



OPEN

# Genetic tracing of the illegal trade of the white-bellied pangolin (*Phataginus tricuspis*) in western Central Africa

Alain Din Dipita<sup>1,2,✉</sup>, Alain Didier Missoup<sup>1,✉</sup>, Samantha Aguillon<sup>2</sup>, Emilie Lecompte<sup>2</sup>, Brice Roxan Momboua<sup>3,9</sup>, Anne-Lise Chaber<sup>4</sup>, Katharine Abernethy<sup>5,10</sup>, Flobert Njiokou<sup>6</sup>, Maurice Tindo<sup>1</sup>, Stephan Ntie<sup>7</sup> & Philippe Gaubert<sup>2,8,✉</sup>

The white-bellied pangolin is subject to intense trafficking, feeding both local and international trade networks. In order to assess its population genetics and trace its domestic trade, we genotyped 562 pangolins from local to large bushmeat markets in western central Africa. We show that the two lineages described from the study region (WCA and Gab) were overlapping in ranges, with limited introgression in southern Cameroon. There was a lack of genetic differentiation across WCA and a significant signature of isolation-by-distance possibly due to unsuspected dispersal capacities involving a Wahlund effect. We detected a c. 74.1–82.5% decline in the effective population size of WCA during the Middle Holocene. Private allele frequency tracing approach indicated up to 600 km sourcing distance by large urban markets from Cameroon, including Equatorial Guinea. The 20 species-specific microsatellite loci provided individual-level genotyping resolution and should be considered as valuable resources for future forensic applications. Because admixture was detected between lineages, we recommend a multi-locus approach for tracing the pangolin trade. The Yaoundé market was the main hub of the trade in the region, and thus should receive specific monitoring to mitigate pangolins' domestic trafficking. Our study also highlighted the weak implementation of CITES regulations at European borders.

**Keywords** Microsatellites, Demographic history, Molecular tracing, White-bellied pangolin, Wildlife trade, Western Central Africa

The illegal wildlife trade is a flourishing and parallel economy, representing a global annual revenue of US\$7–23 billion<sup>1</sup>. It is fed by uncontrolled hunting activities mostly affecting the tropics, with a significant and deleterious impact on the survival and abundance of terrestrial vertebrates. Pangolins (Mammalia, Pholidota) have been spotlighted as one of the emblems of over-harvesting due to the illegal wildlife trade, being considered the most trafficked mammals in the world<sup>2,3</sup>. With approximately 900,000 individuals seized in the last 20 years<sup>2</sup>, the eight species of African and Asian pangolins are threatened with extinction, through the cumulative effect of illegal hunting and deforestation<sup>4</sup>. As a consequence, they have all been listed on Appendix I of the Convention on

<sup>1</sup>Zoology Unit, Laboratory of Biology and Physiology of Animal Organisms, Faculty of Science, University of Douala, 24157 Douala, Cameroon. <sup>2</sup>Centre de Recherche sur la Biodiversité et l'Environnement (CRBE), Université de Toulouse, CNRS, IRD, Toulouse INP, Université Toulouse 3 – Paul Sabatier (UT3), Toulouse, Bâtiment 4R1, 118 route de Narbonne, 31062 Toulouse Cedex 9, France. <sup>3</sup>Agence Nationale des Parcs Nationaux, Libreville, Gabon. <sup>4</sup>School of Animal and Veterinary Sciences, University of Adelaide, Adelaide SA, Australia. <sup>5</sup>African Forest Ecology Group, Biological and Environmental Sciences, University of Stirling, Stirling, UK. <sup>6</sup>Laboratoire de Parasitologie et Ecologie, Faculté des Sciences, Université de Yaoundé I, Yaoundé 812, Cameroon. <sup>7</sup>Laboratoire de Biologie Moléculaire et Cellulaire (LABMC), Département de Biologie, Université des Sciences et Techniques de Masuku, Franceville, Gabon. <sup>8</sup>Centro Interdisciplinar de Investigação Marinha e Ambiental (CIIMAR), Universidade do Porto, Terminal de Cruzeiros do Porto de Leixões, Av. General Norton de Matos, s/n, 4450-208 Porto, Portugal. <sup>9</sup>Département de Biologie, Faculté des Sciences, Université des Sciences et Techniques de Masuku, Franceville, Gabon. <sup>10</sup>Institut de Recherche en Ecologie Tropicale, Libreville, Gabon. ✉email: adipita1@gmail.com; admissoup@ymail.com; philippe.gaubert@ird.fr

International Trade in Endangered Species of Wild Fauna and Flora (CITES)<sup>5</sup>, and rated as Vulnerable, Endangered or Critically Endangered on the IUCN Red List of Threatened Species<sup>6</sup>.

In Africa, pangolins have long been hunted as part of the traditional bushmeat spectrum used by local communities<sup>7</sup>. However, the apparent decrease in numbers of Asian pangolins<sup>3,5</sup> and the large demand from the Chinese traditional medicine market<sup>5</sup>, have recently set up a global trade network where African pangolins are being massively exported –mostly as scales– to South-East Asia<sup>8,9</sup>. Challender et al.<sup>2</sup> estimated that > 400,000 African pangolins were bound for Asian markets between 2015 and 2019. In Africa, a 42 yr-records survey suggests that harvest volumes of pangolins were multiplied by nine between 2005 and 2014<sup>10</sup>. For central Africa alone, the amount of yearly extirpated pangolins has been estimated to increase from 0.42 to 2.71 M<sup>11</sup>. Despite pangolins having been –erroneously– blamed for their role as potential intermediate hosts of the COVID-19 pandemic<sup>12</sup>, African pangolins continue to be harvested at high rates<sup>13</sup>.

DNA-based approaches have proven useful in tracing the illegal wildlife trade, both in terms of species identification and delineation of source populations<sup>14,15</sup>. Most of the efforts to trace the pangolin trade have so far been focused on species identification, using different types of material, samples and DNA quality (see<sup>16</sup>). The genetic tracing of the pangolin trade has proven efficient in species identification<sup>17</sup> but has remained limited in reaching lineage or population-level assignment, notably because of incomplete DNA registers<sup>18</sup>. Genetic and genomic resources recently produced for the white-bellied pangolin (WBP; *Phataginus tricuspis* (Rafinesque, 1821)), the commonest and most traded species in Africa<sup>19</sup>, have provided a unique opportunity to trace the pangolin trade at both the local and global scales. Using mitochondrial and nuclear DNA sequences, Gaubert et al.<sup>20</sup> delineated six cryptic, divergent and geographically isolated lineages across the species' range, that were later traced by Zhang et al.<sup>9</sup> from large scale seizures made in Hong Kong. Further development of specific microsatellites markers<sup>21</sup> and genomic markers allowed to trace the WBP trade at the scale of a sub-region (in West and central Africa), demonstrating medium-distance trading routes for domestic trade networks<sup>22</sup> and a western-to-central shift in source populations feeding the Asian market<sup>23</sup>.

Cameroon has a long history of bushmeat consumption and trade, and pangolins are commonly offered for sale in bushmeat markets and restaurants<sup>24</sup>, even since the advent of the COVID-19 pandemic<sup>25</sup>. Although pangolins have been nationally protected since 2016, Cameroon remains one of the hotspots of the illegal pangolin trade, exporting a large amount of scales from Africa to Asia<sup>8</sup>, with WBP being by far the main contributing species<sup>9</sup>.

Cameroon is home to the Western Central Africa (WCA) WBP lineage<sup>20</sup> (WCA also extending southward to Equatorial Guinea and northern Gabon), which was the most represented lineage in recent Asian seizures<sup>9</sup>. Given the uncontrolled domestic and international trade affecting the species, tracing the domestic trade of WBP in Cameroon and neighboring countries is a stake with obvious conservation implications. WCA is co-occurring in Gabon with the Gab lineage, which was temporarily considered as endemic to the country pending more extensive geographic sampling<sup>20</sup>. Because (i) admixture between pangolin lineages may jeopardize efforts at tracing the pangolin trade<sup>26</sup> and (ii) both WCA and Gab have been detected in seizures from Hong Kong<sup>9</sup>, we propose to build on recently developed genetic resources to investigate both maternal and bi-parental population genetic signatures in WBP from the subregion. Our specific objectives were to (i) assess the population structure and genetic diversity of WBP from Cameroon, Equatorial Guinea and Gabon, (ii) characterize their demographic history and current effective population size, and (iii) trace the scale of the sub-regional WBP trade.

## Results

### Mitochondrial DNA

The NJ phylogenetic tree based on 641 cyt *b* sequences supported the structuring into six WBP geographic lineages, including Western Africa, Ghana, Dahomey Gap, WCA, Gab and Central Africa (Supplementary Fig. S2 online). A total of 579 individuals from Cameroon, Equatorial Guinea and northern Gabon (Bitam) clustered within WCA, whereas ten samples from southern and northern Gabon and three samples from central-south Cameroon (Yaoundé and Abong Mbang) clustered within Gab (Fig. 1).

A total of 67 and eight haplotypes were identified in WCA and Gab, respectively. WCA and Gab were differentiated by 23 mutations in the median-joining haplotype network, whereas no particular geographic structuring could be observed within the two lineages (Supplementary Fig. S4 online). Three haplotypes were dominant in WCA, with H2 (28%) and H3 (25%) located only in Cameroon, and H5 (17%) in Cameroon and Bioko Island (Supplementary Fig. S5 online). In Gab, the dominant haplotype (H70; 50%) was distributed from northern to southeastern Gabon. Douala markets recruited 8–17 haplotypes (dominated by H5 and H2 at 27% and 23%, respectively), while the Yaoundé market was characterized by 35 haplotypes (dominated by H3 at 32%). Haplotype (*Hd*) and nucleotide ( $\pi$ ) diversity were respectively 0.830 and 0.010 in WCA, and 0.897 and 0.019 in Gab (Supplementary Table S2 online).

