# Serial dilution shapes genetic variation and defines conservation units in Asian elephants

Anubhab Khan*[1,2,&], Maitreya Sil*[1,3], Tarsh Thekaekara[1,4], Ishani Sinha[1], Rupsy Khurana[5], Raman Sukumar[5,&], Uma Ramakrishnan[1,&]

[1] National Centre for Biological Sciences, TIFR, GKVK Campus, Bangalore, India
[2] School of Biodiversity, One Health and Veterinary Medicine, University of Glasgow, Glasgow, UK
[3] National Institute of Science Education and Research, Bhabaneshwar, India
[4] The Shola Trust, Gudalur, India
[5] Centre for Ecological Sciences, Indian Institute of Science, Bangalore, India
*These authors contributed equally

[&] Corresponding authors: anubhabkhan@gmail.com, rsuku@iisc.ac.in, uramakri@ncbs.res.in

Keywords: Genomics, genetic load, extinction, megaherbivores, South Asia

## Abstract

Megaherbivores are primary consumers who provide unique ecosystem services. Given their body size, they are disproportionately threatened in the Anthropocene. Asian elephants are the largest extant terrestrial megaherbivores native to Asia, with 60% of the population found in India. Despite their ecological and cultural importance, the management/conservation units, genetic history, diversity and threats remain understudied. We re-sequenced 31 whole genomes (between 11X - 32X) from all known elephant landscapes in India and identified five management/conservation units corresponding to elephants in northern India, central India and three in southern India. The genetic data reveal signatures of serial colonisation, and a dilution of diversity from north to south of India. The northern populations diverged from other populations more than 70,000 years ago, and have higher genetic diversity, with low inbreeding/high effective size (Pi = 0.0016±0.0001; $F_{ROH>1MB}$= 0.09±0.03). Two of three populations in southern India have low diversity and are inbred with much lower effective sizes than current populations sizes (Pi = 0.0014±0.00009and 0.0015±0.0001; $F_{ROH>1MB}$= 0.25±0.09 and 0.17±0.02). Additionally, future generations are expected to be more inbred since pairs of extant elephants have large tracts of the genome that are already identical. Analyses of genetic load reveals purging of potentially high-effect deleterious alleles in the southern populations and potential dilution of all deleterious alleles from north to south in India. However, southern Indian elephants are highly homozygous for all the deleterious alleles that persist, despite dilution and purging. High homozygosity of deleterious alleles, coupled with low neutral genetic diversity make them high priority for conservation and management attention. Most surprisingly, our study suggests that patterns of genetic diversity and genetic load can correspond to geographic signatures of serial founding events even in large mobile endangered species.

## Introduction

Megaherbivores are ecological engineers responsible for maintaining several ecosystem functions (Owen-Smith 1987, Waldram et al. 2008, Coverdale et al. 2016, Berzaghi et al. 2023). They are uniquely responsible for nutrient distribution, seed dispersal and germination, and vegetation/habitat modification (Le Roux et al. 2018, Harich et al. 2016). In the Anthropocene they have been affected disproportionately and face various threats across their range (Enquist et al. 2020). In Asia, megaherbivores have lost 56% (for gaur, *Bos gaurus*) to near 100% (for Javan rhino, *Rhinoceros sondaicus*) of their historic range (Mahmood et al. 2021). They are more inbred and lack genetic diversity compared to their small-bodied counterparts (Brüniche-Olsen et al. 2018), and several megaherbivores have gone extinct in the last few centuries. While megaherbivore genomes are being extensively studied in African landscapes, such investigations for Asian counterparts have been lagging, leading to challenges in identification of genetic threats relevant to their conservation and management.

Asian elephants (*Elephas maximus*) are charismatic megaherbivores distributed across South and Southeast Asia but are culturally important across the globe (Sukumar 2011). They are found in a variety of natural ecosystems from tropical evergreen forests to grasslands at various elevations. India harbours at least 60% of the population of wild Asian elephants (Sukumar 2011, Menon and Tiwari 2019). Increase in human footprint and land use change over the past two centuries has impacted elephants significantly, resulting in population isolation even at regional scales. Today, Asian elephant habitats are extremely fragmented and interspersed with farmland, human settlement, commercial plantations and linear transport infrastructure throughout their range (Liu et al., 2017; Padalia et al., 2020) leading to extensive and, often, intense human-elephant conflicts (Sukumar, 1989, 2003; Gubbi et al., 2014). Despite these challenges, camera trap data (Srinivasaiah et al., 2022; Srinivasaiah et al., 2021) and radiotelemetry studies (Baskaran et al. 1995, Venkataraman et al. 2005, Sukumar 2003, Sukumar et al. 2003) show that elephants have annual home ranges of several hundreds of kilometres in India, often encompassing dense human habitation. Such movement through human-dominated areas might offset impacts of fragmentation, and the associated loss of genetic

variation and inbreeding (Parida et al. 2022). Hence, understanding their phylogeographic history and current population genetic structuring, coupled with possible effects of recent fragmentation could be useful in defining population units for conservation and management.

Phylogeographic studies of elephants point towards a tangled history of demographic change, range expansion/contraction, as well as population admixture (Vidya et al., 2009). Using a few microsatellite and mitochondrial markers, Vidya et al. (2005) reported four Indian elephant population genetic clusters, one each in North-Northeastern India and Central India, and two in Southern India separated by the Palghat gap, broadly corresponding to their regional population distributions. However, a more recent study by De et al. (2021) suggest only three major genetic clusters corresponding to Northern India, Northeastern India, and a combined Southern and Central Indian population (with the Northern population appearing to be admixed). The patterns of genetic diversity also varied across different ecoregions. Vidya et al. (2005) showed that the Southern populations harboured lower genetic diversity (haplotype diversity) than the other populations, while according to De et al. (2021), the Northeast Indian populations showed low heterozygosity. The only genome-wide study of elephants globally (Palkopoulou et al., 2018), without range-wide sampling of Asian elephants, show old divergence between populations in southern India across the Palghat Gap and also between northern and the southern populations. The limited published genomic data from Asian elephants, hence, does not allow unequivocal inference of population structure, divergence times, conservation units and conservation challenges.

To address these gaps, we used whole genome sequences of wild-caught Asian elephants from all the major landscapes, encompassing the known biogeographic barriers in India, to assess their genetic structure, demographic history and genetic variation. We infer the predominant factors that have shaped the observed genetic diversity and its structuring across the country. Further, we use this information to suggest population management units and challenges that could be important to elephant conservation.

## Results and Discussion

### Population structure

We collected 31 blood samples from wild born captive elephants of known origin from almost all known elephant landscapes in India (Supplementary Table 1, Parida et al. 2022) and re-sequenced whole genomes (figure 1a). A set of 2,675,655 SNP markers revealed distinct population structure within elephant landscapes in India. The PCA plot revealed that populations separate from the south to north direction along PC1 axis (figure 1b). Elephants from Northern India (Terai) and Northeastern India (Ne & N) form a single cluster (figure 1b and c). Populations from Central India and Northern India cluster together in the PCA but resolve into separate clusters in the STRUCTURE plot (figure 1c) with the most optimum support for five population clusters (North-Northeast, Central and three clusters in the Western Ghats: North of the Palghat gap (NPG), South of the Palghat gap (SPG) and South of Shencottah gap (SSG), EvalAdmix, Supplementary figures 1 and 2). The admixture graph that includes a Bornean elephant (ERR2260499) as an outgroup with branch lengths qualitatively adjusted to the drift parameter (figure 1d) suggests that the central Indian lineage (SWB) derive their ancestry from a lineage of elephants in northern India (N&NE) and Myanmar. Similarly, in southern India (NPG, SPG and SSG) the populations derive their ancestry from a lineage closely related to the NPG population and the SSG population derives its ancestry from a lineage closely related to the SPG population. These patterns suggest a North-northeastern ancestry of elephants in India and colonisation of Central India, with a deeply diverged lineage and sequential colonisation of the Western Ghats and adjoining Eastern Ghats of Southern India.

