Genomic evidence for inbreeding depression and purging of deleterious genetic variation in Indian tigers

Anubhab Khan[a,1], Kaushalkumar Patel[a], Harsh Shukla[a], Ashwin Viswanathan[a,b], Tom van der Valk[c], Udayan Borthakur[d], Parag Nigam[e], Arun Zachariah[f], Yadavendradev V. Jhala[e], Marty Kardos[g,1], and Uma Ramakrishnan[a,h,1]

[a]National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bangalore 560065, India; [b]Nature Conservation Foundation, Mysore 570017, India; [c]Swedish Museum of Natural History, Stockholm 4011418, Sweden; [d]Aranyak, Guwahati 781028, India; [e]Wildlife Institute of India, Dehradun 248001, India; [f]Kerala Forest Department, Wayanad 670645, India; [g]Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, WA 98112; and [h]Department of Biotechnology-Wellcome Trust India Alliance, Hyderabad 500034, India

Increasing habitat fragmentation leads to wild populations becoming small, isolated, and threatened by inbreeding depression. However, small populations may be able to purge recessive deleterious alleles as they become expressed in homozygotes, thus reducing inbreeding depression and increasing population viability. We used whole-genome sequences from 57 tigers to estimate individual inbreeding and mutation load in a small–isolated and two large–connected populations in India. As expected, the small–isolated population had substantially higher average genomic inbreeding ($F_{ROH} = 0.57$) than the large–connected ($F_{ROH} = 0.35$ and $F_{ROH} = 0.46$) populations. The small–isolated population had the lowest loss-of-function mutation load, likely due to purging of highly deleterious recessive mutations. The large populations had lower missense mutation loads than the small–isolated population, but were not identical, possibly due to different demographic histories. While the number of the loss-of-function alleles in the small–isolated population was lower, these alleles were at higher frequencies and homozygosity than in the large populations. Together, our data and analyses provide evidence of 1) high mutation load, 2) purging, and 3) the highest predicted inbreeding depression, despite purging, in the small–isolated population. Frequency distributions of damaging and neutral alleles uncover genomic evidence that purifying selection has removed part of the mutation load across Indian tiger populations. These results provide genomic evidence for purifying selection in both small and large populations, but also suggest that the remaining deleterious alleles may have inbreeding-associated fitness costs. We suggest that genetic rescue from sources selected based on genome-wide differentiation could offset any possible impacts of inbreeding depression.

large carnivore | purifying selection | natural selection | conservation genomics | small population size

A large proportion of Earth's biodiversity persists in small and isolated populations in today's anthropogenically modified world (1). Such populations may suffer from decreased genetic variation and increased inbreeding (2), which together lead to decreased fitness and increased extinction risk (3). Several theoretical (4–7), experimental (8), and empirical studies (9) reveal that species surviving in small and isolated populations are at the greatest risk of extinction.

While species exist in nature along a continuum from small to large populations with different levels of isolation, populations of endangered species often tend to be small and isolated. African wild dog, Ethiopian wolf, and great Indian bustard are examples of species where all populations are small and isolated (10–12). The "small population paradigm" of conservation biology suggests that such smaller and more isolated populations are at a higher risk of extinction due to inbreeding depression and demographic stochasticity (13–15).

Inbred individuals express deleterious, partially recessive alleles that are inherited identically by descent (IBD) from related parents, leading to inbreeding depression (16). Such inbreeding depression can reduce the average fitness of a population, eventually leading to reduced population size and possibly extinction (17). A commonly adopted strategy to conserve inbred populations is genetic rescue (18), which aims to increase average fitness by decreasing the frequency of deleterious mutations and increase heterozygosity at loci harboring deleterious alleles, via translocations of individuals from genetically differentiated populations. A meta-analysis of empirical data from wild populations showed broadly consistent positive effects of genetic rescue on fitness (15, 19).

Population genetic theory (20–22) predicts that purifying selection can reduce inbreeding depression by purging deleterious alleles from inbred populations in the absence of immigration. Whether isolated populations are likely to purge a substantial fraction of the mutation load has been of longstanding interest in evolutionary biology and conservation. Early empirical data from pedigreed captive populations suggested that purging either was absent or resulted only in slight decreases in inbreeding

Significance

Habitat fragmentation is sequestering species into small and isolated populations, with high chances of extinction. Are small and isolated populations at risk for extinction from inbreeding depression? Or does inbreeding and purging of deleterious alleles reduce such threat? Using whole genomes from several wild Indian tiger populations, we provide evidence supporting purging of highly deleterious variants in a small–isolated population. However, our analyses also indicate that the remaining highly deleterious alleles are at high frequencies, suggesting continued inbreeding depression despite some successful purging. We discuss the implications of our results for conservation, including possible genomics-informed genetic rescue strategies.

depression (23, 24). However, several experimental studies based on model organisms reveal substantial purging and significant reduction in inbreeding depression in small populations (25–27). Recent molecular and population genetic studies have found genomic evidence for purging in wild populations (28–30). Despite broad empirical support for the efficacy of genetic rescue (15, 19), genomic evidence for purging and the long-term persistence of some small–isolated populations have been cited to question the small population paradigm and to argue that standard genetic rescue practices are likely to be counterproductive (29, 31–33). Whether purging removes enough deleterious alleles to improve the viability of small, isolated populations (contradicting the small population paradigm) remains an open question. We address this question by contrasting genomic inbreeding and mutation load in small–isolated versus large–connected populations of wild tigers.

We use Bengal tigers (*Panthera tigris tigris*) from India as a model to investigate levels of inbreeding and relative mutation loads in small–isolated and large–connected populations and examine the potential for genetic rescue. Tigers are large, endangered carnivores, but Bengal tigers have high genetic variation compared to other subspecies, with some subpopulations showing high inbreeding indicating isolation (34). All Bengal tiger populations have been through historic bottlenecks, but inbreeding and genetic variation vary among populations. Also, some populations are relatively large and connected while others are small and isolated from both genetic (35) and demographic perspectives (36, 37), making them an ideal system to investigate inbreeding, mutation load, and possible genetic rescue strategies.

We use genomic data to measure the impact of historically declining population sizes and connectivity on inbreeding and mutation load in three wild Bengal tiger (*P. tigris tigris*) populations that are small–isolated (SI) and large–connected (LC). The small–isolated population is from northwestern India and the large–connected populations are from southern India (s-LC) and central India (c-LC). We expected large–connected populations to be the least inbred and to have the lowest mutation load. Alternatively, purging could result in lower mutation load in this small–isolated population. We also explore strategies for genetic rescue that might effectively decrease inbreeding depression. For tigers, we specifically suggest strategies for identifying populations that may benefit from genetic rescue and how such strategies may be effectively implemented.

Results and Discussion

Genomic Estimates of Inbreeding. Tigers exist as several population genetic clusters within India with the total population size of these clusters ranging from 60 to 1,000 individuals. Each genetic cluster consists of several protected areas, with varying connectivity with other such protected areas and clusters (35, 36, 38). For the purpose of this paper, we defined populations and whether they are "isolated" or "connected" based on population genetic and gene flow analyses using markers across the genome (34, 35, 39). Here, the "large" populations currently have hundreds of tigers (minimum 300 tigers), while "small" populations currently have fewer than 100 tigers (36). While both large populations are part of connected landscapes, they do not have identical histories of connectivity. The southern large–connected population is disconnected from other tiger population genetic clusters in India (e.g., from central India and northwestern India), while the central large–connected population (central Indian landscape) was connected by gene flow to other tiger genetic clusters until recently (35, 39). Site-frequency spectra (SFS)-based demographic history models in Armstrong et al. (34) suggest that most Indian tigers diverged from each other recently. Increased agriculture, bounty hunting

and illegal poaching, habitat loss, and fragmentation together have led to local extinctions of tiger populations (40, 41). The small–isolated population we studied here has fewer than 100 individuals, but was connected (until very recently) to the central large–connected population (42) and potentially to the now extinct tiger population in Afghanistan (43).

We quantify inbreeding by identifying long stretches of the genome that are homozygous and identical by descent (i.e., runs of homozygosity [ROH]). The summed length of ROH divided by total autosomal size is an accurate genomic measure of inbreeding (F_{ROH}) (44). We found that the mean F_{ROH}, measured using ROH longer than 100 kb ($F_{ROH>100kb}$), 1 Mb ($F_{ROH>1Mb}$), and 5 Mb ($F_{ROH>5Mb}$) of all individuals in our dataset (regardless of population) was above 0.48, 0.31, and 0.23, respectively. Tigers from the small–isolated population were most inbred with an average $F_{ROH>100kb}$ of 0.57, while average inbreeding due to recent ancestors (on average up to five generations ago, contingent on assumed recombination rate from domestic cat) (45) was above $F_{ROH>5Mb} = 0.37$ (Fig. 1A). The F_{ROH} was proportional to effective population size (N_e), with the small–isolated population having lower N_e at all time intervals compared to both large and connected populations. This level of inbreeding is empirically associated with high risk of extinction (46–48). However, shorter ROH could be due to shared demographic history of bottlenecks (e.g., $F_{ROH} = 100$ kb) and inbreeding depression may be observed at higher F values for large carnivores (49). Pedigree-based studies (50) have reported that an inbreeding coefficient of 0.47 is lethal for tigers.

