5E4F60EA8B64E

Reyes, A., Gissi, C., Pesole, G., Catzeflis, F.M., Saccone, C., 2000. Where Do Rodents Fit? Evidence from the Complete Mitochondrial Genome of Sciurus vulgaris. Molecular Biology and Evolution 17, 979–983. https://doi.org/10.1093/oxfordjournals.molbev.a026379

Rhie, A., McCarthy, S.A., Fedrigo, O., Damas, J., Formenti, G., Koren, S., Uliano-Silva, M., Chow, W., Fungtammasan, A., Kim, J., Lee, C., Ko, B.J., Chaisson, M., Gedman, G.L., Cantin, L.J., Thibaud-Nissen, F., Haggerty, L., Bista, I.,

Smith, M., Haase, B., Mountcastle, J., Winkler, S., Paez, S., Howard, J., Vernes, S.C., Lama, T.M., Grutzner, F., Warren, W.C., Balakrishnan, C.N., Burt, D., George, J.M., Biegler, M.T., Iorns, D., Digby, A., Eason, D., Robertson, B., Edwards, T., Wilkinson, M., Turner, G., Meyer, A., Kautt, A.F., Franchini, P., Detrich, H.W., Svardal, H., Wagner, M., Naylor, G.J.P., Pippel, M., Malinsky, M., Mooney, M., Simbirsky, M., Hannigan, B.T., Pesout, T., Houck, M., Misuraca, A., Kingan, S.B., Hall, R., Kronenberg, Z., Sović, I., Dunn, C., Ning, Z., Hastie, A., Lee, J., Selvaraj, S., Green, R.E., Putnam, N.H., Gut, I., Ghurye, J., Garrison, E., Sims, Y., Collins, J., Pelan, S., Torrance, J., Tracey, A., Wood, J., Dagnew, R.E., Guan, D., London, S.E., Clayton, D.F., Mello, C.V., Friedrich, S.R., Lovell, P.V., Osipova, E., Al-Ajli, F.O., Secomandi, S., Kim, H., Theofanopoulou, C., Hiller, M., Zhou, Y., Harris, R.S., Makova, K.D., Medvedev, P., Hoffman, J., Masterson, P., Clark, K., Martin, F., Howe, Kevin, Flicek, P., Walenz, B.P., Kwak, W., Clawson, H., Diekhans, M., Nassar, L., Paten, B., Kraus, R.H.S., Crawford, A.J., Gilbert, M.T.P., Zhang, G., Venkatesh, B., Murphy, R.W., Koepfli, K.-P., Shapiro, B., Johnson, W.E., Di Palma, F., Marques-Bonet, T., Teeling, E.C., Warnow, T., Graves, J.M., Ryder, O.A., Haussler, D., O'Brien, S.J., Korlach, J., Lewin, H.A., Howe, Kerstin, Myers, E.W., Durbin, R., Phillippy, A.M., Jarvis, E.D., 2021. Towards complete and error-free genome assemblies of all vertebrate species. Nature 592, 737–746. https://doi.org/10.1038/s41586-021-03451-0

Rhie, A., Walenz, B.P., Koren, S., Phillippy, A.M., 2020. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. Genome Biol 21, 245. https://doi.org/10.1186/s13059-020-02134-9

Ryu, S.H., Kwak, M.J., Hwang, U.W., 2013. Complete mitochondrial genome of the Eurasian flying squirrel Pteromys volans (Sciuromorpha, Sciuridae) and revision of rodent phylogeny. Mol Biol Rep 40, 1917–1926. https://doi.org/10.1007/s11033-012-2248-x

Santangeli, A., Wistbacka, R., Hanski, I.K., Laaksonen, T., 2013. Ineffective enforced legislation for nature conservation: A case study with Siberian flying squirrel and forestry in a boreal landscape. Biological Conservation 157, 237–244. https://doi.org/10.1016/j.biocon.2012.09.012

Sayyari, E., Whitfield, J.B., Mirarab, S., 2018. DiscoVista: Interpretable visualizations of gene tree discordance. Molecular Phylogenetics and Evolution 122, 110–115. https://doi.org/10.1016/j.ympev.2018.01.019

Selonen, V., Mäkeläinen, S., 2017. Ecology and protection of a flagship species, the Siberian flying squirrel. Hystrix, the Italian Journal of Mammalogy 28. https://doi.org/10.4404/hystrix-28.2-12328

Selonen, V., Wistbacka, R., 2016. Siberian flying squirrels do not anticipate future resource abundance. BMC Ecol 16, 51. https://doi.org/10.1186/s12898-016-0107-7