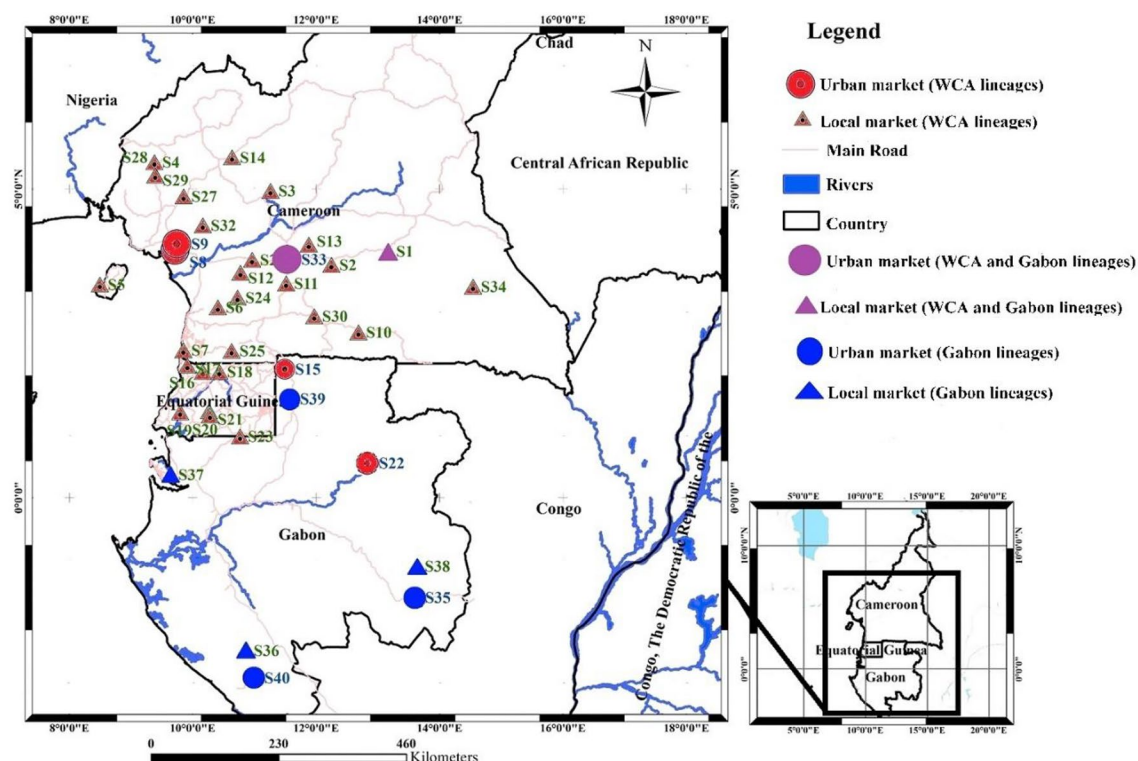
Based on mismatch analysis, the hypotheses of spatial and sudden demographic expansions could not be rejected ( $P(\text{Sim. Rag.} \geq \text{Obs. Rag.}) = 0.559$  and  $P(\text{Sim. Rag.} \geq \text{Obs. Rag.}) = 0.424$  for WCA (see Supplementary Fig. S6 online).

Bayesian skyline plots showed a c. 22-fold increase in the median effective population size of WCA from  $N_{e0} = \text{c. } 1900$  (HPD 95% = c. 750 – 4200) to  $N_{e1} = \text{c. } 41,500$  (HPD 95% = c. 11,000 – 206,000), abruptly starting c. 275 kya (Fig. 2).

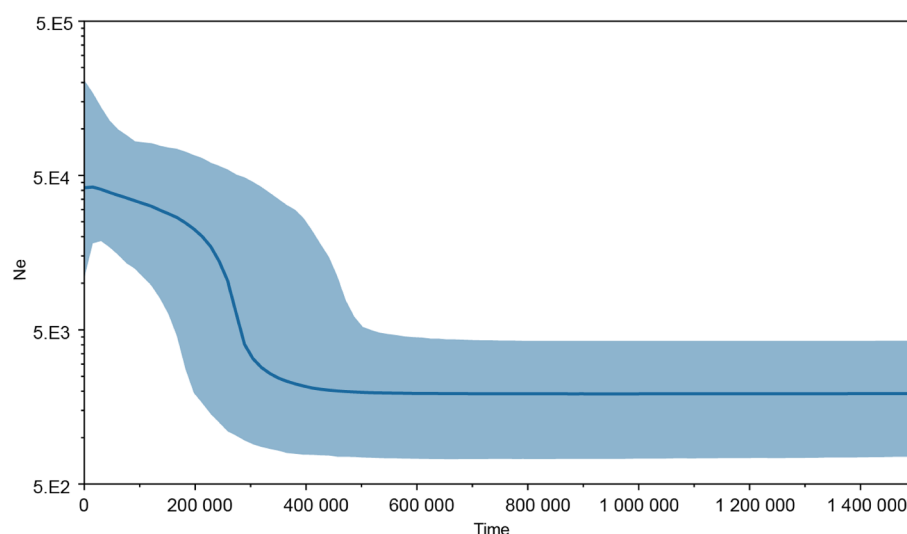
### Microsatellites

#### Genetic diversity

In the reference population that we used to assess locus equilibrium parameters (Foumbot), we detected potential evidence for eight loci affected by null alleles (Supplementary Table S3 online) and significant deviation of



**Figure 1.** Reassessed distribution of the Western Central Africa and Gabon mitochondrial lineages in white-bellied pangolins. Lineage assignment was based on NJ tree clustering (see Results and Supplementary Fig. S2). Red = Western Central Africa lineage; Blue = Gabon lineage; Purple = sites where the two lineages co-occur. S1–Abong Mbang; S2–Akonolinga; S3–Bangong; S4–Bayib Assibong; S5–Bioko; S6–Bipindi; S7–Campo; S8–Douala Central Market; S9–Douala Dakat Market; S10–Djoum; S11–Ekombitié; S12–Eseka; S13–Esse; S14–Foumbot; S15–Bitam; S16–Anguma; S17–Bongoro; S18–Ebenguan; S19–Emangos; S20–Misergue; S21–Taguete; S22–Makokou; S23–Medouneu; S24–Lolodorf; S25–Maan; S26–Mamb; S27–Manengole; S28–Manyemen; S29–Nditam; S30–Sangmelima; S31–Seizure; S32–Yabassi; S33–Yaoundé Nkolndongo Market; S34–Yokadouma; S35–Franceville; S36–Mocabe; S37–Nkoltang; S38–Okoumbi; S39–Oyem; S40–Tchibanga.



**Figure 2.** Bayesian skyline plots showing the demographic history of white-bellied pangolins from the Western Central Africa lineage. Y-axis = Effective population size ( $N_e$ ); X-axis = Time in years before present.

Populations	N	Na	Ho	He	uHe	F <sub>IS</sub>	A <sub>R</sub>	No. private alleles (without urban markets)	No. private alleles (with urban markets)	Private allele frequency (without urban markets)	Private allele frequency (with urban markets)
Abong Mbang	13	6.200	0.571	0.685	0.713	0.207	3.997	5	1	0.320	0.196
Bayomen	5	4.000	0.590	0.593	0.659	<b>0.116</b>	3.645	0	0	0.200	0.140
Eseka	23	7.150	0.548	0.687	0.703	<b>0.217</b>	3.890	4	2	0.217	0.134
Foumbot	36	7.400	0.561	0.657	0.666	<b>0.159</b>	3.664	4	1	0.162	0.097
GES	12	5.850	0.610	0.657	0.685	0.112	3.812	5	2	0.211	0.147
Manengole	19	6.250	0.608	0.671	0.690	<b>0.122</b>	3.786	2	1	0.186	0.113
Mt Cam	6	4.500	0.609	0.628	0.690	0.130	3.850	2	1	0.240	0.139
PNCM	11	5.200	0.580	0.650	0.682	<b>0.149</b>	3.625	1	1	0.127	0.071
Sangmelima	45	8.350	0.587	0.704	0.712	<b>0.170</b>	3.955	11	4	0.244	0.160
Yabassi	9	5.700	0.595	0.647	0.676	0.138	3.632	2	1	0.144	0.073
Gabon	5	4.850	0.575	0.638	0.710	<b>0.210</b>	4.332	15	8	0.932	0.922
Douala markets	84	9.300	0.546	0.687	0.691	-	4.857	-	8	-	0.111
Yaoundé market	248	12.700	0.534	0.710	0.712	-	4.979	-	42	-	0.141
Seizure (Europe)	10	6.000	0.591	0.680	0.716	-	5.284	-	1	-	0.166

**Table 1.** Genetic diversity across reference populations, urban bushmeat markets and seizures of white-bellied pangolins from Cameroon, Equatorial Guinea and Gabon, as estimated from 20 microsatellites loci. Populations are based on geographic proximity and biogeographic continuum among forests, lineage or seizure/selling places (urban markets): Gabon (corresponding to Gab lineage), Bayomen (Bayomen and Bangong [Bafia forest]), Mt Cam (Manyemen and Bayib-Assibong [forest around National Park of Mt Cameroon]), Eseka (Eseka, Mamb, Lolodorf, Bipindi and Eseka [Bassa's forest]), GES (forests of southern Equatorial Guinea), PNCM (Campo, Ma'an [forest of National Park of Campo Ma'an] and northern Equatorial Guinea), Foumbot (Foumbot [Noun/Tikar forests]), Yabassi (Ebo Forest), Manengole (Manengumba Bakossi Forest), Abong Mbang (Lomié [forests of East Cameroon]), and Sangmelima (Sangmelima and Djoum [forests around Dja Biosphere Reserve]). N: number of samples; Na: number of alleles; Ho: observed heterozygosity; He: expected heterozygosity; uHe: unbiased expected heterozygosity; F<sub>IS</sub>: inbreeding coefficient (values in bold are significant); A<sub>R</sub>: allelic richness.

genotypic frequencies from those expected under HWE at four loci (Supplementary Table S4 online). No linkage disequilibrium between pairs of loci could be identified.