We show that inbreeding was low in the large–connected populations compared to the small–isolated population (Fig. 1 A and B). We compared our estimates for zoo individuals with their known pedigree values to authenticate these results and found that estimates of recent inbreeding indeed correspond with pedigree values (51) (*SI Appendix*, Table 1). Overall, the long ROH of ≥ 10 Mb, signifying inbreeding due to ancestors as recent as three generations ago, are rare in large populations but frequent in the small–isolated population. The mean inbreeding due to recent ancestors in the small–isolated population ($F_{ROH>10Mb}$) was above 0.28, which could lead to several negative effects (46, 48). We found that the large–connected populations showed inbreeding due to recent ancestors (on average up to five generations ago) of less than $F_{ROH>5Mb} = 0.1$. The large southern population has more inbred individuals than the large central population, potentially due to its geographical placement, historic size (Fig. 1D), and connectivity (35, 36, 39). Overall, although the theoretical expectation is that populations with $N_e < 1,000$ are unlikely to persist (7, 30, 52, 53), this result indicates that a population with a census size of ~600 tigers may be sufficiently large to buffer short-term inbreeding effects given some gene flow. The long-term viability of these populations remains to be tested. While there are instances of small tiger (and other carnivore) populations persisting (36, 37), this could possibly be due to their ability to purge some deleterious alleles. This implies that other large felid species like lions (54), snow leopards (55, 56), and jaguars (57) that often persist in population sizes close to 600 may not necessarily experience the detrimental effects of inbreeding depression over the short term. We caution that this is likely to depend on founder effects and variation in the mutation load among populations of a species (but see ref. 30) and does not address the likelihood for fitness decline over the long term (7).

In addition to having the highest F_{ROH}, we found that pairs of individuals from the small–isolated population also shared large tracts of genome that were IBD (Fig. 1C). This could be a worrying sign for its long-term future as the pairwise IBD sharing translates into inbreeding in the offspring. On average, pairs of individuals from the small–isolated population shared
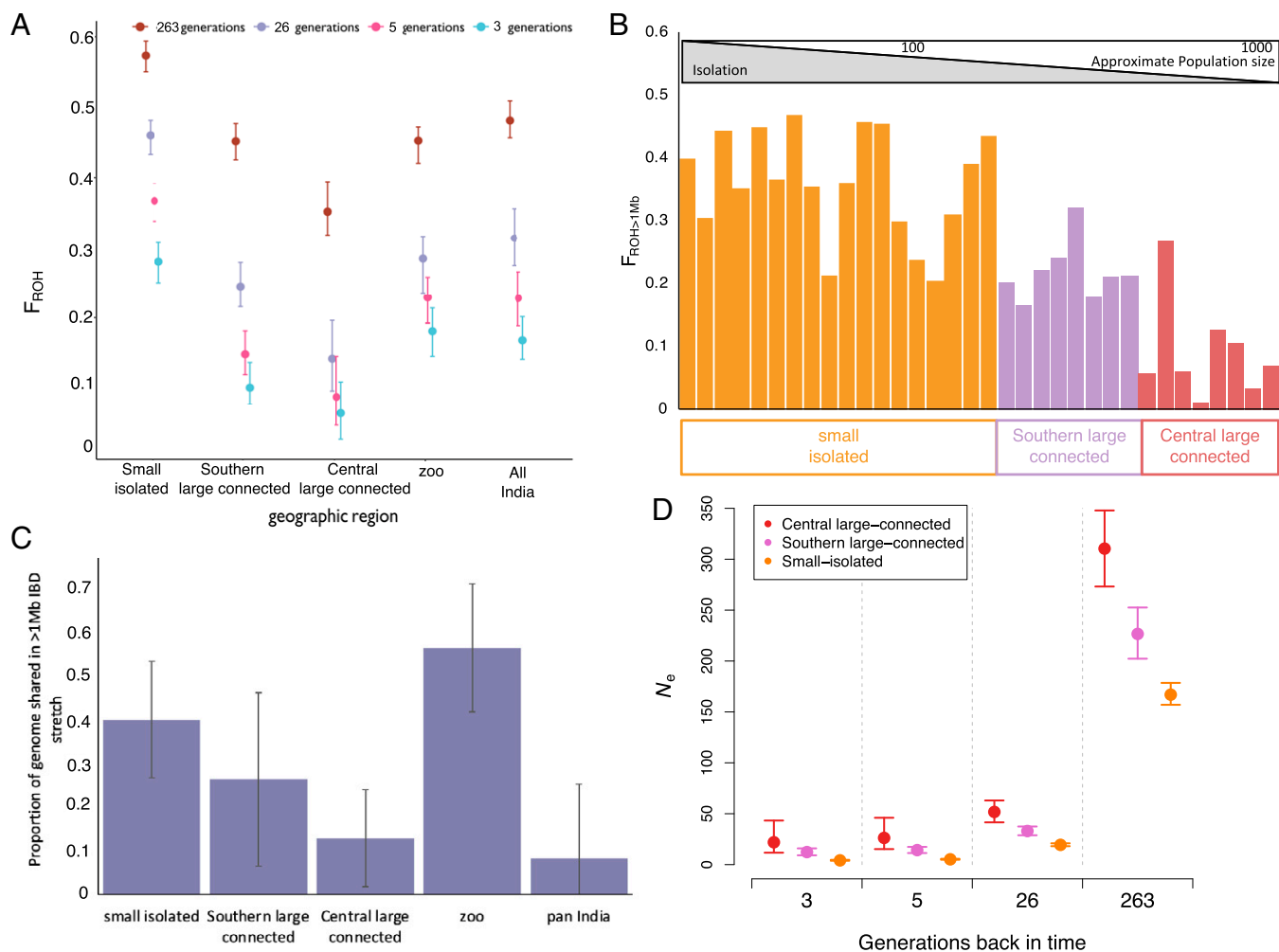
Fig. 1. Inbreeding and shared ancestry in Indian tigers. Here we show cumulative inbreeding due to parental shared ancestry due to ancestors at different historical time ranges as estimated by different lengths of ROH; that is, ROH ≥ 10 Mb indicate shared ancestry on average up to 3 generations ago, 5-Mb stretches indicate on average up to 5 generations ago, 1-Mb stretches indicate on average up to 26 generations ago, and 0.1-Mb stretches indicate on average up to 263 generations ago. Shown is the cumulative inbreeding in each population due to shared ancestry between parents at each time range. Error bars are 95% CI. (*A*) Inbreeding in individuals due to shared ancestry 26 generation ago. (*B*) The proportion of the genome shared in 1-Mb or longer tracts between pairs of individuals from a population of individuals belonging to central large–connected, southern large–connected, and small–isolated populations. Pan-India comparisons in *C* include all pairwise comparisons of wild individuals. Error bars are SEs. N_e over time is estimated from F_{ROH} for the different populations. (*D*) Error bars are SDs.

about 40% of their genome (as more than 1-Mb–long IBD stretches), while pairs from the southern large–connected and the central large–connected populations share about 25 and 15%, respectively, slightly higher than the corresponding mean F_{ROH} values (Fig. 1*A*). As with inbreeding in offspring, pairwise IBD sharing is expected to be influenced by founding bottlenecks in addition to recent small N_e (Fig. 1*D*). The small–isolated population had the lowest mitochondrial haplotype diversity and the lowest number of mitochondrial haplotypes (*SI Appendix*, Fig. 1). This lack of diverse lineages, large tracts of shared IBD genome between individuals, and the recent small population size may therefore require continued immigration of individuals to sustain genetic variation (58) or, in the absence of natural immigration, may require genetic rescue.

Mutation Loads in Small–Isolated and Large–Connected Tiger Populations. We evaluated the relationships between F_{ROH} and the frequency and genomic distribution of putatively deleterious genotypes identified using the Ensembl Variant Effect Predictor (VEP) (59). Under the assumption that most deleterious alleles

are partially recessive (28, 60, 61), the number of homozygous damaging alleles is likely to be informative of the fitness cost of individual inbreeding. First, we found that the number of homozygous putatively damaging alleles (loss-of-function [LOF] and missense alleles together) was proportional to F_{ROH} in all populations. The number of putatively damaging homozygotes increased from <1,500 for individuals with $F_{ROH} \approx 0.3$ to ~2,200 for individuals with $F_{ROH} \approx 0.6$ ($P < 0.01$ for each population, linear regressions; Fig. 2). This genomic evidence suggests a fitness cost associated with increased inbreeding. Additionally, the burden of homozygous putatively damaging mutations was substantially larger on average in the small–isolated population (mean ≈ 2,000 per individual) than in the central large–connected population (mean < 1,600 per individual, $P < 0.0001$, randomization test) or the southern large–connected population (mean > 1,800 per individual, $P = 0.0025$, randomization test; Fig. 2). This is likely due to higher N_e in the central large–connected population compared to the southern large–connected population (Fig. 1*D*).

