javascript:void(0) Scat DNA reveals distinct lineage of Chinese pangolin in Nepal: Implications for Conservation and Wildlife Forensics.

Fiona Hogan[1], Faye Wedrowicz[1], Ambika Pd. Khatiwada[2], Janardan Dev Joshi[3], Samuel Wasser[4], and Wendy Wright[1]

[1]Federation University Australia
[2]University of Oviedo
[3]National Trust for Nature Conservation
[4]University of Washington

September 04, 2024

Abstract

Pangolins are under severe threat from illegal trade and habitat loss. Despite their conservation significance, these elusive creatures remain poorly understood, particularly in the context of phylogeography and genetic diversity. In this study DNA was isolated from two types of pangolin scat (faeces) samples (whole scat and swabs) which had been stored frozen for up to 3 years between the time of collection and our analyses. A mitochondrial sequence for the cytochrome b gene (424 bp) region was reliably generated from both sample types, for scats which appeared in 'good' or 'poor' condition at the time of collection. Sanger sequencing revealed four new cytochrome b haplotypes for pangolin, with variations in distinct sampling regions in Nepal (central and east). Comparison to reference sequences for pangolin originating from China, Taiwan and Thailand, indicated a distinct variant of Chinese pangolin in Nepal (15-19 base pair difference). Genetic differentiation of Chinese pangolin in Nepal has broader conservation significance for this species. Methods described here are robust and could be applied to broader genetic studies of pangolin populations across Nepal, allowing for genetic mapping of pangolin variations to assist in identifying populations of significance, evaluation of conservation interventions, and forensic applications supporting the fight against illegal poaching.

javascript:void(0)**Scat DNA reveals distinct lineage of Chinese pangolin in Nepal: Implications for Conservation and Wildlife Forensics.**

Authors

Fiona Hogan[A], Faye Wedrowicz[A] , Ambika Pd. Khatiwada[CDE], Janardan Dev Joshi[D], Sam Wasser[B] and Wendy Wright[A]

[A] Future Regions Research Centre, Federation University Australia, Victoria, Australia

[B] Department of Biology, University of Washington

[C]Biodiversity Research Institute (University of Oviedo, Principado of Asturias, Spanish National Research Council), Mieres Campus, University of Oviedo, 33600 Mieres, Asturias, Spain

javascript:void(0)**[D] National Trust for Nature Conservation, Khumaltar, Lalitpur 44700, Nepal**

[E] IUCN SSC Pangolin Specialist Group, Zoological Society of London, Regent's Park, NW1 4RY, United Kingdom

Corresponding author

Fiona E. Hogan Address: Victoria, Future Regions Research Centre, Federation University Australia, Victoria, Australia Email: fiona.hogan@federation.edu.au

Keywords

Pangolin, non-invasive genetic sampling, scat, mitochondrial DNA

Abstract

Pangolins are under severe threat from illegal trade and habitat loss. Despite their conservation significance, these elusive creatures remain poorly understood, particularly in the context of phylogeography and genetic diversity. In this study DNA was isolated from two types of pangolin scat (faeces) samples (whole scat and swabs) which had been stored frozen for up to 3 years between the time of collection and our analyses. A mitochondrial sequence for the *cytochrome b* gene (424 bp) region was reliably generated from both sample types, for scats which appeared in 'good' or 'poor' condition at the time of collection. Sanger sequencing revealed four new *cytochrome b* haplotypes for pangolin, with variations in distinct sampling regions in Nepal (central and east). Comparison to reference sequences for pangolin originating from China, Taiwan and Thailand, indicated a distinct variant of Chinese pangolin in Nepal (15-19 base pair difference). Genetic differentiation of Chinese pangolin in Nepal has broader conservation significance for this species. Methods described here are robust and could be applied to broader genetic studies of pangolin populations across Nepal, allowing for genetic mapping of pangolin variations to assist in identifying populations of significance, evaluation of conservation interventions, and forensic applications supporting the fight against illegal poaching.

javascript:void(0)

javascript:void(0)**Introduction**

The illegal wildlife trade is a major global conservation issue that threatens the future survival of many species, including all eight pangolin species. Pangolins, or scaly anteaters (Order: Pholidota), are typically solitary, nocturnal mammals, found across Asia and Africa. Pangolins have been described as the world's most trafficked mammal group (Aisher 2016), driven by demand for their scales and meat through the illegal wildlife trade. Combating pangolin poaching is challenging due to ease of capture, the secretive nature of this lucrative trade, associated corruption (Anagnostou and Doberstein 2022) and difficulties in monitoring and protecting pangolins in remote habitats.

Four Asian pangolin species are currently recognised: the Sunda Pangolin (*Manis javanica*), Philippine Pangolin (*Manis culionensis*), Chinese Pangolin (*Manis pentadactyla*) and Indian Pangolin (*Manis crassicaudata*). Genetic data have recently suggested the existence of a fifth Asian pangolin species (tentatively named*Manis mysteria*), however, these data were derived from seized pangolin material, so the geographic origin is unknown (Hu *et al.* 2020b; Gu *et al.* 2023). Two species of pangolin occur in Nepal, the Chinese pangolin and the Indian pangolin (Khatiwada *et al.* 2020). Distribution records suggest that Chinese pangolins are distributed across eastern, central and mid-western Nepal at elevations of up to 2,000 m while Indian pangolins are found in Nepal's western regions, at lower elevations (below 500 m) (Baral and Shah 2008; Jnawali*et al.* 2011; Sharma *et al.* 2020; Figure 1). There may be some overlap between the ranges of the two species, though this has not been confirmed (Khatiwada *et al.* 2020). Nepal, and especially eastern Nepal (Ghimire *et al.* 2020) has long been considered a major hotspot for pangolin poaching and trafficking (Thapa *et al.* 2014; Katuwal *et al.* 2015).

Given the threats faced by pangolins, there is a need to develop efficient survey and monitoring methods to determine species status, identify priority populations and evaluate the impact of conservation interventions on populations. Pangolins are typically rare, cryptic creatures that are difficult to survey and monitor. Willcox *et al.* (2019) provides an overview of survey methods for pangolins, including burrow counts, nocturnal and diurnal surveys, camera trapping, community interviews, telemetry, use of detection dogs and

molecular techniques. The authors discuss the effectiveness of combinations of the above listed methods, noting the potential for molecular techniques to confirm species identification from scats (faecal samples), where, for example, scats are used to confirm pangolin presence. Willcox *et al.* (2019) also lament that there is limited genetic material available from pangolins of known geographic origin. This is especially true of wild pangolin populations and, along with low coverage across their large geographic range, limits capacity to infer the origin of seized pangolin or pangolin material using genetic analyses. Collection of samples of known origin would facilitate the establishment of a reference sample database that could assist with increasing our understanding of pangolin population structure while also facilitating genetic tracing.

Calls for research to generate knowledge regarding pangolins, to inform policy makers and law enforcement agencies, include the need to identify the geographical sources of trafficked pangolins, at national and local scales (Pietersen and Challender 2020). Accurate, georeferenced DNA databases are required to inform an understanding of trafficking dynamics and origins of trafficked individuals, and to guide conservation action and policy. Molecular methods that identify genetic differences between populations in different geographic regions can be used to create large genetic databases that characterise populations across a species' range (Wasser*et al.* 2004; Ogden *et al.* 2009). Such an approach was reported by Wasser *et al.* (2007) for elephants, and by Ghobrial*et al.* (2010) for chimpanzees and can be applied to pangolins. For example, studies of mitochondrial and nuclear DNA by Gaubert*et al.* (2016) revealed six geographic lineages of the African common pangolin (*Manis tricuspis*). This information could help track the global trade of this species at a sub-regional scale. Similarly, using samples from known origin (sampling of wild pangolins), Nash *et al.* (2018) identified three main clusters of Sunda pangolin across Indonesia and subsequently used that information (and DNA from seized pangolin material) to infer possible trade routes.

The establishment of a georeferenced DNA database for pangolins in Nepal relies upon sampling of wild populations, which can be challenging for rare and cryptic species, especially where field locations are remote and difficult to access. Non-invasive genetic sampling of scats is now a commonly used method to obtain DNA from cryptic, threatened species for population genetic investigations and has many advantages over invasive sampling strategies as it allows for wide landscape surveys to be conducted by non-experts e.g., natural resource managers or citizen scientists. Non-invasive genetic sampling can also be used for opportunistic or targeted sampling, and can be far more cost effective and ethical, as DNA can be collected without the need of capturing or even observing a target. Scat collection is therefore a good option for genetic studies of pangolin, which are elusive and highly threatened creatures. If scat can be preserved without affecting the presence or quality of the DNA, sample banks can be built up over time from various locations.

