Supporting Information (SI) Appendix

Coprolites reveal ecological interactions lost with the extinction of New Zealand birds

Authors: Alexander Boast[abc], Laura S. Weyrich[a], Jamie R. Wood[c], Jessica L. Metcalf[d], Rob Knight[e], Alan Cooper[a]

Affiliations: [a]Australian Centre for Ancient DNA, University of Adelaide, Adelaide SA 5005, Australia; [b]School of Environment, The University of Auckland, Auckland 1142, New Zealand; [c]Landcare Research, Lincoln, Canterbury 7640, New Zealand; [d]Department of Animal Sciences, Colorado State University, 1171 Pitkin Avenue, Fort Collins, CO 80526, USA; [e]Department of Pediatrics and Department of Computer Science & Engineering and Center for Microbiome Innovation, University of California San Diego, 9500 Gilman Drive, San Diego CA 92093, USA

This pdf file includes:

Results

Sequence diversity. The numbers of sequence reads obtained (after quality-filtering) per sample ranged from 26,102 – 257,186 (Table S2). Overall, read depth varied between samples, but was highest in giant and upland moa from southern beech sites (Table S2). Taxonomic identities assigned to OTUs (Operational Taxonomic Units) included a variety of fungi, land plants, parasites and other eukaryotes (Fig. 2, S1-S4, Table S3). OTUs that did not assign to eukaryote reference data were considered unidentified (Fig. 2, S1) and were filtered prior to all subsequent analyses. These unidentified sequences also comprised a relatively high proportion of coprolite sequence reads (max. 85%; mean 28%), which may not be unexpected due to variable rates of endogenous DNA degradation that are specific to environment. Sequences from the mammal samples (1) were relatively homogenous, similar, and dominated by a small number of abundant, microfungal taxa (Fig. S1)). In contrast, both the modern and ancient bird samples showed greater sequence heterogeneity. Coprolites demonstrated significantly higher sequence diversity than all modern samples ($p=0.003$) (Fig. S5, Dataset S2). Moa coprolites from southern-beech forest contained higher sequence diversity than those from semi-arid shrubland / grassland sites ($p=0.03$, Fig S6, Dataset S2), possibly reflecting poorer DNA preservation in the latter's hotter and more seasonally variable environment (supported by very low read counts of expected taxa such as plants or parasites from semi-arid coprolites). Within moa species, upland moa had the highest OTU diversity, while heavy-footed moa had the lowest, although the difference was not statistically significant (Fig. S5, Dataset S4).

PCoA analysis of jackknifed Unifrac distance values clustered mammals separate to birds (100%) (Fig S7, S8). Coprolite samples fell within two groups; all kakapo, and moa from southern beech sites (100%) clustered into a distinct region relative to moa from shrubland/grassland sites and modern birds (86%) (Fig. S7, S9). However, modern ostrich and kiwi failed to cluster with either group (Fig. S7, S9).

Methods

Site and specimen details. Coprolites used in this study came from eight sites in New Zealand's South Island, covering a variation of palaeoecologies which can be

roughly categorised as follows: one subalpine Southern beech forest (Euphrates Cave), three southern beech forest (Dart River Valley, Honeycomb Cave, Hodge Creek), one mixture southern beech / conifer-broadleaf forest (Mount Nicholas), and three semi-arid shrubland / grassland (Kawarau Gorge, Roxburgh Gorge, Shepherd's Creek) (see Fig. 1, Table S1). Full site descriptions can be found in previous publications (2–17) (See Table S1). Species included South Island giant moa (*Dinornis robustus*, 4 from Dart River Valley), upland moa (*Megalapteryx didinus*, 4 from Dart River Valley, 2 Euphrates Cave, 2 Shepherd's Creek), *Pachyornis elephantopus* (heavy-footed moa, 2 from Dart River, 1 Karawua Gorge, 1 Roxburgh Gorge), little bush moa (*Anomalopteryx didiformis*, 1 from Dart River, 2 Mount Nicholas), and kakapo (*Strigops habroptilus*, 1 Hodge Creek, 3 Honeycomb Cave). Full descriptions are in Table S2.

All samples were confirmed to species by aDNA analyses, most of which are in previous publications (see Table S2). The only exception was a single Mt. Nicholas coprolite, which DNA identification supported either coastal moa (*Euryapteryx curtis*) or little bush moa as the depositor. However, known biogeography of coastal moa makes an association with this species extremely unlikely. In addition, little bush moa coprolites have been confirmed from this site (6). AMS radiocarbon C^{14} dates were obtained for 9 of the 23 coprolite samples at the Waikato radiocarbon dating laboratory, Waikato University, New Zealand. (http://www.radiocarbondating.com/), and were calibrated using the Southern Hemisphere calibration curve (18). Most dates have been previously published (6, 10, 17), except dates for two kakapo coprolites from Honeycomb Cave which are new to this study (see Table S2). More generally, five of the eight coprolite sites (Euphrates Cave, Dart River Valley, Honeycomb Cave, Mt Nicholas, and Kawarau Gorge) have had radiocarbon dates estimated of their coprolites (including coprolites not used in this study). Non-dated coprolites from these five sites are expected to have ages falling within, or near-to, these ranges (see Table S1). No dates are available of coprolites from Hodge Creek, Shepherd's Creek or Roxburgh Gorge, although coprolites from these sites are likely to be late Holocene in age (7). Fresh faecal samples were obtained from captive ratites from one North Island brown kiwi (*Apteryx mantelli*) and two ostrich (*Struthio camelus*), from San Diego and Houston zoos respectively, and were extracted following the standard Earth Microbiome

Project protocols (http://www.earthmicrobiome.org/emp-standard-protocols/dna-extraction-protocol/). In addition, Illumina 18S rRNA sequences from of 23 captive and wild modern mammals were included in the analyses from the study by Parfrey *et al*. (1).

DNA extraction and amplification. Coprolite samples were extracted in an isolated, fully contained, ancient DNA laboratory at the Australian Centre for Ancient DNA (ACAD), using the PowerSoil DNA extraction kit (Mo Bio Laboratories, Carlsbad, CA, USA). Most extractions used were from past studies by Wood *et al.* (4, 10, 12, 14, 15), in which information on host ID determination, radiocarbon dating, and extraction protocols can be found. DNA extractions new to this study included all four kakapo coprolites, and two upland moa coprolites from Shepherd's creek, each of which had host DNA amplified by primers specific either to moa or kakapo (4, 10). Extracts (including extraction blanks) were amplified by PCR (in triplicate) by the universal 18S rRNA gene-targeting, Eukaryote-specific, primers Euk1391f and EukBr based on Amaral-Zettler *et al.* (19), and can be found on the Earth Microbiome Project (EMP) webpage (http://www.earthmicrobiome.org/emp-standard-protocols/). All sets of PCR reactions were pooled, had DNA concentrations quantified, and were purified using the Mobio UltraClean PCR Clean-up kit (Mo Bio laboratories, Carlsbad, CA, USA). High-throughput sequencing (HTS) reads were generated on a single lane of the Illumina HiSeq platform at BioFrontiers Institute Next-Generation Genomics Facility at University of Colorado, Boulder.

