High quality genome of the brown hare (*Lepus europaeus*) with chromosome-level scaffolding

Craig Michell[1,2], Joanna Collins[3], Pia K. Laine[4], Zosia Fekete[1,5], Riikka Tapanainen[1], Jonathan M. D. Wood[3], Steffi Goffart[1] and Jaakko L. O. Pohjoismäki*[1]

[1]Department of Environmental and Biological Sciences, University of Eastern Finland – Joensuu, Finland
[2]Red Sea Research Center, Biological and Environmental Science and Engineering Division, King Abdullah University of Science and Technology, – Thuwal, Kingdom of Saudi Arabia
[3]Tree of Life, Wellcome Sanger Institute, Hinxton, Cambridge, UK.
[4]Insitute of Biotechnology, University of Helsinki, Finland.
[5]Hungarian University of Agriculture and Life Sciences, Institute of Genetics and Biotechnology, Gödöllő, Hungary

*Corresponding author
Correspondence: jaakko.pohjoismaki@uef.fi

Abstract

We present here a high-quality genome assembly of the brown hare (*Lepus europaeus* Pallas), based on a fibroblast cell line of a male specimen from Liperi, Eastern Finland. This brown hare genome represents the first Finnish contribution to the European Reference Genome Atlas pilot effort to generate reference genomes for European biodiversity.

The genome was assembled using 25X PacBio HiFi sequencing data and scaffolded utilizing a Hi-C chromosome structure capture approach. After manual curation, the assembled genome length was 2,930,972,003 bp with N50 of 125.8Mb. 93.16% of the assembly could be assigned to 25 identified chromosomes (23 autosomes plus X and Y), matching the published karyotype. The chromosomes were numbered according to size. The genome has a high degree of completeness based on the BUSCO score (mammalia_odb10 database), Complete: 96.1% [Single copy: 93.1%, Duplicated: 3.0%], Fragmented 0.8%, and Missing 2.9%. The mitochondrial genome of the cell line was sequenced and assembled separately.

The brown hare genome is particularly interesting as this species readily hybridizes with the mountain hare (*Lepus timidus* L.) at the species contact zone in northern Eurasia, producing fertile offspring and resulting in gene flow between the two species. In addition to providing a

useful comparison for population studies, the genome can offer insight into the chromosomal evolution among Glires in general and Lagomorpha in particular. The chromosomal assembly of the genome also demonstrates that the cell line has not acquired karyotypic changes during culture.

Keywords: Lagomorpha, reference genome, chromosome structure, phylogeny, fibroblast cell line

Introduction

The brown hare (*Lepus europaeus* Pallas), also known as the European hare, is a widespread species in the western parts of Eurasia (Bock, 2020). Besides its native range, the brown hare has been introduced to numerous regions, including the British Isles, the Falkland Islands, Canada, South America, Australia and New Zealand. In many places the brown hare is regarded as an invasive species and a threat to the local ecosystems (Stott, 2003) or native species (Reid, 2011).

The brown hare is essentially a steppe-adapted species thriving in open grasslands and avoiding forested regions. Its colonization history in Europe is complex. The species has had apparent glacial refugia during the Pleistocene in the Italian peninsula, the Balkans and Asia Minor (Fickel *et al*, 2008), with a very interesting pre-glacial diversity hotspot in the archipelago of Greece (Minoudi *et al*, 2018). After the ice age, the species has been expanding its range both naturally as well as facilitated by human-caused changes in the landscape, especially the expansion of agricultural lands and pastures. As mentioned above, the brown hare has been also frequently introduced by humans to new areas from antiquity to the present day (Petrovan, 2013), resulting in potential mixing of ancestral populations. This is true also for Finland, where the species has arrived naturally from the southeast through the Karelian isthmus, but also *via* introductions to the southwestern parts of the country in the late 19th century.

Curiously, Linné was unaware of the existence of the brown hare, despite having described the mountain hare and the South African Cape hare (*Lepus capensis* L.). In fact, the brown hare is not native to Linné's home country Sweden and was introduced there only a century after his death (Thulin, 2003). As a reflection of the early neglection of this species, Peter Simon Pallas actually never published a formal description of the brown hare (Pallas, 1778). The species authority has been attributed to him based on the inclusion of *Lepus europaeus* into his identification table of hare species of the world (Pallas, 1778: 30) (Figure 1A). Consequently, there is also no information of the type locality or any other information of the species in the original publication. The type locality of the brown hare has been retrospectively assigned – without too convincing arguments – to Burgundy, France or Poland (Holden, 2005). Identifying Poland as a possible type locality can be traced to the Russian zoologist Sergey Ognev, and is based on his interpretation of Pallas mentioning in the 1778 text the hybrid forms of *Lepus variabilis* (synonym of *L. timidus*) and *L. europaeus* from Poland and Lithuania, as well as citing later works of Pallas (Ognev, 1940). In fact, Ognev (1940: 141) boldly states that "we consider southwestern Poland as a typical terrain for the nominal breed; there are hares with the characteristics of the basic form, expressed with full clarity" (translation from Russian).

While being widespread in the temperate regions of Western and Central Europe to the Caspian steppes in historical times, the brown hare began an impressive northward expansion in

the 19th century. In the Fennoscandian region, the species reached the St Petersburg area by 1820s (Ognev, 1940), arriving in Finland at the turn of the century (Thulin, 2003). By 1930, the species had become established throughout southern Finland, with the range expansion stagnating southwest of the N64°-N62° -line until the 1990s. Unlike its northern relative, the mountain hare (*Lepus timidus* L.), the brown hare struggles in snowy winter conditions, with 150 days of annual snow cover limiting its northward expansion (Levanen *et al*, 2018a).

Quite likely benefitting from the ongoing anthropogenic climate change, the brown hare has had another impressive range expansion during the last three decades, in some places reaching the polar circle (Levanen *et al.*, 2018a) (Figure 1B). This range expansion together with the introductions have brought the brown hare increasingly in contact with the mountain hare, especially at the northern edges of the species distribution, which constitutes a threat especially to the more temperate climate adapted populations and subspecies of the mountain hare (Reid, 2011; Reid & Montgomery, 2007; Thulin, 2003; Thulin *et al*, 2021). Besides competing, brown hares also hybridize with mountain hares, producing fertile offspring and resulting in gene flow between the species. Interestingly, this gene flow is biased towards the brown hare (Ferreira *et al*, 2021; Levanen *et al.*, 2018a; Levanen *et al*, 2018b; Thulin & Tegelström, 2002), which obtains genetic variation from mountain hares, some of which might have adaptive significance (Pohjoismaki *et al*, 2021). In contrast to the brown hare's expansion towards north, many of the Central European brown hare populations are contracting in range and numbers, likely driven by changes in agricultural practices and land use (Schai-Braun *et al*, 2013; Smith *et al*, 2005).

Overall, the brown hare's features as an invasive species outside Europe and expansive species in the north, its hybridization tendency and complex population ancestries make the species highly interesting not only to study population genetics and adaptation mechanisms, but also for the understanding the genomic makeup of species boundaries (Gaertner *et al*, 2022). These studies would be greatly facilitated by the availability of high quality, chromosomally scaffolded reference genomes (Blaxter *et al*, 2022).

Over the last decade we have witnessed a boom in sequencing technologies. This has led to many sequencing initiatives, such as the European Reference Genome Atlas (ERGA) and the Darwin Tree of Life (DToL), that aim to produce high quality reference genomes for a variety of organisms (Blaxter *et al.*, 2022). One recent technological advance is third generation sequencing or high-throughput sequencing of long molecules, providing the core of many of these assemblies. With read length N50's > 19 kbp, many genomes are being released with high contiguity compared to the early 2000's that were based on second generation short read sequencing technologies. Similarly, Hi-C techniques have further propelled these genomes into the realm of chromosome-scale assemblies, a level previously only held by model organisms. Coupled with these advances is also a decrease in price per base pair, making sequencing available to many laboratories and research groups. Combining the technologies of third generation and Hi-C sequencing, it is feasible to assemble reference genomes also for non-model organisms (Lawniczak *et al*, 2022). Here we apply these two technologies to produce a high-quality reference genome of the brown hare (*Lepus europaeus*).

