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

Anubhab Khan[1]*, Kaushal Patel[1], Harsh Shukla[1], Ashwin Viswanathan[1,2], Tom van der Valk [3], Udayan Borthakur[4], Parag Nigam[5], Arun Zachariah[6], Yadavendradev V. Jhala[5], Marty Kardos[7]*, Uma Ramakrishnan[1,8]*

Supplementary Methods:

DNA extraction and sequencing

We extracted DNA using Qiagen blood and tissue extraction kit (Cat. 69504) as per the manufacturer's instructions. We prepared DNA whole genome sequencing libraries using NEBNext UltraTM II DNA Library Prep Kit (Cat. E7645L, NEB Inc.). We quantified DNA on QubitTM 3.0 fluorometer using Qubit High sensitivity dsDNA Assay (#Q32854, Thermo Fisher Scientific). We then fragmented the quantified DNA by sonication using Covaris LE220 ultrasonicator and Covaris microTUBE (#520053, Covaris® Inc.) to obtain a final insert size of 250–350 bp. Next, we performed end-repair on the DNA fragments where blunt ends are created on either side of the fragments. We added a single "A" nucleotide on the 3′ ends of the fragments to facilitate the ligation of NEB stem-loop adapters. We cleaned the ligated products and size selected using Agencourt AMPure XP beads (#A63882, Beckman Coulter). We amplified these size selected products using an eight-cycle PCR during which indices (Barcodes) and flow-cell binding sequences were added. After a final clean-up with Agencourt AMPure XP beads, we quantified the libraries using Qubit DNA Assay and assessed the fragments using DNA TapeStation D1000 Screen Tape (#5067-5582,5583, Agilent Technologies). We finally clonally amplified the quantified libraries on a cBOT and sequenced them on the HiSeq X with 150bp paired end chemistry.

Identifying potential phenotypic effects of deleterious mutations

We aimed to use detailed functional annotations for the domestic cat genome to identify putatively deleterious mutations that may affect tiger fitness via diseases with a known genetic basis in domestic cats. We therefore mapped the tiger whole genome sequencing reads to the domestic cat reference genome. The domestic cat genome is annotated functionally and the disease consequences of several mutations are already known (Supplementary Table 2). We then identified putatively deleterious tiger alleles in genes with known disease-causing mutations in domestic cats.

Mapping to domestic genome and variant discovery for predicting potential diseases

We mapped trimmed FASTQ reads to the *Felis catus* reference genome version 9 (GenBank: GCA_000181335.4) using BWA-MEM[1]. The alignments were then saved in a binary format (BAM) using SAMTOOLS1.9[2]. 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[3]. The variants were filtered with VCFtools[4] to retain biallelic sites with a minimum minor allele count of 3, minimum base quality 30, minimum depth 10, minimum genotype quality 30, removed Indels and had less than 20% missing data.

We used Ensembl Variant Effect Predictor[5] on SNPs identified with domestic cat reference genome to identify consequence of a mutation and classify them as missense, loss-of-function (LOF) or intergenic as described by Xue et al.[6]. We then compared these mutations with an annotated dataset of known diseasing cause mutations in domestic cats.

GERP based Mutation load

We estimated GERP scores for each locus on our reference tiger genome as described in van der Valk et al.[7]. We selected the loci with the top 0.1 percent GERP scores from the variant call file filtered from the tiger genome. These are potentially loci hosting highly deleterious alleles. We selected the loci with the least 1 percent GERP scores from the same file. These are potentially neutral loci. Next we used the R_{XY} method described by Do et al.[8] and implemented by Xue et al.[6] to estimate relative excess of mutation loads as described in the main text.