Our results support previous conclusions that the North and Northeast Indian populations of elephants are different from other Indian populations (Vidya et al. 2005a,b; De et al. 2021) and that the Ganges river has acted as a potential barrier to geneflow (figure 1a). Consistent with movement ecology inferences (Koirala et al. 2016) the elephants in the Northern population (described as Northwest Indian population in De et al. 2021) are connected to the Northeast Indian elephants. This is a large landscape running west to east along the Himalayan foothills and potentially the elephant habitat connectivity here is fragile (Koirala et al. 2016) or even

completely broken in recent times (Sukumar 2011, Menon and Tiwari 2019). Our genetic data also suggest that the Brahmaputra river acts as a barrier to geneflow in the Northeast population. Previous studies suggested that the river acts only as an incomplete barrier, with female-led family groups unable to cross except perhaps at the higher reaches, but not a barrier to adult male elephants (Vidya et al., 2005). Anecdotal information suggests that male elephants swimming across the Brahmaputra (RS and AK independent personal observations), suggesting it is not a barrier to dispersal and movement. De et al., 2021 suggest that the Torsa river is also a barrier to elephant movement. However, radio-telemetry and observational studies in this region have shown that elephants certainly cross the river (Sukumar et al. 2003). We were unable to explicitly investigate the role of the Torsa river as a barrier because of inadequate geographic sampling. While geographically proximate, we find that Central Indian populations are genetically distinct as suggested previously (Vidya et al. 2005).

Our data and analyses allow identification of a novel genetic cluster in Southern India and suggest three genetically differentiated populations in the Western Ghats. Along the Western Ghats, certain mountain passes divide the elephant population with the Palghat gap being the most prominent barrier to elephant dispersal. Further south, the Shencottah Gap also acts as a previously unknown impediment to elephant movement (though anecdotal information suggests that elephants moved across this gap until a few decades ago). Interestingly the population south of Shencottah Gap (SSG) is indistinguishable from the population north of Shencottah (SPG) when SNPs are filtered for LD (Supplementary figures 3-4). This probably shows that the differentiation is potentially because of founder effects and inbreeding combined with recent isolation and small population size (Khan et al. 2022). Geneflow between north and south of the Shencottah Gap may have reduced further recently (compared to across the Palghat gap) largely due to a railway line, a highway and associated development along these transportation infrastructures. While nested biogeographic implications of these gaps on population structure and phylogeography has been highlighted for smaller species (e.g. montane birds, Robin et al., 2015; bush frogs, Vijaykumar et al., 2017, geckoes, Chaitanya et al., 2019) and some mammals (e.g. Lion Tailed Macaque, Ram et al., 2015), that they result in such deeply divergent lineages in large, mobile species like the elephant is

surprising. Alternatively, founding events followed by minimal geneflow might have resulted in the observed patterns of divergence. Additionally, elephant numbers and densities south of the Shencottah Gap have always remained small (on the order of 100 individuals). Such small populations are subject to genetic drift, which could accentuate the observed patterns. Pairwise $F_{ST}$ supports these inferences (Supplementary Table 2) and we find no significant geneflow between the clusters based on F3 statistics (Supplementary Table 3).

Interestingly, the results obtained from the haplotype network analysis based on the mitogenome paints a slightly different picture (Supplementary figure 5). Similar to the nuclear genome, the Northern populations were embedded within the Northeastern populations. However, the Central population was allied to the Southern populations, unlike the results obtained from the nuclear data, more in line with some previous studies (Vidya et al., 2009, De et al., 2021).

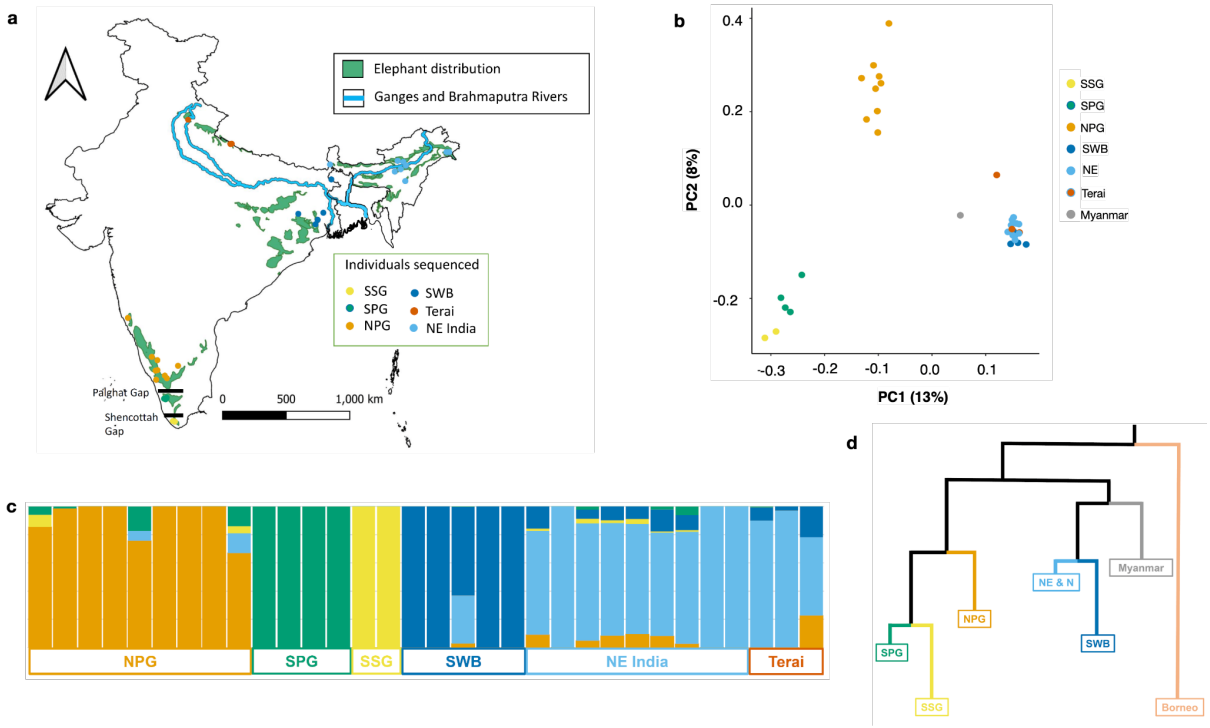


**Figure 1 Population structure (a) sampling locations and the geographic barriers to dispersal (b) PCA (c) ADMIXTURE plot at K=5 and (d) qpgraph with no admixture events with the branch length qualitatively adjusted to the drift parameter. NE&N (n=12) indicates samples from Northeastern India (n=9) and Northern India (n=3), CI (n=5) from Central India, NPG (n=9) from north of**

**Palghat Gap, SPG (n=4) from south of Palghat Gap but north of Shencottah Gap, and SSG (n=2) from south of Shencottah Gap.**

## Demographic History

We investigated recent demographic history for clusters with more than nine samples (N&NE and NPG), and we find that they had undergone a recent bottleneck around 1500-1000 years ago (figure 2a). Since the taming of elephants in the subcontinent during Harappan times, at least four millennia ago, there has been regular exploitation of wild stocks of the species for military and domestic use (Sukumar 2011). The historical accounts suggest that the armies of ancient kingdoms and republics in the north (the Gangetic basin) maintained several thousand captive elephants since as early as the 3rd century BCE, suggestive of overexploitation of wild populations for several hundred years (Sukumar 2011). This is the first report of a bottleneck in Asian elephant populations in recent historical times, mediated perhaps by human exploitation, unlike earlier bottlenecks of a variety of larger mammals globally during Pleistocene glaciations (e.g. Menotti-Raymond and O'Brien 1993, Dalui et al. 2021), Our results also suggest that both populations (N&NE and NPG) may have started recovering from the bottlenecks around 300-500 years ago. Often population structure can dictate the models of demographic history (Mazet et al. 2016). Additionally, we find signatures of population bottlenecks around 100,000 years ago (figure 2b). However, further investigation with hPSMC (Cahill et al. 2016) suggests that this coincides with populations differentiating from each other (figure 2c). We find that the Northern elephant population (NE&N) separated from all other populations about 70,000-100,000 years ago. Central Indian elephants separated from the rest around 50,000-80,000 years ago, while the Southern Indian populations separated from each other only around 20,000-30,000 years ago.