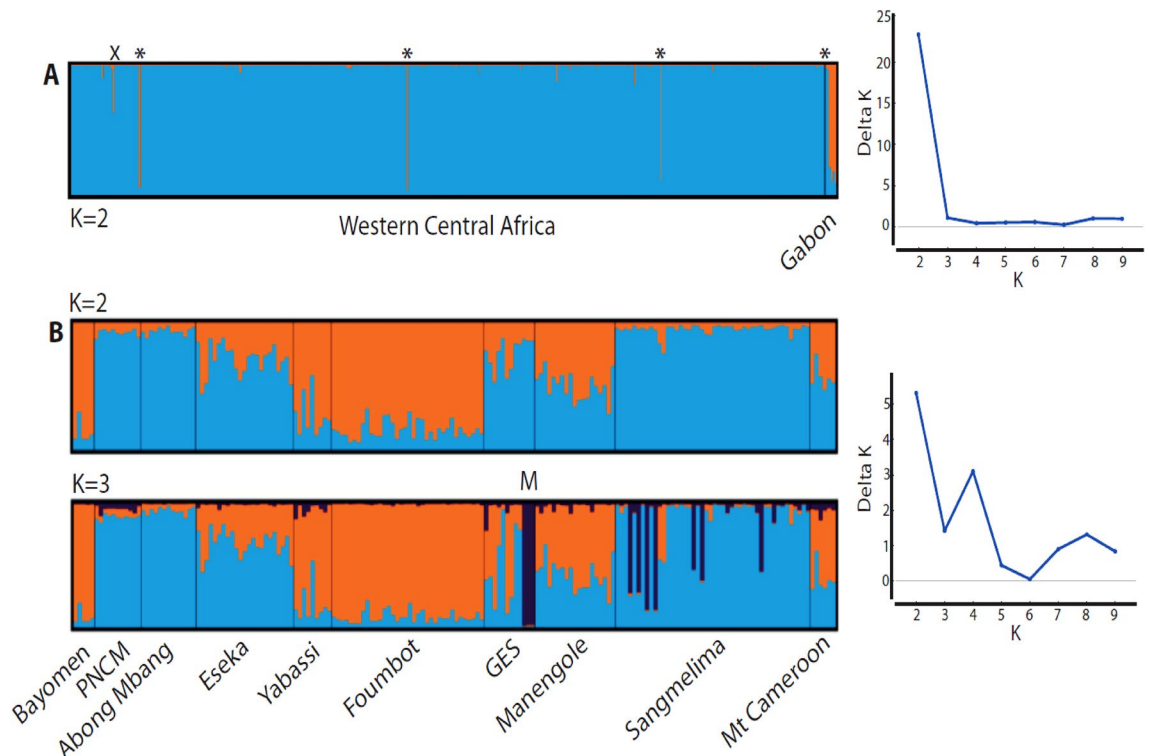
Among populations, mean number of alleles per locus (Na) was 6.4 and ranged from 4.0 (Bayomen) to 12.7 (Yaoundé market). The overall mean observed heterozygosity (Ho) was 0.579 (0.534–0.610) and average expected heterozygosity (He) was 0.664 (0.593–0.710) (Table 1). Overall, inbreeding coefficient (F<sub>IS</sub>) values were positive (F<sub>IS</sub> = 0.168) and ranged from 0.112 in GES to 0.217 in Eseka. Seven F<sub>IS</sub> values were significant. The mean allelic richness (A<sub>R</sub>) was 4.093 and was the highest in airport seizure (5.285) and urban markets (Table 1). Numbers of private alleles were the greatest for Gabon (15) and Sangmelima (11) when the urban bushmeat markets from Yaoundé and Douala were not included. When included, Yaoundé showed the highest number of private alleles (42). Observed frequencies of private alleles as extracted from ADZE were the greatest for Gabon (0.922–0.932) whether urban bushmeat markets were included or not. Frequencies of private alleles were low for the Douala (0.111) and Yaoundé (0.141) markets.

#### Genetic structure

The PCoA did not show any clear structuring of the genetic space within WBP from WCA and Gab (Supplementary Figure S7 online).

Clustering analysis with STRUCTURE, within all individuals and with prior information on lineage detected K = 2 as the most likely number of clusters (Fig. 3A). The two clusters did not totally correspond to the mitochondrial geographic assignment of the individuals regarding WCA and Gab lineages (Fig. 3A). In this case, 98% of all WBP individuals of WCA clustered into one population. Five out of eight mtDNA-assigned to Gab individuals clustered into a single group (Gab lineage). Three individuals (Y110, Ab10 and Ab2) from Yaoundé and Abong Mbang in Cameroon that were mtDNA-assigned to Gab clustered with WCA. Three individuals from Yaoundé and Douala (Y360, Y101 and D1aB103) that were mtDNA-assigned to WCA clustered with Gab. One individual from Yokadouma (Ou4) in Cameroon showed an admixed pattern between WCA and Gab (c. 60/40% assignment, respectively). From K = 3, no specific population structure was found (Supplementary Fig. S8 online).

Within WCA and among reference populations, STRUCTURE detected K = 2 as the most likely number of clusters (Fig. 3B). Two 'pure' genetic clusters regrouped Bayomen, Yabassi, and Foumbot from western and central Cameroon on one side, and PNCM, Abong Mbang and Sangmelima from North of Equatorial Guinea and southern / central Cameroon on the other side. Several reference populations from western, central and southern Cameroon, North of Gabon and South of Equatorial Guinea were admixed (Eseka, GES, Manengole and Mt Cam), and did not correspond to coherent geographic delineations (Fig. 3B). At K = 3, the four individuals from Medouneu (northern Gabon; GES population) clustered into an exclusive group. Seven individuals



**Figure 3.** Assignment plots among individuals of white-bellied pangolins from western central Africa as obtained with STRUCTURE (left) and most likely number of population clusters ( $K = 2$ ) following the  $\Delta K$  method (right). Each individual is represented by a vertical bar. A- including all the samples from urban markets, seizures and reference populations ( $N = 558$ ), sorted according to their mtDNA-based lineage assignment (WCA vs. Gab); asterisks indicate cyto-nuclear hybrids; X is an admixed individual ( $> 20\%$  of genomic ancestry shared with Gab). B- only retaining the samples from 10 reference populations ( $N = 181$ ); M = cluster from Medouneu (northern Gabon).

from Sangmelima showed full to admixed assignment to the Medouneu cluster. From  $K = 4$ , no further coherent geographic structuring was found (Supplementary Fig. S9 online).

Geneland analysis within WCA did not allow to geographically delineate consistent populations. Optimal number of populations varied between 10 and 13 clusters (Supplementary Fig. S10 online), with 11 as the most probable number of clusters according to the second-step iterative run (Supplementary Fig. S10 online). Assignment of individuals to populations among the best five runs was greatly unstable (data not shown).

$F_{ST}$  values were significantly different ( $p \leq 0.001$ ) between all WCA populations and Gab (from 0.184 to 0.225; Table S5). Pairwise differentiation between WCA populations was low, ranging from 0.006 (PNCM–Sangmelima) to 0.048 (Bayomen–Southern Equatorial Guinea).

We detected a significant IBD pattern among WBP individuals across the studied range ( $r = 0.099$ ,  $p = 0.001$ ), whereas there was no significant effect among populations ( $r = 0.298$ ,  $p = 0.094$ ; Supplementary Fig. S11 online).

#### Demographic history

We detected a significant signal of bottleneck in WCA with the S.M.M. model (Wilcoxon sign-rank test;  $p < 0.05$ ), whereas the T.P.M. model did not ( $p = 0.061$ ).

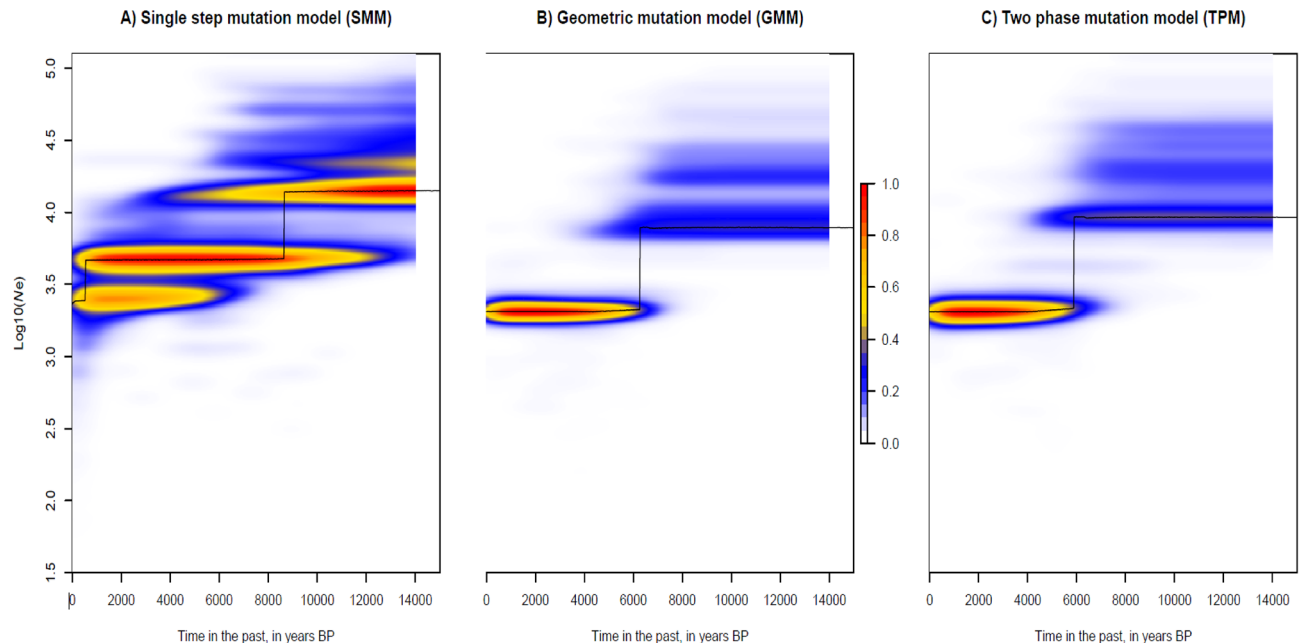
The assessment of the historical demography of WCA with *VarEff* showed that effective population size ( $N_e$ ) drastically declined –regardless of the models– over the last c. 9,840–6,140 years (4920–3070 generations; Fig. 4). Harmonic means estimated a 74.1–82.5%  $N_e$  reduction, from 7773–13,826 (ancestral  $N_e$ ) to 2,311–5,027 (contemporaneous  $N_e$ ).

