Posted on Authorea 21 Jan 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164275202.25941201/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Chromosome-level Genome Assembly for Takin (Budorcas taxicolor) Provides Insights into Its Taxonomic Status and Genetic Diversity

Anning Li[1], Qimeng Yang[1], Ran Li[1], Xuelei Dai[1], Keli Cai[1], Yinghu Lei[2], Kangsheng Jia[2], Yu Jiang[1], and Linsen Zan[3]

[1]Northwest A&F University
[2]Shaanxi Academy of Forestry Sciences
[3]Northwest Agriculture and Forestry University

January 21, 2022

Abstract

The takin (Budorcas taxicolor) is one of the largest bovid herbivores across caprinae subfamily. The takin is at high risk of extinction, however, its taxonomic status is still unclear. In this study, we constructed the first reference genome of B. taxicolor using PacBio long High-Fidelity reads and Hi-C technology. The assembled genome is ˜2.95 Gb with a contig N50 of 68.05 Mb and a scaffold N50 of 101.27 Mb, which were anchored onto 25+XY chromosomes. Compared to the common ancestral karyotype of bovidae (2n=60), we found the takin (2n=52) experienced four chromosome fusions and one large translocation. We also found that the takin was most closely related to muskox, not other caprinae species. Further, we re-sequenced nine golden takins from the main distribution area, Qinling Mountains, and identified 3.3 million SNPs. The genetic diversity of takin was very low ($\vartheta\pi$=0.00028 and heterozygosity=0.00038), which was among the lowest detected in the domestic and wild mammals. We also found takin genomes showed high inbreeding coefficient (FROH=0.217) suggesting severe inbreeding depression. The genome analysis show that the effective population size of takins declined significantly from ˜100,000 years ago. Our results provide valuable information for protection of takins and insights into its evolution.

Chromosome-level Genome Assembly for Takin (*Budorcas taxicolor*) Provides Insights intoIts Taxonomic Status and Genetic Diversity

Running title: *De novo* assembly of Takin (*B. taxicolor*)

Anning Li[1#], Qimeng Yang[1,2#], Ran Li[1,2], Xuelei Dai[1,2], Keli Cai[1], Yinghu Lei[3], Kangsheng Jia[3], Yu Jiang[1,2*], Linsen Zan[1,3*]

[1] College of Animal Science and Technology, Northwest A&F University, Yangling 712100, Shaanxi, P. R. China

[2] Center for Ruminant Genetic and Evolution, Northwest A&F University, Yangling 712100, Shaanxi, P. R. China

[3] Research Center for the Qinling Giant Panda (Shaanxi Rare Wildlife Rescue Base), Shaanxi Academy of Forestry Sciences, Zhouzhi 710402, Shaanxi, P. R. China

[#] These authors contributed equally to this study.

[*] Corresponding E-mail: *yu.jiang@nwafu.edu.cn* ;*zanlinsen@163.com* .

Posted on Authorea 21 Jan 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164275202.25941201/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Abstract: The takin (*Budorcas taxicolor*) is one of the largest bovid herbivores across caprinae subfamily. The takin is at high risk of extinction, however, its taxonomic status is still unclear. In this study, we constructed the first reference genome of *B. taxicolor* using PacBio long High-Fidelity reads and Hi-C technology. The assembled genome is ˜2.95 Gb with a contig N50 of 68.05 Mb and a scaffold N50 of 101.27 Mb, which were anchored onto 25+XY chromosomes. Compared to the common ancestral karyotype of bovidae (2n=60), we found the takin (2n=52) experienced four chromosome fusions and one large translocation. We also found that the takin was most closely related to muskox, not other caprinae species. Further, we re-sequenced nine golden takins from the main distribution area, Qinling Mountains, and identified 3.3 million SNPs. The genetic diversity of takin was very low ($\vartheta\pi$=0.00028 and heterozygosity=0.00038), which was among the lowest detected in the domestic and wild mammals. We also found takin genomes showed high inbreeding coefficient (FROH=0.217) suggesting severe inbreeding depression. The genome analysis show that the effective population size of takins declined significantly from ˜100,000 years ago. Our results provide valuable information for protection of takins and insights into its evolution.

Keywords : Takin; PacBio HiFi; Hi-C; Chromosomal evolution; Inbreeding depression

Introduction

The takin (*Budorcas taxicolor*) is a large bovid herbivore, belonging to the Caprinae subfamily. The number of takins was estimated only about 7,000 ˜ 12,000 in the world (Cheng et al. 2020). Takin has been listed as vulnerable by the International Union for Conservation of Nature (IUCN) (Zeng and Song 2002). It was divided into four subspecies according to the physiological characteristics and geographical location (Wu 1986). Qinling takin (*B. t. bedfordi*) and Sichuan takin (*B. t. tibetana*) are only confined in China, while the Mishmi takin (*B. t. taxicolor*) has a distributed area from Gaoligong Mountains in northwestern Yunnan province, China to India and Myanmar (Wu 1986). In addition, Bhutan takin (*B. t. whieti*) is found in Bhutan and the Yarlung Zangbo River in Tibet, China (Li et al. 2003). Qinling takin (*B. t. bedfordi*), also known as golden takin, is mainly distributed in the Qinling Mountains of China (**Figure 1a**) (Li et al. 2021).

Compared to other Caprinae animals, the morphology, ecological traits and G-banded karyotype of takin were found to be similar to muskox (*Ovibos moschatus*) (Pasitschniak-Arts et al. 1994). However, the analysis of mitochondrial cytochrome b genes (*Cytb*) sequences showed that muskox and takin were not close with each other but under convergent evolution (Groves and Shields 1997, Ren et al. 2012). Recently, based on the complete mitochondrial genome, takin was found to be closely related to goat (*Capra hircus*) rather than sheep (*Ovis aries*) (Feng et al. 2016, Kumar et al. 2019, Zhou et al. 2019). In our previous study, takin was found to be closely related to sheep rather than goat according to the transcriptome analysis (Qiu et al. 2021). Thus, to clarify on its taxonomic status, further analysis is needed at the whole genomic level.

Recently, highly accurate long high-fidelity (HiFi) reads was generated by PacBio single-molecule real-time (SMRT) sequencing with circular consensus sequencing (CCS) (Wenger et al. 2019). HiFi reads combined with Hi-C technology have been used to construct chromosome-level reference genome of various animals (Wu et al. 2021, Zhou et al. 2021) and plants (Chen et al. 2021, Huang et al. 2021, Ma et al. 2021, Sharma et al. 2021, Wang et al. 2021). In this study, we performed the PacBio long HiFi reads and Hi-C technology to construct a high-quality chromosome-level reference genome of takin. The phylogenetic relationship, chromosomal evolution, genetic diversity, demographic history and inbreeding depression of takin were analyzed. These information will be helpful for understanding the evolution of the takin and working towards its conservation.

Materials and Methods

2.1 Sample collection and ethics statement

The liver sample from a dead adult male golden takin was used for *de novo* genome sequencing. Seven Blood samples (alive) and two muscle samples (dead) from nine golden takins were used for genome resequencing.

Posted on Authorea 21 Jan 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164275202.25941201/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

All of the samples were collected from the Rare Wildlife Rescue and Breeding Research Center in Qinling Mountains. All the experimental procedures were carried out according to the guidelines of the China Council on Animal Care and approved by the Experimental Animal Management Committee (EAMC) of Northwest A&F University.