Supplementary table 1: Sample sequencing and location. north-west Indian: 1=Ranthambore Tiger Reserve (n=11), 2=Sariska Tiger Reserve (n=7), south Indian population: 3=Bandipur Tiger Reserve (n=2), 4=Wayanad Wildlife Sanctuary (n=6), 5=Periyar Tiger Reserve (n=1), central Indian Population: 6=Kanha Tiger Reserve (n=8), 7=Corbett Tiger Reserve (n=2), 8=Lalgarh Range (n=1), 10=Sunderban Tiger Reserve (n=2), 11=Bor Tiger Reserve (n=1),north-east Indian population: 9=Kaziranga Tiger Reserve (n=3) , Zoo: Zoo=Nanadankanan Zoo (n=5).

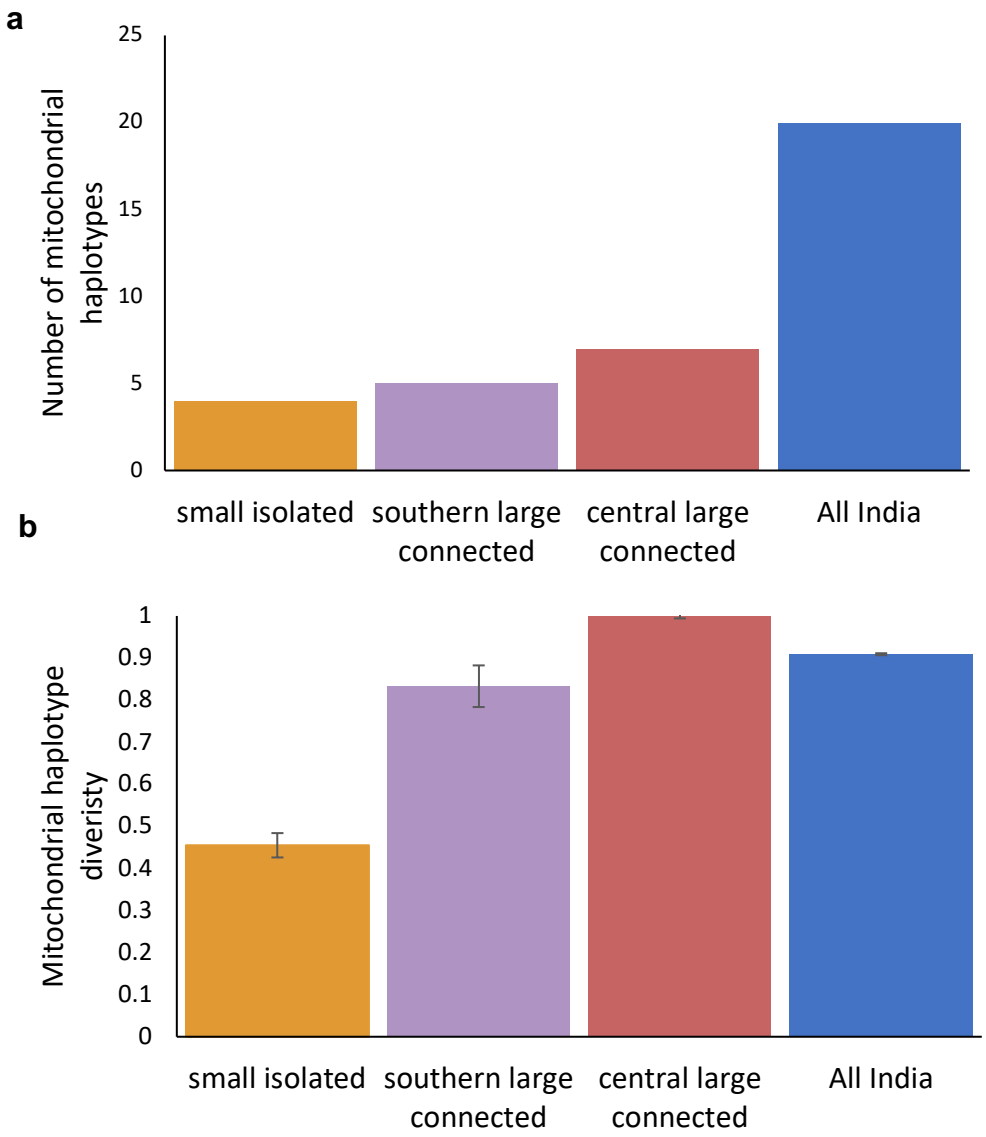
Individual	Location	Sequencing depth	F_{PED}	Proportion/genome [>100Kb]	Proportion/genome [>1Mb]	Biosample accession
BEN_CI15	6	24		0.316	0.091	SRR15369235[1]
BEN_CI16	6	15		0.476	0.304	SRR15369234[1]
BEN_CI18	6	26		0.303	0.080	SRR15369223[1]
BEN_CI19	6	13		0.346	0.146	SRR15369214[1]
BEN_CI2	14	8		0.398	0.207	SRR13500762[2]
BEN_CI21	11	18		0.411	0.197	SRR15369213[1]
BEN_CI3	6	22		0.345	0.123	SRR13500761[2]
BEN_CI4	6	13		0.332	0.125	SRR13500750[2]
BEN_CI5	6	15		0.293	0.063	SRR13500748[2]
BEN_CI6	6	11		0.328	0.105	SRR13500747[2]
BEN_CI7	6	8		0.428	0.253	SRR13500746[2]
BEN_NE1	9	22		0.330	0.097	SRR13500744[2]
BEN_NE2	9	18		0.605	0.469	SRR13500743[2]
BEN_NE3	9	12		0.523	0.363	SRR13500742[2]
BEN_NE4	9	8		0.373	0.229	SRR15369212[1]
BEN_NOR_SJ1	13	12		0.414	0.202	SRR7152401[3]
BEN_NOR1	7	23		0.400	0.159	SRR13500760[2]
BEN_NOR2	7	24		0.453	0.234	SRR13500759[2]
BEN_NW10	1	27		0.620	0.519	SRR15369211[1]
BEN_NW11	2	23		0.495	0.360	SRR15369210[1]
BEN_NW12	2	16		0.608	0.499	SRR15369209[1]
BEN_NW13	2	23		0.577	0.470	SRR15369208[1]
BEN_NW14	2	23		0.586	0.478	SRR15369233[1]
BEN_NW17	1	14		0.556	0.430	SRR15369232[1]
BEN_NW18	1	21		0.633	0.530	SRR15369231[1]
BEN_NW19	1	23		0.524	0.386	SRR15369230[1]
BEN_NW20	1	18		0.554	0.445	SRR10013450[4]
BEN_NW24	1	23		0.569	0.454	SRR15369229[1]
BEN_NW3	2	20		0.609	0.495	SRR13500757[2]
BEN_NW4	2	20		0.592	0.490	SRR13500756[2]
BEN_NW5	1	34		0.483	0.342	SRR9944891[4]
BEN_NW55	2	18		0.571	0.456	SRR15369228[1]
BEN_NW57	1	15		0.625	0.519	SRR15369227[1]
BEN_NW6	1	15		0.491	0.356	SRR15369226[1]

