Knox *et al. Genome Biology* (2025) 26:405
https://doi.org/10.1186/s13059-025-03878-y

Genome Biology

RESEARCH Open Access

Mitochondrial diversity of Bwindi Impenetrable National Park Mountain Gorillas

Matthew A. Knox[1*], Valter Almeida[1], Gladys Kalema-Zikusoka[2], Stephen Rubanga[2], Alex Ngabirano[3] and David T. S. Hayman[1]

*Correspondence:
Matthew A. Knox
m.knox@massey.ac.nz
[1]School of Veterinary Science, Massey University, Palmerston North, New Zealand
[2]Conservation Through Public Health, Entebbe, Uganda
[3]Mubare Biodiversity Conservation, Bwindi, Uganda

Abstract

Background Mitochondrial DNA is a key marker for assessing genetic diversity, critical for the conservation of endangered species. This study investigates the mitochondrial diversity of the Bwindi Impenetrable National Park (BINP) mountain gorilla population (*Gorilla beringei beringei*), one of the most endangered primate subspecies.

Results Using pooled sequencing of 200 faecal samples collected from both habituated and wild gorillas, we identify ten mtDNA variants exceeding a 20% threshold across the population mitogenome. Comparisons with previously sequenced individual BINP gorilla mitogenomes corroborates these findings and reveals additional putative haplotypes, potential heteroplasmy and nuclear mitochondrial DNA segments. Our approach overcomes challenges associated with pooled samples, distinguishing sequencing noise from biological variation. The observed diversity suggests that mitochondrial variability in mountain gorillas is comparable to the higher levels reported in the closely related Grauer's gorilla (*G. beringei graueri*).

Conclusions This study demonstrates the utility of non-invasive faecal sampling and pooled sequencing for assessing genetic diversity in challenging field conditions, highlighting its potential for population-level genetic monitoring of non-human primates. Our findings provide valuable insights into the genetic makeup of this critically endangered population, contributing to future conservation efforts, and supporting the recovery of mountain gorillas.

Background

Biodiversity encompasses ecosystems, species and within-species (intraspecific) genetic diversity which is declining globally, driven by habitat loss, pollution, climate change, overexploitation of resources, and the introduction of invasive species [1]. The loss of intraspecific genetic diversity limits evolutionary responses and reduces the tolerance of populations and species to stressors [2, 3], ultimately impacting both species and ecosystem biodiversity. Over the last century vertebrate species loss is up to 100 times higher than the background rate [4], and remaining populations of wild animals are experiencing declining genetic diversity [5, 6]. Clearly, action is needed to stop genetic diversity loss [7], especially since a time lag exits between declines in genetic diversity and

population collapse [8]. Accurately assessing present-day genetic diversity in endangered species is essential for prioritising and implementing effective conservation measures [9] and halting the ongoing extinction crisis. However, sampling endangered species using blood or biopsies can be challenging and potentially harmful, making the extraction of DNA from faecal samples a valuable non-invasive alternative [10]. In addition, as microbiome sequencing becomes increasingly commonplace, the incidental presence of host DNA in such datasets allows insight into host – microbiome interactions directly from metagenomic samples [11] but also creates valuable opportunities to study genetic diversity in samples from endangered species. Here we utilize these approaches to investigate the mitochondrial diversity of the Bwindi Impenetrable National Park (BINP) mountain gorilla population.

Mitochondrial DNA is a widely used measure of genetic diversity [12] and potentially useful in conservation studies of endangered species [13, 14]. The critically endangered eastern lowland gorilla (*Gorilla beringei*) have experienced population decline over the past 100,000 years [15], cumulating more recently with losses from habitat loss, poaching and disease outbreaks [16, 17]. Eastern lowland gorilla consist of two subspecies; Grauer's gorilla (*G. beringei graueri*) in eastern Democratic Republic of Congo forests and the mountain gorilla (*G. beringei beringei*) which occupies two nearby territories of Virunga Massif and BINP (Fig. 1). While both subspecies are endangered, mountain gorillas number only around 1,000 wild individuals, whereas Grauer's consist of around 6,800 [18]. Of the two mountain gorilla populations, Virunga Massif has 639–669 individuals [19] while the BINP population was recently estimated at 459 individuals [20]. Despite recent and promising signs of population growth for mountain gorillas [19], DNA analyses show that all eastern lowland gorilla populations have decreased in effective population size over the past 100,000 years [21], with a 10-fold reduction in mountain gorillas in the past 5,000–10,000 years [16]. The most recent and drastic loss of genetic diversity over the past century has been attributed to overall population decline including the loss of peripheral populations [22], which can disproportionately impact the overall genetic diversity of a species [23]. Therefore, knowledge of genetic diversity within all remaining populations is essential for halting further decline and setting future goals.