While there is no existing genome assembly for the brown hare, a mountain hare assembly (GCA_009760805) exists, representing the Irish subspecies *L. t. hibernicus* and is based on a female specimen, thus lacking the Y-chromosome. This assembly is a so-called "pseudoreference", scaffolded using the rabbit (*Oryctolagus cuniculus* L.) reference genome

(Marques *et al*, 2020). The original rabbit reference genome had been established using whole-genome shotgun sequencing of females representing the inbred Thorbecke New Zealand White line and was quite fragmented (Carneiro *et al*, 2014), updated only recently with long-read sequencing data (Bai *et al*, 2021). In addition, the diploid chromosome number in domestic rabbits is 2n = 44 (Korstanje *et al*, 1999), whereas it is 48 in brown hare and mountain hare (Gustavsson, 1972).

To obtain a reference genome conforming to current standards, we used a fibroblast cell line (LE1) of a male specimen from Liperi, Eastern Finland (Gaertner *et al.*, 2022), as a source of high-molecular weight (HMW) DNA and fresh RNA. Compared to solid tissues, fibroblasts are optimal for Hi-C, as the method was originally developed for cells growing in a monolayer (Lieberman-Aiden *et al*, 2009).

We assembled the genome of *L. europaeus* using 25X genome coverage of PacBio HiFi read data and then further scaffolded the assembly with Hi-C sequencing data. The assembled genome is 2.96 Gbp in length with a contig N50 of 58 Mbp and a scaffold N90 of 25. The genome has a high degree of completeness based on the BUSCO score, Complete: 96.1% [Single copy: 93.1%, Duplicated: 3.0%] Fragmented 0.8%, and Missing 2.9% based on the mammalia_odb10 database. As the scaffold N90 is 25, the genome assembly is near-chromosome scale as there is an expected 23+X+Y karyotype for this species. By using RNA-seq data as extrinsic evidence for the genome annotation, we were able to produce 19,906 gene models for this assembly. We trust that the produced genome will prove to be valuable resource for future studies with hares, as well as provide interesting insight into the genome evolution among Glires.

Methods

Sampling and confirming of the species identity

A young male brown hare was hunted in October 2018 in Kuorinka, Liperi, Finland (62.6207 N, 29.4478 E, Figure 1B). The collection location is agricultural area with a mosaic of fields, shrubs and mixed forest with a strong brown hare population and only occasional mountain hares. In Finland, brown hares are highly dependent on similar man-made environments (Levanen *et al.*, 2018a; Levänen *et al*, 2019). The sampled population is sympatric with mountain hares, as is the case for almost all Finnish brown hares. However, the population represents a geographic continuum through Russian Karelia and the Baltic states to Poland, making our specimen more representative of the suggested type population sampled by Pallas, compared to western European brown hares. Although a trivial point, it is recommendable that the species reference genome would represent or be closely related to the type locality population (Lawniczak *et al.*, 2022). As pointed out earlier, Pallas was also aware of hybrids between mountain hares and brown hares in Poland, making it unlikely to find eastern examples of brown hares devoid of ancestral admixture with mountain hares.

The sampling had minimal impact on the local brown hare population and no impact on the habitat. As brown hares are legal game animals in Finland and the hunting followed the regional hunting seasons and legislation (Metsästyslaki [Hunting law] 1993/615/5§), the sampling adheres to the ARRIVE guidelines and no ethical assessment for the sampling was required. The sampling also did not involve International Trade in Endangered Species of Wild Fauna and Flora (CITES) or other export of specimens, as defined by the Convention on Biological Diversity (CBD).

The species identity was confirmed at site based on the morphological features distinguishing the brown hare from the mountain hare, the only other hare species in northern Europe. Further analysis of the coding part of the genome and mitochondrial DNA haplotyping showed minimal ancestral admixture with mountain hares (Gaertner *et al.*, 2022). In fact, unlike many other brown hares in the region, showing adaptive introgression of mountain hare specific UCP1 alleles, the specimen is homozygous for the common ancestral brown hare allele, UCP02 (Pohjoismaki *et al.*, 2021).

Generation and vouchering of the cell line

The fibroblast cell line (LE1) was isolated from the specimen as described earlier (Gaertner *et al.*, 2022) and is deposited as cryopreserved living cells under voucher number ZFMK-TIS-69747 into the biobank of the Stiftung Leibniz-Institut zur Analyse des Biodiversitätswandels, Zoological Research Museum Alexander König (ZFMK), Bonn, Germany. Additional identifiers for the sample are mLepEur2 and ERGA specimen ID ERGA_FI_3610_002 (COPO portal, https://copo-project.org).

HMW DNA extraction and PacBio HiFi sequencing

A salting out-method was used to extract high molecular weight DNA from cells grown to confluency on a 10 cm cell culture dish. Briefly, the cells were detached from the dish using trypsin, followed by centrifugation and two washes with PBS. The salting out-method then followed the 10X Genomics "Salting Out Method for DNA extraction from cells" protocol. The high molecular weight DNA was quantified using the Qubit 3.0, followed by qualification with an 0.8 % Agarose gel. The DNA was analyzed by DNA Sequencing and Genomics Laboratory, Institute of Biotechnology, University of Helsinki on the PacBio Sequel II. Sequence data from two flow cells was produced.

Mitochondrial DNA (mtDNA) sequencing

The mitochondrial genomes of our hare cell lines (Gaertner *et al.*, 2022) have been sequenced for the purposes of another study using approximately 2 kb overlapping PCR-amplified fragments of mtDNA. The primers used to amplify the mountain hare mtDNA were as follows:

Le93F: TTGTTTTGTAGCAAGTTTACACATGC
Le184R: GCTTAATACCTGCTCCTCTTGATCTA
Le1580F: TTAAACCCATAGTTGGCCTAAAAGC
Le1635R: TTGAGCTTTAACGCTTTCTTAATTGA
Le3045F: AGGCGTATTATTTATCCTAGCAACCT
Le3175R: CCTCATAAGAAATGGTCTGTGCGA
Le3921F: CCCCCTAATCTTTTCCATCATCCTAT
Le4482R: TCATCCTATATGGGCAATTGAGGAAT
Le4689F: AGGCTTTATTCCAAAGTGAATTATTATTCA
Le5417R: AGGCTCCAAATAAAAGGTAGAGAGTT
Le6696F: ATACCGTCTCATCAATAGGCTCCTTC
Le6756R: ATAAAGATTATTACTATTACAGCGGTTAGA
Le8603F: AGCCTATATCTACATGATAATACTTAATGA
Le8698R: CGGATAAGGCCCCGGTAAGTGG

Le10552F: TTGAAGCAACACTAATCCCTACACTA
Le10613R: TCGTTCTGTTTGATTACCTCATCGT
Le11301F: ACCATTAACCTTCTAGGAGAGCTTCT
Le11807R: AGGATAATGATTGAGACGGCTATTGA
Le12407F: GTCTAATCCTAGCTGCTACAGGTAAG
Le12791R: GAGCATAAAAAGAGTATAGCTTTGAA
Le14204F: ATTGTTAACCACTCTCTAATCGACCT
Le14514R: CCAATGTTTCAGGTTTCTAGGTAAGT
Lt16056F: TGGGGTATGCTTGGACTCAAC
Le16119R: TCGTCTACAATAAGTGCACCGG

In total, 12 separate reactions were prepared to cover the mitochondria genome:

1. Lt16056F + Le184R: 1871 bp
2. Le93F + Le1635R: 1543 bp
3. Le1580F + Le3175R: 1596 bp
4. Le3045F + Le4482R: 1438 bp
5. Le3921F + Le5417R: 1497 bp
6. Le4689F + Le6756R: 2068 bp
7. Le6696F + Le8698R: 2003 bp
8. Le8603F + Le10613R: 2011 bp
9. Le10552F + Le11807R: 1256 bp
10. Le11301F + Le12791R: 1491 bp
11. Le12407F + Le14514R: 2108 bp
12. Le14204F + Le16119R: 1916 bp

(Expected fragment size based on the published *Lepus europaeus* mtDNA sequence from Sweden [NC_004028.1]).