QIIME analyses. Combined reads were analysed in the QIIME software package (20), of which raw barcoded sequences were demultiplexed and quality filtered using default parameters, resulting in reads roughly 100bp in length. Sequences were open-reference clustered into OTUs (operational taxonomic units) using an RDP classifier towards the PR^2 database (Protist 2, http://ssu-rrna.org), modified to be QIIME compatible. For a clustering threshold, we elected to utilise 97% pairwise nucleotide similarity between reads. A 97% threshold is most often utilised by similar studies, and is widely considered to account for predicted sequencing error whilst still identifying high-resolution taxonomic diversity (21, 22). OTUs that clustered with the reference database inherited the reference taxonomy. Non-assigned hits were clustered de-novo, and were blasted to the PR^2 database using an e-value cut-off of

$1e^{-25}$. For OTUs that fell below the threshold, the RDP classifier retrained on PR[2] using a 50% confidence threshold. In all instances the most abundant sequence for each cluster was selected as the representative sequence.

All OTUs sequenced from extraction blanks were filtered from the remaining biom file to control for *in-vitro* contaminants in QIIME. All circumstances of <5 reads per OTU per sample (maximum 0.015% of total reads and considered a sequencing error risk) were additionally removed. This was achieved by splitting all files into separate biom files for each individual sample, filtering out all reads <5 per OTU, and re-merging all into a single biom file. All OTUs that could not be assigned to a particular Eukaryote kingdom were then removed from any subsequent analyses. To improve specificity and flexibility of downstream analyses, five biom files were generated representing different sample groups: all samples, all bird taxa, ancients only (all moa and kakapo), ratites only (including modern kiwi and ostrich, but not kakapo), and moa only. Each was subsequently split into groupings specified by OTU taxonomy including all samples, fungi, plants and parasites (which included all taxa identified as vertebrate parasites by BLAST and phylogenetic inferences, see below).

All split biom files were subsequently analysed using a range of available scripts in QIIME. Multiple rarefactions with minimum reads (-m) 100, maximum reads (-x) 10000, and number of repetitions (–n) 10), were run on each biom file and were used to calculate Shannon diversity and t-test differences using the script “compare_alpha.py”. A phylogenetic tree of the entire dataset was calculated using a fasta file of all the reads present in each biom file, and used as an estimation of difference matrices in downstream analyses. Sequences were aligned using MUSCLE (23), through the script “align_seqs.py”, and a UPGMA phylogenetic tree constructed using FastTree (24) via the script “make_phylogeny.py” (both using default parameters). The script “jackknifed_beta.diversity” was used to calculate UniFrac distance matrices, using custom parameters with a rarefaction depth of 2,500 (all), 250 (fungi only) or 100 (plants or parasites only), and the number of rarefactions set to 1,000. This script then generated preliminary UPGMA UniFrac distance-matrix trees, and unweighted UniFrac distances were used to generate PCoA plots using the script “make_2d_plots.py”. The script “beta_diversity_through_plots” was used to create unweighted distance matrix files (not jackknifed), used for statistical analyses (“compare_categories” adonis test

statistics, and “group_significance” g-test and Kruskal Wallis test statistics). All statistical tests were run independently on a diversity of categories specified in a mapping file, including coprolites vs. moderns, environment type, host genus, local site, host genus / local site, lower host taxonomy (moa, kiwi, ostrich, kakapo and mammals) and higher host taxonomy (ratites, kakapo and mammals).

BLASTn identifications. Although higher-level taxonomic identifications are sufficient for most QIIME analyses, ecologically informative inferences required the deepest taxonomic identification possible. Despite the usefulness of curated databases such as PR^2, reference data is missing for a diversity of taxa below a certain taxonomic threshold. We therefore sought to improve taxon identification with alternative methods and databases. All eukaryotic 18S rRNA reads available were downloaded from Genbank (http://www.ncbi.nlm.nih.gov/genbank/) into a fasta file, which was converted into a local, custom BLAST database. To counteract genbank’s lack of curation, any sequences with the tags “uncultured”, “clonal”, “environmental” and “unidentified” were filtered from the search terms. The fasta file generated from QIIME representing all collapsed reads was then blasted to the custom database (BLASTn, maximum target sequences 25, minimum identity percentage match 80%). The output file was opened and processed in MEGAN5 (http://ab.inf.uni-tuebingen.de/software/megan5/), and was collapsed into reliable taxonomic identifications using custom parameters (Min score 150, top 5%, minimum percentage support 0.1%, minimum support 5). The results from MEGAN compared closely to the earlier results using QIIME, and were used as the basis for most final IDs (to genus level or higher), although some sequences were further analysed phylogenetically (Table S4).

Due to a paucity of plant sequences on PR^2, QIIME IDs of plants were restricted to high taxonomic levels, and provided a poor comparison for plant identifications in MEGAN (dataset S1). Furthermore, limited 18S rRNA reference data or barcoding studies of plants potentially made these ID’s spurious. Therefore we elected to restrict most plant IDs to the Order level or higher, which is conservative (25). The only plant ID lower than Order level, was a single OTU identified in the family Ophioglossaceae (the only family in the Ophioglossales) (26).

Phylogenetic identifications. Sequences from coprolites that were identified from QIIME and BLASTn analyses as taxa of ecological importance, except for plants, were analysed phylogenetically (Table S4). Firstly, these were OTUs identified as plant-symbiotic Agaricales fungi considered likely to be dietary, which included the genera *Armillaria* (Physalacriaceae), *Cortinarius* (Cortinariacae) or *Inocybe* (Inocybaceae). Secondly, these were OTUs identified as parasites, including the nematode order Ascaridida, Trematoda or the apicomplexan suborder Eimeriorina. These OTUs were grouped into seven different phylogenetic datasets of varying taxonomic rank (fungal families Physalacriaceae, Cortinariacae, Inocybaceae; the nematode order Ascaridida; the trematode superfamily Pronocephaloidea; and the Eimeriorina families Eimeriidae and Sarcocystidae).

All representative OTU sequences for each group were filtered into individual fasta files and aligned using ClustalW (27) if multiple sequences were present. All available genbank 18S rRNA sequences from each taxonomic group were downloaded and assembled either to the single OTU sequence (if only one OTU per taxon), or alternatively a consensus sequence each alignment (if several OTUs per taxon), in Geneious v. 7.0.5 (28) (one iteration using “medium sensitivity” parameters, modified to have a “minimum overlap” of 80 base pairs, thus leaving only sequences for which at least 80 base pairs of the barcoding sequence was present). Base pairs overlapping the barcoding region were removed, duplicate sequences were removed, and sequences were re-aligned using ClustalW. The fungus *Lepiota cristata*, the apicomplexans *Calpytospora funduli* and *Isospora parvum*, the Guinea worm (*Dracunculus mediensis*), and the Common Liver Fluke (*Fasciola hepatica*) were used as outgroups for all three fungal families, Eimeriidae, Sarcocystidae, Ascaridida and Pronocephaloidea respectively. The most appropriate substitution model for each alignment was estimated in jModelTest 2 v. 2.1.7 (29) using AIC (Akaike’s Information Criterion) calculations, using default parameters (Table S6). Each full alignment was then subjected to phylogenetic analyses in MrBayes v. 3.2 (30–32) using between 10-50 $\times$ 10^6 generations (generations were increased until the average standard deviation of split frequencies reached <0.01), with the first 25% of trees used as a burn-in. Trees were visualised in FigTree v. 1.4 (http://tree.bio.ed.ac.uk/software/figtree/).