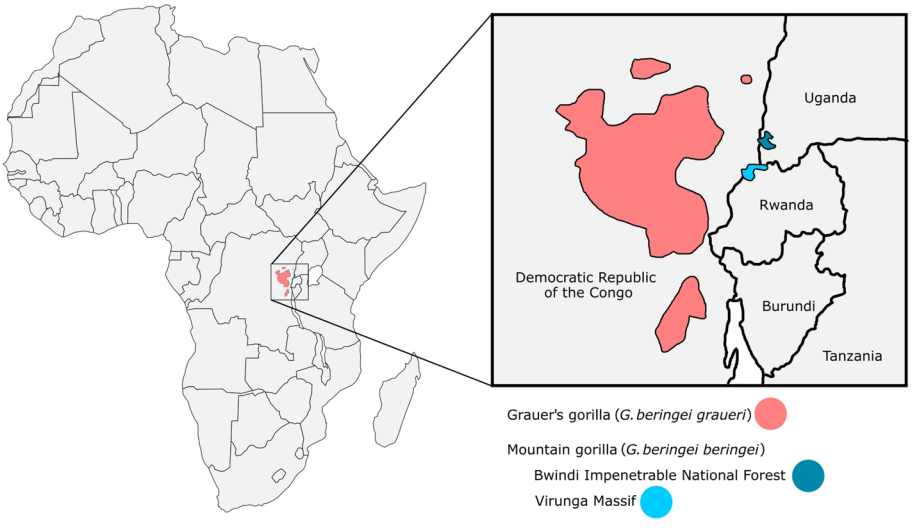

Fig. 1 Distribution of eastern lowland gorilla (*Gorilla beringei*) populations. Sourced from https://www.iucnredlist.org/species/39994/115576640#external-data, accessed 05/08/2025

Currently 130 complete *G. beringei* mitochondrial sequences are available on the National Center for Biotechnology Information (NCBI) database, of which only 20 are from mountain gorillas. In a recent analysis, five haplotypes were identified from the mitogenomes of mountain gorilla [15, 22], however these did not include any samples from the BINP population. In comparison, the 110 complete mitochondrial sequences from Grauer's gorilla have 50 variable locations across the mitogenome (not including the highly polymorphic, but potentially error prone [22, 24] D-loop region). At the subspecies level, consensus sequences of mountain and Grauer's gorilla mitogenomes differ by 76 variable locations. Currently, information on the mitogenome of mountain gorillas is lacking relative to Grauer's, especially within the smaller BINP population. Our study addresses this gap in understanding by interrogating the bulk sequence from 200 faecal samples to identify variable positions with the mtDNA of Bwindi gorillas and hence increase the knowledge of mtDNA genetic diversity in this understudied population.

Results

Initial mapping with Bowtie extracted 111,290 gorilla mitochondrial reads from the metagenomic dataset (min 5,527, max 20,383 in each of the ten pooled samples). Within the complete combined dataset, we observed clear nucleic acid diversity at a threshold of > 20% of total reads per loci at 10 locations across the mitogenome (Table 1, Fig. S1). Background sequence error rates were calculated and 15,263/15,488 loci were homologous at 99% or higher (and (10401/15488) at 99.7%). Of the remaining 225 sites with variants exceeding 1% frequency, the majority (177/225) contained less than 3% variants. Loci with 3% or more sequence variation ($n = 47$) are presented as these may represent rare haplotypes in the BINP mountain gorilla population (Table S1). Analysis of the individual pooled samples (UG01-10) identified the same ten variant loci as in the complete combined dataset (see Table S1). In addition, ten loci with variant frequencies of > 10% were identified (Table 2) from the individual pooled samples. Of these, eight were among the complete combined dataset > 3% variants, while the remaining two (UG09_10543 and UG06_12858) had overall variant frequencies of 2.1% and 2.9% respectively. Contamination with human DNA was ruled out after mapping the study set of reads against a human reference sequence. The majority of reads (82%) were unmapped, and of those that were mapped the resulting consensus differed substantially from the human reference and matched 100% with gorilla mitochondrial DNA after BLAST analyses.

