Online ISSN: 2588-3526
Journal of Wildlife and Biodiversity
*Volume 9(3): 278-295 (2025) (*http://www.wildlife-biodiversity.com/*)*
**Research Article**

# NUMTs as diagnostic markers for hybrid identification in Tigers: A novel reference-free approach

**Wlaa Mohamed[1], Yan Chun Xu[1,2]*, Ahmed Fadol[1]**

[1]College of Wildlife and Protected Area, Northeast Forestry University, Harbin, 150040, China
[2]National Forestry and Grassland Administration Research Center of Engineering Technology for Wildlife Conservation and Utilization, Harbin, 150040, China
University, Harbin 150040, China
*Email: wanghongshu88@outlook.com

Received: 25 June 2025 / Revised: 17 October 2025 / Accepted: 12 October 2025/ Published online: 14 October 2025.
How to cite: Mohamed, W.., Xu, Y.C., Fadol, A. (2025). NUMTs as diagnostic markers for hybrid identification in Tigers: A novel reference-free approach, Journal of Wildlife and Biodiversity, 9(3), 278-295. DOI: https://doi.org/10.5281/zenodo.17386708

## Abstract

Maintaining genetic purity is critical in conservation breeding programs, especially for endangered subspecies like the South China tiger (*Panthera tigris amoyensis*) and the Amur tiger (*Panthera tigris altaica*), where hybridization threatens subspecies integrity and reintroduction success. However, identifying hybrids is challenging due to unreliable morphological traits and limited access to verified purebred genomes. Here, we present Auto-Ref HybrID, a novel reference-free method that leverages nuclear mitochondrial DNA segments (NUMTs) as genetic markers to distinguish purebred and hybrid individuals using whole genome sequencing (WGS) data. We analyzed 400 fecal samples from captive tigers and found that hybrid individuals exhibit significantly higher NUMT counts, elevated mutation rates, and greater divergence from the maternal mitogenome compared to purebreds. Genetic distance metrics and phylogenetic clustering confirmed the presence of hybridization signals, even at shallow sequencing depths. Our results establish NUMTs as robust, scalable markers for hybrid detection and demonstrate that Auto-Ref HybrID is a powerful tool for assessing gene pool contamination in conservation programs where reference genomes are lacking.

**Keywords:** Nuclear mitochondrial DNA (NUMTs): Hybrid identification; Panthera *tigris altaica*, *Panthera tigris amoyensis*; Gene pool contamination

## Introduction

The Amur tiger (*Panthera tigris altaica*), also known as the Siberian tiger, is the largest subspecies of tiger found mainly in the Russian Far East and north-eastern China. It is adapted to cold climates and has thick fur and a large body circumference to retain heat (Mazak, 1981)Despite conservation efforts, habitat loss and poaching remain a major threat to its survival(Miquelle et al., 2005).The South China tiger (*Panthera tigris amoyensis*), once native to southern China, is now considered functionally extinct in the wild due to extensive habitat destruction and hunting(Tilson et al., 2004) It is one of the most critically endangered tiger subspecies, with only a small population maintained in captivity. Conservation efforts focus on breeding programs with the intention of reintroductions into the wild.(Tilson & Nyhus, 2009)The Amur tiger (*Panthera tigris altaica*) is designated as endangered by the International Union for Conservation of Nature (IUCN) Red List(Goodrich et al., 2022)Poaching and habitat damage have historically caused its population to drop, but conservation efforts have helped stabilize numbers, and today there are an estimated 500–600 individuals in the wild, mostly in northeastern China and Russia. However, ongoing threats such as habitat fragmentation and human-wildlife conflict continue to pose risks.(Goodrich et al., 2022; Harihar et al., 2014; Hebblewhite et al., 2011; Zhang et al., 2023).

The last confirmed sighting of a wild individual occurred in the 1990s, and extensive surveys have failed to locate any remaining wild populations(Tilson et al., 2004)Fewer than 200 individuals exist in captivity, and conservation programs focus on breeding and potential reintroduction efforts(Guoliang et al., 2001; SJ, 2004; Tilson et al., 1997; Traylor-Holzer et al., 2010; Wang et al., 2023; Zhang et al., 2019)Conservation breeding programs have played a critical role in maintaining the genetic diversity and survival of both the Amur and South China tigers, particularly as their wild populations face severe threats. The South China tiger (*Panthera tigris amoyensis*) is now considered functionally extinct in the wild, with all known individuals residing in captivity.

(Maguire & Lacy, 1990; Todesco et al., 2016) Around 150–200 South China tigers are kept in 26 zoos across China, with efforts focused on breeding and potential reintroduction programs. However, the population suffers from inbreeding due to the small founder population, raising concerns about genetic health and long-term viability. Many individuals in captivity descend from just six wild-caught founders, resulting in a limited gene pool that compromises adaptability. Genetic drift and accumulation of deleterious alleles may further reduce fitness, making successful

reintroduction more challenging. Molecular genetic monitoring has become essential to track diversity and manage mating strategies. Recent efforts have also explored advanced reproductive technologies and genome-based selection to optimize genetic outcomes for the species.

The Amur tiger (*Panthera tigris altaica*) has a much larger captive population, with approximately 5,000 individuals housed in breeding facilities across China(Luo et al., 2019). These captive programs aim to support genetic diversity and provide individuals for potential rewilding or conservation initiatives. Despite these efforts, the reintroduction of captive-bred Amur tigers into the wild remains complex due to challenges such as human-wildlife conflict and habitat availability(Arumugam & Annavi, 2019; Ballou, 1992; Earnhardt, 1999; Ning et al., 2024; Russello et al., 2004; Veasey, 2020; Wang et al., 2016; Yachmennikova et al., 2022)Hybridization in captive breeding should be avoided to preserve the genetic purity and ecological adaptations of distinct tiger subspecies. Mixing subspecies, such as Amur and South China tigers, can lead to the loss of unique traits necessary for survival (Asa et al., 2011; Russello et al., 2004; Snyder et al., 1996)in specific environments, reducing the fitness of individuals for potential reintroduction(Dinerstein et al., 2007)Hybrids may also complicate conservation goals, as they do not align with species recovery programs and are often ineligible for release into the wild. Additionally, maintaining hybrids in breeding centers diverts limited resources from genetically pure individuals, which are crucial for conservation efforts(Arumugam & Annavi, 2019; Ballou, 1992; Iucn, 2013).