We further investigated differential mutation load between populations for LOF and missense mutations identified with
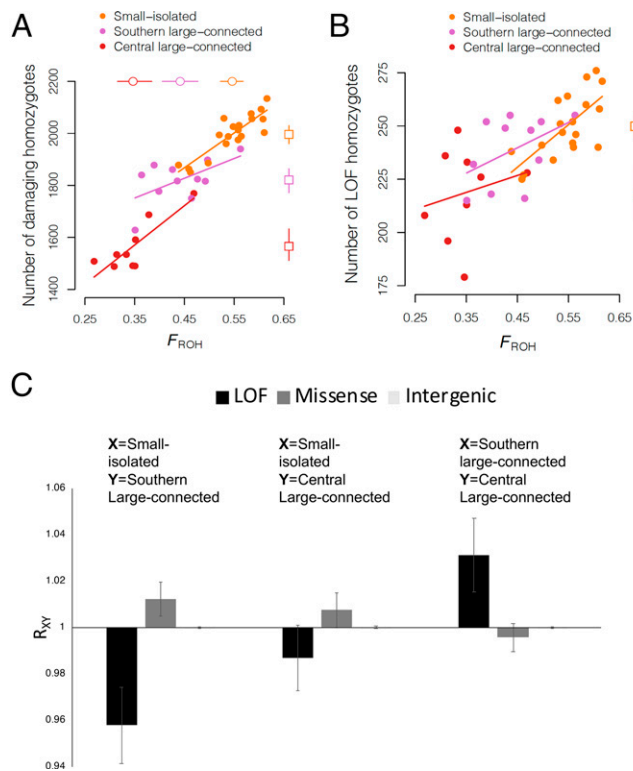
Khan et al.
Genomic evidence for inbreeding depression and purging of deleterious genetic variation in Indian tigers

PNAS | **3 of 10**
https://doi.org/10.1073/pnas.2023018118

EVOLUTION

Fig. 2. (A–C) Relationships between inbreeding and number of homozygous damaging alleles (LOF plus missense) (A) and homozygous LOF mutations (B) and relative mutation loads in populations of various sizes and levels of isolation (C). Solid circles represent the values of F_{ROH} (x axis) for different individuals and the number of homozygous damaging mutations (y axis), with each population represented by color as indicated in the key. Open circles and squares represent the mean F_{ROH} and mean number of homozygous damaging mutations, respectively, for each population, with the error bars representing 95% percentile bootstrap confidence intervals. R_{XY} indicates excess of deleterious alleles in population X relative to Y. All populations had similar numbers of neutral (intergenic) mutations, and error bars are SEs from jackknife in C.

VEP and used the R_{XY} method (28, 62) to estimate relative excess of number of putatively deleterious alleles in one population compared to another. R_{XY} compares the number of derived alleles of a particular category (LOF, missense, or putatively neutral) found in one population relative to the other. If $R_{XY} = 1$, both populations X and Y have the same load of derived alleles, whereas if $R_{XY} > 1$, then population X has more derived alleles than Y and vice versa if $R_{XY} < 1$ (28, 62). As above, we found a higher load associated with missense mutations in the small–isolated population (Fig. 2C). This is likely explained by more efficient purifying selection against missense mutations of relatively small effect, which are typically only slightly recessive (61), due to larger N_e in the large populations than in the small population. We also found a lower load of LOF mutations in the small–isolated population compared to the large populations (Fig. 2C). This suggests that higher inbreeding in the small–isolated population facilitated purging of relatively large-effect LOF mutations—which are more likely to be highly recessive than small-effect deleterious mutations (61, 63). Together, these different approaches suggest a higher putatively damaging allele load in the small–isolated population coupled with purging of loss-of-function alleles in the small–isolated (and inbred) tiger population.

Despite this evidence for purging of LOF alleles in the small–isolated population, it appears that high inbreeding led to individuals in this population having a higher number of homozygous LOF alleles on average (n = 250) compared to the southern

large–connected (n = 238) and central large–connected (n = 218) populations (Fig. 2B). High homozygosity for remaining LOF alleles suggests that the fitness cost due to LOF mutations may be higher in the small–isolated population compared to the large populations despite partial purging. Importantly, while it is impossible to know the realized fitness effects of homozygosity for LOF and missense mutations, this analysis suggests that purging was not sufficient to result in a smaller fitness cost of inbreeding in the small–isolated population compared to the larger populations. We observe similar results when we used a Genomic Evolutionary Rate Profiling (GERP) (64, 65) analysis to identify potentially highly deleterious alleles by keeping loci with GERP scores in the top 0.1% (*SI Appendix*, Fig. 3). For this analysis and those below, it important to recognize that our measure of mutation load is not a direct measure of fitness. Rather, our inferences rest on the assumption that missense and LOF mutations and mutations at conserved sites are deleterious on average, which is well supported empirically (65, 66). Definitively determining whether the bottlenecks faced by these populations (40, 41) and the additional bottleneck faced by northwestern Indian population three generations ago in 2004 (67) purged part of the mutation load would require historical genetic data to quantify pre- versus postbottleneck mutation loads. Additionally, better estimates of very recent demographic history would help to predict the potential for purging during the bottlenecks.

Annotated sets of deleterious variants are rarely available for endangered species. We compared our missense mutations to annotated sets of domestic cat mutations (*SI Appendix, SI Methods* and Table 2). These missense mutations mapped to genes responsible for disease states such as hypertrophic cardiomyopathy, progressive retinal atrophy, polycystic kidney disease, cystinuria, gangliosidosis, hyperoxaluria, hypothyroidism, mucolipidosis, and mucopolysaccharidosis. Further studies are needed to functionally validate the effect of these mutations in tigers. Studies on inbred/small populations of Asiatic lions revealed several abnormalities (68), and genomic studies here might provide further insights on deleterious mutations in large cats. In the small–isolated population at least one tiger has a potential eye condition (*SI Appendix*, Fig. 2), although no skeletal defects were observed in an earlier study (69). Additional quantitative genetic studies on inbred zoo individuals can aid in understanding the possible phenotypic effects of inbreeding and inbreeding depression for endangered species.

SFS revealed that nearly 14% of both putatively damaging and neutral derived alleles were fixed in the small–isolated population (Fig. 3). A smaller proportion of derived alleles (both damaging and neutral) were fixed in the large populations (<0.06). However, there were slightly more fixed derived alleles in the southern large–connected population than in the central large–connected population. The higher abundance of fixed derived alleles (both damaging and neutral) in the small–isolated population, and to a lesser extent in the southern large–connected population, suggests that genetic drift (e.g., due to historical population bottlenecks and founder effects) drove previously rare alleles to fixation in these populations (Fig. 3) (70). The fixation of deleterious alleles is consistent with theoretical predictions for very small populations (4, 7) and empirical observations of wolves (32), Apennine brown bears (71), and other species (31). In the Isle Royale wolves, high-frequency deleterious alleles were consistent with frequent bone deformities in the population (32, 72). The large fraction of fixed putatively damaging mutations in the small–isolated population suggests that partial purging of the mutation load due to LOF mutations (Fig. 2C) was not sufficient to erase a substantial fraction of the mutation load in this population. While the SFS was essentially flat for polymorphic loci (i.e., excluding loci with derived allele frequencies of 0 or 1) in the small–isolated population (as expected following a population
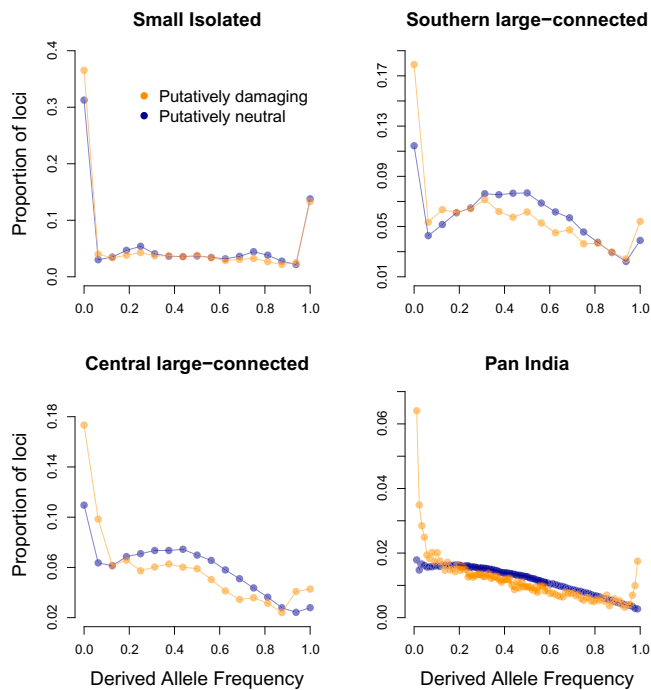
Fig. 3. Site-frequency spectrum for damaging (missense and LOF mutations combined, orange) and neutral mutations (blue) for each population. The proportion of loci (y axis) is shown for each possible derived allele frequency (x axis). Sample sizes were equalized among the three populations by randomly subsampling 16 nonmissing alleles for each locus in each of the three populations. Fixed (frequency = 1) and missing (frequency = 0) alleles are included in the SFS for the three individual populations. We subsampled 88 nonmissing alleles at each locus to calculate the SFS for pan India to account for missing genotypes.

bottleneck), there was a larger fraction of rare derived alleles in the large populations, as expected for large populations at mutation drift equilibrium (70). A larger proportion of putatively damaging than neutral alleles were at or near fixation in each of the large populations and across all of India (Fig. 3), possibly due to positive selection driving selective sweeps on a small fraction of nonsynonymous mutations. The hump-shaped SFS are likely the result of a complex demographic history for Indian tigers, including strong population structure and historical population bottlenecks (70, 73, 74).