Table 1 Summary of variants within and Bwindi impenetrable National park mountain Gorilla mitogenomes relative to reference NC_037853.1. The WGS column shows how many of the previously published Gorilla with whole genome sequences from Bwindi have variant mitochondrial haplotypes

Nucleotide position	Substitution	Percent variant	Coverage	Region	WGS
5462	A ->G	42.0	962	COX1	2/6
5694	G ->A	42.3	894	COX1	3/6
7864	G ->A	42.0	735	ATP8	3/6
8320	G ->A	39.1	1004	ATP6	2/6
8587	G ->A	56.7	1050	ATP6	3/6
8684	G ->A	43.0	882	COX3	4/6
11,374	A ->G	20.0	963	ND4	2/6
12,487	A ->G	99.6	832	ND5	6/6
15,197	G ->A	40.8	850	CYTB	2/6
15,326	G ->A	40.3	737	-	2/6

Table 2 Variant frequencies relative to reference NC_037853.1 of > 10% from individual pooled samples not included in Table 1

Nucleotide position	Coverage	Substitution	Percent Variant	Sample
4928	51	T ->C	31.4	UG04
4928	90	T ->C	12.2	UG07
6339	48	G ->A	14.6	UG01
6339	51	G ->A	11.8	UG04
6339	78	G ->A	10.3	UG05
6339	62	G ->A	25.8	UG06
6339	88	G ->A	14.8	UG10
7925	94	A ->C	11.7	UG02
7925	49	A ->C	14.3	UG09
8025	54	A ->C	11.1	UG09
9010	55	T ->A	10.9	UG10
9465	134	A ->G	11.2	UG02
9465	54	A ->G	13	UG04
9465	56	A ->G	17.9	UG06
9627	49	A ->C	12.2	UG01
9627	44	A ->C	11.4	UG04
10,543	71	C ->T	16.9	UG09
12,858	49	A ->C	10.2	UG06
13,705	50	A ->C	8	UG01
13,705	50	A ->G	12	UG01
13,705	132	A ->C	6.8	UG02
13,705	132	A ->G	6.1	UG02

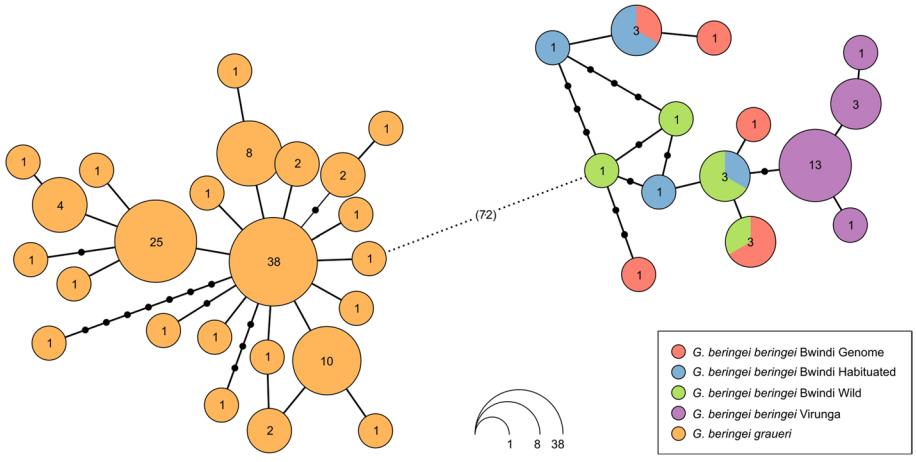

Fig. 2 Minimum spanning haplotype network of *Gorilla beringei* populations and subspecies based on complete mitochondrial genomes. Dots on the haplotype edges correspond to the number of substitutions. The total number of substitutions between subspecies is shown in brackets

Our minimum spanning haplotype network (Fig. 2) revealed seven distinct BINP gorilla mitochondria haplotypes from our study samples (ten including the sequences derived from previously sequenced whole genomes) and places the BINP population close to but distinct from the previously sampled Virunga mountain gorilla population. Wild and habituated groups have overlapping haplotypes, as do the mitochondrial sequences mined from whole genome samples.

