



OPEN

## Population genetics and evolutionary history of the endangered Eld's deer (*Rucervus eldii*) with implications for planning species recovery

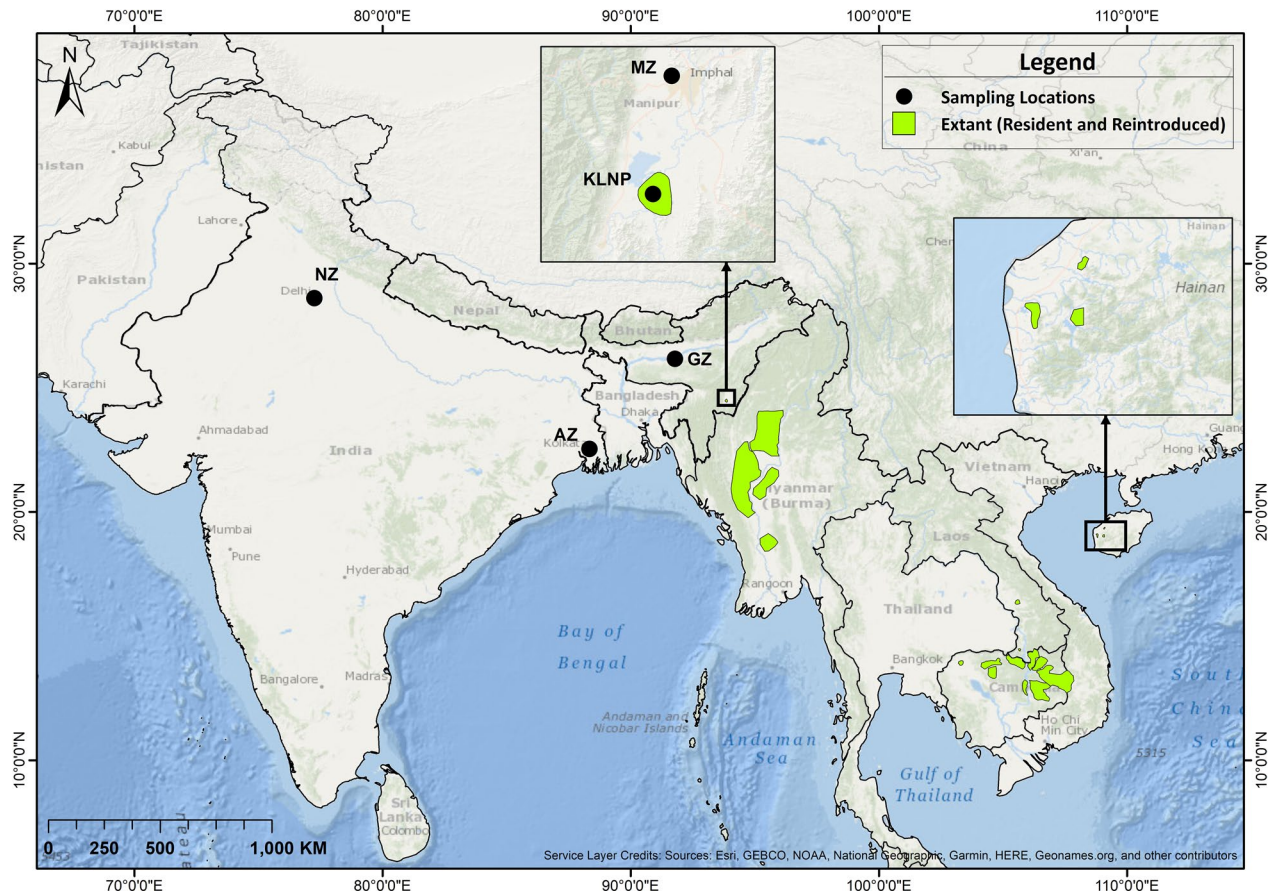
Mirza Ghazanfarullah Ghazi, Surya Prasad Sharma, Chongpi Tuboi, Sangeeta Angom, Tennison Gurumayum, Parag Nigam & Syed Ainul Hussain

Eld's deer (*Rucervus eldii*) with three recognised subspecies (*R. e. eldii*, *R. e. thamin*, and *R. e. siamensis*) represents one of the most threatened cervids found in Southeast Asia. The species has experienced considerable range contractions and local extinctions owing to habitat loss and fragmentation, hunting, and illegal trade across its distribution range over the last century. Understanding the patterns of genetic variation is crucial for planning effective conservation strategies. This study investigated the phylogeography, divergence events and systematics of Eld's deer subspecies using the largest mtDNA dataset compiled to date. We also analysed the genetic structure and demographic history of *R. e. eldii* using 19 microsatellite markers. Our results showed that *R. e. siamensis* exhibits two divergent mtDNA lineages (mainland and Hainan Island), which diverged around 0.2 Mya (95% HPD 0.1–0.2), possibly driven by the fluctuating sea levels of the Early Holocene period. The divergence between *R. e. eldii* and *R. e. siamensis* occurred around 0.4 Mya (95% HPD 0.3–0.5), potentially associated with the adaptations to warm and humid climate with open grassland vegetation that predominated the region. Furthermore, *R. e. eldii* exhibits low levels of genetic diversity and small contemporary effective population size (median = 7, 4.7–10.8 at 95% CI) with widespread historical genetic bottlenecks which accentuates its vulnerability to inbreeding and extinction. Based on the observed significant evolutionary and systematic distance between Eld's deer and other species of the genus *Rucervus*, we propose to classify Eld's deer (*Cervus eldii*) in the genus *Cervus*, which is in congruent with previous phylogenetic studies. This study provides important conservation implications required to direct the ongoing population recovery programs and planning future conservation strategies.

Small populations are susceptible to loss of genetic diversity as a consequence of founder events, genetic bottlenecks, and inbreeding, accelerated by the cumulative impacts of natural and anthropogenic factors<sup>1</sup>. Species characterised by fragmented and isolated populations are increasingly prone to factors that can lead to reduced levels of genetic diversity and loss of evolutionary potential which, in turn, limit the populations' ability to adapt to the changing environment<sup>2,3</sup> and as a consequence, elevate the risk of local extinctions<sup>4,5</sup>. Therefore, preservation and maintenance of adequate amount of genetic diversity in natural populations of threatened species is crucial for their long-term survival in the wild<sup>6</sup>.

Eld's deer or brow-antlered deer (*Rucervus eldii*) is a localised and endangered tropical Southeast Asian cervid<sup>7</sup>. It was once widely distributed across South and Southeast Asia, extending from Manipur in Northeast India to Myanmar, Thailand, Cambodia, Vietnam up to Hainan Island in China<sup>8,9</sup>. Over the past 200 years of known history, the species has witnessed severe range contraction across its historical distribution range. Over-exploitation of natural resources, habitat loss and fragmentation, hunting, and illegal trade are the major cause of its population decline<sup>7,10</sup>. The species at present is listed as endangered by the IUCN Red List<sup>7</sup>. It is classified into three subspecies; Manipur's brow-antlered deer (*R. e. eldii*), Burmese brow-antlered deer (*R. e. thamin*),

Wildlife Institute of India, Chandrabani, Post Box #18, Dehra Dun, Uttarakhand 248002, India. email: hussain@wii.gov.in

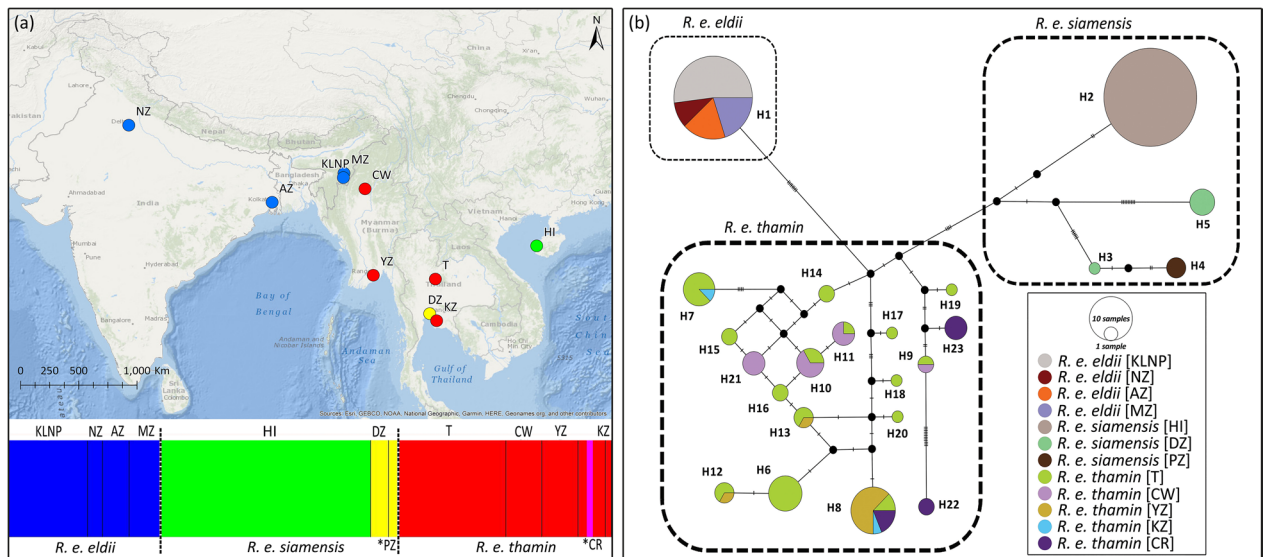


**Figure 1.** Map showing the geographical distribution of Eld's deer (*Rucervus eldii*) obtained from the IUCN Red List<sup>7</sup>. Sampling locations are represented as NZ National Zoological Park, New Delhi; AZ Alipore Zoological Garden, West Bengal; GZ Assam State Zoo, Assam; MZ Manipur Zoological Garden, Manipur; KLNP Keibul Lamjao National Park, Manipur. The map was created using ArcGIS v.10.3.1 software developed by ESRI (<https://www.esri.com>).

and Siamese Eld's deer (*R. e. siamensis*)<sup>7</sup>. However, the Hainan Island (HI) population of *R. e. siamensis* is, at times, considered to be the fourth subspecies of Eld's deer<sup>11,12</sup>. The current distribution of Eld's deer subspecies is highly fragmented (Fig. 1), with several regional extirpations from its former distribution range<sup>7</sup>. The Manipur's brow-antlered deer (*R. e. eldii*) was declared extinct in 1951, and the current wild population in Keibul Lamjao National Park (KLNP) in Manipur, India has recovered from 14 individuals rediscovered in 1975<sup>13</sup>. It represents the most threatened subspecies of Eld's deer which is highly susceptible to local extinction compared to *R. e. thamin* and *R. e. siamensis*, due to its small population size and restricted habitat in KLNP<sup>14</sup>. Despite its vulnerability to inbreeding and extinction due to stochastic and demographic factors, limited information is available on the genetic status of this subspecies to assist its population management programmes. *R. e. siamensis* underwent a significant decline across its distribution range during the last century and was declared extinct from Thailand and Vietnam. Presently, the subspecies occurs in scattered populations in Lao PDR and Cambodia and Hainan Island in China<sup>7</sup>. *R. e. thamin* has also witnessed severe range contractions and is considered extinct from Thailand. Presently, this subspecies represents the largest population of Eld's deer in the wild with localised populations in Myanmar and introduced populations in Thailand<sup>7,15</sup>.