Additionally, our results emphasise the antiquity of the Northern populations of elephants consistent with Vidya et al. (2005, 2009). Palkopoulou et al. (2018) had suggested more recent divergence which could be due to the differences in the mutation rates used for making inferences (Prado et al. 2023). Vidya et al. (2005, 2009) made inferences based on mitochondrial sequences which are expected to provide older estimates of divergence times (Moriyama & Powell 1997, Hugall et al.

2007). However, consistent with Palkopoulou et al. (2018) we find that the Southern populations of elephants split from each other only around 20,000 years ago, or the time of the Last Glacial Maximum when southern India, in particular the Western Ghats, was more arid (Sukumar et al. 1993, Rajagopalan et al. 1997). These findings further support the recognition of five elephant Mus in India, emphasising their antiquity and unique evolutionary histories.

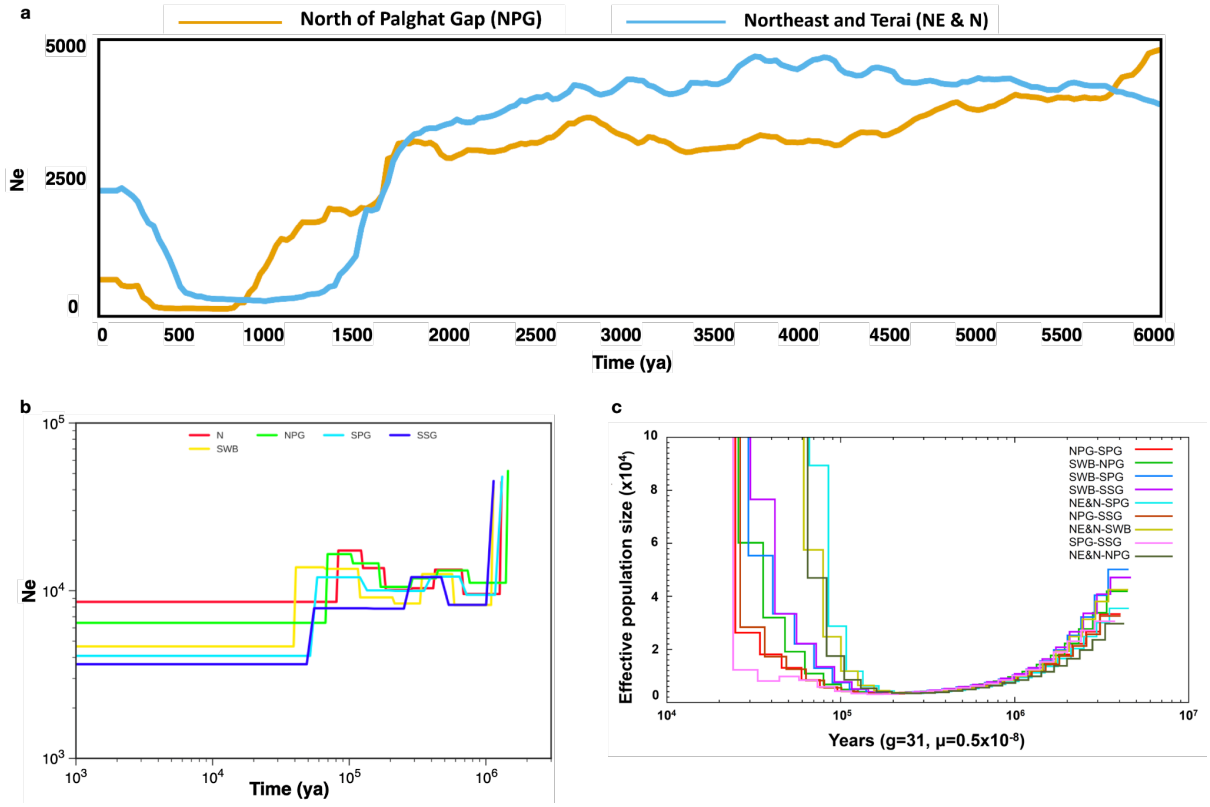


**Figure 2 Demographic history (a) recent demographic history within in the last 100 generations from GONE (b) old demographic history from 100 generations to 40,000 generations from SMC++ and (c) population divergence estimated from hPSMC. The point where population sizes start increasing exponentially is the time to population split in hPSMC plot.**

## Genetic Diversity

But how do these management units compare with each other? Within each cluster we estimated the genetic diversity as the number of average pairwise differences in sequences (pi). We find that the clusters that are from southern India (SPG, SSG and NPG) have lower average pi than the northern Indian and central Indian clusters. Interestingly, there is no visible difference in pi within the southern and

between the northern and central (N&NE and SWB) clusters (figure 3a) (Supplementary Table 4). This suggests that all the clusters within southern have a similar number of haplotypes and the central and Northern cluster have a similar number of haplotypes. About 75% of the discovered SNPs are shared, or present in all populations (Supplementary table 5). However, the number of heterozygous sites encountered per Mb is the highest for the northern population (NE&N) while it is lowest for the southernmost population (SSG) (figure 3b). We find no difference in heterozygous SNV per Mb for NPG and SWB. Overall, our results suggest that most populations of elephants have similar nucleotide diversity (pi) but, in the population south of Shencottah gap (SSG), similar nucleotides are often paired together indicating lower effective population size, a pattern not investigated earlier, while in the populations in Northern India (NE&N) most of the nucleotides pair with a different one indicative of higher effective size.

Genetic variation is related to effective population size which, in turn, is dependent not only on the total population size of a species but other variables such as the sex ratio (Frankel and Soulé 1981). Although we cannot speculate on the historical population sizes of Asian elephants across the various regions we have investigated, a cursory look at the recent population sizes of the five management/conservation units we have inferred from the genomic data shows that population size per se is not indicative of the observed genetic variation (pi and heterozygosity) (Supplementary Table 6). For instance, the genetic variation of the Northeastern population is much higher than that of the Southern population north of the Palghat Gap. Both these clusters have similar current population sizes on the order c. 10,000 individuals. Central Indian elephants have higher variation than Southern populations even though the former has distinctly lower population numbers. This might be indicative of recent bottlenecks in central India, as heterozygosity decays slowly while southern populations, especially those south of Palghat gap may have had historically smaller populations. The highly female-biased sex ratios in Southern populations, especially those to the South of the Palghat Gap, in recent historical times (1970s-1990s) from selective poaching of male elephants for ivory (Sukumar 1989, Ramakrishnan et al. 1998), could have decreased the effective population size (compared to census size), but this is unlikely to have decreased genetic variation to the extent observed. Elephants South of Shencottah Gap were also connected to the

SPG population in recent history. Overall, our results are thus also consistent with a serial dilution of variation that could be the result of sequential colonization (Hellenthal et al., 2008; Pierce et al., 2014; Pless et al., 2022) from north to south: N&NE, Central, Southern, with North of PG, South of PG and South of Shencottah, in that order.

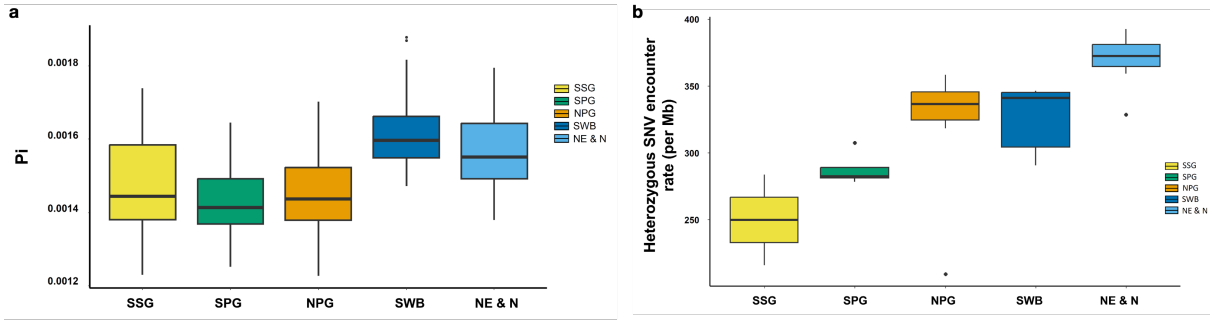


**Figure 3 Genetic diversity (a) boxplots of pairwise nucleotide differences per site (pi) in the populations (error bars are variance in mean pi/site across scaffolds) and (b) heterozygous SNV encounter rate per Mb of the genome.**