To further investigate sequence error rates in mitochondrial sequences from individual mountain gorillas (versus our pooled samples) we accessed mitogenomes from six

G. beringei beringei whole genome samples, also from BINP. These samples were analysed by mapping to mitogenome reference sequence NC_037853.1 and assessing variant frequencies. Variant haplotypes were identified in the same ten loci as in the pooled samples (Table 1). The variant identified at loci 12,487 at nearly 100% was also found in all six individuals from BINP. In addition to these loci, one individual had two variants at position 4928 and 9465. Interestingly, these were loci that both had > 5% variant rates (SI) in our pooled samples, suggesting that these (and others) may represent genuine variants in the BINP gorilla population. Heteroplasmy appeared to be present in one individual (ERR2300764, loci 10543, 74.5% T, 25.6% C). The individual mitogenome samples from whole genome sequences had variable numbers of loci with > 3% variants (10–70 loci), comparable to the pooled study samples ($n = 37$ loci with between 3 and 12% variants). Noticeably, several of these loci were shared between samples. For example, consistent high variant rates (> 3% in both 5 or more individual WGS samples and in our pooled study samples) appear at loci 2991, 2996, 4855, 5841, 8955, 10,359, 10,370.

The mapping of our raw reads to a complete genome assembly revealed gorilla mapping rates of 0.32–0.79% (284,255–685,677 reads) per pooled sample. No significant difference in gorilla read count was found between wild and habituated groups ($p = 0.66$).

Discussion

Our study uses pooled sequence data to analyse host mitochondrial diversity from an underrepresented population of the endangered subspecies of eastern gorillas. Our data is unusual, in that we have pooled samples rather than individuals. However, by separating sequencing noise from biological variation and comparison with mitogenomes available from individuals, we are confident that the ten variants we have identified at the >20% threshold are genuine. Furthermore, the additional loci presented in Table S1 are likely to represent extra diversity within *G. beringei beringei*, comparable to levels seen within the *G. beringei graueri* population [22] (Fig. 2), either as rare haplotypes as is likely the case with the ten variants identified in Table 2, or as mitochondrial heteroplasmy, where certain loci contain variants within individuals. Our findings and approach are corroborated by comparison with individual mitochondrial genomes from the same population, which contained all ten of the >20% variants as well as some other haplotypes which occurred at lower frequencies in our pooled data.

The minimum spanning haplotype network clearly shows the BINP population (including mitogenomes from individuals [25]) is genetically diverse and separate from the nearby Virunga population. We used strict thresholds to avoid nucleotide ambiguity in the analyses and because of this, missing haplotypes in our network may represent minority haplotypes present in individuals from our pooled samples. Molecular studies suggest that the split between the two mountain gorilla populations of BINP and the Virunga Massif occurred 5,000 years ago [16]. Our findings show a single fixed nucleotide difference (nt 12,487) using mitochondrial sequences. In contrast, *G. beringei graueri*, which was estimated to have split 10,000 years ago [16] has >70 mitochondrial differences from mountain gorillas, in agreement with previous analyses [22]. From a conservation perspective, our study has shown that the mountain gorilla population at BINP represents a genetically distinct population, with several unique mitochondrial haplotypes, underscoring the importance of preserving this population for future recovery efforts.

Previous studies have estimated error rates in Illumina next generation sequencing to be 0.24 ± 0.06% per base [26]. The majority of loci in our study (10401/15488) had error rates below 0.3%. However, this leaves approximately a third of all loci with more diversity than can be explained by typical error rates alone. Allowing for an order of magnitude increase of 3% error, we have used the pooled samples available to us to identify 47 potential variants among the BINP gorilla population. The majority of known haplotypes in the Grauer's gorilla dataset (n = 110) are rare and found in single individuals. That is, they occur in less than 1% of sequences and would be difficult to detect against background sequencing noise in a pooled dataset such as in this study. Our analysis of individual pooled samples revealed additional potential variants, possibly representing rare haplotypes, present in a few individuals.