BEN_NW63	1	15		0.632	0.529	SRR15369225[1]
BEN_NW8	1	16		0.589	0.474	SRR15369224[1]
BEN_NW9	1	21		0.593	0.482	SRR15369222[1]
BEN_SI1	4	21		0.461	0.249	SRR13500755[2]
BEN_SI11	4	12		0.423	0.220	SRR15369221[1]
BEN_SI18	3	16		0.436	0.232	SRR15369220[1]
BEN_SI19	3	17		0.447	0.238	SRR15369219[1]
BEN_SI2	3	6		0.432	0.244	SRR13500754[2]
BEN_SI3	4	16		0.487	0.287	SRR13500753[2]
BEN_SI5	5	10		0.668	0.532	SRR13500751[2]
BEN_SI6	4	20		0.479	0.280	SRR13500749[2]
BEN_SI8	4	25		0.552	0.376	SRR13500752[2]
BEN_SI9	4	23		0.442	0.218	SRR15369218[1]
BEN_SI_SJ1	12	11		0.423	0.226	SRR7152380[3]
BEN_SI_SJ2	12	11		0.370	0.144	SRR7152381[3]
BEN_SU2	10	15		0.470	0.273	SRR15369217[1]
BEN_SU3	10	41		0.565	0.403	SRR15369216[1]
LGS1	8	15		0.408	0.208	SRR15369215[1]
ZSB_01	Zoo	20	0.21	0.389	0.190	SRR15292763[5]
ZSB_02	Zoo	20	0.28	0.482	0.329	SRR15292762[5]
ZSB_03	Zoo	29	0.28	0.468	0.307	SRR15292761[5]
ZSB_04	Zoo	26	0.28	0.466	0.309	SRR15292760[5]
ZSB_05	Zoo	30	0.26	0.451	0.281	SRR15292759[5]