Thailand is considered as the historical centre of the geographical distribution range of Eld's deer<sup>16</sup>. It is also an area of biogeographical importance as it facilitated the historical migration of large mammals between South China and Java. The oldest fossil record of Eld's deer is known from the Quaternary reddish clay deposits of Middle Pleistocene in north-eastern Thailand<sup>16,17</sup>. Another fossil form of Eld's deer was discovered in the Island of Java, Indonesia, dating back to around 3000 years before present (Ybp)<sup>18</sup> indicating the historical distribution of Eld's deer in this region which is presently isolated due to the rise in eustatic sea levels. Evidence of faunal compressions and local extinctions due to the increased sea levels has been reported in the Isthmus of Kra region, which divides the Indochinese and Sundaic sub-regions<sup>18,19</sup>. These changes together with dynamic climatological and vegetational conditions facilitated the diversification of many ruminant species<sup>20–23</sup> and would have influenced the patterns of genetic diversity, divergence and evolutionary history of Eld's deer subspecies.

A taxonomic conundrum of assigning Eld's deer to either genus *Cervus* or *Rucervus* still persists. Eld's deer share morphological similarities with swamp deer (*Rucervus duvaucelii*) and the extinct Schomburgk's deer (*R. schomburgki*)<sup>24</sup>. Based on these affinities, all three species are currently classified together in the genus



**Figure 2.** (a) Bayesian analysis of population structure (BAPS) of Eld's deer (*Rucervus eldii*) subspecies identified four genetic clusters using 200 mitochondrial control region sequences. Clusters are represented by colours. The dashed black lines separate subspecies and solid black lines separate populations. The map was created using ArcGIS v.10.3.1 software developed by ESRI (<https://www.esri.com>). (b) Median-joining haplotype network for 23 control region haplotypes identified among the three Eld's deer subspecies. Three major clusters were recovered and are indicated by subspecies names. Circle size represents the frequency of individuals, and colour defines the populations. Small black circles represent median vectors. Mutational steps are shown by bar on the respective branches. The populations are abbreviated as KLNP Keibul Lamjao National Park, Manipur; NZ National Zoological Park, New Delhi; AZ Alipore Zoological Garden, West Bengal; MZ Manipur Zoological Garden, Manipur; HI Hainan Island, China; DZ Dusit Zoo, Thailand; \*PZ Paris Zoo, France (not shown in map); T Thailand; CW Chatthin Wildlife Sanctuary, Myanmar; YZ Yangon Zoo, Myanmar; KZ Khao Khew Open Zoo, Thailand; \*CR Conservation & Research Centre, USA (not shown in map).

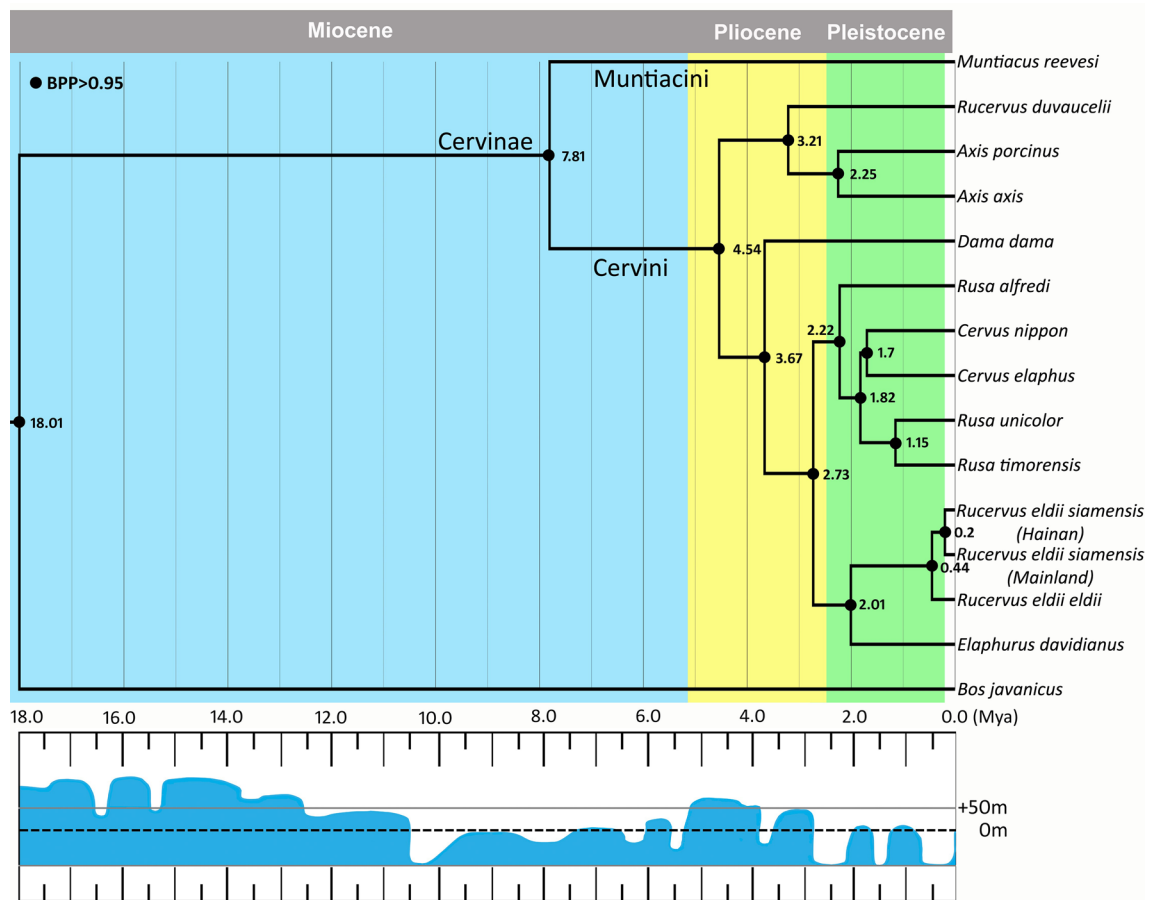
*Rucervus*<sup>24,25</sup>. However, previous molecular studies assigned Eld's deer to the genus *Cervus*<sup>26–29</sup> but Grubb<sup>8</sup> revived Thomas's<sup>30</sup> classification and assigned the species to genus *Rucervus* which is currently adopted by the IUCN Species Survival Commission. Despite the species being globally endangered, few studies have attempted to assess its genetic status<sup>12,29,31</sup>. Studies on microsatellite genetic variation are limited to the HI population of *R. e. siamensis*<sup>11</sup>. Lack of adequate information on the spatial genetic structure and levels of genetic variability among different populations of Eld's deer subspecies is an impediment for planning effective conservation breeding and population management strategies.

In this study, we integrated the available mitochondrial DNA (mtDNA) control region dataset to assess the patterns of genetic diversity, phylogeography and phylogenetic relationships among the Eld's deer subspecies. We analysed the complete mitochondrial genomes of *R. e. eldii* and *R. e. siamensis* to estimate the age of these subspecies and the major divergence events within the Eld's deer group. To resolve the taxonomic dilemma of assigning Eld's deer to either genus *Rucervus* or *Cervus*, we reconstructed the phylogenetic association of species of tribe *Cervini* using Bayesian analysis of complete mitochondrial genomes. We also estimated the microsatellite genetic diversity, contemporary and ancestral effective population size ( $N_e$ ), and demographic history of *R. e. eldii* using microsatellite markers. Due to logistical and legal constraints of range-wide biological sampling of Eld's deer subspecies, we confined the microsatellite analyses to *R. e. eldii*. The findings of this study will provide insights into the patterns of genetic association among the subspecies with a deeper understanding of the evolutionary history of Eld's deer, which is essential for planning conservation and management strategies.

## Results

**Phylogeography and genetic differentiation.** A total of 23 haplotypes were observed in the control region (447 bp) of three subspecies of Eld's deer segregated by 126 polymorphic sites with an overall haplotype diversity; (Mean  $\pm$  SD)  $hd = 0.805 \pm 0.020$  and low nucleotide diversity;  $Pi = 0.023 \pm 0.0005$ . All the three subspecies of Eld's deer contain unique control region haplotypes localized within the populations of the respective subspecies (Fig. 2b). The captive and wild populations ( $n = 71$ ) of *R. e. thamin* contain 18 haplotypes ( $hd = 0.908 \pm 0.018$ ) out of the total 23 haplotypes observed in all the three subspecies. *R. e. siamensis* exhibited four control region haplotypes ( $n = 9$ ) with  $hd = 0.639 \pm 0.126$ . Interestingly, we observed no haplotype sharing among the populations of *R. e. siamensis*. All the identified haplotypes in *R. e. siamensis* were unique to their respective populations with one haplotype (H2) observed in the Hainan Island, two haplotypes (H3 and H5) in Dusit Zoo, Thailand, and one haplotype (H4) in Paris Zoo, France. All the unique individuals of the captive ( $n = 23$ ) and wild populations of *R. e. eldii* ( $n = 27$ ) identified using 19 microsatellite loci shared a single control region haplotype.