Additionally, maintaining hybrids in breeding centers diverts limited resources from genetically pure individuals, which are crucial for conservation efforts. Ethical and legal concerns further discourage hybridization, as international conservation guidelines prioritize maintaining subspecies integrity to ensure the long-term survival of tigers in their natural habitats.

Potential hybrid chance in captivity for tigers, mostly due to ambiguity and misidentification of tiger subspecies. And the need to investigate gene pool contamination of tigers and sort out the pure tigers with conservation values(Jackiw et al., 2015; Wang et al., 2023).

Identifying hybrid tigers is challenging because morphological characteristics are unreliable due to the high similarity among subspecies. Traditional identification based on physical traits, such as coat patterns and body size, cannot accurately distinguish hybrids from purebred individuals. Genetic approaches using genome analysis provide a more reliable method, but they typically require known purebred reference samples for comparison. However, obtaining such purebred

references is nearly impossible due to historical interbreeding and the lack of verified genetic lineages in captivity. To overcome this limitation, the AutoRef approach developed in this study offers a novel method to identify hybrids without the need for pre-established reference genomes. The purpose of this study 1 is to evaluate the effectiveness of NUMTs in separating subspecies of tigers. (2) To evaluate the effectiveness of Auto-Ref HybrID to identify tiger subspecies.

## Material and methods

### Collection of samples and data

The collection of samples for this study involved gathering 100 fecal samples from South China tigers and 300 fecal samples from Amur tigers. The South China tiger samples were obtained from various zoos across China and selected based on their breeding programs and the availability of purebred individuals. Each zoo provided detailed records of the tigers' lineage and breeding history to ensure the validity of the samples. The fecal samples were collected non-invasively to minimize stress and disruption to the animals. In contrast, the Amur tiger samples were collected from the Heilongjiang Siberian Tiger Park, a key conservation population known for its breeding and preservation efforts.

### DNA extraction

Sample treatment and DNA extraction were performed using **PEERS** (Peri-Extraction Enrichment of host DNA from fecal Samples), a method developed in our laboratory, technique is optimized to enrich host DNA while minimizing environmental and dietary contaminants present in fecal matter. The procedure included the following steps:

**Sample Collection and Preparation** Fresh tiger fecal samples were collected from the field using sterile gloves and instruments to prevent cross-contamination. Each sample was divided into two portions for preservation: one portion was placed in 95% ethanol (v/v) to inhibit microbial activity and preserve DNA integrity, and the second portion was stored at −80 °C immediately after collection to prevent enzymatic degradation. Samples preserved in ethanol were kept at 4 °C during transport, while frozen samples were maintained on dry ice and stored at −80 °C until processing.

**Pre-treatment with Sodium Dodecyl Sulfate (SDS) and Phosphate-Buffered Saline (PBS):** To each tube containing fecal material, 500 μL of 10 mmol/L phosphate-buffered saline (PBS) was added to rehydrate the sample. Next, sodium dodecyl sulfate (SDS), an anionic detergent used to disrupt cell membranes and solubilize proteins, was added to achieve a final concentration of

approximately 5% (w/v). The tubes were vigorously vortexed for 30 seconds to ensure homogeneous mixing. The mixture was then briefly centrifuged and allowed to stand at room temperature for 15–30 minutes to facilitate efficient cell lysis and release of host DNA into solution.

**Centrifugation and Supernatant Collection:** The centrifuge tubes were spun at 12,000 ×g for 10 min to pellet debris and intact cells. The supernatant, containing solubilized DNA, was carefully transferred to new 1.5 ml centrifuge tubes, avoiding carryover of any pellet material.

**DNA Extraction:** DNA extraction was performed using the AxyPrep Genomic DNA Mini Kit according to the manufacturer's instructions with minor modifications. To the supernatant in each 1.5 ml tube, 150 µL Buffer C-L and 20 µL Proteinase K were added and mixed thoroughly. After brief centrifugation, the tubes were incubated at 56°C for 10 minutes to facilitate protein digestion and DNA release.

**Purification and Elution:** Following digestion, 350 µL Buffer PD was added to each tube, vortexed for 20 s, and then centrifuged at 12,000 × g for 10 min to bind DNA to the membrane in the spin column. The flow-through was discarded, and the column was washed twice with Buffer W1 (500 µL each wash) and twice with Buffer W2 (700 µL each wash), followed by centrifugation at 12,000 × g for 1 min each. DNA was finally eluted from the column membrane by adding 100 µL Elution Buffer to the center of the membrane and centrifuging at 12,000 × g for 1 min.