The fragments were amplified from total DNA preparations using a PCR program with a 1 min 94 °C denaturing step, followed by 35 cycles of 94 °C for 15 s, 56 °C for 15 s and 72 °C for 2 min and a final 3 min elongation step at 72 °C. The obtained products were gel purified using the GeneJET gel extraction kit (Thermo Scientific) and sent for sequencing using Illumina MiniSeq at the Genome Center of Eastern Finland.

The sequence of the non-coding region-containing PCR fragment (Lt16056F + Lt184R) was further validated by Sanger sequencing, applying also the following additional primer:
Le101F: TATAAATTCCTGCCAAACCCCAAAAA

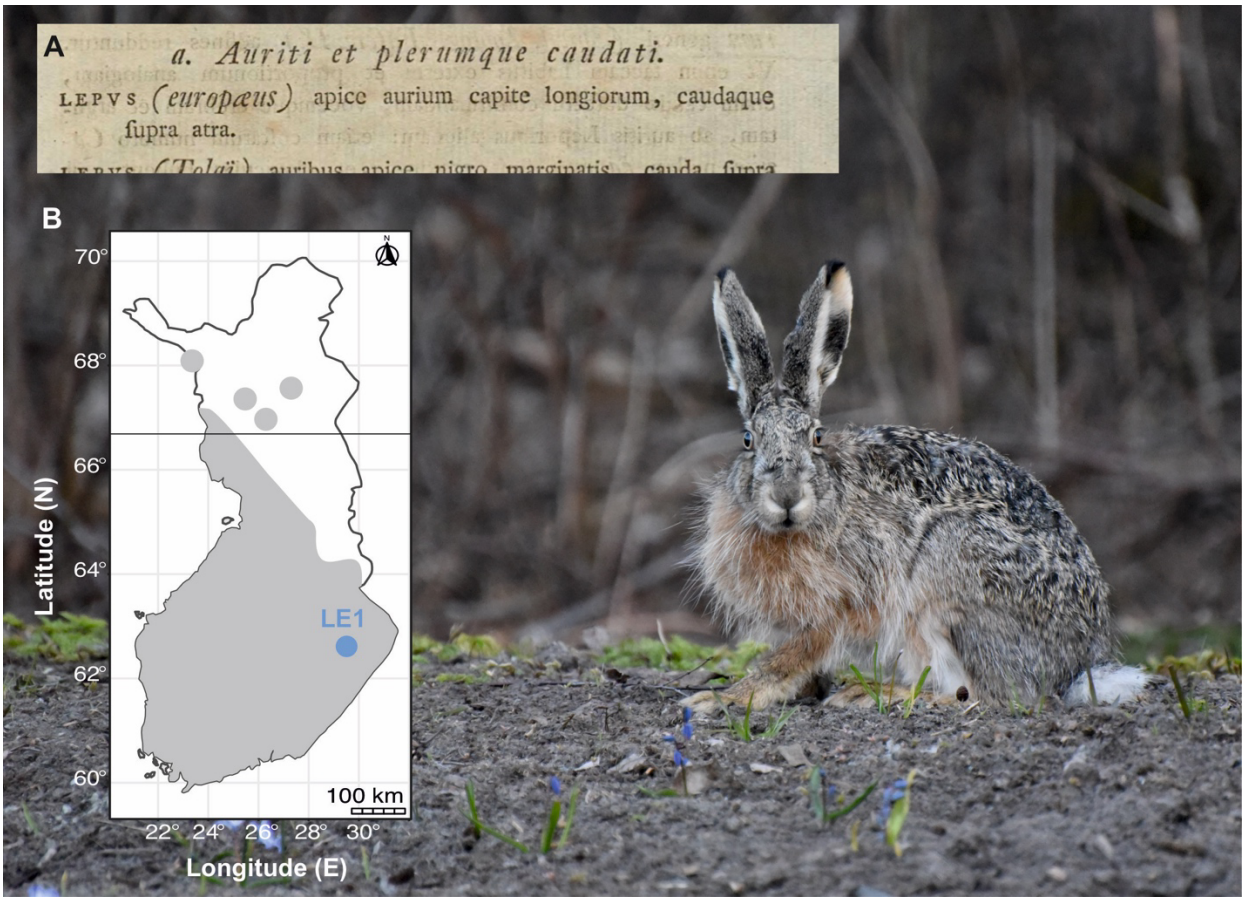

Figure 1. The brown hare. (A) An excerpt from the identification table by Pallas (1778: 30), which has been recognized as the species description for *Lepus europaeus*. "Tips of the ears longer than the head, tail black above." (translation from Latin). (B) The geographic location of the LE1 sample used in this study. The grey-shaded area shows the approximate current distribution of the brown hare in Finland, based on the data in the Finnish Biodiversity Information Facility (FinBIF, https://laji.fi/en). Individual records from the north are from settlemens or towns. The arctic circle (black line) runs along the 66.6 ° parallel. Background photo shows a typical Finnish male brown hare. Photo taken in Joensuu, 20 km East from the LE1 sampling site by Dr. Mervi Kunnasranta.

Hi-C library preparation

Hi-C sequencing libraries were prepared following the protocol of (Belaghzal *et al*, 2017) with the following changes: 1.) To prepare a diverse sequencing library, we performed the Hi-C protocol in triplicates, 2.) Size fractionation was performed using the NEBNext Ultra II FS DNA module, 3.) Library preparation was performed using the NEBNext Ultra II library preparation kit for Illumina and 4.) Library enrichment was performed using triplicate PCR reactions with six cycles of PCR. The PCR reactions were then purified using Ampure XP beads at a ratio of 0.9X. The final clean libraries were quantified using Qbit, followed by agarose gel electrophoresis to confirm the fragment size. The sequencing was performed on a single lane of the Illumina NovaSeq 6000 using the SP flowcell with paired-end chemistry 2 x 150bp.

Genome assembly

HiFiasm version 0.18.7 (Cheng *et al*, 2021) was used to assemble the 25X (rq >0.99) PacBio HiFi reads using the arguments *-l3 –h1 and –h2 –primary* to integrate the Hi-C read data and produce a primary assembly. We then continued with the scaffolding of the primary assembly and annotation. To process the HiC data, we first mapped the Hi-C data to the primary genome assembly using BWA-mem version 0.7.17 (Li & Durbin, 2009) with the arguments *-SP*. The mapped reads were then parsed and filtered using pairtools version 1.0.2 (Open2C *et al*, 2023). To parse the aligned Hi-C reads, we used the options *–no-flip –dd-columns mapq –walks-policy mask*. The parsed pair data was then sorted and deduplicated using default arguments. Finally, we selected unique-unique and unique-rescued pairs and split these into the pairs file and bam file for input in YaHS version 1.1 (Zhou *et al*, 2023). YAHS was run using the default parameters with the primary contig assembly and the filtered Hi-C bam file. Contiguity and general genome statistics were calculated using QUAST version 5.2.0 (Mikheenko *et al*, 2018). We assessed the completeness of the genome by calculating the number of complete single copy orthologs with BUSCO version 5.1.2 (Manni *et al*, 2021), using the mammalia_odb10 database as well as the more lineage specific glires_odb10 database.