We further test this by estimating $F_{ROH,}$ the proportion of the genome in homozygous stretches, length of which is an indicator of older inbreeding/recent bottlenecks. We observe that individuals from south of Shencottah gap (SSG) have high average $F_{ROH>0.1Mb}$ of 0.4 (40% of the genome is in homozygous stretches) while individuals from northern India (NE&N) have the least average $F_{ROH>0.1Mb}$ of 0.2 (20% of the genome is in homozygous stretches, Figure 4a). There is no significant difference between individuals from SPG, NPG and SWB populations (average $F_{ROH>0.1Mb}$ is 0.25). Longer homozygous stretches are indicative of recent inbreeding (Curik et al., 2014; Sumreddee, 2021). ROH longer than 1Mb and 10Mb shows few differences between populations although, the northern population (NE&N) has the least recent inbreeding. Additionally, the SPG and NE&N populations have lower variance in $F_{ROH}$ compared to other populations, potentially indicating geneflow and connectivity within these large landscapes, which will result in lower variance in parental relatedness. We caution that these results do not directly imply choosing to mate between relatives, behavioral studies suggest ongoing inbreeding avoidance in elephants (Sukumar 2006), but more likely suggest founder effect induced drift and mating between relatives, especially south of the Shencottah gap. Founder effects

will result in a small set of individuals as founders, consequentially leading to the individuals in the population soon being related. Mating within this set of individuals then leads to inbreeding.

We measured the percent genome that is identical by descent (IBD) shared between pairs of individuals in each population (Figure 4b). We observe that elephants in the northern Indian cluster (NE&N) share very few IBD stretches of genome (on an average 1.6% in more than 10Mb long and 6.6% in more than 0.1Mb long IBD stretches) with each other while the two individuals we had from the south of Shencottah Gap (SSG) shared about 43% of their genome in stretches longer than 10Mb and about 52% of their genome in stretches longer than 0.1Mb. Apart from the one outlier pair in the SPG sample set we observe very negligible difference in SPG, NPG and SWB population with regards to the proportion IBD stretches of genome shared between pairs of individuals. On an average these three clusters share more than 10% of their genome in IBD longer than 10Mb (NPG: 8%, SPG: 14%, SWB:12%) and 19% of their genome in IBD longer than 0.1Mb (NPG: 15%, SPG: 25%, SWB:17%). These results indicate possible threats from loss of fitness due to inbreeding in upcoming generations. Since the present individuals already share large IBD stretches of genome, impacts on fitness will be maximal, especially if these individuals have higher numbers of deleterious alleles. However, since there are a few outlier pairs with lower amounts of shared genomic IBD stretches (figure 4b) there is hope that maintaining gene flow within these population clusters could sustain extant genetic variation in these populations into the future.

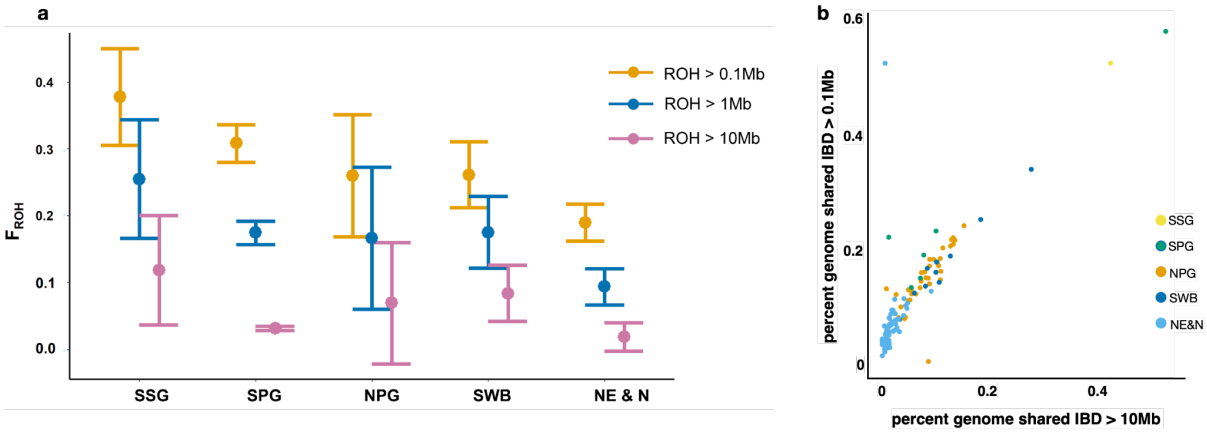


**Figure 4 Present and future inbreeding (a) inbreeding measured as FROH of individuals in each population based on ROH stretches longer than 0.1, 1 and**

**10 Mb of genome and (b) identical by descent (IBD) stretches of genome longer than 0.1Mb vs 10Mb shared between pairs of individuals in each population.**

## Mutation load

We attempted to understand whether individuals, from certain populations, that have high homozygosity could suffer fitness effects due to inbreeding depression by investigating mutation load, or putative deleterious mutations. This depends on the number of derived deleterious mutations harboured across the genome, how many are homozygous, and the putative magnitude of their fitness effects (indel Loss of function > Loss of function > missense mutation). We normalised this by derived SNPs in neutral/intergenic regions, to account for possible missingness/uneven coverage across genomes. As expected, populations with lowest genetic variation (SSG) had fewer LOF mutations (5a, x axis). This could be because of the serial founding demographic history from north to south (see Figure 1d). However, LOFs in SSG were not entirely a subset of those in N&NE, supportive of emergence of *de-novo* mutations in the SSG population due to long history of isolation between the two populations. SSG population has the highest homozygosity for the LOF mutations. Interestingly, SPG also had fewer homozygous LOF mutations (Figure 5a), potentially due to purging and *de-novo* mutation accumulation due to historical isolation.

To better understand whether populations like SSG are endangered by imminent fitness effects, we compared indel LOFs. While the total indel LOFs did not vary considerably between populations (x axis, Figure 5b). The observation that all populations have a similar indel LOF load (compared to LOF load, Figure 5a, x axis) suggests that they are under strong purifying selection, and that these putative high negative effect mutations could have already been purged from these populations. In contrast, the homozygous indel LOFs remain higher in SSG. This suggests that despite purging, elephants south of the Shencottah Gap could experience negative fitness consequences of indel LOFs first. This population may be threatened with inbreeding depression, or lower fitness of individuals, ensuing negative feedback on population growth rates, and a potential for further increase in homozygosity due to inbreeding in the future generations (Robinson et al. 2023). However, African

elephant populations (such as at Addo National Park) grew rapidly with no apparent deleterious effects after a severe bottleneck only about a century ago (Whitehouse and Hall-Martin 2000). Genetically small populations are expected to be especially efficient at purging strongly deleterious alleles (Kyriazis et al. 2023, Khan et al. 2021), this does seem to be happening in the isolated population south of the Shencottah Gap. But the question that remains is whether these endangered populations with long generation times and low population growth rates can tolerate the selection pressure brought about by highly deleterious alleles.

Southern populations (NPG, SPG and SSG) have fewer derived missense mutations than the central (SWB) and the northern (NE&N) populations (Supplementary figure 6) but the populations south of Palghat Gap (SPG and SSG) have higher homozygous missense mutation load than the NPG, SWB and NE&N populations, suggesting lack of purging. The fewer number of mildly deleterious missense alleles in the southern Indian population north of Palghat Gap (NPG) compared to SWB and N&NE is counterintuitive, as current literature suggests larger populations host high numbers of mildly deleterious alleles (Kyriazis et al. 2023, Khan et al. 2021). We suggest serial dilution of missense variants combined with a long isolation from the larger and connected north Indian population (NE&N) may have led to restricted immigration of mild impact deleterious alleles in the NPG population while the LOF are selected out equally well in the large NPG, SWB and NE&N populations. Additionally, the low heterozygosity in the NPG, SWB and NE&N populations cautions that all populations have potential for inbreeding depression (Robinson et al. 2022).