**Whole genome sequencing**

For each DNA sample, a sequencing library was prepared using the MGIEasy Universal DNA Library Prepared Kit (MGI, China) according to the protocol provided by the manufacturer. The libraries were then sequenced on a DNBSEQ-T1 sequencer, generating pair-end reads each 150 base pairs long. The raw sequencing reads were processed using a standard quality control pipeline. First, low-quality bases and adapter sequences were removed using Trimmomatic v0.39. Reads (Bolger et al., 2014)with a quality score below Q30 or shorter than 36 base pairs after trimming were discarded. Quality metrics of the processed reads were assessed using FastQC v0.11.9(Andrews et al., 2010) to ensure the removal of artifacts, sequencing biases, and low-complexity regions. Subsequently, duplicate reads were filtered using Picard Tools v2.27.1[(Toolkit, 2019)]

**Assembly of the mitochondrial genome**

All tiger species ' cytochrome b (Cyt b) sequences were downloaded from NCBI. The accession numbers were 6262392 (*Panthera tigris altaica*) and 10020650 (*Panthera tigris amoyensis*). These sequences were used as seeds for *de novo* assembly of mitogenome contigs from whole genome sequencing data using NOVOplasty v4.3.1. A k-mer size of 33 bp was used, optimized for the 100bp read length. The Cyt b contigs were then used to recruit unmapped reads from the whole genome data to iteratively refine and assemble complete *de novo* mitogenomes for each sample.

**Identification of NUMTs**

Whole genome sequencing reads were mapped to the assembled mitogenomes using BWA. After the initial alignment, quality control steps were applied to extract potential NUMTs. The process began with aligning reads to the reference mitogenome using bwa mem to ensure high-quality and well-annotated references for accurate alignments. The resulting SAM file was then converted to BAM format using SAMtools to facilitate efficient downstream processing. Following this, samtools view was used to filter the BAM file, removing low-quality reads (mapping quality < 30), unmapped reads, and secondary alignments, retaining only reliable data. Extracted reads were then converted to BED format using bedtools bamtobed, facilitating the identification of NUMT regions. Redundant hits were minimized by clustering overlapping reads within 1000 bp using Bedtools Cluster and then converting them back to FASTQ format using Bedtools Bamtofastq. A BLAST search (blastn) against the mitogenome was performed to mask targeted reads, using the "-dust no" option to preserve informative regions. Finally, a Python script (parse_blast.py) extracted the top BLAST hits, ensuring that only the most relevant NUMT sequences were retained for subsequent analyses(Bazinet et al.).

**Hybrid signal detection**

To illustrate the diversification of NUMTs between hybridizable species, we take a sample from each of the hybridizable pairs and hypothetically assign it as the dam and the other as the sire. The NUMTs identified from each sample were aligned with the mitogenome assembly of the dam sample using MAFFT v7.475. A neighbor-joining tree was constructed using MEGA X with a p-distance model and 500 bootstrap replicas. At the same time, we constructed an NJ-tree for NUMTs in each hybrid sample using the same method, but all NUMTs were from the same sample (a part of which are maternal and the other paternal). Branch length shows the degree of differentiation among clades.

Secondly, the genetic distances of NUMTs to the mitogenome assembly (maternal) were computed using the Kimura 2-parameter (K2P) model in MEGA X, taking both transition and transversion substitutions into account. The genetic distance of purebred samples was defined as the intraspecific distance (intra-ds), and the distance of hybrid samples was defined as the interspecific distance (inter-ds). The 0.95 confidence intervals of intra-ds and inter-ds were estimated using 1000 times bootstrapping in R. The species resolution of NUMTs was assessed by comparing the upper limit of intra-ds and the lower limit of inter-ds; the greater the difference is, the higher the resolution to separate species.

### Mutation analysis

To detect mutations within NUMTs, sequences were aligned to the mitochondrial genome to identify substitutions, insertions, and deletions. Mutation rates were calculated as the number of mutations per site per generation. Comparisons were made between hybrids and purebred individuals of both Amur and South China tigers, focusing on understanding how mutations in hybrid genomes differ from those in the parental species. The analysis also considered the contribution of maternal and paternal mitochondrial DNA to NUMT mutations in the hybrid genomes. Mutation rates across different groups (hybrids, pure Amur tigers, and pure South China tigers) were compared using non-parametric statistical methods, such as Kruskal-Wallis and Mann-Whitney U tests. Correlation analyses explored potential relationships between hybrid indices (percentage of Amur tiger ancestry) and mutation rates in NUMTs. Additionally, phylogenetic analyses were performed to assess whether NUMTs in hybrids clustered differently from those in purebred individuals, indicating possible evolutionary shifts in mitochondrial integration into the nuclear genome.

## Results

### Characteristics of NUMTs

The mitogenomes assembled from re-sequencing data represented the parental species in purebred samples and the maternal species in hybrid samples, revealing significant differences in the characteristics of NUMTs. In purebred Amur tigers, the number of NUMTs ranged from 77 to 267 (mean = 322.25 ± 88.49), while in purebred South China tigers, the range was 50 to 146 (mean = 78.53 ± 32.12). Hybrid individuals showed higher NUMT counts, with 87 to 763 (mean = 525.10 ± 100.60) for Amur tiger NUMTs mapped to the South China tiger assembly and 73 to 313 (mean

= 193.23 ± 51.72) for South China tiger NUMTs mapped to the Amur tiger assembly. NUMT lengths were also variable, with the shortest mean length in purebred South China tigers (122.30 ± 87.58 bp) and the longest in hybrid Amur tiger NUMTs mapped to the South China tiger assembly (558.00 ± 200.60 bp). Sequence identity between NUMTs and homologous sequences on reference mitogenomes was highest in purebred Amur tigers (97–100%) and lowest in hybrids (65–98%).

The K2P genetic distance between NUMTs and homologous sequences was significantly smaller in purebred individuals, ranging from 0.105 ± 0.138 in Amur tigers to 0.256 ± 0.200 in South China tigers, compared to hybrids, where the mean distances were 0.164 ± 0.150 for Amur tiger NUMTs mapped to the South China tiger assembly and 0.350 ± 0.200 for South China tiger NUMTs mapped to the Amur tiger assembly. For purebred samples, the independent t-test comparing the K2P genetic distances between NUMTs and homologous sequences on reference mitogenomes resulted in a t-value of 2.334 and a p-value of 0.0006, suggesting that purebred animals have NUMTs that are closer in genetic distance to the mitogenome of their parental species. Similarly, for hybrid samples, the independent t-test resulted in a t-value of 4.564 and a p-value of 0.0009, indicating a statistically significant difference and confirming that hybrid animals have NUMTs less homologous to the maternal mitogenome. These findings demonstrate the potential of NUMT characteristics to distinguish between purebred and hybrid individuals.