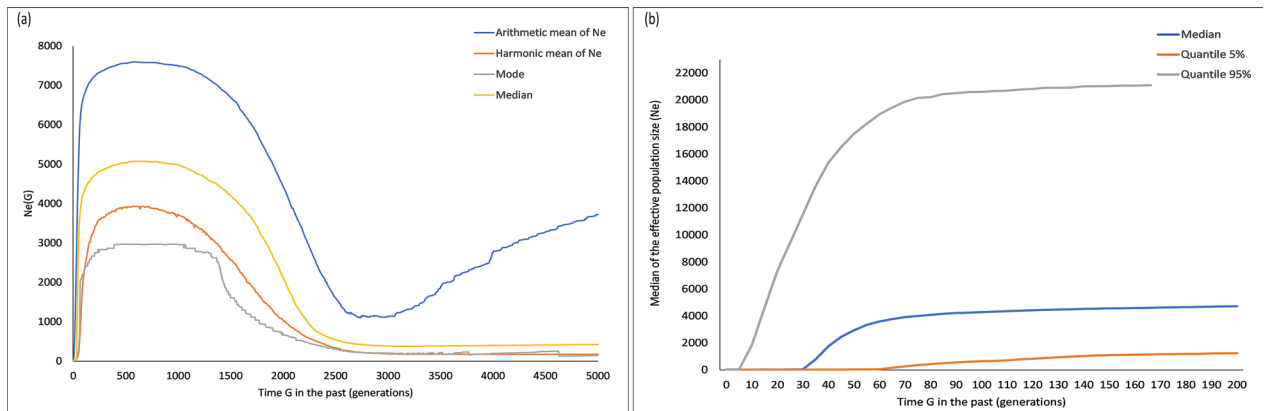
**Figure 4.** Time-calibrated phylogenetic tree of the tribe *Cervini* extending from 18 Mya to present based on Bayesian analysis of mitochondrial genomes. Approximate estimates of the eustatic sea level changes are presented under the chronogram<sup>35</sup>. The current sea levels are represented by the dashed line (0 m). The geological time scale was acquired from the Geological Society of America.

assigned two mitogenomes to the HI population (NC014701 and HM138200) and one to the mainland population based on the similarity indices of the blast search and estimated the divergence between these populations of *R. e. siamensis*. The obtained value for the index of substitution saturation (*I<sub>ss</sub>*)<sup>32</sup> for the analysed mitogenome dataset was 0.4093 and the critical *I<sub>ss.c</sub>* values for the symmetrical and asymmetrical trees were 0.8424 ( $P < 0.001$ ) and 0.6743 ( $P < 0.001$ ), respectively. Both the values of *I<sub>ss.c</sub>* were significantly higher than the observed *I<sub>ss</sub>* values. Therefore, the complete mitogenome dataset was found suitable for phylogenetic and divergence analysis.

**Divergence times.** The divergence time estimation based on Bayesian analysis of 13 PCGs, two rRNA genes (12S and 16S), and control region sequences obtained a strongly supported time-calibrated tree, which is consistent with the previous phylogenetic studies<sup>26,27</sup> (Fig. 4). Approximations of time to the most recent common ancestor ( $T_{\text{MRCA}}$ ) with 95% Highest Posterior Densities (HPD) is provided in Supplementary Table S2 online. The details of parameter convergence of divergence analyses are provided in Supplementary Fig. S2 online. The estimates of molecular divergence revealed simultaneous diversification of two evolutionary lineages within the tribe *Cervini*. We recovered an early divergence of common ancestors of the genera *Axis* and *Rucervus* ~ 4.5 million years ago (Mya) (95% HPD 3.6–5.4) from the main phylogenetic stock of tribe *Cervini*. The split between Eld's deer and the *Cervus*-*Rusa* group occurred around 2.7 Mya (95% HPD 2.1–3.3) during the Late Pliocene. Subsequently, the diversification of *Cervus* and *Rusa* species took place around 1.8 Mya (95% HPD 1.4–2.2). However, early separation of Philippine spotted deer (*Rusa alfredi*) from the common ancestors of the *Cervus*/*Rusa* group around 2.2 Mya (95% HPD 1.7–2.7) indicates that *R. alfredi* is the oldest form of living rutine deer.

Simultaneously, the divergence of *Elaphurus davidianus* and *R. eldii* took place ~ 2.0 Mya (95% HPD 1.5–2.5) during the Early Pleistocene period. Within *R. eldii*, we obtained two diversification events. The divergence of *R. e. eldii* from *R. e. siamensis* which occurred ~ 0.4 Mya (95% HPD 0.3–0.5), followed by the divergence of HI and the mainland populations of *R. e. siamensis* ~ 0.2 Mya (95% HPD 0.1–0.2).

**Phylogenetic analysis.** Phylogenetic reconstruction of tribe *Cervini* based on 13 PCGs, two rRNA genes and control region sequences using the Bayesian approach yielded a statistically supported phylogenetic tree of tribe *Cervini* (see Supplementary Fig. S3 online). *Bos javanicus* was nested at the root of the phylogenetic tree as



**Figure 5.** Estimates of effective population size ( $N_e$ ) of *R. e. eldii* in KLNP based on 19 microsatellite loci using approximate likelihood MCMC approach in VarEff v.1.2 package in R (a)  $N_e$  estimates from present to 5000 generations in the past and (b) present to 200 generations in the past.

an out-group. Within tribe *Cervini*, we recovered the genera *Axis*, *Cervus*, and *Dama* to be monophyletic (Bayesian Posterior Probability, BPP = 1). Fallow deer (*Dama dama*) and Persian fallow deer (*Dama mesopotamica*), chital (*Axis axis*) and hog deer (*Axis porcinus*) were allied together as sister species in their respective clades. Within the genus *Cervus*, all three species; red deer (*Cervus elaphus*), sika deer (*Cervus nippon*) and white-lipped deer (*Cervus albirostris*) were monophyletic (BPP = 1). However, the genus *Rucervus* and *Rusa* were observed to be paraphyletic. *R. duvaucelii* was nested with the genus *Axis* (BPP = 1) while *R. eldii* was associated with Père David's deer (*Elaphurus davidianus*) as a sister group. The monophyly of genus *Rusa* was disrupted by the disjunct position of *R. alfredi*, which was allied to the clade comprised of species of the genus *Cervus* and *Rusa* (*R. unicolor/R. timorensis*).

**Microsatellite genetic variation in *R. e. eldii*.** We successfully genotyped 66 out of total 79 faecal pellet samples, with 48 samples from KLNP and 18 samples from captive populations including Manipur Zoological Garden (n = 6), Assam State Zoo (n = 8) and National Zoological Garden (n = 4). The remaining samples from Manipur Zoological Garden (n = 4) and Alipore Zoological Garden (n = 9) failed to amplify, therefore excluded from the microsatellite analysis (see Supplementary Table S3 online). Based on multi-locus match, we identified 24 unique individuals out of 48 samples genotyped from KLNP and 18 individuals from 18 samples from captive populations. Therefore, a total of 42 samples (Wild = 24 and Captive = 18) were further used to perform population genetic analyses. We obtained low levels of microsatellite genetic variation in captive populations (Mean  $\pm$  SE) ( $H_o = 0.13 \pm 0.04$ ;  $H_e = 0.31 \pm 0.05$ ) compared to the wild population of *R. e. eldii* ( $H_o = 0.31 \pm 0.05$ ;  $H_e = 0.47 \pm 0.03$ ) in KLNP (see Supplementary Table S4 online). The mean allelic richness for the wild population was  $2.68 \pm 0.15$  and for the captive population was  $2.11 \pm 0.21$ . All the 19 microsatellite loci were polymorphic for the wild population of KLNP and obtained a total of 51 alleles. However, four loci (RT1, DF/R, BM4107, and AF232760) were monomorphic in the captive populations with a total of 40 alleles observed across 19 microsatellite loci. Most of the loci showed significant deviation from the Hardy–Weinberg equilibrium (HWE), which can be attributed to heterozygosity excess and inbreeding. The Bayesian clustering analysis using STRUCTURE revealed the presence of two genetic clusters based on  $\Delta K$  statistics (see Supplementary Fig. S4 online). The KLNP individuals were assigned to the first cluster whereas all the individuals of captive origin were assigned to the second cluster (see Supplementary Fig. S5 online). The genetic differentiation estimates between the captive and wild populations of *R. e. eldii* obtained low to moderate levels of differentiation based on the fixation measures ( $F_{st} = 0.26$ ;  $P < 0.001$ );  $G_{st} = 0.15$  ( $P < 0.001$ ) and allelic differentiation (Jost's  $D = 0.12$ ;  $P < 0.001$ ). The estimates of the inbreeding coefficient ( $F$ ) were  $0.342 \pm 0.100$  and  $0.566 \pm 0.098$  in the wild and captive populations, respectively.

**Past population changes in *R. e. eldii*.** The bottleneck detection using heterozygosity excess showed strong signature of genetic bottleneck. The one-tailed Wilcoxon test for heterozygosity excess was significant for all the three mutation models ( $P < 0.001$ ), viz. Infinite Allele Model (IAM); Two-phase Model (TPM) and Stepwise Mutation Model (SMM) (see Supplementary Table S5 online). The allelic frequency distribution also recovered a mode-shift curve, which confirms that the wild population of *R. e. eldii* has gone through a genetic bottleneck.

Estimates of past effective population size based on the approximate likelihood Markov chain Monte Carlo (MCMC) method obtained low  $N_e$  estimates until  $\sim 3000$ – $5000$  generations ago (median = 383–420) with a stable population trend until  $\sim 3000$  generations (Fig. 5a). Subsequently, the population showed a massive increase in the effective population size for  $\sim 3000$  generations until the recent bottleneck event was detected at  $\sim 300$  generations ago (median = 4876) (Fig. 5a). The ancestral effective population size as inferred from the estimates of the posterior distribution revealed a large population size during the past  $\sim 300$ – $1000$  generations (median 4281–4978) which corresponds to around 1500–5000 Ybp, considering the generation time of *R. e. eldii* as 5 years<sup>7</sup>. The posterior distribution of time to the most recent common ancestor ( $T_{MRC A}$ ) supported the

demographic estimates with the first peak observed at ~ 50 generations and a second peak at ~ 3000 generations, indicating the time at which the population underwent genetic bottlenecks (see Supplementary Fig. S6 online).