Finally, the frequency distributions of putatively deleterious and neutral derived alleles reveal signatures of purging via natural selection. Some previously rare, deleterious alleles are expected to be removed from a population by random genetic drift alone, during periods of small N_e (e.g., population bottlenecks or founding events). Effective natural selection through purging is expected to eliminate deleterious derived alleles more frequently than neutral alleles or to keep them at lower frequencies than neutral derived alleles. We expected putatively deleterious derived alleles to either be absent or be less frequent (on average) than putatively neutral derived alleles if purging has occurred in our study populations. This prediction was supported via a randomization test in each population and for all populations combined ($P < 2e-4$, two-sided randomization test). This suggests that purifying selection has successfully removed a small fraction of the mutation load in each of the three study populations and across India. The higher proportion of fixed and very high-frequency derived putatively damaging alleles in the large populations and across India may be caused by historical selective sweeps on a small fraction of positively selected nonsynonymous mutations, the details of which we leave to future research.

Interestingly, both inbreeding and mutation load for the southern large–connected population (from Western Ghats in India) are higher than for the central large–connected population (from central India, Figs. 1 and 2). This could reflect the central location of this population (35, 38, 39). Additionally, the central large–connected population may have had higher historical connectivity (with other population clusters), a two-dimensional network of local populations (versus a linear array in southern large–connected populations), and other ecological factors that allow higher population densities to be achieved in all habitats of central large–connected populations compared to southern large–connected populations.

Genetic Rescue Strategies. Our data suggest that genomic inbreeding (F_{ROH} up to 0.6 in the small–isolated population) is comparable to populations experiencing inbreeding depression [e.g., Florida puma, F_{ROH} up to 0.6 (75), and Isle Royale wolves, F_{ROH} between 0.1 and 0.5 (32)]. While we also demonstrate the effects of purifying selection, homozygous LOF alleles are more frequent and mildly deleterious missense mutations are in excess in the small–isolated population.

Gene flow is the proposed strategy to sustain small and isolated populations (18). Gene flow that results in masking of deleterious alleles responsible for genetic load leading to increased population growth rate is genetic rescue. Evolutionary rescue is gene flow that increases genetic variation (76) and thus adaptive potential. High F_{ST} at loci with high-frequency damaging alleles in the receiving population (frequency >0.9) should be indicative of good source populations for genetic rescue. In practice, for both evolutionary rescue and genetic rescue, source populations are identified as those that have high genome-wide genetic differentiation with the receiving population (to maximize the heterozygosity of admixed offspring) (reviewed in ref. 19). Gene flow from populations with high genome-wide F_{ST} in theory should result in masking deleterious alleles, allowing both genetic rescue and evolutionary rescue. However, the smaller number of loci with high-frequency damaging alleles compared to all polymorphic loci could lead to discordance between these two proxies for genetic and evolutionary rescue (i.e., genetic differentiation at loci with putatively damaging alleles may not be a good predictor of genome-wide genetic differentiation).

We investigated whether evolutionary rescue is a good proxy for genetic rescue. We propose that certain source populations may be ideal if they result in masking of damaging alleles and increased genome-wide heterozygosity (Fig. 4 *A, Top Right*). Alternatively, other source populations (Fig. 4 *A, Bottom Right*) will increase heterozygosity but not mask damaging alleles, while those in Fig. 4 *A, Top Left* mask damaging alleles but may not increase heterozygosity substantially genome-wide. Our data and estimates reveal a strong and positive correlation between the two proxies (Fig. 4B, adjusted $r^2 = 0.91$), reinforcing that selecting source populations based on high genome-wide F_{ST} would result in both genetic and evolutionary rescue. For the small–isolated population, our data suggest that Kaziranga (proportion of damaging loci rescued = 0.74, genome-wide $F_{ST} = 0.26$, damaging loci $F_{ST} = 0.26$) is the best source population, while ecologically more similar populations like Kanha (proportion of damaging loci rescued = 0.69, genome-wide $F_{ST} = 0.22$, damaging loci $F_{ST} = 0.23$) and Corbett (proportion of damaging loci rescued = 0.65, genome-wide $F_{st} = 0.26$, damaging loci $F_{ST} = 0.28$) may be more practical and almost as effective at genetic rescue. For the large–isolated population in southern India, both proxies are lower on average (average proportion of damaging loci rescued = 0.65, average genome-wide $F_{ST} = 0.195$, average damaging allele $F_{ST} = 0.21$, across source populations). Note that high genome-wide F_{ST} between a donor and a recipient population could also be due to intense bottlenecks and resulting drift in endangered species,
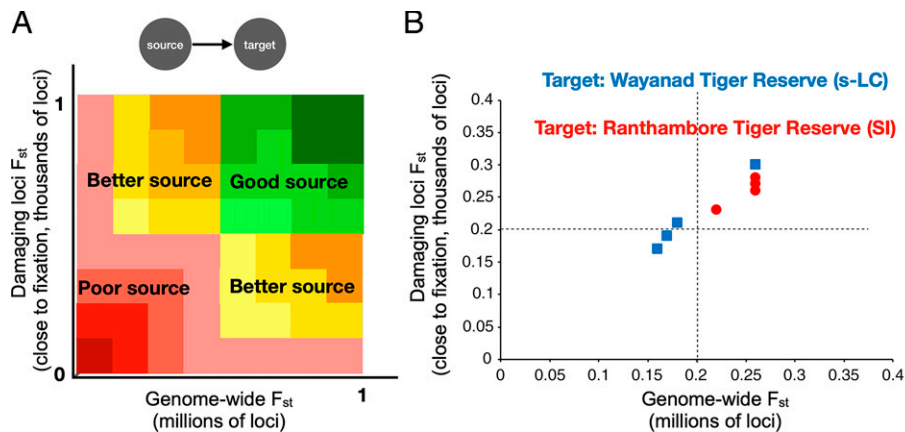
Fig. 4. (*A and B*) A schematic showing the possible distribution of genome-wide F_{ST} (*A*) and the observed values for our data (*B*), assuming two target populations, Ranthambore Tiger Reserve (in red circles, the small–isolated population) and Wayanad Tiger Reserve (blue squares, a protected area within the southern large–connected population). F_{ST} on the *y* axis is for loci where the putatively damaging allele is at a frequency >0.9 in the target population and for all loci across the genome on the *x* axis.

and such population pairs may not result in effective genetic rescue. We caution that these inferences are subject to assumed fitness effects of damaging alleles. Most genome-wide predictions of deleterious alleles in wild species await annotation with phenotypic data (77); such a strategy could be implemented when the loci affecting fitness are identified as in the case of Tasmanian devils (78) and Soay sheep (79).

Whether there is a substantial advantage to selecting source populations for genetic rescue based on differentiation for putatively damaging loci (versus across the whole genome) remains to be tested empirically. Genome-wide surveys would help in identifying source populations that best mask damaging alleles and increase genome-wide heterozygosity; however, there are other considerations (as highlighted in ref. 33) that could impact selection of source populations. A denser sampling of tigers across the Indian tiger range is necessary to predict best rescue individuals for particular target populations. Such dense sampling of genomes can benefit from sequencing of noninvasive samples since acquiring blood samples from endangered species can be difficult (80). Methods involving sequencing genomes from shed hair (81) or fecal samples (82) will be useful in creating such translocation networks and monitoring the populations for gene flow.

Implications for Conservation of Endangered Species. We find that small and isolated populations, in this case tigers of Ranthambore Tiger Reserve, have high inbreeding arising from both recent ancestors and ancestors in deep history. Our analyses revealed a signature of purging for putatively large effect (LOF) deleterious alleles in this population. However, this small and isolated population also had the highest average number of homozygous damaging alleles of the three study populations, suggesting a potential fitness cost of inbreeding despite purging of some large effect mutations and compared to the large and connected populations. Together, these results demonstrate that purging (as inferred from genomic signatures) does not eliminate all damaging alleles and hence does not contradict the small population paradigm of conservation biology (19). Careful predictions of putative phenotypic effects of the existing deleterious alleles at high frequencies might be instructive to understand future inbreeding depression, and these alleles could potentially be used as early warning signs for population decline. However, caution is warranted in this regard because it is still unclear how effectively molecular predictions (i.e., identification of loss-of-function or derived alleles at conserved sites) can translate into predictions of fitness differences among extant individuals and populations.

Historical isolation and bottlenecks due to overharvest appear to have a significant effect on missense (potentially mild effect) mutation load. Our analyses suggest that purging can also occur in large–connected populations in the context of endangered species (Fig. 3). This could be because all tiger populations are in general small, and the range of population sizes is not very wide (36). Continued isolation of these populations will increase mutation load in the future. Thus, maximizing connectivity might be the best strategy to minimize extinction (from an ecological perspective), increase genome-wide genetic variation, and also minimize mutation load in the future. Maximizing connectivity in the wild would require increasing population abundance since large populations are sources of dispersing individuals, but also planning for habitat corridors, since demographics of tiger landscapes are dominated by extinction and recolonization dynamics (83, 84). Such strategies may be relevant to several endangered species with wide ranges but fragmented populations, like elephants, lions, and wild dogs.