**Table 3**. Comparison of nuclear mitochondrial DNA segments (NUMTs) between two tiger subspecies using self and cross-species mitochondrial genome (mitogenome) references

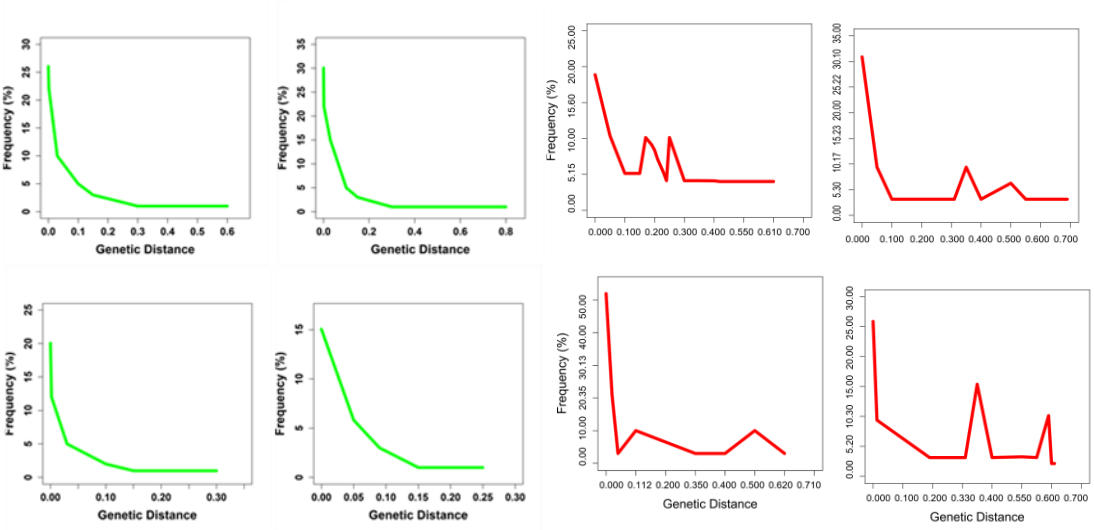
| Species | Mitogenome Reference | Number of NUMTs per individual | Mean number of NUMTs per individual | Length of NUMTs (bp) | Mean length of NUMTs (bp) | Identity between NUMTs and homologous sequence on reference | K2P distance between NUMTs and homologous sequence on | Mean of K2P distance |
|---|---|---|---|---|---|---|---|---|

| | | | | | | mitogenome (%) | reference mitogenome | |
|---|---|---|---|---|---|---|---|---|
| Amur tiger | Amur tiger assembly | 77- 267 | 322.25± 88.49 | 121-552 | 386.50± 200.00 | 97-100 | 0 -0.800 | 0.105 ± 0.138 |
| South China tiger | South China Tiger assembly | 50-146 | 78.53 ± 32.12 | 66-436 | 122.30 ± 87.58 | 90-100 | 0-0.630 | 0.256 ± 0.200 |
| Amur tiger | South China Tiger assembly | 87 -763 | 525.10±100.60 | 153-663 | 558.00± 200.60 | 65-98 | 0 -0.600 | 0.164± 0.150 |
| South China tiger | Amur tiger assembly | 73-313 | 193.23± 51.72 | 173-563 | 368.15± 150.20 | 76-93 | 0-0.750 | 0.350± 0.200 |

**Hybridization signals**

To investigate the presence of hybridization signals, we plotted the frequency of NUMTs against their genetic distance to the mitogenome assembly for both purebred and hybrid samples. As depicted in Figure3-1 and Figure S2, the purebred samples from both the Amur and South China tigers exhibited a clear, monotonically declining trend in the frequency of NUMTs as the genetic distance increased. This decline was marked by a steep initial phase and a more gradual, gentle decrease. In contrast, the hybrid samples displayed a distinct pattern with multiple peaks occurring in the more gradual declining phase, following the steep drop. These peaks, representing NUMTs, correspond to the paternal NUMTs significantly diverged from the maternal mitogenome. This divergence indicates that these paternal NUMTs are genetically distinct from the maternal mitogenome, serving as a clear hybridization signal. The presence of these peaks in hybrid individuals provides strong evidence of hybridization events and reinforces the ability to distinguish hybrid individuals from purebreds based on their NUMT profiles. This pattern underscores the utility of NUMTs as genetic markers in identifying hybridization events and distinguishing between purebred and hybrid subspecies in wildlife conservation and genetic studies. Besides the degree of differentiation between parents, the sensitivity of detection is also influenced by WGS depth. We tested the least detectable sequencing depth using serial subsampling of WGS data. As shown in Figure 4B, $\lg(D_{h/p})$ across all taxa was positively correlated with the number of reads in the manner $\lg(D_{h/p}) = 0.7509 + 0.2773 \times \lg(\text{Reads})$ ($R^2 = 0.7572$, $p = 0.00031$), and the average depth of NUMTs in the manner $\lg(D_{h/p})$ $0.2767 \times \lg(N) - 0.951$ ($R^2 = 0.7648$, $p = 0.0000291$). The lower limit of 0.95 CI of $D_{h/p}$ for hybridization signal peaks 0.2 and 0.55. Taking this value as the criterion to define a hybridization signal (valid peak), the corresponding average depth of NUMTs was ≥ 12×and the total data quantity was 1800 reads.

