RESEARCH ARTICLE

Open Access

Population genomic analysis provides evidence of the past success and future potential of South China tiger captive conservation

Chen Wang[1†], Dong-Dong Wu[2,3,4†], Yao-Hua Yuan[5], Meng-Cheng Yao[2,3,4], Jian-Lin Han[6,7], Ya-Jiang Wu[1], Fen Shan[1], Wan-Ping Li[1], Jun-Qiong Zhai[1], Mian Huang[1], Shi-Ming Peng[1], Qin-Hui Cai[1], Jian-Yi Yu[5], Qun-Xiu Liu[5], Zhao-Yang Liu[8], Lin-Xiang Li[9], Ming-Sheng Teng[10], Wei Huang[11], Jun-Ying Zhou[12], Chi Zhang[13], Wu Chen[1*] and Xiao-Long Tu[2,3,4*]

Abstract

Background Among six extant tiger subspecies, the South China tiger (*Panthera tigris amoyensis*) once was widely distributed but is now the rarest one and extinct in the wild. All living South China tigers are descendants of only two male and four female wild-caught tigers and they survive solely in zoos after 60 years of effective conservation efforts. Inbreeding depression and hybridization with other tiger subspecies were believed to have occurred within the small, captive South China tiger population. It is therefore urgently needed to examine the genomic landscape of existing genetic variation among the South China tigers.

Results In this study, we assembled a high-quality chromosome-level genome using long-read sequences and re-sequenced 29 high-depth genomes of the South China tigers. By combining and comparing our data with the other 40 genomes of six tiger subspecies, we identified two significantly differentiated genomic lineages among the South China tigers, which harbored some rare genetic variants introgressed from other tiger subspecies and thus maintained a moderate genetic diversity. We noticed that the South China tiger had higher F_{ROH} values for longer runs of homozygosity (ROH > 1 Mb), an indication of recent inbreeding/founder events. We also observed that the South China tiger had the least frequent homozygous genotypes of both high- and moderate-impact deleterious mutations, and lower mutation loads than both Amur and Sumatran tigers. Altogether, our analyses indicated an effective genetic purging of deleterious mutations in homozygous states from the South China tiger, following its population contraction with a controlled increase in inbreeding based on its pedigree records.

Conclusions The identification of two unique founder/genomic lineages coupled with active genetic purging of deleterious mutations in homozygous states and the genomic resources generated in our study pave the way for a

†Chen Wang and Dong-Dong Wu contributed equally to this work.

*Correspondence:
Wu Chen
guangzhouchenwu@sina.com
Xiao-Long Tu
tuxiaolong@mail.kiz.ac.cn
Full list of author information is available at the end of the article

Wang *et al. BMC Biology*   (2023) 21:64

Page 2 of 16

Background

The tiger (*Panthera tigris*) is one of the largest felids and a widely recognized flagship species of wildlife conservation in the world. There are six commonly accepted living tiger subspecies, including the South China tiger (*P. t. amoyensis*), Amur tiger (*P. t. altaica*), Indochinese tiger (*P. t. corbetti*), Malayan tiger (*P. t. jacksoni*), Bengal tiger (*P. t. tigris*), and Sumatran tiger (*P. t. sumatrae*) [1–3]. Among them, the Sumatran tiger is the only island population to be distinctive from all continental tigers and the Amur tiger splits last within mainland Asia [4], while the South China tiger is the rarest tiger subspecies [5]. In the 1950s, around 4000 South China tigers were found in 13 provinces in China [6]. Unfortunately, up to 3000 tigers were hunted as a pest and killed mercilessly during that period. Habitat loss/fragmentation further accelerated the decline of these tigers [7]. Only in 1979, the Chinese government banned the hunting of the tigers, whereas the number of the South China tiger was estimated to be only 30–80 in 1996. No South China tiger was directly sighted in the wild since the 1990s [7], it was therefore believed to be "functionally extinct" by many scientists, followed by an official declaration of its extinction in the wild in 2012.

Considerable efforts have been made to rescue the South China tiger through a captive breeding program in China [8]. Since 1955, a total of 27 male and 20 female South China tigers were kept in captivity according to the studbook [9]. In 1963, the first success of captive breeding of this tiger subspecies was achieved in the Guiyang Qianling Zoo. Detailed pedigree records indicated that all captive South China tigers were the descendants of two male and four female wild-caught tigers, which were managed in the Shanghai Zoo (one male and one female from Guizhou as well as one female from Fujian provinces) and Guiyang Qianling Zoo (one male and two females from Guizhou province) [9]. These descendants formed two founder lineages managed independently by the Shanghai and Guiyang Qianling Zoos over the 1970s. During 1972-1984, the captive population expanded quickly from 13 to 49 tigers. However, fewer cubs were born during 1985–1995 and the average survival rate of newborn cubs was low [10, 11], following the rapid loss of genetic variations and possible inbreeding depression within this small population.

The Chinese Association of Zoological Gardens commenced a coordinated South China tiger captive breeding and management program in 1994. From 1995, the South China Tiger Committee (renamed as Tiger Taxonomy Advisory Group, the Tiger TAG, in 2015) has been organizing an annual review of the captive population and designing breeding and exchange plans [12]. The Tiger TAG set a goal to maintain 90% of the genetic variations present in the captive population in 1995 and targeted to sustain 70% of the variations over the next 100 years. Since then, the breeding tigers began to be exchanged between the zoos based on their pedigree records and the number of the South China tigers increased rapidly. There were 205 South China tigers managed in captivity by 16 Chinese zoos and 18 individuals in the Laohu Valley Reserve in South Africa by 2020.

Because of the speculated inbreeding depression and extinction of the South China tiger in the wild, some experts suggested to introduce genetic supplementation from other tiger subspecies into the captive South China tiger population to enlarge its gene pool [8]. Hybridization between the South China and Amur tigers was believed to have occurred in some Chinese zoos [13], which was supported by genetic characterization using mitochondrial and microsatellite DNA markers [14, 15]. It is evident that one allele at individual genetic loci throughout the genomes of a small and managed population can be fixed rapidly, leading to an increase in genomic homozygosity and subsequently the inbreeding load [16, 17]. Considering the positive correlation between genetic heterozygosity and fitness [18, 19], it is urgently needed to examine the genomic landscape of existing genetic variations that have driven the past rescue of the South China tiger, in reference to large-scale population genomic studies on other tiger subspecies [1, 4, 20].

In this study, we assembled a de novo chromosome-level genome and re-sequenced 29 whole genomes of the captive South China tigers collected from four major zoos. Combined with other whole-genome data of six tiger subspecies, we characterized the genetic diversity, population genetic structure, demographic history, genomic inbreeding, and deleterious mutation load of the captive South China tiger population. These findings do not only explain the successful breeding history of the South China tigers in captivity, but also pave the way for

Wang *et al. BMC Biology*   (2023) 21:64

Page 3 of 16

a genomics-informed management by applying genome-wide markers to routinely monitor and sustain their critical genetic variations in the future.

Results

De novo genome assembly and whole-genome re-sequencing of the South China tigers

To investigate the genetic variations of the South China tiger (Fig. 1a), we first constructed a high-quality de novo assembly of the South China tiger genome using a combination of high-fidelity short-read sequencing [21], long-read single-molecule real-time sequencing [22], optical mapping [23], and Hi-C [24] technologies. We generated 122.31 Gb (50.96×) of PacBio long reads, 1,011.73 Gb (421.55×) of Illumina paired-end short reads, 440.32 Gb (183.47×) of Bionano optical molecules, and 532.46 Gb (221.86×) of Hi-C data (Additional file 1: Tables S1 and S2). The K-mer [25] analysis revealed its genome size to be 2.47 Gb (Additional file 1: Table S3 and Additional file 2: Figure S1). After being polished with the PacBio long reads and corrected using the Illumina short reads [26], the PacBio-based initial assembly resulted in a contig N50 at 6.20 Mb. We scaffolded the PacBio contigs using the Bionano optical mapping data.
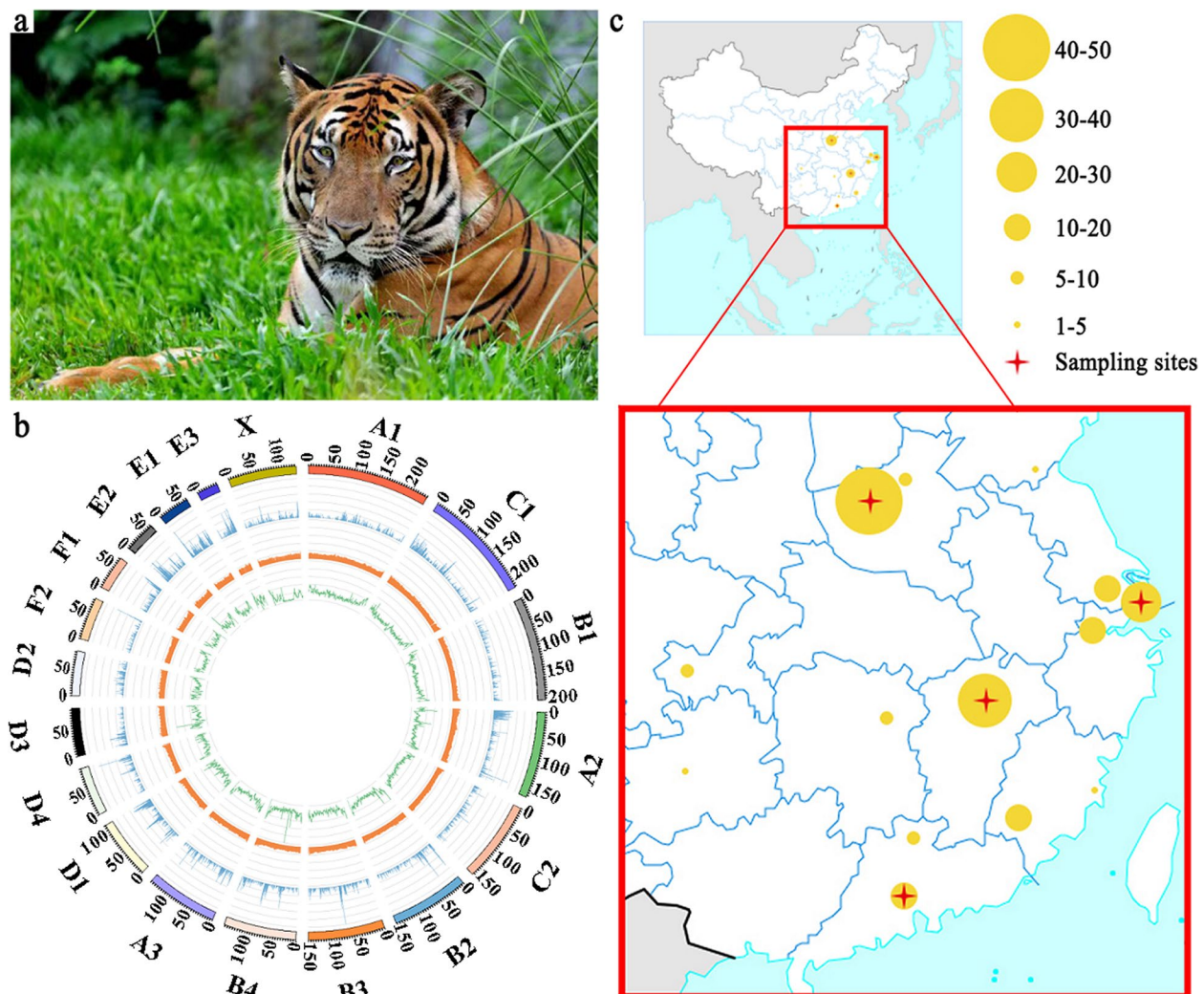