Critical to providing recommendations for management is understanding ongoing evolutionary trajectories for endangered species. Given the large number of frequent deleterious alleles in the small–isolated population, management could aim to decrease the number of close to fixation damaging alleles, potentially maximizing mean absolute fitness. Our data reveal that source populations with high genome-wide differentiation (with the target population) would adequately mask deleterious alleles and increase heterozygosity. Gene flow from even the relatively proximate large–connected population in central India would decrease the frequency of damaging alleles and allow genetic rescue. We caution that such gene flow could also introduce other damaging alleles (e.g., in wolves) (85).

Overall, genomic data and analyses provide richer, more nuanced ways to address inbreeding and genetic rescue of small and isolated populations of endangered species (i.e., the small population paradigm in conservation). We demonstrate effective purging of some loss-of-function alleles, but not mutations with smaller effects, highlighting the need to characterize detrimental mutations and how they will be addressed through genetic rescue in future conservation efforts. Conservation management based on predictive models for fitness effects in the context of observed mutations will aid long-term persistence of populations in the wild.

Methods

Sample Collection. We collected tissue from tranquilized or dead tigers from several tiger reserves (14 protected areas across India); zoo individuals from Sagar et al. (51) were also analyzed (Fig. 5 and *SI Appendix,* Table 1).
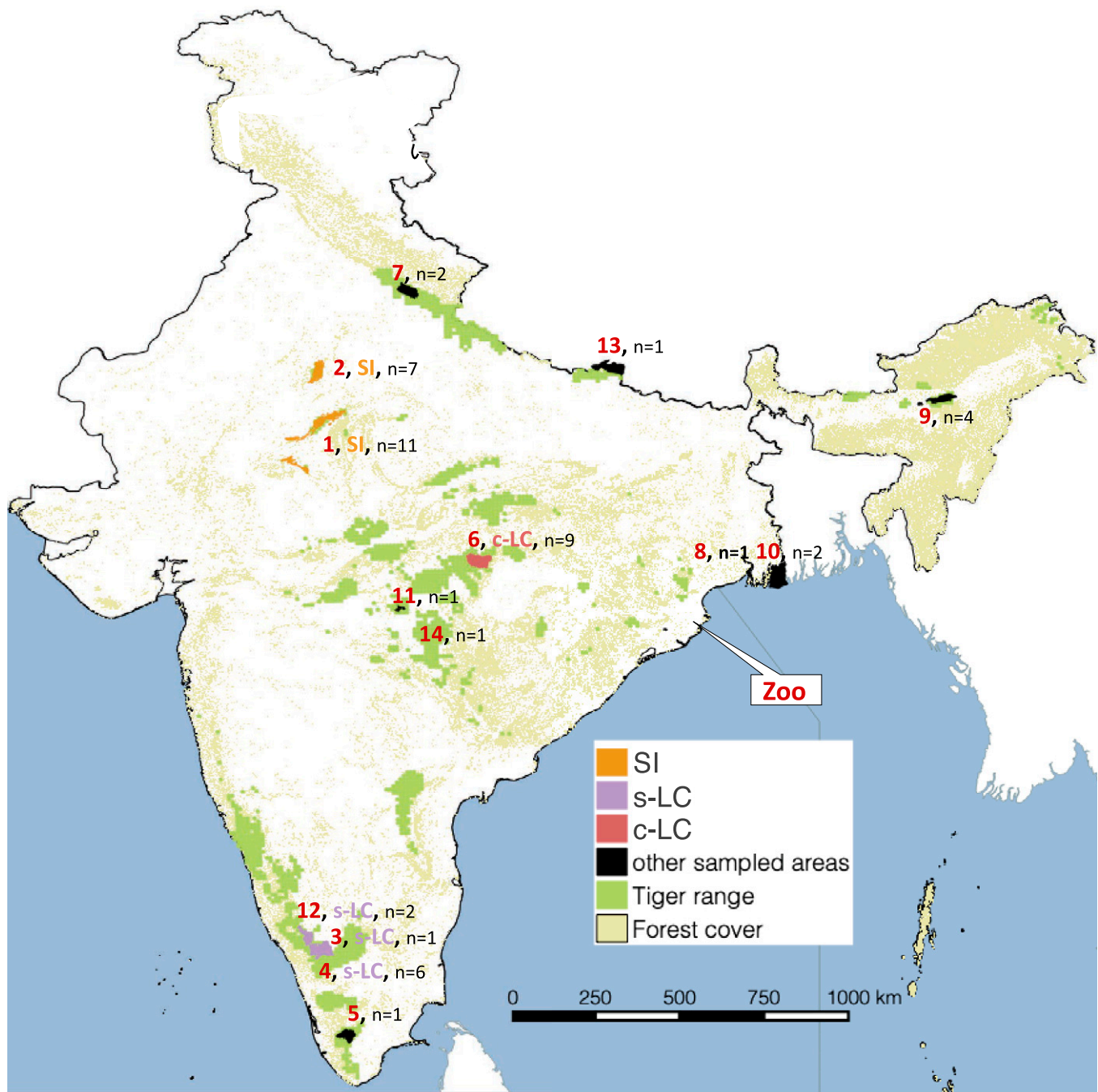
Fig. 5. Sample locations and populations. The numbers represent the protected area where, in the northwestern population, 1 is Ranthambore Tiger Reserve ($n = 11$), 2 is Sariska Tiger Reserve ($n = 7$); in the southern India population, 3 is Bandipur Tiger Reserve ($n = 3$), 4 is Wayanad Wildlife Sanctuary ($n = 6$), 12 is Nagarhole Tiger Reserve ($n = 2$), 5 is Periyar Tiger Reserve ($n = 1$); in the central cluster, 6 is Kanha Tiger Reserve ($n = 9$), 7 is Corbett Tiger Reserve ($n = 2$), 8 is Lalgarh Range ($n = 1$), 10 is Sunderban Tiger Reserve ($n = 2$), 11 is Bor Tiger Reserve ($n = 1$), 13 is Chitwan National Park ($n = 1$), 14 is Chandrapur ($n = 1$); in the northeastern population, 9 is Kaziranga Tiger Reserve ($n = 4$); and Zoo (51) is Nanadankanan Zoo ($n = 5$).

Individuals belonged to one of five population clusters: 1) From the small–isolated population in northwestern India (NW) ($n = 19$), we obtained tissues from 15 NW individuals and used data from 2 NW individuals in Armstrong et al. (34) and 2 NW individuals in Khan et al. (81). Our NW samples consist only of individuals that were born in Ranthambore Tiger Reserve (RTR) but could have been translocated to other protected areas (for example, Sariska Tiger Reserve) according to forest department records. 2) From the central large–connected population in central India (C) ($n = 9$), we obtained tissue from four Kanha Tiger Reserve (KTR) individuals and used data from Armstrong et al. (34) for five other KTR individuals. 3) From the southern large–connected population in southern India ($n = 11$), we obtained tissue from two Wayanad Wildlife Sanctuary (WWS) individuals and two Bandipur

Tiger Reserve (BTR) individuals. We additionally used data from Armstrong et al. (34) for five WWS individuals and Liu et al. (86) for two Nagarhole Tiger Reserve (NTR) individuals. Since WWS, BTR, and NTR populations are effectively contiguous, we refer to them together as Wayanad (WAY) tigers. 4) From Nandankanan Zoo, Odisha, India, Sagar et al. (51) obtained five blood samples from inbred individuals, and these genomes were used as a control to test our methods. The pedigree-based inbreeding coefficients of these individuals vary from 0.21 ($n = 1$) to 0.26 ($n = 1$) and 0.28 ($n = 3$) (*SI Appendix*, Table 1). Such pedigree-based estimates are expected to be an underestimate of the actual inbreeding due to the low depth of pedigrees (44). 5) From other protected areas, we also sampled one individual tiger from Lalgarh Division Forest, two individuals from Corbett Tiger Reserve (COR), one from Chitwan

EVOLUTION

National Park, four from Kaziranga Tiger Reserve, one from Periyar Tiger Reserve, one from Bor Tiger Reserve, one from Chandrapur district in Maharashtra, and two from Sundarban Tiger Reserve. These individuals combined with others are used for drawing pan-India inferences. The sampling locations are depicted in Fig. 5.

DNA extraction and sequencing were done using the tissue library preparation protocol described in Khan et al. (81) (*SI Appendix, SI Methods*).