Figure A




Figure(B)

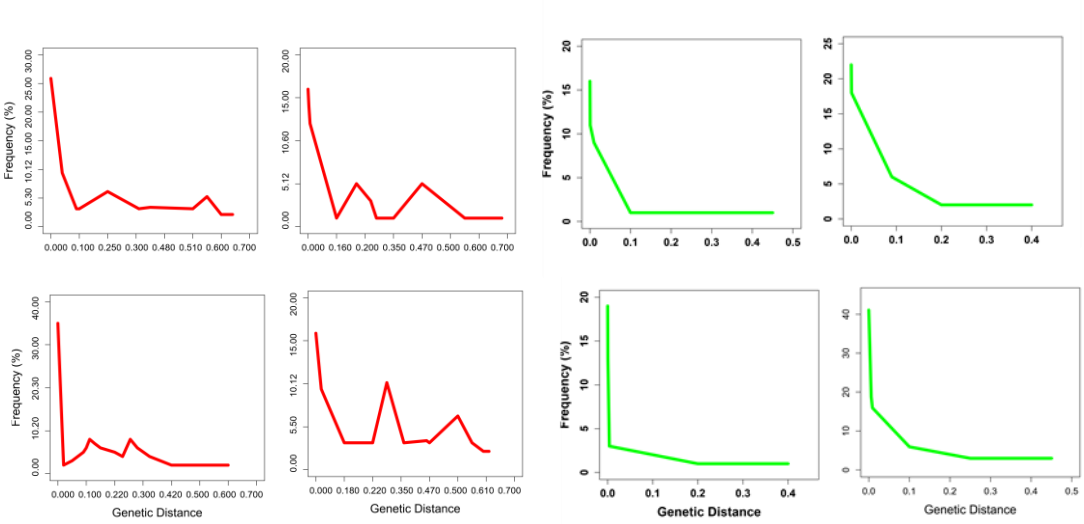



**Figure 3.** Frequency Distribution of Genetic Distance of NUMTs to the Mitogenome in Purebred Amur Tigers, Purebred South China Tigers, and Their Hybrids

The figure illustrates the frequency distribution of genetic distances of NUMTs to the mitogenome for pure and hybrid individuals of two tiger populations: (A) Amur tigers and (B) South China tigers. In purebred individuals, NUMTs exhibit a declining trend in genetic distance (green lines). In hybrid individuals, distinct signal peaks are observed (red lines), representing admixture patterns in the NUMT pool.

### Ability for paternal subspecies identification using NUMTs

The genetic distance results and their 0.95 confidence intervals (CIs) provide valuable insights into the maternal inheritance patterns and genetic diversity of the Amur tiger and South China tiger. For the Amur tiger, the CI for genetic distance to the maternal mitogenome (Inter-d) ranged from [0.35, 0.51], while for the South China tiger, it was slightly higher at [0.39, 0.54], indicating greater divergence in NUMTs among hybrids in the South China tiger. In terms of genetic distance to the mitogenome itself (Intra-d), the Amur tiger exhibited a narrower CI of [0.18, 0.22], compared to the South China tiger's broader range of [0.18, 0.25], suggesting slightly higher variability in the latter's mitogenome. These results highlight distinct patterns of NUMT integration and maternal lineage divergence between the two subspecies.

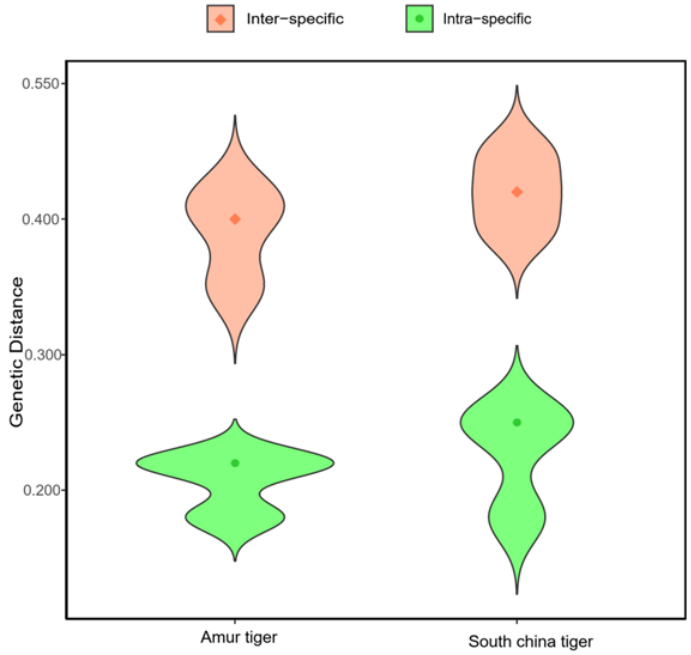



**Figure 3.** The species resolution of NUMTs in the Amur and South China tigers revealed differences in their genetic distances to the consubspecific mitogenome (intra-d) compared to the genetic distances to the mitogenome of hybridizable subspecies (inter-d).

### Comparison of NUMT mutation rates in hybrid and purebred tigers

The comparison of mutation rates in NUMTs between hybrid and purebred individuals of both Amur and South China tigers revealed significant differences, as shown in the boxplot (Figure 3-3). Hybrid individuals exhibited higher mutation rates compared to their purebred counterparts.