Assessment of the contemporary effective population size of wild population of *R. e. eldii* showed that the population underwent a steady decline during the past ~ 100–200 generations (median 4281–4717) (Fig. 5b) around 500–1000 Ybp. During the last ~ 30–100 generations, the decline accelerated and the population experienced a sudden, steep and continuous drop in the population size (median 45–4281), which accounts for a substantial decline in the effective population size of *R. e. eldii* during this period. However, the estimates of contemporary effective population size based on the linkage disequilibrium method were comparatively lower (median = 7.5, 95% CI 4.7–10.8) than those obtained using the MCMC approximation method (median = 20) (see Supplementary Table S6 online).

## Discussion

Demographic, environmental, and genetic factors profoundly affect the viability of populations<sup>33</sup>. Genetic diversity has potential impacts on the persistence and adaptation of populations to environmental changes<sup>34</sup>. Small populations generally exhibit low genetic variation due to population bottleneck and founder effects induced by natural as well as anthropogenic factors<sup>1</sup>. Eld's deer occur in small and fragmented populations, and its management largely depends on human-assisted maintenance of genetic diversity through population reintroductions. Hence, information on the genetic variability of Eld's deer populations across its distribution range is of prime importance for planning conservation strategies. This study examines the most extensive mtDNA control region dataset available to date, to provide information on the levels of genetic diversity, phylogeographic and phylogenetic patterns of Eld's deer subspecies.

Additionally, this is the first attempt to estimate the divergence events among the Eld's deer subspecies using complete mitochondrial genomes. Due to the unavailability of biological samples from other subspecies, we restricted the microsatellite analyses to *R. e. eldii* and delivered a comprehensive insight into the demographic history and past effective population size of this subspecies. The results of this study, combined with low population densities and increasing anthropogenic threats, raise serious conservation concerns and indicate increased susceptibility of Eld's deer to genetic, demographic, and stochastic factors.

**Phylogeography, divergence times and phylogenetics of *R. eldii*.** The phylogeographic analyses based on the control region (447 bp) indicated the presence of genetic structure among the Eld's deer subspecies. Consistent with the wide geographical distribution, persistence in multiple populations, and comparatively stable demographic history, *R. e. thamin* exhibited the highest number of haplotypes (18) followed by *R. e. siamensis* (4) and *R. e. eldii* (1) which occurs in small and isolated populations (Fig. 2b). We observed no segregation of haplotypes based on sampling locations in *R. e. thamin* and *R. e. eldii*. However, in *R. e. siamensis* the haplotypes were unique to their respective analysed populations. Further, the captive and wild populations of *R. e. eldii* showed complete homogeneity in 447 bp of the control region. Contrary to previous phylogenetic studies<sup>29,31</sup>, our phylogenetic analysis did not support the monophyly of Eld's deer subspecies (Fig. 3). Instead, we observed a close phylogenetic association between *R. e. eldii* and *R. e. thamin*, which is consistent with the close geographical distribution of these two subspecies. Despite its phylogenetic affinity to *R. e. thamin*, *R. e. eldii* possesses splayed hooves, a unique morphological adaptation for marshy habitats<sup>24</sup>. The presence of unique control region haplotypes and the observed patterns of phylogeographic structuring (Fig. 2) indicate a substantial amount of genetic differentiation between the mainland and HI populations of *R. e. siamensis*. Phylogenetic reconstruction based on the control region also supported the reciprocal monophyly of HI and mainland populations of *R. e. siamensis* (Fig. 3). Additionally, the genetic differentiation measures (Fst) and nucleotide sequence variation indicated significant differentiation between these isolated populations of *R. e. siamensis* (see Supplementary Table S1 online).

Divergence analysis based on mitochondrial genomes also supported the differentiation patterns observed between the HI and mainland populations of *R. e. siamensis*. This subspecies exhibits two divergent mitochondrial lineages corresponding to the HI and mainland populations, which diverged during the Late Pleistocene ~ 0.2 Mya (95% HPD 0.1–0.2) (Fig. 4). The low sea levels during this period<sup>35</sup> resulted in the emergence of land bridges, which would have facilitated the dispersal of *R. e. siamensis* from mainland China to Hainan Island. A rise in the eustatic sea levels during Holocene<sup>35</sup> interrupted the terrestrial connectivity and would have resulted in isolating the Hainan Island from the mainland populations of *R. e. siamensis* (Fig. 4). During the periods of low sea levels of Plio-Pleistocene, the emergence of land bridges comprising of tropical forests and savannah corridors facilitated faunal dispersion between Indochina, Southeast Asia and the Sundaic islands<sup>36</sup> possibly through the Siva-Malayan, Sino-Malayan and Taiwan-Philippine Archipelago corridors<sup>37</sup>. These events are supported by the patterns of genetic differentiation observed in this study between the isolated HI and mainland populations of *R. e. siamensis*. Therefore, intermixing the HI and mainland populations of *R. e. siamensis* to ameliorate the levels of genetic diversity is not a pragmatic proposal; instead, increasing the population size and establishment of meta-populations in the wild would practically benefit its long-term survival in the wild.

The divergence between *R. e. eldii* and *R. e. siamensis* took place around 0.4 Mya (95% HPD 0.3–0.5). During the Late Pleistocene and Early Holocene period, Northeast India was dominated by open grassland habitats<sup>38,39</sup>, which would have facilitated the dispersal of grassland dependent species during the period. The dynamic influences of climatic and environmental factors during the Pleistocene have significantly driven the divergence and phylogeographic structure of several species<sup>40</sup>. We hypothesise that different ecological niches occupied by the Eld's deer subspecies might have facilitated their diversification. Overall, the mitochondrial genomes investigated in this study revealed two major divergence events among the Eld's deer subspecies during Late Pleistocene.

However, one limitation of our study is the unavailability of the complete mitochondrial genome of *R. e. thamin*, which would have provided more profound insights into the evolutionary history of Eld's deer.

Further, our results indicated diversification of two evolutionary lineages in tribe Cervini which occurred around 4.5 Mya (95% HPD 3.6–5.4) during Pliocene. The first lineage gave rise to species of the genus *Rucervus* and *Axis*, and the latter radiated into the genera *Rusa*, *Cervus*, *Dama*, and *Elaphurus* including the Eld's deer (Fig. 4). Eld's deer is presently classified in the genus *Rucervus* based on the morphological similarities with *R. duvaucelii* and the now-extinct *R. schomburgki*<sup>24</sup>. However, these morphological similarities possibly characterise convergence rather than phylogenetic relatedness<sup>41</sup>. Moreover, our divergence and phylogenetic analyses revealed an ancestral divergence of *Axis* and *Rucervus* sp. from the main phylogenetic stock of tribe Cervini (Fig. 4 and Supplementary Fig. S3 online). Therefore, based on the significant evolutionary and systematic distance of Eld's deer from other species of the genus *Rucervus*, we propose to assign Eld's deer to genus *Cervus* as *Cervus eldii* together with *E. davidianus* and all four species of genus *Rusa* (*R. unicolor*, *R. timorensis*, *R. alfredi*, and *R. marianna*) (see Supplementary Fig. S3 online). These recommendations are in agreement with the previous molecular phylogenetic studies<sup>26,27</sup>. The classification proposed in this study will resolve the taxonomic perplexity of Eld's deer and other species of tribe Cervini to its phylogenetically most appropriate genus.

**Microsatellite genetic variation in *R. e. eldii*.** The microsatellite analyses revealed low levels of genetic diversity in both captive and wild populations of *R. e. eldii* based on 19 microsatellite loci (see Supplementary Table S4 online). However, the wild population exhibits relatively higher levels of microsatellite genetic diversity compared to the captive populations, which is consistent with KLNP being the source population of all the captive stock. The population structure analysis using microsatellite revealed the presence of two genetic clusters ( $K=2$ ) and assigned the captive and wild populations to separate genetic clusters (see Supplementary Fig. S5 online). The captive population of *R. e. eldii* was established at Alipore Zoological Garden, West Bengal in 1956, by acquiring a pair of wild individuals from KLNP. All the existing populations in captivity are derived from the captive stock of Alipore Zoological Garden and have never been supplemented with individuals from the wild<sup>42</sup>. Hence, the assignment of all the captive individuals in a single cluster supports the evidence that all the captive populations are derived from the initial captive stock of Alipore Zoo. Further, no gene flow between the wild and captive populations and a small number of founder individuals of captive stock would have resulted in the observed patterns of genetic distinctiveness between the captive and wild populations of *R. e. eldii*.

The populations of *R. e. siamensis* in Datian, Bangxi, and Ganshiling Nature Reserves of Hainan Island were reported to exhibit low mean heterozygosity and genetic polymorphism estimates with  $H_o = 0.34 \pm 0.03$ ,  $0.33 \pm 0.01$  and  $0.36 \pm 0.03$  and  $H_e = 0.28 \pm 0.01$ ,  $0.32 \pm 0.00$  and  $0.36 \pm 0.03$ , respectively<sup>11</sup>. The observed  $H_o$  and  $H_e$  for *R. e. eldii* in KLNP were estimated to be  $H_o = 0.31 \pm 0.05$  and  $H_e = 0.471 \pm 0.033$  (see Supplementary Table S4 online) which indicates that *R. e. eldii* in KLNP and *R. e. siamensis* in Hainan Island exhibit similar levels of microsatellite diversity. Both these subspecies have experienced severe range contraction during the last century and recovered from a small number of founder individuals<sup>7</sup>, therefore exhibit low levels of genetic polymorphism. Moreover, a recent study<sup>43</sup> based on similar microsatellite markers estimated the mean heterozygosity ( $H_o = 0.42 \pm 0.02$  and  $H_e = 0.51 \pm 0.03$ ) of *A. porcinus*, which co-occurs with *R. e. eldii* in KLNP sympatrically. Both *A. porcinus* and *R. e. eldii* exhibit low to moderate levels of microsatellite genetic variability, which can be attributed to their single, small and isolated populations, inhabiting the space-constrained habitat in KLNP. Most of the selected microsatellite loci significantly deviated from the HWE, possibly due to the small population size, non-random mating, and impacts of inbreeding<sup>44</sup> in *R. e. eldii*. In summary, the existing wild and captive populations of *R. e. eldii* exhibit low levels of microsatellite genetic diversity. The wild population in KLNP is highly susceptible to inbreeding because of the increased probability of interbreeding of related individuals due to its small and geographically isolated population.