#### Tracing the pangolin trade

A total of 535 (96%) of 557 WBP samples had unique genotypes. Eleven pairs of samples from Djoum, Manengole, Yaoundé and Medouneu shared a same genotype. The null hypothesis of encountering the same genotypes more than once by chance was rejected for all pairs of individuals ( $P < 0.0001$ ). Unbiased probability of identity (uPI) and probability of identity among siblings (PIsibs) were both very low (uPI =  $0.00125 \pm 0.00538$ ; PIsibs =  $0.0236 \pm 0.07236$ ; Supplementary Fig. S12 online).

Seven of a total of 11 loci that presented appropriate private allele signatures for one or several populations using private allele rarefaction, could be related to 18 private alleles observed in the reference populations (WCA; Supplementary Table S6 online). On this basis, we could trace a total of 47 (12.8%) individuals sold as part of the





**Figure 4.** Temporal change in effective population size ( $N_e$ ) of white-bellied pangolins from Western Central Africa, as estimated using VarEff under three different mutation models. Harmonic means (black line) and kernel density (color scale) of  $N_e$  posterior distributions are given in years BP.

bushmeat trade in the study region (Supplementary Table S7 online) to seven potential reference populations (see Fig. 5). Thirty-five WBP from the Yaoundé market were harvested from Sangmelima, Eseka, Mt Cameroon, Abong Mbang, Manengole and South of Equatorial Guinea (GES). Ten individuals sold in Douala markets came from Abong Mbang, Sangmelima and Yabassi. Two pangolins seized at Brussels airport were traced to South of Equatorial Guinea and Abong Mbang.

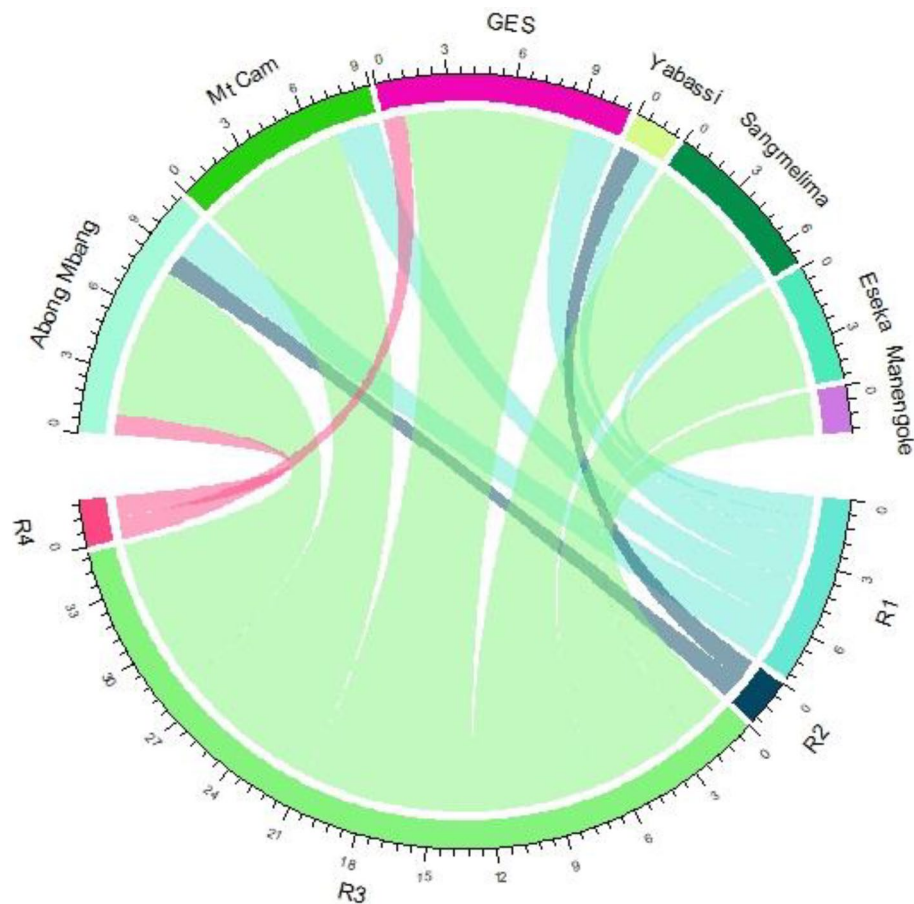
## Discussion

### White-bellied pangolins from western central Africa: range-overlapping lineages and weak geographic structuring

Our exhaustive sample set across three bordering countries, representing 563 newly collected samples of WBP from Cameroon, Equatorial Guinea and Gabon, allowed refining the lineage distribution of WBP from western central Africa. We show that the two mitochondrial lineages described from the study region are overlapping in ranges, with Gab being present all over Gabon and reaching southern Cameroon, while the southern extension of WCA reaches northern Gabon. Thus, our study clarifies the geographic distribution of the two lineages in western central Africa as previously based on a reduced number of samples<sup>20</sup>, and suggests that Gab is not endemic to Gabon. The three individuals representing Gab in Cameroon were found in the Yaoundé market and in the reference population of Abong Mbang. The latter corresponds to a local market that is off the main road network connecting northern Gabon to southern Cameroon<sup>27</sup>. The involved market stakeholder indicated that the two pangolins originated from the Lomié forest (c. 100 km from the selling site), thus suggesting that the Gab lineage naturally occurs in Cameroon.

Microsatellites genotyping showed that WCA and Gab corresponded to differentiated evolutionary units (mean  $F_{ST} = 0.192$ ), in line with previously published nDNA sequencing data<sup>20</sup> and with the mtDNA pattern found in the present study. However, because the three mtDNA-assigned Gab individuals from Cameroon were cyto-nuclear hybrids (i.e., assigned to the WCA cluster using microsatellites), ancient gene flow between the two lineages might be involved and is blurring their geographic delineation, notably in their supposed range of co-occurrence. All the more since an admixed individual from Yokadouma, in south-eastern Cameroon (Yokadouma), provides evidence for ongoing (but probably limited) introgression between the two lineages. Further sampling in forests from southern Cameroon and northern Gabon will have to be conducted in order to achieve a more comprehensive picture of the respective ranges of the WCA and Gab genetic entities, and their potential zones of admixture.

Despite the use of hypervariable microsatellites markers and a fair level of genetic diversity observed across populations<sup>21</sup>; this study) along the c. 1,000 km extension of our study zone, genetic structuring and differentiation within WCA was weak. Structure was the only analysis capable of detecting population clusters, using the available signal of geographic structuring based on prior delineation of populations ( $r < 1$ ; data not shown). However, the two 'pure' genetic clusters did not correspond to anything coherent in space, with a first cluster located strictly south of the Sanaga river (PNCM + Abong Mbang + Sangmelima) and the other spread on both sides of the river (Bayomen [north] vs. Yabassi and Foubot [south]), while the admixed individuals were also present north and south of the river (Manengole and Mt Cam [north] vs. Eseka and GES [south]).



**Figure 5.** Chord diagram tracing the geographic origin of white-bellied pangolins from urban bushmeat markets and seizures. GES = South of Equatorial Guinea; R1 = Douala Dacat market; R2 = Douala Central market; R3 = Yaoundé Nkolndongo market; R4 = Seizures at Brussels airport.

At  $K = 3$ , WBP from northern Gabon (Medouneu), which corresponds to the southernmost limit of WCA, were identified as a distinct population cluster. Because Medouneu is located c. 50 km from Monts de Cristal, a putative Pleistocene rainforest refugium<sup>28</sup>, pangolins from this area may hold the signature of a relict population, as previously observed in mammals and plants. However, the lack of population structuring in other areas of putative refugia such as southern Gabon and western Cameroon<sup>28</sup>, both sampled in our study, questions this interpretation; especially since Medouneu is located at the bottom end of the North–South IBD pattern that we observed in WCA (see below), which could be the cause of such apparent clustering (see<sup>29</sup>). The refugium hypothesis is also challenged by the presence of Medouneu-related individuals (pure or admixed) further north in southern Cameroon (Sangmelima), suggesting a more extended range for this population.

The lack of genetic structure and population differentiation across WCA is to be interpreted in the light of the IBD pattern that we detected among WBP individuals. Although our study did not quantify dispersal rate and range, the pattern that we observed fits a stepwise migration model where dispersal distances are large enough to lead to near-panmictic genetic patterns<sup>30</sup>. Our results confirm the non-influential role of the Sanaga river on the biogeography and population delimitation of several African taxa (see<sup>31</sup>), although its role as a limiting –i.e., not prohibitive– factor to gene flow (e.g.,<sup>32</sup>) was not tested here. They also suggest that there is no specific barrier to gene flow for WBP in the study zone (WCA), in line with the lack of structure found in another lineage of WBP investigated with the same markers (Dahomey Gap;<sup>22</sup>). Our population genetic study again provides evidence for unsuspected dispersal capacities in WBP, which is also coherent with a recent, spatial and/or demographic expansion across the WCA range following forest post-glacial expansion during the Late Pleistocene (e.g.,<sup>33</sup>), as suggested by the mtDNA analyses. Despite habitat fragmentation and hunting activities occurring in the study zone<sup>14</sup>, dispersal may still be maintained given the capacity of the species to cross and occur in anthropized habitats<sup>34</sup> and the still prevalent, large forested areas available<sup>35</sup>. Field surveys are blatantly needed to document the dispersal behavior of WBP, notably during their early phase of supposed vagrancy<sup>36</sup>.