Specifically, hybrids displayed a broader range of mutation rates, with some showing higher mutation frequencies than those observed in purebred tigers. The mutation rates in purebred Amur and South China tigers were relatively stable, with similar median values between these groups. The increased variation in mutation rates observed in hybrid individuals suggests that hybridization may contribute to greater genomic instability, which could be associated with the interaction between mitochondrial and nuclear genomes. This variability in NUMTs may reflect the impact of interspecific gene flow on the integrity of mitochondrial DNA fragments integrated into the nuclear genome. The differences in mutation rates between hybrids and purebred tigers provide valuable insights into the potential evolutionary consequences of hybridization, particularly concerning the stability and evolution of NUMTs in these endangered species

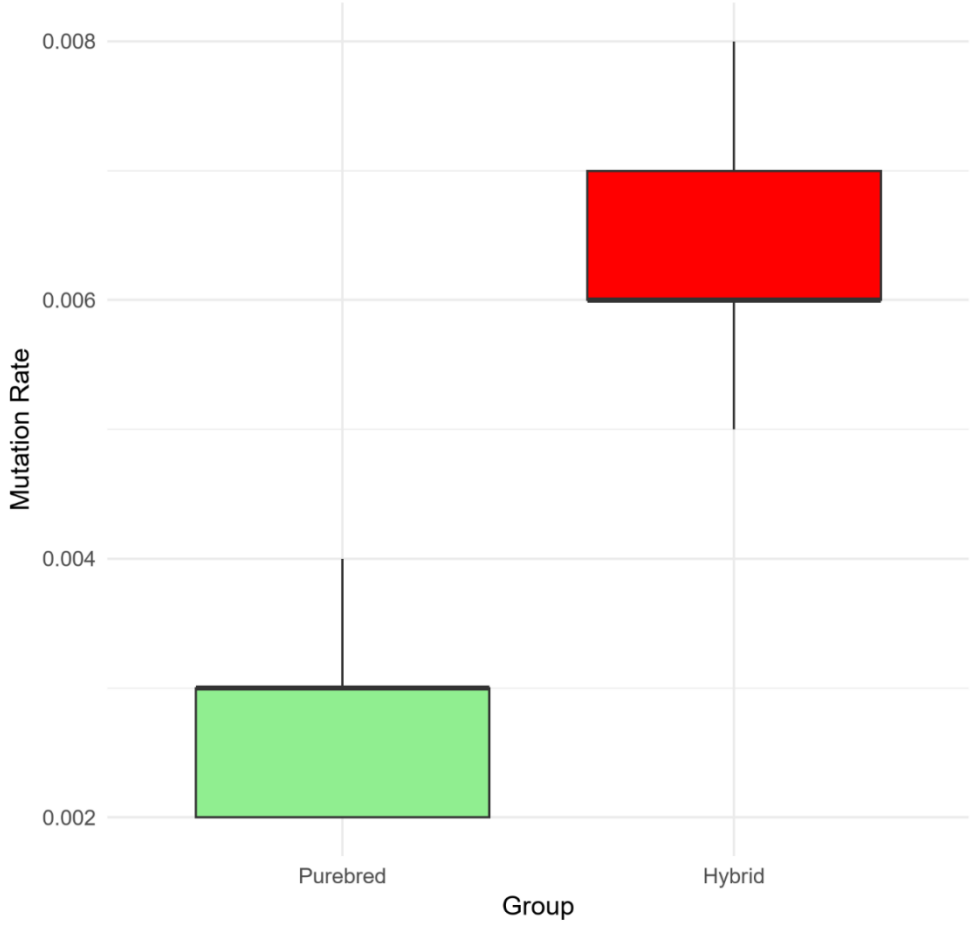



**Figure 3.** Mutation Rates of NUMTs in Purebred vs. Hybrid Tigers

### Mutation rates, hybrid index correlation of NUMTs in tigers

Mutation rates in NUMTs were significantly different across the three groups (hybrids, pure Amur tigers, and pure South China tigers), as revealed by the Kruskal-Wallis test ($\chi^2 = 12.45$, $p = 0.002$). Pairwise comparisons using the Mann-Whitney U test demonstrated that hybrids exhibited significantly higher mutation rates compared to both pure Amur tigers ($p = 0.008$) and pure South China tigers ($p = 0.005$). No significant difference was observed between the mutation rates of pure Amur tigers and pure South China tigers ($p = 0.35$). Spearman's correlation analysis indicated a strong positive relationship between the hybrid index (percentage of Amur tiger ancestry) and NUMT mutation rates ($\rho = 0.82$, $p = 0.001$). This suggests that NUMT mutation rates increase as hybridization levels rise.

## Discussion

Maintaining genetic integrity is a cornerstone of conservation breeding programs, particularly for critically endangered taxa such as tigers. The South China tiger (*Panthera tigris amoyensis*) and the Amur tiger (*Panthera tigris altaica*) represent two genetically and geographically distinct subspecies, each facing unique conservation challenges. While captive breeding is essential for their survival, it carries a significant risk of gene pool contamination through unintended hybridization. Our study addresses this issue by evaluating the utility of NUMTs (nuclear mitochondrial DNA segments) in identifying hybrid individuals and assessing gene purity using the Auto-Ref HybrID approach.NUMTs have historically been regarded as genomic artifacts(Richly & Leister, 2004)Our findings support this emerging view, showing that NUMT count and length vary significantly among purebred Amur and South China tigers and are markedly elevated in hybrids. These results align with prior studies showing that hybridization can affect mitochondrial-nuclear genome interactions, leading to increased NUMT integration and mutation rates(Calabrese et al., 2012)Specifically, we observed significantly higher NUMT counts in Amur tigers than in South China tigers, with hybrids showing even greater counts. This pattern suggests the accumulation of NUMTs from both parental genomes, consistent with reports in other hybrid systems, such as in canids (Pollinger et al., 2011)and felids(Li et al., 2016)where interspecific gene flow led to mosaic mitochondrial-nuclear patterns. The NUMT length differences we observed—longer fragments in hybrids—may reflect reduced constraints on NUMT integration during

genomic incompatibility, as previously suggested in hybrid fish and primate models (Skowronek et al., 2012).

