
Supplementary information

Trait-based sensitivity of large mammals to a catastrophic tropical cyclone

In the format provided by the
authors and unedited

Supplementary information

Trait-based sensitivity of large mammals to a catastrophic tropical cyclone

In the format provided by the
authors and unedited

Supplementary Methods:

DNA metabarcoding analysis of wild dog diets. Samples were collected on an opportunistic basis with unused nitrile gloves and placed in an unused plastic bag. Sampling date and location were recorded on the bag, which was placed on ice for transport to the field laboratory. There, samples were frozen at -20° C until processing. Every three months, all samples collected from the intervening period were thawed and preprocessed for DNA extraction. Sample preprocessing involved homogenizing each thawed sample by massaging the bag between thumb and forefinger. A pea-sized subsample was then transferred into a plastic tube containing silica bashing beads and 750uL of buffer (Xpedition Lysis/Stabilization Solution; Zymo Research, CA, USA). We then capped the tube and vortexed it for 30 s to break apart the subsample, exposing it to the buffer. Samples were heat-treated at 72° C for 30 min as an antiviral precaution (a step mandated by the US Department of Agriculture for most herbivore species, described below) before being refrozen at -20° C for transport to Princeton University.

In a Biosafety Level 2 facility at Princeton University dedicated to fecal DNA analysis, we extracted DNA from the samples using Zymo Quick-DNA Soil/Fecal Microbe MiniPrep kits (Zymo Research, CA, USA) following manufacturer instructions. DNA extractions comprised batches of 29 samples (wild dog samples were extracted as part of a broader extraction session involving other species) and one negative extraction control (750uL DNA lysis buffer). Extracted DNA was compiled onto two 96-well plates for amplification. Before proceeding with DNA amplification, we first standardized DNA concentrations across samples using the Quant-iT Picogreen assay (ThermoFisher Scientific) to quantify DNA concentrations and adding nucleic-acid-free water (Qiagen, MD) to dilute high-concentration samples. We amplified the

mitochondrial 16S gene to amplify mammal DNA using an established primer pair (MamP007F, 5'-CGAGAAGACCCTATGGAGCT-3'; MamP007R, 5'-CCGAGGTCRCCCCAACC-3')¹. To multiplex PCR products, we tagged the forward and reverse primers with 8-nt tags that each differed by at least 4 nt to distinguish samples after sequencing. To limit amplification of wild dog DNA, we used a blocking primer (5'-GGAGCTTTAATTAACCAAGCTTACGG-3') with a C3-spacer added to the 3' end and a six-base overlap with MamP007F (first six bases of primer sequence, underlined). PCRs contained 2μL of extracted DNA, 0.5U of AmpliTaq Gold (Applied Biosystems, MA), 0.2μg BSA (New England Biolabs, MA), 0.2μM of each primer, 0.2mM of each dNTP (New England Biolabs, MA), 2mM MgCl₂ (Applied Biosystems, MA), 1X GeneAmp PCR Buffer II (Applied Biosystems, MA), and 2μM of the blocking primer, with a final volume of 20μL. Thermocycling conditions were: initial denaturing (95° C, 10 min), 45 cycles of denaturing (95° C, 30 s), annealing (52° C, 30 s), elongation (72° C, 30 s), and a final extension (72° C, 7 min). All PCRs were performed in triplicate and included negative extraction controls, negative PCR controls, and positive PCR controls (where DNA template was provided as a sequence designed *in silico* to comprise: 5'-10 random bases, MamP007F, a 75-base random barcode, MamP007R, 10 random bases-3'). We confirmed successful amplification of the barcode using gel electrophoresis. We pooled PCR products by plate and purified them with a MinElute PCR Purification kit (Qiagen, MD). Purified PCR products from each plate were submitted for sequencing as equimolar libraries to the Lewis-Sigler Institute for Integrative Genomics at Princeton University, where Illumina tags were appended with a low-cycle PCR approach and libraries were sequenced in paired-end (2×150bp) on a NovaSeq SP 300-nt.