Functionally, the derived missense mutations are related to sensory perception, detection of chemical stimulus and especially olfaction (Supplementary figure 7). This could also be due to independent evolution of olfactory genes in Asiatic elephants (Reddy et al. 2015); however, since we also polarized the alleles with Bornean elephant genome, the missense mutations can be confidently classified to be recently derived and maybe potentially deleterious. The loss-of-function mutations mostly affect protein, ions and nucleic acid binding abilities along with transferases and transporters (Supplementary figure 8). Given that elephants rely heavily on chemical signals, the deleterious mutations may impair crucial olfactory functions in

the elephants. Although inbred individuals are more homozygous for these mutations and maybe expected to be severely affected, it is possible that sociality and herd living allows reduced sensory abilities in individuals. However, this needs further investigation.

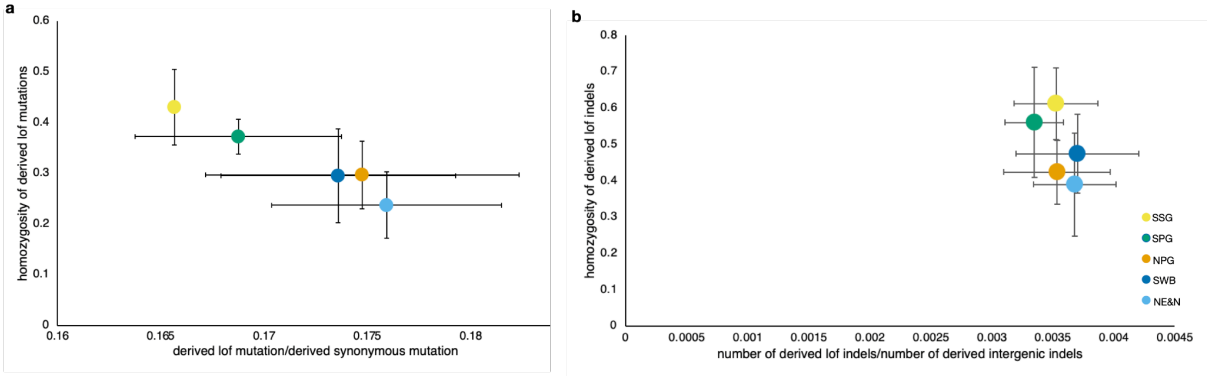


**Figure 5 Mutation load measured as a function of homozygosity vs number of derived (a) indel mutations and (b) loss-of-function (lof) mutation. The number of derived deleterious alleles/number of derived neutral alleles is a proxy for number of deleterious alleles. The error bars indicate standard deviations.**

## Conservation implications

Overall, our results support five management units of Asian elephants in India. Elephants in the Himalayan foothills from the Northern population (in the west) to the Northeast of India are a single cluster that diverged from other Indian populations more than 70,000 years ago. This makes the N&NE cluster the oldest and, hence, the most evolutionarily unique population. Elephant populations south of the Shencottah Gap need conservation attention through detailed genomic research at fine spatial scales, conservation action to protect remaining habitats and minimise conflict associated mortality, and serious consideration about translocations into and out of this population. Our results suggest that further decline in population numbers here could result in inbreeding depression. Long-term studies on inbreeding and possible associated phenotypes will be important to understand future fitness trajectories here. Populations north and south of the Palghat Gap are distinct management units, and animals should be moved across this biogeographic divide only after very careful evaluations of the consequences. Further, given the high missense mutation load in these two management units, connectivity must be maintained within these two landscapes. The Ganga and Brahmaputra river in West

Bengal function as biogeographic divides, making elephants south of these rivers a unique management/conservation unit. Elephants here are facing significant challenges because of negative human impacts, and their conservation must take into account their unique 50,000 year old evolutionary history.

Our descriptive population genomic approach allowed for unique and important conservation insights. To the best of our knowledge, elephants in India present one of the first clear examples of serial founding of a large landscape at a subcontinental scale, with decreasing founder size from the source. Our data allows us to detect recent declines, potentially mediated by historic elephant captures on a large scale. Further, elephants provide an excellent system to better understand the interplay between mutation load and inbreeding in a set of five populations, a rare set up in endangered species. While on-ground conservation challenges for elephants are mitigation of human-associated conflict and human infrastructure associated mortality, conservation genomics insights provide long term conceptual guidance for future survival of these populations.

## Methods

**Sample collection:** State Forest Department databases of captive elephants were examined from 6 states, to locate individual elephants that were captured or rescued from all of the known four major disjunct populations across India. Within these populations, individuals representing habitats across major barriers such as Palghat and the Shencottah Gaps in Southern India and the Ganges and the Brahmaputra rivers in northern and Northeast India were also located. Blood samples and exact capture locations were collected from 28 such individuals and included in our final dataset (see Supplementary Table 1). Out of those samples, 12 were collected from the Southern eco-region, four were collected from the Central eco-region, three were collected from the Northern eco-region, and lastly nine were collected from the Northeastern eco-region. Out of the 12 samples collected from the Southern eco-region, six were collected from North of Palghat Gap, four from South of Palghat Gap (but north of Shecottah Gap), and two from South of Shencottah Gap. We additionally included genomic data from five more individuals obtained from online sources — one individual each from Borneo, Myanmar and Northeastern India, and

three individuals from Southern India (two North of Palghat Gap and one South of Palghat Gap) (See Supplementary Table 1).

**DNA extraction, library preparation and sequencing**

Genomic DNA was extracted from blood samples using Qiagen DNeasy Blood & Tissue Kit. The library preparation and whole genome resequencing was carried out at Medgenome INC. The DNA extraction and sequencing were carried out following Khan et al., 2020.

**Variant calling and filtering**

The raw sequencing reads were trimmed using the default settings in TrimGalore-0.4.5. The trimmed reads were mapped to the *Elephas maximus* reference genome (https://dnazoo.s3.wasabisys.com/index.html?prefix=Elephas_maximus/) using default settings of BWA *mem* (https://github.com/lh3/bwa). The mapped reads were converted to binary format and sorted using Samtools-1.9 (Li et al. 2009). The PCR and optical duplicates were marked using Picardtools MarkDuplicates (http://broadinstitute.github.io/picard) or Samtools-markdup. SNPs were identified using Strelka variant caller 2.9.10 (Saunders et al. 2012). The variants were filtered using VcftoolsV13 (Danecek et al. 2011). We removed indels and retained only SNP loci with minimum phred scaled base quality 30, minimum genotype quality 30, a minimum minor allele count of 3, did not deviate from Hardy-Weinberg equilibrium with chi square p value of 0.05. All sites with the FILTER flag other than PASS were removed. Sites that showed mean depth across individuals more than the 97.5$^{th}$ percentile and less than 2.5$^{th}$ percentile or were missing in more than 20% individuals , were removed. We identified the sex chromosome scaffold as described in Armstrong et al. (2021) and removed the sex chromosome scaffolds.

**Population genetic structure**:

*PCA*

We first employed PCA to partition the data along their main axes of variation. The PCA was carried out using the *--pca* function in Plink (Purcell et al. 2007) based on the final filtered variants. The resulting eigenvectors were plotted in R using ggplot2.

*STRUCTURE*

Thereafter, a maximum likelihood-based method, ADMIXTURE (Alexander et al. 2009) was employed to investigate the population genomic structure of Asian elephants. We estimated the number of clusters across *K* values ranging from 1 to 6, based on the number of clusters obtained from PCA. The evaluation of the most optimal number of clusters was carried out using EvalAdmix (Garcia-Erill et al. 2020). EvalAdmix estimates the correlation of the residual matrices of the individuals from the Admixture analyses. The resulting correlation matrices were plotted in R.

*qpgraph*

We assigned a population to each individual based on the results from admixture. We converted the .vcf file to .ped format using VCFtools and then .ped to .bed format using Plink. Employed the *find_graphs* function in ADMIXTOOLS2 (https://github.com/uqrmaie1/admixtools) to automatically find the optimum graphs using various values for the number of admixture events. The model with zero admixture had the best statistical support as tested using the protocol suggested in ADMIXTOOLS2 with functions *qpgraph_resample_multi* and *compare_fits*.