**Past population changes in *R. e. eldii*.** The low estimates of effective population size are congruent with the genetic bottleneck events detected in this study. The one-tailed Wilcoxon test for heterozygosity excess and allelic frequency distribution indicated that the wild population of *R. e. eldii* has gone through genetic bottleneck (see Supplementary Table S5 online). Posterior distribution of time to the most recent common ancestor ( $T_{MRC(A)}$ ) indicated two genetic bottleneck events (see Supplementary Fig. S6 online). The first event occurred ~ 3000 generations ago, which resulted in a mild fluctuation in the number of individuals (Fig. 5a). Subsequently, a massive upsurge in the effective population size was observed, indicating a range expansion event. Based on the generation time of Eld's deer (5 years)<sup>7</sup>, we assume that this event took place ~ 15,000 Ybp.

During the Late Pleistocene–Early Holocene, Northeast India had experienced dynamic fluctuations in climatic conditions, significantly affecting the past vegetation structure of the region<sup>45–47</sup>. The existing population of *R. e. eldii* inhabits the floating meadows of KLNP situated in the southern fringes of Loktak Lake in Manipur. The existing lake is the remnant of the original lake, which once occupied almost the entire Manipur valley<sup>39</sup> and is known to have originated around 25,000 Ybp<sup>48</sup>. Since the course of its formation, the Loktak Lake underwent a prolonged phase of hydrodynamic and morphological changes<sup>48</sup>. Therefore, we presume that the lake during this period would have remained uninhabitable for *R. e. eldii*. The low estimates of effective population size during ~ 3000–5000 generations ago (median 383–420) around 15,000–25,000 Ybp are therefore consistent with the biogeographical history of the region. Subsequently, the appearance of open grasslands with aquatic herbaceous associates in the Loktak Lake region during Late Holocene<sup>38,39</sup> indicates the emergence of extensive habitat conditions favourable for grassland dependent species, such as *R. e. eldii*. As shown in our results, an increase in the ancestral population size during the past ~ 2500–300 generations (median 540–4876) (Fig. 5a), signifies a range expansion event which would have been triggered due to the increased habitat availability during this period. Range expansion events and genetic drift often have negative impacts on the genetic diversity of the



species<sup>49,50</sup>. Furthermore, our results indicated a second genetic bottleneck ~ 1500 Ybp (~ 300 generations ago; median = 4876), which caused a steady decline in the population size till past ~ 100 generations (median = 4281) (Fig. 5b). Subsequently, the population underwent a rapid reduction in the effective population size during the past ~ 35–100 generations (median 791–4281). During this period, the Manipur valley underwent a drastic climatic transition, and warm and dry conditions replaced the prevailing warm and humid climatic conditions. Consequently, a reduction in the Loktak Lake dimensions along with the associated aquatic vegetation<sup>38</sup> possibly reduced the habitat available for *R. e. eldii* which potentially resulted in a reduction in the effective population size. The intensification of human activities during the Late Holocene period is likely to accelerate the reduction and fragmentation of the suitable habitats of Eld's deer in the region. We hypothesise that the cumulative impacts of these factors would have influenced the demographic history of Eld's deer and consequently resulted in the reduction of genetic diversity.

**Implications for conservation and management.** The findings of this study provide significant implications for planning conservation of Eld's deer across its distribution range. Our results demonstrate that *R. e. siamensis* exhibits two divergent mitochondrial lineage, with long divergence history and substantial amount of genetic differentiation. However, due to sampling restrictions, we were unable to either confirm or invalidate the subspecies status of Hainan's Eld's deer. We recommend a combined assessment of genetic variability between these populations to confirm the patterns of mtDNA variation highlighted in this study. If confirmed, the HI populations might represent a separate subspecies of Eld's deer. Until then, we recommend to identify the HI and mainland populations of *R. e. siamensis* as distinct units for conservation and suggest to prevent any intermixing of these lineages. As proposed by Zhang et al.<sup>11</sup>, we support the reintroduction and recovery programs for *R. e. siamensis* within the historical distribution range throughout the Hainan Island.

Low levels of genetic diversity, past effective population size, and widespread genetic bottleneck events in *R. e. eldii* accentuate its vulnerability to loss of evolutionary potential due to resulting genetic erosion, therefore necessitating immediate conservation interventions. The contemporary estimates of the effective population size of *R. e. eldii* are low (median = 7.5, 4.7–10.8 at 95% CI), far below the minimum threshold required to prevent inbreeding depression ( $Ne > 100$ ) and to maintain long-term evolutionary potential ( $Ne > 1000$ ) as proposed by Frankham et al.<sup>51</sup>. The HI populations of *R. e. siamensis* also contain low genetic diversity<sup>11</sup>, but large population size, which will undoubtedly benefit its population recovery in the wild. Unlike the HI populations, *R. e. eldii* exists as a single isolated wild population in a space-constrained habitat in KLNP with low population abundance in the wild<sup>14</sup>. Therefore, *R. e. eldii* is exceedingly susceptible to extinction as any further reduction in the number of breeding individuals will have detrimental effects on the viability of the population. Given these factors, it is crucial that improved conservation efforts with overarching strategies are implemented with immediate objectives to maintain population abundance in the wild as well as restoration of genetic diversity. Based on the observed patterns of genetic diversity and population genetic structure of *R. e. eldii*, we recommend initiating the conservation breeding using individuals from both captive and wild populations. Subsequently, the captive-bred founder stock can be used to ameliorate the levels of genetic diversity of the wild population by restocking the existing population in KLNP and establishing alternative populations to ensure its long-term persistence in the wild.

Furthermore, high human population density and increasing resource dependence, together with encroachment and changing land-use patterns in and around KLNP and Loktak Lake are leading to overexploitation of natural resources, thereby affecting the ecological integrity of the Park<sup>14</sup>. Therefore, planning immediate conservation strategies to mitigate the prevalent anthropogenic pressures, and arrest any further change in the existing land-use regime will benefit in protecting the ecological integrity of the Loktak Lake including KLNP, thereby ensuring the long-term survival of *R. e. eldii* in the wild. Any further delay in mitigating these conservation challenges may lead to an irreversible damage to the genetic diversity of the species.

The long-term survival of Eld's deer in the wild largely relies on human-assisted maintenance of genetic diversity through population reintroductions, restocking, and conservation breeding. However, the lack of range-wide information on genetic variation raises apprehensions on the effectiveness of these strategies. The need for mapping the spatial genetic variation using an extensive set of mitochondrial and nuclear markers cannot be overstated, especially given the critical status of Eld's deer across its distribution range. The information generated in this study will aid in evaluating the efficacy of the ongoing population recovery programs adopted for Eld's deer subspecies. Moreover, a collaborative approach among various institutions and policymakers across Eld's deer range countries is crucial to support the conservation efforts for ensuring the long-term survival of this species in the wild.

## Materials and methods

**Sampling and DNA extraction.** We collected non-invasive biological samples of *R. e. eldii* including freshly defecated faecal pellets from Keibul Lamjao National Park (KLNP), Manipur (n = 48) and captive populations (n = 31) of Alipore Zoological Garden, West Bengal (n = 9), Manipur Zoological Garden, Manipur (n = 10), Assam State Zoo, Assam (n = 8) and National Zoological Park, New Delhi (n = 4) during 2006–2009 (see Supplementary Table S3 online). We also collected tissue samples (n = 3) from the remains of dead animals from KLNP during 2019 (see Supplementary Table S7 online) (Fig. 1). Tissue and faecal pellet samples were stored in absolute ethanol at –20 °C for DNA extraction. Total genomic DNA from tissue and faecal samples were extracted using QIAamp DNA Blood and QIAamp DNA Stool kit (QIAGEN) respectively following the manufacturer's protocols. Tissue samples were used to sequence the complete mitochondrial genomes and faecal pellets for sequencing the mtDNA control region and microsatellite analyses of *R. e. eldii*. The tissue samples were obtained from remains of naturally dead animals and therefore, did not require the approval of the animal ethical committee. The samples were collected with due permissions from the Principal Chief Conservator of

Forests and the Park administration of Manipur. All the field and laboratory experiments were executed under the pertinent guidelines.

**PCR and sequencing.** We sequenced the control region (447 bp) using *Cerv.tPro*: 5'-CCACYATCAACA CCCAAAGC-3' and *Cerv.CRH*: 5'-GCCCTGAARAAAGAACCAGATG-3'<sup>29</sup>. Sequencing of the mitochondrial genome was performed using 23 pairs of primers (see Supplementary Table S8 online) described by Hassanin et al.<sup>52</sup> with an overlapping region of 80–120 base pairs to remove any gaps among the amplicons. Due to the hyper-variable nature of the control region, bi-directional sequencing using both forward and reverse primers was performed. Polymerase Chain Reactions (PCR) for both control region and mitogenome amplifications were carried out in 20 µL volumes containing 10–20 ng of the total genomic DNA, 10 µL of 2× PCR buffer (Applied Biosystems), 0.2 mg/mL Bovine Serum Albumin (BSA), 0.3 µM dNTPs, 0.3 µM of forward and reverse primers with 0.1 µL (0.5 units) of DreamTaq DNA Polymerase (ThermoFisher Scientific). PCR conditions followed an initial denaturation at 94 °C for 5 min, followed by 35 cycles at 95 °C for 35 s, 55 °C for 45 s, and 72 °C for 1 min 30 s. Reactions were terminated with a final extension of 10 min at 72 °C. We used positive and negative controls to monitor the efficacy and reliability of the reactions. PCR products were visualised for positive amplicons using 2% agarose gel electrophoresis using ethidium bromide and were examined under UV illuminator. Positive amplicons of control region and 23 fragments of the mitochondrial genomes were sequenced using Sanger sequencing in ABI 3500XL Genetic Analyser using BigDye v.3.1 sequencing kit (Applied Biosystems).

**Microsatellite amplification and genotyping.** We used 19 microsatellite loci<sup>53–63</sup> (see Supplementary Table S4 online) for assessing the genetic variability in the wild (n = 48) and captive populations of *R. e. eldii* (n = 31) using non-invasively collected fresh faecal pellet samples. PCR was performed in 10 µL reaction volumes with 2.0–3.0 µL of total genomic DNA, 5 µL of QIAGEN Multiplex PCR Buffer Mix (QIAGEN) and 0.2 µM of each labelled forward and 0.2 µM of the non-labelled reverse primers. PCR cycles of the microsatellite loci were performed at an initial denaturation at 94 °C for 15 min followed by 40 cycles of 94 °C for 30 s with annealing temperatures ranging from 48 to 60 °C for 45 s, 72 °C for 40 s, and final elongation at 72 °C for 30 min. The products were then genotyped in ABI 3130 Genetic Analyser using LIZ 500 Size Standard (Applied Biosystems). Positive and negative controls were used to monitor the efficacy and reliability of the reactions.