Genome annotation

Repeat annotation of the genome was performed with EDTA version 2.1.0 (Ou *et al*, 2019), a *de novo* repeat identification pipeline. Using the repeat library produced by EDTA, we masked the scaffolded genome using RepeatMasker version 4.1.1 (Smit *et al*, 2013-2015). RNA-seq data from the same cell line (SRA accession number: SRR18740842) as well as RNA-seq data from other *L. timidus* libraries (SRA accession number: SRR10020054, SRR10020055, SRR10020060, SRR10491719, SRR18740839, SRR18740840, and SRR18740841) was collected from the sequence read archive (SRA). The RNA-seq data was trimmed using fastp version 0.23.2 (Chen *et al*, 2018) and mapped against the masked genome using HISAT2 version 2.1.0 (Kim *et al*, 2019) with the default parameters. Furthermore, we included the protein sequences of the assembled transcripts from this cell line as further evidence for the genome annotation. These lines of gene evidence were included in the annotation using BRAKER3 version 3.02 (Bruna *et al*, 2021). Telomeric sequences [AACCCT]n were identified using a Telomere Identification toolkit (tidk) version 0.2.31. The telomeric sequence copy number was then calculated in windows of 200kb for visualization using Circos (Krzywinski *et al*, 2009).

Mitochondrial DNA assembly and annotation

Mitochondrial DNA was assembled from the PCR-amplified and Illumina sequenced mtDNA using the MitoZ pipeline (Meng *et al*, 2019). After comparing the results of the pipeline's outputs with different kmer options, we selected the best assembly as final. Run options used in the final assembly were --clade Chordata –fastq_read_length 150, --requiring_taxa Chordata --genetic_code 2 --kmers_megahit 21 29 39 59 79 99 119 141. The tools invoked by the pipeline included fastp (Chen *et al.*, 2018) for cleaning the raw data, MEGAHIT (Li *et al*, 2015) for assembly, after which sequences were filtered to ensure the correct taxa by HMMER (Wheeler & Eddy, 2013) and further filtered for completeness of protein-coding genes. Annotation steps were done using TBLASTN (Gertz *et al*, 2006), GeneWise (Birney *et al*, 2004) and MiTFi (Juhling *et al*, 2012). Final manual curation, the annotation of the non-coding region (NCR) as well as the illustration of the

316 mitochondrial genome was done using Geneious® 10.2.6. The functional loci on the NCR were
317 identified based on the similarity with the human (NC_012920) and mouse (FJ374652) NCR
318 sequences.
319
320 **Manual curation**
321 The assembled and annotated genome was manually curated to further improve its
322 quality as described in (Howe *et al*, 2021).
323
324 **Comparison with previous assembly**
325 We performed a comparison of our scaffolded assembly to the current *L. timidus* genome
326 assembly (NCBI Accession number: GCA_009760805). Mapping to the genome was performed
327 using minimap2 version 2.21 (Li, 2018) with the arguments *-asm5*. A dot plot of the alignment
328 was created using the R script pafCoordsDotPlotly.R (https://github.com/tpoorten/dotPlotly).
329

330 # Results

331
332 The genome assemblies can be accessed via BioSample accession SAMN33984520 as well as
333 BioProject accession PRJNA1009711 for the primary assembly and PRJNA1009710 for the
334 alternative assembly.
335
336 **Genome assembly**
337 The expected haploid genome size of the *L. europaeus* is similar to *L. timidus*, which has
338 been published to be 3.25 pg (3,178,500,000 bp) (Vinogradov, 1998), containing 23 autosomal
339 chromosomes and two sex chromosomes (Gustavsson, 1972). PacBio HiFi sequencing with two
340 flow cells resulted in 25 X coverage of the expected genome size. The sequence N50 of the HiFi
341 data was 19.97 Kb. Based on the PacBio HiFi data, the resulting genome size using k-mer (K = 21)
342 coverage is 2.4 Gbp (Figure 2A). The illumina sequencing of the Hi-C data produced 494,285,585
343 paired reads, representing about 51 X coverage of the genome. The duplication rate of the Hi-C
344 data was 17 %. Assembly with HiFiasm yielded a contig assembly of 2.96 Gbp made up of 671
345 contigs with a contig N50 of 58 Mbp. The longest contig was 149 Mbp. Using the uniquely mapped
346 Hi-C data, we were able to scaffold the contigs and fix misassembled contigs. The Hi-C scaffolded
347 assembly (Figure 2B) was 2.96 Gbp in size, similar to the contig assembly, while the largest scaffold
348 was 149 Mbp. The scaffold N50 was 124 Mbp (N=10) and the scaffold N90 was 21 Mbp (N=25).
349 The BUSCO scores of the *L. europaeus* assembly suggest it to be near-complete, with the following
350 results: Complete: 96.1% [Single copy: 93.1 %, Duplicated: 3.0 %] Fragmented 0.8 %, and Missing
351 2.9 % based on the mammalia_odb10 database.
352 To further improve the assembly, manual curation was performed using Hi-C maps
353 (Table 1), resulting in 76 scaffold breaks, 81 joins and removal of 62 erroneously duplicated
354 contigs. These interventions led to an increase in scaffold N50 of 1.69 % from 123.7 Mb to
355 125.8 Mb, and a reduction in scaffold count of 9.21 % from 716 to 650. Of the finalized assembly,
356 93.16 % could be assigned to 25 identified chromosomes (23 autosomes plus X and Y), matching
357 the published karyotype. Chromosomes were numbered according to size. The curated genome
358 length was 2,930,972,003 bp.