Genome sequencing and assembly

A 15 Kb DNA SMRTbell library was constructed for SMRT sequencing according to a standard protocol (Ardui et al. 2018) using PacBio Sequel II platform with circular consensus sequencing (CCS). HiFi reads were assembled in *de novo* using Hifiasm (v0.13) (Cheng et al. 2021). Then the restriction enzyme *Mbo* I was used to digest cross-linked chromatin to construct the Hi-C library, which was sequenced on an Illumina NovaSeq6000 platform. The Hi-C contigs were anchored onto the chromosomes using Juicer (v1.6) (Durand et al. 2016) and 3D-DNA (v201008) (Dudchenko et al. 2017) combined with Juicebox (*https://github.com/theaidenlab/juicebox*). BUSCO (v3.0.2) was used to assess the completeness of assembled genome (Simao et al. 2015). Genome resequencing was performed by the Illumina NovaSeq6000 platform. All of the sequencing services were provided by the Berry Genomics Biotechnologies Co., Ltd. (Beijing, China). A resequencing genome of golden takin (TX-11) was selected to estimate the genome size and heterozygosity rate by gce (v1.0.2) (*https://arxiv.org/abs/1308.2012v2*).

Repeats and gene annotation

The RepeatMasker software (v4.1.2) (http://www.repeatmasker.org) and the RepBase library (2018.10.26) were used to identify the repeats in takin genome. Homology-based method was used to predict the protein-coding genes with RNA-seq data. Firstly, we annotated the protein-coding genes based on the protein sequences from *Bos taurus* , *C. hircus* ,*O. aries* and *Homo sapiens* using the BRAKER (v2.1.5) pipeline (Hoff et al. 2016, Hoff et al. 2019). Secondly, we aligned three RNA-seq datasets (PRJNA720167) to the takin's genome using STAR software (v2.7.1a) (Dobin et al. 2013), which were predicted by the BRAKER pipeline. Next, the BRAKER with TSEBRA module (Gabriel et al. 2021) was used to integrate the prediction results. Finally, the prediction results were used to align against the non-redundant (Nr) database using the DIAMOND (v2.0.11.149) software (Buchfink et al. 2015, Buchfink et al. 2021). GC content distribution of genome was performed by bedtools (v2.26.0) with a window of 500 kb (Quinlan 2014).

Mitochondrial genome assembly

All the HiFi reads (Average length 14 kb) were blasted to the mitochondrial genome of Sichuan takin (GenBank accession No. NC_039686.1) and assembly reference genome of golden takin using the Minimap2 (v2.22-r1101) (Li 2018). The HiFi reads aligned to NC_039686.1 were selected to assemble the mitochondrial genome of golden takin. The sequence similarity was analyzed using the BLAST search program. The annotation of mitochondrial genome was completed by the AGORA online platform (Jung et al. 2018).

Phylogenetic analysis and divergence time estimation

We downloaded nine mitochondrial genome from GenBank database to construct a phylogenetic tree, including *B. taurus* (GenBank accession No. NC_006853.1), *C. hircus* (GenBank accession No. NC_005044.2), *O. aries* (GenBank accession No. NC_001941.1),*Pantholops hodgsonii* (GenBank accession No. NC_007441.1),*O. moschatus* (GenBank accession No. NC_020631.1), *Oreamnos americanus* (GenBank accession No. NC_020630.1), *Ammotragus lervia* (GenBank accession No. NC_009510.1), *Pseudois nayaur*(GenBank accession No. NC_020632.1), and *B. taxicolor* (GenBank accession No. NC_039686.1, NC_013069.1, NC_043930.1, KY399869.1, KU361169.1 and OM237313). Firstly, the mitochondrial genomes were aligned by MUSCLE (v3.8.31) (Edgar 2004) to exclude the ambiguous regions. Then, the neighbor-joining (NJ) method was selected using MEGA (v11) (Tamura et al. 2021) with 1000 bootstrap replicates.

The genome of *B. taurus* (ARS-UCD1.2) (Rosen et al. 2020) was selected as reference, and that of *O. americanus* (ASM975805v1) (Martchenko et al. 2020),*O. aries*(Oar_rambouillet_v1.0),*C. hircus* (ARS1) (Worley 2017),*P. hodgsonii* (PHO1.0) (Ge et al. 2013),*A. lervia*(ALER1.0) (Chen et al. 2019),*Pseudois nayaur* (ASM318257v1) (Chen et al. 2019), *O. moschatus* (ASM2146233v1) and *B. taxicolor*(Takin1.0) were

Posted on Authorea 21 Jan 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164275202.25941201/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

separately aligned to the reference genome using LAST software (v942) (Kielbasa et al. 2011). The pair-wise alignments were merged into multiple genome alignments using the MULTIZ software (v11.2) (Blanchette et al. 2004). The consensus coding sequences of nine species were extracted using the Perl scripts from RGD (v2.0). 5, 862 single-copy homologous genes were identified using the DIAMOND (v2.0.11.149) and RAxML (v8.2.12) (Stamatakis 2014) with Maximum Likelihood (ML) method to construct the phylogenetic tree. The estimations of divergence time were calculated using PAML MCMCtree (v4.9j) (Yang 2007) with calibrated time from the previous study (Chen et al. 2019).

Chromosomal evolution

The ancestral chromosome karyotype was reconstructed using the genomes of *B. taurus* (Rosen et al. 2020), *C. hircus* (Li et al. 2021), *O. aries* (Davenport et al. 2021), *B. taxicolor* and*Physeter catodon* (Fan et al. 2019) (as outgroup). The genome of*B. taurus* was selected as the reference, and other genomes were aligned to the reference genome using LAST (v942) (Kielbasa et al. 2011). Next, "chain" and "net" files were generated from axtChain and ChainNet, which were used as input for DESCHRAMBLER at a 350 kb resolution (Kim et al. 2017). Lastly, we analyzed the collinearity of the genomes of *B. taurus* , *O. aries* and *B. taxicolor*using mummer (v4.0beta2) (Marcais et al. 2018). We visualized the collinearity region and detected the chromosome fusion events using the RectChr (*https://github.com/BGI-shenzhen/RectChr*).

Heterozygosity estimation

Genome resequencing datasets from nine golden takins were aligned with the assembly reference genome using BWA-MEM (v0.7.17) (*http://arxiv.org/abs/1303.3997*). The SAMtools (v1.7) (Li et al. 2009) was used to convert sam to bam files. The single nucleotide polymorphisms (SNPs) were called and filted using the GATK (v4.1.7.0) (McKenna et al. 2010) with parameters "QD $<$ 2.0, QUAL $<$ 30.0, FS $>$ 60.0, MQ $<$ 40.0, MQRankSum $<$ -12.5, ReadPosRankSum $<$ -8.0". The nucleotide diversity ($\vartheta\pi$) was calculated using the VCFtools (v0.1.16) with a 50 kb window (Danecek et al. 2011). Genome resequencing datasets 8 giant pandas (Zhao et al. 2013) were used to calculate $\vartheta\pi$. The genome-wide heterozygosity rate was calculated as previously described (Liu et al. 2021). Genome resequencing datasets from 26 snub-nosed monkeys (Yu et al. 2016) were used to calculate heterozygosity. The heterozygosity rates of other species were obtained from previous reports (Li et al. 2010, Cho et al. 2013, Dobrynin et al. 2015, Liu et al. 2021).

ROH estimation and inbreeding coefficient

The runs of homozygosity (ROH) were estimated across autosomes for nine golden takins using PLINK (v1.90b6.21) (Chang et al. 2015). The following PLINK parameters (Purcell et al. 2007) were applied to define a ROH: (i) a 50 SNPs sliding window across the genome; (ii) the proportion of homozygous overlapping windows was 0.05; (iii) minimum number of consecutive SNPs included in a ROH was 50; (iv) required minimum density was set at one SNP per 50 kb; (v) maximum gap between consecutive homozygous SNPs was 1000 kb; and (vi) maximum of five SNPs with missing genotypes and up to three heterozygous genotype were allowed in a ROH. Inbreeding coefficient based on ROH (FROH) for each individual was calculated according to previous study (McQuillan et al. 2008). The number of generations was estimated to the common ancestor of these homologous sequences as previously described (Thompson 2013).