## Data analyses

**Phylogeography and genetic differentiation.** We obtained a global alignment of 200 control region sequences (447 bp) representing all the three subspecies of Eld's deer (see Supplementary Table S3 online). The dataset was comprised of published control region sequences downloaded from GenBank (n = 150) along with 24 unique individuals identified in the wild population of *R. e. eldii* out of total 48 faecal pellet samples collected from KLNK and captive populations (n = 23) from Alipore Zoological Garden (n = 9) Manipur Zoological Garden (n = 10) and National Zoological Park (n = 4) using 19 microsatellite markers. We also included the control region sequences (n = 3) obtained from the complete mitochondrial genomes sequenced in this study. The faecal pellet samples obtained from Assam State Zoo (n = 8) were excluded from phylogeographic analysis because they failed to amplify at the control region. The control region sequences were aligned using the CLUSTAL W algorithm<sup>64</sup> in program MEGA X<sup>65</sup>. Haplotype (*hd*) and nucleotide diversities (*Pi*) were calculated using program DnaSP v.5.10<sup>66</sup>. Haplotypes were derived by removing non-informative sites, and gaps were considered. The spatial relationship among the haplotypes was reconstructed using the median-joining network in PopART v.1.7<sup>67</sup>. Bayesian Analysis of Population Structure (BAPS) v.6.0<sup>68,69</sup> was used to identify the number of genetic clusters. We performed the non-spatial clustering with admixture analysis using an assumed number of clusters (*K*) ranging from 1 to 30 with ten replicates for each value of *K* and 500 iterations. Pairwise genetic differentiation (*Fst*) of the identified clusters were calculated using program DnaSP v.5.10<sup>66</sup>. Phylogenetic relationships among the Eld's deer subspecies were reconstructed using the partial fragment of the control region based on the Bayesian approach as implemented in program MrBayes v.3.2<sup>70</sup>. The most appropriate model for nucleotide substitution was selected based on the Akaike Information Criterion (AIC)<sup>71</sup> values using program jModelTest v.2.1.3<sup>72,73</sup>. We performed two MCMC chains of 30 million simulations sampling at every 10,000 generations and discarded the initial 3 million runs as burn-in. The output tree topologies were edited using FigTree v.1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree/>).

**Divergence times and phylogenetic analyses.** The complete mitochondrial genome of 16,357 bp sequenced from three individuals of *R. e. eldii* were obtained by aligning the 23 overlapping sequence fragments using SeqScape software v.3.0 (Applied Biosystems). The alignments were visualised manually and Phred Quality Score of each nucleotide variation ( $Q \geq 20$ ) was selected as a threshold for considering the variable sites (<https://apps.thermofisher.com/apps/spa/#/apps>). The mitochondrial genome annotation was performed in CGView Server<sup>74</sup>. The complete mitochondrial sequences were submitted to GenBank (Accession No: MT555112–MT555114). Further, we included the published complete mitogenomes of 10 species of tribe *Cervini*, *Bos javanicus*, and *Muntiacus reevesi* obtained from GenBank to estimate the divergence events and reconstruct the accurate phylogenetic position of Eld's deer (see Supplementary Table S7 online). We used 13 PCGs, two rRNA genes and control region sequences to perform the divergence and phylogenetic analyses. All the analysed mitochondrial genes were aligned separately using CLUSTAL W algorithm<sup>64</sup> in program MEGA X<sup>65</sup>. Gaps were considered in the sequence alignments of each gene because of the variable lengths of PCGs and control region sequences among the analysed species. The aligned sequences of the mitochondrial genes were concatenated in MEGA X<sup>65</sup> to obtain the final alignment for performing further data analyses. The degree of substitution saturation was calculated using program DAMBE to test whether the observed entropy is significantly lower than the

entropy of full substitution saturation in the analysed sequences<sup>32,75</sup>. Divergence times were estimated using the Bayesian approach as implemented in program BEAST v.1.8.2<sup>76</sup>. PartitionFinder<sup>77</sup> was used for selecting the most appropriate data partitioning scheme of nucleotide substitution for evolutionary rate heterogeneity. Greedy algorithm with linked branch length along with Bayesian Information Criterion as implemented in PartitionFinder was used. Substitution models were selected with separate model estimation for 13 PCGs using different codon positions and individually for rRNA and control region sequences. We used fossil information of the oldest known fossils of (1) Cervidae (18.4 Mya)<sup>78</sup> for the Bovidae/Cervidae split and tribe *Muntiacini* (8 Mya)<sup>79</sup> as calibration points. We used the lognormal uncorrelated relaxed-clock model and Yule speciation process with normally distributed prior parameters. We performed two independent divergence analyses with 50 million generations, sampling at every 1000 generations with a 10% burn-ins.

Phylogenetic analysis to reconstruct the molecular phylogeny of tribe *Cervini* based on 13 PCGs, two rRNA and control region sequences was executed in MrBayes v.3.2<sup>70</sup>. The best suitable model for nucleotide substitution was obtained based on the lowest AIC values using jModelTest v.2.1.3<sup>72,73</sup>. We ran two MCMC chains of 50 million simulations sampling every 10,000 generations. We discarded the initial 5 million runs as burn-ins. Parameter convergence and effective sample size (> 400) for divergence and phylogenetic analyses were analysed in program TRACER v.1.6. TREEANNOTATOR v.1.8.2 (BEAST package) was used for annotating the output tree with a burn-in of 10%. The output tree topologies were visualised and edited in FigTree v.1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree/>).

**Microsatellite genetic variation.** We performed the identification of alleles using GeneMarker v.2.7.4<sup>80</sup>. The likelihood of two unrelated individuals and siblings sharing the same genotype (probability of identity,  $P_{ID}$  and probability of identity for siblings,  $P_{ID-Sibs}$ ) was executed in Gimlet v.1.3.3<sup>81</sup>. The overall summary statistics comprising of the total number of alleles per locus ( $N_a$ ), observed ( $H_o$ ) and expected ( $H_e$ ) heterozygosity, and deviations from the Hardy–Weinberg equilibrium (HWE) were calculated in GenAlEx v.6.0<sup>82</sup>. The Polymorphic Information Content (PIC) values were estimated in program CERVUS v.3.0.7<sup>83</sup>. We used model-based hierarchical Bayesian clustering approach implemented in program STRUCTURE v.2.3.4<sup>84</sup> to identify the presence of genetic structure in *R. e. eldii* populations. We used an assumed number of clusters ( $K$ ) ranging from 1 to 10. For each  $K$ , we executed ten iterations with a burn-in of 50,000 followed by 5,00,000 MCMC simulations under the admixture model with correlated allele frequencies. The optimal number of clusters in the sampled populations were determined following Evanno et al.<sup>85</sup> approach executed in the web version of STRUCTURE HARVESTER v.0.6.94<sup>86</sup>. The assignment probability plot using individual membership coefficient was produced using program DISTRUCT v.1.1<sup>87</sup>. Further, the genetic differentiation estimates  $F_{st}$ ,  $G_{st}$  and Jost's  $D$  among the identified clusters were estimated using program *strataG*<sup>88</sup> package in R with 1000 bootstrap iterations.

**Past effective population size.** To detect the signatures of bottleneck events, we performed the transient heterozygosity excess and mode-shift tests in program BOTTLENECK v.1.2.02<sup>89</sup> assuming that microsatellite mutations followed the Infinite Allele Model (IAM), Stepwise Mutation Model (SMM) or Two-phase Model (TPM). In TPM we used 95% single-step mutational events at 12% variance<sup>89</sup>. We also performed the one-tailed Wilcoxon test to determine the significance of the heterozygosity excess. The allelic distribution curve was analysed to determine whether the population is under mutation-drift equilibrium (L shaped curve) or not (mode shift)<sup>90,91</sup>.

We implemented the linkage disequilibrium method<sup>92–94</sup> to estimate the contemporary effective population size using the bias-corrected version in program NeEstimator v.2<sup>95</sup>. The linkage disequilibrium approach is the most suited method for microsatellite data which has been proven to be relatively robust in assessing the effective population size<sup>94</sup>. We also implemented the approximate likelihood MCMC approach using VarEff v.1.2 package in R (<https://qgsp.jouy.inra.fr/>)<sup>96</sup> to estimate the past changes in effective population size. This approach implements coalescence theory to generate posterior distributions of effective population size using an approximation of likelihood based on different *prior* demographic parameters. It implements the distribution of motif distance frequencies between alleles at the tested microsatellite loci<sup>96,97</sup>. This method allows recovering the posterior distribution of the time to the most recent common ancestor ( $T_{MRC A}$ ) between two alleles and provides the time of the bottleneck events in the number of generations in the past. We implemented a Two-phase Mutation Model (TPM) to estimate the past changes in the population size with proportion for multi-step mutations of  $c = 0.2$  and an average microsatellite mutation rate ( $\mu$ ) of  $2.0 \times 10^{-3}$  per generation<sup>98–103</sup>. We estimated the effective population size from present to 5000 generations in the past, assuming the generation time ( $G$ ) of 5 years for Eld's deer<sup>7</sup>. Priors for past effective population size were derived from global estimates of Theta ( $\theta = 4N_e\mu$ ) executing the Theta function in VarEff<sup>96</sup>. MCMC runs were performed using 10,000 batches with sampling at every ten batches and a burn-in of 10,000. We performed trial runs by changing the analysis parameters and finally selected the parameters recommended by Nikolic and Chevalet<sup>96</sup> (see Supplementary Table S9).

## Data availability

The sequence data generated in this study is submitted to the GenBank database, National Centre for Biotechnology Information under Accession Nos. MT555259–MT555301, MW033296–MW033299 and MT555112–MT555114. The genotyping data used in this study is available in the supplementary section.