Fig. 1 Genome of South China tiger. **a** A male South China tiger. **b** Circos plot of South China tiger genomic features. Outer to inner: pseudo-chromosome, gene density (500-kb window) (with higher gene density indicated by higher blue column), GC content (500-kb window), and SNP density (500-kb window). **c** Distribution of the captive South China tigers in China. Yellow circles show 15 city locations of the captive South China tigers, including Luoyang, Zhengzhou, Linyi, Suzhou, Shanghai, Hangzhou, Nanchang, Chongqing, Chengdu, Guiyang, Changsha, Fuzhou, Longyan, Shaoguan, and Guangzhou. Circle size is proportional to the number of South China tigers in each city. All data are from the South China tiger studbook (2020). Red cross represents the sampling site of the South China tigers in our study

The resulting scaffolds were further clustered into chromosome-scale scaffolds using the Hi-C data (Additional file 2: Figure S2). Finally, the de novo assembly yielded 2.44 Gb of genomic sequences with a contig N50 at 6.13 Mb and a scaffold N50 at 150.19 Mb (Fig. 1b and Additional file 1: Table S4). The de novo assembly contained 19 pseudo-chromosomes anchored with 2.40 Gb of contigs (99.35%) and 2.42 Gb of scaffolds (99.36%), showing a high collinearity with the domestic cat (*Felis catus*) reference genome (FelCat9.0, Ensembl release 98, last access in September 2019), except for E3 chromosome (Additional file 1: Table S5 and Additional file 2: Figure S3). In the analysis of complete Benchmarking Universal Single-Copy Orthologs (BUSCO), our assembled genome covered 95.5% of the BUSCO genes [27] (Additional file 1: Table S6). By integrating the homology- and de novo-based predictions, 20,908 protein-coding genes were annotated (Additional file 1: Tables S7-S9). There were 844.92 Mb (34.98% of the genome size) of repetitive elements as well as 568 microRNAs (miRNA), 6,309 transfer RNAs (tRNA), 993 ribosomal RNAs (rRNA), and 1410 small nuclear RNAs (snRNA) (Additional file 1: Table S10). Altogether, we assembled and annotated the South China tiger genome (Amotig1.0) [28] as one of the top-quality chromosome-level genomes of all big cats (Additional file 1: Table S4) [29–36].

To explore genetic variations in the captive South China tiger population, we re-sequenced the whole genomes of 29 South China tigers collected from four major zoos in China (Fig. 1c). A domestic cat was also re-sequenced and used as an outgroup. We generated around 1200 Gb of whole-genome sequencing data with an average coverage depth of 15.63× per genome [37]. We combined our data with the 40 published genomes of six tiger subspecies (10.38-29.59× coverage depths) [38, 39] (Additional file 1: Tables S11-S13 and Additional file 2: Figures S4 and S5). Altogether, we identified 10.21 million high-quality single nucleotide polymorphisms (SNPs) in these 69 tiger genomes after stringent quality control and alignment of all sequencing data against the Amotig1.0 genome (Additional file 1: Tables S14 and S15 and Additional file 2: Figure S6).

Genetic variations among and demographic history of tiger subspecies

All tiger subspecies have experienced severe population bottlenecks due to human hunting and habitat loss/fragmentation, we thus compared the levels of their genetic variations. It was evident that the Sumatran tiger carried the least genetic diversity in terms of genome-wide heterozygosity and nucleotide diversity ($\pi = 0.553 \times 10^{-3}$) (Fig. 3a, Additional file 1: Table S16, and Additional file 2: Figures S7 and S8), but the highest genetic differentiation

from other tiger subspecies ($F_{ST} = 0.324$–0.459) (Additional file 1: Table S17 and Additional file 2: Figure S9), confirming the observations of previous studies [1, 4]. However, the genetic diversity of the South China tiger was not as low ($\pi = 0.657 \times 10^{-3}$) (Additional file 1: Table S16 and Additional file 2: Figure S7) as what was inferred from its pedigree records [9], and the genomic heterozygosity of the South China tiger was moderate among six tiger subspecies (Additional file 1: Table S15 and Additional file 2: Figure S8). Nonetheless, the South China tiger also showed a significant genetic differentiation from other tiger subspecies ($F_{ST} = 0.278$–0.459) (Additional file 1: Table S17 and Additional file 2: Figure S9).

To clarify the phylogenetic relationships among six tiger subspecies, we performed neighbor-joining (NJ) phylogenetic reconstructions, principal component analysis (PCA), and a model-based ancestry estimation using ADMIXTURE software to infer their population genetic structure. We reconstructed the NJ trees based on pairwise genetic distances with the domestic cat as an outgroup, which supported the taxonomic status of six distinct tiger subspecies [1, 3, 40, 41] (Fig. 2a). The PCA (Additional file 2: Figure S10), maximum-likelihood tree (Additional file 2: Figure S11), and identity-by-state analysis (Additional file 2: Figures S12 and S13) all verified their phylogenetic relationships at the subspecies level. However, the South China tigers were further differentiated into two lineages (e.g., lineages 1 and 2), except one particular tiger labeled as the ptam_1 to be a potential hybrid, while the Amur, Sumatran, and Malayan tigers formed three additional genetic clusters at $K = 5$ (Fig. 2b). All six tiger subspecies were differentiated from each other at $K = 9$, despite potential gene flow among some tiger subspecies (Fig. 2b). Additionally, the population genetic structure of only captive South China tigers verified the two genomic lineages at $K = 2$ (Additional file 2: Figure S14), which mirrored the Shanghai and Guiyang founder lineages [10, 11]. Meanwhile, the pairwise sequential Markovian coalescent (PSMC) plots showed that all tigers experienced continuous bottlenecks since the onset of the Last Glacial Period (Fig. 2c and Additional file 2: Figure S15), a pattern that was observed in a previous study [1].

Limited gene flow from other tiger subspecies into the captive South China tiger population

To detect a signal of potential genetic admixture/introgression from other tiger subspecies into the South China tiger population (e.g., the lineages 1 and 2 as well as the ptam_1 tiger), we applied several population genetic analyses, including the ABBA-BABA
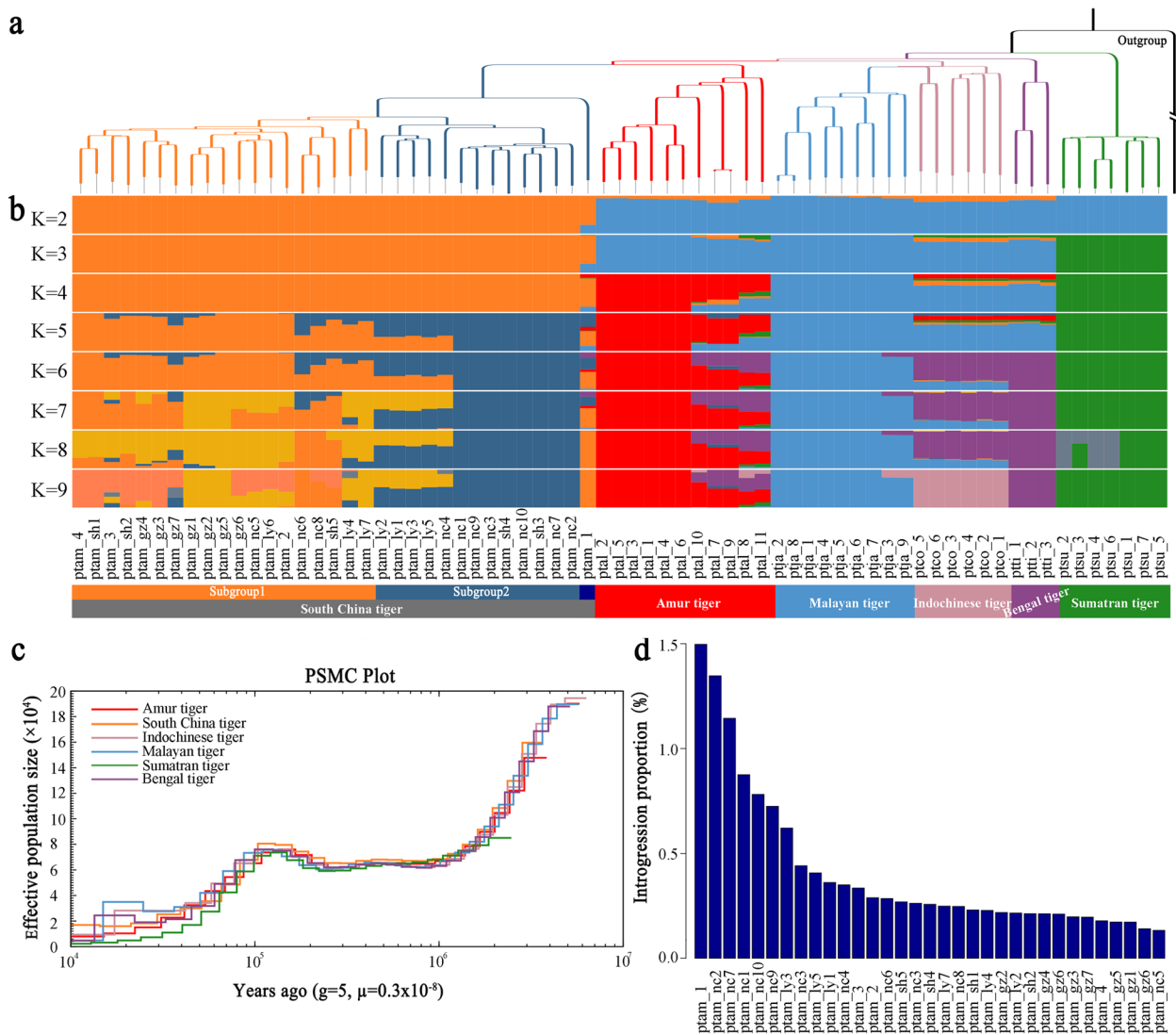
Fig. 2 Genetic components of extant South China tigers. **a** Phylogenetic relationship of the South China tiger with other tiger subspecies, with a domestic cat as an outgroup, all nodes are 100% of reliability. **b** Population genetic structuring of different tiger subspecies. **c** PSMC plot of the inferred demographic histories of each tiger subspecies. **d** Average ratio of introgression between the South China tiger and other tiger subspecies