During sample collection, data such as nest number, family group, life stage, date and location were collected (Table S2). Efforts were made to avoid re-sampling the same individual. In the laboratory, DNA was extracted from samples individually, normalised based on DNA quantity and pooled equally in batches of 20 for sequencing. However, since the proportion of host DNA may vary in each sample, there are likely some which were under or overrepresented. Due to these factors, we cannot be sure that the data we analysed accurately represented 200 individual BINP mountain gorillas as would be the case with individually sequenced samples. Nonetheless, our analyses of the pooled data have identified variable loci in the BINP population and targeted studies using the individual samples may further investigate patterns among groups or sectors.

The individual mitochondrial sequences from whole genome samples we analysed were useful for comparison with our pooled samples. Interestingly, the individual samples contained loci with high heterogeneity, that is mixtures of nucleotides above the error thresholds discussed above. These sites were shared between individual samples as well as in our pooled ones. Mitochondrial heteroplasmy (the presence of two or more types of mtDNA in the same individual) [27], has been reported in a range of animal species including hominids [28, 29]. While not observed previously in eastern gorilla [30] it is possible that some of the heterogeneous loci identified in our study, particularly those observed in the individual samples are examples of mitochondrial heteroplasmy. Nuclear mitochondrial DNA (NUMT) may also contribute to the patterns of diversity we observed. Since mitochondria are present in higher copy numbers, these nuclear copies (and their mutations) may appear as low frequency variants in mitochondrial studies.

Conclusions

Microbiome studies of wild animals are becoming increasingly common, giving insight into microbial diversity, host resistome and evolutionary patterns [31–33]. These datasets provide opportunities for extraction of data and study of host DNA in endangered species [34, 35]. Our analyses found that we can use environmentally collected faecal samples to detect DNA which can be good enough for population studies. Other previous work has used this approach for analysis of whole genomes in great ape populations using exome-capture sequencing to focus sequencing effort to capture host DNA [36]. Pooling of samples allows greater sequencing depth, is suitable for population level analyses, and can potentially cut sequencing costs 10-fold [37], allowing for greater sequencing depth. Overall, with the microbiome focussed approach we found that average of 0.6% of reads were from gorillas when processing DNA from faecal samples. We

successfully retrieved 4,787,633 sequences, resulting in a total of 670,063,562 bp of gorilla DNA from the nuclear genome data to be analysed. This opens possibilities for future research on wild species populations through non-invasive sampling in difficult field conditions.

Methods

Sample collection in BINP was conducted by Conservation Through Public Health (CTPH) and collaborators in Bwindi, Uganda. In total, 200 faecal samples were collected from the night nests of habituated gorillas ($N = 100$), which have been gradually accustomed to the presence of humans near the park's northern border with Buhoma, and wild (unhabituated) mountain gorillas ($N = 100$) from further within the park. The sample collection was conducted between May 2017 and May 2018 by collaborators in Uganda and was designed to maximise the number of different individuals. Where possible, the park sector, gorilla family group, nest number and estimated age (silverback, subadults, adult female, juvenile and infant) were recorded (Table S2). Samples were placed in RNAlater® (1:5 ratio) within 24 h of collection and stored in a -20 °C freezer until shipment to New Zealand, which was made in refrigerated conditions. The faecal samples arrived on October 1st, 2018, under the Ministry for Primary Industries permit number 2,018,068,860. They were stored at -20 °C until analysis in a PC2 containment facility at the Hopkirk Research Institute, Massey University, New Zealand.