2.9 Demographic history analysis

The golden takin (TX-11, ˜42X), Milu (*Elaphurus davidianus* , SRR5762659- SRR5762666, ˜50X) (Chen et al. 2019) and Qinling snub-nosed monkey (SRR2017686, ˜15X) (Yu et al. 2016) were selected for inferring the effective population size trajectory using the Pairwise Sequentially Markovian Coalescence (PSMC) model (v0.6.5-r67) (Li and Durbin 2011). The neutral mutation rate (μ) was estimated using r8s (v1.81) (Sanderson 2003). The remaining steps were performed as previously described (Yin et al. 2021).

Results

Chromosome-level de novo genome assembly of golden takin

Posted on Authorea 21 Jan 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164275202.25941201/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

To estimate the genome size of *B. taxicolor* , 124.37 Gb clean reads were used for k-mer analysis. The genome size was estimated at 2.79 Gb (k=17, **Figure 1b**). We constructed a chromosome-level reference genome for *B. taxicolor* by PacBio HiFi combined with Hi-C sequencing technology (**Figure 1c**). A total of 68.56 Gb HiFi reads (˜23×) were generated by PacBio Sequel II platform (**Table S1**). We assembled 4671 contigs using hifasm software. The *de novo*assembled reference genome was about 2.95 Gb with a contig N50 of 68.05 Mb (**Table 1**), which had a longer contig N50 as compared to goat, sheep and cattle reference genomes (**Table S2**). Further, 344.7 Gb Hi-C reads (˜117×) spanned to 4,512 scaffolds with a N50 of 101.27 Mb (**Table 1**), which anchored onto 25+XY chromosomes (**Figure 1d**). The genome of takin had a good collinearity relationship with that of the goat (Saanen_v1) (**Figure S1**) and sheep (ARS-UI_Ramb_v2.0) (**Figure S2**). Based on the assembled genome, we evaluated its completeness with mammalia_odb9 database using BUSCO program. The completeness of assembled genome reached 94.2%, of which a single-copy reached 93.1% (**Table S3**).

Genome annotation

A total of 49.48% of the assembled genome was repeat sequences. Long interspersed nuclear elements (LINEs) accounted for 25.08%, which was the highest among all types of repeats, followed by short interspersed nuclear elements (SINEs, 9.92%), satellites (6.97%), long terminal repeats (LTRs, 4.33%), and DNA elements (1.97%) (**Table S4**). In addition, 554,037 simple sequence repeats (SSRs, 0.82%) were identified (**Table S4**). The dimers were the largest proportion (29.70%) of SSRs, followed by quadmers (14.72%), monomers (13.75%), pentamers (9.73%), hexamers (9.20%) and trimers (8.90%) (**Table S5**). 21,301 protein-coding genes were predicted by the homology-based method combined with RNA-seq data. The average gene length was 26,238 bp with 8.82 exons on average. 99.15% (20,282) and 87.69% (18,679) were functionally annotated in the Nr and InterPro databases, respectively.

Phylogenetic analysis at the mitochondrial and genomic level

The assembled mitochondrial genome was about 16,584 bp (GenBank accession No. OM237313), which shared 98.10%, 99.05% and 99.69% sequence similarity with Mishmi (GenBank accession No. NC_043930.1), Sichuan (GenBank accession No. NC_039686.1) and golden takin (GenBank accession No. KY399869.1 and KU361169.1), respectively. At the mitochondrial genome level, takin was closely related to goat branch as compared to muskox (**Figure 2a**), which is consistent with previous studies (Feng et al. 2016, Kumar et al. 2019, Zhou et al. 2019).

To reveal the evolution of takin genome, we constructed the phylogenetic tree with representative subfamily Caprinae species. Interestingly, takin was most closely related to muskox (**Figure 2b and Figure S3**). Molecular dating using PAML MCMCtree showed that takin had split from the ancestor of sheep and goat at ˜10.06 million years ago (Mya), while takin showed a split from the ancestor of muskox at ˜8.26 Mya (**Figure 2b**).

Chromosome evolution analysis

Using the sperm whale as outgroup, we reconstructed the ancestral karyotype (2n=60, **Table S6**) of the cattle, goat, sheep and takin. 559 conserved segments were identified and 30 ancestral chromosomes with a total length of ˜2.60 Gb were predicted. Compared to the ancestral karyotype of bovid (2n=60), the karyotype of cattle (2n=60) was most conserved. The karyotype of goat (2n=60) occured only one translocation, however the karyotype of sheep (2n=54) occured three fusions and one translocation. The karyotype of takin decreased to 2n=52 after occurring four fusions and one translocation (**Figure 3a**). Further, we constructed the collinear relationship among takin, sheep and cattle. There were three chromosome (Chr) fusions from cattle to sheep including Chr 1 and 3 to Chr 1, Chr 2 and 8 to Chr 2, and Chr 5 and 11 to Chr 3. In contrast, there were four fusions from cattle to takin including Chr 1 and 22 to Chr 1, Chr 2 and 25 to Chr 2, Chr 5 and 28 to Chr 5, and Chr 11 and 23 to Chr 11. Interestingly, the translocation occurred from bovine Chr 9 to takin's Chr 14 and sheep's Chr 9 (**Figure 3b**).

Genetic diversity and population history of golden takins

Posted on Authorea 21 Jan 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164275202.25941201/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Nine additional golden takins were carried out by genome resequencing, 442.36 Gb clean reads were generated (**Table S7**). In total of 3,328,260 SNPs were identified by GATK software. To examine the extent of inbreeding of golden takins, we estimated the $\vartheta\pi$, heterozygosity and identified the runs of homozygosity (ROH). Golden takins displayed much lower single nucleotide variants (SNVs) ($\vartheta\pi$=0.00028) as compared to Qinling giant pandas ($\vartheta\pi$=0.00133). The average heterozygosity of nine golden takin genomes was about 0.00038, which was lower than domestic and most protected wild animals including Qinling giant pandas and snub-nosed monkeys, while was a little higher than that of snow leopards and cheetahs (**Figure 4a**).

The heterozygosity of golden takin genomes and proportion of the genome in ROH were shown as **Figure 4b.** Based on the length of ROH > 1 Mb (eliminating the effects of linkage disequilibrium (LD)), the average of genomic inbreeding coefficient (FROH) was found to be 0.217 (**Table S8**). The number of short ROH (2.5-4.17 Mb) was more than 20, while the number of long ROH (6.25-10 Mb) was at least one in all of the nine golden takins (**Figure 4c**). The result indicated that inbreeding occurred from 5^{th} to 20^{th} generations.

To investigate the demographic history of golden takins, the PMSC was performed. Effective population size dropped significantly from ˜1 million years ago (the Xixiabangma Glaciation) for golden takins, Milu and snub-nosed monkeys. The golden takins declined more significantly than Milu and snub-nosed monkeys from ˜100,000 years ago (**Figure 4d**).

Discussion

4.1 High-quality chromosome-level assembly reference genome of takin

The genome size of takin was about 2.95 Gb, which was slightly larger than that of the cattle (˜2.72 Gb) (Rosen et al. 2020), goat (˜2.70 Gb) (Li et al. 2021) or sheep (˜2.63 Gb) (Davenport et al. 2021). The contig N50 of takin assembly genome reached 68.05 Mb, which was also larger compared to cattle (˜25.90 Mb), goat (˜46.21 Mb) and sheep (˜43.18 Mb). Moreover, the takin assembly genome, which only has 151 gaps, had a good collinearity relationship with goat and sheep. Further, Hi-C reads with a N50 of 101.27 Mb were anchored onto 25+XY chromosomes. These results showed a high-quality chromosome-level reference genome of takin was assembled. 49.48% of reference genome was repeat sequences. Among these repeat sequences, 554,037 SSRs were identified which were more than that of the takin transcriptome (Qiu et al. 2021). These SSRs could be used as molecular markers for population genetic analysis and conservation of takins.