Pangolin scat DNA can be a valuable resource which can be used to improve our understanding and knowledge about these vulnerable species. DNA isolated from scats can be used to generate unique DNA profiles, which allows for individual identification. Genetic data can provide insights into the genetic diversity within populations, whilst genotyping individuals from different geographic locations, can be used to infer differentiation, gene flow between populations and to monitor population trends, all of which are critical for conservation management and prioritising populations for protection. Extensive DNA profiling can allow the characterisation of pangolin populations which can assist forensic investigations allowing the geographic origin and / or source population of confiscated pangolin products to be identified.

Pangolin scats largely consist of grit (˜50%) and undigested insect matter (˜30%) (Karawita *et al.* 2020; Mahmood *et al.* 2021) and contain DNA from numerous sources, including from the animal itself (target DNA in this context), its intestinal microbiome, pathogens, parasites and plant or animal species it has ingested. Target DNA may therefore represent a limited proportion of the total DNA that is obtained from a scat sample and may be present in low amounts. Further to this, Pangolin DNA, isolated from scats is likely to be degraded due to exposure to environmental factors such as sunlight, moisture / precipitation and microbial activity. DNA degradation, low DNA yield and contamination can limit the use of scats for genetic studies. Despite these challenges, careful selection and optimisation of methods for sample collection and genetic analysis can produce reliable genetic data (e.g. Piggott and Taylor 2003; Luikart et al. 2008; Wedrowicz et al. 2013).

In this study, we used DNA isolated from pangolin scats, using both surface swabs and whole scat material to generate genetic sequence suitable for use in establishing a georeferenced DNA database for wild pangolin populations in Nepal. Such data will be especially useful for mapping the geographic extent of the species and its populations, providing a reference against which genetic evidence for investigations and monitoring of crimes such as poaching and illicit wildlife trade can be compared. This pilot study aimed to determine whether pangolin DNA of sufficient quality and quantity to be useful for genetic analyses can be obtained from pangolin scat samples.

Methods

Sample origin and DNA isolation

Both scat and blood samples, previously collected between 2017 and 2020 from two broad geographic regions in Nepal were provided as stored, frozen samples by Nepal's National Trust for Nature Conservation (NTNC). Scat samples originated from east (Taplejung, n =23) and central (Kathmandu, n =4; Chitwan, n =16) Nepal (Fig. 1) and were mostly collected in 2017 as part of another study (Kim 2021). Two additional samples were from an unknown location in Nepal. Scat samples were collected whole, and the surface of the scat swabbed at the time of collection (n =26). Swabs and whole scats (n =37,) were both stored frozen (–20°C) for up to 3 years. Samples were collected from 40 putative individuals and included 23 paired scat and swab samples, 14 scats without a paired swab and 3 swabs without a paired scat. Some of the scat samples had been visually assessed at the time of collection as being in either 'good' or 'poor' condition; intact scats appearing fresh (e.g. signs of moisture) were classed as good quality while those appearing dry, degraded or mouldy were classified as poor quality. NTNC also provided access to stored blood samples (n =5) which had been taken from pangolins rescued within the Kathmandu (n =4) and Chitwan (n =1) areas and stored at –20°C for approximately 1.5 years.

DNA was isolated from blood samples using the DNeasy Blood & Tissue Kit (Qiagen) following the manufacturer's instructions, except DNA was eluted in 50 µL. Swab samples were prepared for DNA isolation by allowing the samples to thaw at room temperature and then vortexing the tubes for approximately 3 minutes, ensuring that material adhering to the swab was released into the liquid. Swabs were removed from each tube followed by centrifugation at 10,000 rpm for 10 minutes. A pipette was used to remove most of the supernatant, leaving the pellet undisturbed. Samples were then vortexed and DNA isolated using the QIAamp DNA Stool Mini Kit (Qiagen). DNA was isolated following the manufacturers protocol except that the sample–buffer ASL mixture was incubated for one hour at 35°C to ensure that a homogenous solution was obtained. DNA was eluted using two separate aliquots of 50 µL and combined.

Frozen whole scat samples were thawed at room temperature. Pieces of scat material were transferred from each whole scat to a 2 mL microcentrifuge tube, which was loosely filled to approximately half of the containers volume (this equated to 200 – 400 mg of scat material). Scat material was then washed by addition of 1.5 mL of phosphate buffered saline (PBS) to each tube, vortexed thoroughly and placed on a tube rotator for 10 minutes. The liquid was then transferred to a new tube and centrifuged at 7,500 rpm for 5 minutes. Most of the supernatant was discarded and DNA was then extracted using the QIAamp DNA Stool Mini Kit (Qiagen) with a one-hour lysis incubation.

javascript:void(0)*DNA quality and sequencing*

A section of the mitochondrial *cytochrome b* (*cytB*) gene was amplified for DNA sequencing using PCR and universal primers L14724 (5'–CGAAGCTTGATATGAAAAACCATCGTTG–3') and H15149 (5'–AAACTGCAGCCCCTCAGAATGATATTTGTCCTCA–3') (Kocher and White 1989; Irwin *et al.* 1991). PCRs were carried out using 10 µM of each primer, 0.1 µg/µL of bovine serum albumin (BSA) and 1 X MyTaq Red Mix (Bioline) in a total volume of 20 µL. Thermocycling conditions began with an initial denaturation cycle of 3 mins at 94°C, followed by 40 cycles of denaturation (93°C, 1 min), annealing (50°C, 1 min) and extension (72°C, 1 min) and finished with a final extension cycle of 5 mins at 72°C. Amplicons were separated on a 2% agarose gel stained with SYBR safe (Invitrogen) and visualised under UV light.

A subset of samples from those that successfully amplified the *cytB* gene were selected for Sanger sequencing (n =37, blood n =5, scat material n =12 and scat swab n =20). PCR products were purified using ExoSAP-IT express reagent (Applied Biosystems) following the manufacturer's instructions. DNA sequencing of PCR products was conducted in both directions and sequences aligned, trimmed and compared to reference sequences (NCBI BLAST) using Geneious Prime (Geneious 2023.2.1). A minimum spanning haplotype network was produced using the PopART program (Leigh and Bryant 2015) alongside reference haplotypes for *cytB* sequence from the Chinese (MT335859), Indian (MG196306) and Malaysian (MG196309) pangolin (Hari *et al.* 2016; Gaubert *et al.* 2018). A phylogenetic tree was also produced to visualise relationships between Chinese pangolin from Nepal and those sampled in other studies/regions. The phylogenetic tree included an additional 14 reference sequences that were retrieved from GenBank to include representatives from all eight species of pangolin and Asian pangolin samples originating from various geographical locations (Supplementary material, Table S1). The 17 reference sequences were aligned with the haplotypes identified in this study, along with *cytB* sequence from the gray wolf (HG998573) as an outgroup, trimmed and used to produce a neighbour joining tree with 1000 bootstrap replicates in Geneious Prime.

Results

Frozen swab and scat samples provide sufficient DNA for mtDNA sequencing

Mitochondrial (*cytB* region) amplification success was high with 100% of blood samples (5/5) and 71% (43/63) of scat samples producing clear bands of the expected size (˜420 bp, Table 1). Amplification success was greater for DNA isolated from scat swabs (81%, 21/26) than from DNA isolated from scat material taken from whole scats (59%, 22/37; Table 1). This is probably because swab samples were first collected from many of the scats which comprised the whole scat samples; and suggests that compared to the scat interior, DNA quality may be greater when sampled from the scat surface.

Scat quality assessments made at the time of sample collection were available for approximately 70% (28/40) of the scat samples used in this study, with 22% (9/40) and 49% (19/40) being regarded as good or poor quality, respectively. For scats that were categorised as good at the time of collection, swab DNA amplified 100% (6/6) of samples, while 63% (5/8) of the whole scat samples resulted in successful amplification of isolated DNA (Table 1). Scats that were classified as poor quality performed worse than good quality scats, however, amplification was still achieved for 78% (11/14) of DNA samples isolated from swabs, and 47% (9/19) samples isolated from scat material (Table 1).

A subset of 37 samples that successfully amplified a segment of the *cytochrome b* gene (blood n =5, scat material n =12 and scat swab n =20) were selected for Sanger sequencing of which 65% (24/37) samples produced readable data. Sequencing success was, as expected, high for DNA isolated from blood, with all five samples producing good sequence data. For scats, 59% (19/32) of samples produced reliable DNA sequence data, where success was similar for both scat material (58%, 7/12) and scat swabs (60%, 12/20; Table 1). Overall, scats classed as being in good condition provided sequence data more reliably (75%, 6/8) than those classed as being in poor condition (11%, 2/18; Table 1). DNA isolated from scat material and sent for sequencing had an associated scat quality assessment for only five samples (one classed as being in good condition and four in poor condition) with all of these samples failing to produce quality sequence data (Table 1). DNA isolated from scat swabs described as being in good condition produced better sequencing results (100%, 6/6) than those described as being in poor condition (20%, 2/10). These data indicate that a visual assessment score of scats could be useful for identifying pangolin scat samples that would have a greater chance of providing DNA of sufficient quality for genetic analysis.