DNA from approximately 150 mg of faecal sample was extracted using the Zymo Quick-DNA™ Faecal/Soil Microbe Miniprep Kit following the manufacturer's instructions. Extractions were quantified with Qubit and pooled equally in groups of 20 for deep sequencing for community level microbiome analyses and sent to Australian Genome Research Facility Ltd (AGRF) for library preparation (XGen Prism) and sequencing (NovaSeq 150 bp PE). Approximately 1×10^9 reads were generated from the ten sample pools (complete metagenomes available from bioproject accession PRJNA1213876). These reads were comprised of microbial, dietary and host DNA, including nuclear and mitochondrial sequences analysed here. Sequence reads were quality-checked using FastQC (0.11.9) [38]. Adapter sequences were removed, and reads were trimmed with Trimmomatic (0.39) [39]. A mapping index file of the mitogenome reference sequence NC_037853.1 was created using Bowtie2 (2.4.5) [40]. The reference sequence does not contain any D-loop sequence, and this was also excluded from our analyses in accordance with previous studies [22]. After mapping the reads, SAMtools (1.15.1) [41] were used to sort and compress the resulting file of mapped reads (BioSample accessions SAMN46534863- SAMN46534872). Full code for these steps is provided as Additional file 3 as well as summary sequencing and mapping statistics (Table S3).

The subset of mapped reads were then remapped to the original reference sequence and we calculated variant frequency against sequencing background error rates within pooled samples and for the overall dataset using Geneious 10.2.6 [42]. Reads were mapped to the reference using low sensitivity default settings but with a reduced maximum number of mismatches per read of 3% to increase mapping stringency. Due to our pooling strategy, individual samples were indistinguishable from each other, and sequence variants presented themselves as elevated non-consensus base frequencies at each locus. Because of this it was important to clearly distinguish underlying sequencing error from potential biological variation. Potential variants and baseline sequence error

rates were examined using Geneious by calculating the sequence identify statistics and coverage at each location on the genome. The minimum variant frequency setting was set to zero in order to detect any differences in all reads, and the output was exported as a csv (see Additional file 4) for analysis and comparison among samples. This was undertaken using both the complete sample set of 10 pools and each pool separately (UG01-10, consisting of 20 individuals in each) with minimum coverage of 30x. To rule out contamination from human sources, we also mapped reads against a human mitochondrial reference sequence (NC_012920.1).

We also compared our results with previously published mitogenomes from mountain and Grauer's gorilla populations and extracted mitochondrial sequences from six whole genome sequences sequenced previously from BINP [25] and available from (https://www.ebi.ac.uk/ena/browser/view/PRJEB12821). Mitochondrial reads from these individual samples were accessed and assembled using the same reference and mapping strategy as for our study samples and were used for comparative purposes. Being from individual animals they provide an opportunity to separate baseline sequencing error from genetic diversity in the population, including mitochondrial heteroplasmy and nuclear mitochondrial DNA (NUMTs). Consensus sequences from the six whole genome samples and individual pools (UG01-10) were created using a strict 50% threshold, and aligned with mitochondrial sequences from a previous analysis [22] using MAFFT [43]. A minimum spanning haplotype network was constructed in Hapsolutely [44].

Finally, we mapped our raw pooled sample reads to a genome assembly (PGDP_GorBer - GCA_963575185.1) as a reference to assess host nuclear genome coverage from the samples. Using BBMap (version 39.01), an index of the reference genome was created, and the pooled sample reads were mapped with default parameters. The differences between the mapped and unmapped output files were then used to calculate the percentage of gorilla host DNA present in the samples. We also tested whether there were differences in the number of reads from wild versus habituated gorilla pooled samples using a t-test.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13059-025-03878-y.

Additional file 1:Table S1-3.

Additional file 2: Figure S1.

Additional file 3: Bioinformatic scripts. Code used in DNA processing: Trimmomatic, Bowtie, read mapping, BBMap.

Additional file 4: Variant frequency data. DNA substitutions and coverage at all loci for complete sample set and individually pooled samples.

Acknowledgements
Tom van der Valk and Katerina Guschanski for helpful discussion and support during preparation of the manuscript.

Peer review information
Jeffrey Rogers and Tim Sands were the primary editors of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team. The peer-review history is available in the online version of this article.

Authors' contributions
MK led the writing and analysis, VA contributed to analysis, GKZ contributed to sampling , SR contributed to sampling , AN contributed to sampling , DTSH designed the study. All authors contributed to editing of the final manuscript.