4.2 Taxonomic status and chromosome evolutionof takin

The mitochondrial genome was assembled from HiFi reads, which showed high similarity with that of the GenBank database. The HiFi reads can reach to an average length of 14 kb, thus the assembled mitochondrial genome was highly credible. Phylogenetic results of subfamily Caprinae showed that takin was closely related to goat and Bharal at the mitochondrial genome level. However, it was more closely related to muskox at the genomic level. It indicated that mitochondrial evolution is different from genome evolution. Previous studies showed that takin was most closely related to goat (Feng et al. 2016, Kumar et al. 2019, Zhou et al. 2019) at the mitochondrial genome level, which was consistent with the present study. However, based on the morphology, ecological traits and G-banded karyotype, takin showed similarity to muskox (Pasitschniak-Arts et al. 1994). Combined with the phylogenetic analysis at the genomic level, we speculated that it is more reasonable to classify takin and muskox into subfamily Ovibovini.

Chromosome evolution is linked to phenotypic evolution, gene family evolution and speciation (Eichler and Sankoff 2003, Damas et al. 2021). In the Bovidae family, the number of chromosomes dramatically varied in case of cattle (2n=60), goat (2n=60), sheep (2n=54), takin (2n=52) and muskox (2n=48). Based on the chromosome-level reference genome in muskox, the ancestral karyotype (2n=60) was reconstructed with cattle, goat, sheep and takin. The karyotype of takin (2n=52) occurred when there were four chromosome fusions and one translocation from the ancestor (2n=60). As a result, the number of chromosomes in takin has been reduced to 2n=52. Previous studies showed chromosome evolution was related to chromosome fission and fusion events (Chen et al. 2019, Liu et al. 2021, Yin et al. 2021). Interestingly, a functional

Posted on Authorea 21 Jan 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164275202.25941201/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

single-chromosome yeast was successfully created by end-to-end chromosome fusions and centromere deletions (Shao et al. 2018). Therefore, further research on chromosome arrangements is required in the Bovidae family to reveal the relation of chromosome evolution as well as rapid evolution of genomic sequences.

4.3 Genetic diversity and conservation status of takin

Being a vulnerable species (Zeng and Song 2002), the genetic diversity of golden takin was very low ($\vartheta\pi$=0.00028 and heterozygosity=0.00038), which was even lower than that of the Qinling giant pandas and snub-nosed monkeys who live in the same habitat. The high inbreeding coefficient (FROH=0.217) indicated that golden takins presented inbreeding depression. The long ROH represents recent inbreeding events, while the shorter ROH stems from an ancient process or mating between distant relatives (van der Valk et al. 2019). We speculated that the golden takins in our study had always been inbreeding from 5^{th} to 20^{th} generations. These results indicated that the golden takin was in a highly inbreeding state (Kardos et al. 2021). It has been reported that the number of golden takin is only about 3, 500 in Qinling mountains (Zeng et al. 2003). It is probably that the golden takin is currently endangered. Therefore, it is urgent to establish the protection scheme of golden takin.

5. Conclusion

The high-quality chromosome-level reference genome of the takin was *de novo* assembled for the first time, with its phylogenic relationship resolved. The karyotype of takin (2n=52) was formed via four chromosome fusions and one translocation from the common ancestral karyotype of bovidae family (2n=60). The very low genetic diversity ($\vartheta\pi$=0.00028 and heterozygosity=0.00038) and high inbreeding coefficient (FROH=0.217) indicated that takins were in severe inbreeding depression and endangered. These results provide valuable information to protect the takin and insights into its evolution.

Acknowledgments: We are very grateful to the Rare Wildlife Rescue and Breeding Research Center, Shaanxi Academy of Forestry Sciences for providing the samples. This research was funded by Shaanxi Province Forestry Science and Technology Innovation Special Project (No. SXLK2021-0102 and SXLK2020-0301).

References

Ardui, S., A. Ameur, J. R. Vermeesch and M. S. Hestand (2018). "Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics." *Nucleic Acids Res* **46** (5): 2159-2168.

Blanchette, M., W. J. Kent, C. Riemer, L. Elnitski, A. F. Smit, K. M. Roskin, R. Baertsch, K. Rosenbloom, H. Clawson, E. D. Green, D. Haussler and W. Miller (2004). "Aligning multiple genomic sequences with the threaded blockset aligner." *Genome Res* **14** (4): 708-715.

Buchfink, B., K. Reuter and H. G. Drost (2021). "Sensitive protein alignments at tree-of-life scale using DIAMOND." *Nat Methods* **18** (4): 366-+.

Buchfink, B., C. Xie and D. H. Huson (2015). "Fast and sensitive protein alignment using DIAMOND." *Nat Methods* **12** (1): 59-60.

Chang, C. C., C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell and J. J. Lee (2015). "Second-generation PLINK: rising to the challenge of larger and richer datasets." *Gigascience* **4** : 7.

Chen, F., L. Y. Su, S. Y. Hu, J. Y. Xue, H. Liu, G. H. Liu, Y. F. Jiang, J. K. Du, Y. S. Qiao, Y. N. Fan, H. Liu, Q. Yang, W. J. Lu, Z. Q. Shao, J. Zhang, L. S. Zhang, F. Chen and Z. M. Cheng (2021). "A chromosome-level genome assembly of rugged rose (Rosa rugosa) provides insights into its evolution, ecology, and floral characteristics." *Hortic Res-England* **8** (1).

Chen, L., Q. Qin, Y. Jiang, K. Wang, Z. S. Lin, Z. P. Li, F. Bibi, Y. Z. Yang, J. H. Wang, W. H. Nie, W. T. Su, G. C. Liu, Q. Y. Li, W. W. Fu, X. Y. Pan, C. Liu, J. Yang, C. Z. Zhang, Y. Yin, Y. Wang, Y. Zhao, C. Zhang, Z. K. Wang, Y. L. Qin, W. Liu, B. Wang, Y. D. Ren, R. Zhang, Y. Zeng, R. R. da Fonseca,

Posted on Authorea 21 Jan 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164275202.25941201/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

B. Wei, R. Li, W. T. Wan, R. P. Zhao, W. B. Zhu, Y. T. Wang, S. C. Duan, Y. Gao, Y. E. Zhang, C. Y. Chen, C. Hvilsom, C. W. Epps, L. G. Chemnick, Y. Doug, S. Mirarab, H. R. Siegismund, O. A. Ryder, M. T. P. Gilbert, H. S. Lewin, G. J. Zhang, R. Heller and W. Wang (2019). "Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits." *Science* **364** (6446): 1152-+.

Cheng, G., H. Ren, J. Chen, S. Yang, F. Long and L. Zan (2020). "Status, research progress and countermeasures for protection of endangered and rare wildlife resources in Qinling mountains area." *Shaanxi Forset Science and Technology***48** (6): 74-79.

Cheng, H. Y., G. T. Concepcion, X. W. Feng, H. W. Zhang and H. Li (2021). "Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm." *Nat Methods***18** (2): 170-+.

Cho, Y. S., L. Hu, H. Hou, H. Lee, J. Xu, S. Kwon, S. Oh, H. M. Kim, S. Jho, S. Kim, Y. A. Shin, B. C. Kim, H. Kim, C. U. Kim, S. J. Luo, W. E. Johnson, K. P. Koepfli, A. Schmidt-Kuntzel, J. A. Turner, L. Marker, C. Harper, S. M. Miller, W. Jacobs, L. D. Bertola, T. H. Kim, S. Lee, Q. Zhou, H. J. Jung, X. Xu, P. Gadhvi, P. Xu, Y. Xiong, Y. Luo, S. Pan, C. Gou, X. Chu, J. Zhang, S. Liu, J. He, Y. Chen, L. Yang, Y. Yang, J. He, S. Liu, J. Wang, C. H. Kim, H. Kwak, J. S. Kim, S. Hwang, J. Ko, C. B. Kim, S. Kim, D. Bayarlkhagva, W. K. Paek, S. J. Kim, S. J. O'Brien, J. Wang and J. Bhak (2013). "The tiger genome and comparative analysis with lion and snow leopard genomes." *Nat Commun* **4** : 2433.