Sequence data identified four unique cytB haplotypes

BLAST comparisons of sequence data indicated that all samples (blood and scat) originated from pangolin and were most similar to the Chinese pangolin. Four new *cytB* haplotypes, cytB-01 (PQ261028), cytB-02 (PQ261029), cytB-03 (PQ261030) and cytB-04 (PQ261031), were detected within the 24 pangolin *cytB* sequences identified in this study (Fig. 2). The most common haplotype detected was cytB-01, with 54% (13/24) individuals sequenced in this study having this haplotype (Fig. 2). Compared to cytB-01, haplotypes

detected in this study had between one and four base pair differences. Other haplotypes were less common and unique to areas in east (Taplejung, cytB-02 and cytB-04) or central Nepal (Kathmandu, cytB-03), however, further sampling would be needed to determine whether these haplotypes may be distributed more broadly.

javascript:void(0)*Chinese pangolins found in Nepal have highly divergent cytB sequence*

Comparison of *cytB* sequence obtained from Chinese pangolin originating from Nepal with other publicly available Chinese pangolin sequences showed that Chinese pangolins from Nepal, group together in their own sub clade (Fig. 3). Across 424 bp of *cytB* sequence, the pangolin samples sequenced in this study had between 15 and 19 nucleotide differences (95.5 – 96.5% similarity) compared to reference Chinese pangolin sequences sampled in China (Hua *et al.* 2020), Taiwan (Sun *et al.* 2021) and Thailand (Gaubert *et al.* 2018). A neighbour joining tree grouped pangolin from Nepal as a sister clade to Chinese pangolin from China (*M. p. aurita*) and Taiwan (*M. p. pentadactyla*) (Fig. 3). This suggests that Chinese pangolin from Nepal represent a genetic lineage distinct from those found in China and Thailand, but the geographic extent across which pangolins with these haplotypes is distributed requires further investigation.

javascript:void(0)*Potential wildlife forensic application*

A BLAST search identified five *cytB* sequences on GenBank that were identical or almost identical (1 bp difference) to those in this study (across 304 bp of overlapping sequence, Fig. A1). These sequences were reported by Xie *et al.* (2021) for Chinese pangolin samples that were obtained from a pharmaceutical company located in China. Such high similarity indicates these samples may have been derived from the Nepal region and demonstrates the potential power of mtDNA sequencing for wildlife forensic applications. This finding, however, must be considered with caution, due to a lack of knowledge regarding the extent of the genetic haplotype variants revealed in this study.

javascript:void(0)**Discussion**

Techniques described in this study are useful to genetically characterise wild populations over large spatial scales, and therefore allow the identify of locations from which seized/traffic individuals have come, supporting antipoaching measures. Current and future genetic technologies will be crucial in the combat against illicit wildlife trade and help protect endangered species, including pangolins (Nash et al. 2018; Hu et al. 2020a; Heighton et al. 2023; Priyambada et al. 2023). In this study we have demonstrated that DNA, isolated from scats, can be used to identify the geographic origins of specimens. Such information could be pivotal in identifying poaching hotspots and may lead to more effective law enforcement efforts in areas where illegal trade is most prevalent. Genetic profiling can also be used to identify individuals which is useful for wildlife forensic investigations as it can be used to link seized wildlife products to specific crimes and individuals (Wasser *et al.* 2018; Wasser *et al.* 2022). The use of genetic identification in anti-poaching efforts can serve as a deterrent, discouraging poachers, whilst also providing an opportunity to increase public awareness and education about the consequences of wildlife trade.

Determining the geographic origin of threatened species is not possible without knowledge of the genetic differences which exist within a species across a landscape. Using non-invasive genetic sampling is a valuable tool as it allows for the collection of genetic material from species without harming or disturbing them. It also provides a mechanism to increase sampling efforts by engaging communities and other non-specialists in the sample collection process. Isolating high-quality DNA from scats can be challenging due to the presence of inhibitors, degradation and low DNA concentration. This study demonstrates that DNA useful for genetic analysis of pangolin can be successfully sourced from scat samples, even where those samples have been previously frozen. Scats appearing in good condition (fresher scats) at the time of collection provided better quality DNA than samples that appeared in poor condition, which is consistent with other genetic studies of mammals that have used DNA isolated from scats (e.g. Piggott 2004; Schultz *et al.* 2018). It is also important to note, however, that viable DNA was also obtained from some scats appearing in poor condition and attempting analysis from such scats may be important where samples are from underrepresented regions. Compared to the initial trial results reported here, further method optimisation may increase DNA quantity

and quality that can be obtained from pangolin scats. For example, this trial used a very small amount of scat starting material. Increasing the amount of scat that is washed may raise the amount of pangolin DNA obtained. Different DNA isolation kits can also have a substantial impact on DNA quantity and quality (e.g. Wedrowicz *et al.* 2019), therefore trialling various methods may be useful to maximise the quantity and quality of DNA that can be obtained from pangolin scats.

The ability to obtain pangolin DNA from their scats has also been demonstrated by Priyambada *et al.* (2023) who used faecal DNA isolates to sequence a portion of the mtDNA *cytB* gene and to genotype samples using 20 microsatellites, highlighting eight microsatellite loci with high amplification success (91 – 100%) and low rates of genotyping error (< 6%). Confidently inferring the location of origin for seized pangolins or pangolin products relies heavily on having access to reference data from individuals sampled from across their distribution. Widespread sampling of live, wild individuals would be very difficult to achieve and also raises ethical concerns. DNA sourced non-invasively from pangolin scats offers the opportunity to be able to produce such a database, by allowing large numbers of samples to be sourced without having to interfere with live animals. A coordinated program for collecting pangolin scat samples across their range for the purpose of building a DNA database would be a significant benefit to pangolin conservation efforts.

Both mtDNA and microsatellites may be useful markers from which to obtain data for building a genetic map of Chinese pangolin across their range. Both are PCR based markers which is important for maximising the chance of obtaining data from low quality DNA sources such as confiscated scales (Hsieh *et al.* 2011). Compared to mtDNA, nuclear DNA markers such as microsatellites or SNPs may provide more fine scale resolution of population structure and hence a better ability to pinpoint the origins of seized pangolins. Microsatellite genotyping using next generation sequencing may be a useful way by which to increase genotyping success from low quality sample sources and decrease costs (De Barba *et al.* 2016). Due to high copy numbers per cell, mtDNA markers are more likely to successfully amplify with increasing degradation and would still indicate geographic regions of origin (although perhaps at a broader scale). Amplicon sequencing of mtDNA markers using high throughput platforms may increase sensitivity and sequencing success while also reducing costs where large numbers of samples are to be processed (Andrews *et al.* 2018).

Evidence for distinct lineages of Chinese pangolins has been reported previously (see Appendix, Fig. A2). Hu *et al.* (2020a) reported two strongly differentiated lineages of Chinese pangolin, with one group representing samples from China (Yunnan, Guangdong, Hunan, Hainan and Taiwan) and Thailand and the second group represented by samples of uncertain origin as they were obtained from seizures on the Sino-Burmese border. Hu *et al.* (2020a) estimated that the two lineages of Chinese pangolin diverged 130,000 years ago at the time of the most recent uplift of Tibetan plateau. Such events resulting in changes in topography and climate are important factors contributing to currently observed distributions of diversity across landscapes (He and Jiang 2014; Hu *et al.* 2020a).

Three subspecies of Chinese pangolin are currently recognised and include *Manis pentadactyla aurita* distributed across the Asian mainland and two island subspecies, *Manis pentadactyla pentadactyla* from Taiwan and *Manis pentadactyla pusilla* from Hainan, China (Sun *et al.* 2021). The amount of divergence between Chinese pangolins sampled in Nepal and those sampled in China was found to be significantly greater than the amount of divergence observed between subspecies from the Chinese mainland and Taiwan (Fig. 3). This was also reported by Hu *et al.* (2020a) who showed that Chinese pangolin, putatively from Myanmar, were substantially more divergent from both island (Taiwan and Hainan) and mainland pangolins (detected in China and Thailand) than mainland pangolins were from island pangolins (based on COI sequences; Hu *et al.* 2020a).