Shar, S., Lkhagvasuren, D., Henttonen, H., Maran, T., Hanski, I.K., 2016. Pteromys volans (errata version published in 2017). The IUCN Red List of Threatened Species 2016: e.T18702A115144995. https://doi.org/10.2305/IUCN.UK.2016-3.RLTS.T18702A22270935.en

Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J.M., Birol, İ., 2009. ABySS: A parallel assembler for short read sequence data. Genome Res. 19, 1117–1123. https://doi.org/10.1101/gr.089532.108

Smit, A.F.A., Hubley, R., Green, P., 2023. RepeatMasker.

Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J., Prins, P., 2015. Sambamba: fast processing of NGS alignment formats. Bioinformatics 31, 2032–2034. https://doi.org/10.1093/bioinformatics/btv098

The UniProt Consortium, 2019. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Research 47, D506–D515. https://doi.org/10.1093/nar/gky1049

Thorington, R.W., Koprowski, J.L., Steele, M.A., Whatton, J.F. (Eds.), 2012. Squirrels of the world. Johns Hopkins University Press, Baltimore.

Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E.S., Fischer, A., Bock, R., Greiner, S., 2017. GeSeq – versatile and accurate annotation of organelle genomes. Nucleic Acids Research 45, W6–W11. https://doi.org/10.1093/nar/gkx391

Törmälä, T., Vuorinen, H., Hokkanen, H., 1980. Timing of circadian activity in the flying squirrel in central Finland. Acta Theriologica 25, 461–474.

Uliano-Silva, M., Ferreira, J.G.R.N., Krasheninnikova, K., Darwin Tree of Life Consortium, Blaxter, M., Mieszkowska, N., Hall, N., Holland, P., Durbin, R., Richards, T., Kersey, P., Hollingsworth, P., Wilson, W., Twyford, A., Gaya, E., Lawniczak, M., Lewis, O., Broad, G., Martin, F., Hart, M., Barnes, I., Formenti, G., Abueg, L., Torrance, J., Myers, E.W., Durbin, R., Blaxter, M., McCarthy, S.A., 2023. MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio high fidelity reads. BMC Bioinformatics 24, 288. https://doi.org/10.1186/s12859-023-05385-y

Ungerer, M.C., Johnson, L.C., Herman, M.A., 2008. Ecological genomics: understanding gene and genome function in the natural environment. Heredity 100, 178–183. https://doi.org/10.1038/sj.hdy.6800992

Wistbacka, R., Orell, M., Santangeli, A., 2018. The tragedy of the science-policy gap – Revised legislation fails to protect an endangered species in a managed boreal landscape. Forest Ecology and Management 422, 172–178. https://doi.org/10.1016/j.foreco.2018.04.017

Wolf, J.F., Bowman, J., Keobouasone, S., Taylor, R.S., Wilson, P.J., 2022. A *de novo* genome assembly and annotation of the southern flying squirrel (*Glaucomys volans*). G3 Genes|Genomes|Genetics 12, jkab373. https://doi.org/10.1093/g3journal/jkab373

Xu, M., Guo, L., Gu, S., Wang, O., Zhang, R., Peters, B.A., Fan, G., Liu, X., Xu, X., Deng, L., Zhang, Y., 2020. TGS-GapCloser: A fast and accurate gap closer for large genomes with low coverage of error-prone long reads. GigaScience 9, giaa094. https://doi.org/10.1093/gigascience/giaa094

Zagorodniuk, I., 2022. The Siberian flying squirrel (Pteromys volans) in south of Eastern Europe: distribution boundaries and its changes. Theriol. Ukr. 2022, 66–77. https://doi.org/10.15407/TU2308

Zelditch, M.L., Li, J., Tran, L.A.P., Swiderski, D.L., 2015. Relationships of diversity, disparity, and their evolutionary rates in squirrels (Sciuridae): DIVERSITY AND DISPARITY IN SQUIRRELS. Evolution 69, 1284–1300. https://doi.org/10.1111/evo.12642

Zhang, C., Rabiee, M., Sayyari, E., Mirarab, S., 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics 19, 153. https://doi.org/10.1186/s12859-018-2129-y

Zhou, C., McCarthy, S.A., Durbin, R., 2023. YaHS: yet another Hi-C scaffolding tool. Bioinformatics 39, btac808. https://doi.org/10.1093/bioinformatics/btac808

Table 1. Assembly statistics table on the scaffold-level as well was on the contig-level. (K)bp = (kilo) base pairs; # = number; GC content = Guanine-Cytosine content; auN = area under the Nx curve. A more detailed assembly statistics table can be found in the Supplementary Information (Supplementary Table S1).