The only individual from Bioko Isl. that we could sequence and genotype did not show any genetic differentiation from the continent. Thus, the question of whether WBP sold on Bioko Isl. are imported from the continent (Cameroon) or not cannot be answered here<sup>8</sup>, and as a corollary, that of a surviving population endemic to the island. Additional genetic surveys in Bioko Isl., both from bushmeat markets and field sites, will have to be

conducted in order to assess the status of WBP on the island, an urgent matter of conservation as island populations are often subject to rapid genetic drift, accumulation of deleterious variants and inbreeding.

### Highest levels of genetic diversity and Wahlund effect in the context of a recent demographic decline

Overall, populations from the WCA and Gab lineages were characterized by the highest levels of genetic diversity described in WBP. This is important for the conservation of the species in western central Africa, as fair levels of genetic diversity are expected to promote population fitness and ensure adaptability to environmental change<sup>37</sup>. MtDNA nucleotide diversity was by far the highest for WCA (largest sampling effort;  $N > 550$ ) and Gab (reduced sampling effort, similar to other lineages;  $N = 13$ ), c. 1.7–9.5 times greater compared to the other lineages. Mean allelic richness in WCA was also higher than in the Dahomey Gap lineage ( $A_R = 10.79$  vs. 4.27; minimum sample size = 169;<sup>21,22</sup>), the only comparable study available.

The generally significant deficit of heterozygotes observed among populations ( $F_{IS} = 0.112–0.217$ ) could lead to the conclusion that WBP populations from WCA are inbred. However, inbreeding would come in contradiction with the fairly high levels of heterozygosity observed among populations (mean  $H_o = 0.579$ ; mean  $H_e = 0.664$ ) and the weak geographic structuring observed across WCA (see<sup>38</sup>). Because population differentiation ( $F_{ST}$ ) across WCA was low whereas  $F_{IS}$  values were high, our results fit the expectation of a Wahlund effect, i.e. where a same population regroups individuals from genetically differentiated entities<sup>39</sup>. Such signature could be reached through regular, long-range inter-connectivity among WBP populations, implying again unsuspected dispersal capacities of the species. However, given the knowledge gaps affecting the natural history of WBP and African pangolins in general<sup>4</sup>, we cannot discard the fact that some unknown breeding strategy of the species could participate to this Wahlund effect.

We identified an important decline in the effective population size of WCA (as mirrored by the bottleneck signature detected in one of our analyses) occurring c. 4,300–7,440 ya, a time of climatic stability favourable to rainforest cover and that predated a major period of savannah expansion and forest fragmentation in western central Africa (starting c. 4000 ya). Rather than a paleoclimatic influence, we hypothesize that the arising of the Ceramic Late Stone Age culture in the region from c. 7000 BP onward –characterized by bifacial macrolithic and polished stone tools<sup>40</sup>–could have constituted the technological change (from “Stone to Metal Age”;<sup>41</sup>) that made hunting more efficient and responsible for the decline of WBP. Such decline, although drastic (loss of c. 74.10–82.5% of the effective population size), places the effective populations size of WBP from WCA ( $2161 \pm 219$ ) within the conservative threshold range of minimum viable population size (500–5000). Because viable population size levels are paramount to maintaining acceptable levels of genetic diversity, it will be important to create or maintain forest corridors among WBP populations and mitigate deforestation and illegal hunting activities in order to preserve the genetic diversity levels of pangolin populations from the study region.

### An endemic, trans-national trade of white-bellied pangolins from western central Africa, with Yaoundé as a main domestic hub

We showed that the 20 microsatellite loci specifically developed for WBP<sup>21</sup> provided the necessary power to confidently distinguish among all the WCA and Gab individuals. Only five microsatellite loci were needed to reach the conservative value of probability of identity  $< 0.01$ <sup>42</sup>, compared to the seven loci required for distinguishing pangolins from the Dahomey Gap<sup>22</sup>. This has important implications for the implementation of a convenient genetic tracing tool applied to the pangolin trade in western central Africa. Indeed, even a subset of our markers would be capable of estimating the exact number of individuals from scale seizures and trace the parallel trafficking networks of scale and body trade where same individuals can be dispatched<sup>8</sup>. Microsatellites genotyping also allowed identifying eleven cases where a same pangolin had been sampled twice, illustrating the difficulties of relying on third parties such as trained local assistants who are part of the bushmeat chain<sup>14</sup>, but at the same time showing the usefulness of our markers to counter such sampling bias.

Although mtDNA was not resolute enough to trace the trade at the local scale, our results showed that the trade of WBP was endemic to western central Africa, involving both WCA and Gab lineages but excluding the other lineages described from central and West Africa. This is in line with the trade endemism also found for the Dahomey Gap lineage<sup>22</sup>, and suggests that large, urban bushmeat markets (here, Douala and Yaoundé for the most sampled) are not necessarily hubs for long-range, international trade of WBP. Rather, they represent domestic hubs that source large volumes of WBP at the sub-regional level (i.e. country-wide and neighboring countries), but are not directly involved in global trade networks such as those identified from harbor or airport seizures in Africa and Asia<sup>9</sup>.

The private allele frequency approach proposed by Zanvo et al.<sup>22</sup> allowed us to infer geographic origin hypotheses for 47 WBP sold from three urban bushmeat markets in Cameroon and seized at a European airport, representing c. 13% of the traceable individuals from WCA. Despite this relative success and the fact that our sample set covered quite exhaustively the WCA range, such inferences should be considered with caution as private allele frequency estimates will depend on the genetic representativeness of the sampled populations and the dispersal dynamics among populations<sup>43,44</sup>. Although we carefully controlled for the sampling effort (number of individuals) bias within populations using rarefaction, a denser geographic sampling of reference populations might dilute the tracing signal that we obtained in this study. In order to validate the hypotheses of trading sources that we generated, more powerful -genomic- approaches will be needed<sup>23,26</sup>.

Our results suggest that the Yaoundé market is the main national hub for the trade of bushmeat in Cameroon<sup>27</sup>, recruiting among six of the seven geographic sources that were traced as part of this study. Sourcing occurs in the western, southern and eastern part of the country, but also in Equatorial Guinea, representing a distance network



between c. 120 and 370 km. Such inferences contrast with the more localized sources of the international trade that were recently inferred for the species<sup>23</sup>.

One geographic sourcing North of Douala and exclusive to the Yaoundé market, Manengolé, may provide evidence for transit of pangolins through Douala to Yaoundé, urban bushmeat markets in this case playing both a role of drop-off point and transit zone<sup>27</sup>. The two bushmeat markets from Douala showed contrasting spheres of influence, with Dacat being fed by a minimum of five geographic sources *contra* solely two sources for the Central market. Having said that, both showed a wide range of sourcing, respectively extending c. 600 and 500 km from Douala, beyond the maximal range delineated for Yaoundé. The two markets shared a unique supply source through the Yabassi population, situated c. 90 km north of Douala, in line with the recognized role of the unprotected Ebo forest as a feeding source of the Douala bushmeat market places<sup>14</sup>.

Eventually, we could trace the presence of WBP from the forests of southern Equatorial Guinea in Douala and Yaoundé markets, and Brussels airport (with flights originating from Cameroon). This result suggests that trans-national trade across neighbouring countries occur in the study region, potentially following the main road network connecting Equatorial Guinea to Cameroon, and that Cameroon constitutes a major hub of this sub-regional trade through the large feeding network of its urban bushmeat markets (<sup>8</sup>; <https://cameroonvoice.com/news/2021/07/22/cameroun-trafiquants-de-pangolin-en-afrique-centrale-demasques/>; this study). On the other hand, the likely occurrence of the Gab lineage and admixture between WCA and Gab or individuals from northern Gabon (Medouneu) in Cameroon blurred our tracing ability for a potential WBP trade occurring between Gabon and Cameroon, as previously mentioned in the literature<sup>46</sup> and recent media covers (<https://www.gabonmediatime.com/216-pointes-divoire-saisies-cameroun-provenant-gabon-reactions-de-lanpn-gouvernement-attendues/>).

The number of haplotypes and number of private alleles were strikingly high in the Yaoundé market, supporting the hypothesis that Yaoundé is one of the main pangolin trade hubs of the study region. This was confirmed by the 12 different geographic sources mentioned by bushmeat vendors when asked about the origin of the pangolins sold in Yaoundé ( $N = 59$ ; data not shown), including Sangmelima (and its surroundings), Mekin, Djoum, Akonolinga, Makenene, Bafia, Ma'an, Campo Ayos, Lomié, Yokadouma, Foubot, Mbouda, Magba, Eseka, Mboumyebel and Bankim. Such wide geographic network, obviously minimized by the genetic data set at hands, suggests a feeding network with a maximum sourcing distance of c. 600 km from the city, more in line with the large market's network range expected from the literature (see above). The great number of private alleles detected for the Yaoundé market (42 *vs.* 0 to 8 in reference populations and Douala markets, and 15 in Gabon) could suggest that, despite our considerable sampling effort, major reference populations remain unsampled. However, the relatively low private allele frequency calculated from Yaoundé supports the view that such high numbers of private alleles are an effect of sampling bias (> 250 samples). We believe that additional sampling of reference populations will have to be conducted across western central Africa to reach higher accuracy in the tracing of the WBP trade, although the theoretical numbers of sampled individuals for obtaining confident estimates of allele frequencies among populations for forensic purpose might be hardly achievable (see<sup>48</sup>).