Another study utilising pangolins sampled within Nepal and the cytochrome oxidase I gene (COI) found that Chinese pangolin sampled in Nepal clustered separately from those sampled elsewhere in Asia (Shrestha *et al.* 2020). Chinese pangolins sampled in the Darjeeling district of northwest Bengal, India (less than 100 km east from the Taplejung region where pangolins were sampled for this study) identified four *cytB* haplotypes (Priyambada *et al.* 2023). The Priyambada *et al.* (2023) study utilised a different portion of the *cytB* gene than used in this study, so data were unable to be directly compared, however, the data presented also

appear to suggest that pangolin from the Darjeeling district are also clearly distinct from Chinese pangolins sampled in China and Thailand (Priyambada*et al.* 2023).

Together, these data suggest that at least one distinct lineage of Chinese pangolin is distributed from Nepal in the west across to at least Myanmar in the east (Appendix, Fig. A2). Further work is needed to clarify the distribution of this distinct lineage, whether there are additional distinct groups of Chinese Pangolin and if morphological differences between groups exist suggesting that the detected lineages may represent separate subspecies or species.

Further diversity within the mainland Chinese pangolin may be identified with more thorough sampling across this species range, suggestions of which have been previously reported. Chromosome diploid numbers of both 2n = 38 and 2n = 40 have been reported for the Chinese pangolin located in southern China (Wu *et al.* 2007; Nie *et al.* 2009). Two forms of Chinese pangolin have been reported (both occurring in Yunnan, China) which differ in scale shape and colour and are referred to as dusky and brown Chinese pangolins (Zhang and Shi 1991). Dusky Chinese pangolin have opaque, grey-black scales while the brown Chinese pangolin has transparent, yellowish-brown scales (Zhang and Shi 1991). Dusky and brown Chinese pangolins were found to have very low levels of differentiation using protein electrophoresis (Su *et al.* 1994) and greater levels of differentiation using mtDNA restriction enzyme digest patterns (Zhang and Shi 1991). More recent genetic studies comparing the two forms and reports of their distributions are lacking. Further work is needed to ascertain whether these observations align with one another and the divergent lineages or whether these might represent additional diversity within the Chinese pangolin population.

The value in establishing a pangolin DNA database to support conservation and to assist in detecting and solving wildlife crime is high and significant. The ability to source DNA from scats provides the means to achieve this. A comprehensive DNA database for pangolins in Nepal will allow genetic characterisation of sampled populations, allowing identification of significant populations and documentation of the distribution and amount of diversity currently present within pangolin populations. Such data would provide a valuable resource to allow the origins of seized pangolin or pangolin products to be inferred, which may aid managers in decisions about where best to focus protection efforts. Developing a genetic database for Chinese pangolin across its extensive range is a substantial undertaking that would require investment and collaboration between research groups and governments in the Asian region.

Author Contributions

Fiona Hogan: Conceptualization (equal); writing original draft (equal); writing – review and editing (equal). **Faye Wedrowicz:** Conceptualization (equal); investigation (lead); writing original draft (equal); formal analysis (lead); writing – review and editing (equal).**Ambika Khatiwada** : Project administration (equal); resources; writing – review and editing (supporting). **Janardan Dev Joshi:** Investigation (supporting); writing – review and editing (supporting).**Sam Wasser:** Resources (equal); writing – review and editing (supporting). **Wendy Wright:** Conceptualization (equal); project administration (equal); writing original draft (supporting); writing – review and editing (equal).

Acknowledgements

We acknowledge HJ (full name obviously), Jennifer Hartman and Athena (detection dog) who originally collected the samples that were subsequently stored at NTNC BCC and which we then accessed for this study. Local communities from Taplejung, Makwanpur, Chitwan, Kanchanpur, and Bhaktapur districts of Nepal supported the collection of those original pangolin samples. We particularly acknowledge the support provided by Dr. Chiranjibi Prasad Pokharel from Nepal's National Trust for Nature Conservation (NTNC), Mr. Nabin Khatiwada from Taplejung district and Harkaman Lama, Tirtha Lama and Bishnu Lama from NTNC.

Data Accessibilty New *cytB* sequence data are available on GenBank (accession numbers: PQ261028, PQ261029, PQ261030, PQ261031).

Competing Interests Statement

The authors declare no conflicts of interest.

References

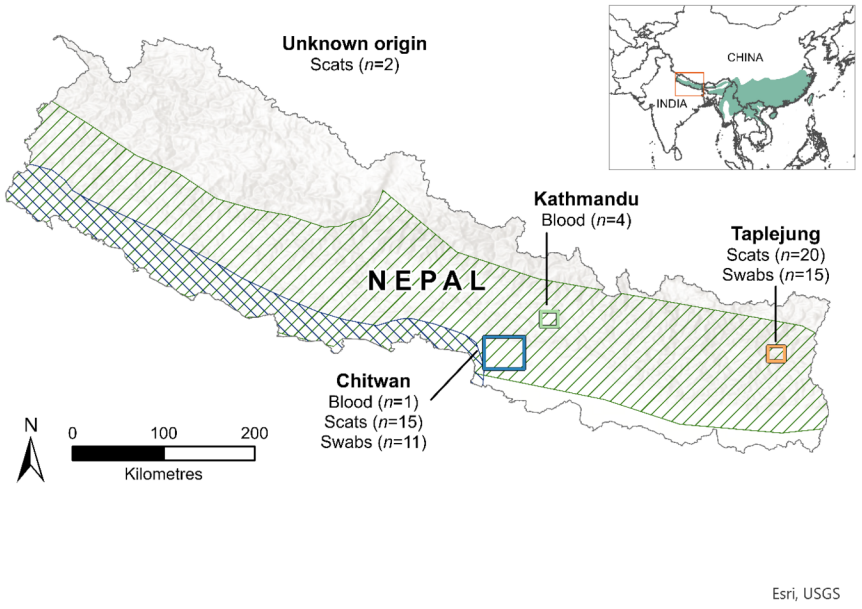
https://www.geneious.com

Aisher, A. (2016). Scarcity, alterity and value: Decline of the pangolin, the world's most trafficked mammal. *Conservation and Society* **14** , 317-329. doi: 10.4103/0972-4923.197610.Anagnostou, M. and Doberstein, B. (2022). Illegal wildlife trade and other organised crime: A scoping review. *Ambio* **51** , 1615-1631. doi: 10.1007/s13280-021-01675-y.Andrews, K. R., De Barba, M., Russello, M. A., and Waits, L. P. (2018). Advances in Using Non-invasive, Archival, and Environmental Samples for Population Genomic Studies. In 'Population Genomics: Wildlife'. (Eds P. A. Hohenlohe and O. P. Rajora) pp. 63-99. (Springer International Publishing: Cham.)Arnason, U., Adegoke, J. A., Bodin, K., Born, E. W., Esa, Y. B., Gullberg, A., Nilsson, M., Short, R. V., Xu, X., and Janke, A. (2002). Mammalian mitogenomic relationships and the root of the eutherian tree. *Proceedings of the National Academy of Sciences* **99** , 8151-8156. doi: 10.1073/pnas.10216429.Baral, H. S. and Shah, K. (2008) 'Wild mammals of Nepal.' (Himalayan Nature.)De Barba, M., Miquel, C., Lobreaux, S., Quenette, P. Y., Swenson, J. E., and Taberlet, P. (2016). High-throughput microsatellite genotyping in ecology: improved accuracy, efficiency, standardization and success with low-quantity and degraded DNA. *Molecular Ecology Resources* **17** , 492-507. doi: 10.1111/1755-0998.12594.du Toit, Z., du Plessis, M., Dalton, D. L., Jansen, R., Paul Grobler, J., and Kotze, A. (2017). Mitochondrial genomes of African pangolins and insights into evolutionary patterns and phylogeny of the family Manidae. *BMC Genomics* **18** , 746. doi: 10.1186/s12864-017-4140-5.du Toit, Z., Grobler, J. P., Kotze, A., Jansen, R., Brettschneider, H., and Dalton, D. L. (2014). The complete mitochondrial genome of Temminck's ground pangolin (*Smutsia temminckii* ; Smuts, 1832) and phylogenetic position of the Pholidota (Weber, 1904). *Gene* **551** , 49-54. doi: 10.1016/j.gene.2014.08.040.Gaubert, P. and Antunes, A. (2015). What's behind these scales? Comments to "The complete mitochondrial genome of Temminck's ground pangolin (*Smutsia temminckii* ; Smuts, 1832) and phylogenetic position of the Pholidota (Weber, 1904)". *Gene* **563** , 106-108. doi: 10.1016/j.gene.2015.03.021.Gaubert, P., Antunes, A., Meng, H., Miao, L., Peigne, S., Justy, F., Njiokou, F., Dufour, S., Danquah, E., Alahakoon, J., *et al.* (2018). The complete phylogeny of pangolins: Scaling up resources for the molecular tracing of the most trafficked mammals on earth. *Journal of Heredity* **109** , 347-359. doi: 10.1093/jhered/esx097.Gaubert, P., Njiokou, F., Ngua, G., Afiademanyo, K., Dufour, S., Malekani, J., Bi, S. G., Tougard, C., Olayemi, A., Danquah, E., *et al.* (2016). Phylogeography of the heavily poached African common pangolin (Pholidota, *Manis tricuspis*) reveals six cryptic lineages as traceable signatures of Pleistocene diversification. *Molecular Ecology* **25** , 5975-5993. doi: 10.1111/mec.13886.Gaubert, P., Njiokou, F., Olayemi, A., Pagani, P., Dufour, S., Danquah, E., Nutsuakor, M. E. K., Ngua, G., Missoup, A.-D., Tedesco, P. A., *et al.* (2015). Bushmeat genetics: setting up a reference framework for the DNA typing of African forest bushmeat. *Molecular Ecology Resources* **15** , 633-651. doi: 10.1111/1755-0998.12334.Geneious 2023.2.1 () Ghimire, P., Raut, N., Khanal, P., Acharya, S., and Upadhaya, S. (2020). Species in peril: assessing the status of the trade in pangolins in Nepal. *Journal of Threatened Taxa* **12** , 15776-15783.Ghobrial, L., Lankester, F., Kiyang, J. A., Akih, A. E., de Vries, S., Fotso, R., Gadsby, E. L., Jenkins, P. D., and Gonder, M. K. (2010). Tracing the origins of rescued chimpanzees reveals widespread chimpanzee hunting in Cameroon. *BMC Ecology* **10** , 2. doi: 10.1186/1472-6785-10-2.Gu, T.-T., Wu, H., Yang, F., Gaubert, P., Heighton, S. P., Fu, Y., Liu, K., Luo, S.-J., Zhang, H.-R., Hu, J.-Y., *et al.* (2023). Genomic analysis reveals a cryptic pangolin species. *Proceedings of the National Academy of Sciences* **120** , e2304096120. doi: 10.1073/pnas.2304096120.Hari, R., Paterson, I. C., and Choo, S. W. (2016). A new complete mitogenome of the critically endangered Chinese pangolin *Manis pentadactyla* .*Conservation Genetics Resources* **8** , 423-426. doi: 10.1007/s12686-016-0586-z.Hassanin, A., Hugot, J.-P., and van Vuuren, B. J. (2015). Comparison of mitochondrial genome sequences of pangolins (Mammalia, Pholidota). *Comptes Rendus Biologies* **338** , 260-265. doi: 10.1016/j.crvi.2015.02.003.He, K. and Jiang, X. (2014). Sky islands of southwest China. I: an overview of phylogeographic patterns. *Chinese Science Bulletin* **59** , 585-597. doi: 10.1007/s11434-013-0089-1.Heighton, S. P., Allio, R., Murienne, J., Salmona, J., Meng, H., Scornavacca, C., Bastos, A. D. S., Njiokou, F., Pietersen, D. W., Tilak, M.-K., *et al.* (2023). Pangolin genomes offer key insights and resources for the world's most trafficked wild mammals. *Molecular*