Forward and reverse reads were paired using OBITools' *illuminapairedend* command (minimum score = 40)². Sequences were then assigned to the samples they came from (*ngsfilter*; up to two errors allowed), while sequences that were unaligned, contained ambiguous bases, or were outside the expected barcode length (<40 or >140 bp) were removed. Identical sequences were aggregated (*obiuniq*) and any sequences with only one read in the dataset (i.e., singletons) were removed (*obigrep -p 'count>1'*). Taxonomic identifiers were assigned to each sequence using the *ecotag* command and a 16S reference database created using *in silico* PCR and the EMBL taxonomic library (release 143). Only sequences belonging to Mammalia (taxonomic ID: 40674) were retained in the dataset. We removed putative chimeras (sequences with <80% match to reference database) and PCR errors, which were identified using the *obiclean* command (parameters: -d 1; -r 0.25). Sequences with highest abundance in controls were considered contaminants and removed. PCR replicates were removed if their read depth fell below a critical threshold (1000 reads), or if they were above the 95th quantile for contaminant read abundance. Consistency among PCR replicates was confirmed by comparing their composition. If inter-replicate dissimilarity was greater than the 95th quantile of inter-replicate dissimilarity, or if an inter-replicate comparison fell within the distribution of inter-sample dissimilarity, it was removed. When two or more PCR replicates remained in the dataset, their composition was averaged to give the mean scat-level sequence composition. Finally, we limited the data to sequences that had a perfect match to the reference database and removed sequences accounting for <1% of each sample's relative read abundance (RRA) to reduce the likelihood of false positives^{3,4}. Secondary predation (i.e., detecting prey of prey) is unlikely in this dataset because wild dog prey were almost exclusively plant-eating ungulates.

DNA metabarcoding analysis of herbivore diets. Fecal samples were collected by driving Gorongosa's road network and collecting samples from defecating ungulates. Additional samples were collected from animals immobilized for GPS-collaring⁵⁻⁸. For each sample, we recorded the GPS coordinates of the defecation site and a classification of the surrounding habitat type. Other field handling and preprocessing protocols were as described above for wild dogs. Samples were subjected to an antiviral treatment (72°C for 30 minutes) before import into the United States, as mandated by the US Department of Agriculture (USDA/AHPIS/VS permit #130123 to R.M.P.).

On arrival at Princeton University, we extracted DNA from each sample as described above for wild dogs (batches of 29 samples, one extraction control containing lysis buffer only) using Zymo Quick-DNA Fecal/Soil Microbe Mini Prep kits (Zymo Research, CA, USA). DNA extraction, amplification, and sequencing was done in batches corresponding to each sampling period. Extracted DNA from samples was organized into 96-well PCR plates. To limit large differences in post-PCR DNA concentrations, we quantified and standardized the concentration of DNA in samples collected in 2018 and 2019 using Quant-iT Picogreen (ThermoFisher Scientific). Samples with high concentrations were diluted using nucleic-acid-free water. In triplicate for each sample, we amplified the P6 loop of the chloroplast *trnL*(UAA) intron (and the *g* and *h* primers designed for it), a widely used barcode for degraded plant DNA, using primers with a unique 8-nt tag at the 5' end that enabled pooling of uniquely identifiable PCR products for sequencing in a single high-throughput run⁹. PCRs included negative controls (nucleic-acid-free water), negative extraction controls, and positive controls (2016 only). PCR mixtures comprised 2 µl of DNA extract, 0.5U AmpliTaq Gold DNA Polymerase, 0.17 mg ml⁻¹ of BSA, 0.2 µM of each primer, 0.2 mM of each dNTP, 2.5 mM of MgCl₂, 0.4% dimethyl sulfoxide, 1X

GeneAmp PCR buffer II in a final reaction volume of 20 μ L. Thermocycling conditions were: initial denaturing (95°C for 10 min); 35 cycles of denaturing (95°C, 30 seconds), annealing (55°C, 30 seconds), and extension (72°C, 30 seconds); final elongation (72°C, 2 min). We confirmed successful amplification with gel electrophoresis, and PCR products were pooled by plate and purified using MinElute PCR Purification kits. Purified PCR products were then pooled into sequencing libraries, ensuring equimolarity. Sequencing libraries were prepared with a PCR-free approach (2016) or low-cycle PCR (2018, 2019) by the Lewis Sigler Institute for Integrative Genomics at Princeton University, and paired-end sequencing (2 $\times$ 150bp) was performed on Illumina MiSeq and HiSeq 2500 platforms.

We processed sequence data for each collection period using the *OBITools* package². These procedures were similar as described above for wild dogs. Briefly, we assembled paired-end reads (*illumina-pairedend*; arguments: --sanger --score-min=40), assigned reads to the samples they came from (*ngsfilter*; arguments: -e 2), and merged unique sequences (*obiuniq*). We discarded sequences with a low alignment score, ambiguous nucleotides, unexpected barcode length (<8 or >180 nt), and <10 reads across the dataset. We then aligned the remaining sequences to identify sequence variants (*obiclean*; arguments: -d 1 -r 0.25) and removed sequences that were identified more often as variants of dominant sequences rather than dominant sequences themselves or singular sequences (without variants present). We assigned taxonomic identities to the remaining sequences with primary reference a local plant DNA database based on vouchered specimens collected in Gorongosa¹⁰. When the local-library assignment score was <98%, we used secondary comparisons to a global database compiled from the European Molecular Biology Laboratory (release 143) and local databases of plants from