**Demographic history**

*Recent Demographic History*

We used the LD based GONE (Santiago et al. 2020) for estimating recent demography history up to a couple of hundred generations ago. We used the default settings of the parameters and set *PHASE* as 0. We plotted the results assuming a generation time of 31 years. This is a multiple sample analysis and requires several samples for accurate estimates. Hence, we performed this analysis only for the NE&N and the NPG cluster which have at least nine samples. The results from this analysis better predict recent events but are not very reliable for estimating events older than a hundred generations (Santiago et al. 2020).

*Old demographic History*

We used the coalescent based SMC++ (Terhorst et al. 2017) method for estimating older demographic history. We used the protocol described for the implementation of

the method (https://github.com/popgenmethods/smcpp) as is. We set the mutation rate to 5.3*10^-9/base/generation (Prado et al. 2023) and a generation time of 31 years. The estimates from this method are reliable for estimating demographic events between a hundred generations ago to forty thousand generations ago.

*Divergence time*

We estimate divergence times between populations using hPSMC (). We used a single individual selected at random from each population (Supplementary table 1). We created a pseudo-diploid individual using samples from two populations for each scaffold except for the sex chromosome scaffold (https://github.com/jacahill/hPSMC). Then performed PSMC with the default settings. We plotted the results assuming a mutation rate of 5.3*10^-9/base/generation (Prado et al. 2023) and a generation time of 31 years. The time point where the effective population size estimated from the pseudo-diploid individual rises exponentially is the point where the two haplotypes do not coalesce and hence signify population divergence.

**Genetic diversity**

*Nucleotide diversity: pi*

We randomly subsampled four individuals from N&NE, SWB, NPG and SPG sample set and retained the SSG dataset (n=2) as is. We estimated the average number of pairwise differences per site for each population cluster as described in Wang et al. (2020). Briefly, we set the function *dosaf* to 1 ANGSD (Korneliussen et al. 2014) to estimate allele frequency likelihood for each site. The we used the used the function *realSFS* and estimated a folded site frequency spectrum. We set the function *doThetas* to 1 and estimated pi per site and used the *ThetaStat* function the summarise the pi value for each scaffold. We used only the chromosomal scaffolds from this summary and divided the “tP” values by the number of sites for each scaffold to obtain the average pi per site. Furthermore, we statistically compared the significance of the estimated values between populations (see Supplement).

*Heterozygous SNV encounter rate*

We estimated the number of heterozygous SNPs for each individual using the *vcfstats* function of RTGtools (https://github.com/RealTimeGenomics/rtg-tools). We then divided the number of heterozygous sites by the total number of sites genotyped for the individual and multiplied by 10^6 to obtain SNV encounter rate/Mb.

## Inbreeding

### *FROH*

We estimated runs of homozygosity (ROH) using the *roh* function in BCFtools version 1.3 (Narasimhan et al. 2016) as described previously (Shukla et al. 2023). We classified the ROH stretches into three size classes: the more than 0.1 Mb (100 Kb) stretches show cumulative inbreeding due to old and recent bottlenecks, the more than 1Mb stretches show cumulative inbreeding in the recent past and the more than 10Mb stretches show recent inbreeding. FROH was estimated as described previously (Khan et al. 2021).

### *Shared IBD stretches*

We used IBDseq (Browning & Browning, 2013) to detect the stretches of genome that are identical-by-descent in pairs of individuals. For each pair of individuals in a population, we summed the lengths of IBD stretches longer than 0.1Mb and 10Mb and divided by the total autosomal length.

## Genetic load

### *Filtering variant sites*

We filtered the raw variants obtained from Strelka variant caller 2.9.10 again in VCFtools using the parameters with minimum phred scaled base quality 30, minimum genotype quality 30, minor allele count 1, FILTER flag as PASS. We also removed sited with -0.5>Fis>0.95 as described previously (Khan et al. 2021) and removed sites missing in more than 20% of the individuals. We filtered out the sex chromosomes as well. We call this set of loci as deleterious_allele_set.

### *Identifying ancestral alleles*

We defined the ancestral allele as the allele present most commonly in sister species of Asiatic elephants. For this we used the genome of African elephant (genbank accession: GCA_000001905.1) and dugong (genbank accession: GCA_015147995). We converted these assemblies to 100bp long FASTQ reads and mapped to the Asiatic elephant reference genome and removed reads mapping to multiple regions as described in Khan et al. (2021). We converted the mapped reads to consensus fasta files and estimated the majority allele at a locus. We further filtered these alleles with the genome of Bornean elephant (SRA accession: ERR2260499). Any loci where the Bornean elephant was not homozygous for the ancestral allele was removed from the analysis thus ensuring that the derived alleles are new to the population which is an important assumption for the identification of deleterious alleles (Khan 2023).

*Derived indel load*

We further filtered the loci in the deleterious_allele_set using the *--keep-only-indels* tag in VCFtools. We annotate this set of loci with the Asiatic elephant genome annotation (https://dnazoo.s3.wasabisys.com/index.html?prefix=Elephas_maximus/) using Ensembl Variant Effect Predictor (VEP, McLaren et al. 2016). For indels, the ancestral state cannot be determined using distant species. Hence, the indel allele that was homozygous in the Bornean elephant was referred to as the ancestral allele and any indel site where the Bornean elephant was not homozygous, was removed from the analysis. We then estimated the number of indels predicted to cause transcript_ablation, splice_donor_variant, splice_acceptor_variant, stop_gained, frameshift_variant, inframe_insertion, inframe_deletion, splice_region_variant were classified as lof causing indels. We also counted the derived indels in the intergenic regions of the genome. We divided the number of lof indels with the number of intergenic indels to control for differences in depth leading to missingness in the data.

*Derived lof load*

We filtered the loci in the deleterious_allele_set to remove indels and that had mean depth across individuals more than the 97.5$^{th}$ percentile and less than 2.5$^{th}$ percentile. This set was annotated on VEP and lof mutations were identified as

591 described earlier. We then counted the number of derived lof alleles and divided
592 them with the number of synonymous mutations.

593 *Derived missense load*

594 We used the same set of loci used for estimating the lof load but chose those set of
595 loci that caused non-synonymous changes in the genome and followed the same
596 procedure as described for the lof load.

597 **Acknowledgements**
598 This work was supported by a Department of Biotechnology grant
599 [BT/PR23223/BCE/8/1490/2018] awarded to RS and UR. Permissions for sample
600 collection were granted in letters C-66011/02/2019 dated 09/07/2019 (West Bengal),
601 1729/5-6 dated 09/12/2020 (Uttarakhand), 4838/2019-CWW/WL10 dated 06/10/2019
602 (Kerala), PCCF(WL)/E2/CR-79/2018-19 dated 26/05/2018 (Karnataka),
603 WL/FG31/Technical Committee/2019 dated 26/08/2019 (Assam) and
604 CWL/G/173/2018-19/Pt.VII/1695-702 dated 10/10/2019 (Arunachal Pradesh). We
605 would like to thank the respective forest veterinary doctors of all the states,
606 particularly Dr. Aditi Sharma, Dr. Arun Zachariah and Dr. Shweta Mandal. RS was
607 supported as National Science Chair by Science and Engineering Research Board
608 (SERB), Government of India. Gratitude to NCBS IT team support during data
609 analysis; NCBS data cluster used is supported under project no. 12-R&D-TFR-5.04-
610 0900, Department of Atomic Energy, Government of India). VV Robin and Ansil BR
611 provided helpful comments on drafts of the manuscript. Gratitude to field assistants
612 for their continuous presence in the field and assisting in sample collection for this
613 work.
614

615 **Data availability**

616
617 All sequencing data have been deposited in BioProject SUB13810487 in NCBI.
618
619

## References

Alexander, D.H., Novembre, J. and Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome research, 19(9), pp.1655-1664.

Baskaran, N., Balasubramanian, M., Swaminathan, S. and Desai, A.A., 1995. Home range of elephants in the Nilgiri Biosphere Reserve, south India. A week with elephants. Bombay Natural History Society and Oxford University Press, Bombay, pp.296-313.