Importantly, Auto-Ref HybrID was sensitive enough to detect genetic structure within the Amur tiger population, revealing NUMT profile differences among individuals from distinct captive lineages. This finding parallels studies in other endangered mammals where captive lineages have diverged genetically due to geographic separation or founder effects(Frankham et al., 2002)In this context, NUMTs offer a valuable resolution to detect subtle genetic divergence that conventional markers may overlook.Our genetic distance analysis revealed that South China tiger NUMTs have greater divergence from the mitogenome compared to Amur tigers. This may reflect the historical bottlenecks and prolonged isolation of the South China tiger, whose population has not been refreshed with wild individuals for over half a century. Similar patterns have been observed in captive populations of the Przewalski's horse and cheetahs, where mitochondrial-nuclear discordance increases over time due to small effective population sizes and inbreeding (Dobrynin et al., 2015).

Furthermore, we detected significantly higher NUMT mutation rates in hybrids, suggesting increased genomic instability. This observation is consistent with the "hybrid breakdown" model, where incompatibilities between parental genomes disrupt regulatory and repair pathways(Orr & Turelli, 2001). It is plausible that such disruptions affect mitochondrial-nuclear communication, facilitating the insertion and mutation of NUMTs. This phenomenon has been proposed as a postzygotic barrier in many taxa, reinforcing the need to avoid hybridization in conservation contexts(Turelli & Moyle, 2007)Our study is the first to propose a reference-free, NUMT-based method for simultaneously identifying hybrids and assessing gene pool contamination in tigers. Compared to traditional markers like microsatellites or SNP panels(Luo et al., 2004)NUMTs offer unique advantages: they capture both mitochondrial lineage history and nuclear genomic integration, offering insights into past hybridization events and recent genomic instability.

The conservation implications are profound. As ex-situ populations of both subspecies are considered for future reintroduction, understanding the genetic purity of candidates is essential. Our findings highlight the risk of overlooking hidden hybridization and underscore the need for genome-wide tools like Auto-Ref HybrID in pre-release screening protocols.

## Acknowledgements

We want to express our deep gratitude to Dr. Boyang Liu and the staff of the National Forestry and Grassland Administration Engineering & Technology Research Centre for the Conservation and Utilisation of Wildlife, China, for their support in the experiments, bioinformatics analysis, and other aspects of the research. We also thank all the people and institutions who provided samples and helped with the sampling. In addition, we gratefully acknowledge the support from the CSC grant from the Chinese government. Finally, we thank the anonymous reviewers for their valuable comments and recommendations.

## References

Andrews, S., Krueger, F., Segonds-Pichon, A., Biggins, L., Krueger, C., & Wingett, S. (2010). FastQC. *A quality control tool for high throughput sequence data*, *370*.

Arumugam, K. A., & Annavi, G. (2019). Captive breeding of threatened mammals native to Southeast Asia–a review on their ex-situ management, implications, and reintroduction guidelines. *Annual Research & Review in Biology*, *30*, 1-16.

Asa, C., Traylor‐Holzer, K., & Lacy, R. (2011). Can conservation‐breeding programmes be improved by incorporating mate choice? *International Zoo Yearbook*, *45*(1), 203-212.

Ballou, J. D. (1992). Genetic and demographic considerations in endangered species captive breeding and reintroduction programs. In *Wildlife 2001: populations* (pp. 262-275). Springer.

Bazinet, A., Ondov, B., Sommer, D., & Ratnayake, S. BLAST-based validation of metagenomic sequence assignments. PeerJ. 2018; 6: e4892. In.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *bioinformatics*, *30*(15), 2114-2120.

Calabrese, F. M., Simone, D., & Attimonelli, M. (2012). Primates and mouse NumtS in the UCSC Genome Browser. *BMC bioinformatics*, *13*, 1-9.

Dinerstein, E., Loucks, C., Wikramanayake, E., Ginsberg, J., Sanderson, E., Seidensticker, J.,…Klenzendorf, S. (2007). The fate of wild tigers. *BioScience*, *57*(6), 508-514.

Dobrynin, P., Liu, S., Tamazian, G., Xiong, Z., Yurchenko, A. A., Krasheninnikova, K.,…Johnson, W. (2015). Genomic legacy of the African cheetah, Acinonyx jubatus. *Genome biology*, *16*, 1-20.

Earnhardt, J. M. (1999). Reintroduction programmes: genetic trade‐offs for populations. *Animal Conservation*, *2*(4), 279-286.

Frankham, R., Briscoe, D. A., & Ballou, J. D. (2002). *Introduction to conservation genetics*. Cambridge university press.

Goodrich, J., Wibisono, H., Miquelle, D., Lynam, A., Sanderson, E., Chapman, S.,…Harihar, A. (2022). Panthera tigris. The IUCN red list of threatened species 2022: E. T15955A214862019. In.

Guoliang, C., Zhongkui, L., Peilin, S., Kun, J., & Qingyong, S. (2001). The current situation and the pedigree analysis of South China tiger. *Chinese Journal of Zoology*, *36*(4), 45-48.

Harihar, A., Ghosh-Harihar, M., & MacMillan, D. C. (2014). Human resettlement and tiger conservation–socio-economic assessment of pastoralists reveals a rare conservation opportunity in a human-dominated landscape. *Biological conservation*, *169*, 167-175.

Hebblewhite, M., Miquelle, D. G., Murzin, A. A., Aramilev, V. V., & Pikunov, D. G. (2011). Predicting potential habitat and population size for reintroduction of the Far Eastern leopards in the Russian Far East. *Biological conservation*, *144*(10), 2403-2413.

Iucn, S. (2013). Guidelines for reintroductions and other conservation translocations. *Gland Switz Camb UK IUCNSSC Re-Introd Spec Group*, *57*.

Jackiw, R. N., Mandil, G., & Hager, H. A. (2015). A framework to guide the conservation of species hybrids based on ethical and ecological considerations. *Conservation biology*, *29*(4), 1040-1051.