Damas, J., M. Corbo and H. A. Lewin (2021). "Vertebrate Chromosome Evolution." *Annu Rev Anim Biosci***9** : 1-27.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin and G. Genomes Project Analysis (2011). "The variant call format and VCFtools." *Bioinformatics* **27** (15): 2156-2158.

Davenport, K. M., D. M. Bickhart, K. Worley, S. C. Murali, N. Cockett, M. P. Heaton, T. P. L. Smith, B. M. Murdoch and B. D. Rosen (2021). "An Improved, High-quality Ovine Reference Genome Assembly." *J Anim Sci* **99** : 23-23.

Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson and T. R. Gingeras (2013). "STAR: ultrafast universal RNA-seq aligner." *Bioinformatics* **29** (1): 15-21.

Dobrynin, P., S. Liu, G. Tamazian, Z. Xiong, A. A. Yurchenko, K. Krasheninnikova, S. Kliver, A. Schmidt-Kuntzel, K. P. Koepfli, W. Johnson, L. F. Kuderna, R. Garcia-Perez, M. Manuel, R. Godinez, A. Komissarov, A. Makunin, V. Brukhin, W. Qiu, L. Zhou, F. Li, J. Yi, C. Driscoll, A. Antunes, T. K. Oleksyk, E. Eizirik, P. Perelman, M. Roelke, D. Wildt, M. Diekhans, T. Marques-Bonet, L. Marker, J. Bhak, J. Wang, G. Zhang and S. J. O'Brien (2015). "Genomic legacy of the African cheetah, Acinonyx jubatus." *Genome Biol* **16** : 277.

Dudchenko, O., S. S. Batra, A. D. Omer, S. K. Nyquist, M. Hoeger, N. C. Durand, M. S. Shamim, I. Machol, E. S. Lander, A. P. Aiden and E. L. Aiden (2017). "De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds." *Science* **356** (6333): 92-95.

Durand, N. C., M. S. Shamim, I. Machol, S. S. Rao, M. H. Huntley, E. S. Lander and E. L. Aiden (2016). "Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments." *Cell Syst* **3** (1): 95-98.

Edgar, R. C. (2004). "MUSCLE: multiple sequence alignment with high accuracy and high throughput." *Nucleic Acids Res* **32** (5): 1792-1797.

Eichler, E. E. and D. Sankoff (2003). "Structural dynamics of eukaryotic chromosome evolution." *Science***301** (5634): 793-797.

Fan, G. Y., Y. L. Zhang, X. C. Liu, J. H. Wang, Z. G. Sun, S. Sun, H. Zhang, J. W. Chen, M. Q. Lv, K. Han, X. X. Tan, J. Hu, R. Guan, Y. Y. Fu, S. S. Liu, X. Chen, Q. W. Xu, Y. T. Qin, L. Q. Liu, J. Bai, O. Wang, J. B. Tang, H. R. Lu, Z. C. Shang, B. Wang, G. H. Hu, X. Zhao, Y. Zou, A. Chen, M. H. Gong, W. W.

Posted on Authorea 21 Jan 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164275202.25941201/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Zhang, S. M. Y. Lee, S. H. Li, J. N. Liu, Z. Li, Y. S. Lu, J. S. M. Sabir, M. J. Sabir, M. Khan, N. H. Hajrah, Y. Yin, K. Kristiansen, H. M. Yang, J. Wang, X. Xu and X. Liu (2019). "The first chromosome-level genome for a marine mammal as a resource to study ecology and evolution." *Mol Ecol Resour* **19** (4): 944-956.

Feng, H., C. Feng, Y. Huang and L. Wang (2016). "Analysis on complete mitochondrial genome sequence of golden Takin (Budorcas taxicolor)." *Genomics and Applied Biology***35** (9): 2299-2305.

Gabriel, L., K. J. Hoff, T. Bruna, M. Borodovsky and M. Stanke (2021). "TSEBRA: transcript selector for BRAKER." *BMC Bioinformatics* **22** (1): 566.

Ge, R. L., Q. Cai, Y. Y. Shen, A. San, L. Ma, Y. Zhang, X. Yi, Y. Chen, L. Yang, Y. Huang, R. He, Y. Hui, M. Hao, Y. Li, B. Wang, X. Ou, J. Xu, Y. Zhang, K. Wu, C. Geng, W. Zhou, T. Zhou, D. M. Irwin, Y. Yang, L. Ying, H. Bao, J. Kim, D. M. Larkin, J. Ma, H. A. Lewin, J. Xing, R. N. Platt, 2nd, D. A. Ray, L. Auvil, B. Capitanu, X. Zhang, G. Zhang, R. W. Murphy, J. Wang, Y. P. Zhang and J. Wang (2013). "Draft genome sequence of the Tibetan antelope." *Nat Commun* **4** : 1858.

Groves, P. and G. F. Shields (1997). "Cytochrome B sequences suggest convergent evolution of the Asian takin and Arctic muskox." *Molecular phylogenetics and evolution***8** (3): 363-374.

Hoff, K. J., S. Lange, A. Lomsadze, M. Borodovsky and M. Stanke (2016). "BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS." *Bioinformatics***32** (5): 767-769.

Hoff, K. J., A. Lomsadze, M. Borodovsky and M. Stanke (2019). "Whole-Genome Annotation with BRAKER."*Methods Mol Biol* **1962** : 65-95.

Huang, Y., M. Escalona, G. Morrison, M. P. A. Marimuthu, O. Nguyen, E. Toffelmier, H. B. Shaffer and A. Litt (2021). "Reference genome assembly of the big berry Manzanita (Arctostaphylos glauca)." *The Journal of heredity* .

Jung, J., J. I. Kim, Y. S. Jeong and G. Yi (2018). "AGORA: organellar genome annotation from the amino acid and nucleotide references." *Bioinformatics* **34** (15): 2661-2663.

Kardos, M., E. E. Armstrong, S. W. Fitzpatrick, S. Hauser, P. W. Hedrick, J. M. Miller, D. A. Tallmon and W. C. Funk (2021). "The crucial role of genome-wide genetic variation in conservation." *Proc Natl Acad Sci U S A* **118** (48).

Kielbasa, S. M., R. Wan, K. Sato, P. Horton and M. C. Frith (2011). "Adaptive seeds tame genomic sequence comparison." *Genome Res* **21** (3): 487-493.

Kim, J., M. Farre, L. Auvil, B. Capitanu, D. M. Larkin, J. Ma and H. A. Lewin (2017). "Reconstruction and evolutionary history of eutherian chromosomes." *Proc Natl Acad Sci U S A* **114** (27): E5379-E5388.

Kumar, A., K. B. Gautam, B. Singh, P. Yadav, G. V. Gopi and S. K. Gupta (2019). "Sequencing and characterization of the complete mitochondrial genome of Mishmi takin (Budorcas taxicolor taxicolor) and comparison with the other Caprinae species." *International journal of biological macromolecules***137** : 87-94.

Li, A., J. Qiu, Y. Zheng, Y. Li, R. Guo, K. Jia, Y. Lei and L. Zan (2021). "Research Progress in Conservation Biology of Qinling Takin (Budorcas taxicolor bedfordi)."*Acta Agriculturae Boerali-occidentalis Sinica* **30** (8): 1113-1121.

Li, H. (2018). "Minimap2: pairwise alignment for nucleotide sequences." *Bioinformatics***34** (18): 3094-3100.

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

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and S. Genome Project Data Processing (2009). "The Sequence Alignment/Map format and SAMtools." *Bioinformatics* **25** (16): 2078-2079.