References (SI Appendix Text)

1. Parfrey LW, et al. (2014) Communities of microbial eukaryotes in the mammalian gut within the context of environmental eukaryotic diversity. *Front Microbiol* 5:298.
2. Worthy TH (1997) Fossil deposits in the Hodges Creek Cave System, on the northern foothills of Mt Arthur, Nelson. South Island, New Zealand. *Notornis* 44:111–108.
3. Horrocks M, et al. (2008) Plant microfossil analysis of coprolites of the critically endangered kakapo (*Strigops habroptilus*) parrot from New Zealand. *Rev Palaeobot Palynol* 149(3–4):229–245.
4. Wood JR, et al. (2008) Coprolite deposits reveal the diet and ecology of the extinct New Zealand megaherbivore moa (Aves, Dinornithiformes). *Ice Age Refug Quat Extinctions Issue Quat Evol Palaeoecol* 27(27–28):2593–2602.
5. Wood JR, Wilmshurst JM (2013) Pollen analysis of coprolites reveals dietary details of heavy-footed moa (*Pachyornis elephantopus*) and coastal moa (*Euryapteryx curtus*) from Central Otago. *N Z J Ecol* 37(1):151–155.
6. Wood JR, Wilmshurst JM, Worthy TH, Cooper A (2012) First coprolite evidence for the diet of *Anomalopteryx didiformis*, an extinct forest ratite from New Zealand. *N Z J Ecol* 36(2):164.
7. Wood JR, Wilmshurst JM (2014) Late Quaternary terrestrial vertebrate coprolites from New Zealand. *Quat Sci Rev* 98(0):33–44.
8. White T (1875) Notes on moa caves, etc., in the Wakatipu District. *Trans Proc R Soc N Z* 8:97–102.
9. Trotter MM (1970) Archaeological investigations in the Aviemore area, South Island. *Rec Canterb Mus* 8(5):439–453.
10. Wood JR, Wilmshurst JM, Worthy TH, Holzapfel AS, Cooper A (2012) A Lost Link between a Flightless Parrot and a Parasitic Plant and the Potential Role of Coprolites in Conservation Paleobiology. *Conserv Biol* 26(6):1091–1099.
11. Worthy TH (1993) *Fossils of Honeycomb Hill* (Museum of New Zealand).
12. Wood JR, et al. (2012) High-Resolution Coproecology: Using Coprolites to Reconstruct the Habits and Habitats of New Zealand's Extinct Upland Moa (*Megalapteryx didinus*). *PLoS ONE* 7(6):e40025.
13. Rowe P, Millar I, Worthy T (1994) Exploration on Garibaldi Ridge–Euphrates Cave, Kahurangi National Park. *N Z Speleol Bull* 9:271–290.
14. Wood JR, et al. (2013) Resolving lost herbivore community structure using coprolites of four sympatric moa species (Aves: Dinornithiformes). *Proc Natl Acad Sci* 110(42):16910–16915.
15. Wood JR, et al. (2013) A Megafauna's Microfauna: Gastrointestinal Parasites of New Zealand's Extinct Moa (Aves: Dinornithiformes). *PLoS ONE* 8(2):e57315.
16. Wood JR (2008) Pre-settlement Paleoecology of Central Otago's Semi-arid Lowlands, with Emphasis on the Pre-settlement Role of Avian Herbivory in South Island Dryland Ecosystems, New Zealand.
17. Wood JR, Wilmshurst JM, Rawlence NJ (2011) Radiocarbon-dated faunal remains correlate very large rock avalanche deposit with prehistoric Alpine fault rupture. *N Z J Geol Geophys* 54(4):431–434.
18. McCormac F, et al. (2004) SHCal04 Southern Hemisphere calibration, 0–11.0 cal kyr BP. *Radiocarbon* 46(3):1087–1092.
19. Amaral-Zettler LA, McCliment EA, Ducklow HW, Huse SM (2009) A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS One* 4(7):e6372.
20. Caporaso JG, et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Meth* 7(5):335–336.
21. Bokulich NA, et al. (2013) Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat Methods* 10(1):57–59.

22. Orsi W, Biddle JF, Edgcomb V (2013) Deep sequencing of subseafloor eukaryotic rRNA reveals active fungi across marine subsurface provinces. *PLoS One* 8(2):e56335.
23. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792–1797.
24. Price MN, Dehal PS, Arkin AP (2009) FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 26(7):1641–1650.
25. Wood JR, Crown A, Cole TL, Wilmshurst JM (2016) Microscopic and ancient DNA profiling of Polynesian dog (kurī) coprolites from northern New Zealand. *J Archaeol Sci Rep* 6:496–505.
26. Smith AR, et al. (2006) A classification for extant ferns. *Taxon* 55(3):705–731.
27. Larkin MA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947–2948.
28. Kearse M, et al. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28(12):1647–1649.
29. Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* 9(8):772–772.
30. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17(8):754–755.
31. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19(12):1572–1574.
32. Ronquist F, et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61(3):539–542.

Figures

Fig. S1. Stacked column graph of raw (unfiltered) of total read proportions per sample, to taxonomic and ecological groups as determined through QIIME. Note read counts non-equal. Includes extraction blanks, duplicates, and all modern samples.