Received: 17 July 2020; Accepted: 18 January 2021

Published online: 28 January 2021

## References

1. Banks, S. C. *et al.* How does ecological disturbance influence genetic diversity?. *Trends Ecol. Evol.* **28**, 670–679 (2013).
2. Coltman, D. W., Pilkington, J. G., Smith, J. A. & Pemberton, J. M. Parasite-mediated selection against inbred soay sheep in a free living island population. *Evolution* **53**, 1259–1267 (1999).
3. Hedrick, P. W. & Fredrickson, R. Genetic rescue guidelines with examples from Mexican wolves and Florida panthers. *Cons. Genet.* **11**, 615–626 (2010).
4. Frankham, R. Genetics and extinction. *Biol. Conserv.* **126**, 131–140 (2005).
5. Markert, J. A. *et al.* Population genetic diversity and fitness in multiple environments. *BMC. Evol. Biol.* **10**, 205 (2010).
6. Willoughby, J. R. *et al.* The reduction of genetic diversity in threatened vertebrates and new recommendations regarding IUCN conservation rankings. *Biol. Conserv.* **191**, 495–503 (2015).
7. Gray, T. N. E. *et al.* *Rucervus eldii*. The IUCN red list of threatened species. e.T4265A22166803 (2015). <https://dx.doi.org/10.2305/IUCN.UK.2015-2.RLTS.T4265A22166803.en>. Downloaded on 19 January 2020.
8. Grubb, P. Artiodactyla. In *Mammal Species of the World* (eds Wilson, D. E. & Reeder, D. M.) 637–722 (Johns Hopkins University Press, Baltimore, 2005).
9. Salter, R. E. & Sayer, J. A. The brow-antlered deer in Myanmar—Its distribution and status. *Oryx*. **20**, 241–245 (1986).
10. McShea, W. J., Leimgruber, P., Aung, M., Monfort, S. L. & Wemmer, C. Range collapse of a tropical cervid (*Cervus eldi*) and the extent of remaining habitat in central Myanmar. *Anim. Conserv.* **2**, 173–183 (1999).
11. Zhang, Q., Zeng, Z., Ji, Y., Zhang, D. & Song, Y. Microsatellite variation in China's Hainan Eld's deer (*Cervus eldi hainanus*) and implications for their conservation. *Cons. Genet.* **9**, 507–514 (2008).
12. Zhang, Q., Zeng, Z., Sun, L. & Song, Y. The origin and phylogenetics of Hainan Eld's deer and implications for Eld's deer conservation. *Acta. Ther. Sin.* **29**, 365–371 (2009).
13. Ranjitsinh, M. K. Keibul Lamjao Sanctuary and the Brow-antlered deer—1972 with notes on a visit in 1975. *J. Bom. Nat. His. Soc.* **72**, 243–255 (1975).
14. Hussain, S. A. & Badola, R. *Conservation Ecology of Sangai and Its Wetland Habitat. Study Report Vol. I* (Wildlife Institute of India, Dehra Dun, 2013).
15. McShea, W. J., Aung, M., Songer, M. & Connette, G. M. The challenges of protecting an endangered species in the developing world: A case history of Eld's Deer conservation in Myanmar. *Case Stud. Environ.* **2**, 1–9 (2018).
16. Ginsburg, L., Ingavat, R. & Sen, S. A Middle Pleistocene (Loagian) cave fauna in Northern Thailand. *Comptes Rendus de l'Académie des Sciences Paris.* **294**, 295–297 (1982).
17. Tougaard, C. Y., Chaimanee, V., Sutheethron, S. & Triamwichanon, Jaeger, J. J. Extension of the geographic distribution of the giant panda (*Ailuropoda*) and reasons for its progressive disappearance in Southeast Asia during the Latest Middle Pleistocene. *C. R. Acad. Sci. Paris.* **323**, 973–979 (1996).
18. Corbett, G. B. & Hill, J. E. *The Mammals of the Indomalay Region: A Systematic Review. Natural History Museum Publications* (Oxford University Press, Oxford, 1992).
19. Woodruff, D. S. & Turner, L. M. The Indochinese-Sundaic zoogeographic transition: A description and analysis of terrestrial mammal species distributions. *J. Biogeogr.* **36**, 803–821 (2009).
20. Hassanin, A. & Ropiquet, A. Molecular phylogeny of the tribe Bovini (Bovidae, Bovinae) and the taxonomic status of the kouprey, *Bos sauveli* Urbain, 1937. *Mol. Phylo. Evol.* **33**, 896–907 (2004).
21. Meijaard, E. Solving mammalian riddles. A reconstruction of the Tertiary and Quaternary distribution of mammals and their palaeoenvironments in island South-East Asia. PhD Thesis, *The Australian National University, Canberra* (2004).
22. Ropiquet, A. & Hassanin, A. Molecular evidence for the polyphyly of the genus *Hemitragus* (Mammalia, Bovidae). *Mol. Phylo. Evol.* **36**, 154–168 (2005).
23. Bird, M. I., Taylor, D. & Hunt, C. Palaeoenvironments of insular Southeast Asia during the Last Glacial Period: A Savanna corridor in Sundaland?. *Quat. Sci. Rev.* **24**, 2228–2242 (2005).
24. Geist, V. *Deer of the World: Their Evolution, Behaviour, and Ecology* (Stackpole Books, Mechanicsburg, 1998).
25. Ellerman, J. R. & Morrison-Scott, T. C. S. *Checklist of Palaearctic and Indian Mammals, 1758 to 1947* (British Museum Natural History, London, 1951).
26. Gilbert, C., Ropiquet, A. & Hassanin, A. Mitochondrial and nuclear phylogenies of Cervidae (Mammalia, Ruminantia): Systematics, morphology, and biogeography. *Mol. Phylo. Evol.* **40**, 101–117 (2006).
27. Hassanin, A. *et al.* Pattern and timing of diversification of cetartiodactyla (mammalia, laurasiatheria), as revealed by a comprehensive analysis of mitochondrial genomes. *C. R. Biol.* **335**, 32–50 (2012).
28. Pitra, C., Fickel, J., Meijaard, E. & Groves, C. P. Evolution and phylogeny of old world deer. *Mol. Phyl. Evol.* **33**, 880–895 (2004).
29. Balakrishnan, C. N., Monfort, S. L., Gaur, A., Singh, L. & Sorenson, M. D. Phylogeography and conservation genetics of Eld's deer (*Cervus eldi*). *Mol. Ecol.* **12**, 1–10 (2003).
30. Thomas, O. The nomenclature and the geographical forms of the panolia deer (*Rucervus eldi*) and its relatives. *J. Bom. Nat. His. Soci.* **23**, 363–367 (1918).
31. Angom, S., Kumar, A., Gupta, S. K. & Hussain, S. A. Analysis of mtDNA control region of an isolated population of Eld's deer (*Rucervus eldii*) reveals its vulnerability to inbreeding. *Mito. DNA. Part B.* **2**, 277–280 (2017).
32. Xia, X., Xie, Z., Salemi, M., Chen, L. & Wang, Y. An index of substitution saturation and its application. *Mol. Phylo. Evol.* **26**, 1–7 (2002).
33. Lande, R. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *Am. Nat.* **142**, 911–927 (1993).
34. Hughes, A. R., Inouye, B. D., Johnson, M. T. J., Underwood, N. & Vellend, M. Ecological consequences of genetic diversity. *Ecol. Lett.* **11**, 609–623 (2008).
35. Haq, B. U., Hardenbol, J. & Vail, P. R. The chronology of fluctuating sea level since the Triassic. *Sci.* **235**, 1156–1165 (1987).
36. Suraprasit, K., Jongautcharyakul, S., Yamee, C., Pothichaiya, C. & Bocherens, H. New fossil and isotope evidence for the Pleistocene zoogeographic transition and hypothesized savanna corridor in peninsular Thailand. *Quat. Sci. Rev.* **221**, 105861 (2019).
37. Suraprasit, K. *et al.* The middle Pleistocene vertebrate fauna from Khok Sung (Nakhon Ratchasima, Thailand): Biochronological and paleobiogeographical implications. *Zoo Keys.* **613**, 1–157 (2016).
38. Nautiyal, C. M. & Chauhan, M. S. Late Holocene vegetation and climate change in Loktak Lake region, Manipur, based on pollen and chemical evidence. *Palaeb.* **58**, 21–28 (2009).
39. Tripathi, S., Singh, Y. R., Nautiyal, C. M. & Thakur, B. Vegetation history, monsoonal fluctuations and anthropogenic impact during the last 2330 years from Loktak Lake (Ramsar site), Manipur, Northeast India: A pollen-based study. *Palynology* **42**, 406–419 (2017).
40. Leonard, J. A. *et al.* Phylogeography of vertebrates on the Sunda Shelf: A multi-species comparison. *J. Biogeogr.* **42**, 871–879 (2015).
41. Naish, D. Eld's deer: Endangered, persisting in fragmented populations, and morphologically weird... but it wasn't always so. Scientific American Blog Network (2015). <https://blogs.scientificamerican.com/tetrapod-zoology/elds-deer-endangered-fragmented-weird/>. Accessed on 20 April, 2020.