Funding
This work was supported by funding from the Royal Society Te Aparangi (Royal Society of New Zealand) [contracts: RDF-MAU1701; MAU1503], the Percival Carmine Chair in Epidemiology and Public Health (DTSH/MK/VA), Massey University Research Fund and McGeorge Research Grant (DTSH).

Data availability
The sequencing data generated in the broader metagenomic study have been deposited in the NCBI Sequence Read Archive (SRA) under the BioProject accession code PRJNA1213876 [45]. Reads mapped used in this study are available as BioSample accessions SAMN46534863- SAMN46534872. These data are also available from the authors upon request. Additional data used in analyses are available from the European Nucleotide Archive, https://www.ebi.ac.uk/ena/browser/view/PRJEB12821.

Declarations

Ethics approval and consent to participate
Massey University Animal Ethics Committee advice was sought and official approval not deemed to be needed due to the absence of manipulations of animals performed. The Ministry for Primary Industries Animal Welfare Senior Policy Advisor was consulted because of ape involvement, but no further consultation was necessary because of the same reason. Kaupapa Kura Taiao (KKT) was consulted via the Environmental Protection Agency as part of the discussions relating to the importation of potentially infectious materials into New Zealand. Permission to import samples was given and Uganda National Council for Science and Technology committee approval was given locally. The samples were imported to New Zealand under the Ministry for Primary Industries Restricted Biological Products Import permit # 2014053241.

Competing interests
The authors declare no competing interests.

Received: 4 February 2025 / Accepted: 19 November 2025
Published online: 28 November 2025

References

1. Jaureguiberry P, et al. The direct drivers of recent global anthropogenic biodiversity loss. Sci Adv. 2022;8(45):eabm9982.
2. Bijlsma R, Loeschcke V. Genetic erosion impedes adaptive responses to stressful environments. Evol Appl. 2012;5(2):117–29.
3. Spielman D, Brook BW, Frankham R. Most species are not driven to extinction before genetic factors impact them. Proc Natl Acad Sci. 2004;101(42):15261–4.
4. Ceballos G, et al. Accelerated modern human–induced species losses: entering the sixth mass extinction. Sci Adv. 2015;1(5):e1400253.
5. Schmidt C, et al. Continent-wide effects of urbanization on bird and mammal genetic diversity. Proc Biol Sci. 2020;287(1920):p20192497.
6. Leigh DM, et al. Estimated six per cent loss of genetic variation in wild populations since the industrial revolution. Evol Appl. 2019;12(8):1505–12.
7. Shaw RE, et al. Global meta-analysis shows action is needed to halt genetic diversity loss. Nature. 2025;638(8051):704–10.
8. Liu X, et al. Time-lagged genomic erosion and future environmental risks in a bird on the Brink of extinction. Proc Biol Sci. 2025;292(2043):20242480–p.
9. Baas P, et al. Population-level assessment of genetic diversity and habitat fragmentation in critically endangered grauer's gorillas. Am J Phys Anthropol. 2018;165(3):565–75.
10. Beja-Pereira A, et al. Advancing ecological Understandings through technological transformations in noninvasive genetics. Mol Ecol Resour. 2009;9(5):1279–301.
11. Kolde R, et al. Host genetic variation and its Microbiome interactions within the human Microbiome project. Genome Med. 2018;10(1):6.
12. James J, Eyre-Walker A. Mitochondrial DNA sequence diversity in mammals: A correlation between the effective and census population sizes. Genome Biol Evol. 2020;12(12):2441–9.
13. Margaryan A, et al. Recent mitochondrial lineage extinction in the critically endangered Javan rhinoceros. Zool J Linn Soc. 2020;190(1):372–83.
14. Zhou Y, et al. Mitogenomics reveals extremely low genetic diversity in the endangered Jilin clawed salamander: implications for its conservation. Ecol Evol. 2024;14(3):e11132.
15. Xue Y, et al. Mountain Gorilla genomes reveal the impact of long-term population decline and inbreeding. Science. 2015;348(6231):242–5.
16. Roy J, et al. Recent divergences and size decreases of Eastern Gorilla populations. Biol Lett. 2014;10(11):20140811.
17. Dunay E, et al. Pathogen transmission from humans to great apes is a growing threat to primate conservation. EcoHealth. 2018;15(1):148–62.
18. Plumptre AJ, et al. Changes in grauer's Gorilla (Gorilla Beringei graueri) and other primate populations in the Kahuzi-Biega National park and Oku community Reserve, the heart of grauer's Gorilla global range. Am J Primatol. 2021;83(7):e23288.
19. Granjon A-C, et al. Estimating abundance and growth rates in a wild mountain Gorilla population. Anim Conserv. 2020;23(4):455–65.
20. Hickey JR, et al. Bwindi-Sarambwe 2018 surveys monitoring mountain gorillas, other select mammals, and human activities. Bwindi Impenetrable National park; 2019.
21. van der Valk T, et al. Comparative genomic analyses provide new insights into evolutionary history and conservation genomics of gorillas. BMC Ecol Evol. 2024;24(1):14.
22. Van Der Valk T, et al. Significant loss of mitochondrial diversity within the last century due to extinction of peripheral populations in Eastern gorillas. Sci Rep. 2018;8(1):6551.
23. Macdonald SL, et al. Peripheral isolates as sources of adaptive diversity under climate change. Front Ecol Evol. 2017;5:88.
24. Guschanski K, et al. Next-generation Museomics disentangles one of the largest primate radiations. Syst Biol. 2013;62(4):539–54.