Li, G., Davis, B. W., Eizirik, E., & Murphy, W. J. (2016). Phylogenomic evidence for ancient hybridization in the genomes of living cats (Felidae). *Genome research*, *26*(1), 1-11.

Luo, S.-J., Kim, J.-H., Johnson, W. E., Walt, J. v. d., Martenson, J., Yuhki, N.,…Quigley, H. B. (2004). Phylogeography and genetic ancestry of tigers (Panthera tigris). *PLoS biology*, *2*(12), e442.

Luo, S.-J., Liu, Y.-C., & Xu, X. (2019). Tigers of the world: Genomics and conservation. *Annual review of animal biosciences*, *7*(1), 521-548.

Maguire, L. A., & Lacy, R. C. (1990). Allocating scarce resources for conservation of endangered subspecies: Partitioning zoo space for tigers. *Conservation biology*, *4*(2), 157-166.

Mazak, V. (1981). Panthera tigris. *Mammalian species*(152), 1-8.

Miquelle, D. G., Stephens, P. A., Smirnov, E. N., Goodrich, J. M., Zaumyslova, O. Y., & Myslenkov, A. E. (2005). Tigers and wolves in the Russian Far East: competitive exclusion, functional redundancy and conservation implications. *Large carnivores and the conservation of biodiversity*, 179-207.

Ning, Y., Liu, D., Gu, J., Zhang, Y., Roberts, N. J., Guskov, V. Y.,…Qi, J. (2024). The genetic status and rescue measure for a geographically isolated population of Amur tigers. *Scientific reports*, *14*(1), 8088.

Orr, H. A., & Turelli, M. (2001). The evolution of postzygotic isolation: accumulating Dobzhansky‐Muller incompatibilities. *Evolution*, *55*(6), 1085-1094.

Pollinger, J. P., Earl, D. A., Knowles, J. C., Boyko, A. R., Parker, H., Geffen, E.,…Sidorovich, V. (2011). A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. *Genome research*, *21*(8), 1294-1305.

Richly, E., & Leister, D. (2004). NUMTs in sequenced eukaryotic genomes. *Molecular Biology and Evolution*, *21*(6), 1081-1084.

Russello, M. A., Gladyshev, E., Miquelle, D., & Caccone, A. (2004). Potential genetic consequences of a recent bottleneck in the Amur tiger of. *Conservation Genetics*, *5*, 707-713.

SJ, L. (2004). Phylogeography and genetic ancestry of tigers (Panthera tigris). *PLoS Bio*, *2*, 2275-2293.

Skowronek, K., Boniecki, M. J., Kluge, B., & Bujnicki, J. M. (2012). Rational engineering of sequence specificity in R. MwoI restriction endonuclease. *Nucleic acids research*, *40*(17), 8579-8592.

Snyder, N. F., Derrickson, S. R., Beissinger, S. R., Wiley, J. W., Smith, T. B., Toone, W. D., & Miller, B. (1996). Limitations of captive breeding in endangered species recovery. *Conservation biology*, *10*(2), 338-348.

Tilson, R., Defu, H., Muntifering, J., & Nyhus, P. J. (2004). Dramatic decline of wild South China tigers Panthera tigris amoyensis: field survey of priority tiger reserves. *Oryx*, *38*(1), 40-47.

Tilson, R., & Nyhus, P. J. (2009). *Tigers of the world: the science, politics and conservation of Panthera tigris*. Academic Press.

Tilson, R., Traylor‐Holzer, K., & Jiang, Q. M. (1997). The decline and impending extinction of the South China tiger. *Oryx*, *31*(4), 243-252.

Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hübner, S.,…Bock, D. G. (2016). Hybridization and extinction. *Evolutionary applications*, *9*(7), 892-908.

Toolkit, P. (2019). Broad institute. *(No Title)*.

Traylor-Holzer, K., Zhong, X., & Yuzhong, Y. (2010). The struggle to save the last South China tigers. In *Tigers of the World* (pp. 457-461). Elsevier.

Turelli, M., & Moyle, L. C. (2007). Asymmetric postmating isolation: Darwin's corollary to Haldane's rule. *Genetics*, *176*(2), 1059-1088.

Veasey, J. S. (2020). Can zoos ever be big enough for large wild animals? A review using an expert panel assessment of the psychological priorities of the amur tiger (Panthera tigris altaica) as a model species. *Animals*, *10*(9), 1536.

Wang, C., Wu, D.-D., Yuan, Y.-H., Yao, M.-C., Han, J.-L., Wu, Y.-J.,…Huang, M. (2023). Population genomic analysis provides evidence of the past success and future potential of South China tiger captive conservation. *BMC biology*, *21*(1), 64.

Wang, T., Feng, L., Mou, P., Wu, J., Smith, J. L., Xiao, W.,…Cheng, Y. (2016). Amur tigers and leopards returning to China: direct evidence and a landscape conservation plan. *Landscape ecology*, *31*, 491-503.

Yachmennikova, A., Zhu, S., Kotlov, I., Sandlersky, R., Yi, Q., & Rozhnov, V. (2022). Is the Lesser Khingan suitable for the Amur tiger restoration? Perspectives with the current state of the habitat and prey base. *Animals*, *13*(1), 155.

Zhang, L., Lan, T., Lin, C., Fu, W., Yuan, Y., Lin, K.,⋯Chen, D. (2023). Chromosome‐scale genomes reveal genomic consequences of inbreeding in the South China tiger: A comparative study with the Amur tiger. *Molecular Ecology Resources*, *23*(2), 330-347.

Zhang, W., Xu, X., Yue, B., Hou, R., Xie, J., Zou, Z.-T.,…Xie, Z. (2019). Sorting out the genetic background of the last surviving South China tigers. *Journal of Heredity*, *110*(6), 641-650.