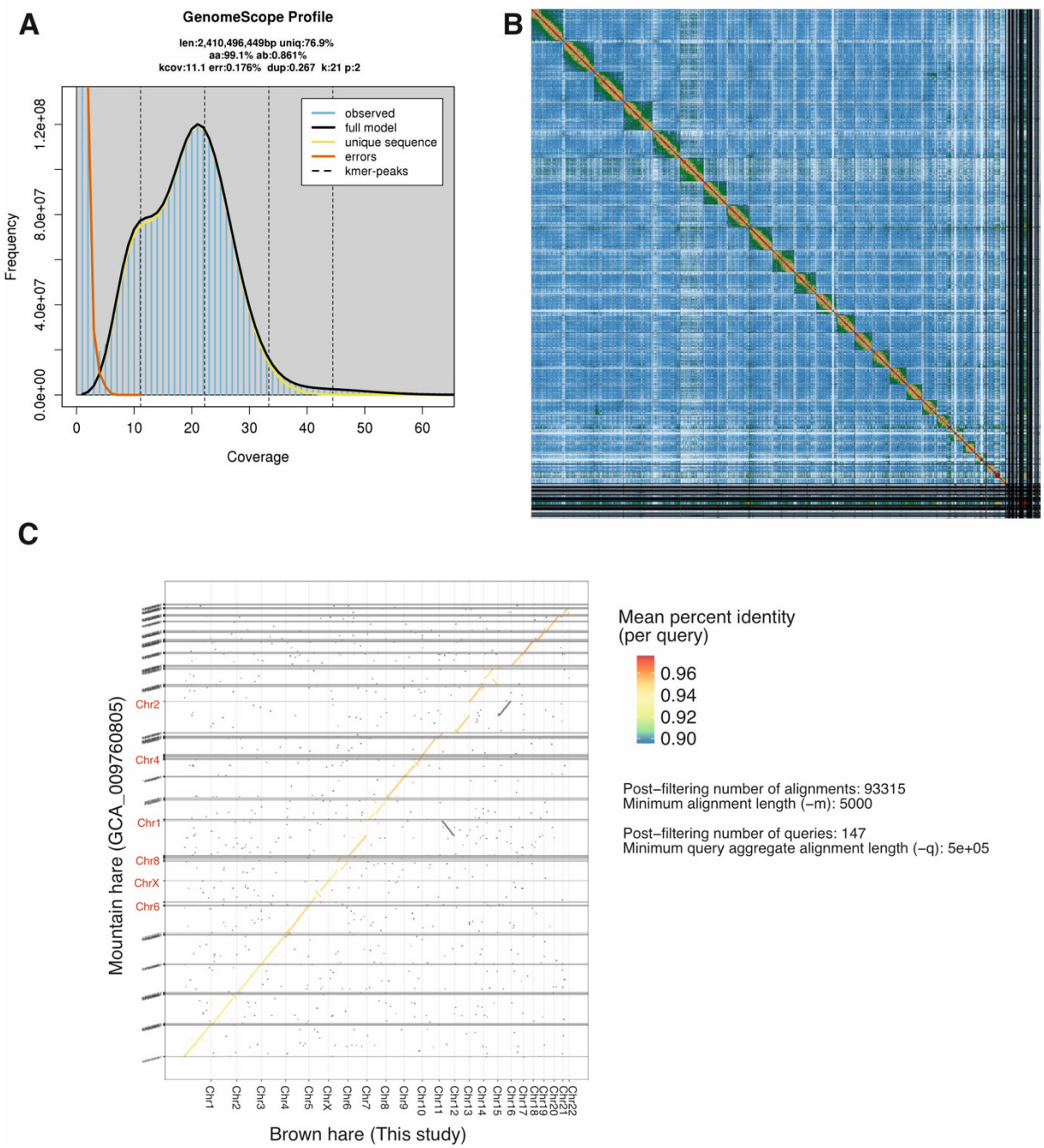

Figure 2. A) Genomescope2 profile of the PacBio HiFi data. B) Pretext map of Hi-C scaffolding after manual curation. C) Dot plot comparison with the previous assembly. Some examples of the previous chromosome assemblies are highlighted.

Table 1. The brown hare reference genome chromosome assignment and assembly statistics

Chromosome assignment	
Chr length	2,730,543,680
Chr length %	93.16
Autosomes	23

X	1
Y	1
Unlocalised	19

Statistics for primary assembly after manual curation	
# scaffolds	650
Total scaffold length	2,930,972,003
Average scaffold length	4,509,187.70
Scaffold N50	125,776,599
Scaffold auN	117,151,857.99
Scaffold L50	10
Largest scaffold	184,151,585
Smallest scaffold	1000
# contigs	1068
Total contig length	2,930,889,045
Average contig length	2,744,278.13
Contig N50	43,409,466
Contig auN	47,572,960.95
Contig L50	21
Largest contig	141,329,000
Smallest contig	1000
# gaps in scaffolds	418
Total gap length in scaffolds	82,958
Average gap length in scaffolds	198.46
Gap N50 in scaffolds	200
Gap auN in scaffolds	199.70
Gap L50 in scaffolds	208
Largest gap in scaffolds	200
Smallest gap in scaffolds	31
Base composition	
A	821,587,687
C	645,897,412
G	644,604,632
T	818,799,314
GC content	44.03 %
# soft-masked bases	0
# segments	1,068
Total segment length	2,930,889,045

Average segment length	2,744,278.13
# gaps	418
# paths	650

365
366
367 **Genome annotation**
368 EDTA produced a curated custom repeat library of 7,045 repetitive elements. Interestingly,
369 the proportion of the genome masked as repetitive elements was higher than expected at 46.8 %
370 of the genome. In the previous genome assembly, only 28 % of the genome was masked, and
371 similarly the *k*-mer based estimate of the repetitive elements was 26.7 %. To annotate the
372 genome, we mapped ~1.6 billion RNA-seq reads from previously published RNA-seq libraries,
373 including this cell line. The average mapping rate of all libraries was 89.95 %, which we consider
374 a good indicator of the quality of this genome. Using BRAKER3, we were able to annotate 19,906
375 gene models in the genome. We then ran BUSCO on the predicted genes and the results show a
376 good level of completeness: Complete: 81.1 % [Single copy: 63.3 %, Duplicate: 17.8 %],
377 Fragmented: 1.1 %, Missing: 17.8 %, Total: 9226 when compared to the mammalia_odb10
378 database. Telomeric sequences on both ends of the chromosome were detected on 13
379 chromosomes (Figure 3A). Furthermore, telomeric sequences were found in high copy number
380 throughout the chromosome length, a feature previously noted using FISH in mountain hares
381 (Forsyth *et al*, 2005).
382
383 **Comparison to the mountain hare assembly**
384 Minimap2 was able to align 97.3 % of the contig sequences from the existing mountain
385 hare genome assembly (GCA_009760805) to this genome (Figure 2C). As pointed out earlier, the
386 mountain hare genome assembly has been scaffolded against the rabbit reference genome (Giska
387 *et al*, 2022), which has a different chromosome count, among other differences. Despite these
388 chromosome assembly and species differences, we have a high level of sequence similarity with
389 the mountain hare genome assembly as well as a high degree of synteny. However, there a few
390 minor notable inversions and structural differences between the two genomes (Figure 2C). These
391 differences are mostly explained by the fact that the previous genome assembly retains the
392 chromosomal arrangement of the rabbit genome (Beklemisheva *et al*, 2011). For example, in
393 assembly presented in this study, chr7 and chr 12 have a good degree of synteny with the Chr1
394 (Chr1 - Rabbit) of the mountain hare assembly, demonstrating that the rabbit chromosome 1 is a
395 fusion of two ancestral lagomorph chromosomes. Similarly, also the mountain hare assembly
396 Chr2 – Rabbit is a fused chromosome and has good synteny and orientation with chr13 and chr16
397 in this genome. This exemplifies the importance of high-quality reference genomes to analyze
398 chromosome evolution across lineages and opens doors for future studies of chromosomal
399 evolution among mammals.
400
401 **Mitochondrial genome**
402 While it is possible to recover the entire mitochondrial genome from the HiFi sequencing
403 data (Uliano-Silva *et al*, 2023), we had already previously sequenced the mitochondrial genomes
404 of our hare cell lines for the purposes of another study. The assembly of the Illumina data

provided a 16,836 bp mtDNA sequence, which was complete and circular, with a 142 bp non-repetitive overlap between the 3´ and 5´ ends of the original assembly removed from the final version. All the expected vertebrate mitochondrial genes, 37 in total, were found (Figure 3B).

Interestingly, the primary assembly was notably smaller than expected based on the original PCR amplicons, with a notable difference between the expected and obtained non-coding region (NCR) length. Sanger sequencing of the PCR-fragment containing the NCR (Lt16056F + Le184R) revealed that it contained a sequence run consisting of six head-to-tail repeats of a 140 bp element between OriH and tRNA-Phe, explaining the incorrect assembly of the NCR sequence from the short-read Illumina data. With the addition of the repeat elements, the final mitochondrial genome size of our specimen is 17,540 bp.

Based on the sequence similarity with human and mouse mitochondrial genomes, we were able to tentatively identify the termination associated sequence (TAS, 2 nt difference to human), required for the generation of the 7S DNA and thus the displacement loop (D-loop) structure on the NCR (Crews *et al*, 1979). Interestingly, assuming that the 7S DNA is roughly the same size as in other mammals (600-650 nt), its 5'-end would map to the beginning of another highly repetitive 240 bp region, likely to have structural significance. This is a notably longer sequence run than the tRNA-like sequence reported at the replication origin of *L. timidus* mtDNA (Melo-Ferreira *et al*, 2014). We have preliminary assigned this structure to correspond to the heavy-strand origin of replication (OriH), supported by the fact that a similar, tRNA sequence-derived hairpin structure is required for the priming of the origin of the light-strand replication (OriL) (Fuste *et al*, 2010). Experimental validation of the proposed OriH is required.