Genome statistic	scaffold-level	contig-level
total length	2,848,339,732 bp	2,848,327,632 bp
# of sequences	638	759
# sequences (≥ 0 bp)	638	-
# sequences (≥ 1000 bp)	638	758
Total length (≥ 0 bp)	2,848,339,732 bp	-
Total length (≥ 1000 bp)	2,848,339,732 bp	2,848,326,863 bp
Largest sequence	214,577,690 bp	136,295,669 bp
GC level	40.77%	40.77%
N50	157,387,023 bp	60,698,384 bp
N90	42,248,379 bp	3,984,759 bp
auN	147,030,830.8 bp	62,127,644.0 bp
L50	8	16
L90	18	73
# N's per 100 kbp	0.42	0.00

Table 1. Busco gene scores generated with *compeasm* and *BUSCO* both based on the total number of 9,226 expected orthologs.

software	***compleasm***		***BUSCO***	
unit	***n***	**%**	***n***	**%**
Complete buscos (C = S + D)	9,217	99,90%	8,947	97%
Complete and single-copy buscos (S)	9,145	99.12%	8,763	95%
Complete and duplicated buscos (D)	72	0.78%	184	2%
Fragmented buscos, subclass 1 (F)	6	0.07%	51	0.6%
Fragmented buscos, subclass 2 (I)	0	0%	-	-
Missing buscos (M)	3	0.03%	228	2.4%

Table 2. Assembly assignment table estimated by *Repeatmodeler* and *Repeatmasker*. Most repeats fragmented by insertions or deletions have been counted as one element (marked with an asterisk).

Type of elements	number of elements [*n*]*	length occupied [bp]	percentage of sequence [%]
Retroelements	3,255,026	886,574,991 bp	31.13%
SINEs:	1,483,628	169,453,665 bp	5.95%
Penelope:	62	11,402 bp	0.00%
LINEs:	878,225	522,064,913 bp	18.33%
CRE/SLACS	0	0 bp	0.00%
L2/CR1/Rex	86,215	23,004,669 bp	0.81%
R1/LOA/Jockey	0	0 bp	0.00%
R2/R4/NeSL	246	56,708 bp	0.00%
RTE/Bov-B	5,179	1,231,196 bp	0.04%
L1/CIN4	786,498	497,749,583 bp	17.48%
LTR elements:	893,173	195,056,413 bp	6.85%
BEL/Pao	0	0 bp	0.00%
Ty1/Copia	0	0 bp	0.00%
Gypsy/DIRS1	6,436	1,386,734 bp	0.05%
Retroviral	869,037	191,862,047 bp	6.74%
DNA transposons	224,420	46,280,608 bp	1.62%
hobo-Activator	163,126	32,188,953 bp	1.13%
Tc1-IS630-Pogo	55,186	13,103,816 bp	0.46%
En-Spm	0	0 bp	0.00%
MULE-MuDR	251	41,323 bp	0.00%
PiggyBac	521	160,128 bp	0.01%
Tourist/Harbinger	323	50,596 bp	0.00%
Other (Mirage, P-element, Transib)	0	0 bp	0.00%
Rolling-circles	879	195,645 bp	0.01%
Unclassified:	59,407	3,502,033 bp	0.12%
Total interspersed repeats:		936,369,034 bp	32.87%
Small RNA:	50,111	3,274,868 bp	0.11%
Satellites:	49,554	2,663,208 bp	0.09%
Simple repeats:	530,473	22,897,566 bp	0.80%
Low complexity:	106,166	0 bp	0.19%