## Conclusion

Our study provides an unprecedented assessment of the genetic status of WBP from western central Africa. Combining mtDNA sequencing and microsatellites genotyping of several hundreds of pangolins, we identified range overlap between the WCA and Gab mitochondrial lineages with limited introgression in southern Cameroon, so the two lineages can still be considered separate evolutionary units. Further sampling and sequencing efforts –particularly through high-throughput sequencing approaches– will have to be conducted across the study region, notably in Gabon where the number of genetic samples was low, in order to assess more accurately the admixture pattern between WCA and Gab and geographic structuring within WCA. Ideally, sampling will have to be oriented towards forest sites in order to minimize the potential bias of inaccurate sourcing radius of the local markets herein used as reference populations (notably relative to a potentially artifactual Wahlund effect). Such refined data will be crucial to establish a genetic-based management unit delineation strategy for guiding the conservation of WBP in the study region. As a corollary, new knowledge on the dispersal and breeding behavior of WBP, which remains literally unknown, will have to be produced urgently.

Our set of markers provides a unique opportunity to trace the WBP trade at the individual and population levels across western central Africa, and should be considered a valuable resource for future forensic applications at reasonable costs. For the geographic tracing of WBP from western central Africa, and because single locus (mtDNA)-typing has recently been applied to trace the pangolin trade<sup>9</sup>, we recommend a multi-locus approach given the admixture pattern observed between WCA and Gab. Because introgression may occur between other neighboring lineages of WBP<sup>49</sup>, it will be important in the future to rely on a global reference database where all the WBP lineages will have been genotyped.

According to our results and previous investigations, Yaoundé is a main hub for the pangolin trade in the sub-region, where WBP are openly sold on the stalls despite being nationally protected<sup>14</sup>. The Yaoundé market is driven by a large urban demand for bushmeat consumption, and as such should receive urgent attention from researchers, conservationists and anti-poaching services as a crucial point of entry to mitigate the WBP trade in Cameroon and, more globally, western central Africa. A similar conclusion applies to European points of entry of WBP carcasses, where CITES regulations apply but are in practice not efficiently implemented because of a global lack of prioritization, means and training at European borders<sup>14</sup>. As a general conclusion, better sensitization of range-state wildlife management services and foreign customs agencies to the regional and global trade of WBP should help prioritizing the conservation of the species.

## Methods

### Genetic sampling and wet laboratory procedures

We collected 571 genetic samples of WBP from the Western Central Africa lineage (WCA), across Cameroon, Equatorial Guinea (including Bioko Island) and Gabon. Because the endemic Gabon lineage (Gab) co-occurs with WCA in Gabon<sup>20</sup>, we assume that samples from Gab were also collected. We surveyed and collected 204 samples from 37 local bushmeat markets and forest sites, and 358 samples from three large urban bushmeat markets in Cameroon (Douala and Yaoundé; see Supplementary Fig. S1 and Supplementary Table S1 online). We collected through upstream interviews with urban market sellers' information on the geographic origin of the pangolins sold. Using the same approach with local market vendors and cross-checking information with local authorities and NGOs involved in wildlife trade surveys, we could confirm that all the pangolins collected on the local bushmeat markets were originating from the vicinity (< c. 100 km) of the surveyed sites, and thus could be used as proxies of "populations" within the large-scale geographic framework (c. 1,000 km North–South) of our study. Free consent from hunters and market sellers was obtained before collecting samples. We relied on an opportunistic sampling strategy (see<sup>27</sup>), without financial incentives. Sample types consisted of fresh skin, tissue (muscle) and tongue taken from dead animals. In each case, samples were taken beneath the tissue surface, so to prevent any potential exogenous contamination. A series of 12 samples were also collected from seized luggage originating from Cameroonian flights at Roissy (Paris, France) and Brussels (Belgium) airports.

Genomic DNA extraction was performed using the NucleoSpin® Tissue Kit (Macherey–Nagel, Hoerd, France), following manufacturer's recommendations. Final elution step was repeated twice in 50 µl elution buffer to increase DNA yield. DNA concentration was quantified using the NanoDrop 1000 Spectrophotometer (ThermoFisher Scientific, Illkirch–Graffenstaden, France).

We amplified a mitochondrial fragment of 402 bp from cytochrome *b* (cyt *b*) able to differentiate among the six WBP lineages, following Gaubert et al.<sup>20</sup>. PCR products were sequenced at Genoscreen (<https://www.genoscreen.fr/en>; Lille, France) and Macrogen Europe (<https://dna.macrogen-europe.com/en>; Amsterdam, the Netherlands). Sequences were manually aligned in BioEdit v7.0.5<sup>50</sup>. All the sequences produced in this study were deposited in Genbank under accession numbers OP524750–OP525226.

We amplified 20 microsatellite markers developed from the genome of WBP using four multiplexes (4 to 6 loci) following Aguillon et al.<sup>21</sup>. PCR products were run on an ABI 3730 DNA Analyser (Thermo Fisher Scientific) at Genoscreen and GeT-PlaGe (<https://get.genotoul.fr/>; INRAE, Toulouse, France).

### Data analysis–Cytochrome *b*

#### Clustering analysis

We used cyt *b* to phylogenetically assign the collected samples to their respective mitochondrial lineages, according to Gaubert et al. (2016). For that purpose, we combined our samples from Cameroon (N = 505), Equatorial Guinea (N = 15), Bioko Island (N = 1), Gabon (N = 18) and European seizures (N = 12) to the cyt *b* sequences available for WBP at the time of the study (N = 76). A clustering distance tree was obtained with MEGA-X v10.2.2 using Neighbor Joining and the Kimura 2-parameter model<sup>51</sup>. Node support was assessed through 500 bootstrap replicates.

#### Genetic diversity and structure

Based on the new cyt *b* sequences that we produced, we reassessed the levels of genetic variation in WCA and Gab, after removing sequences with missing data (remaining N = 563). We used DnaSP v6.12 to compute number of polymorphic sites (*S*), haplotype number (*h*), haplotype diversity (*Hd*) and nucleotide diversity ( $\pi$ ) (Supplementary Table S2 online). We mapped the geographic distribution of WCA and Gab haplotypes using ArcGIS 10.1 (Esri France). We ran Network v10.2.0.0 to build a median-joining haplotype network with  $\epsilon = 0$ , so to minimize the presence of alternative median networks.

#### Demographic history

Because of the low number of samples attributed to Gab (N = 11; see Results), we explored the demographic history of WCA only (N = 552). We performed mismatch analysis using Arlequin 3.5.2 to test for signatures of demographic and spatial expansions, by calculating the sum of squared deviations (SSD) between observed and expected distributions using 1,000 bootstrap replicates.

Past variation in effective population size was assessed through a Bayesian skyline plot analysis in BEAST 2.6.3. We used the Bayesian Model Test option (bModelTest;) for site modeling, a coalescent Bayesian skyline prior and a relaxed clock (log normal-distributed). We fixed the mean mutation rate (*ucl*d mean) to 0.026 substitution/site/Myrs with a normal distribution ranging from 0.010 to 0.045, as extracted from the posterior distribution values of the calibrated tree of Gaubert et al.<sup>20</sup> (data not shown). We ran 500.10<sup>6</sup> MCMC iterations, with trees and model parameters sampled every 10,000 generations and a 10% burnin. Parameter convergence and Bayesian skyline plot reconstruction were, respectively, visualized and performed in Tracer 1.7.1. Effective population size ( $N_e$ ) was estimated considering a generation time of 2 years in WBP<sup>22</sup>.

### Data analysis–Microsatellites

#### Genetic diversity

We used Geneious 9.0.5<sup>52</sup> and the Microsatellites plugin (<https://www.geneious.com/features/microsatellite-genotyping/>) for allele scoring and genotype extraction. Only WCA and Gabon individuals with at least 75% of genotyping success were considered for the analyses (N = 558). The 75% threshold coincides with  $\geq 15$  microsatellite markers, which triples the minimum number of loci needed to discriminate among WCA individuals<sup>21</sup>.

Detection of null alleles in the 20 loci was performed with Microchecker 2.2.3<sup>53</sup>, assuming population at equilibrium, on the Foubot population ( $N = 31$ ). Deviation from Hardy–Weinberg Equilibrium (HWE) was calculated in GenAlEx 6.5<sup>54</sup>. Linkage Disequilibrium (LD) was assessed using FSTAT 2.9.4<sup>55</sup> after 1000 randomizations ( $P\text{-value} \geq 0.00005$ ). The Bonferroni correction was applied to null hypothesis testing in those three analyses.

Genetic diversity was assessed through the number of alleles ( $N_a$ ), observed ( $H_o$ ), expected ( $H_e$ ) and unbiased expected ( $uH_e$ ) heterozygosity (in GenAlEx), allelic richness ( $A_R$ ) and deficit of heterozygotes ( $F_{IS}$ ) (in FSTAT), as mean values per population (see below).

#### Genetic structure

Global genetic variance among all genotyped individuals was assessed through pairwise population matrix unbiased genetic distances, and was visualized in GenAlEx using Principal Coordinates Analysis (PCoA).

For downstream population-based analyses, we defined the following 11 populations based on geographic proximity and biogeographic continuum among forests: Bayomen (Bayomen and Bangong [Bafia forest];  $N = 5$ ), Mt Cam (Manyemen and Bayib-Assibong [forest around National Park of Mt Cameroon];  $N = 6$ ), Eseka (Eseka, Mamb, Lolodorf, Bipindi and Eseka [Bassa's forest]  $N = 22$ ), GES (forests of southern Equatorial Guinea;  $N = 12$ ), PNCM (Campo, Ma'an [forest of National Park of Campo Ma'an] and northern Equatorial Guinea;  $N = 11$ ), Foubot (Foubot [Noun/Tikar forests];  $N = 36$ ), Yabassi ([Ebo Forest];  $N = 9$ ), Manengole ([Manengumba Bakossi Forest];  $N = 19$ ), Abong Mbang (Lomié [forests of East Cameroon];  $N = 15$ ), Sangmelima (Sangmelima and Djoum [forests around Dja Biosphere Reserve];  $N = 45$ ), and Gab (North and South of Gabon;  $N = 5$ ). Samples from Yokadouma ( $N = 3$ ), Ekombitie ( $N = 1$ ), Esse ( $N = 2$ ), Akonolinga ( $N = 1$ ), Nditam ( $N = 3$ ), Yokadouma ( $N = 2$ ) and Bioko Isl. ( $N = 1$ ) (see Supplementary Table S1 online) were excluded because not proximate enough to any site to be merged and reach a sufficient number of samples.