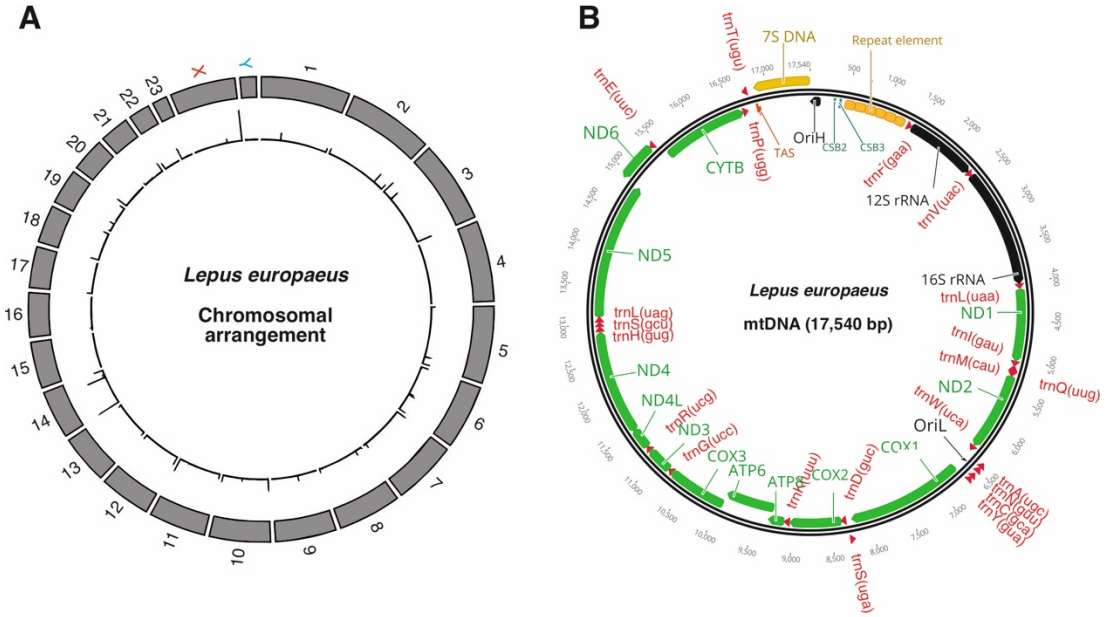

Figure 3. A reference genome for the brown hare. A) A size map of 23 autosomes as well as X and Y sex chromosomes of *Lepus europaeus*. Telomeric sequences and their relative lengths are indicated on the inner circle. B) A schematic illustration of the mountain hare mitochondrial genome. Genes encoded by the H-strand are placed inside and the L-strand genes outside of the circle. NCR = non-coding region.

434 # Discussion

435

436 Given the challenges of assembling a large mammalian genome, we have produced a high-quality
437 genome of the brown hare. The genome has a high degree of contiguity (Contig N50 43 Mbp) and
438 completeness (BUSCO complete 96.1 %). Interestingly, the genome contains a high number of
439 repetitive elements, a fact that warrants further investigation to elucidate their identity. The
440 genome is in chromosome scale, and all 23 autosomes as well as X and Y chromosomes could be
441 scaffolded. Compared with the previously published mountain hare genome, the brown hare
442 genome is much more contiguous and also resolves chromosome structure including telomeric
443 repeats on both ends of several chromosomes (Figure 3A).

444 Together, this genome and the mountain hare genome will provide a solid base for future
445 studies. Having a high-quality genome for both species has important implications for the fields
446 of evolutionary and conservation biology. Due to the high contiguity of the genome, it will be
447 possible to perform accurate analyses of linkage disequilibrium as well as the identification of
448 runs of homozygosity (ROH). Accurate ROH analysis is useful for identifying and understanding
449 the impact of inbreeding burden on the genome of increasingly threatened and isolated species,
450 as quantitative traits can be associated with ROH burden (Ceballos *et al*, 2018). Furthermore, this
451 information might provide important insights into the hybridization of the brown hare with the
452 mountain hare, as well as help to pinpoint genomic regions that might be helping the species to
453 adapt to local environmental conditions during the ongoing range expansion (Pohjoismaki *et al.*,
454 2021).

455 The obtained 17,540 bp long mitochondrial genome also provides some interesting
456 aspects of mtDNA sequence variation within mammalian species. The size difference to human
457 (16,569 bp) and mouse (16,298 bp) mitochondrial genomes is caused by a rather long non-coding
458 region (2,102 bp vs. 1,123 bp and 879 bp in humans and mice, respectively), longer rRNA and
459 tRNA genes, as well as additional non-coding nucleotides between genes. Interestingly, the
460 mitochondrial genome of our specimen is slightly smaller than the 17,734 bp long previous NCBI
461 Reference Sequence for brown hare (NC_004028.1) from Sweden. The main size difference can
462 be attributed to a copy number difference in the repeat elements within the non-coding region,
463 with seven in the previous reference mtDNA vs. six in our specimen. Indels elsewhere in the NCR
464 explain why the length difference does not add up to simple multiplies of 140 bp repeat element.
465 Repeat length polymorphisms seem to be a common theme in hare mtDNA and we have noted
466 similar within-species size variation in the NCR also among Finnish mountain hares, with some
467 evidence of these variants occurring in multiple heteroplasmy within an individual (a manuscript
468 in preparation). To our understanding, such within-population and within-individual length
469 variation in the non-coding region is unusual and warrants for further investigation.

470 The main technical challenge for highly continuous, good quality reference genomes Is the
471 availability of relatively large quantities of intact HMW DNA for HiFi sequencing. This typically
472 requires snap freezing of tissue samples and assuring an intact cold chain of below -80 °C for their
473 sending, handling, and long-term storage. While development in sequencing technologies can
474 reduce the required DNA amounts, obtaining high sequencing coverage over a large vertebrate
475 genome is easier when material is plentiful. Similarly, RNA is very sensitive to degradation in post-
476 mortem tissue samples, also requiring immediate preparation or snap-freezing to be suitable for

RNA-seq. Although transcript data is not obligatory, it is highly useful for the purpose of genome annotation. The requirement of large chunks of fresh tissue for -omics purposes complicates the sampling of most vertebrates species. For example, in Finland all land vertebrates are protected by law, with very few exceptions, such as game animals and pest species. Capturing and lethal sampling is especially problematic for endangered species. Not only will the sampling require extensive permits, but it is also difficult to justify and conduct ethically.

Although the source specimen for the cell line utilized in this study was hunted, fibroblasts can be isolated from relatively small skin biopsies, such as ear clippings (Seluanov *et al*, 2010), with minimally invasive sampling and harm to the individual. These cells can be expanded in tissue culture and, when immortalized, provide a highly scalable source of fresh DNA and RNA. Furthermore, living cells can be stored as frozen stocks in biobanks for decades (Fazekas *et al*, 2017). Tissue sampling can be done directly into the growth medium in ambient temperature, allowing several days for the transport to the laboratory without a need for dry ice or N_2-cryocontainers, which greatly simplifies collection and logistics. The same applies to the isolated cells. As an example, we sent living cell suspensions from Joensuu, Finland to the ZFMK biobank in Bonn, Germany, over regular airmail in ambient temperature. The receiving laboratory can then amplify the cells by culturing and cryopreserve them in sufficient replicate stocks. It is worth to point out that the obtained genome scaffolded to 23 autosomes, X and Y chromosomes, as expected for the species, shows no evidence of karyotypic drift caused by cell culture, dispelling possible concerns for the use of cell lines for such work (Gaertner *et al.*, 2022).