Biology and Evolution **40** . doi: 10.1093/molbev/msad190.Hsieh, H.-M., Lee, J. C.-I., Wu, J.-H., Chen, C.-A., Chen, Y.-J., Wang, G.-B., Chin, S.-C., Wang, L.-C., Linacre, A., and Tsai, L.-C. (2011). Establishing the pangolin mitochondrial D-loop sequences from the confiscated scales. *Forensic Science International: Genetics* **5** , 303-307. doi: 10.1016/j.fsigen.2010.06.003.Hu, J.-Y., Hao, Z.-Q., Frantz, L., Wu, S.-F., Chen, W., Jiang, Y.-F., Wu, H., Kuang, W.-M., Li, H., Zhang, Y.-P.*, et al.* (2020a). Genomic consequences of population decline in critically endangered pangolins and their demographic histories.*National Science Review* **7** , 798-814. doi: 10.1093/nsr/nwaa031.Hu, J.-Y., Roos, C., Lv, X., Kuang, W., and Yu, L. (2020b). Molecular Genetics Supports a Potential Fifth Asian Pangolin Species (Mammalia, Pholidota, *Manis*). *Zoological Science* **37** , 538-543, 536.Hua, Y., Wang, J., An, F., Xu, J., Zhang, H., and Gu, H. (2020). Phylogenetic relationship of Chinese pangolin (*Manis pentadactyla aurita*) revealed by complete mitochondrial genome. *Mitochondrial DNA Part B* **5** , 2523-2524. doi: 10.1080/23802359.2020.1772693.Irwin, D. M., Kocher, T. D., and Wilson, A. C. (1991). Evolution of the cytochrome b gene of mammals.*Journal of Molecular Evolution* **32** , 128-144. doi: 10.1007/BF02515385.Jnawali, S., Baral, H., Lee, S., Acharya, K., Upadhyay, G., Pandey, M., Shrestha, R., Joshi, D., Lamichhane, B., and Griffiths, J. (2011). The Status of Nepal's Mammals: The National Red List Series-IUCN.Karawita, H., Perera, P., Dayawansa, N., and Dias, S. (2020). Dietary composition and foraging habitats of the Indian Pangolin (*Manis crassicaudata*) in a tropical lowland forest-associated landscape in southwest Sri Lanka. *Global Ecology and Conservation* **21** , e00880. doi: 10.1016/j.gecco.2019.e00880.Katuwal, H. B., Neupane, K. R., Adhikari, D., Sharma, M., and Thapa, S. (2015). Pangolins in eastern Nepal: trade and ethno-medicinal importance.*Journal of Threatened Taxa* **7** , 7563-7567.Khatiwada, A. P., Suwal, T. L., Wright, W., Roe, D., Kaspal, P., Thapa, S., and Paudel, K. (2020). Community conservation in Nepal – opportunities and challenges for pangolin conservation In 'Pangolins: Science, Society and Conservation'. (Eds D. W. Challender, H. C. Nash, and C. Waterman). (Academic Press: London, UK.)Kim, H. J. (2021) Past and Present: What Museum Specimens and Detection Dogs Bring to Pangolin Conservation. (University of Washington, Dept of Biology.)Kocher, T. D. and White, T. J. (1989). Evolutionary Analysis via PCR. In 'PCR Technology: Principles and Applications for DNA Amplification'. (Ed. H. A. Erlich) pp. 137-147. (Palgrave Macmillan UK: London.)Leigh, J. W. and Bryant, D. (2015). POPART: full-feature software for haplotype network construction.*Methods in Ecology and Evolution* **6** , 1110-1116.Luikart, G., Zundel, S., Rioux, D., Miquel, C., Keating, K. A., Hogg, J. T., Steele, B., Foresman, K., and Taberlet, P. (2008). Low genotyping error rates and noninvasive sampling in bighorn sheep. *The Journal of Wildlife Management* **72** , 299-304.Mahmood, A., Li, J., Wu, S., Zhang, F., Xu, N., and Wang, W. (2021). Faecal composition: an observation of food habits of Chinese pangolin (*Manis pentadactyla*) in a subtropical forest associated countryside in South China. *Chiang Mai Journal of Science* **48** , 373-381.Nash, H. C., Wirdateti, Low, G. W., Choo, S. W., Chong, J. L., Semiadi, G., Hari, R., Sulaiman, M. H., Turvey, S. T., Evans, T. A.*, et al.* (2018). Conservation genomics reveals possible illegal trade routes and admixture across pangolin lineages in Southeast Asia. *Conservation Genetics* **19** , 1083-1095. doi: 10.1007/s10592-018-1080-9.Nie, W., Wang, J., Su, W., Wang, Y., and Yang, F. (2009). Chromosomal rearrangements underlying karyotype differences between Chinese pangolin (*Manis pentadactyla*) and Malayan pangolin (*Manis javanica*) revealed by chromosome painting. *Chromosome Research* **17** , 321-329. doi: 10.1007/s10577-009-9027-0.Ogden, R., Dawnay, N., and McEwing, R. (2009). Wildlife DNA forensics—bridging the gap between conservation genetics and law enforcement. *Endangered Species Research* **9** , 179-195.Pietersen, D. W. and Challender, D. W. S. (2020). Research needs for pangolins. In 'Pangolins: Science, Society and Conservation'. (Eds D. W. Challender, H. C. Nash, and C. Waterman). (Academic Press: London, UK.)Piggott, M. P. (2004). Effect of sample age and season of collection on the reliability of microsatellite genotyping of faecal DNA. *Wildlife Research* **31** , 485-493. doi: 10.1071/WR03096.Piggott, M. P. and Taylor, A. C. (2003). Extensive evaluation of faecal preservation and DNA extraction methods in Australian native and introduced species. *Australian Journal of Zoology* **51** , 341-355. doi: 10.1071/ZO03012.Priyambada, P., Singh, V. K., Singh, A., Joshi, B. D., Dalui, S., Wangmo, L. K., Ghosh, A., Thapa, A., Singh, S. K., Chandra, K.*, et al.* (2023). Noninvasive genetic monitoring of Chinese pangolin (*Manis pentadactyla*) from Darjeeling district of Northwest Bengal, India.*Wildlife Letters* **1** , 15-22. doi: 10.1002/wll2.12005.Schultz, A. J., Cristescu, R. H., Littleford-Colquhoun, B. L., Jaccoud, D., and Frere, C. H. (2018). Fresh is best: Accurate SNP genotyping from koala scats. *Ecology and Evolution*