Serengeti National Park, Tanzania, and Laikipia, Kenya¹¹ for taxonomic identification. Sequence variants identified by *obiclean* were retained in the dataset if they perfectly matched a reference sequence in the local Gorongosa reference database. Finally, we removed putative contaminants (sequences most abundant in negative controls) and chimeras (sequences that had a <80% match in the reference databases). Last, PCR replicates were retained if they were within the 95th quantile of inter-replicate dissimilarity, did not fall within the distribution of inter-sample dissimilarities, and were below the 95th quantile of putative contaminant and chimera abundance. Samples were retained if they had >1 replicate remaining after this procedure. We then averaged the number of reads across all retained PCR replicates for each sample and removed molecular operational taxonomic units (mOTUs) accounting for <1% of reads per sample, following the approach that we have used in previous studies from this system^{10,12}. While there are multiple approaches to dealing with low-abundance reads in metabarcoding, per-sample proportional abundance thresholds are one of the simplest and most effective⁴. We tested the sensitivity of our results to removing this threshold and found that while this increased the number of unique sequences by two orders of magnitude (implausibly relative to the estimated number of plant taxa in our study area), it did not qualitatively alter our results for the abundance-weighted dietary metrics that we analyze. Finally, we rarified sample read depth to 1,250 reads to facilitate comparisons among samples and converted mOTU-by-sample matrices into proportional abundance (relative read abundance, RRA) of each plant mOTU per sample. We use RRA data for inference given that RRA of the *trnL*-P6 marker (i) provides a reasonable approximation of proportional consumption¹³⁻¹⁵, (ii) has yielded inferences in our previous studies that qualitatively match those based on presence-absence data^{10,12,15}, and (iii) provides a fuller and more accurate characterization of population-level diets than presence-absence data^{16,17}.

Confirming species identities for a subset of herbivore fecal samples. For a subset of fecal samples, our field notes recorded uncertainties in the herbivore species' identity. Typically this arose from unusual dung morphology or difficulty locating a dung pile after directly observing defecation. For this subset of samples, we amplified the 16S mitochondrial barcode (MamP007F, MamP007R primers) used to analyze wild dog diets to identify the herbivore species of the sample. As described above, the primers were tagged on the 5' end with 8-nt identifiers that each differed by at least 4-nt. We used the same PCR mixture and thermocycling conditions reported above for wild dog diet analysis. After confirming amplification with gel electrophoresis, we pooled and purified the PCR products using the MinElute PCR Purification kit. Purified DNA was combined in equimolar proportions into two sequencing libraries and submitted to the Lewis Sigler Institute for Integrative Genomics, where Illumina tags were appended with a low-cycle PCR approach and libraries were sequenced in paired-end (2×150bp) on a NovaSeq SP 300 nt.

As for diet datasets, sequences were filtered using OBITools following the same procedure outlined above (paired-end reads assembled with *illumina-paired-end*, sequences assigned to samples with *ngsfilter*, sequences dereplicated with *obiuniq*)². As for the wild dog diet filtering, sequences outside the expected 40–140 bp range, with low alignment scores, or with ambiguous nucleotides were discarded. Sequence variants were identified with *obiclean* and only dominant or singular variants were retained. Taxonomic identifiers were associated with sequences using the global reference database for mammals that we compiled with *in silico* PCR from the European Molecular Biology Laboratory release 143. Following ref. ¹⁰, we used five criteria to assign identities to the remaining samples: (1) the sample had >350 mammalian reads, (2) the most common sequence accounted for ≥50% reads, (3) most common sequence matched a

reference database sequence with $\geq 98\%$, (4) only one Latin binomial name was associated with that most common sequence, (5) that most common sequence was at least twice as common as the second most common. Samples were discarded if they failed two of these criteria; for the remainder, the most common sequence was assigned as the sample identity.