[42, 43], Dsuite [44], and TreeMix methods [45]. All results clearly indicated the highest admixture in the ptam_1 tiger [1] as a hybrid from the Indochinese tiger (Fig. 2b, d, and Additional file 2: Figures S16 and S20). Although genetic admixture occurred among tiger subspecies, there was limited gene flow from other tiger subspecies into the lineages 1 and 2 of the South China tiger (Additional file 1: Tables S16 and S17 and Additional file 2: Figures S17-S19). Altogether, we observed very limited introgression (0.13–1.50%) among the South China tiger genomes (Fig. 2d), indicating their genetic uniqueness to be warranted for a full protection.

Genomic inbreeding among tiger subspecies

To characterize the speculated inbreeding depression among the South China tigers, we determined pairwise relatedness between individuals within each tiger subspecies using allelic identity-by-descent (IBD) [46] (Fig. 3b and Additional file 1: Table S20). Further, we calculated individual inbreeding based on genome-wide SNPs using the inbreeding coefficients F_H (a measure of the increase in individual SNP homozygosity compared with mean Hardy-Weinberg expected homozygosity) and F_{ROH} [47] (based on the runs of homozygosity (ROH) $\geq$ 100 kb). Among all tiger subspecies, the highest levels of the relatedness, F_H, and F_{ROH} in ROH < 1 Mb were observed

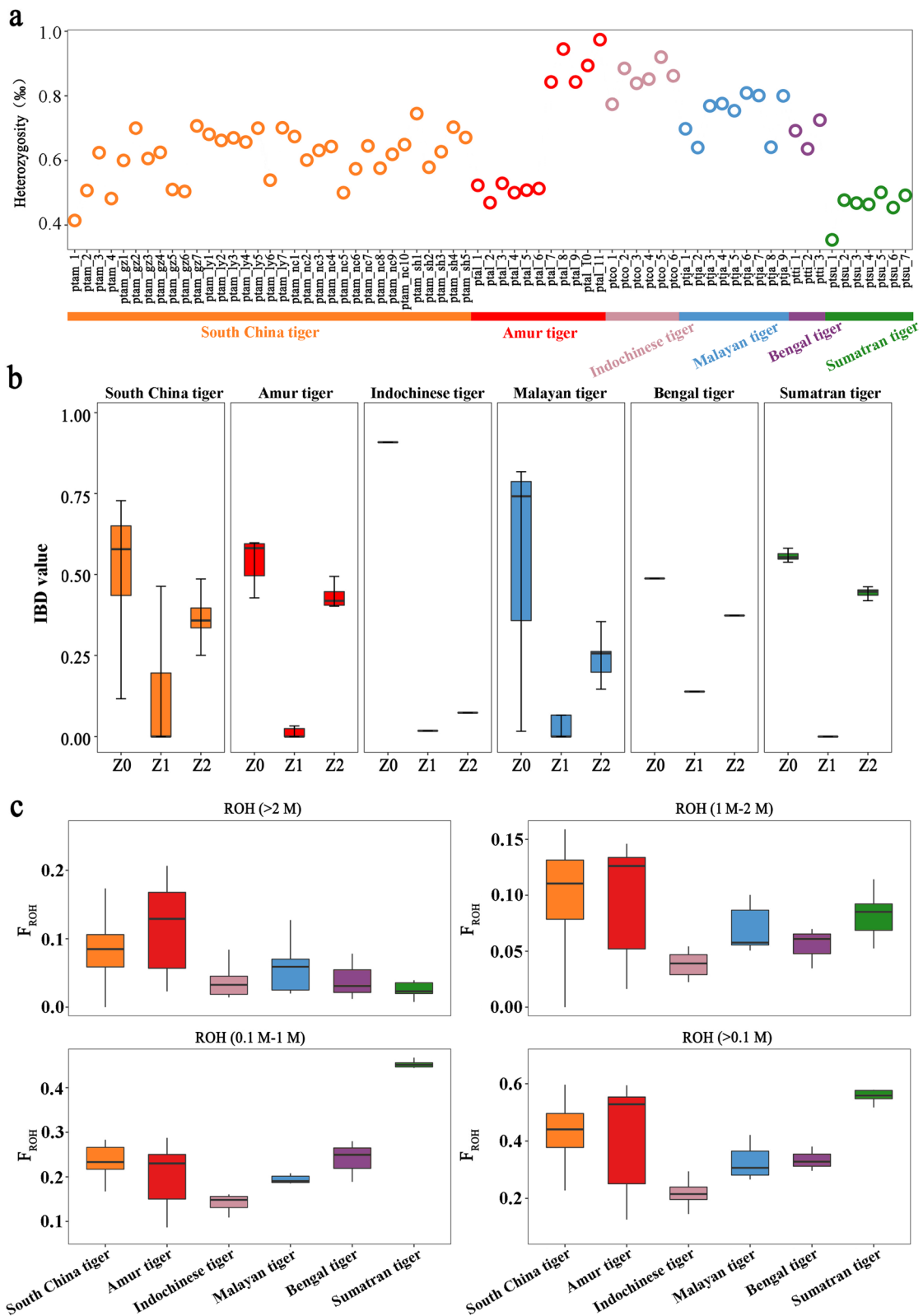Wang *et al. BMC Biology*   (2023) 21:64

Page 6 of 16

Fig. 3 Comparison of genetic variations among different tiger subspecies. **a** Genome-wide heterozygosity per individual. **b** Pairwise relatedness based on allelic identity-by-descent (IBD), that is, genetic identity because of a recent common ancestor. Z0, Z1, and Z2 are the probabilities that two individuals share neither, one or two alleles IBD. **c** Genomic inbreeding coefficients (F_{ROH}) based on different lengths of runs of homozygosity (ROH), with a minimum length of 100 kb

Wang *et al. BMC Biology*   (2023) 21:64

Page 7 of 16

in the Sumatran tiger, however, its F_{ROH} values were reduced down to the lowest in ROH > 2 Mb, indicating the effects of strong founder and/or bottleneck events occurred in the recent past (Additional file 1: Table S21 and Additional file 2: Figure S21), while the Amur tiger seemed to have experienced constant inbreeding pressure as evidenced by their second highest relatedness, F_H, and F_{ROH} values in ROH > 100 kb, which were increased to the highest in ROH > 2 Mb. On the other hand, the South China tigers shared a similar pattern of their relatedness as what was observed among other tiger subspecies, but they had relatively high F_{ROH} values in ROH > 100 kb and 1 Mb, compared with the Bengal, Indochinese, and Malayan tigers, which were likely the impact of recent inbreeding/founder events coupled with contraction of their effective population size (Fig. 3c). Among the four major zoos, the South China tigers in the Shanghai Zoo had the most number of ROH (average number = 3293, ROH = 100 kb–1 Mb), but only all the South China tigers in the Shanghai Zoo and the hybrid ptam_1 did not have any long ROH (> 2 Mb) (Additional file 1: Table S21), suggesting the captive tigers in the Shanghai Zoo to be least inbred. Furthermore, the South China tiger had the lowest F_{ROH} (average value = 0.33, ROH ≥ 100 kb) (Additional file 1: Table S21). Although the pedigree-based inbreeding coefficients (F_p) among the South China tigers were as high as 0.1796–0.5048 [9], their relatedness was more strongly correlated with both F_H and F_{ROH} than F_p when high-density SNPs were available [48] (Additional file 2: Figure S22). The genome-wide SNP-based individual inbreeding estimates are thus recommended for assisting the decision-making of captive breeding of the South China tiger.