8 , 3139-3151. doi: 10.1002/ece3.3765.Sharma, H. P., Rimal, B., Zhang, M., Sharma, S., Poudyal, L. P., Maharjan, S., Kunwar, R., Kaspal, P., Bhandari, N., Baral, L.*, et al.* (2020). Potential Distribution of the Critically Endangered Chinese Pangolin (*Manis pentadactyla*) in Different Land Covers of Nepal: Implications for Conservation. *Sustainability* **12** , 1282.Shrestha, S., Bashyal, A., Dhakal, A., McGreevy, T. J., Buffum, B., Joshi, J., Chaudhary, H. K., and Khanal, S. N. (2020). Mitochondrial DNA analysis of critically endangered Chinese Pangolins (*Manis pentadactyla*) from Nepal. *Mitochondrial DNA Part B* **5** , 3257-3261. doi: 10.1080/23802359.2020.1811174.Su, B., Liu, R.-Q., Wang, Y.-x., and Shi, L.-m. (1994). Genetic diversity in the Chinese pangolin (*Manis pentadactyla*) inferred from protein electrophoresis. *Biochemical Genetics* **32** , 343-349. doi: 10.1007/BF02426896.Sun, N. C.-M., Huang, C.-C., Tseng, Y.-W., Suwal, T. L., Chi, M.-J., Jang-Liaw, N.-H., and Hung, K.-H. (2021). Complete mitochondrial genome of *Manis pentadactyla pentadactyla* (Mammalia: Pholidota), an endemic subspecies of Chinese pangolin: mitogenome characterisation and phylogenetic implications. *Biodiversity Data Journal* **9** .Thapa, P., Khatiwada, A. P., Nepali, S. C., and Paudel, S. (2014). Distribution and conservation status of Chinese pangolin (*Manis pentadactyla*) in Nangkholyang VDC, Taplejung, Eastern Nepal. *American Journal of Zoological Research* **2** , 16-21.Wasser, S. K., Mailand, C., Booth, R., Mutayoba, B., Kisamo, E., Clark, B., and Stephens, M. (2007). Using DNA to track the origin of the largest ivory seizure since the 1989 trade ban. *Proceedings of the National Academy of Sciences* **104** , 4228-4233. doi: 10.1073/pnas.0609714104.Wasser, S. K., Shedlock, A. M., Comstock, K., Ostrander, E. A., Mutayoba, B., and Stephens, M. (2004). Assigning African elephant DNA to geographic region of origin: Applications to the ivory trade. *Proceedings of the National Academy of Sciences* **101** , 14847-14852. doi: 10.1073/pnas.0403170101.Wasser, S. K., Torkelson, A., Winters, M., Horeaux, Y., Tucker, S., Otiende, M. Y., Sitam, F. A. T., Buckleton, J., and Weir, B. S. (2018). Combating transnational organized crime by linking multiple large ivory seizures to the same dealer. *Science Advances* **4** , eaat0625. doi: 10.1126/sciadv.aat0625.Wasser, S. K., Wolock, C. J., Kuhner, M. K., Brown, J. E., Morris, C., Horwitz, R. J., Wong, A., Fernandez, C. J., Otiende, M. Y., Hoareau, Y.*, et al.* (2022). Elephant genotypes reveal the size and connectivity of transnational ivory traffickers. *Nature Human Behaviour* **6** , 371-382. doi: 10.1038/s41562-021-01267-6.Wedrowicz, F., Karsa, M., Mosse, J., and Hogan, F. E. (2013). Reliable genotyping of the koala (*Phascolarctos cinereus*) using DNA isolated from a single faecal pellet. *Molecular Ecology Resources* **13** , 634-641. doi: 10.1111/1755-0998.12101.Wedrowicz, F., Mosse, J., Wright, W., and Hogan, F. E. (2019). Isolating DNA sourced non-invasively from koala scats: a comparison of four commercial DNA stool kits.*Conservation Genetics Resources* **11** , 219-229. doi: 10.1007/s12686-018-0994-3.Willcox, D., Nash, H. C., Trageser, S., Kim, H. J., Hywood, L., Connelly, E., Ichu Ichu, G., Kambale Nyumu, J., Mousset Moumbolou, C. L., Ingram, D. J.*, et al.* (2019). Evaluating methods for detecting and monitoring pangolin (Pholidata: Manidae) populations. *Global Ecology and Conservation* **17** , e00539. doi: 10.1016/j.gecco.2019.e00539.Wu, S.-H., Chen, M., Chin, S.-C., Lee, D.-J., Wen, P.-Y., Chen, L.-W., Wang, B.-T., and Yu, H.-T. (2007). Cytogenetic analysis of the Formosan pangolin,*Manis pentadactyla pentadactyla* (Mammalia: Pholidota).*Zoological Studies* **46** , 389-396.Xie, X., Ye, H., Cai, X., Li, C., Li, F., Tian, E., and Chao, Z. (2021). DNA mini-barcodes, a potential weapon for conservation and combating illegal trade of pangolin. *Tropical Conservation Science* **14** , 19400829211017361. doi: 10.1177/19400829211017361.Zhang, Y.-P. and Shi, L.-M. (1991). Genetic diversity in the Chinese pangolin (*Manis pentadactyla*): Inferred from restriction enzyme analysis of mitochondrial DNAs. *Biochemical Genetics* **29** , 501-508.

javascript:void(0)**Tables and Figures**



javascript:void(0)**Figure 1** Map of Nepal showing the broad sampling locations for scats (and scat swabs) and the origin of pangolins entering care from which blood samples were taken. Green (forward) diagonal lines indicate the distribution of the Chinese pangolin (*M. pentadactyla*) in Nepal, while blue (reverse) diagonal lines show the distribution of Indian pangolin (*M. crassicaudata*) which overlaps with the southern range of the Chinese pangolin in Nepal. The green shading on the inset map shows the full distribution of the Chinese pangolin across Asia.

Table 1 Summary of PCR and sequencing success for the samples used in this study according to sample type. *Quality data for sequenced scat samples were only available for five samples.

	Scat quality	**Blood**	**Scat**	**Swab**
PCR success	Good		63% (5/8)	100% (6/6)
	Poor		47% (9/19)	78% (11/14)
	All	**100% (5/5)**	**59% (22/37)**	**81% (21/26)**
Sequencing success	Good		0% (0/1)*	85% (6/7)
	Poor		0% (0/4)*	14% (2/14)
	All	**100% (5/5)**	**58% (7/12)**	**60% (12/20)**