We used STRUCTURE 2.3.4<sup>56</sup> to conduct a clustering analysis, following a two-step dataset refinement procedure, where (i) all the individuals from the study area were included (WCA + Gab;  $N = 558$ ), and (ii) only the WCA individuals were included (WCA;  $N = 181$ ), discarding large urban markets and cyto-nuclear hybrids or admixed individuals (who with at least 20% of shared ancestry). The first scheme allowed screening the global genetic structure of WBP at the largest study scale. The second scheme aimed at assessing the genetic structure among WCA reference populations. The first and second schemes were run with the LOCPRIOR model, notably to manage the sampling bias among lineages (first scheme) and populations (second scheme;<sup>57</sup>). We performed for each scheme 20 independent runs for  $K = 2$  to 10 with  $10^5$  Markov chain Monte Carlo (MCMC) iterations (burnin =  $10^4$ ), assuming admixture and uncorrelated (first scheme) or correlated (second scheme) allele frequencies. We used STRUCTURE HARVESTER 0.6.94 to estimate the most likely number of populations ( $K$ ) using the  $\Delta K$  method<sup>58</sup>. Graphic outputs from independent runs were summarized with CLUMPP and CLUMPAK server (<http://clumpak.tau.ac.il/>) was used to generate the bar plots of population assignment.

We also inferred the most likely number of geographic populations using WCA georeferenced individuals, through the *Geneland* package<sup>59</sup> in R 4.0.5 (*R Team Development Core* 2021). We followed Coulon et al.<sup>60</sup> by first allowing  $K$  to vary from 1 to 20 and launched five runs of  $5 \cdot 10^5$  MCMC iterations (thinning = 500; burnin = 500) under the frequency-correlated model and 1 km of uncertainty for spatial coordinates. Second, we fixed the number of estimated populations to  $K = 10$ –13 after the results of the first analysis [highest mean posterior densities for:  $K = 10$  (0.21),  $K = 11$  (0.30),  $K = 12$  (0.26) and  $K = 13$  (0.15)], and performed 20 independent runs with the same parameters. In the final analysis, the stability of individual assignment to populations was assessed among the best five runs (*i.e.* with the highest posterior probabilities). We used  $500 \times 500$  pixels to map the posterior probabilities of population membership.

Pairwise differentiation ( $F_{ST}$ ) among the 11 populations as defined above was computed in Arlequin 3.5<sup>61</sup>. Significant differentiation values ( $P \leq 0.05$ ) were estimated using  $10^5$  MCMC iteration chains and 10,000 dememorization steps.

We tested isolation-by-distance (IBD) within WCA individuals and among WCA populations through the R package *pegas*<sup>62</sup> using a Mantel test. The significance of the correlation ( $r$ ) between individual-based genetic (Edward's) and geographic (Euclidean) distances was estimated through 10,000 permutations.

#### Tracing the pangolin trade

The discriminative power of our 20 microsatellite markers among WBP individuals was evaluated after three different approaches: (i) counting the number of identical genotypes among samples using the Multilocus tagging option in GenAlEx (suboption Matches), (ii) computing the probability of encountering the same genotype more than once by chance using the R package *poppr* (method = single;<sup>63</sup>), and (iii) calculating the indices of unbiased probability of identity and probability of identity among siblings ( $uPI$  and  $PI_{sibs}$ ) using Gimlet 1.3.3<sup>64</sup>.

Because we could not find a deep level of genetic structuring within WCA, we followed the protocol based on private allele frequency rarefaction described by Zanco et al.<sup>22</sup> to trace the trade of WBP in our study area. We used the generalized rarefaction approach implemented in ADZE<sup>65</sup> to compute private allele frequencies (paf) among various combinations of populations from the 10 source populations as defined above (removing Gab). Because the original scheme of 10 populations yielded the greatest inferred values of paf (data not shown), we graphed the paf rarefaction curves (sample size rarefied from 2 to 5, 5 being the lowest number of individuals among populations) for each locus in these 10 populations. Only loci for which paf were  $> 40\%$  and reached a plateau or showed an increasing trend at  $N = 5$  for a given population were considered as potentially useful for tracing the origin of WBP found in the urban markets. We then cross-filtered these results with the private alleles observed in the actual populations (using GenAlEx), and eventually retained the loci that showed both high potential for tracing (ADZE output) and observed private alleles (GenAlEx output) for the given populations.

As a final step, we manually screened the private alleles present in WBP sampled from the Cameroonian large urban markets (Douala and Yaoundé) and international seizures to trace their geographical origins.

#### Demographic history

We tested for the signature of demographic bottlenecks in WCA (excluding hybrid and admixed individuals) using both the Single Mutation Model (SMM) and the Two Phase Model (TPM) in BOTTLENECK 1.2.02<sup>66</sup>. We applied the Wilcoxon sign-rank test to estimate heterozygote excess/deficit using 10,000 replications.

We assessed temporal changes in the effective population size of WCA through an approximate-likelihood approach implemented in the R package *VarEff*<sup>67</sup>. We followed Zanzo et al.<sup>22</sup> by fixing a generation time of 2 yrs and a microsatellites mutation rate of  $5.10^{-4}$  per generations<sup>68</sup>. We ran the analysis with three mutation models (single, geometric and two phases), using 10,000 MCMC batches (length = 1, thinned every 100 batches; JMAX = 3). The first 10,000 batches were discarded as burnin.

#### Data availability

The sequence data obtained in this study have been deposited at GenBank of NCBI at OP524750–OP525226 (<https://www.ncbi.nlm.nih.gov/>). We confirm that this work is original and has not been published elsewhere. We declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Received: 14 April 2023; Accepted: 30 May 2024

Published online: 07 June 2024

#### References

- Morton, O., Scheffers, B. R., Haugaasen, T. & Edwards, D. P. Impacts of wildlife trade on terrestrial biodiversity. *Nat. Ecol. Evol.* **5**, 540–548 (2021).
- Challender, D. W. S., Heinrich, S., Shepherd, C. R. & Katsis, L. K. D. International trade and trafficking in pangolins, 1900–2019. In *pangolins* (eds Challender, D. W. S. et al.) 259–276 (Academic Press, New York, 2020).
- Heinrich, S., Wittman, T. A., Ross, J. V., Shepherd, C. R. & Challender, D. W. S. Cassey, P. *The global trafficking of pangolins: a comprehensive summary of seizures and trafficking routes from 2010–2015*. (TRAFFIC Report, 2017).
- Heighton, S. P. & Gaubert, P. A timely systematic review on pangolin research, commercialization, and popularization to identify knowledge gaps and produce conservation guidelines. *Biol. Conserv.* **256**, 109042 (2021).
- Challender, D. W. S. & Watermann, C. Implementation of CITES decisions 17.239 b) and 17.240 on pangolins. *CITES SC69 Doc 57 Annex*, (2017).
- Hoffmann, R. & Challender, D. W. S. Conservation strategies and priority actions for pangolins. In *Pangolins* 531–535 (eds Challender, D. W. S. et al.) (Academic Press, New York, 2020). <https://doi.org/10.1016/B978-0-12-815507-3.00033-2>.
- Zanzo, S. et al. Ethnozoological and commercial drivers of the pangolin trade in Benin. *J. Ethnobiol. Ethnomed.* **17**, 1–11 (2021).
- Ingram, D. J., Cronin, D. T., Challender, D. W. S., Venditti, D. M. & Gonder, M. K. Characterising trafficking and trade of pangolins in the Gulf of Guinea. *Glob. Ecol. Conserv.* **17**, e00576 (2019).
- Zhang, H. et al. Genetic identification of African pangolins and their origin in illegal trade. *Glob. Ecol. Conserv.* **23**, e01119 (2020).
- Ingram, D. J. & Coad, L. *Hunting and sale of pangolins across sub-saharan Africa: a preliminary analysis. Offtake Working Paper 1* (2016).
- Ingram, D. J. et al. Assessing Africa-wide pangolin exploitation by scaling local data. *Conserv. Lett.* **11**, 1–9 (2018).
- Frutos, R., Serra-Cobo, J., Chen, T. & Devaux, C. A. COVID-19: time to exonerate the pangolin from the transmission of SARS-CoV-2 to humans. *Infect. Genet. Evol.* **84**, 104493 (2020).
- Saylors, K. E. et al. Market characteristics and zoonotic disease risk perception in Cameroon Bushmeat markets. *Soc. Sci. Med.* **268**, 113358 (2021).
- Din Dipita, A., Missoup, A. D., Tindo, M. & Gaubert, P. DNA-typing improves illegal wildlife trade surveys: Tracing the Cameroonian bushmeat trade. *Biol. Conserv.* **269**, 109552 (2022).
- Smart, U., Cihlar, J. C. & Budowle, B. International wildlife trafficking: A perspective on the challenges and potential forensic genetics solutions. *Forensic Sci. Int. Genet.* **54**, 102551 (2021).
- Kotze, A. et al. Understanding illegal trade in pangolins through forensics: applications in law enforcement. In *Pangolins* 321–332 (eds Challender, D. W. S. et al.) (Academic Press, New York, 2020). <https://doi.org/10.1016/B978-0-12-815507-3.00020-4>.
- Gaubert, P. & Antunes, A. 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 (2015).
- Baker, C. S. A truer measure of the market: The molecular ecology of fisheries and wildlife trade. *Mol. Ecol.* **17**, 3985–3998 (2008).
- Heinrich, S. et al. Where did all the pangolins go? International CITES trade in pangolin species. *Glob. Ecol. Conserv.* **8**, 241–253 (2016).
- Gaubert, P. et al. Phylogeography of the heavily poached African common pangolin (*Pholidota*, *Manis tricuspis*) reveals six cryptic lineages as traceable signatures of Pleistocene diversification. *Mol. Ecol.* **25**, 5975–5993 (2016).
- Aguiillon, S. et al. Development and characterization of 20 polymorphic microsatellite markers for the white-bellied pangolin *Phataginus tricuspis* (Mammalia, Pholidota). *Mol. Biol. Rep.* **47**, 4827–4833 (2020).
- Zanzo, S. et al. Can DNA help trace the local trade of pangolins? Conservation genetics of white-bellied pangolins from the Dahomey Gap (West Africa). *BMC Ecol. Evol.* **22**, 1–14 (2022).
- Tinsman, J. C. et al. Genomic analyses reveal poaching hotspots and illegal trade in pangolins from Africa to Asia. *Science*. **382**, 1282–1286 (2023).
- Cabana, F. et al. Feeding Asian pangolins: An assessment of current diets fed in institutions worldwide. *Zoo Biol.* **36**, 298–305 (2017).
- Harvey-Carroll, J. et al. Continued availability and sale of pangolins in a major urban bushmeat market in Cameroon despite national bans and the COVID-19 outbreak. *Afr. J. Ecol.* **00**, 1–7 (2022).
- Nash, H. C. et al. Conservation genomics reveals possible illegal trade routes and admixture across pangolin lineages in Southeast Asia. *Conserv. Genet.* **19**, 1083–1095 (2018).
- Randolph, S. G., Ingram, D. J., Curran, L. M., Holland Jones, J. & Durham, W. H. Urban wild meat markets in Cameroon: actors and motives. *World Dev.* **160**, 106060 (2022).
- Maley, J. The African rain forest - main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. *Proc. R. Soc. Edinburgh Sect. B Biol. Sci.* **104**, 31–73 (1996).