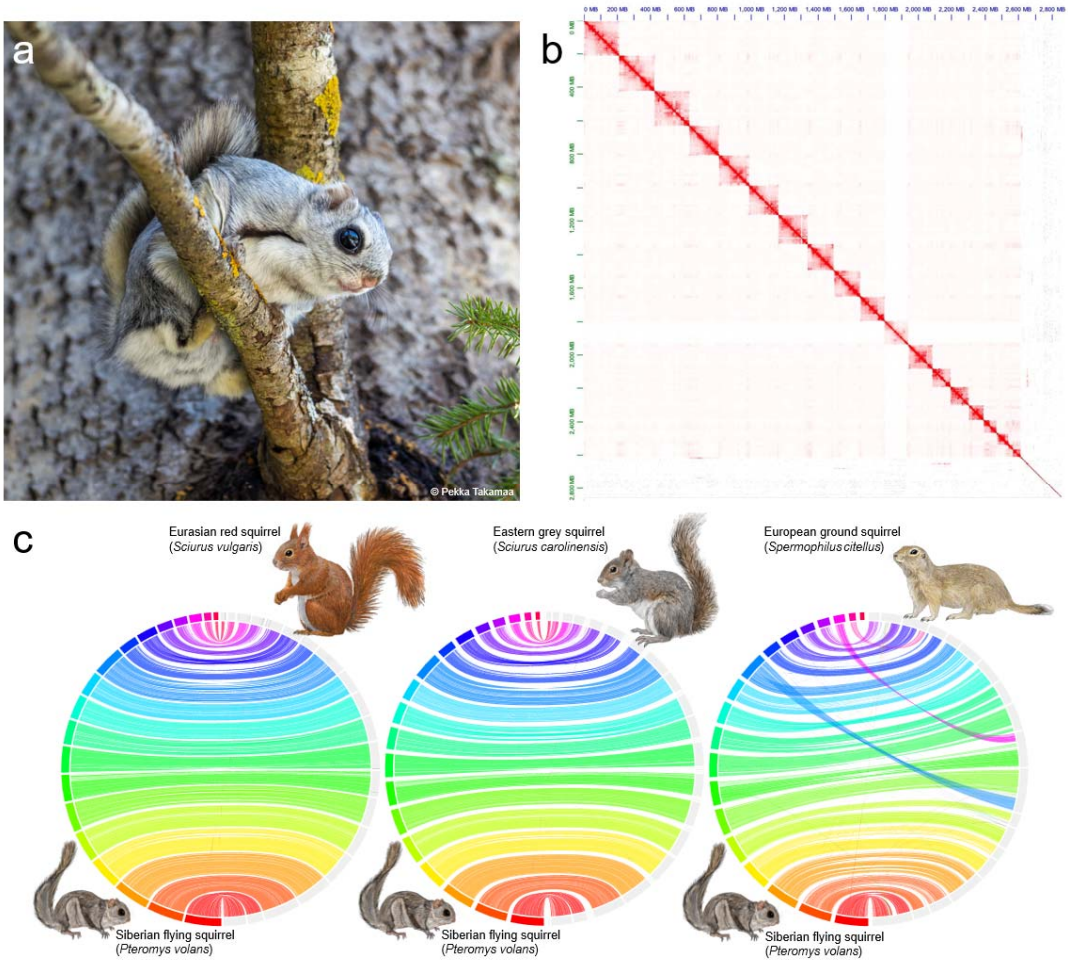

Figure 1. (a) Adult specimen of *Pteromys volans volans* in Lapua, Finland. **(b)** Omni-C contact density map depicting the 19 distinct chromosome-level scaffolds. **(c)** Circos plots comparing the synteny of the chromosome-scale genome assembly of *Pteromys volans* (Uoulu_pteVol_1.0) with chromosome-scale assemblies of a Eurasian red squirrel (*Sciurus vulgaris*), an eastern grey squirrel (*Sciurus carolinensis*), and a European ground squirrel (*Spermophilus citellus*). All three synteny plots are available in a higher resolution in the supplementary information (Supplementary Figures S1-S3).