Whole-Genome Resequencing and Variant Discovery. We trimmed the raw reads using TRIMMOMATIC (87) to have a mean phred-scaled quality of 30 in a sliding window of 15 bp and removed any read that was shorter than 36 bp after trimming from further analysis. We aligned these reads to a Bengal tiger reference genome (JAHFZI000000000) using BOWTIE2 (88). The alignments were then saved in a binary format (BAM) using SAMTOOLS1.9 (89). We marked duplicate reads with the Picard Tools "MarkDuplicates" command (http://broadinstitute.github.io/picard). We called variants from the BAM files using Strelka with default options (90). The variants were filtered with VCFtools (91) to retain biallelic sites with a minimum minor allele count of 3 and remove indels and loci with mean depth across individuals below the 2.5th percentile and above the 97.5th percentile across all loci. We removed all sites with missingness >20% after removing genotypes with genotype quality (GQ) less than 30. We removed the X chromosome from the analysis. A total of 1,601,148 single-nucleotide polymorphisms (SNPs) with mean read depth = 17.2 ($\pm$ 1.8) remained for analyses after these filtering steps.

Runs of Homozygosity. We identified ROH using a sliding-window, likelihood-ratio method (92, 93). For this analysis we used SNP data without the minimum genotype quality filter described above and then used genotype likelihoods, rather than called genotypes, as input to maximize the genomic information available for detecting ROH. We split each chromosome into sliding windows of 200 SNPs and step size of 10 SNPs. For each individual i and window j, we calculated the probability of the observed genotype likelihood at each SNP k (G_k) assuming the two alleles were IBD and also assuming they were non-IBD (92). We calculated the logarithm of odds (LOD) score for each window as

$$LOD(j, i) = \sum_{k=1}^{k_j} log_{10}\left(\frac{Pr(G_k|IBD)}{Pr(G_k|non-BD)}\right),$$

where LOD scores across all windows and individuals form a bimodal distribution. Windows with LOD score ≥ 0 were called IBD, while windows in the left-hand side of the distribution were called non-IBD as in Kardos et al. (44, 94). Consecutive IBD windows were joined to map the starting and stop positions of ROH for each individual. F_{ROH} was calculated as the fraction of the autosomal genome in ROH; for example,

$$F_{ROH>100kb} = \frac{\sum ROH_{>100kb}}{Total\ autosomal\ size}.$$

We calculated the percentage of the genome in ROH above a particular size class to estimate inbreeding arising from ancestors in different historical time periods. The coalescent times of ROH were estimated as $g = 100/(2rL)$, where g is the expected time (in generations) back to the parental common ancestor where the IBD haplotypes forming an ROH coalesce (95), r is the recombination rate (here, 1.9 cM/Mb based on domestic cat) (45), and L is the length of the ROH in megabases. For example, a 100-kb ROH is estimated to arise from a single haplotype in an ancestor 263 generations ago on average. An individual's genomic inbreeding arising from ancestors up to 263 generations ago is then estimated using only ROH longer than 100 kb. Similar calculations were performed for other ROH lengths. Table 1 lists the lengths of ROH and corresponding average estimated coalescent times for the IBD haplotypes that form the ROH.

We constructed percentile bootstrap confidence intervals for mean F_{ROH} in each population (96). We randomly resampled individuals within a population with replacement 10,000 times, each time calculating the mean F_{ROH} as

described above. The 0.025 and 0.975 quantiles defined the 95% confidence interval for mean F_{ROH} in each population. We used a randomization procedure to test for a difference in the mean F_{ROH} for each pair of populations as follows. First, we randomized the population affiliation for each individual 10,000 times. For each randomization replicate, we calculated the mean F_{ROH} for each of the two populations. We calculated the P value as the proportion of 10,000 randomization replicates where the difference in F_{ROH} was at least as large as the empirical difference in F_{ROH}.

IBD Stretches of Genome Shared between Pairs of Individuals. We estimated stretches of genome IBD between pairs of individuals using IBDseq default parameters (97). We filtered for shared stretches longer than 1 Mb and estimated the sum of shared lengths of more than 1 Mb per autosomal genome size. We estimated this for all pairs of individuals within a population and all pairs across populations.

Estimating Effective Population Size from ROH. We used the estimates of F_{ROH} using ROH arising from ancestors in four different estimated time frames (Fig. 1) to estimate N_e over each respective time frame. If the estimated coalescent times for ROH are unbiased, then the average F_{ROH} based on ROH with estimated maximum coalescent times less than t generations back in time is an estimator of the inbreeding accumulated in the population from the time of sampling back to t generations ago. We estimated N_e from the following expression of mean expected individual inbreeding as a function of N_e over t generations:

$$F_{ROH,t} = 1 - \left(1 - \frac{1}{2N_e}\right)^t.$$

This approach assumes that the probability of inferring an ROH segment due to several smaller homozygous segments is very low. We calculated 95% confidence intervals for N_e by bootstrapping over individuals within each population as described above for our estimates of mean F_{ROH}.

Mitochondrial Haplotype Diversity. We used the reads aligned to the mitochondrial scaffold of BenTig1.0. We called consensus fasta files for each individual from this as described in Khan et al. (81). These multiple consensus fasta files were aligned using clustal omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) and written to nexus output files. We then used DnaSP 6 (98) to estimate haplotype diversity.

Identifying Effect of Mutations. We refiltered the SNPs as described above with VCFtools (91), except here we removed genotypes with phred-scaled genotype quality <30 and then removed loci that were monomorphic, had >20% missing data, or mean read depth in the bottom and top 2.5% quantiles across individuals. We also removed loci with extreme values of F_{IS} ($F_{IS} < -0.5$ or $F_{IS} > 0.95$) estimated across all Indian tigers to remove loci that likely have frequent genotyping or read mapping errors; this expectedly resulted in mean $F_{IS} > 0$ (0.14).

We used Ensembl Variant Effect Predictor (59) on SNPs to identify the consequence of a mutation and classify them as missense, LOF, or intergenic as described by Xue et al. (28). We used annotation files generated for the Bengal tiger assembly using RNA seq.

Mutation Load. We used the R_{XY} method described by Do et al. (62) and implemented by Xue et al. (28) to estimate relative excess of mutation loads in a population X with respect to another population Y. We randomly selected seven individuals from each population and estimated the relative mutation load for each of the three possible pairwise combinations of the SI, s-LC, and c-LC populations. SDs were obtained by 100 rounds of jackknife. For each round of jackknife for missense and LOF mutations, we randomly excluded 10 to 15% of loci and repeated the estimation. For intergenic variants, 100,000 loci were randomly excluded for each round.

To infer the ancestral state of each site in the tiger reference we converted the references for the domestic cat (GenBank: GCA_000181335.4), cheetah

Table 1. ROH lengths and the corresponding expected mean time of their origin due to inbreeding where number of years assumes a generation length of 5 y

ROH length(s)	Generations since parents shared ancestor	Years since parents shared ancestor
<100 kb	More than 263	More than 1,315
>100 kb	Up to 263	Up to 1,315
>1 Mb	Up to 26	Up to 130
>5 Mb	Up to 5	Up to 25
>10 Mb	Up to 3	Up to 15

(GCA_003709585.1), and lion (GCA_008795835.1) into FASTQ reads by sliding across the genome in nonoverlapping windows of 100 bp and transforming each window into a separate FASTQ read. The resulting FASTQ reads were then mapped to the tiger reference genome with bwa mem v0.7.17, slightly lowering the mismatch penalty ($-B$ 3) and removing reads that mapped to multiple regions. Mapped reads were realigned around indels using GATK (https://gatk.broadinstitute.org/hc/en-us) IndelRealigner. Next, we converted the mapped reads into a haploid FASTA consensus sequence, excluding all sites with depth above one (as such sites contain at least one mismapped read) using ANGSD (http://www.popgen.dk/angsd/index.php/ANGSD) -dofasta. The ancestral allele at a locus was then determined as the majority allele found in genomic alignment of the domestic cat, cheetah, and lion. Sites where the ancestral allele could not be identified (e.g., where the ancestral allele identified in the cat–cheetah–lion alignment was not present in our sample of tigers) were excluded.

We counted the number of homozygous deleterious alleles per individual as number of derived missense and LOF mutations. We quantified uncertainty in the mean number of homozygous deleterious alleles for each population using percentile bootstrap confidence intervals as described above. We used randomization tests as described above to test for differences in the mean number of homozygous damaging and LOF alleles per individual for each pair of populations.

Site-Frequency Spectrum. We equalized the sample sizes across the populations and across loci within each population by randomly subsampling 16 non-missing alleles from each locus in each population before estimating the derived neutral and derived damaging allele frequencies. For the SFS analysis of all populations combined, we subsampled 88 nonmissing alleles from each locus. Repeating the subsampling process and analysis had no effect on the results.

We used randomization tests to determine whether the mean putatively damaging derived allele frequency was statistically significantly different from the mean putatively neutral derived allele frequency in each population and for all populations combined (Fig. 3). We randomly reassigned the neutral versus damaging status for each allele frequency 5,000 times. For each of the 5,000 replicates, we recalculated the randomized mean derived neutral and

damaging allele frequency. A P value was calculated as the proportion of the 5,000 permutations that resulted in a difference in the mean neutral versus damaging derived allele frequency that was larger than observed in the nonrandomized data.

Genetic Rescue and F_{ST}. We estimated genome-wide genetic differentiation as F_{ST} with VCFtools. For loci with high-frequency (>0.85, closer to fixation) deleterious (missense and LOF) alleles in a population, we listed all the loci in each target population (population in need of rescue). For every possible source population, we estimated the proportion of these loci with frequency <0.85. We then calculated F_{ST} at only these loci to estimate damaging loci F_{ST}. The proportion of loci rescued was calculated as the fraction of loci that had <0.85 frequency in the possible source population.