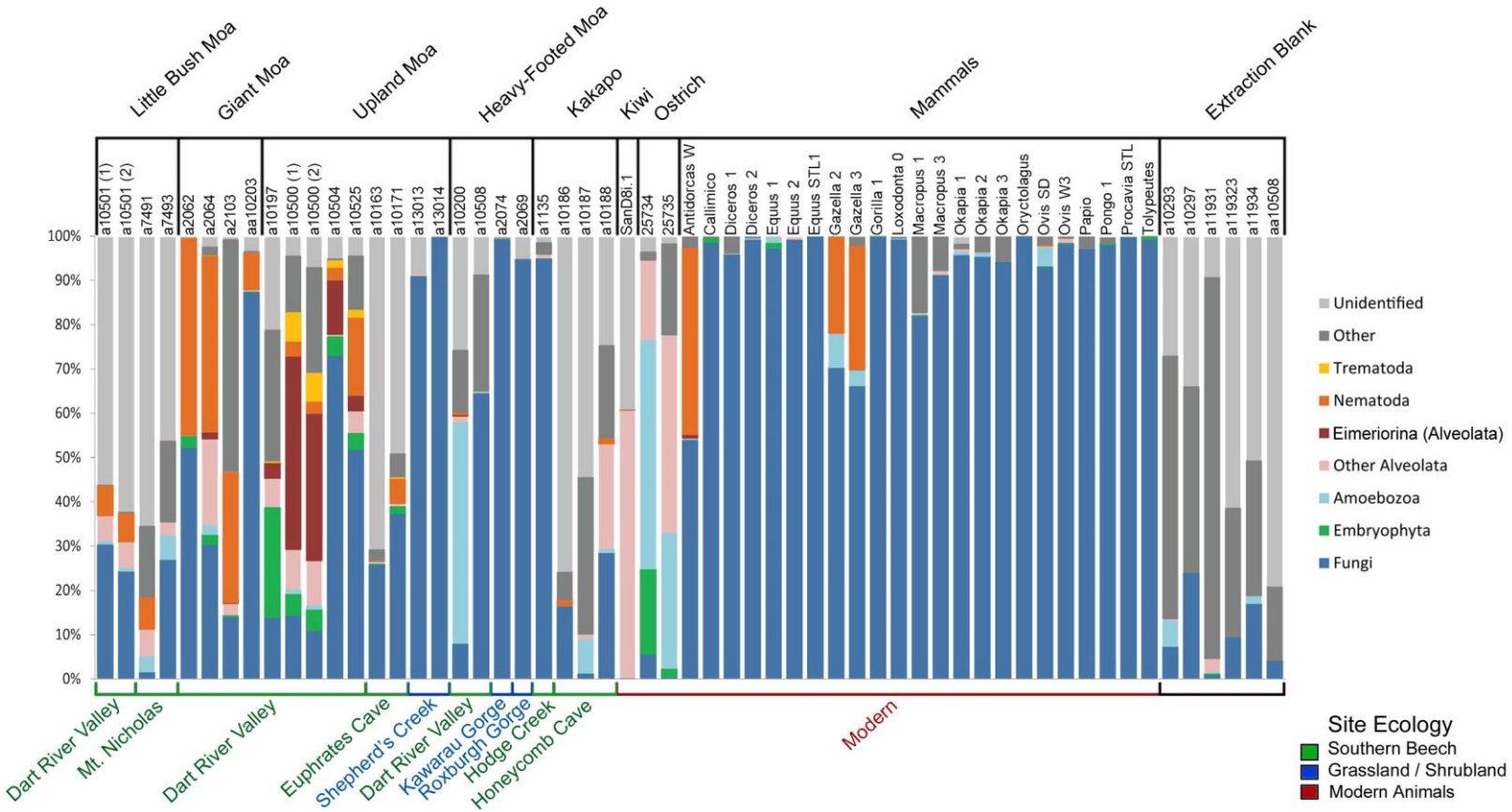


Fig. S2. (Above) Stacked column graph of fungal (kingdom: Fungi) proportions post-filtering, using QIIME (PR2 database) and MEGAN (custom database) based taxonomic associations. (Below) bar graph of proportion of total reads determined as fungal by each taxonomic assignment.

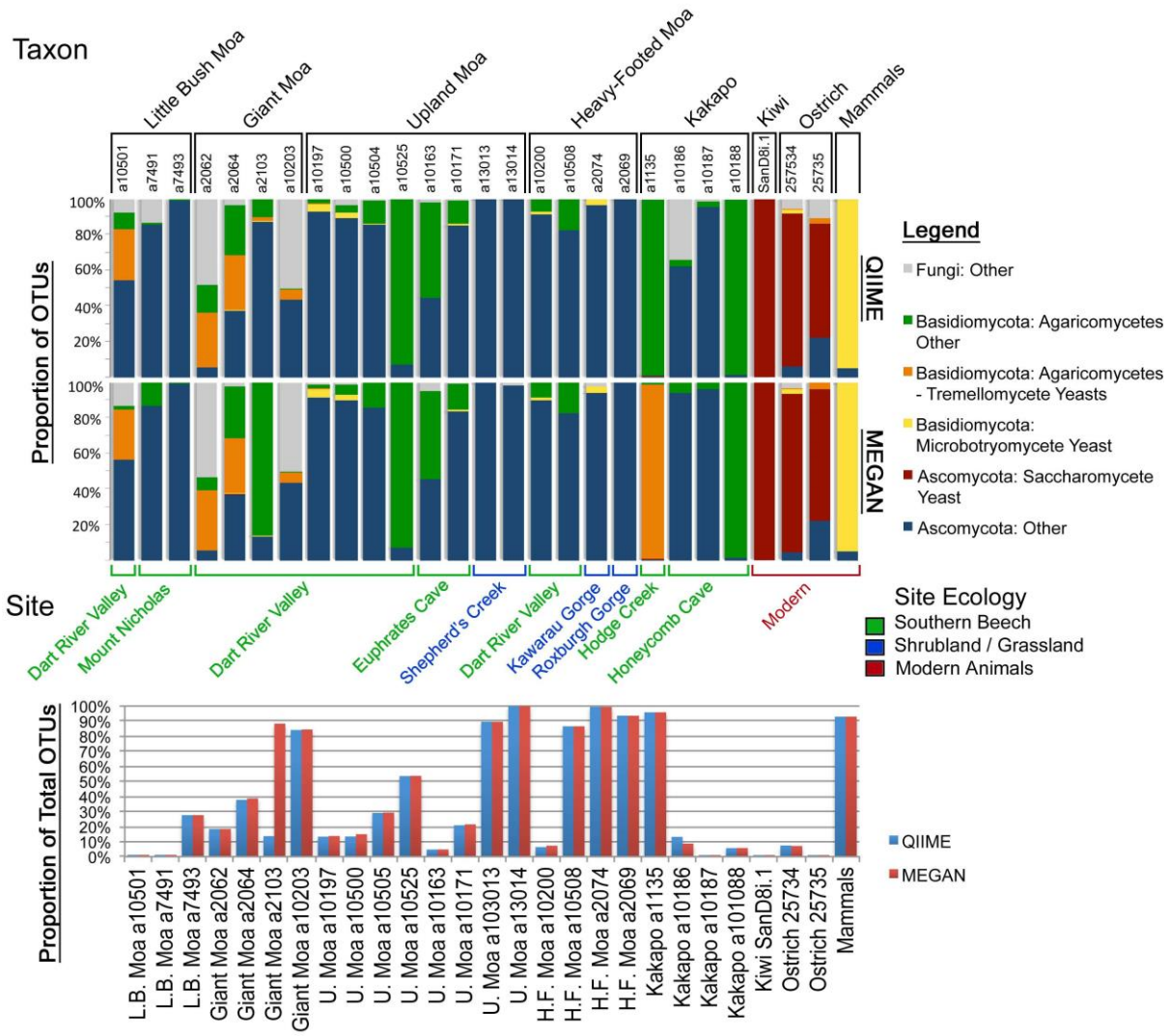


Fig. S3. (Above) Stacked column graph of Embryophyta (land plant) proportions post-filtering, using QIIME (PR2 database) and MEGAN (custom database) based taxonomic associations. (Below) bar graph of proportion of total reads determined as Embryophyta by each taxonomic assignment.