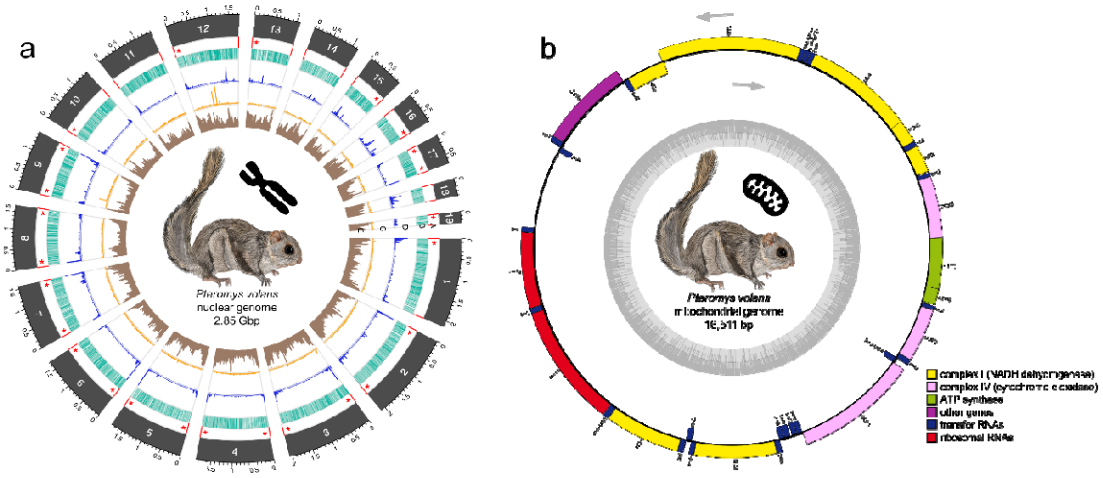

Figure 2. (a) Circos plot of the chromosome-level genome assembly of *Pteromys volans* (Uoulu_pteVol_1.0). The tracks show the following information for 19 scaffolds (length in Gbp) : Red [A]: Position of the telomeres; red asterisks indicate telomeric regions with length of > 100 copies measured with *quarTeT*. Turquoise [B]: Gene content within a 500,000 base pair windows. Blue [C]: Simple repeats. The length of the bars indicates the relative length of the simple repeats after clustering. Orange [D]: Transposable elements (TE) clustered by type. The length of the bars indicates the relative length of the TE after clustering. Brown [E]: GC content. **(b)** Circular gene map of the mitogenome of *Pteromys volans*. The different functional genes groups are shown in different colours, which are shown on the bottom right. The genes transcribed in clockwise and counterclockwise (indicated by arrows) are shown inside and outside of the external circle, respectively. The dark grey colour of inner circle shows the GC content. Figures 2a and 2b as well as the telomere overview plot from *quarTeT* can be found in the Supplementary Information (Supplementary Figures S4 – S6).