Berzaghi, F., Bretagnolle, F., Durand-Bessart, C. and Blake, S., 2023. Megaherbivores modify forest structure and increase carbon stocks through multiple pathways. *Proceedings of the National Academy of Sciences*, *120*(5), p.e2201832120.

Browning, B.L. and Browning, S.R., 2013. Detecting identity by descent and estimating genotype error rates in sequence data. The American Journal of Human Genetics, 93(5), pp.840-851.

Brüniche-Olsen, A., Kellner, K.F., Anderson, C.J. and DeWoody, J.A., 2018. Runs of homozygosity have utility in mammalian conservation and evolutionary studies. *Conservation Genetics*, *19*(6), pp.1295-1307.

Cahill, J.A., Soares, A.E., Green, R.E. and Shapiro, B., 2016. Inferring species divergence times using pairwise sequential Markovian coalescent modelling and low-coverage genomic data. Philosophical Transactions of the Royal Society B: Biological Sciences, 371(1699), p.20150138.

Chaitanya, R., Giri, V.B., Deepak, V., Datta-Roy, A., Murthy, B.H.C.K. and Karanth, P., 2019. Diversification in the mountains: a generic reappraisal of the Western Ghats endemic gecko genus Dravidogecko Smith, 1933 (Squamata: Gekkonidae) with descriptions of six new species. Zootaxa, 4688(1), pp.1-56.

Coverdale, T.C., Kartzinel, T.R., Grabowski, K.L., Shriver, R.K., Hassan, A.A., Goheen, J.R., Palmer, T.M. and Pringle, R.M., 2016. Elephants in the understory: opposing direct and indirect effects of consumption and ecosystem engineering by megaherbivores. Ecology, 97(11), pp.3219-3230.

Curik, I., Ferenčaković, M. and Sölkner, J., 2014. Inbreeding and runs of homozygosity: A possible solution to an old problem. Livestock Science, 166, pp.26-34.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T. and McVean, G., 2011. The variant call format and VCFtools. Bioinformatics, 27(15), pp.2156-2158.

De, R., Sharma, R., Davidar, P., Arumugam, N., Sedhupathy, A., Puyravaud, J.P., Selvan, K.M., Rahim, P.A., Udayraj, S., Parida, J. and Digal, D.K., 2021. Pan-India population genetics signifies the importance of habitat connectivity for wild Asian elephant conservation. Global Ecology and Conservation, 32, p.e01888.

Enquist, B.J., Abraham, A.J., Harfoot, M.B., Malhi, Y. and Doughty, C.E., 2020. The megabiota are disproportionately important for biosphere functioning. *Nature Communications*, *11*(1), p.699.

Frankel, O.H. and Soulé, M.E. 1981. Conservation and evolution. Cambridge University Press, Cambridge, U.K.

Harich, F.K., Treydte, A.C., Ogutu, J.O., Roberts, J.E., Savini, C., Bauer, J.M. and Savini, T., 2016. Seed dispersal potential of Asian elephants. Acta Oecologica, 77, pp.144-151.

Gubbi, S., Swaminath, M.H., Poornesha, H.C., Bhat, R. and Raghunath, R., 2014. An elephantine challenge: human–elephant conflict distribution in the largest Asian elephant population, southern India. Biodiversity and conservation, 23, pp.633-647.

Hedrick, P.W., 1999. Perspective: highly variable loci and their interpretation in evolution and conservation. Evolution, 53(2), pp.313-318.

Hellenthal, G., Auton, A. and Falush, D., 2008. Inferring human colonization history using a copying model. PLoS genetics, 4(5), p.e1000078.

Khan, A., Krishna, S.M., Ramakrishnan, U. and Das, R., 2022. Recapitulating whole genome based population genetic structure for Indian wild tigers through an ancestry informative marker panel. Heredity, 128(2), pp.88-96.

Khan, A., Patel, K., Shukla, H., Viswanathan, A., van der Valk, T., Borthakur, U., Nigam, P., Zachariah, A., Jhala, Y.V., Kardos, M. and Ramakrishnan, U., 2021. Genomic evidence for inbreeding depression and purging of deleterious genetic variation in Indian tigers. *Proceedings of the National Academy of Sciences*, *118*(49), p.e2023018118.

693 Khan, A., Patel, K., Bhattacharjee, S., Sharma, S., Chugani, A.N., Sivaraman, K.,
694 Hosawad, V., Sahu, Y.K., Reddy, G.V. and Ramakrishnan, U., 2020. Are shed hair
695 genomes the most effective noninvasive resource for estimating relationships in the
696 wild?. *Ecology and Evolution*, *10*(11), pp.4583-4594.

697 Khan, A., 2023. The year of the tiger and the year of tiger genomes!. *Molecular*
698 *Ecology Resources*, *23*(2), pp.327-329.

699 Koirala, R.K., Ji, W., Aryal, A., Rothman, J. and Raubenheimer, D., 2016. Dispersal
700 and ranging patterns of the Asian Elephant (Elephas maximus) in relation to their
701 interactions with humans in Nepal. Ethology Ecology & Evolution, 28(2), pp.221-231.
702
703 Korneliussen, T.S., Albrechtsen, A. and Nielsen, R., 2014. ANGSD: analysis of next
704 generation sequencing data. *BMC bioinformatics*, *15*(1), pp.1-13.
705
706 Kyriazis, C.C., Beichman, A.C., Brzeski, K.E., Hoy, S.R., Peterson, R.O., Vucetich,
707 J.A., Vucetich, L.M., Lohmueller, K.E. and Wayne, R.K., 2023. Genomic
708 underpinnings of population persistence in Isle Royale moose. Molecular biology and
709 evolution, 40(2), p.msad021.
710
711 Le Roux, E., Kerley, G.I. and Cromsigt, J.P., 2018. Megaherbivores modify trophic
712 cascades triggered by fear of predation in an African savanna ecosystem. *Current*
713 *Biology*, *28*(15), pp.2493-2499.
714
715 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G.,
716 Abecasis, G., Durbin, R. and 1000 Genome Project Data Processing Subgroup,
717 2009. The sequence alignment/map format and SAMtools. bioinformatics, 25(16),
718 pp.2078-2079.
719
720 Liu, P., Wen, H., Harich, F.K., He, C., Wang, L., Guo, X., Zhao, J., Luo, A., Yang, H.,
721 Sun, X. and Yu, Y., 2017. Conflict between conservation and development: cash
722 forest encroachment in Asian elephant distributions. Scientific Reports, 7(1), p.6404.
723 Mahmood, T., Vu, T.T., Campos-Arceiz, A., Akrim, F., Andleeb, S., Farooq, M.,
724 Hamid, A., Munawar, N., Waseem, M., Hussain, A. and Fatima, H., 2021. Historical
725 and current distribution ranges and loss of mega-herbivores and carnivores of Asia.
726 *PeerJ*, *9*, p.e10738.
727

728 Mazet, O., Rodríguez, W., Grusea, S., Boitard, S. and Chikhi, L., 2016. On the
729 importance of being structured: instantaneous coalescence rates and human
730 evolution—lessons for ancestral population size inference?. Heredity, 116(4),
731 pp.362-371.

McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R., Thormann, A., Flicek, P. and Cunningham, F., 2016. The ensembl variant effect predictor. Genome biology, 17(1), pp.1-14.

Menon, V. and Tiwari, S.K., 2019. Population status of Asian elephants Elephas maximus and key threats. *International Zoo Yearbook*, *53*(1), pp.17-30.

Menotti-Raymond, M. and O'Brien, S.J., 1993. Dating the genetic bottleneck of the African cheetah. *Proceedings of the National Academy of Sciences*, *90*(8), pp.3172-3176.

Narasimhan, V., Danecek, P., Scally, A., Xue, Y., Tyler-Smith, C. and Durbin, R., 2016. BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. Bioinformatics, 32(11), pp.1749-1751.

Owen-Smith, N., 1987. Pleistocene extinctions: the pivotal role of megaherbivores. Paleobiology, 13(3), pp.351-362.

Padalia, H., Ghosh, S., Reddy, C.S., Nandy, S., Singh, S. and Kumar, A.S., 2019. Assessment of historical forest cover loss and fragmentation in Asian elephant ranges in India. Environmental Monitoring and Assessment, 191, pp.1-13.