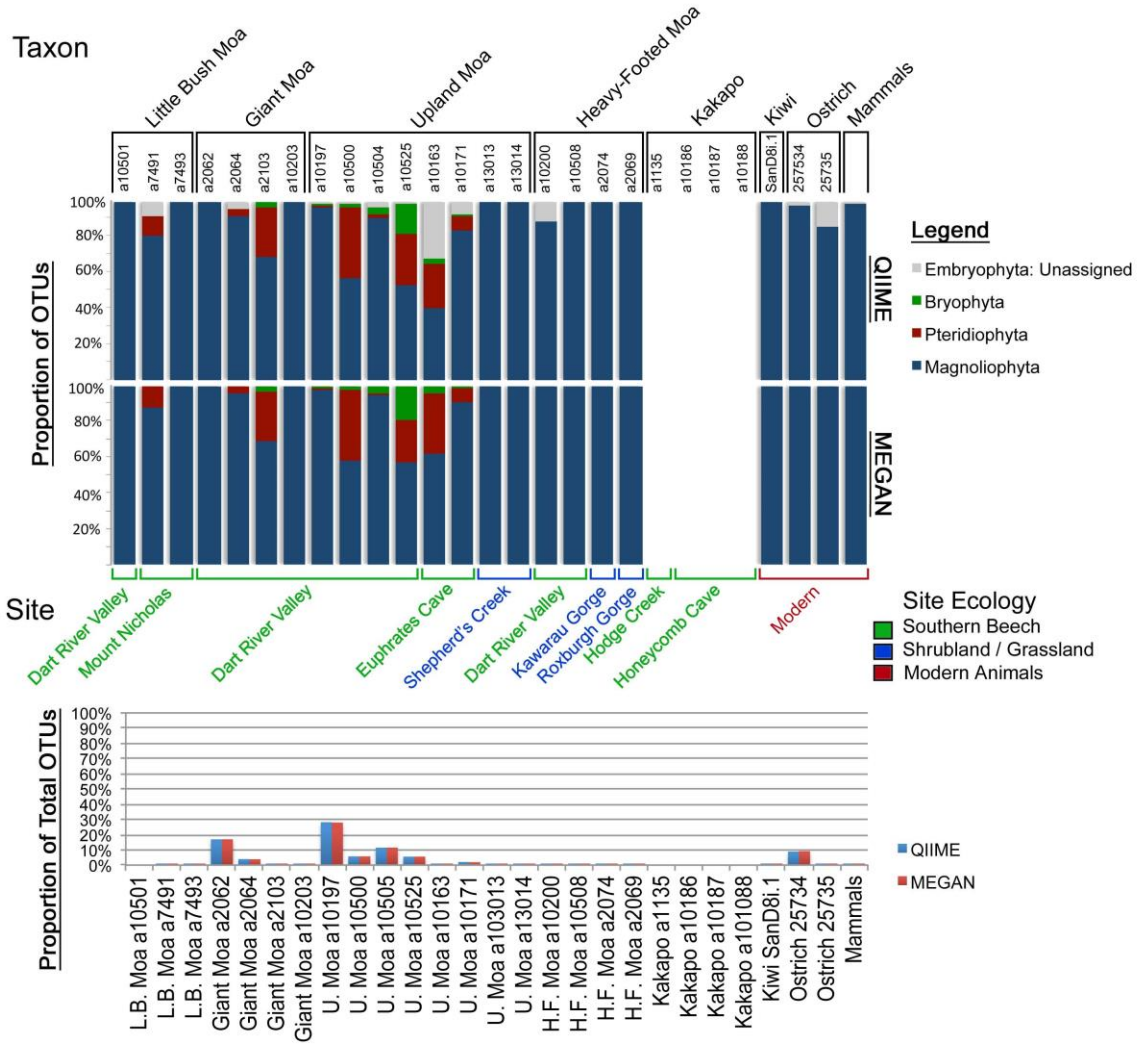


Fig. S4. (Above) Stacked column graph of vertebrate parasite proportions post-filtering, using QIIME (PR2 database) and MEGAN (custom database) based taxonomic associations. Total Nematoda counts (for comparisons with Fig. 2, S2) are included. (Below) bar graph of proportion of total reads determined as parasitic / Nematoda by each taxonomic assignment.

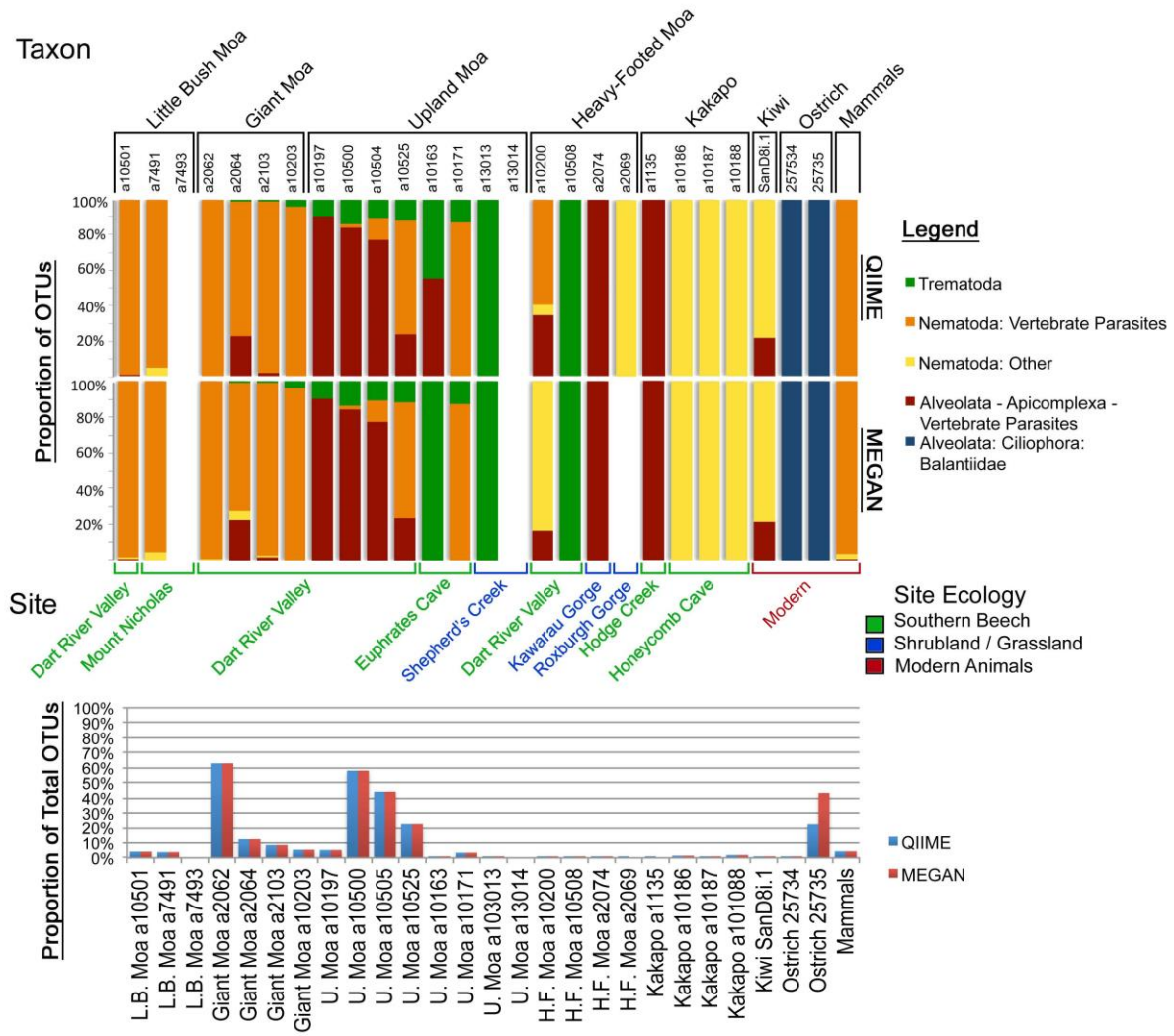


Fig. S5. Alpha diversity (Shannon's diversity) box and whisker plots of taxonomic groups.