Accumulation and purging of deleterious mutations among tiger subspecies

Because of the importance of managing deleterious mutations in conserved genomic elements for species conservation [49], we computed the proportion of deleterious mutations retained in potential coding regions of six tiger subspecies and obtained a total of 70,273 SNPs in three categories of high-, moderate-, and low-impact (nearly neutral) mutations, of which only 0.51% were highly deleterious as most of them were stop-gain mutations across six tiger subspecies (Additional file 1: Table S22). All tiger subspecies shared a similar distribution pattern in the proportions of these three categories of mutations (Fig. 4a and Additional file 1: Table S22). Nevertheless, the average number of homozygous sites with high- and moderate-impact mutations of the South China tiger (9.27 and 632.51, respectively) was significantly lower than other tiger subspecies (19.17–29.66 and 1255.67–2198.42, respectively) (*t* test, *P* < 0.01), while the average

number of homozygous sites with low-impact mutations of the South China tiger (13131.88) was close to other tiger subspecies (11127.83–12914.29) (Additional file 1: Table S23), an indication of a stronger genetic purging of the high- and moderate-impact mutations from the South China tiger. Additionally, the average proportions of high- and moderate-impact mutations in a homozygous state were the lowest (20.24% and 20.47%, respectively) among the South China tiger compared with other tiger subspecies (23.59–42.94% and 25.29–48.34% for high- and moderate-impact mutations, respectively) (Fig. 4b and Additional file 1: Table S23), indicating a relatively effective genetic purging of homozygous genotypes of such mutations. Moreover, this was particularly true that the average proportion of homozygous genotypes of both high- and moderate-impact mutations to all deleterious mutations was much lower than that of the low-impact mutations in the South China tiger, a pattern largely shared by other tiger subspecies (Additional file 1: Table. S23 and Additional file 2: Figure S23). The distribution of derived alleles of such high- and moderate-impact mutations displayed a downward shift compared to the low-impact mutations in the South China tiger (Fig. 4c). Besides, we calculated *Rxy* [50] to estimate whether there was an excess or deficit of deleterious mutations in the South China tiger compared with the other tiger subspecies. The relative mutation loads of *Rxy* were reduced below 1 in both high- and moderate-impact mutations in the South China tiger compared with both Amur and Sumatran tigers (Fig. 4d). Altogether, these results indicated a relatively effective genetic purging of both high- and moderate-impact mutations in a homozygous state from the South China tiger following its population contraction with a controlled increase in inbreeding based on pedigree records, a phenomenon that was observed in other species [51–54].

Discussion

To facilitate the best practice of conservation genomics based on high-density markers across the entire genomes [55] to mitigate potential inbreeding loads associated with deleterious mutations of all endangered tigers, we first assembled the best South China tiger genome and applied it as a reference to the analyses of all tiger genomic data [13, 20–23]. We analyzed the largest population of the South China tiger combined with other tiger subspecies and found that all six tiger subspecies phylogenetically differentiated from each other, despite gene flow between some tiger subspecies [1–3]. The admixture from the Amur, Indochinese, and Bengal tigers was found in the captive South China tiger population by previous genomic studies [1, 15], so the South China tiger was thought to be the
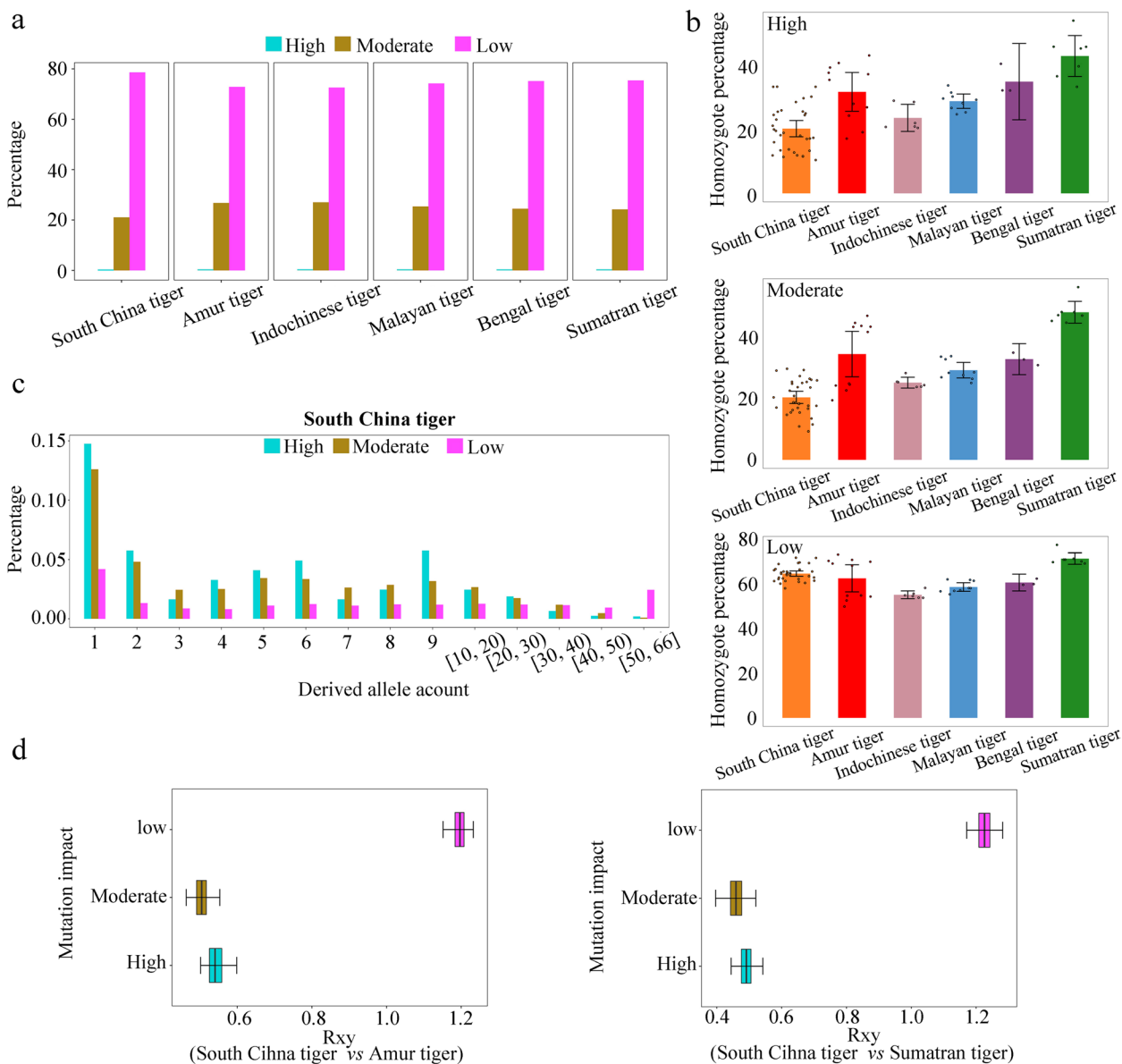
Wang *et al. BMC Biology*   (2023) 21:64

Page 8 of 16

Fig. 4 Comparison of deleterious mutations among different tiger subspecies. **a** Percentage of three categories of nearly neutral (low), mildly (moderate), and highly deleterious (high impact) mutations in different tiger subspecies. **b** Individual homozygote percentage per mutation category of different tiger subspecies. **c** Site frequency spectra for SNPs per mutation category in the South China tiger. Derived allele counts ≥ 10 are displayed as mean counts per interval. **d** *Rxy* analysis compares the South China tiger with the Sumatran and Amur tigers per mutation category. *Rxy* < 1 indicates a relative frequency deficit of the corresponding category

most taxonomically controversial among all the tiger subspecies. From 2004, the Tiger TAG collected blood samples from all newborn tigers for genetic analysis to remove apparently admixed tigers from breeding, including the ptam_1 from the Chongqing Zoo [1] that was obviously a hybrid from the Indochinese tiger (Additional file 2: Figures S16 and S20). In this study, we found that the South China tiger had very limited genetic admixture from other tiger subspecies,

validating very little or negligible genetic contamination in the captive South China tiger population. However, we only re-sequenced the whole genomes of 29 South China tigers collected from the four major zoos in China. As such, we shall continue and focus our efforts on collecting and re-sequencing additional whole genomes of most, if not all, reproductive South China tigers, to further fine-map and effectively manage their viable genomic landscapes along with

Wang *et al. BMC Biology*   (2023) 21:64

Page 9 of 16

unique genetic variations and putatively deleterious mutations.

Evaluation of genomic variations in all tiger subspecies showed that the Sumatran tiger had the lowest genetic diversity, while the South China tiger harbored moderate genetic variations. We also found genomic inbreeding in the Sumatran, Amur, and South China tigers. Moreover, the accumulation of large ROH (> 1 Mb) since the beginning of captive breeding has been resulting in relatively high genomic inbreeding and genetic burden in the South China tiger [56]. It was reported that the captive South China tiger population suffered from inbreeding depression in terms of high juvenile mortality and impaired adult fertility [12]. To control the inbreeding based on pedigree records, the captive South China tigers were permitted to mate only if their mate suitability indexes (MSI, ranging from 1 to 6 across all tigers) calculated by PMx software [57], were below 4. The MSI is a composite score that integrates four genetic components into a single index, including Delta GD, differences in Mk values, inbreeding coefficient, and unknown ancestry [58]. Therefore, the breeding plan based on the MSI may have assisted most deleterious mutations to be inherited as recessive alleles masked in heterozygous states and thus free of purifying selection [59, 60]. Our genomic analysis, however, revealed an efficient genetic purging of both high- and moderate-impact deleterious mutations from the South China tiger genomes. It is known that the purging reduces the frequency of deleterious mutations, depending on the degree of dominance and the magnitude of the deleterious effects. Because most high-impact mutations were stop-gain variants leading to loss of gene functions, the high-impact mutations were exposed to the highest strength of purging [51]. It was true that the average number of homozygous sites with high-impact mutations per individual (9.27) was much lower than those of the moderate- (632.51) and low-impact (13,131.88) mutations in the South China tiger genomes. Surprisingly, the lowest proportion of homozygous genotypes and relatively low Rxy estimates (< 1) were associated with both high- and moderate-impact mutations in the South China tiger genomes. This verified that both high- and moderate-impact mutations were more likely to be inherited as recessive alleles than low-impact mutations in the captive South China tiger population [12]. Based on recent founder effect and ongoing inbreeding, these deleterious mutations are expected to be continuously accumulated in the South China tiger genomes, therefore their potential impacts on fitness should be evaluated across generations of the captive South China tiger population [10, 11].