1) This study
2) 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).
3) Y.-C. Liu, *et al.*, Genome-Wide Evolutionary Analysis of Natural History and Adaptation in the World's Tigers. *Current Biology* **28**, 3840–3849.e6 (2018).
4) 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).
5) V. Sagar, *et al.*, High frequency of an otherwise rare phenotype in a small and isolated tiger population. *PNAS* **118**, 39 e2025273118 (2021).

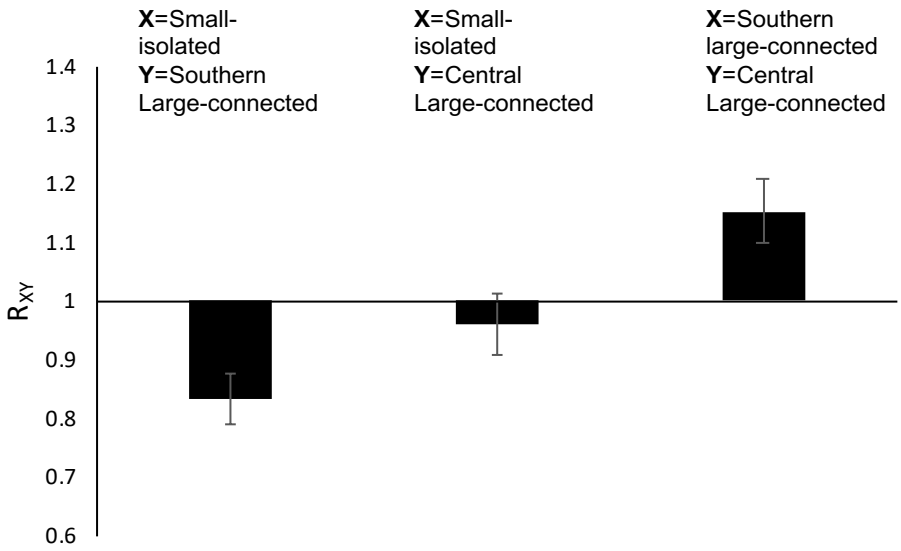
Supplementary Table 2: Genes with derived missense mutations and potential diseases predicted with domestic cat annotations

Gene	Disease	Reference
MYBPC3	Hypertrophic Cardiomyopathy	9,10
MYBPC2	Hypertrophic Cardiomyopathy	9,10
WNK4	Hypokalemia	11
CEP290	Progressive Retinal Atropy	12
PKD1	Polycystic Kidney Disease	13
DPYS	Dihydropyrimidiname Deficiency	14
SLC3A1	Cystinuria, Type 1A	15
HEXB	Gangliosidosis	16,17
GRHPR	Hyperoxaluria	18
TPO	Hypothyroidism	19
GNPTAB	Mucolipidosis II	20
GUSB	Mucopolysaccharidosis VII	21, 22
UROS	Porphyria	23,24



Supplementary figure 1: (a) Number of mitochondrial haplotypes detected and (b) mitochondrial haplotype diversity

Supplementary figure 2: Photos of tigeress T99 from Ranthambore Tiger Reserve (a,b). Close up of her eyes (c). Her litter-mate BEN_NW63 has 35% of her genome in more than 10Mb long ROH.



Supplementary figure 3: Relative mutation loads based on loci with 0.1 percent of the top GERP scores. Error bars are standard errors from jacknife.