We investigate this for five possible source populations, including tigers from KTR, WAY, RTR, COR, and Kaziranga Tiger Reserve (KAZ).

1. N. M. Haddad et al., Habitat fragmentation and its lasting impact on Earth's ecosystems. *Sci. Adv.* **1**, e1500052 (2015).
2. D. R. Schlaepfer, B. Braschler, H.-P. Rusterholz, B. Baur, Genetic effects of anthropogenic habitat fragmentation on remnant animal and plant populations: A meta-analysis. *Ecosphere* **9**, e02488 (2018).
3. D. Spielman, B. W. Brook, R. Frankham, Most species are not driven to extinction before genetic factors impact them. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 15261–15264 (2004).
4. R. Lande, Risk of population extinction from fixation of new deleterious mutations. *Evolution* **48**, 1460–1469 (1994).
5. R. Lande, Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *Am. Nat.* **142**, 911–927 (1993).
6. R. Lande, G. Barrowclough, "Effective population size, genetic variation, and their use in population" in *Viable Populations for Conservation*, M. E. Soulé, Ed. (Cambridge University Press, Cambridge, 1987), pp. 87–123.
7. M. Lynch, J. Conery, R. Burger, Mutation accumulation and the extinction of small populations. *Am. Nat.* **146**, 489–518 (1995).
8. D. Newman, D. Pilson, Increased probability of extinction due to decreased genetic effective population size: Experimental populations of *Clarkia pulchella. Evolution* **51**, 354–362 (1997).
9. R. Burrows, H. Hofer, M. L. East, Demography, extinction and intervention in a small population: The case of the Serengeti wild dogs. *Proc. Biol. Sci.* **256**, 281–292 (1994).
10. R. Woodroffe, C. Sillero-Zubiri, *Lycaon Pictus (Amended Version of 2012 assessment). The IUCN Red List of Threatened Species 2020* (IUCN, 2020).
11. J. Marino, C. Sillero-Zubiri, *Canis simensis. The IUCN Red List of Threatened Species 2011* (IUCN, 2011).
12. S. Dutta, A. R. Rahmani, Y. V. Jhala, Running out of time? The great Indian bustard *Ardeotis nigriceps*—status, viability, and conservation strategies. *Eur. J. Wildl. Res.* **57**, 615–625 (2011).
13. G. Caughley, Directions in conservation biology. *J. Anim. Ecol.* **63**, 215 (1994).
14. P. W. Hedrick, R. C. Lacy, F. W. Allendorf, M. E. Soule, Directions in conservation biology: Comments on Caughley. *Conserv. Biol.* **10**, 1312–1320 (1996).
15. K. Ralls, P. Sunnucks, R. C. Lacy, R. Frankham, Genetic rescue: A critique of the evidence supports maximizing genetic diversity rather than minimizing the introduction of putatively harmful genetic variation. *Biol. Conserv.* **251**, 108784 (2020).
16. D. Charlesworth, B. Charlesworth, Inbreeding depression and its evolutionary consequences. *Annu. Rev. Ecol. Syst.* **18**, 237–268 (1987).
17. L. F. Keller, D. M. Waller, Inbreeding effects in wild populations. *Trends Ecol. Evol.* **17**, 230–241 (2002).
18. A. R. Whiteley, S. W. Fitzpatrick, W. C. Funk, D. A. Tallmon, Genetic rescue to the rescue. *Trends Ecol. Evol.* **30**, 42–49 (2015).
19. R. Frankham, Genetic rescue of small inbred populations: Meta-analysis reveals large and consistent benefits of gene flow. *Mol. Ecol.* **24**, 2610–2618 (2015).
20. P. W. Hedrick, Purging inbreeding depression and the probability of extinction: Full-sib mating. *Heredity* **73**, 363–372 (1994).
21. P. W. Hedrick, A. Garcia-Dorado, Understanding inbreeding depression, purging, and genetic rescue. *Trends Ecol. Evol.* **31**, 940–952 (2016).
22. A. García-Dorado, Understanding and predicting the fitness decline of shrunk populations: Inbreeding, purging, mutation, and standard selection. *Genetics* **190**, 1461–1476 (2012).
23. J. D. Ballou, Ancestral inbreeding only minimally affects inbreeding depression in mammalian populations. *J. Hered.* **88**, 169–178 (1997).
24. E. H. Boakes, J. Wang, W. Amos, An investigation of inbreeding depression and purging in captive pedigreed populations. *Heredity* **98**, 172–182 (2007).
25. V. Avila, C. Amador, A. García-Dorado, The purge of genetic load through restricted panmixia in a *Drosophila* experiment. *J. Evol. Biol.* **23**, 1937–1946 (2010).
26. N. Pekkala, K. E. Knott, J. S. Kotiaho, K. Nissinen, M. Puurtinen, The benefits of interpopulation hybridization diminish with increasing divergence of small populations. *J. Evol. Biol.* **25**, 2181–2183 (2012).
27. E. López-Cortegano, A. Vilas, A. Caballero, A. García-Dorado, Estimation of genetic purging under competitive conditions. *Evolution* **70**, 1856–1870 (2016).
28. Y. Xue et al., Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding. *Science* **348**, 242–245 (2015).
29. J. A. Robinson, C. Brown, B. Y. Kim, K. E. Lohmueller, R. K. Wayne, Purging of strongly deleterious mutations explains long-term persistence and absence of inbreeding depression in island foxes. *Curr. Biol.* **28**, 3487–3494.e4 (2018).
30. C. Grossen, F. Guillaume, L. F. Keller, D. Croll, Purging of highly deleterious mutations through severe bottlenecks in Alpine ibex. *Nat. Commun.* **11**, 1001 (2020).
31. J. A. Robinson et al., Genomic flatlining in the endangered island fox. *Curr. Biol.* **26**, 1183–1189 (2016).
32. J. A. Robinson et al., Genomic signatures of extensive inbreeding in Isle Royale wolves, a population on the threshold of extinction. *Sci. Adv.* **5**, eaau0757 (2019).
33. C. C. Kyriazis, R. K. Wayne, K. E. Lohmueller, Strongly deleterious mutations are a primary determinant of extinction risk due to inbreeding depression. *Evol. Lett.* **5**, 33–47 (2020).
34. E. E. Armstrong et al., Recent evolutionary history of tigers highlights contrasting roles of genetic drift and selection. *Mol. Biol. Evol.* **38**, 2366–2379 (2021).
35. N. Alcala, A. Goldberg, U. Ramakrishnan, N. A. Rosenberg, Coalescent theory of migration network motifs. *Mol. Biol. Evol.* **36**, 2358–2374 (2019).

EVOLUTION

Khan et al.
Genomic evidence for inbreeding depression and purging of deleterious genetic variation in Indian tigers

www.manaraa.com

36. Y. V. Jhala, Q. Qureshi, A. K. Nayak, *Status of Tigers, Co-Predators and Prey in India* (National Tiger Conservation Authority, Government of India, New Delhi, and Wildlife Institute of India, Dehradun, India, 2020).

37. Q. Qureshi *et al.*, *Connecting Tiger Populations for Long-Term Conservation* (National Tiger Conservation Authority & Wildlife Institute of India, Dehradun, India, 2014).

38. V. Kolipakam, S. Singh, B. Pant, Q. Qureshi, Y. V. Jhala, Genetic structure of tigers (*Panthera tigris tigris*) in India and its implications for conservation. *Glob. Ecol. Conserv.* 20, e00710 (2019).

39. M. Natesh *et al.*, Conservation priorities for endangered Indian tigers through a genomic lens. *Sci. Rep.* 7, 9614 (2017).

40. M. Rangarajan, *India's Wildlife History: An Introduction* (Orient Blackswan, 2005).

41. J. Sramek, "Face him like a Briton": Tiger hunting, imperialism, and British masculinity in colonial India, 1800-1875. *Vic. Stud.* 48, 659–680 (2006).

42. P. Singh, G. V. Reddy, Lost Tigers Plundered Forests: A *Report Tracing* the *Decline* of the *Tiger Across* the *State* of Rajasthan (1900 to Present) (WWF India, (2016).

43. S. Mondol, M. W. Bruford, U. Ramakrishnan, Demographic loss, genetic structure and the conservation implications for Indian tigers. *Proc. Biol. Sci.* 280, 20130496 (2013).

44. M. Kardos *et al.*, Genomic consequences of intensive inbreeding in an isolated wolf population. *Nat. Ecol. Evol.* 2, 124–131 (2018).

45. G. Li *et al.*, A high-resolution SNP array-based linkage map anchors a new domestic cat draft genome assembly and provides detailed patterns of recombination. *G3 (Bethesda)* 6, 1607–1616 (2016).

46. K. Ralls, K. Brugger, J. Ballou, Inbreeding and juvenile mortality in small populations of ungulates. *Science* 206, 1101–1103 (1979).

47. K. Ralls, J. D. Ballou, A. Templeton, Estimates of lethal equivalents and the cost of inbreeding in mammals. *Conserv. Biol.* 2, 185–193 (1988).