The brown hare is an iconic European mammal, locally familiar to a wider public of nature-goers, farmers and hunters, while being at the same time an invasive species on many other continents. It has populations with complex ancestral makeup due to isolated Pleistocene refugia, frequent introductions and cross-species hybridization, making it an exceptionally interesting mammal for genetic studies. Furthermore, it can be both an endangered, as well as a rapidly expanding species, depending on the local environmental context, prompting opportunities for both conservation as well as invasion genetics studies. A high-quality reference genome will allow us also to peer deeper into genome structure and chromosome evolution, which will be particularly interesting from the viewpoint of genetic introgression (Juric *et al*, 2016) and maintenance of species boundaries. Any issues of genetic compatibility between species should leave their fingerprint in the genomes (Skov *et al*, 2023), providing interesting insight into the speciation mechanisms (Wolf *et al*, 2010). Other *Lepus* reference genomes available soon will help to complete this picture.

Acknowledgements

We acknowledge the DNA Sequencing and Genomics Laboratory, Institute of Biotechnology, University of Helsinki for the PacBio and Hi-C sequencing, with special thanks to Dr Petri Auvinen and Dr Martyn James for helpful discussions and personal interest towards the project. Dr Seppo Helisalmi and Dr Joose Raivo from the Gene diagnostics lab at Genome Center of Eastern Finland are thanked for providing the mtDNA sequencing service. We are grateful to the European Reference Genome Atlas (ERGA) for the opportunity to include our mountain hare genome into their pilot project. Dr Ann McCartney is thanked for her kind help throughout the process and Mr Jukka Pusa for his contribution in obtaining the brown hare specimen. We thank Ms Anita

Kervinen for her valuable laboratory assistance and Dr. Mervi Kunnasranta for kindly allowing to use her hare photo in the publication.

Data, scripts, code, and supplementary information availability

The genome assembly is available from the NCBI database under BioProject ID PRJNA950335. The cell line is available through the corresponding author as well as from the ZFMK biobank.

Conflict of interest disclosure

The authors declare that they comply with the PCI rule of having no financial conflicts of interest in relation to the content of the article.

Funding

This study belongs to the xHARES consortium funded by the R'Life initiative of the Academy of Finland, grant no. 329264.

References

Bai Y, Lin W, Xu J, Song J, Yang D, Chen YE, Li L, Li Y, Wang Z, Zhang J (2021) Improving the genome assembly of rabbits with long-read sequencing. *Genomics* 113: 3216-3223

Beklemisheva VR, Romanenko SA, Biltueva LS, Trifonov VA, Vorobieva NV, Serdukova NA, Rubtsova NV, Brandler OV, O'Brien PC, Yang F *et al* (2011) Reconstruction of karyotype evolution in core Glires. I. The genome homology revealed by comparative chromosome painting. *Chromosome Res* 19: 549-565

Belaghzal H, Dekker J, Gibcus JH (2017) Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods* 123: 56-65

Birney E, Clamp M, Durbin R (2004) GeneWise and Genomewise. *Genome Res* 14: 988-995

Blaxter M, Archibald JM, Childers AK, Coddington JA, Crandall KA, Di Palma F, Durbin R, Edwards SV, Graves JAM, Hackett KJ *et al* (2022) Why sequence all eukaryotes? *P Natl Acad Sci USA* 119

Bock A (2020) Lepus europaeus (Lagomorpha: Leporidae). *Mammalian Species* 52: 125-142

Bruna T, Hoff KJ, Lomsadze A, Stanke M, Borodovsky M (2021) BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genom Bioinform* 3: lqaa108

Carneiro M, Rubin CJ, Di Palma F, Albert FW, Alfoldi J, Martinez Barrio A, Pielberg G, Rafati N, Sayyab S, Turner-Maier J *et al* (2014) Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication. *Science* 345: 1074-1079

Ceballos FC, Joshi PK, Clark DW, Ramsay M, Wilson JF (2018) Runs of homozygosity: windows into population history and trait architecture. *Nat Rev Genet* 19: 220-234

Chen SF, Zhou YQ, Chen YR, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34: 884-890

558 Cheng H, Concepcion GT, Feng X, Zhang H, Li H (2021) Haplotype-resolved de novo assembly using
559 phased assembly graphs with hifiasm. *Nat Methods* 18: 170-175
560 Crews S, Ojala D, Posakony J, Nishiguchi J, Attardi G (1979) Nucleotide sequence of a region of
561 human mitochondrial DNA containing the precisely identified origin of replication. *Nature*
562 277: 192-198
563 Fazekas J, Grunt TW, Jensen-Jarolim E, Singer J (2017) Long term storage in liquid nitrogen leads
564 to only minor phenotypic and gene expression changes in the mammary carcinoma model
565 cell line BT474. *Oncotarget* 8: 35076-35087
566 Ferreira MS, Jones MR, Callahan CM, Farelo L, Tolesa Z, Suchentrunk F, Boursot P, Mills LS, Alves
567 PC, Good JM *et al* (2021) The Legacy of Recurrent Introgression during the Radiation of
568 Hares. *Syst Biol* 70: 593-607
569 Fickel J, Hauffe HC, Pecchioli E, Soriguer R, Vapa L, Pitra C (2008) Cladogenesis of the European
570 brown hare (Lepus europaeus Pallas, 1778). *Eur J Wildlife Res* 54: 495-510
571 Forsyth NR, Elder FF, Shay JW, Wright WE (2005) Lagomorphs (rabbits, pikas and hares) do not
572 use telomere-directed replicative aging in vitro. *Mech Ageing Dev* 126: 685-691
573 Fuste JM, Wanrooij S, Jemt E, Granycome CE, Cluett TJ, Shi Y, Atanassova N, Holt IJ, Gustafsson
574 CM, Falkenberg M (2010) Mitochondrial RNA polymerase is needed for activation of the
575 origin of light-strand DNA replication. *Mol Cell* 37: 67-78
576 Gaertner K, Michell C, Tapanainen R, Goffart S, Saari S, Soininmaki M, Dufour E, Pohjoismaki JLO
577 (2022) Molecular phenotyping uncovers differences in basic housekeeping functions among
578 closely related species of hares (Lepus spp., Lagomorpha: Leporidae). *Mol Ecol*
579 Gertz EM, Yu YK, Agarwala R, Schaffer AA, Altschul SF (2006) Composition-based statistics and
580 translated nucleotide searches: improving the TBLASTN module of BLAST. *BMC Biol* 4: 41
581 Giska I, Pimenta J, Farelo L, Boursot P, Hacklander K, Jenny H, Reid N, Montgomery WI, Prodohl
582 PA, Alves PC *et al* (2022) The evolutionary pathways for local adaptation in mountain hares.
583 *Mol Ecol* 31: 1487-1503
584 Gustavsson I (1972) Mitotic and meiotic chromosomes of the variable hare (Lepus timidus L.), the
585 common hare (Lepus europaeus Pall.) and their hybrids. *Hereditas* 67: 27-34
586 Holden ME (2005) Family Gliridae. In: *Mammal Species of the World: A Taxonomic and Geographic*
587 *Reference*, Wilson D.E., Reeder D.M. (eds.) pp. 819-841. Johns Hopkins University Press:
588 Baltimore, USA
589 Howe K, Chow W, Collins J, Pelan S, Pointon D-L, Sims Y, Torrance J, Tracey A, Wood J (2021)
590 Significantly improving the quality of genome assemblies through curation. *GigaScience* 10
591 Juhling F, Putz J, Bernt M, Donath A, Middendorf M, Florentz C, Stadler PF (2012) Improved
592 systematic tRNA gene annotation allows new insights into the evolution of mitochondrial
593 tRNA structures and into the mechanisms of mitochondrial genome rearrangements.
594 *Nucleic Acids Res* 40: 2833-2845
595 Juric I, Aeschbacher S, Coop G (2016) The Strength of Selection against Neanderthal Introgression.
596 *PLoS Genet* 12: e1006340
597 Kim D, Paggi JM, Park C, Bennett C, Salzberg SL (2019) Graph-based genome alignment and
598 genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 37: 907-915
599 Korstanje R, O'Brien PC, Yang F, Rens W, Bosma AA, van Lith HA, van Zutphen LF, Ferguson-Smith
600 MA (1999) Complete homology maps of the rabbit (Oryctolagus cuniculus) and human by
601 reciprocal chromosome painting. *Cytogenet Cell Genet* 86: 317-322