29. Perez, M. F. *et al.* Assessing population structure in the face of isolation by distance: are we neglecting the problem?. *Divers. Distrib.* **24**, 1883–1889 (2018).
30. Kinitz, T. *et al.* Strong isolation-by-distance in the absence of genetic population structure in the eelpout (*Zoarces viviparus*, Linnaeus 1758). *Ecol. Indic.* **27**, 116–122 (2013).
31. Fuchs, J. & Bowie, R. C. K. Concordant genetic structure in two species of woodpecker distributed across the primary West African biogeographic barriers. *Mol. Phylogenet. Evol.* **88**, 64–74 (2015).
32. Gonder, M. K. *et al.* Evidence from Cameroon reveals differences in the genetic structure and histories of chimpanzee populations. *Proc. Natl. Acad. Sci. USA* **108**, 4766–4771 (2011).
33. Piñeiro, R., Dauby, G., Kaymak, E. & Hardy, O. J. Pleistocene population expansions of shade-tolerant trees indicate fragmentation of the African rainforest during the ice ages. *Proc. R. Soc. B Biol. Sci.* **284**, 20171800–20171800 (2017).
34. Jansen, R. *et al.* White-bellied pangolin *Phataginus tricuspis* (Rafinesque, 1820). In *Pangolins 139–156* (eds Challender, D. W. S. *et al.*) (Academic Press, New York, 2020).
35. Gillet, P., Vermeulen, C., Feintrenie, L., Dessard, H. & Garcia, C. Quelles sont les causes de la déforestation dans le bassin du Congo ? Synthèse bibliographique et études de cas. *BASE* **20**, 183–194 (2016).
36. Gaubert, P. Family manidae (pangolins). *Handb. Mamm. World* **2**, 82–103 (2011).
37. Habel, J. C. & Schmitt, T. The burden of genetic diversity. *Biol. Conserv.* **147**, 270–274 (2012).
38. Wright, S. The theory of gene frequencies. *Evol. Genet. Popul.* **2**, 290–344 (1969).
39. De Meeüs, T. Revisiting FIS, FST, Wahlund effects, and null alleles. *J. Hered.* **109**, 446–456 (2018).
40. Grollemund, R. *et al.* Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 13296–13301 (2015).
41. Lavachery, P. The Holocene archaeological sequence of shum laka rock shelter (grassfields, western Cameroon). *African Archaeol. Rev.* **18**, 213–247 (2001).
42. Waits, L. P., Luikart, G. & Taberlet, P. Estimating the probability of identity among genotypes in natural populations: Cautions and guidelines. *Mol. Ecol.* **10**, 249–256 (2001).
43. Slatkin, M. Rare alleles as indicators of gene flow. *Evolution* **39**, 53–65 (1985).
44. Schwartz, M. K. & McKelvey, K. S. Why sampling scheme matters: The effect of sampling scheme on landscape genetic results. *Conserv. Genet.* **10**, 441–452 (2009).
45. Cameroon Voice. Cameroun: trafiquants de pangolin en Afrique centrale démasqués. <https://cameroonvoice.com/news/2021/07/22/cameroun-trafiquants-de-pangolin-en-afrique-centrale-demasesques/> (2021).
46. Dobson, A. D. M., Milner-Gulland, E. J., Ingram, D. J. & Keane, A. A Framework for assessing impacts of wild meat hunting practices in the Tropics. *Hum. Ecol.* **47**, 449–464 (2019).
47. Gabon Media Time. 216 pointes d'ivoire saisies au Cameroun provenant du Gabon: les réactions de l'ANPN et du gouvernement attendues. <https://www.gabonmediatime.com/216-pointes-divoire-saisies-cameroun-provenant-gabon-reactions-de-lanpn-gouvernement-attendues/> (2018).
48. Chakraborty, R. Sample size requirements for addressing the population genetic issues of forensic use of DNA typing. *Hum. Biol.* **141**–159 (1992).
49. Gossé, K. J., Gonedélé-Bi, S., Dufour, S., Danquah, E. & Gaubert, P. Conservation genetics of the white-bellied pangolin in West Africa: a story of lineage admixture, declining demography, and wide sourcing by urban bushmeat markets. *Ecol. Evol.* **14**, e11031 (2024).
50. Hall, T. Bioedit: a user friendly biological sequence alignment editor and analysis programme for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**, 95–98 (1999).
51. Kimura, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120 (1980).
52. Kearse, M. *et al.* Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012).
53. Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M. & Shipley, P. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* **4**, 535–538 (2004).
54. Peakall, R. & Smouse, P. E. GenAlEx 6.5: Genetic analysis in Excel. Population genetics software for teaching and research-an update. *Bioinformatics* **28**, 2537–2539 (2012).
55. Goudet, J. FSTAT ver. 2.9. 4, a program to estimate and test population genetics parameters. (2003).
56. Pritchard, J. K., Wen, X., Falush, D. Documentation for structure software: version 2.3. <http://pritch.bsd.uchicago.edu/structure.html> (2009).
57. Hubisz, M. J., Falush, D., Stephens, M. & Pritchard, J. K. Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Resour.* **9**, 1322–1332 (2009).
58. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **14**, 2611–2620 (2005).
59. Guillot, G., Mortier, F. & Estoup, A. Geneland: A computer package for landscape genetics. *Mol. Ecol. Notes* **5**, 712–715 (2005).
60. Coulon, A. *et al.* Genetic structure is influenced by landscape features: Empirical evidence from a roe deer population. *Mol. Biol. Rep.* **15**, 1669–1679 (2006).
61. Excoffier, L. & Lischer, H. E. L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 564–567 (2010).
62. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).
63. Kamvar, Z. N., Tabima, J. F. & Grunwald, N. J. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *Peer J.* **2**, e281 (2014).
64. Valière, N. gimlet: A computer program for analysing genetic individual identification data. *Mol. Ecol. Notes* **2**, 377–379 (2002).
65. Szpiech, Z. A., Jakobsson, M. & Rosenberg, N. A. ADZE: A rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics* **24**, 2498–2504 (2008).
66. Piry, S., Luikart, G. & Cornuet, J. M. Bottleneck: A program for detecting recent effective population size reductions from allele data frequencies. *J. Hered.* **90**, 502–503 (1999).
67. Nikolic, N. & Chevalet, C. Estimation of effective population size and variation in the time scale. 1–26 (2014).
68. Schlotterer, C. Evolutionary dynamics of microsatellite DNA. *Chromosoma* **109**, 365–371 (2000).

## Acknowledgements

We are grateful to the Ministère de la Recherche Scientifique et de l'Innovation for providing the research permit (000094/MINRESI/B00/C00/C10/nye). We thank the staff of the “Plateau technique–Biologie moléculaire et microbiologie” at EDB for assistance during lab work. We thank Franklin and Jean for their help during field work. The local hunters and vendors gently provided us with the necessary tissue samples by free consent. We are grateful to V. Benoît and L. Cervantes (Institut de Recherche Criminelle de la Gendarmerie Nationale, Cergy

Pontoise, France) for giving us access to the samples seized at Roissy airport (France). ADD, SA and PG received support from the Agence Nationale de la Recherche (ANR-17-CE02-0001; PANGO-GO). ADD and PG were also supported by the Fundação para a Ciência e Tecnologia (FCT IC&DT 02/SAICT/2017–n° 032130; BUSHRISK). ADD received a grant « Séjours scientifiques de haut niveau » (No. 968962A) from the Service de Coopération et d'Action Culturelle (SCAC) of the Ambassade de France in Cameroon. We thank three anonymous reviewers for greatly improving the initial version of the draft.

### Author contributions

PG, ADM, EL, KA and ADD conceived the ideas; ADD, ADM, FN, PG, ALC, SN and BRM collected the samples; ADD, SA and PG conducted the molecular laboratory work; ADD, SA and PG analyzed the data; ADD and PG led the writing; ADM, SA, EL, MT, FN, BRM, KA, SN and ALC read and revised the drafted work. All the co-authors have read and approved the final version of the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-63666-9>.

**Correspondence** and requests for materials should be addressed to A.D.D., A.D.M. or P.G.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024