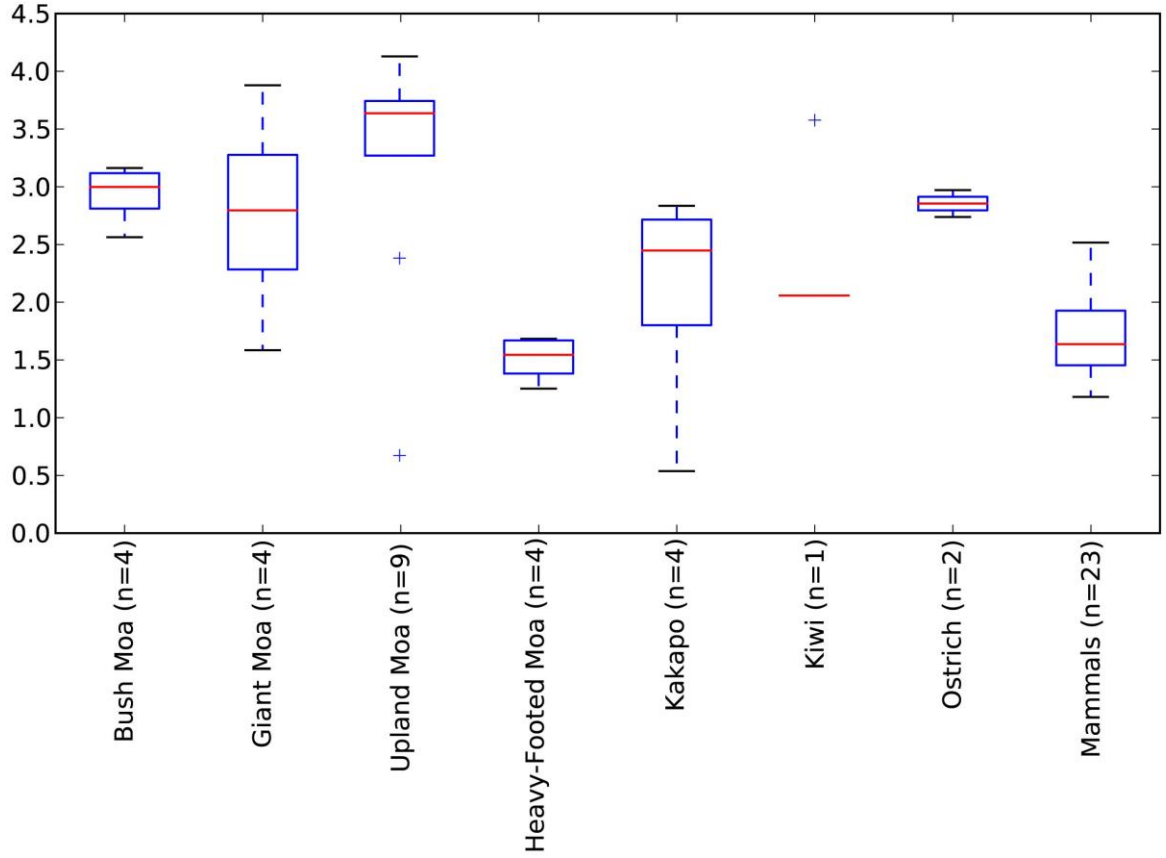


Fig. S6. Alpha diversity (Shannon's diversity) box and whisker plots of collection sites.

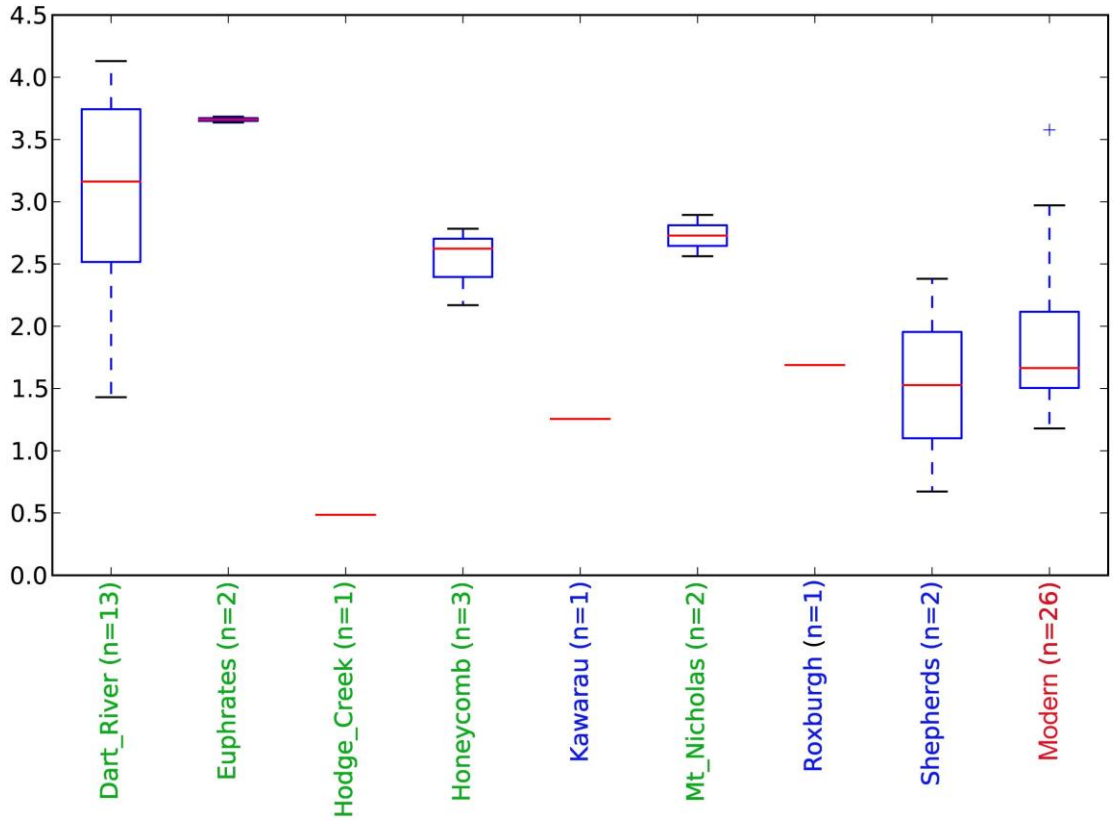


Fig. S7. UPGMA tree of all samples based on UniFrac distance matrices.