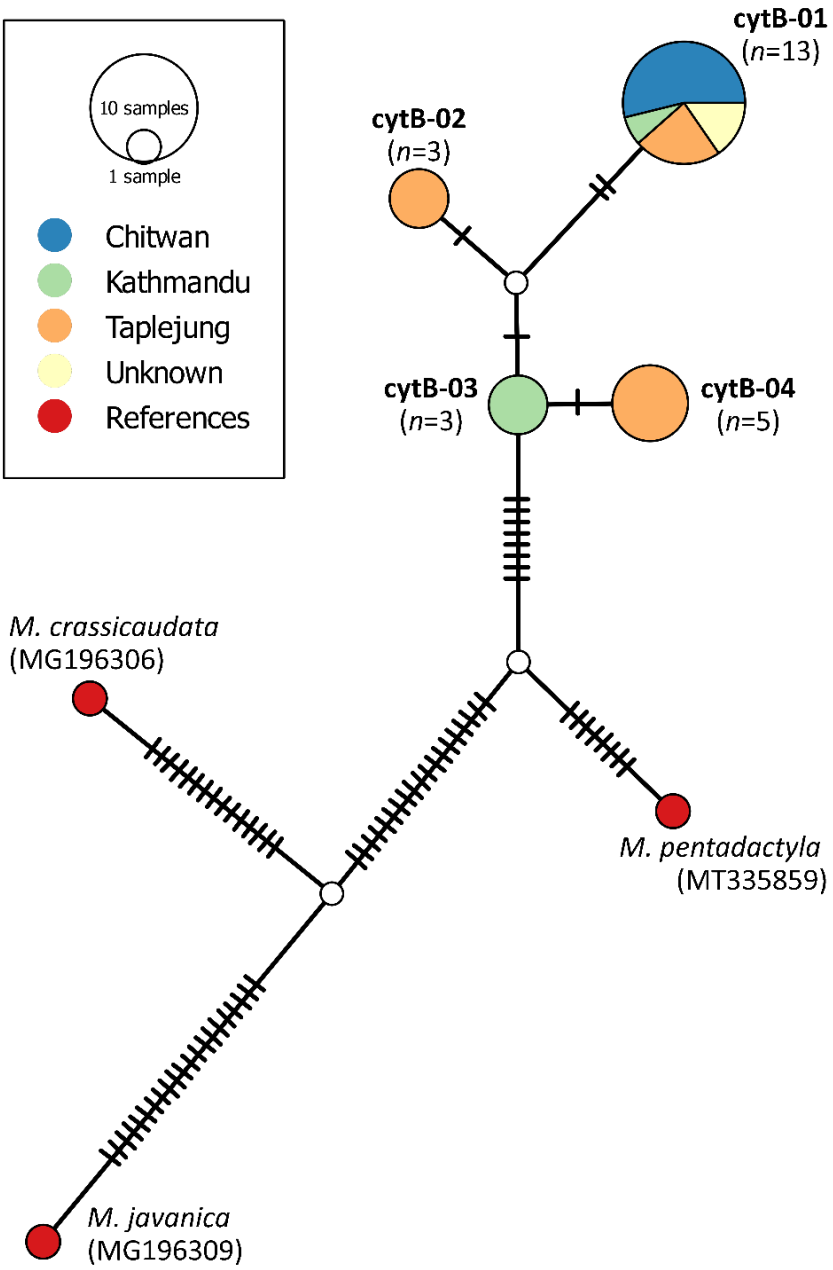


Figure 2 TCS haplotype network showing relationships between the *cytB* haplotypes detected in Nepal (PQ261028, PQ261029, PQ261030, PQ261031) alongside reference haplotypes for the Chinese Pangolin (Guangdong province, China, MT335859), Indian pangolin (Karnataka, India, MG196306) and Malayan pangolin (Guangxi province, China, MG196309).

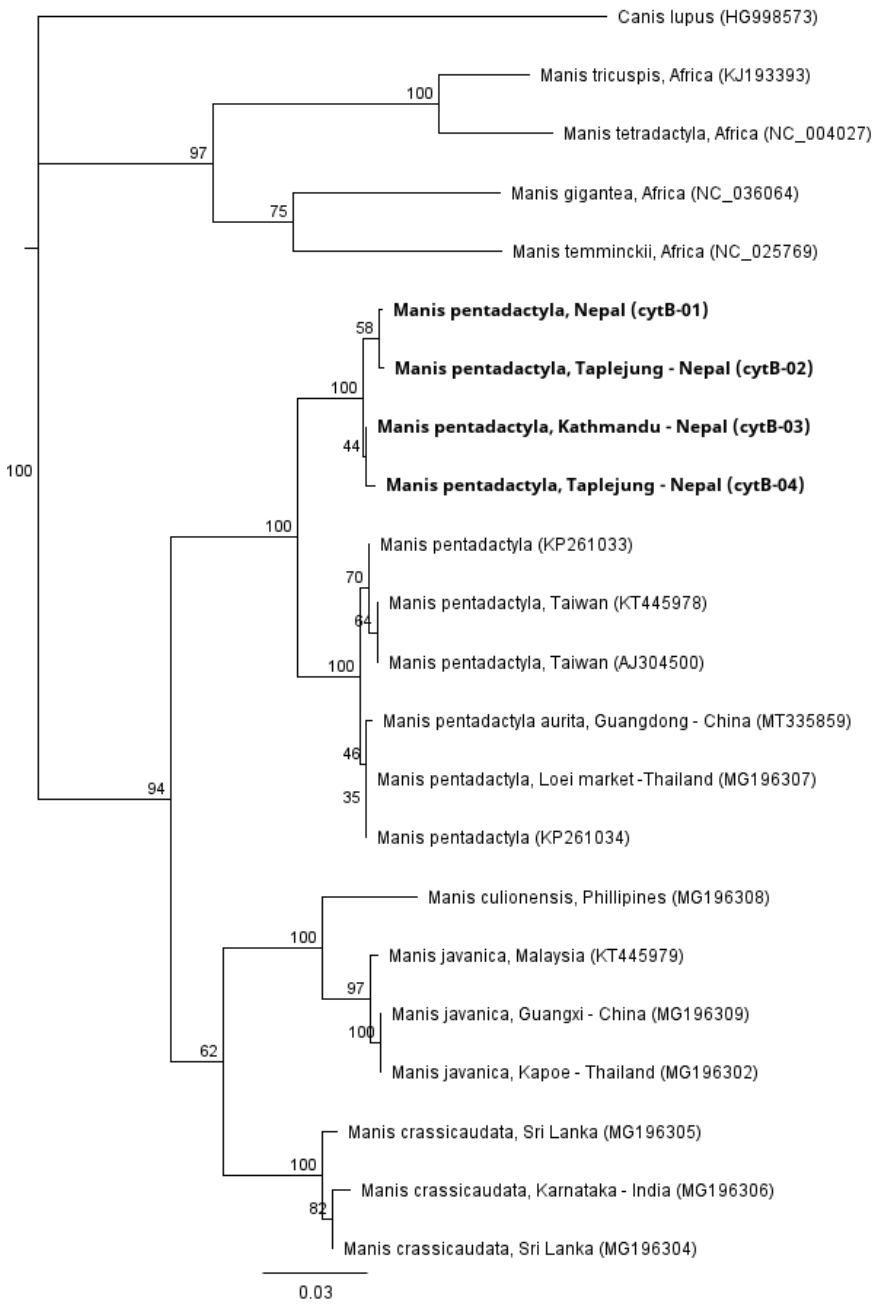


Figure 3 Consensus neighbour joining tree based on 1000 bootstraps illustrating the relationships between the partial Chinese pangolin *cytochrome B* haplotypes (424 bp) identified in this study (shown in bold, cytb01 PQ261028, cytb02 PQ261029, cytb03 PQ261030, cytb04 PQ261031) and reference sequences for the eight pangolin species publicly available on Genbank (Supplementary Table S1). Gray wolf (*Canis lupus*) sequence (HG998573) was used as the outgroup.