References

1. Giguet-Covex, C., J. et al. Long livestock farming history and human landscape shaping revealed by lake sediment DNA. *Nature Comm.* **5**, 3211 (2014).
2. Boyer, F., et al. OBITOOLS: a UNIX-inspired software package for DNA metabarcoding, *Mol. Ecol. Resour.* **16**, 176-182 (2016).
3. Pansu, J., et.al., Generality of cryptic dietary niche differentiation in diverse large-herbivore assemblages. *PNAS* **119**, e2204400119 (2022).
4. Drake, L. E., et al. An assessment of minimum sequence copy thresholds for identifying and reducing the prevalence of artefacts in dietary metabarcoding data. *Methods in Ecol. & Evol.* **13**, 694-710 (2022).
5. Daskin, J. H., et al. Allometry of behavior and niche differentiation among congeneric African antelopes. *Ecol. Monogr.*, e1549 (2022).
6. Atkins, J. L., et al. Cascading impacts of large-carnivore extirpation in an African ecosystem. *Science* **364**, 173-177 (2019).
7. Branco, P. S., et al. Determinants of elephant foraging behavior in a coupled human-natural system: is brown the new green? *J. An. Ecol.* **88**, 780-792 (2019).
8. Walker, R. H., et al. Mechanisms of individual variation in large herbivore diets: roles of spatial heterogeneity and state-dependent foraging. *Ecology* **104**, e3921 (2023).

9. Taberlet, P., et al. Power and limitation of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research* **35**, e14 (2007).
10. Pansu, J., et al. Trophic ecology of large herbivores in a reassembling African ecosystem, *J. Ecol.* **109**, 1355-1376 (2019).
11. Gill, B. A., et al. Plant DNA-barcode library and community phylogeny for a semi-arid Eastern African savanna, *Mol. Ecol. Resour.* **19**, 838-846 (2019).
12. Guyton, J. A., et al. Trophic rewilding revives biotic resistance to shrub invasion, *Nat. Ecol. Evol.* **4**, 712-724 (2020).
13. Willersley, E., et al. Fifty thousand years of Arctic vegetation and megafaunal diet. *Nature* **506**, 47-15 (2014).
14. Craine, J. M., Towne, E.G., Miller, M., & Fierer, N. Climatic warming and the future of bison as grazers. *Sci. Rep.* **5**, 16738 (2015).
15. Kartzinel, T. R., et al. DNA metabarcoding illuminates dietary niche partitioning by African large herbivores, *PNAS* **112**, 8019-8024 (2015).
16. Deagle, B. E., et al. Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? *Mol. Ecol.* **28**, 391-406 (2019).
17. Littleford-Colquhoun, B. L., et al. The precautionary principle and dietary DNA metabarcoding: commonly used abundance thresholds change ecological interpretation. *Mol. Ecol.* **31**, 1615-1626 (2022).

Supplementary Table 1. Summary of large-herbivore fecal sample sizes used in analyses of diet composition, in descending order of total sample size. LWS = late-wet season, EDS = early-dry season, and LDS = late-dry season. The final row reports the mean number of samples per species in each collection period. Late-wet season sample sizes for bushbuck and hartebeest were insufficient for analysis.

Common name	Latin name	Collection periods						
		2016 EDS	2018 LWS	2018 EDS	2018 LDS	2019 LWS	2019 EDS	2019 LDS
Waterbuck	<i>Kobus ellipsiprymnus</i>	40	18	33	16	20	30	17
Warthog	<i>Phacochoerus africanus</i>	19	20	30	12	14	28	14
Impala	<i>Aepyceros melampus</i>	23	25	30	19	22	29	18
Oribi	<i>Ourebia ourebi</i>	16	13	28	16	9	30	16
Reedbuck	<i>Redunca arundinum</i>	12	12	28	15	11	30	15
Elephant	<i>Loxodonta africana</i>	21	15	22	19	15	15	15
Sable	<i>Hippotragus niger</i>	17	6	18	15	22	18	14
Wildebeest	<i>Connochaetes taurinus</i>	29	13	16	15	12	7	11
Kudu	<i>Tragelaphus strepsiceros</i>	13	13	20	14	13	14	15
Nyala	<i>Tragelaphus angasii</i>	10	13	15	10	17	14	13
Buffalo	<i>Syncerus caffer</i>	17	11	12	12	11	9	8
Bushbuck	<i>Tragelaphus sylvaticus</i>	18	0	16	13	0	15	11
Hartebeest	<i>Alcelaphus buselaphus</i>	12	0	14	14	0	8	12
Mean		19	14.5	21.7	14.6	15.1	19	13.8

Supplementary Table 2. Candidate models for predicting proportional changes in herbivore abundance between 2018 (before Cyclone Idai) and 2020 (after the cyclone). Each model is an ordinary least-squares regression, and candidate predictors included log-transformed herbivore body mass ('Mass'), mean proportion of diet composed of grasses in the early dry season of non-cyclone years ('Grass'), and affiliation with floodplain habitat ('Floodplain').