Posted on Authorea 21 Jan 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164275202.25941291/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Li, M., S. Meng, F. Wei, J. Wang and Y. Yong (2003). "Genetic Diversity and Population Genetic Structure of Takin (Budorcastaxicolor)." *Acta Theriologica Sinica***23** (1): 10-16.

Li, R., W. Fan, G. Tian, H. Zhu, L. He, J. Cai, Q. Huang, Q. Cai, B. Li, Y. Bai, Z. Zhang, Y. Zhang, W. Wang, J. Li, F. Wei, H. Li, M. Jian, J. Li, Z. Zhang, R. Nielsen, D. Li, W. Gu, Z. Yang, Z. Xuan, O. A. Ryder, F. C. Leung, Y. Zhou, J. Cao, X. Sun, Y. Fu, X. Fang, X. Guo, B. Wang, R. Hou, F. Shen, B. Mu, P. Ni, R. Lin, W. Qian, G. Wang, C. Yu, W. Nie, J. Wang, Z. Wu, H. Liang, J. Min, Q. Wu, S. Cheng, J. Ruan, M. Wang, Z. Shi, M. Wen, B. Liu, X. Ren, H. Zheng, D. Dong, K. Cook, G. Shan, H. Zhang, C. Kosiol, X. Xie, Z. Lu, H. Zheng, Y. Li, C. C. Steiner, T. T. Lam, S. Lin, Q. Zhang, G. Li, J. Tian, T. Gong, H. Liu, D. Zhang, L. Fang, C. Ye, J. Zhang, W. Hu, A. Xu, Y. Ren, G. Zhang, M. W. Bruford, Q. Li, L. Ma, Y. Guo, N. An, Y. Hu, Y. Zheng, Y. Shi, Z. Li, Q. Liu, Y. Chen, J. Zhao, N. Qu, S. Zhao, F. Tian, X. Wang, H. Wang, L. Xu, X. Liu, T. Vinar, Y. Wang, T. W. Lam, S. M. Yiu, S. Liu, H. Zhang, D. Li, Y. Huang, X. Wang, G. Yang, Z. Jiang, J. Wang, N. Qin, L. Li, J. Li, L. Bolund, K. Kristiansen, G. K. Wong, M. Olson, X. Zhang, S. Li, H. Yang, J. Wang and J. Wang (2010). "The sequence and de novo assembly of the giant panda genome." *Nature***463** (7279): 311-317.

Li, R., P. Yang, X. Dai, H. Asadollahpour Nanaei, W. Fang, Z. Yang, Y. Cai, Z. Zheng, X. Wang and Y. Jiang (2021). "A near complete genome for goat genetic and genomic research." *Genet Sel Evol* **53** (1): 74.

Liu, C., J. Gao, X. Cui, Z. Li, L. Chen, Y. Yuan, Y. Zhang, L. Mei, L. Zhao, D. Cai, M. Hu, B. Zhou, Z. Li, T. Qin, H. Si, G. Li, Z. Lin, Y. Xu, C. Zhu, Y. Yin, C. Zhang, W. Xu, Q. Li, K. Wang, M. T. P. Gilbert, R. Heller, W. Wang, J. Huang and Q. Qiu (2021). "A towering genome: Experimentally validated adaptations to high blood pressure and extreme stature in the giraffe." *Sci Adv***7** (12).

Liu, S., M. V. Westbury, N. Dussex, K. J. Mitchell, M. S. Sinding, P. D. Heintzman, D. A. Duchene, J. D. Kapp, J. von Seth, H. Heiniger, F. Sanchez-Barreiro, A. Margaryan, R. Andre-Olsen, B. De Cahsan, G. Meng, C. Yang, L. Chen, T. van der Valk, Y. Moodley, K. Rookmaaker, M. W. Bruford, O. Ryder, C. Steiner, L. G. R. Bruins-van Sonsbeek, S. Vartanyan, C. Guo, A. Cooper, P. Kosintsev, I. Kirillova, A. M. Lister, T. Marques-Bonet, S. Gopalakrishnan, R. R. Dunn, E. D. Lorenzen, B. Shapiro, G. Zhang, P. O. Antoine, L. Dalen and M. T. P. Gilbert (2021). "Ancient and modern genomes unravel the evolutionary history of the rhinoceros family." *Cell***184** (19): 4874-4885 e4816.

Ma, D. N., S. S. Dong, S. C. Zhang, X. Q. Wei, Q. J. Xie, Q. S. Ding, R. Xia and X. T. Zhang (2021). "Chromosome-level reference genome assembly provides insights into aroma biosynthesis in passion fruit (Passiflora edulis)." *Mol Ecol Resour* **21** (3): 955-968.

Marcais, G., A. L. Delcher, A. M. Phillippy, R. Coston, S. L. Salzberg and A. Zimin (2018). "MUMmer4: A fast and versatile genome alignment system." *PLoS Comput Biol***14** (1): e1005944.

Martchenko, D., R. Chikhi and A. B. A. Shafer (2020). "Genome Assembly and Analysis of the North American Mountain Goat (Oreamnos americanus) Reveals Species-Level Responses to Extreme Environments." *G3 (Bethesda)* **10** (2): 437-442.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly and M. A. DePristo (2010). "The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data." *Genome Res* **20** (9): 1297-1303.

McQuillan, R., A. L. Leutenegger, R. Abdel-Rahman, C. S. Franklin, M. Pericic, L. Barac-Lauc, N. Smolej-Narancic, B. Janicijevic, O. Polasek, A. Tenesa, A. K. Macleod, S. M. Farrington, P. Rudan, C. Hayward, V. Vitart, I. Rudan, S. H. Wild, M. G. Dunlop, A. F. Wright, H. Campbell and J. F. Wilson (2008). "Runs of homozygosity in European populations." *Am J Hum Genet***83** (3): 359-372.

Pasitschniak-Arts, M., P. F. Flood, S. M. Schmutz and B. Seidel (1994). "A comparison of G-band patterns of the muskox and takin and their evolutionary relationship to sheep." *The Journal of heredity* **85** (2): 143-147.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. de Bakker, M. J. Daly and P. C. Sham (2007). "PLINK: a tool set for whole-genome association and

Posted on Authorea 21 Jan 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164275202.25941201/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

population-based linkage analyses."*Am J Hum Genet* **81** (3): 559-575.

Qiu, J., R. Guo, Y. Li, Y. Zhang, K. Jia, Y. Lei, L. Zan and A. Li (2021). "De Novo Transcriptome Assembly, Functional Annotation and SSR Marker Discovery of Qinling Takin (Budorcas taxicolor bedfordi)." *Animals (Basel)* **11** (8).

Quinlan, A. R. (2014). "BEDTools: The Swiss-Army Tool for Genome Feature Analysis." *Curr Protoc Bioinformatics* **47** : 11 12 11-34.

Ren, Y., H. Feng, C. Feng and S. Meng (2012). "Phylogenic analysis of Takin by mtDNA." *Genomics and Applied Biology* **31** (5): 481-484.

Rosen, B. D., D. M. Bickhart, R. D. Schnabel, S. Koren, C. G. Elsik, E. Tseng, T. N. Rowan, W. Y. Low, A. Zimin, C. Couldrey, R. Hall, W. Li, A. Rhie, J. Ghurye, S. D. McKay, F. Thibaud-Nissen, J. Hoffman, B. M. Murdoch, W. M. Snelling, T. G. McDaneld, J. A. Hammond, J. C. Schwartz, W. Nandolo, D. E. Hagen, C. Dreischer, S. J. Schultheiss, S. G. Schroeder, A. M. Phillippy, J. B. Cole, C. P. Van Tassell, G. Liu, T. P. L. Smith and J. F. Medrano (2020). "De novo assembly of the cattle reference genome with single-molecule sequencing." *Gigascience* **9** (3).

Sanderson, M. J. (2003). "r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock." *Bioinformatics* **19** (2): 301-302.