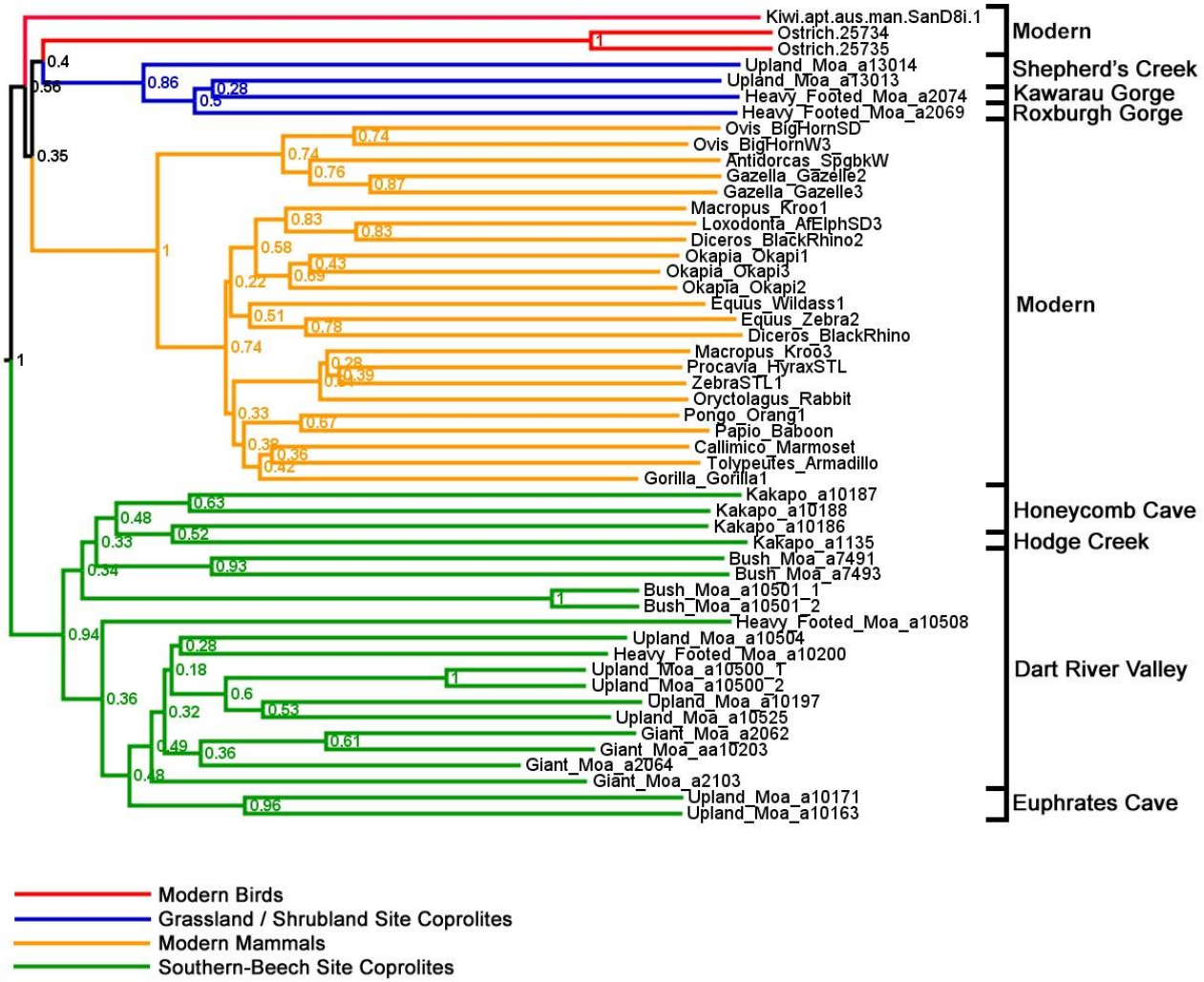


Fig. S8. PCoA (Principal Coordinates Analysis) plot, of jacknifed rarefied sampling of unifrac distances from all filtered reads from all modern and ancient samples. Samples are clustered relative to general similarity between OTU assemblages of each sample. Circles around icons represent error due to differences between the rarefied datasets. Samples are unweighted (number of reads per OTU are not included).

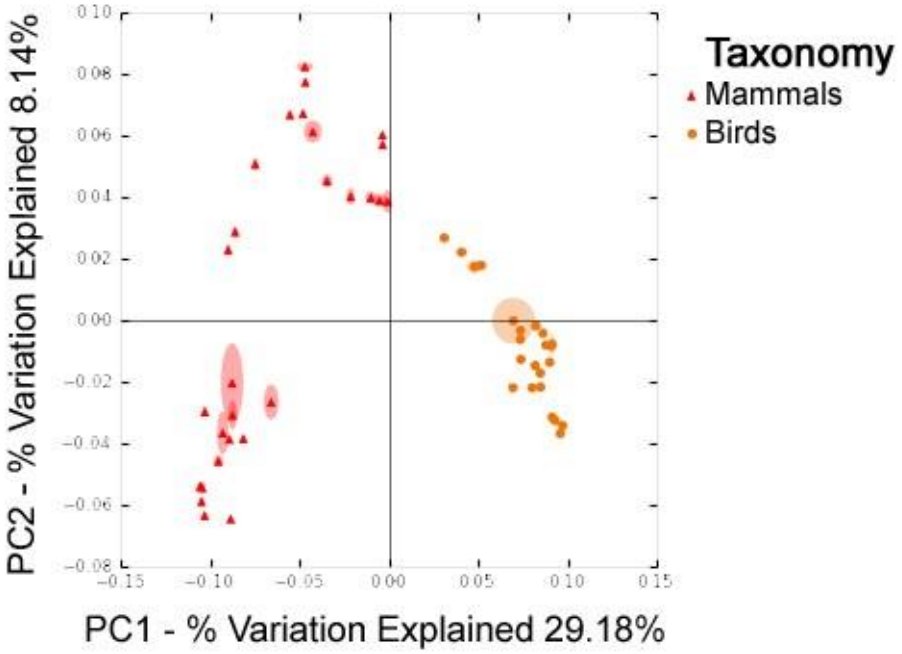


Fig. S9. PCoA (Principal Coordinates Analysis) plots, of jacknifed rarefied sampling from unifrac distances of all filtered reads from bird samples only. Samples are clustered relative to general similarity between OTU assemblages of each sample. Circles around icons represent error due to differences between the rarefied datasets. Samples are unweighted (number of reads per OTU are not included). Grouped to ecology category, host taxon, and collection site.

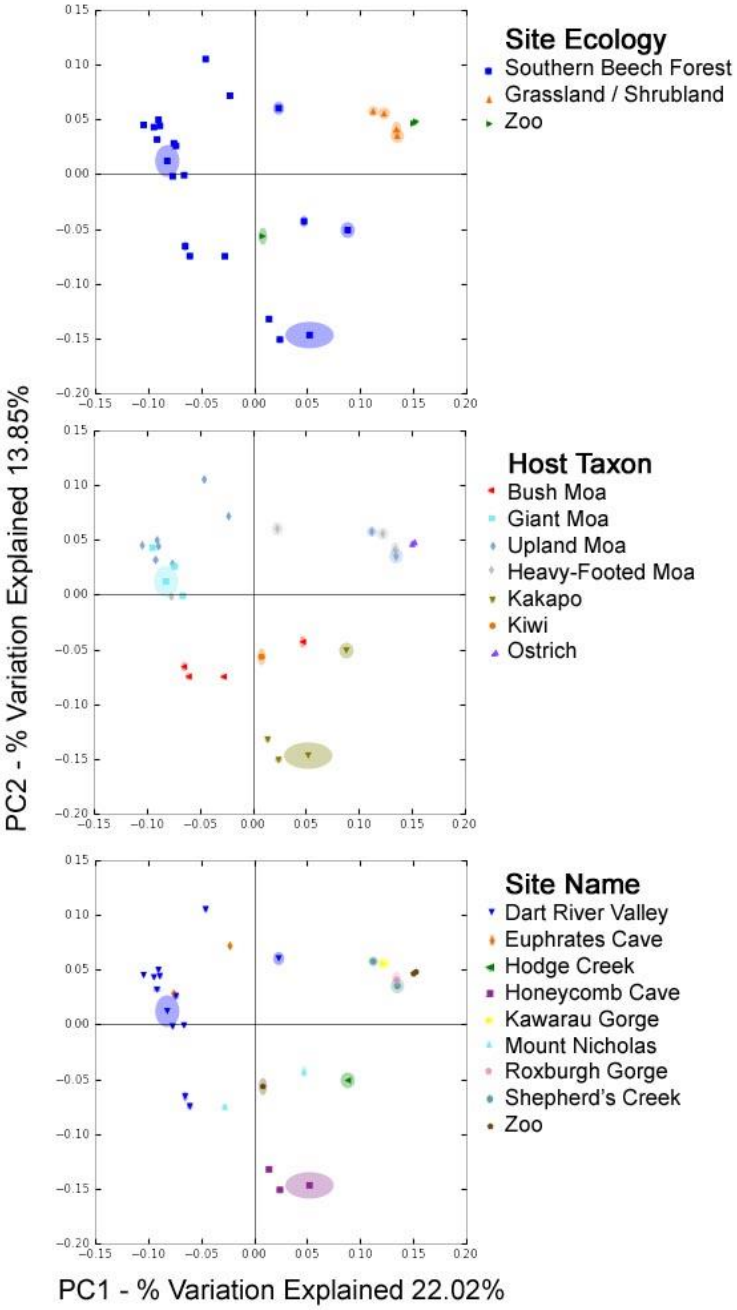


Fig. S10. PCoA (Principal Coordinates Analysis) plots, of jacknifed rarefied sampling of unifrac distances from fungal reads from bird samples only. Samples are clustered relative to general similarity between OTU assemblages of each sample. Circles around icons represent error due to differences between the rarefied datasets. Samples are unweighted (number of reads per OTU are not included). Grouped to ecology category, host taxon, and collection site.