48. L. Laikre, N. Ryman, Inbreeding depression in a captive wolf (*Canis lupus*) population. *Conserv. Biol.* 5, 33–40 (1991).

49. T. van der Valk, M. Manuel, T. Marques-Bonet, K. Guschanski, Estimates of genetic load in small populations suggest extensive purging of deleterious alleles. *bioRxiv* [Preprint] (2019). https://doi.org/10.1101/696831 (Accessed 22 November 2021).

50. A. K. Roychoudhury, K. S. Sankhala, Inbreeding in white tigers. Proc. Indian Acad. Sci. *Anim. Sci.* 88, 311–323 (1979).

51. V. Sagar *et al.*, High frequency of an otherwise rare phenotype in a small and isolated tiger population. *Proc. Natl. Acad. Sci. U.S.A.* 118, e2025273118 (2021).

52. R. Frankham, C. J. A. Bradshaw, B. W. Brook, Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biol. Conserv.* 170, 56–63 (2014).

53. I. G. Jamieson, F. W. Allendorf, How does the 50/500 rule apply to MVPs? *Trends Ecol. Evol.* 27, 578–584 (2012).

54. H. Bauer, S. Van Der Merwe, Inventory of free-ranging lions *Panthera leo* in Africa. *Oryx* 38, 26–31 (2004).

55. J. L. Fox, S. P. Sinha, R. S. Chundawat, P. K. Das, Status of the snow leopard Panthera uncia in Northwest India. *Biol. Conserv.* 55, 283–298 (1991).

56. R. M. Jackson, J. D. Roe, R. Wangchuk, D. O. Hunter, Estimating snow leopard population abundance using photography and capture–recapture techniques. *Wildl. Soc. Bull.* 34, 772–781 (2006).

57. S. C. Silver *et al.*, The use of camera traps for estimating jaguar *Panthera onca* abundance and density using capture/recapture analysis. *Oryx* 38, 148–154 (2004).

58. J. R. Adams, L. M. Vucetich, P. W. Hedrick, R. O. Peterson, J. A. Vucetich, Genomic sweep and potential genetic rescue during limiting environmental conditions in an isolated wolf population. *Proc. Biol. Sci.* 278, 3336–3344 (2011).

59. W. McLaren *et al.*, The Ensembl Variant Effect Predictor. *Genome Biol.* 17, 122 (2016).

60. A. García-Dorado, A. Caballero, On the average coefficient of dominance of deleterious spontaneous mutations. *Genetics* 155, 1991–2001 (2000).

61. A. F. Agrawal, M. C. Whitlock, Inferences about the distribution of dominance drawn from yeast gene knockout data. *Genetics* 187, 553–566 (2011).

62. R. Do *et al.*, No evidence that selection has been less effective at removing deleterious mutations in Europeans than in Africans. *Nat. Genet.* 47, 126–131 (2015).

63. B. E. Housden *et al.*, Loss-of-function genetic tools for animal models: Cross-species and cross-platform differences. *Nat. Rev. Genet.* 18, 24–40 (2017).

64. G. M. Cooper *et al.*, NISC Comparative Sequencing Program, Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res.* 15, 901–913 (2005).

65. C. D. Huber, B. Y. Kim, C. D. Marsden, K. E. Lohmueller, Determining the factors driving selective effects of new nonsynonymous mutations. *Proc. Natl. Acad. Sci. U.S.A.* 114, 4465–4470 (2017).

66. A. Eyre-Walker, P. D. Keightley, The distribution of fitness effects of new mutations. *Nat. Rev. Genet.* 8, 610–618 (2007).

67. S. Narain, H. S. Panwar, M. Gadgil, V. Thapar, S. Singh, *Joining the Dots: The Report of the Tiger Task Force. Project Tiger Directorate* (Union Ministry of Environment, Government of India, 2005).

68. Y. V. Jhala *et al.*, Asiatic lion: Ecology, economics, and politics of conservation. *Front. Ecol. Evol.* 7, 312 (2019).

69. A. Sadhu *et al.*, Demography of a small, isolated tiger (*Panthera tigris tigris*) population in a semi-arid region of western India. *BMC Zool.* 2, 16 (2017).

70. G. Luikart, F. W. Allendorf, J. M. Cornuet, W. B. Sherwin, Distortion of allele frequency distributions provides a test for recent population bottlenecks. *J. Hered.* 89, 238–247 (1998).

71. A. Benazzo *et al.*, Survival and divergence in a small group: The extraordinary genomic history of the endangered Apennine brown bear stragglers. *Proc. Natl. Acad. Sci. U.S.A.* 114, E9589–E9597 (2017).

72. J. Räikkönen, J. A. Vucetich, R. O. Peterson, M. P. Nelson, Congenital bone deformities and the inbred wolves (*Canis lupus*) of Isle Royale. *Biol. Conserv.* 142, 1025–1031 (2009).

73. A. Keinan, J. C. Mullikin, N. Patterson, D. Reich, Measurement of the human allele frequency spectrum demonstrates greater genetic drift in East Asians than in Europeans. *Nat. Genet.* 39, 1251–1255 (2007).

74. G. V. Barroso, A. F. Moutinho, J. Y. Dutheil, "A population genomics lexicon" in *Statistical Population Genomics*, J. Y. Dutheil, Ed. (Humana, New York, 2020), pp. 3–17.

75. N. F. Saremi *et al.*, Author correction: Puma genomes from North and South America provide insights into the genomic consequences of inbreeding. *Nat. Commun.* 10, 5276 (2019).

76. A. A. Hoffmann, A. D. Miller, A. R. Weeks, Genetic mixing for population management: From genetic rescue to provenancing. *Evol. Appl.* 14, 634–652 (2020).

77. M. Kardos, A. B. A. Shafer, The peril of gene-targeted conservation. *Trends Ecol. Evol.* 33, 827–839 (2018).

78. P. A. Hohenlohe *et al.*, Conserving adaptive potential: Lessons from Tasmanian devils and their transmissible cancer. *Conserv. Genet.* 20, 81–87 (2019).

79. M. A. Stoffel, S. E. Johnston, J. G. Pilkington, J. M. Pemberton, Genetic architecture and lifetime dynamics of inbreeding depression in a wild mammal. *Nat. Commun.* 12, 2972 (2021).

80. A. Khan, A. Tyagi, Considerations for initiating a wildlife genomics research project in south and south-east Asia. *J. Indian Inst. Sci.* 101, 243–256 (2021).

81. A. Khan *et al.*, Are shed hair genomes the most effective noninvasive resource for estimating relationships in the wild? *Ecol. Evol.* 10, 4583–4594 (2020).

82. K. L. Chiou, C. M. Bergey, Methylation-based enrichment facilitates low-cost, noninvasive genomic scale sequencing of populations from feces. *Sci. Rep.* 8, 1975 (2018).

83. P. Thatte, A. Joshi, S. Vaidyanathan, E. Landguth, U. Ramakrishnan, Maintaining tiger connectivity and minimizing extinction into the next century: Insights from landscape genetics and spatially-explicit simulations. *Biol. Conserv.* 218, 181–191 (2018).

84. B. Yumnam *et al.*, Prioritizing tiger conservation through landscape genetics and habitat linkages. *PLoS One* 9, e111207 (2014).

85. P. W. Hedrick, J. A. Robinson, R. O. Peterson, J. A. Vucetich, Genetics and extinction and the example of Isle Royale wolves. *Anim. Conserv.* 22, 302–309 (2019).

86. Y. C. Liu *et al.*, Genome-wide evolutionary analysis of natural history and adaptation in the world's tigers. *Curr. Biol.* 28, 3840–3849.e6 (2018).

87. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120 (2014).

88. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359 (2012).

89. H. Li *et al.*, 1000 Genome Project Data Processing Subgroup, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079 (2009).

90. C. T. Saunders *et al.*, Strelka: Accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* 28, 1811–1817 (2012).

91. P. Danecek *et al.*, 1000 Genomes Project Analysis Group, The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158 (2011).

92. S. Wang, C. Haynes, F. Barany, J. Ott, Genome-wide autozygosity mapping in human populations. *Genet. Epidemiol.* 33, 172–180 (2009).

93. T. J. Pemberton *et al.*, Genomic patterns of homozygosity in worldwide human populations. *Am. J. Hum. Genet.* 91, 275–292 (2012).

94. M. Kardos, A. Qvarnström, H. Ellegren, Inferring individual inbreeding and demographic history from segments of identity by descent in *Ficedula* flycatcher genome sequences. *Genetics* 205, 1319–1334 (2017).

95. E. A. Thompson, Identity by descent: Variation in meiosis, across genomes, and in populations. *Genetics* 194, 301–326 (2013).

96. B. Efron, R. J. Tibshirani, *An Introduction to the Bootstrap* (CRC Press, 1994).

97. B. L. Browning, S. R. Browning, Detecting identity by descent and estimating genotype error rates in sequence data. *Am. J. Hum. Genet.* 93, 840–851 (2013).

98. J. Rozas *et al.*, DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34, 3299–3302 (2017).

Khan et al.
Genomic evidence for inbreeding depression and purging of deleterious genetic variation in Indian tigers

www.manaraa.com