javascript:void(0)**Appendix**

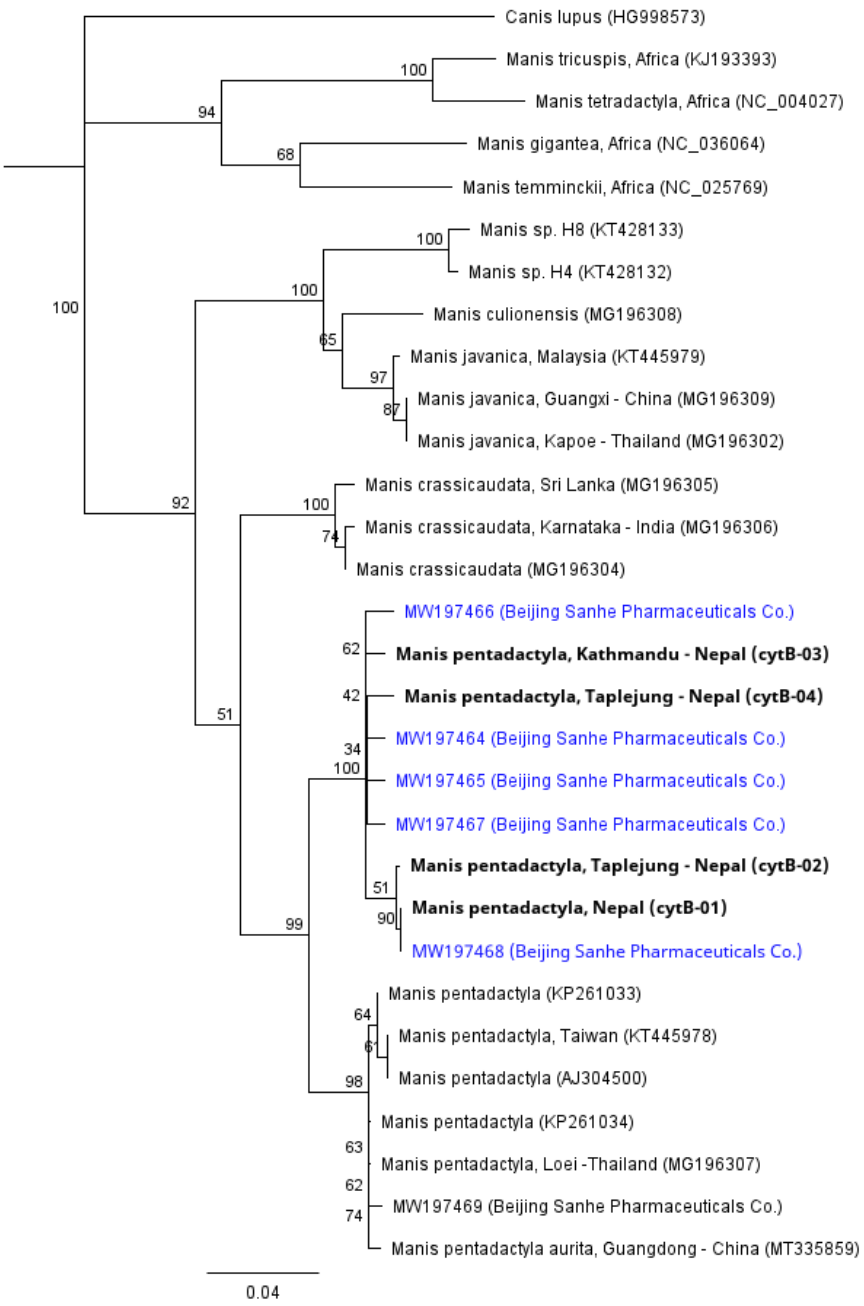


Figure A1 Consensus neighbour joining tree based on 1000 bootstraps illustrating the relationships between the partial Chinese pangolin cytochrome B haplotypes (304 bp) identified in this study (shown in bold) in comparison to closest matches identified on GenBank (shown in blue; Xie *et al.* 2021) and reference sequences. Gray wolf (*Canis lupus*) sequence (HG998573) was used as the outgroup.

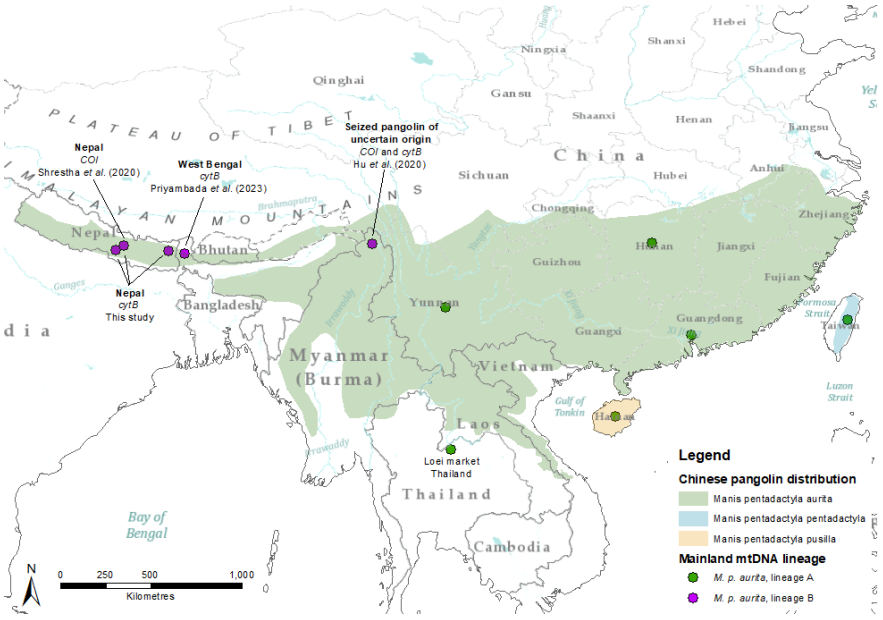


Figure A2 Map illustrating the distribution of the Chinese pangolin and sample sites for studies detecting significantly divergent pangolin lineages. More thorough sampling of Chinese pangolin across their range would be needed to determine whether there are additional lineages and to infer the distributions of the different lineages.

Table A1 Details of the pangolin reference sequences retrieved from NCBI for use in this study.

Accession	Species name	Common name	Origin
KP306514	*Manis tricuspis*	White-bellied pangolin	Africa
KJ193393	*Manis tricuspis*	White-bellied pangolin	Africa
NC_004027	*Manis tetradactyla*	Black-bellied pangolin	Africa
NC_036064	*Manis gigantea*	Giant ground pangolin	Africa
NC_025769	*Manis temminckii*	Temminck's ground pangolin	Africa
MG196304	*Manis crassicaudata*	Indian pangolin	Sri Lanka
MG196305	*Manis crassicaudata*	Indian pangolin	Sri Lanka
MG196306	*Manis crassicaudata*	Indian pangolin	India (Vijayanagara, Ballari, K
MG196308	*Manis culionensis*	Philippine pangolin	Philippines (Iwahig, Palawan Is
MG196302	*Manis javanica*	Malayan pangolin	Thailand (Kapoe)
MG196309	*Manis javanica*	Malayan pangolin	China (Guangxi)
KT445979	*Manis javanica*	Malayan pangolin	Malaysia
KP261033	*Manis pentadactyla*	Chinese pangolin	Southeast Asia
KT445978	*Manis pentadactyla*	Chinese pangolin	Taiwan
AJ304500	*Manis pentadactyla*	Chinese pangolin	Taiwan
MT335859	*Manis pentadactyla*	Chinese pangolin	China (Guangdong)
MG196307	*Manis pentadactyla*	Chinese pangolin	Thailand (Loei market)
KP261034	*Manis pentadactyla*	Chinese pangolin	Southeast Asia
NC_060494	*Manis pentadactyla*	Chinese pangolin	Eastern Taiwan
MW197456	*Manis pentadactyla*	Chinese pangolin	Unknown* (1 bp different to M
MW197457	*Manis pentadactyla*	Chinese pangolin	Unknown* (equivalent to MG1
MW197464 MW197465 MW197467	*Manis pentadactyla*	Chinese pangolin	Unknown* (equivalent to cytB-
MW197466	*Manis pentadactyla*	Chinese pangolin	Unknown* (1 bp diff. to cytB-(

Accession	Species name	Common name	Origin
MW197468	*Manis pentadactyla*	Chinese pangolin	Unknown* (equivalent to cytB-
MW197469	*Manis pentadactyla*	Chinese pangolin	Unknown* (1 bp diff. to MG19
PQ261028	*Manis pentadactyla*	Chinese pangolin	Nepal
PQ261029	*Manis pentadactyla*	Chinese pangolin	Nepal
PQ261030	*Manis pentadactyla*	Chinese pangolin	Nepal
PQ261031	*Manis pentadactyla*	Chinese pangolin	Nepal

* Samples of unknown origin sourced from Beijing Sanhe Pharmaceuticals Co. (Xie *et al.* 2021).

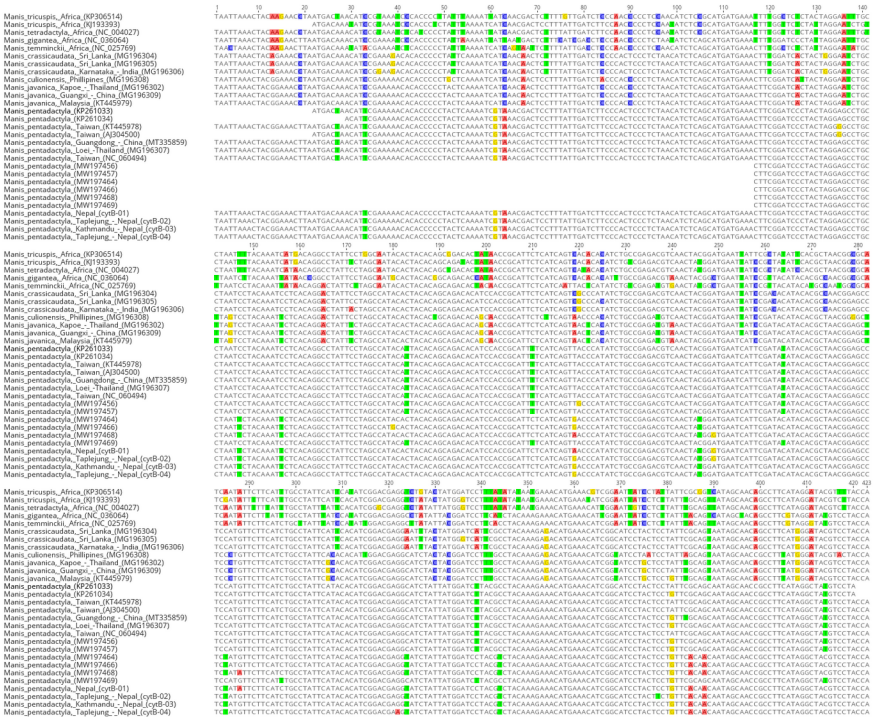


Figure A3 Alignment for the part *cytB* region (423 bp) sequenced in this study along with references obtained from GenBank.