It is certain that all captive South China tigers are the descendants of two males and four female tigers [9]. The Shanghai and Guiyang founder lineages were formed during the 1970s, but the two lineages began to merge for breeding to minimize potential inbreeding depression since the mid-1980s [11]. The two genetic lineages based on current genome-wide SNPs and early mtDNA and microsatellite analyses [15] clearly mirrored the two founder lineages, due probably to historical allelic segregation and/or genetic drift among limited founders of separate geographical origins and relatively independent reproduction of the South China tigers in isolated zoos, as what was observed in the killer whale ecotypes [46] and highly inbred pigs [61]. The Tiger TAG has been in charge of the breeding and transfer plan for South China tiger, following a principle of priority to allow the tigers to breed in their original facilities first, which may have limited the exchange of breeding tigers among the zoos [11]. According to the studbook, the descendants of the Guiyang lineage were much fewer than those of the Shanghai lineage [9]. We also recognized that the tigers of the two lineages were not evenly distributed among the zoos, for example, the seven tigers sampled from the Guangzhou Zoo for this study all belonged to the Shanghai lineage. This calls for a genomics-assisted exchange plan of breeding tigers across all major zoos. In fact, the Shanghai and Guiyang lineages showed different reproductive performance and fitness among the newborn tigers [62]. Our genomic analysis detected the lowest genomic inbreeding of the South China tigers in the Shanghai Zoo. We therefore recommend a genomics-informed exchange of breeding tigers between these two founder/genomic lineages, so that to maximize the benefit of maintaining existing unique and critical genetic variations that are expected to outweigh the cost of increased genetic load within the captive South China tiger population [20].

Conclusions

In this study, we de novo assembled the high-quality chromosome-level reference genome of a South China tiger and re-sequenced the whole genomes of 29 South China tigers collected from four major breeding zoos in China. The results indicated that the captive South China tigers included in our study (expect the pam_1) had limited genetic admixture from other tiger subspecies. The genetic diversity was slightly higher in the South China tiger than in the Sumatran tiger, while the South China tiger had high F_{ROH} values under longer ROH (> 1 Mb), indicating its recent inbreeding/founder events and/or population bottleneck/isolation. Although most high- and moderate-impact deleterious mutations in the South China tiger genomes may be masked as recessive alleles for their inheritance in heterozygous states, such deleterious SNPs have been effectively purged, when they were in homozygous states, from the South China tiger

Wang *et al. BMC Biology*   (2023) 21:64

Page 10 of 16

population. We recommend that individuals of the two genomic lineages could be merged to breed for minimizing a further loss of the unique and critical genetic variations. All the new findings from our present study demonstrated the power and effectiveness of concerted efforts to conserve the captive South China tiger in the past and shed light into a potentially bright future of these critically endangered cats. The captive South China tigers are the last hope of the tiger subspecies, we wish that they would have a success of reintroduction program in the future and to be another case as what is achieved in the protection of the Giant Panda, which was downlisted to 'vulnerable' on the IUCN Red List.

Methods

De novo genome sequencing

Blood samples of two male (Tuantuan and Kangkang) and ione female (Huanhuan) South China tiger were acquired from the Guangzhou Zoo for genome sequencing. All the DNA was extracted from blood using phenol-chloroform method [63]. All libraries were constructed at BGI (Shenzhen, China). For PacBio sequencing, genomic DNA from the Tuantuan was used to construct three libraries of 20-kb insert size using the SMRTbell Template Prep Kit 1.0 (Pacific Biosciences, USA), and the fragment size was selected using Blue Pippin (Labgene Scientific SA, Switzerland). Then, the library sequencing was performed on a PacBio RS II sequencer. For short reads sequencing, three short insert libraries (270, 500, and 800 bp) were constructed for Tuantuan and four long mate-paired insert libraries (2, 5, 10, and 20 kb) were constructed for Huanhuan. The libraries were constructed as per the manufacturer's instructions (MGIEasy universal DNA Library Preparation Kit, BGI, China). All libraries were sequenced on the Illumina HiSeq X Ten Platform using 150-bp paired-end reads according to the Illumina protocols, except for the 800-bp library, which was sequenced on the HiSeq 2500 System using 125-bp paired-end reads.

Estimation of genome size

We applied K-mer (17-mer) distribution analysis to estimate the South China tiger genome size using clean reads from the short-insert libraries. Genome size was calculated using the formula: Genome size = K-mer_number / K_depth of peak.

Genome assembly

The PacBio data were de novo assembled using WTDBG-1.2.8 (https://github.com/ruanjue/wtdbg-1.2.8) and genomic contigs were polished with the Arrow program (https://www.pacb.com/support/software-downloads/) by aligning SMRT reads, which yielded an assembly (2.43 Gb in length) with N50 size at 6.24 Mb. We also aligned the Illumina X Ten data to the assembly using BWA (v0.7.15) [64] for error correction by Pilon (v1.22) [65] with the parameter "--mindepth 6", which is an integrated tool for comprehensive variant detection and genome assembly improvement. The final assembly generated a total length of 2.42 Gb and N50 length at 6.20 Mb.

Construction of optical genome maps

DNA of sufficient quality was extracted and labeled from the Tuantuan blood cells according to standard BioNano protocols (BioNano Genomics), after which nicking, labeling, repairing, and staining processes were implemented. We digested DNA using a specific single-stranded nicking endonuclease (Nt.BspQI). BioNano Solve (v3.0.1) [66] was used to produce optical maps with single molecules above 100 kb in size and six labels per molecule. The scaffold-level assembly had a N50 length at 31.62 Mb.

Hi-C analysis

The Hi-C library was constructed from the blood cells of Kangkang according to the standard procedures of BGI and sequenced using the MGISEQ-2000 Platform. The Bionano-based scaffolds were anchored into a chromosome-scale assembly using a Hi-C proximity-based assembly approach. We aligned the Hi-C reads to scaffolds using bowtie2 (v2.2.5) [67] and interaction maps were generated following HiC-Pro (v2.5.0) [68] pipelines. The uniquely mapped read pairs were used as input for Juicer [69] and 3d-DNA [70] Hi-C analysis and scaffolding pipelines. The resulting Hi-C contact maps were visualized using Juicebox [71], and mis-assemblies and mis-joins were manually corrected based on neighboring interactions. The preliminary chromosome assembly was then generated and named as Amotig1.0. The Amotig1.0 genome was assessed by calculating the number of Hi-C read pairs in any two bins of 500 kb. The synteny analysis between the Amotig1.0 and the domestic cat genomes (Ensembl release 98, accessed in September 2019) was performed using the Mummer-4.0.0 software (https://github.com/gmarcais/mummer). The chromosome IDs in the Amotig1.0 genome were determined according to the synteny relationship with the cat genome. To evaluate the quality of the Amotig1.0 genome, we performed the BUSCO pipeline (version 5.0) [27] with orthologs database of mammalia_odb10.

Genome annotation

We identified repetitive sequences in the Amotig1.0 genome using a combination of homology- and de novo-based methods. First, RepeatMasker [72] and RepeatProteinMask [72] were used to search repeat sequences

against with Repbase database [73, 74]. Second, LTRharvest (GenomeTools v1.5.9) [75] and RepeatModeler (http://www.repeatmasker.org/RepeatModeler/) were used to build a de novo repeat library, and then repeats were annotated using RepeatMasker [72] with default parameters. Last, tandem repeats were detected using Tandem Repeats Finder (TRF) [76].

For gene structure prediction, we used homology-based prediction based on the protein sequences from five species (*Felis catus*, *Homo sapiens*, *Mus musculus*, *Panthera pardus*, and *Panthera tigris altaica*) downloaded from the Ensembl database (release 93). These protein sequences were mapped to the Amotig1.0 genome using TBLASTN (*E*-value cutoff: 1e^{-5}) [77]. High-scoring segment pairs (HSPs) were concatenated using Solar (v0.9.6) [25]. GeneWise (v2.4.1) [78] was used to define accurate gene models. We then merged and filtered redundancy from different homology results based on the GeneWise score ($\geq$ 0.4). To obtain the final gene set, transposons and single-exon genes without functional annotations were filtered out.

Gene functional annotations were assigned using BLASTP (BLAST+ v2.2.26) [77] against public databases, including the Swiss-Prot (release-2017_09) [79], TrEMBL (release-2017_09) [79], KEGG (v84.0) [80], COG [81], and NCBI nucleotide collection nr/nt (https://www.ncbi.nlm.nih.gov/nucleotide/, v20170924). The motifs and domains in the protein sequences were annotated using InterProScan (v5.16-55.0) [82].

The tRNA genes were predicted using tRNAscan-SE (v1.3.1) [83] with eukaryote parameters. The rRNA fragments were identified by aligning the rRNA template sequences from the Human Rfam database [84] using BLASTN (BLAST+ v2.2.26) [77] (*E*-value 1e^{-5}). The snRNAs and miRNAs were searched via a two-step method: i.e., aligned with BLAST and then searched with INFERNAL (infernal-1.1.1) [85] against the Rfam database.

Sampling information and whole genome re-sequencing

We collected a total of 30 specimens, including 29 South China tigers (*P. t. amoyensis*) from four major zoos in China and a domestic cat (*Felis catus*) from the Guangzhou Zoo, China (Supplementary Table S11). The 29 South China tigers were born in 1999 to 2018, and the blood samples of the individuals were collected in four major city zoos in 2018 and 2019. Genomic DNA was extracted from whole blood using the DNeasy Blood & Tissue Kit (QIAGEN, Valencia, California, USA) following the manufacturer's protocols. For each sample, one 350 bp size of DNA library was constructed according to the manufacturer's protocols (Illumina). The constructed libraries were sequenced using the Illumina Hiseq X Ten platform for 150 bp paired-end reads.

Quality control

To ensure reads were reliable and without artificial bias (e.g., low-quality paired reads that result from base-calling duplicates and adapter contamination), we conducted a series of quality control (QC) procedures, as follows:

(1) Removed reads with $\geq$ 1% unidentified nucleotides (*N*).
(2) Removed reads with > 40% bases having phred quality < 20.
(3) Removed reads with > 10 nucleotides overlapping the adapter (allowing $\leq$ 10% mismatches).