Model	Adj. R^2	AICc	Δ AICc	AIC $\hat{w}$
Mass	0.28	7.04	0	0.31
Floodplain	0.22	8.00	0.96	0.19
Mass + Floodplain	0.36	8.54	1.50	0.15
Mass + Grass	0.34	8.86	1.82	0.13
<i>Intercept only</i>	NA	8.90	1.86	0.12
Grass	-0.005	11.29	4.25	0.04
Grass + Floodplain	0.20	11.43	4.39	0.03
Mass + Grass + Floodplain	0.40	11.95	4.91	0.03
Mass + Grass + Floodplain + Mass:Grass	0.42	17.24	10.20	0.00
Mass + Grass + Floodplain + Mass:Floodplain	0.41	17.67	10.63	0.00
Mass + Grass + Floodplain + Grass:Floodplain	0.32	19.34	12.30	0.00
Mass + Grass + Floodplain + Mass:Floodplain + Mass:Grass	0.38	26.79	19.75	0.00
Mass + Grass + Floodplain + Mass:Grass + Grass:Floodplain	0.37	27.12	20.08	0.00
Mass + Grass + Floodplain + Mass:Floodplain + Grass:Floodplain	0.35	27.47	20.43	0.00
Mass + Grass + Floodplain + Mass:Floodplain + Mass:Grass + Grass:Floodplain	0.28	42.37	35.33	0.00
Mass + Grass + Floodplain + Mass:Floodplain + Mass:Grass + Grass:Floodplain + Mass:Grass:Floodplain	0.20	67.44	60.40	0.00

Supplementary Table 3. Summary of principal component analysis of nutritional-condition data from adult bushbuck, nyala, and kudu (*Tragelaphus* spp., $n = 136$; 2014–2019). Each column (PC1–PC11) represents a principal component. The first two principal components explained 80% of the variance among the 11 metrics used in this analysis (body and hind foot length, maximum rump-fat depth, thickness of biceps femoris and longissimus dorsi muscles, palpation scores at the sacrosciatic ligament, lumbar vertebrae, sacrum, base of tail, and caudal vertebrae).

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
Standard deviation	2.40	1.73	0.79	0.68	0.61	0.52	0.50	0.36	0.28	0.22	0.16
Proportion of variance explained	0.52	0.27	0.06	0.04	0.03	0.02	0.02	0.01	0.01	0.00	0.00
Cumulative proportion of variance	0.52	0.80	0.85	0.89	0.93	0.95	0.97	0.99	0.99	1.00	1.00

Supplementary Table 4. Principal component loadings for each nutritional-condition metric from bushbuck, nyala, and kudu (*Tragelaphus* spp., $n = 136$; 2014–2019). Variables associated with body size (i.e., body and hind foot length, chest girth, thickness of biceps femoris and longissimus dorsi muscles) loaded most strongly onto PC1, whereas variables associated with body fat (i.e., maximum rump fat depth, palpation scores at the sacrosciatic ligament, lumbar vertebrae, sacrum, base of tail, and caudal vertebrae) loaded most strongly onto PC2.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
Body length	-0.33	0.32	-0.06	-0.09	0.00	-0.19	-0.09	-0.04	-0.28	-0.57	-0.57
Hind foot length	-0.33	0.30	-0.06	-0.09	0.08	-0.40	-0.20	0.04	-0.26	0.71	-0.03
Chest girth	-0.34	0.31	-0.05	0.00	-0.01	-0.13	-0.07	-0.21	0.02	-0.34	0.78
B. femoris thickness	-0.32	0.26	0.15	0.10	-0.14	0.75	0.28	0.02	-0.32	0.17	0.00
L. dorsi thickness	-0.36	0.26	-0.08	0.04	-0.02	0.11	-0.02	0.23	0.84	0.07	-0.17
Max. rump fat depth	-0.24	-0.23	0.80	-0.49	0.09	-0.06	0.01	0.00	0.08	-0.03	0.01
Score at sacrosciatic	-0.26	-0.31	0.28	0.66	-0.25	-0.03	-0.49	0.14	-0.06	-0.04	-0.01
Score at lumbar	-0.29	-0.26	-0.04	0.36	0.74	-0.08	0.38	-0.13	0.00	-0.01	-0.04
Score at sacrum	-0.23	-0.36	-0.39	-0.37	0.25	0.39	-0.55	-0.10	-0.03	-0.01	0.01
Score at base of tail	-0.29	-0.33	-0.17	-0.03	-0.49	-0.13	0.28	-0.63	0.11	0.12	-0.12
Score at caudal	-0.29	-0.35	-0.27	-0.17	-0.23	-0.16	0.33	0.68	-0.16	-0.09	0.14