42. National Studbook of Sangai (*Rucervus eldii eldii*), Wildlife Institute of India, Dehradun and Central Zoo Authority (2018) New Delhi. TR. No. 2018/07. [https://wii.gov.in/research\\_report2018](https://wii.gov.in/research_report2018).
43. Angom, S., Tuboi, C., Ghazi, M. G. U., Badola, R. & Hussain, S. A. Demographic and genetic structure of a severely fragmented population of the endangered hog deer (*Axis porcinus*) in the Indo Burma biodiversity hotspot. *PLoS ONE* **15**, e0210382 (2020).
44. Hartl, D. L. & Clark, A. G. Organisation of genetic variation. In *Principles of Population Genetics* (eds Hartl, D. L. & Clark, A. G.) 74–110 (Sinauer Associates, Sunderland, 1997).
45. Sharma, C. & Chauhan, M. S. Vegetation and climate since Last Glacial Maxima in Darjeeling (Mirik Lake), Eastern Himalaya. in *Proc. 29th Int. Geol. Congr. Part B*, 279.e288 (1994).
46. Tripathi, S., Thakur, B., Nautiyal, C. M. & Bera, S. K. Floristic and climatic reconstruction in the Indo-Burma region for the last 13,000 cal. yr: A palynological interpretation from the endangered wetlands of Assam, northeast India. *The Holocene* **30**, 1–17 (2019).
47. Mehrotra, N., Shah, S. K. & Bhattacharyya, A. Review of palaeoclimate records from Northeast India based on pollen proxy data of Late Pleistocene-Holocene. *Quat. Inter.* **325**, 41–54 (2014).
48. Singh, N. R. Fluvial regime of the Manipur river basin and Loktak Lake with study of backflow. M. Tech thesis. *Indian Institute of Technology* (2006).
49. Excoffier, L., Foll, M. & Petit, R. J. Genetic consequences of range expansions. *Annu. Rev. Ecol. Evol. Syst.* **40**, 481–501 (2009).
50. Slatkin, M. & Excoffier, L. Serial founder effects during range expansion: A spatial analog of genetic drift. *Genetics* **191**, 171–181 (2012).
51. Frankham, R., Bradshaw, C. J. A. & Brook, B. W. Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biol. Cons.* **170**, 56–63 (2014).
52. Hassanin, A., Ropiquet, A., Couloux, A. & Cruaud, C. Evolution of the mitochondrial genome in mammals living at high altitude: New insights from a study of the tribe Caprini (Bovidae, Antilopinae). *J. Mol. Evol.* **68**, 293–310 (2009).
53. Moore, S. S., Barendse, W., Berger, K. T., Armitage, S. M. & Hetzel, D. J. S. Bovine and ovine DNA microsatellites from the EMBL and GenBank databases. *Anim. Genet.* **23**, 463–467 (1992).
54. Gaur, A. *et al.* Development and characterisation of 10 novel microsatellite markers from chital deer (*Cervus axis*) and their cross-amplification in other related species. *Mol. Ecol. Not.* **3**, 607–609 (2003).
55. Bishop, M. D. *et al.* A genetic linkage map for cattle. *Genet.* **136**, 619–639 (1994).
56. Marshall, T. C., Slate, J., Kruuk, L. E. & Pemberton, J. M. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* **7**, 639–655 (1998).
57. DeWoody, J. A., Honeycutt, R. L. & Skow, L. C. Microsatellite markers in white-tailed deer. *J. Hered.* **86**, 317–319 (1995).
58. Jones, K. C., Levine, K. F. & Banks, J. D. DNA-based genetic markers in black-tailed and mule deer for forensic applications. *California Dept Fish Game.* **86**, 115–126 (2000).
59. Vaiman, D., Osta, R., Mercier, D., Grohs, C. & Leveziel, H. Characterization of five new bovine dinucleotide repeats. *Anim. Genet.* **23**, 537–541 (1992).
60. Brezinsky, L., Kemp, S. J. & Teale, A. J. ILSTS005: A polymorphic bovine microsatellite. *Anim. Genet.* **24**, 75–76 (1993).
61. Zhang, Q., Ji, Y. J., Zeng, Z. G., Song, Y. L. & Zhang, D. X. Polymorphic microsatellite DNA markers for the vulnerable Hainan Eld's deer (*Cervus eldi hainanus*) in China. *Act. Zoo. Sin.* **51**, 530–534 (2005).
62. Buchanan, F. C. & Crawford, A. M. Ovine dinucleotide repeat polymorphism at the MAF70 locus. *Anim. Genet.* **23**, 185 (1992).
63. Poetsch, M., Seefeldt, S., Maschke, M. & Lignitz, E. Analysis of microsatellite polymorphism in red deer, roe deer, and fallow deer possible employment in forensic applications. *Foren. Sci. Int.* **6**, 1–8 (2001).
64. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Aci. Res.* **22**, 4673–4680 (1994).
65. Kumar, S., Stecher, G., Li, M., Nkayaz, C. & Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
66. Librado, P. & Rozas, J. DnaSPv5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452 (2009).
67. Leigh, J. W. & Bryant, D. PopART: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**, 1110–1116 (2015).
68. Cheng, L., Connor, T. R., Sirén, J., Aanensen, D. M. & Corander, J. Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. *Mol. Biol. Evol.* **30**, 1224–1228 (2013).
69. Corander, J., Marttinen, P., Siren, J. & Tang, J. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC. Bioinf.* **9**, 539 (2008).
70. Ronquist, F. *et al.* MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
71. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control.* **19**, 716–723 (1974).
72. Durraba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: More models, new heuristics and parallel computing. *Nat. Meth.* **9**, 772 (2012).
73. Posada, D. jModelTest: Phylogenetic model averaging. *Mol. Biol. Evol.* **25**, 1253–1256 (2008).
74. Grant, J. R. & Stothard, P. The CG View Server: A comparative genomics tool for circular genomes. *Nucl. Aci. Res.* **36**, 181–184 (2008).
75. Xia, X. & Xie, Z. DAMBE: Software package for data analysis in molecular biology and evolution. *J. Hered.* **92**, 371–373 (2001).
76. Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973 (2012).
77. Lanfear, R., Calcott, B., Ho, S. Y. & Guindon, S. PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* **29**, 1695–1701 (2012).
78. Bibi, F. A multi-calibrated mitochondrial phylogeny of extant *Bovidae* (artiodactyla, ruminantia) and the importance of the fossil record to systematics. *BMC. Evol. Biol.* **13**, 166 (2013).
79. Dong, W., Pan, Y. & Liu, J. The earliest *Muntiacus* (Artiodactyla, Mammalia) from the Late Miocene of Yuanmou, southwestern, China. *C. R. Palevol.* **3**, 379–386 (2004).
80. Hulce, D., Li, X., Snyder-Leiby, T. & Liu, C. S. J. GeneMarker® genotyping software: Tools to increase the statistical power of DNA fragment analysis. *J. Biomol. Tech.* **22**, S35–S36 (2011).
81. Valiere, N. GIMLET: A computer program for analysing genetic individual identification data. *Mol. Ecol. Not.* **2**, 377–379 (2002).
82. Peakall, R. & Smouse, P. E. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28**, 2537–2539 (2012).
83. Kalinowski, S. T., Taper, M. L. & Marshall, T. C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* **16**, 099–1106 (2007).
84. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genet.* **155**, 945–959 (2000).
85. 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).

86. Earl, D. A. & von Holdt, B. M. STRUCTURE HARVESTER: A website and program for visualising STRUCTURE output and implementing the Evanno method. *Cons. Genet. Res.* **4**, 359–361 (2012).
87. Rosenberg, N. A. DISTRUCT: A program for the graphical display of population structure. *Mol. Ecol. Not.* **4**, 137–138 (2004).
88. Archer, F. L., Adams, P. E. & Schneiders, B. B. *strataG*: An R package for manipulating, summarising and analysing population genetic data. *Mol. Ecol. Res.* **17**, 5–11 (2017).
89. 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).
90. Cornuet, J. M. & Luikart, G. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genet.* **144**, 2001–2014 (1996).
91. Luikart, G., Allendorf, F. W., Cornuet, J. M. & Sherwin, W. B. Distortion of allele frequency distributions provides a test for recent population bottlenecks. *J. Hered.* **89**, 238–247 (1998).
92. Peel, D., Waples, R. S., Macbeth, G. M., Do, C. & Ovenden, J. R. Accounting for missing data in the estimation of contemporary genetic effective population size ( $N_e$ ). *Mol. Ecol. Res.* **13**, 243–253 (2013).
93. Waples, R. S. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Cons. Genet.* **7**, 167–184 (2006).
94. Waples, R. S. & Do, C. Linkage disequilibrium estimates of contemporary  $N_e$  using highly variable genetic markers: A largely untapped resource for applied conservation and evolution. *Evol. Appl.* **3**, 244–262 (2010).
95. Do, C. *et al.* NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size from genetic data. *Mol. Ecol. Res.* **14**, 209–214 (2014).
96. Nikolic, N. & Chevalet, C. Detecting past changes in effective population size. *Evol. Appl.* **7**, 663–681 (2014).
97. Chevalet, C. & Nikolic, N. The distribution of coalescence times and distances between microsatellite alleles with changing effective population size. *Theor. Popul. Biol.* **77**, 152–163 (2010).
98. Dallas, J. F. Estimation of microsatellite mutation rates in recombinant inbred strains of mouse. *Mam. Gen.* **3**, 452–456 (1992).
99. Weber, J. L. & Wong, C. C. Mutation of human short tandem repeats. *Hum. Mol. Genet.* **2**, 1123–1128 (1993).
100. Brinkmann, B., Klitschar, M., Neuhuber, F., Huhne, J. & Rolf, B. Mutation rate in human microsatellites: Influence of the structure and length of the tandem repeat. *Am. J. Hum. Genet.* **62**, 1408–1415 (1998).
101. Sajantila, A., Lukka, M. & Syvänen, A. Experimentally observed germline mutations at human micro- and minisatellite loci. *Eur. J. Hum. Genet.* **7**, 263–266 (1999).
102. Ellegren, H. Microsatellite mutations in the germline: Implications for evolutionary inference. *Trends Genet.* **16**, 551–558 (2000).
103. Hrbek, T., de Brito, R. A., Wang, B., Pletscher, L. S. & Cheverud, J. M. Genetic characterisation of a new set of recombinant inbred lines (LGXSM) formed from the intercross of SM/J and LG/J inbred mouse strains. *Mam. Gen.* **17**, 417–429 (2006).

## Acknowledgements

This study was sponsored by the Ministry of Environment, Forest and Climate Change, Government of India through the Compensatory Afforestation Fund Management and Planning Authority (CAMPA) under the project “Conservation Action Plan for Manipur’s Brow antlered deer or Sangai: An Integrated Approach”. We thank Dr Ruchi Badola, Scientist G and Co-principal Investigator of the project for facilitating the study. We are grateful to the Director and the Dean, Wildlife Institute of India (WII), and the Nodal Officer, Wildlife Forensic, and Conservation Genetics Cell (WFCGC) at WII for providing logistic support. We are thankful to the Department of Forests, Government of Manipur, for providing permission to conduct the research in Keibul Lamjao National Park. We express our sincere thanks to Mr A. Madhanraj for his support during sequencing work and Dr Shivani Barthwal and Ms Sayanti Basak for their valuable comments on the manuscript.

## Author contributions

S.A.H.: conceptualisation, fund acquisition, project supervision, writing—review and editing. M.G.G.: sample collection, methodology, data curation, data analysis, writing—original manuscript. S.P.S.: data analysis, writing—original manuscript, writing—review and editing. C.T.: sample collection, writing—original manuscript. S.A., T.G.: sample collection, Data curation. P.N.: writing—review and editing. All authors agreed and approved the final 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/s41598-021-82183-7>.

**Correspondence** and requests for materials should be addressed to S.A.H.

**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) 2021