Shao, Y., N. Lu, Z. Wu, C. Cai, S. Wang, L. L. Zhang, F. Zhou, S. Xiao, L. Liu, X. Zeng, H. Zheng, C. Yang, Z. Zhao, G. Zhao, J. Q. Zhou, X. Xue and Z. Qin (2018). "Creating a functional single-chromosome yeast." *Nature* **560** (7718): 331-335.

Sharma, P., A. K. Masouleh, B. Topp, A. Furtado and R. J. Henry (2021). "De novo chromosome level assembly of a plant genome from long read sequence data." *Plant J* .

Simao, F. A., R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva and E. M. Zdobnov (2015). "BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs."*Bioinformatics* **31** (19): 3210-3212.

Stamatakis, A. (2014). "RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies." *Bioinformatics* **30** (9): 1312-1313.

Tamura, K., G. Stecher and S. Kumar (2021). "MEGA11: Molecular Evolutionary Genetics Analysis Version 11."*Molecular biology and evolution* **38** (7): 3022-3027.

Thompson, E. A. (2013). "Identity by descent: variation in meiosis, across genomes, and in populations."*Genetics* **194** (2): 301-326.

van der Valk, T., D. Diez-Del-Molino, T. Marques-Bonet, K. Guschanski and L. Dalen (2019). "Historical Genomes Reveal the Genomic Consequences of Recent Population Decline in Eastern Gorillas." *Curr Biol* **29** (1): 165-170 e166.

Wang, B., X. Yang, Y. Jia, Y. Xu, P. Jia, N. Dang, S. Wang, T. Xu, X. Zhao, S. Gao, Q. Dong and K. Ye (2021). "High-quality Arabidopsis thaliana Genome Assembly with Nanopore and HiFi Long Reads." *Genomics Proteomics Bioinformatics* .

Wenger, A. M., P. Peluso, W. J. Rowell, P. C. Chang, R. J. Hall, G. T. Concepcion, J. Ebler, A. Fungtammasan, A. Kolesnikov, N. D. Olson, A. Topfer, M. Alonge, M. Mahmoud, Y. Qian, C. S. Chin, A. M. Phillippy, M. C. Schatz, G. Myers, M. A. DePristo, J. Ruan, T. Marschall, F. J. Sedlazeck, J. M. Zook, H. Li, S. Koren, A. Carroll, D. R. Rank and M. W. Hunkapiller (2019). "Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome." *Nat Biotechnol***37** (10): 1155-1162.

Worley, K. C. (2017). "A golden goat genome." *Nat Genet* **49** (4): 485-486.

Posted on Authorea 21 Jan 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164275202.25941201/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Wu, B., X. Chen, M. J. Yu, J. F. Ren, J. Hu, C. W. Shao, L. Q. Zhou, X. J. Sun, T. Yu, Y. X. Zheng, Y. Wang, Z. Y. Wang, H. Zhang, G. Y. Fan and Z. H. Liu (2021). "Chromosome-level genome and population genomic analysis provide insights into the evolution and environmental adaptation of Jinjiang oyster Crassostrea ariakensis." *Mol Ecol Resour* .

Wu, J. (1986). "Study of system and distribution of Chinese Takin." *Zoological Research***7** (2): 167-175.

Yang, Z. (2007). "PAML 4: phylogenetic analysis by maximum likelihood." *Molecular biology and evolution* **24** (8): 1586-1591.

Yin, Y., H. Fan, B. Zhou, Y. Hu, G. Fan, J. Wang, F. Zhou, W. Nie, C. Zhang, L. Liu, Z. Zhong, W. Zhu, G. Liu, Z. Lin, C. Liu, J. Zhou, G. Huang, Z. Li, J. Yu, Y. Zhang, Y. Yang, B. Zhuo, B. Zhang, J. Chang, H. Qian, Y. Peng, X. Chen, L. Chen, Z. Li, Q. Zhou, W. Wang and F. Wei (2021). "Molecular mechanisms and topological consequences of drastic chromosomal rearrangements of muntjac deer." *Nat Commun* **12** (1): 6858.

Yu, L., G. D. Wang, J. Ruan, Y. B. Chen, C. P. Yang, X. Cao, H. Wu, Y. H. Liu, Z. L. Du, X. P. Wang, J. Yang, S. C. Cheng, L. Zhong, L. Wang, X. Wang, J. Y. Hu, L. Fang, B. Bai, K. L. Wang, N. Yuan, S. F. Wu, B. G. Li, J. G. Zhang, Y. Q. Yang, C. L. Zhang, Y. C. Long, H. S. Li, J. Y. Yang, D. M. Irwin, O. A. Ryder, Y. Li, C. I. Wu and Y. P. Zhang (2016). "Genomic analysis of snub-nosed monkeys (Rhinopithecus) identifies genes and processes related to high-altitude adaptation." *Nat Genet* **48** (8): 947-952.

Zeng, Z. and Y. Song (2002). "Golden Takin (Budorcas taxicolor bedfordi)." *Chinese Journal of Zoology***37** (1): 5.

Zeng, Z., W. Zhong, Y. Song, J. Li, L. Zhao and H. Gong (2003). "Present status of studies on eco-biology of takin." *Acta Theriologica Sinica* **23** (2): 161-167.

Zhao, S. C., P. P. Zheng, S. S. Dong, X. J. Zhan, Q. Wu, X. S. Guo, Y. B. Hu, W. M. He, S. N. Zhang, W. Fan, L. F. Zhu, D. Li, X. M. Zhang, Q. Chen, H. M. Zhang, Z. H. Zhang, X. L. Jin, J. G. Zhang, H. M. Yang, J. Wang, J. Wang and F. W. Wei (2013). "Whole-genome sequencing of giant pandas provides insights into demographic history and local adaptation." *Nature Genetics***45** (1): 67-U99.

Zhou, J., P. Wu, Z. Xiong, N. Liu, N. Zhao, M. Ji, Y. Qiu and B. Yang (2021). "Chromosome-Level Genome Assembly Reveals Significant Gene Expansion in the Toll and IMD Signaling Pathways of Dendrolimus kikuchii." *Frontiers in genetics* **12** : 728418.

Zhou, M., J. Yu, B. Li, B. Ouyang and J. Yang (2019). "The complete mitochondrial genome of Budorcas taxicolor tibetana (Artiodactyla: Bovidae) and comparison with other Caprinae species: Insight into the phylogeny of the genus Budorcas." *International journal of biological macromolecules* **121** : 223-232.

Data Availability Statement: The data presented in this study are available in the Sequence Read Archive (SRA) database, BioProject: PRJNA776710 and PRJNA778655. The Takin1.0 genome has been submitted to the GenBank database.

Author Contributions: Conceptualization, Y. J. and A. L.; Methodology and Validation, Q. Y., K. C., R. L. and X. D.; Formal Analysis, A. L. and Q. Y.; Resources, Y. L., K. J. and L. Z.; Writing – Original Draft Preparation, A. L.; Writing – Review & Editing, Y. J., R. L., Q. Y. and X. D.; Supervision, Y. J. and L. Z.; Project Administration, Y. J. and A. L.; Funding Acquisition, L. Z. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no competing financial interest.

Supplemental Figures and Tables (Captions)

Figure S1. The collinearity relationship between takin and goat genome.

Figure S2. The collinearity relationship between takin and sheep genome.

Posted on Authorea 21 Jan 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164275202.25941201/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Figure S3. Phylogeny of the subfamily Caprinae in the genomic level with bootstrap.

Table S1 Summary of HiFi reads.

Table S2 Comparison of well-assembled genomes.

Table S3 Assessment of genome completeness.

Table S4 Repeat content of Takin.

Table S5 Simple repeats content of Takin.

Table S6 Lengths of reconstructed ancestral chromosome (chr) karyotypes of Bovidae

Table S7 Genome resequencing of 9 golden takins.

Table S8 The statistics of ROH.