602 Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009)
603 Circos: an information aesthetic for comparative genomics. *Genome Res* 19: 1639-1645
604 Lawniczak MKN, Durbin R, Flicek P, Lindblad-Toh K, Wei X, Archibald JM, Baker WJ, Belov K, Blaxter
605 ML, Marques Bonet T *et al* (2022) Standards recommendations for the Earth BioGenome
606 Project. *Proc Natl Acad Sci U S A* 119
607 Levanen R, Kunnasranta M, Pohjoismaki J (2018a) Mitochondrial DNA introgression at the
608 northern edge of the brown hare (Lepus europaeus) range. *Ann Zool Fenn* 55: 15-24
609 Levänen R, Pohjoismäki J, Kunnasranta M (2019) Home ranges of semi-urban brown hares (Lepus
610 europaeus) and mountain hares (Lepus timidus) at northern latitudes. *Ann Zool Fenn* 56:
611 107–120
612 Levanen R, Thulin CG, Spong G, Pohjoismaki JLO (2018b) Widespread introgression of mountain
613 hare genes into Fennoscandian brown hare populations. *Plos One* 13
614 Li D, Liu CM, Luo R, Sadakane K, Lam TW (2015) MEGAHIT: an ultra-fast single-node solution for
615 large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31:
616 1674-1676
617 Li H (2018) Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34: 3094-
618 3100
619 Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform.
620 *Bioinformatics* 25: 1754-1760
621 Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR,
622 Sabo PJ, Dorschner MO *et al* (2009) Comprehensive mapping of long-range interactions
623 reveals folding principles of the human genome. *Science* 326: 289-293
624 Manni M, Berkeley MR, Seppey M, Zdobnov EM (2021) BUSCO: Assessing Genomic Data Quality
625 and Beyond. *Curr Protoc* 1: e323
626 Marques JP, Seixas FA, Farelo L, Callahan CM, Good JM, Montgomery WI, Reid N, Alves PC, Boursot
627 P, Melo-Ferreira J (2020) An Annotated Draft Genome of the Mountain Hare (Lepus timidus).
628 *Genome Biol Evol* 12: 3656-3662
629 Melo-Ferreira J, Vilela J, Fonseca MM, da Fonseca RR, Boursot P, Alves PC (2014) The elusive
630 nature of adaptive mitochondrial DNA evolution of an arctic lineage prone to frequent
631 introgression. *Genome Biol Evol* 6: 886-896
632 Meng G, Li Y, Yang C, Liu S (2019) MitoZ: a toolkit for animal mitochondrial genome assembly,
633 annotation and visualization. *Nucleic Acids Res* 47: e63
634 Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A (2018) Versatile genome assembly
635 evaluation with QUAST-LG. *Bioinformatics* 34: i142-i150
636 Minoudi S, Papapetridis I, Karaiskou N, Chatzinikos E, Triantaphyllidis C, Abatzopoulos TJ,
637 Triantafyllidis A (2018) Genetic analyses of brown hare (Lepus europaeus) support limited
638 migration and translocation of Greek populations. *Plos One* 13: e0206327
639 Ognev SI (1940) 157. Lepus (Eulagus) europeaus Pall. (1778) заяц-русак. In: *Mammals of the USSR*
640 *and adjacent countries: rodents (In Russian)*, pp. 112-161. Akademiya Nauk SSSR: Moscow
641 Open2C, Abdennur N, Fudenberg G, Flyamer IM, Galitsyna AA, Goloborodko A, Imakaev M, Venev
642 SV (2023) Pairtools: from sequencing data to chromosome contacts. *bioRxiv*
643 Ou S, Su W, Liao Y, Chougule K, Agda JRA, Hellinga AJ, Lugo CSB, Elliott TA, Ware D, Peterson T *et*
644 *al* (2019) Benchmarking transposable element annotation methods for creation of a
645 streamlined, comprehensive pipeline. *Genome Biol* 20: 275

646 Pallas PS (1778) *Novae species quadrupedum e glirium ordine cum illustrationibus variis*
647 *complurium ex hoc ordine animalium*. Wolfgang Walther
648 Petrovan S Lepus europaeus (European hare) (2013) [DATASET]
649 Pohjoismaki JLO, Michell C, Levanen R, Smith S (2021) Hybridization with mountain hares
650 increases the functional allelic repertoire in brown hares. *Sci Rep* 11: 15771
651 Reid N (2011) European hare (Lepus europaeus) invasion ecology: implication for the
652 conservation of the endemic Irish hare (Lepus timidus hibernicus). *Biol Invasions* 13: 559-
653 569
654 Reid N, Montgomery WI (2007) Is naturalisation of the brown hare in Ireland a threat to the
655 endemic Irish hare? *Biol Environ* 107b: 129-138
656 Schai-Braun SC, Weber D, Hacklander K (2013) Spring and autumn habitat preferences of active
657 European hares (Lepus europaeus) in an agricultural area with low hare density. *Eur J*
658 *Wildlife Res* 59: 387-397
659 Seluanov A, Vaidya A, Gorbunova V (2010) Establishing primary adult fibroblast cultures from
660 rodents. *J Vis Exp*
661 Skov L, Coll Macia M, Lucotte EA, Cavassim MIA, Castellano D, Schierup MH, Munch K (2023)
662 Extraordinary selection on the human X chromosome associated with archaic admixture.
663 *Cell Genom* 3: 100274
664 Smit AFA, Hubley R, Green P, 2013-2015. RepeatMasker Open-4.0. http://www.repeatmasker.org.
665 Smith RK, Jennings NV, Harris S (2005) A quantitative analysis of the abundance and demography
666 of European hares Lepus europaeus in relation to habitat type, intensity of agriculture and
667 climate. *Mammal Rev* 35: 1-24
668 Stott P (2003) Use of space by sympatric European hares (Lepus europaeus) and European rabbits
669 (Oryctolagus cuniculus) in Australia. *Mamm Biol* 68: 317-327
670 Thulin C-G (2003) The distribution of mountain hares (Lepus timidus, L. 1758) in Europe: A
671 challenge from brown hares (L. europaeus, Pall. 1778)? *Mammal Rev* 33: 29-42
672 Thulin C-G, Tegelström H (2002) Biased geographical distribution of mitochondrial DNA that
673 passed the species barrier from mountain hares to brown hares (genus Lepus): an effect of
674 genetic incompatibility and mating behaviour? *J Zool* 258: 299-306
675 Thulin CG, Winiger A, Tallian AG, Kindberg J (2021) Hunting harvest data in Sweden indicate
676 precipitous decline in the native mountain hare subspecies Lepus timidus sylvaticus (heath
677 hare). *J Nat Conserv* 64
678 Uliano-Silva M, Ferreira J, Krasheninnikova K, Darwin Tree of Life C, Formenti G, Abueg L, Torrance
679 J, Myers EW, Durbin R, Blaxter M *et al* (2023) MitoHiFi: a python pipeline for mitochondrial
680 genome assembly from PacBio high fidelity reads. *BMC Bioinformatics* 24: 288
681 Vinogradov AE (1998) Genome size and GC-percent in vertebrates as determined by flow
682 cytometry: The triangular relationship. *Cytometry* 31: 100-109
683 Wheeler TJ, Eddy SR (2013) nhmmer: DNA homology search with profile HMMs. *Bioinformatics*
684 29: 2487-2489
685 Wolf JB, Lindell J, Backstrom N (2010) Speciation genetics: current status and evolving
686 approaches. *Philos Trans R Soc Lond B Biol Sci* 365: 1717-1733
687 Zhou C, McCarthy SA, Durbin R (2023) YaHS: yet another Hi-C scaffolding tool. *Bioinformatics* 39
688