Read alignment and variant calling

We used BWA [86] to align the clean reads of each sample against our newly assembled South China tiger genome (settings: mem -t 4 -k 32 -M -R). Alignment files were converted to BAM files using SAMtools (settings: -bS -t) (v-0.1.19) [87]. In addition, potential PCR duplications were removed using Picard (http://broadinstitute.github.io/picard/). We called SNPs using the Haplotype-Caller approach implemented in the Genome Analysis Toolkit (GATK) package [88]. Filtering criteria were as follows:

(1) SNPs with QD < 2.0; FS > 60.0; MQ < 40.0; QUAL < 30; DP < 4.0; MQRankSum < -12.5; and ReadPosRankSum < -8.0 were filtered.
(2) Multi-nucleotide polymorphisms were ignored.
(3) SNPs within 5 bp of a gap were filtered.
(4) Overall depth (for all individuals) was > 1/3× and < 3×.
(5) Unobserved variant allele constituted < 10%.

A total of 54,067,600 high-quality SNPs were retained for subsequent analyses after filtering. Gene-based SNP annotation was performed using ANNOVAR [89].

Phylogenetic and population genetic analysis

We selected genome-wide 54,067,600 SNPs for phylogenetic construction with a cat as outgroup and 10,205,707 SNPs for PCA and population structure analyses components. PCA was performed using PLINK (v1.9) (settings: --bfile --pca –noweb) [90]. Genetic structure was inferred using ADMIXTURE (v1.3) [91], with implementation of a block-relaxation algorithm (settings: --cv -m=block). To explore convergence of individuals, we predefined the number of genetic clusters *K* from 2 to 9. We calculated the *p*-distance matrix using VCF2Dis (https://github.com/BGI-shenzhen/VCF2Dis) and a NJ tree was generated using the R package APE and 100 bootstraps were run for a reliable tree [92].

Wang *et al. BMC Biology*   (2023) 21:64

Page 12 of 16

Calculation of genetic diversity

Pairwise nucleotide diversity θ_π and Watterson's estimator θ_w [93] within a tiger subspecies were calculated using a sliding-window approach (20-kb windows sliding in 10-kb steps). Genetic differentiation between tiger subspecies was calculated using the pairwise fixation index F_{ST} [94].

Identity-by-state (IBS) and identity-by-descent (IBD) analyses

To evaluate the similarity between two tigers within a tiger subspecies, genome-wide IBS pairwise identities were calculated using the toolset SNPRelate [95] in the R package. Based on the matrix of genome-wide IBS pairwise distances, we performed multidimensional scaling and cluster analyses and determined the groups using a permutation score. We also calculated the pairwise IBD using PLINK (v1.9) (settings: --file --genome --min 0.05) [90].

Introgression and demographic history analyses

We separated the South China tiger population into three classes (i.e., lineage 1, lineage 2, and ptam1) based on phylogeny. TreeMix (v1.13) [96] was used to detect gene flow between the tiger subspecies based on genome-wide allele frequency data at individual SNPs. We first constructed a maximum-likelihood tree for six tiger subspecies using blocks of 10,000 SNPs. The number of migration events was set from 1 to 6. We calculated introgression among six tiger subspecies using Patterson's D-statistic (ABBA-BABA test) [97], with the cat as the outgroup, and tested the proportions of admixture events (*f4*-ratio) within each South China tiger using Dsuite (v0.4) [44], we calculated the mean introgression ratio between each South China tiger and other tiger subspecies.

We used the PSMC [98] model to reconstruct demographic history. To estimate effective ancestral population size changes for individual tiger subspecies, we selected individuals with high sequencing depth to ensure the quality of the consensus sequence. SNPs were detected using SAMtools [87], sites were filtered based on a minimum depth (DP = 4) and the highest depth (DP = 50) and mapping quality (Q = 20). We only retained autosomal SNPs. Parameters were set to: -N30, -t15, -r5, -p'4+25*2+4+6'. A mutation rate of 3.5×10^{-9} per base per generation and generation time of 5 years was used.

Analysis of runs of homozygosity

Regions of homozygosity were extracted for all chromosomes of all individuals based on SNP information.

PLINK (v1.9) [90] was used to detect ROH via a sliding-window approach, with the following parameters: '--homozyg-window-snp 100 --homozyg-window-het 2 --homozyg-window-missing 5 --homozyg-snp 100 --homozyg-kb 100 --homozyg-density 10 --homozyg-gap 100'.

Inbreeding coefficient

We measured individual inbreeding using the genomic inbreeding coefficients F_H [48], which is the fraction of IBD of the two alleles in a diploid individual from a common ancestor. F_H was calculated using PLINK (v1.9) [90]. Alternatively, individual genomic inbreeding coefficients was also measured using F_{ROH} [99], which is an estimate of ROH proportion in an individual genome.

Identification of deleterious mutations

The variants leading to functional changes were regarded as candidates of deleterious mutations. Thus, we only analyzed the SNPs in all potential coding regions, while the domestic cat allele was regarded as an ancestral allele. SnpEff (v4.3t) [100] was used for genetic variant annotation and functional effect prediction.

Abbreviations

SNPs	Single nucleotide polymorphisms
PCA	Principal component analysis
NJ	Neighbor-joining
IBD	Identity-by-descent
ROH	Runs of homozygosity
F_p	Pedigree-based inbreeding coefficients
TRF	Tandem Repeats Finder
HSPs	High-scoring segment pairs
QC	Quality control
GATK	Genome Analysis Toolkit
IBS	Identity-by-state

Additional file 1: Table S1. Sample characteristics and genome sequencing information before and after filtering. **Table S2.** Sequencing data statistics. **Table S3.** 17-mer statistic information. **Table S4.** Genome assembly information. **Table S5.** Sizes of individual chromosomes. **Table S6.** Assessment of completeness of the South China tiger genome assembly. **Table S7.** Summary of repeat contents. **Table S8.** Statistics of gene structure prediction. **Table S9.** Result of gene functional annotation. **Table S10.** Summary of non-coding RNA in genome. **Table S11.** Sampling information included in the analyses. **Table S12.** Re-sequencing data statistics in this study. **Table S13.** Genome mapping and coverage information for all accessions analyzed in this study. **Table S14.** Number and distribution of SNPs in each genome. **Table S15.** Distribution of whole-genome SNPs in different subspecies/species. **Table S16.** Average genome-wide nucleotide diversity in six tiger subspecies. **Table S17.** Pairwise F_{ST} values between six tiger subspecies. **Table S18.** ABBA-BABA estimates. Only results with |Z score| > 3 were remained. **Table S19.** *f4*-ratio calculated between individual South China tigers and other tiger subspecies. Only results with FDR < 0.05 were retained. **Table S20.**

Pairwise relatedness estimates based on allelic identity-by-descent (IBD). **Table S21.** Inbreeding coefficients and ROH information. **Table S22.** Deleterious mutations segregating across and within six tiger subspecies. **Table S23.** Individual homozygote and heterozygote SNP counts per impact category in each tiger subspecies.

Additional file 2: Figure S1. Distribution of 17-mer depth for estimating the genome size. The x-axis represents k-mer depths while y-axis represents the proportions. The blue line represents the proportion of 17-mer in each depth. The peak depth is at 25-fold and total number of 17-mer is 61,791,522,108. The South China tiger (*P. t. amoyensis*) genome size was estimated to be 2471.66 Mb from the formula: Genome size = K-mer_number/K_depth of peak. **Figure S2.** Hi-C chromosomal contact heat map. A 500 kb resolution was used to calculate the number of Hi-C read pairs in any two bins. **Figure S3.** The collinearity between the South China tiger and domestic cat (*Felis catus*) genomes. Each dot represents an aligned region while the minimum length is 10 kb. The red dot represents forward comparison and the blue dot reverse comparison. **Figure S4.** Q30 and GC content for each sample. The average Q30 and GC content are 93.24% and 42.04%, respectively. Their very low variation reflected our re-sequencing data to be high quality. **Figure S5.** The distribution of mapping rates and average mapping depths for each individual. The depths ranged from 12.91× to 18.96 × while the rates varied from 96.03% to 98.92%. **Figure S6.** The distribution of the SNP number and frequency. Number of genome-wide SNPs (top) and frequency of SNPs per 1 kb of each tiger sample (bottom). **Figure S7.** Nucleotide diversity π estimates of six tiger subspecies. **Figure S8.** Heterozygosity statistics of genome-wide SNPs. (a) Observed heterozygosity across all individual genomes of each tiger subspecies. (b) Genomic heterozygosity in each tiger subspecies at population level. **Figure S9.** Pairwise F_{ST} values between six tiger subspecies. The weighted F_{ST} values are shown above the diagonal while their standard deviations below the diagonal. **Figure S10.** Principal component analysis using genome-wide SNPs of six tiger subspecies. **Figure S11.** A maximum likelihood tree was built using TreeMix software with whole-genome sequencing data of six tiger subspecies and cat. **Figure S12.** Cluster analysis based on the matrix of genome-wide identity-by-state (IBS) pairwise distances between six tiger subspecies determined by a permutation score. **Figure S13.** Analysis of genome-wide average identity-by-state (IBS) pairwise identities between six tiger subspecies. **Figure S14.** Population genetic structure of the South China tigers estimated by the ADMIXTURE. **Figure S15.** Demographic histories of tiger subspecies. PSMC analysis shows the change in effective population size over time. The dash lines represent 100 bootstraps. The representative individuals sequenced at a high read coverage were selected for each graph, South China tiger (ptam_4, mean 25.38×), Amur tiger (ptal_1, mean 14.76×), Indochinese tiger (ptco_2, mean 13.98×), Malayan tiger (ptja_7, mean 13.13×), Bengal tiger (ptti_3, mean 13.79×), and Sumatran tiger (ptsu_1, mean 13.41×). **Figure S16.** The results of Dsuite. Heatmap showing statistical support for introgression between pairs of tiger subspecies. Cells in the heatmap indicate the pairwise Z score values between the branch b identified on the expanded tree on the Y axis (relative to its sister branch) and the taxa P3 identified on the X-axis. The grey color indicates the none. (a) Result of branch for South China tiger (lineage 1) with the other tiger subspecies populations. (b) Result of branch for South China tiger (lineage 2) with the other tiger subspecies populations. (c) Result of branch for South China tiger (ptam1 individual) with the other tiger subspecies populations. **Figure S17.** The graph of ABBA-BABA test. The values of D (pop1, pop2; pop3, pop4) >0 indicates that there are gene flows between pop1 and pop3. Here, only the introgression occurred in South China tiger will be shown. (a) D-statistic for South China tiger (lineage 1) with other tiger subspecies. (b) D-statistic for South China tiger (lineage 2) with other tiger subspecies. (c) D-statistic for South China tiger (ptam1 individual) with other tiger subspecies. **Figure S18.** Plot of inferred introgression between the South China tiger lineage 1 and other tiger subspecies populations detected by the TreeMix method. The scale bar shows 10 times the average standard error of the entries in the sample covariance matrix. **Figure S19.** Plot of inferred introgression between the South China tiger lineage 2 and other tiger subspecies populations detected by the TreeMix method. The scale bar shows 10 times the average standard error of the entries in the sample covariance matrix. **Figure**