Table 1. Assembly statistics of takin genome

Assembly	**Number/length**
Total assembly length (bp)	2,947,084,417
Gap number	151
Number of contigs	4,671
N50 contig length (bp)	68,053,581
Contig L50	16
Number of scaffolds	4,512
N50 scaffold length (bp)	101,265,939
Scaffold L50	11
Chromosome number	25+XY
GC content (%)	43.22
Protein-coding genes	21,301

Posted on Authorea 21 Jan 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164275202.25941201/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

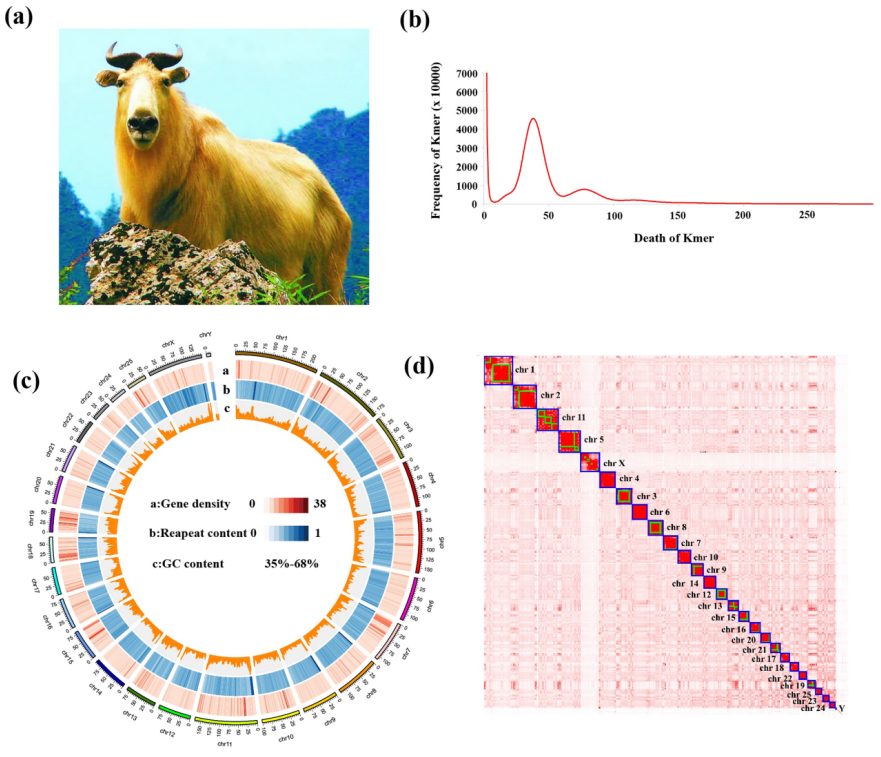

Figure 1. High-quality chromosome-level reference genome assembly.

(a) Golden takin; **(b)** K-mer analysis of the genome size by gce (v1.0.2) (*https://arxiv.org/abs/1308.2012v2*);**(c)** Circos plot showing the visualization of genomic details with a window of 500 kb. a. gene density; b. repeat content; c. GC content; **(d)** The Hi-C interaction heatmap of takin.

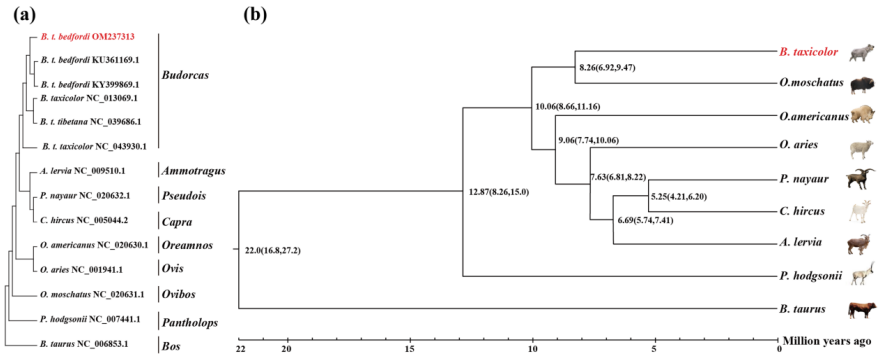

Figure 2. Phylogeny of the subfamily Caprinae in the mitochondrial and genomic level.

(a) Takin was closely related to goat clade, but far to muskox in the mitochondrial genomic level. The neighbor-joining (NJ) method with 1000 bootstrap replicates using MEGA (v11) (Tamura et al. 2021);**(b)** Takin was most closely related to muskox in the genomic level. 5862 single-copy homologous genes were used to construct the phylogenetic tree by RAxML (v8.2.12) with Maximum Likelihood (ML) method. The divergence time were calculated using PAML (v4.9j) with calibrated from divergence time of goat and sheep (3.9-8.1Mya) (Chen et al. 2019). The estimated divergence time (Mya) was marked on each node (95% confidence intervals).

Figure 3. Reconstruction of ancestral chromosomes of cattle, goat, sheep and takin.

(a) The distribution of ancestral chromosome segments in the genomes of cattle, goat, sheep and takin, including chromosome fission and fusion events. The asterisks represent chromosome fission events;**(b)** The collinear relationship of takin, sheep and cattle. Sperm whale as outgroup.

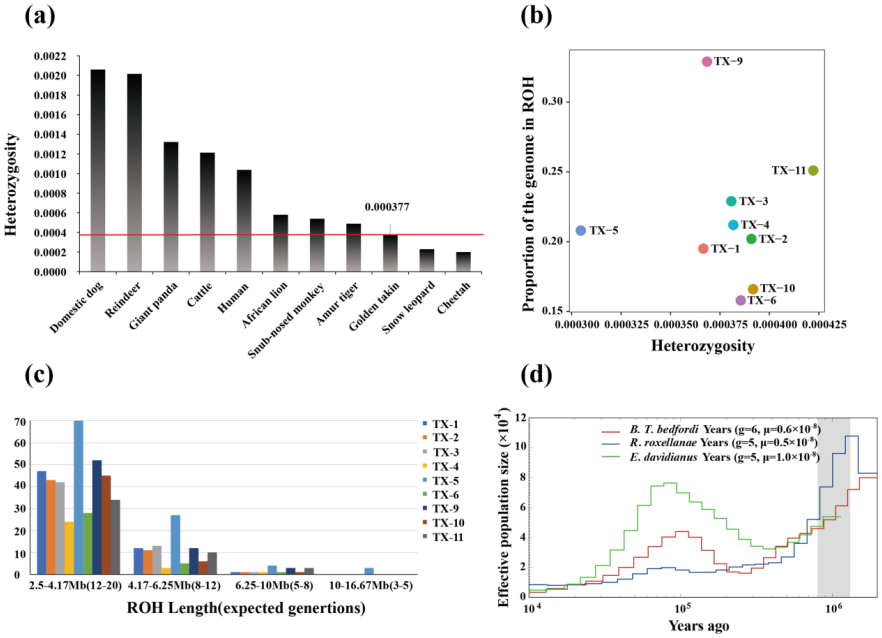

Figure 4. Genetic diversity and demographic history of nine golden takins.

(a) The average heterozygosity of different domestic and wild animals. Genome resequencing datasets from nine golden takins were used to calculate heterozygosity as previously described (Liu et al. 2021). The heterozygosity rates of other species were reported in previous reports (Li et al. 2010, Cho et al. 2013, Dobrynin et al. 2015, Liu et al. 2021). **(b)** The heterozygosity and proportion of the genome in ROH from nine golden takins. **(c)** The distributions of ROH statistics. With the length of ROH increasing, the number of ROH gradually decreased. **(d)** Demographic history analysis of golden takins, Milu and snub-nosed monkeys estimated by PSMC. The gray box marks the time range of the Xixiabangma Glaciation (0.8-1.17 million years ago). "g" represents generation, "μ" represents neutral mutation rate.

Posted on Authorea 21 Jan 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.164275202.25941201/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.