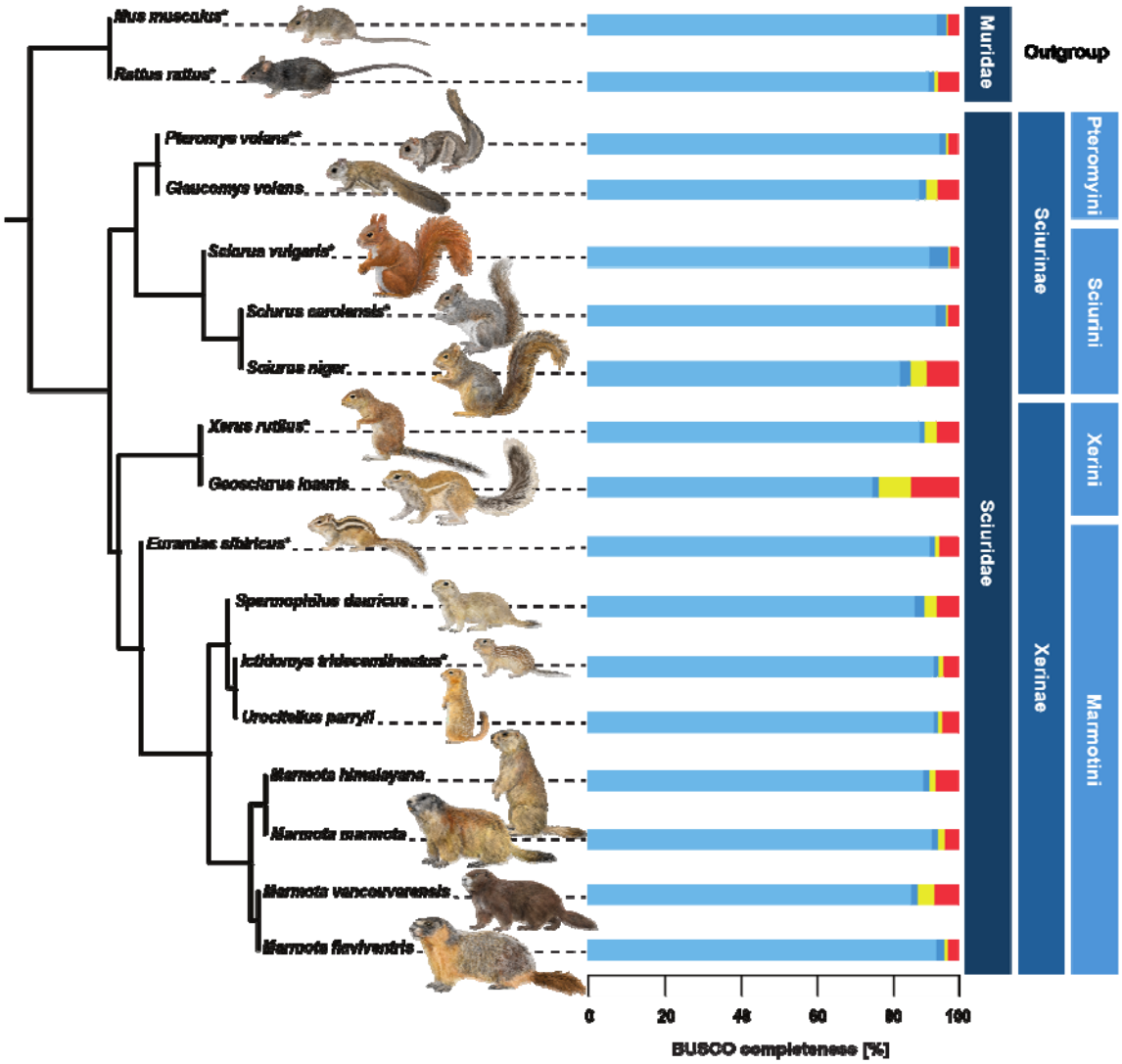

Figure 3. Coalescent phylogenetic tree from 17 rodent species based on 3,269 busco genes each with the assigned busco completeness [%]. Species with available chromosome-level genome assemblies were marked with an asterisk (*). The novel chromosome-level genome assembly from *Pteromys volans* is indicated with two asterisks (**). The outgroup and the according higher taxa are assigned on the right-hand side: families (darkest blue) as well as for the squirrel subfamilies (medium blue) and tribes (lightest blue). The branch quartet frequencies graphs generated by *DiscoVista* can be found in the Supplementary Information (Supplementary Figure S7).

Data Accessibility Statement

All underlying read data and the assembly are available at NCBI GenBank under BioProject PRJNA1141127. The hard-masked genome assembly and gene annotations are available at Dryad (https://doi.org/10.5061/dryad.3xsj3txth). More detailed results are available in the Supplementary Information.

Competing Interests Statement

The authors declare no competing interests.

Author Contributions

Gerrit Wehrenberg: Conceptualisation; Data Curation, Formal Analysis; Methodology; Visualisation; Writing – Original Draft Preparation (lead); Writing – Review & Editing (lead). **Angelika Kiebler**: Formal Analysis; Methodology; Visualisation; Writing – Original Draft Preparation; Writing – Review & Editing. **Carola Greve**: Formal Analysis; Methodology; Writing – Review & Editing. **Núria Beltrán-Sanz**: Data Curation, Formal Analysis; Methodology; Visualisation; Writing – Review & Editing. **Alexander Ben Hamadou**: Methodology. **René Meißner**: Formal Analysis; Methodology; Writing – Review & Editing. **Sven Winter**: Conceptualisation; Data Curation, Formal Analysis; Methodology; Supervision; Visualisation; Writing – Review & Editing. **Stefan Prost**: Conceptualisation; Data Curation, Formal Analysis; Funding Acquisition; Methodology; Project Administration; Supervision; Visualisation; Writing – Original Draft Preparation; Writing – Review & Editing.

Acknowledgments

G.W. and S.P. received funding from Biodiverse Anthropocenes (ANTS) through University of Oulu and the Research Council of Finland PROFI6 funding. We also thank Timo Perätie from Kuopion kaupungin alueelliset ympäristönsuojelupalvelut and the Zoological Museum of the University of Oulu, especially Mikko Vallinmäki, for sampling, sample provision, sample storage and database management. We thank Charlotte Gerheim of the LOEWE Translational Biodiversity Genomics (TBG) project based in the Senckenberg Research Institute and Museum, Frankfurt, for support with lab work. The authors wish to acknowledge CSC – IT Center for Science,

Finland, for computational resources. We also thank the Genome Technology Center (RGTC) at Radboud University Medical Center (Radboudumc; Nijmegen, The Netherlands) and the Bioscientia Institut für Medizinische Diagnostik GmbH (Ingelheim, Germany) for providing the PacBio SMRT sequencing service on the PacBio Revio platform. Yannis Schöneberg for helping with the data analysis. We are grateful for the permission to use a wonderful photograph of a Finnish *Pteromys volans volans* by Pekka Takamaa, as well as the scientific illustrations of various rodent species provided by ALADA BOOKS, S.L. Kiitos paljon to Professor Marko Mutanen for the abstract translation into the Finnish language.