S20. Plot of inferred introgression between the South China tiger ptam_1 individual and other tiger subspecies populations detected by the TreeMix method. The scale bar shows 10 times the average standard error of the entries in the sample covariance matrix. **Figure S21.** Genomic inbreeding coefficients F_H in each tiger subspecies. **Figure S22.** The plot of three inbreeding coefficients (F_P, F_H, and F_{ROH}) of the South China tiger. **Figure S23.** Comparison of homozygote and heterozygote percentage of per mutation category among six tiger subspecies. The significant value was calculated by *t*-test.

Availability of data and materials
Genome assemblies and DNA sequencing data were deposited in the National Genomics Data Center (NGDC, https://ngdc.cncb.ac.cn) under Project no. PRJCA006384. The de novo genome of the South China tiger was deposited under Accession ID GWHBEIN00000000 [28]. The re-sequencing genomic data of 29 South China tigers and a domestic cat were deposited under Accession ID CRA004909 [37].

Declarations

Ethics approval and consent to participate
All necessary research permits and ethics approvals for this study were granted by Guangzhou Zoo, China (Project number GZ_20180216).

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Guangzhou Zoo & Guangzhou Wildlife Research Center, Guangzhou 510070, China. [2]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650201, China. [3]Kunming Natural History Museum of Zoology, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, Yunnan, China. [4]Kunming College of Life Science, University of the Chinese Academy of Sciences, Kunming 650204, China. [5]Shanghai Zoo, Shanghai 200336, China. [6]CAAS-ILRI Joint Laboratory on Livestock and Forage Genetic Resources, Institute of Animal Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100193, China. [7]International Livestock Research Institute (ILRI), Nairobi 00100, Kenya. [8]Wangcheng Park, Luoyang 471000, China. [9]Suzhou Shangfangshan Forest Zoo, Suzhou 215009, China. [10]Chongqing Zoo, Chongqing 401326, China. [11]Nanchang Zoo, Nanchang 330025, China. [12]Chinese Association of Zoological Gardens, Beijing 100037, China. [13]Qinghai Province Key Laboratory of Crop Molecular Breeding, Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining 810008, Qinghai, China.

Wang *et al. BMC Biology*   (2023) 21:64

Page 14 of 16

References

1. Liu YC, Sun X, Driscoll C, Miquelle DG, Xu X, Martelli P, et al. Genome-wide evolutionary analysis of natural history and adaptation in the world's tigers. Curr Biol. 2018;28:3840–9.e6.
2. Luo SJ, Kim JH, Johnson WE, van der Walt J, Martenson J, Yuhki N, et al. Phylogeography and genetic ancestry of tigers (*Panthera tigris*). PLoS Biol. 2004;2:e442.
3. Luo SJ, Liu YC, Xu X. Tigers of the world: genomics and conservation. Annu Rev Anim Biosci. 2019;7:521–48.
4. Armstrong EE, Khan A, Taylor RW, Gouy A, Greenbaum G, Thiéry A, et al. Recent evolutionary history of tigers highlights contrasting roles of genetic drift and selection. Mol Biol Evol. 2021;38:2366–79.
5. IUCN. The IUCN Red List of Threatened Species. 2021. https://www.iucnredlist.org. Accessed on 1 Feb 2022.
6. Li MY, Zhou Q, Huang WQ. Prediction of potential habitat for South China tiger (*Panthera tigris*) based on GIS and web datasets. Procedia. Environ Sci. 2011;10(part B):1613–20.
7. Tilson R, Defu H, Muntifering J, Nyhus PJ. Dramatic decline of wild South China tigers *Panthera tigris amoyensis*: feld survey of priority tiger reserves. Oryx. 2004;38:40–7.
8. Kathy TH, Xie Z, Yin YZ. The struggle to save the last South China tigers. Tigers of the world (second edition): William Andrew Publishing; 2010. p. chapter:457-61.
9. Yin YZ. Studbook of the South China tiger. Chongqing: Chongqing Zoo Press; 2017.
10. Xu YC, Fang SG, Li ZK. Sustainability of the South China tiger: Implications of inbreeding depression and introgression. Conserv Genet. 2007;8:1199–207.
11. Yuan YH, Pei EL, Liu QX. Reproductive parameters of female South China tigers in captivity. Eur J Wildl Res. 2020;66:37.
12. Yuan YH, Yin YZ, Liu QX. Inbreeding depression and population viability analysis of the South China tigers (*Panthera tigris amoyensis*) in captivity. Mamm Biol. 2021;101:803–9.
13. Tilson R, Traylor-Holzer K, Jiang QM. The decline and impending extinction of the South China tiger. Oryx. 1997;31:243–52.
14. Guo J. Year of the tiger. Nature. 2007;449:16–8.
15. Zhang WP, Xu X, Yue BS, Hou R, Xie JJ, Zou ZT, et al. Sorting out the genetic background of the last surviving South China tigers. J Hered. 2019;110:641–50.
16. Keller LF, Waller DM. Inbreeding effects in wild populations. Trends Ecol Evol. 2002;17:230–41.
17. Hedrick PW, Garcia-Dorado A. Understanding inbreeding depression, purging, and genetic rescue. Trends Ecol Evol. 2016;31:940–52.
18. Reed DH, Frankham R. Correlation between fitness and genetic diversity. Conserv Biol. 2003;17:230–7.
19. Vandewoestijne S, Schtickzelle N, Baguette M. Positive correlation between genetic diversity and fitness in a large, well-connected metapopulation. BMC Biol. 2008;6:46.
20. Khan A, Patel K, Shukla H, Viswanathan A, van der Valk T, Borthakur U, et al. Genomic evidence for inbreeding depression and purging of deleterious genetic variation in Indian tigers. Proc Natl Acad Sci U S A. 2021;118:e2023018118.
21. Paszkiewicz K, Studholme DJ. *De novo* assembly of short sequence reads. Brief Bioinform. 2010;11:457–72.
22. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, et al. Real-time DNA sequencing from single polymerase molecules. Science. 2009;323:133–8.
23. Schwartz DC, Li X, Hernandez LI, Ramnarain SP, Huff EJ, Wang YK. Ordered restriction maps of Saccharomyces cerevisiae chromosomes constructed by optical mapping. Science. 1993;262:110–4.
24. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009;326:289–93.
25. Li RQ, Fan W, Tian G, Zhu HM, He L, Cai J, et al. The sequence and *de novo* assembly of the giant panda genome. Nature. 2010;463:311–7.
26. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods. 2013;10:563–9.
27. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31:3210–2.
28. Wang C, Wu DD, Yuan YH, Yao MC, Han JL, Wu YJ, et al. Population genomic analysis provides evidence of the past success and future potential of South China tiger captive conservation. NGDC accession: GWHBEIN00000000. 2022. https://ngdc.cncb.ac.cn/gwh/Assembly/22233/show. Accessed 9 Jan 2023.
29. Zhang L, Lan TM, Lin CY, Fu WY, Yuan YH, et al. Chromosome-scale genomes reveal genomic consequences of inbreeding in the South China tiger: a comparative study with the Amur tiger. China National GenBank accession: CNP0001654. 2022. https://db.cngb.org/search/project/CNP0001654/. Accessed 9 Jan 2023.
30. Bredemeyer KR, Murphy WJ. Unpublished liger assemblies. NCBI accession: GCA_018350195.1. 2021. https://www.ncbi.nlm.nih.gov/assembly/GCA_018350195.1/. Accessed 9 Jan 2023.
31. Bredemeyer KR. Murphy WJ. Unpublished liger assemblies. NCBI accession: GCA_018350215.1. 2021. https://www.ncbi.nlm.nih.gov/assembly/GCF_018350215.1/. Accessed 9 Jan 2023.
32. Armstrong EE, Campana MG, Solari KA, Morgan SR, Ryder OA, Naude VN, et al. Genome report: Chromosome-level draft assemblies of the snow leopard, African leopard, and tiger (*Panthera uncia*, *Panthera pardus pardus*, and *Panthera tigris*). NCBI accession: GCA_024034525.1. 2022. https://www.ncbi.nlm.nih.gov/assembly/GCA_024034525.1/. Accessed 9 Jan 2023.
33. Armstrong EE, Campana MG, Solari KA, Morgan SR, Ryder OA, Naude VN, et al. Genome Report: Chromosome-level draft assemblies of the snow leopard, African leopard, and tiger (*Panthera uncia*, *Panthera pardus pardus*, and *Panthera tigris*). NCBI accession: GCA_023721935.1. 2022. https://www.ncbi.nlm.nih.gov/assembly/GCF_023721935.1/. Accessed 9 Jan 2023.
34. Armstrong EE, Taylor RW, Miller DE, Kaelin C, Barsh G, Hadly EA, et al. Long live the king: chromosome-level assembly of the lion (*Panther leo*) using linked-read, Hi-C, and long read data. NCBI accession: GCA_008795835.1. 2019. https://www.ncbi.nlm.nih.gov/assembly/GCA_008795835.1/. Accessed 9 Jan 2023.
35. Suryamohan K, Shukla H, Khan A, Zachariah A, Ramakrishnan U, Seshagiri S. Sequencing and de novo assembly of two Bengal tiger genomes. NCBI accession: GCA_021130815.1. 2021. https://www.ncbi.nlm.nih.gov/assembly/GCA_021130815.1/. Accessed 9 Jan 2023.
36. Suryamohan K, Shukla H, Khan A, Zachariah A, Ramakrishnan U, Seshagiri S. NCBI accession: GCA_021131075.1. 2021. https://www.ncbi.nlm.nih.gov/assembly/GCA_021131075.1/. Accessed 9 Jan 2023.
37. Wang C, Wu DD, Yuan YH, Yao MC, Han JL, Wu YJ, et al. Population genomic analysis provides evidence of the past success and future potential of South China tiger captive conservation. NGDC accession: CRA004909. 2022. https://ngdc.cncb.ac.cn/gsa/browse/CRA004909. Accessed 9 Jan 2023.
38. Liu YC, Sun X, Driscoll C, Miquelle DG, Xu X, Martelli P, et al. Genome-wide evolutionary analysis of natural history and adaptation in the world's tigers. NCBI accession:PRJNA437782. 2018. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA437782. Accessed 9 Jan 2023.
39. Du H, Yu J, Li Q, Zhang M. New evidence of tiger subspecies differentiation and environmental adaptation: comparison of the whole genomes of the Amur tiger and the South China tiger. NCBI accession: PRJNA484850. 2018. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA484850/. Accessed 9 Jan 2023.
40. Xue HR, Yamaguchi N, Driscoll CA, Han Y, Bar-Gal GK, Zhuang Y, et al. Genetic ancestry of the extinct Javan and Bali tigers. J Hered. 2015;106:247–57.
41. Driscoll CA, Yamaguchi N, Bar-Gal GK, Roca AL, Luo S, Macdonald DW, et al. Mitochondrial phylogeography illuminates the origin of the extinct caspian tiger and its relationship to the amur tiger. PLoS One. 2009;4:e4125.
42. Zeng L, Liu HQ, Tu XL, Ji CM, Gou X, Esmailizadeh A, et al. Genomes reveal selective sweeps in kiang and donkey for high-altitude adaptation. Zool Res. 2021;42:450–60.
43. Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, et al. A draft sequence of the Neandertal genome. Science. 2010;328:710–22.