25. Pawar H et al. Ghost admixture in eastern gorillas. Nat Ecol Evol. 2023;7:1503–14.
26. Pfeiffer F, et al. Systematic evaluation of error rates and causes in short samples in next-generation sequencing. Sci Rep. 2018;8(1):10950.
27. Parakatselaki ME, Ladoukakis ED. mtDNA Heteroplasmy: Origin, Detection, Significance, and Evolutionary Consequences. Life (Basel). 2021;11(7).
28. Stewart JB, Chinnery PF. Extreme heterogeneity of human mitochondrial DNA from organelles to populations. Nat Rev Genet. 2021;22(2):106–18.
29. Jebb D, et al. Population level mitogenomics of long-lived bats reveals dynamic heteroplasmy and challenges the free radical theory of ageing. Sci Rep. 2018;8(1):13634.
30. Das R, et al. Complete mitochondrial genome sequence of the Eastern Gorilla (Gorilla beringei) and implications for African ape biogeography. J Hered. 2014;105(6):846–55.
31. Youngblut ND et al. Large-scale metagenome assembly reveals novel animal-associated microbial genomes, biosynthetic gene clusters, and other genetic diversity. mSystems. 2020;5(6).
32. Tung J, et al. Social networks predict gut Microbiome composition in wild baboons. eLife. 2015;4:e05224.
33. Kumari P, et al. The host-specific resistome in environmental feces of Eurasian otters (Lutra lutra) and Leopard cats (Prionailurus bengalensis) revealed by metagenomic sequencing. One Health. 2022;14:100385.
34. Ang A, et al. Faecal DNA to the rescue: shotgun sequencing of non-invasive samples reveals two subspecies of Southeast Asian primates to be critically endangered species. Sci Rep. 2020;10(1):9396.
35. Baeza JA et al. Mitochondrial genomes assembled from Non-Invasive eDNA metagenomic scat samples in critically endangered mammals. Genes (Basel). 2023;14(3):657.
36. Hayakawa T, et al. Genome-scale evolution in local populations of wild chimpanzees. Sci Rep. 2025;15(1):548.
37. Ray KJ, et al. High-throughput sequencing of pooled samples to determine community-level Microbiome diversity. Ann Epidemiol. 2019;39:63–8.
38. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
39. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics. 2014;30(15):2114–20.
40. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. Nat Methods. 2012;9(4):357–9.
41. Li H, et al. The sequence Alignment/Map format and samtools. Bioinformatics. 2009;25(16):2078–9.
42. Kearse M, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28(12):1647–9.
43. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80.
44. Vences M, et al. Hapsolutely: a user-friendly tool integrating haplotype phasing, network construction, and haploweb calculation. Bioinform Adv. 2024;4(1):vbae083.
45. BioProject. PRJNA1213876. National Center for Biotechnology Information. Available at: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1213876.

Publisher's note