Supplement References

1. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997* (2013).
2. H. Li, *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
3. C. T. Saunders, *et al.*, Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* **28**, 1811–1817 (2012).
4. P. Danecek, *et al.*, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
5. W. McLaren, *et al.*, The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
6. Y. Xue, *et al.*, Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding. *Science* **348**, 242–245 (2015).
7. 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*, 696831 (2019).
8. 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).
9. K.M. Meurs, *et al.*, A cardiac myosin binding protein C mutation in the Maine Coon cat with familial hypertrophic cardiomyopathy. *Human Molecular Genetics* **14**, 3587-3593 (2005).
10. K.M. Meurs, M.M. Norgard, M.M. Ederer, K.P. Hendrix, M.D. Kittleson, A substitution mutation in the myosin binding protein C gene in ragdoll hypertrophic cardiomyopathy. *Genomics* **90**, 261-264 (2007).
11. B. Gandolfi, *et al*., First WNK4-hypokalemia animal model identified by genome-wide association in Burmese cats. *PloS one* **7**, p.e53173 (2012).
12. M. Menotti-Raymond, *et al*., Mutation discovered in a feline model of human congenital retinal blinding disease. *Investigative Ophthalmology & Visual Science* **51**, 2852-2859 (2010).
13. L.A. Lyons, *et al*., Feline polycystic kidney disease mutation identified in PKD1. *J. Am. Soc. Nephrology*, **15**, 2548-2555 (2004).

14. H.S. Chang, *et al*., Dihydropyrimidinase deficiency: the first feline case of dihydropyrimidinuria with clinical and molecular findings. In *JIMD Reports-Case and Research Reports, 2012/3* (pp. 21-26). Springer, Berlin, Heidelberg (2012).
15. K. Mizukami, K. Raj, U. Giger, Feline cystinuria caused by a missense mutation in the SLC3A1 gene. *Journal of Veterinary Internal Medicine* **29**, 120-125 (2015).
16. D.R. Martin, *et al.*, An inversion of 25 base pairs causes feline GM2 gangliosidosis variant 0. *Experimental Neurology* **187**, 30-37 (2004).
17. Y. Kanae, *et al*., Nonsense mutation of feline β-hexosaminidase β-subunit (HEXB) gene causing Sandhoff disease in a family of Japanese domestic cats. *Research in Veterinary Science* **82**, 54-60 (2007).
18. R.E. Goldstein, S. Narala, N. Sabet, O. Goldstein, S.P. McDonough, Primary hyperoxaluria in cats is caused by a mutation in the feline GRHPR gene. *Journal of Heredity* **100**(suppl_1), S2-S7 (2009).
19. U. Giger, *et al*., Congenital hypothyroidism with goiter in cats due to a tpo mutation: ESVE-O-4. *Journal of Veterinary Internal Medicine* **29**, (2015).
20. H. Mazrier, *et al*., Inheritance, biochemical abnormalities, and clinical features of feline mucolipidosis II: the first animal model of human I-cell disease. *Journal of Heredity* **94**, 363-373 (2003).
21. J.C. Fyfe, *et al*., Molecular basis of feline β-glucuronidase deficiency: an animal model of mucopolysaccharidosis VII. *Genomics* **58**, 121-128 (1999).
22. P. Wang, *et al*., Mucopolysaccharidosis VII in a cat caused by 2 adjacent missense mutations in the GUSB gene. *Journal of Veterinary Internal Medicine* **29**, 1022-1028 (2015).
23. S. Clavero, D.F. Bishop, U. Giger, M.E. Haskins, R.J. Desnick, Feline congenital erythropoietic porphyria: two homozygous UROS missense mutations cause the enzyme deficiency and porphyrin accumulation. *Molecular Medicine* **16**, 381-388 (2010).
24. S. Clavero, *et al*., Diagnosis of feline acute intermittent porphyria presenting with erythrodontia requires molecular analyses. *The Veterinary Journal* **198**, 720-722 (2013).