44. Malinsky M, Matschiner M, Svardal H. Dsuite - Fast D-statistics and related admixture evidence from VCF files. Mol Ecol Resour. 2021;21:584–95.
45. Zeng L, Tu XL, Dai H, Han FM, Lu BS, Wang MS, et al. Whole genomes and transcriptomes reveal adaptation and domestication of pistachio. Genome Biol. 2019;20:79.
46. Foote AD, Vijay N, Ávila-Arcos MC, Baird RW, Durban JW, Fumagalli M, et al. Genome-culture coevolution promotes rapid divergence of killer whale ecotypes. Nat Commun. 2016;7:11693.
47. Kardos M, Åkesson M, Fountain T, Flagstad Ø, Liberg O, Olason P, et al. Genomic consequences of intensive inbreeding in an isolated wolf population. Nat Ecol Evol. 2018;2:124–31.
48. Kardos M, Luikart G, Allendorf FW. Measuring individual inbreeding in the age of genomics: marker-based measures are better than pedigrees. Heredity. 2015;115:63–72.
49. van Oosterhout C. Mutation load is the spectre of species conservation. Nat Ecol Evol. 2020;4:1004–6.
50. Do R, Balick D, Li H, Adzhubei I, Sunyaev S, Reich D. No evidence that selection has been less effective at removing deleterious mutations in Europeans than in Africans. Nat Genet. 2015;47:126–31.
51. Grossen C, Guillaume F, Keller LF, Croll D. Purging of highly deleterious mutations through severe bottlenecks in Alpine ibex. Nat Commun. 2020;11:1001.
52. Feng S, Fang Q, Barnett R, Li C, Han S, Kuhlwilm M, et al. The genomic footprints of the fall and recovery of the Crested Ibis. Curr Biol. 2019;29:340–49.e7.
53. Xue Y, Prado-Martinez J, Sudmant PH, Narasimhan V, Ayub Q, Szpak M, et al. Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding. Science. 2015;348:242–5.
54. Xie HX, Liang XX, Chen ZQ, Li WM, Mi CR, Li M, et al. Ancient demographics determine the effectiveness of genetic purging in endangered lizards. Mol Biol Evol. 2022;39:msab359.
55. Supple MA, Shapiro B. Conservation of biodiversity in the genomics era. Genome Biol. 2018;19:131.
56. Zhang L, Lan TM, Lin CY, Fu WY, Yuan YH, et al. Chromosome-scale genomes reveal genomic consequences of inbreeding in the South China tiger: A comparative study with the Amur tiger. Mol Ecol Rescour. 2022. https://doi.org/10.1111/1755-0998.13669.
57. Ballou JD, Lacy RC, Pollak JP. PMx: software for demographic and genetic analysis and management of pedigreed populations (version 1.0). Brookfield: Chicago Zoological Society; 2011.
58. Traylor-Holzer K. PMx users manual, Version 1.0. Apple Valley: IUCN SSC Conservation Breeding Specialist Group; 2011.
59. Wang MS, Zhang JJ, Guo X, Li M, Meyer R, Ashari H, et al. Large-scale genomic analysis reveals the genetic cost of chicken domestication. BMC Biol. 2021;19:118.
60. Housden BE, Muhar M, Gemberling M, Gersbach CA, Stainier DY, Seydoux G, et al. Loss-of-function genetic tools for animal models: cross-species and cross-platform differences. Nat Rev Genet. 2017;18:24–40.
61. Wang L, Mu Y, Xu L, Li K, Han J, Wu T, et al. Genomic analysis reveals specific patterns of homozygosity and heterozygosity in inbred pigs. Animals. 2019;9:314.
62. Chen GL, Li ZK, Song PL, Jin K, Shen QY. The current situation and the pedigree analysis of South China tiger. Chinese J Zool. 2001;36(4):45–8 (in Chinese).
63. Green MR, Sambrook J. Isolation of high-molecular-weight dna using organic solvents. Cold Spring Harb Protoc. 2017;2017(4):pdb.prot093450.
64. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754–60.
65. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One. 2014;9:e112963.
66. Lam ET, Hastie A, Lin C, Ehrlich D, Das SK, Austin MD, et al. Genome mapping on nanochannel arrays for structural variation analysis and sequence assembly. Nat Biotechnol. 2012;30:771–6.
67. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9.
68. Servant N, Varoquaux N, Lajoie BR, Viara E, Chen CJ, Vert JP, et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol. 2015;16:259.
69. Durand NC, Shamim MS, Machol I, Rao SS, Huntley MH, Lander ES, et al. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. Cell Syst. 2016;3:95–8.
70. Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, et al. *De novo* assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. Science. 2017;356:92–5.
71. Robinson JT, Turner D, Durand NC, Thorvaldsdóttir H, Mesirov JP, Aiden EL. Juicebox.js provides a cloud-based visualization system for Hi-C data. Cell Syst. 2018;6:256–8.e1.
72. Tarailo-Graovac M, Chen N. Using RepeatMasker to identify repetitive elements in genomic sequences. Curr Protoc Bioinformatics. 2009;Chapter 4:Unit 4.10.
73. Bao W, Kojima KK, Kohany O. Repbase Update, a database of repetitive elements in eukaryotic genomes. Mob DNA. 2015;6:11.
74. Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. Repbase Update, a database of eukaryotic repetitive elements. Cytogenet Genome Res. 2005;110:462–7.
75. Gremme G, Steinbiss S, Kurtz S. GenomeTools: a comprehensive software library for efficient processing of structured genome annotations. IEEE/ACM Trans Comput Biol Bioinform. 2013;10:645–56.
76. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999;27:573–80.
77. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.
78. Birney E, Clamp M, Durbin R. GeneWise and Genomewise. Genome Res. 2004;14:988–95.
79. Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, et al. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res. 2003;31:365–70.
80. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28:27–30.
81. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, et al. The COG database: an updated version includes eukaryotes. BMC Bioinformatics. 2003;4:41.
82. Mulder N, Apweiler R. InterPro and InterProScan: tools for protein sequence classification and comparison. Methods Mol Biol. 2007;396:59–70.
83. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 1997;25:955–64.
84. Griffiths-Jones S, Moxon S, Marshall M, Khanna A, Eddy SR, Bateman A. Rfam: annotating non-coding RNAs in complete genomes. Nucleic Acids Res. 2005;33:D121–4.
85. Nawrocki EP, Kolbe DL, Eddy SR. Infernal 1.0: inference of RNA alignments. Bioinformatics. 2009;25:1335–7.
86. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv. 2013:arXiv:1303.3997. https://arxiv.org/abs/1303.3997.
87. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25:2078–9.
88. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20:1297–303.
89. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38:e164.
90. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75.
91. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 2009;19:1655–64.
92. Paradis E, Claude J, Strimmer K. APE: Analyses of phylogenetics and evolution in R language. Bioinformatics. 2004;20:289–90.
93. Nei M, Li WH. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci U S A. 1979;76:5269–73.
94. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution. 1984;38:1358–70.

Wang *et al. BMC Biology*   (2023) 21:64

Page 16 of 16

95. Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS. A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics. 2012;28:3326–8.

96. Pickrell JK, Pritchard JK. Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet. 2012;8:e1002967.

97. Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, et al. Ancient admixture in human history. Genetics. 2012;192:1065–93.

98. Li H, Durbin R. Inference of human population history from individual whole-genome sequences. Nature. 2011;475:493–6.

99. Keller MC, Visscher PM, Goddard ME. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. Genetics. 2011;189(1):237–49.

100. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 2012;6:80–92.

Publisher's Note