Palkopoulou, E., Lipson, M., Mallick, S., Nielsen, S., Rohland, N., Baleka, S., Karpinski, E., Ivancevic, A.M., To, T.H., Kortschak, R.D. and Raison, J.M., 2018. A comprehensive genomic history of extinct and living elephants. Proceedings of the National Academy of Sciences, 115(11), pp.E2566-E2574.

Parida, J., Sharma, R., De, R., Kalam, T., Sedhupathy, A., Digal, D.K., Reddy, P.A., Goyal, S.P., Puyravaud, J.P. and Davidar, P., 2022. Genetic characterisation of fragmented Asian elephant populations with one recent extinction in its eastern-central Indian range. Ecological Genetics and Genomics, 24, p.100132.

Pierce, A.A., Zalucki, M.P., Bangura, M., Udawatta, M., Kronforst, M.R., Altizer, S., Haeger, J.F. and de Roode, J.C., 2014. Serial founder effects and genetic differentiation during worldwide range expansion of monarch butterflies. Proceedings of the Royal Society B: Biological Sciences, 281(1797), p.20142230.

Pless, E., Powell, J.R., Seger, K.R., Ellis, B. and Gloria-Soria, A., 2022. Evidence for serial founder events during the colonization of North America by the yellow fever mosquito, Aedes aegypti. Ecology and Evolution, 12(5), p.e8896.

Prado, N.A., Armstrong, E.E., Brown, J.L., Goldenberg, S.Z., Leimgruber, P., Pearson, V.R., Maldonado, J.E. and Campana, M.G., 2023. Genomic Resources for Asian (Elephas maximus) and African Savannah Elephant (Loxodonta africana) Conservation and Health Research. *bioRxiv*, pp.2023-02.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I., Daly, M.J. and Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. The American journal of human genetics, 81(3), pp.559-575.

Rajagopalan, G., Sukumar, R., Ramesh, R., Pant, R.K. and Rajagopalan, G., 1997. Late Quaternary vegetational and climatic changes from tropical peats in southern India–an extended record up to 40,000 years BP. *Current Science*, pp.60-63.

Ram, M.S., Marne, M., Gaur, A., Kumara, H.N., Singh, M., Kumar, A. and Umapathy, G., 2015. Pre-historic and recent vicariance events shape genetic structure and diversity in endangered lion-tailed macaque in the Western Ghats: implications for conservation. PloS one, 10(11), p.e0142597.

Robin, V.V., Vishnudas, C.K., Gupta, P. and Ramakrishnan, U., 2015. Deep and wide valleys drive nested phylogeographic patterns across a montane bird community. Proceedings of the Royal Society B: Biological Sciences, 282(1810), p.20150861.

Robinson, J., Kyriazis, C.C., Yuan, S.C. and Lohmueller, K.E., 2023. Deleterious variation in natural populations and implications for conservation genetics. Annual review of animal biosciences, 11, pp.93-114.

Santiago, E., Novo, I., Pardiñas, A.F., Saura, M., Wang, J. and Caballero, A., 2020. Recent demographic history inferred by high-resolution analysis of linkage disequilibrium. Molecular Biology and Evolution, 37(12), pp.3642-3653.

Saunders, C.T., Wong, W.S., Swamy, S., Becq, J., Murray, L.J. and Cheetham, R.K., 2012. Strelka: accurate somatic small-variant calling from sequenced tumor–normal sample pairs. Bioinformatics, 28(14), pp.1811-1817.

Srinivasaiah NM, Vaidyanathan S, Sukumar R, Sinha A. (2021). Elephants on the move: Implications for human–elephant interactions. In: *The Crisis of Climate Change: Weather Report*. (eds. Agarwal R and Goyal O). Routledge, Abingdon, United Kingdom and New York, pp. 92-103

Srinivasaiah NM, Vaidyanathan S, Sukumar R, Sinha A. (2022). The rurban elephant: Behavioural ecology of Asian elephants in response to large-scale land use change in a human-dominated landscape in peri-urban southern India. In: *New Forms of Urban Agriculture: An Urban Ecology Perspective* (eds. Diehl J A and Kaur H). Springer Nature, Singapore, pp. 289-31

Sukumar, R. 1989. The Asian elephant: ecology and management. Cambridge University Press, Cambridge, U.K.

Sukumar, R., Ramakrishnan, U. and Santosh, J.A., 1998, November. Impact of poaching on an Asian elephant population in Periyar, southern India: a model of demography and tusk harvest. In *Animal Conservation forum* (Vol. 1, No. 4, pp. 281-291). Cambridge University Press.

Sukumar, R., Ramesh, R., Pant, R.K. and Rajagopalan, G., 1993. A δ13C record of late Quaternary climate change from tropical peats in southern India. *Nature*, *364*(6439), pp.703-706.

Sukumar, R. 2003. The Living Elephants: Evolutionary Ecology, Behavior and Conservation. Oxford University Press, New York

Sukumar, R., 2006. A brief review of the status, distribution and biology of wild Asian elephants Elephas maximus. International Zoo Yearbook, 40(1), pp.1-8.

Sukumar, R., 2011. *The story of Asia's elephants*. Mumbai: MARG foundation.

Shukla, H., Suryamohan, K., Khan, A., Mohan, K., Perumal, R.C., Mathew, O.K., Menon, R., Dixon, M.D., Muraleedharan, M., Kuriakose, B. and Michael, S., 2023. Near-chromosomal de novo assembly of Bengal tiger genome reveals genetic hallmarks of apex predation. GigaScience, 12, p.giac112.

Sumreddee, P., Hay, E.H., Toghiani, S., Roberts, A., Aggrey, S.E. and Rekaya, R., 2021. Grid search approach to discriminate between old and recent inbreeding using phenotypic, pedigree and genomic information. BMC genomics, 22, pp.1-17

840 Terhorst, J., Kamm, J.A. and Song, Y.S., 2017. Robust and scalable inference of
841 population history from hundreds of unphased whole genomes. Nature genetics,
842 49(2), pp.303-309.
843
844 Venkataraman, A.B., Saandeep, R., Baskaran, N., Roy, M., Madhivanan, A. and
845 Sukumar, R., 2005. Using satellite telemetry to mitigate elephant–human conflict: an
846 experiment in northern West Bengal, India. Current Science, pp.1827-1831.
847
848 Vidya, T.N.C., Fernando, P., Melnick, D.J. and Sukumar, R. 2005a. Population
849 differentiation within and among Asian elephant (Elephas maximus) populations in
850 southern India. Heredity, 94(1), pp.71-80.
851
852 Vidya, T.N.C., Fernando, P., Melnick, D.J. and Sukumar, R., 2005b. Population
853 genetic structure and conservation of Asian elephant (*Elephas maximus*) across
854 India. Animal Conservation, 8: 377-388.

855 Vidya, T.N.C., Sukumar, R. and Melnick, D.J., 2009. Range-wide mtDNA
856 phylogeography yields insights into the origins of Asian elephants. Proceedings of
857 the Royal Society B: Biological Sciences, 276(1658), pp.893-902.

858 Vijayakumar, S.P., Menezes, R.C., Jayarajan, A. and Shanker, K., 2016. Glaciations,
859 gradients, and geography: multiple drivers of diversification of bush frogs in the
860 Western Ghats Escarpment. Proceedings of the Royal Society B: Biological
861 Sciences, 283(1836), p.20161011.

862 Whitehouse, A.M. and Hall-Martin, A.J., 2000. Elephants in Addo Elephant National
863 Park, South Africa: reconstruction of the population's history. Oryx, 34(1), pp.46-55.

864 Waldram, M.S., Bond, W.J. and Stock, W.D., 2008. Ecological engineering by a
865 mega-grazer: white rhino impacts on a South African savanna. Ecosystems, 11,
866 pp.101-112.

867 Wang, X., Bernhardsson, C. and Ingvarsson, P.K., 2020. Demography and natural
868 selection have shaped genetic variation in the widely distributed conifer Norway
869 spruce (Picea abies). *Genome biology and evolution*, *12*(2), pp.3803-3817.

870

871