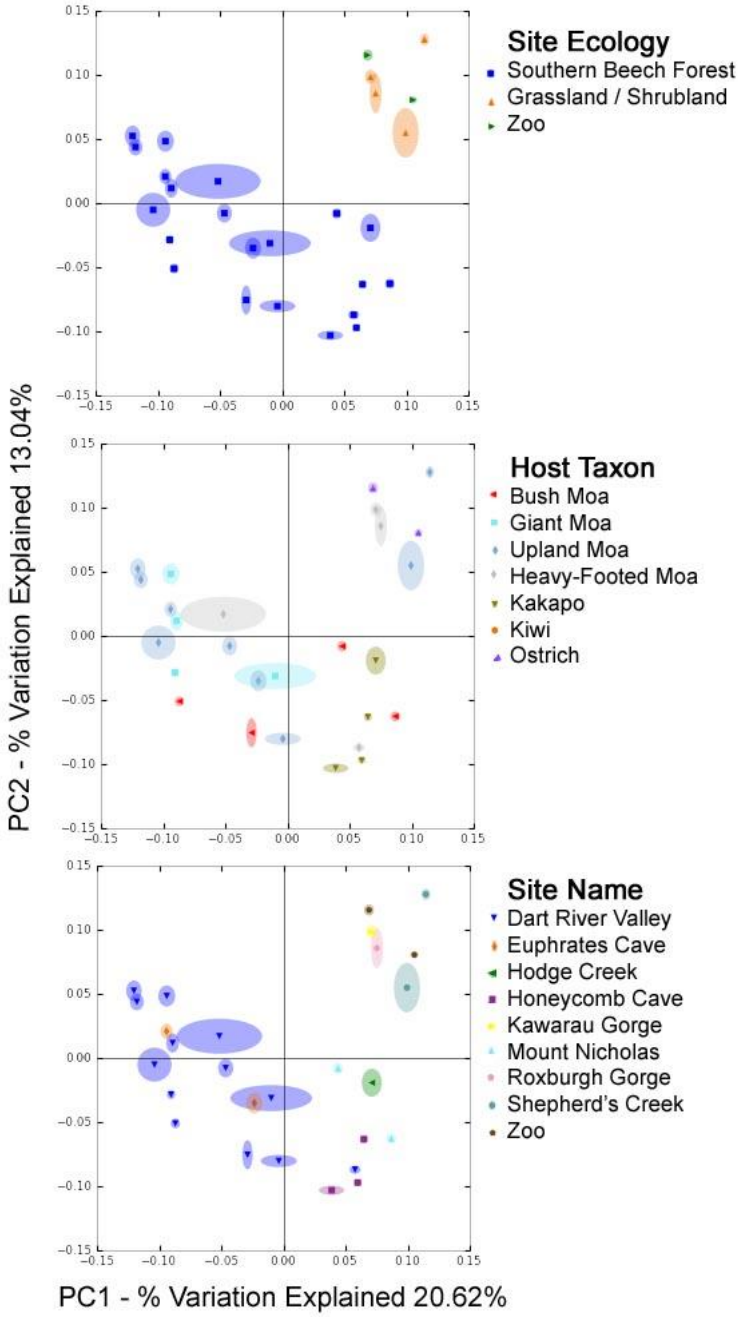


Fig. S11. PCoA (Principal Coordinates Analysis) plots, of jacknifed rarefied sampling of unifrac distances from plant reads from bird samples only. Samples are clustered relative to general similarity between OTU assemblages of each sample. Circles around icons represent error due to differences between the rarefied datasets. Samples are unweighted (number of reads per OTU are not included). Grouped to ecology category, host taxon, and collection site.

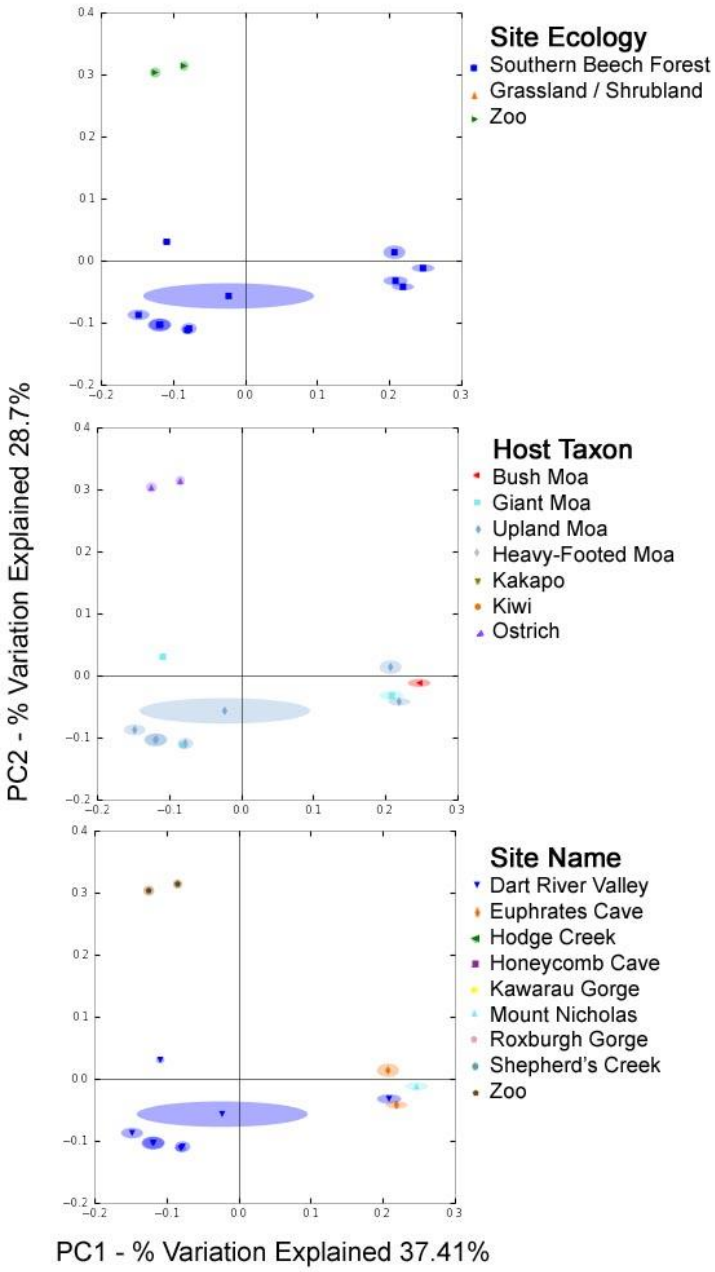


Fig. S12. PCoA (Principal Coordinates Analysis) plots, of jacknifed rarefied sampling of unifrac distances from parasite reads from bird samples only. Samples are clustered relative to general similarity between OTU assemblages of each sample. Circles around icons represent error due to differences between the rarefied datasets. Samples are unweighted (number of reads per OTU are not included). Grouped to ecology category, host taxon, and collection site.

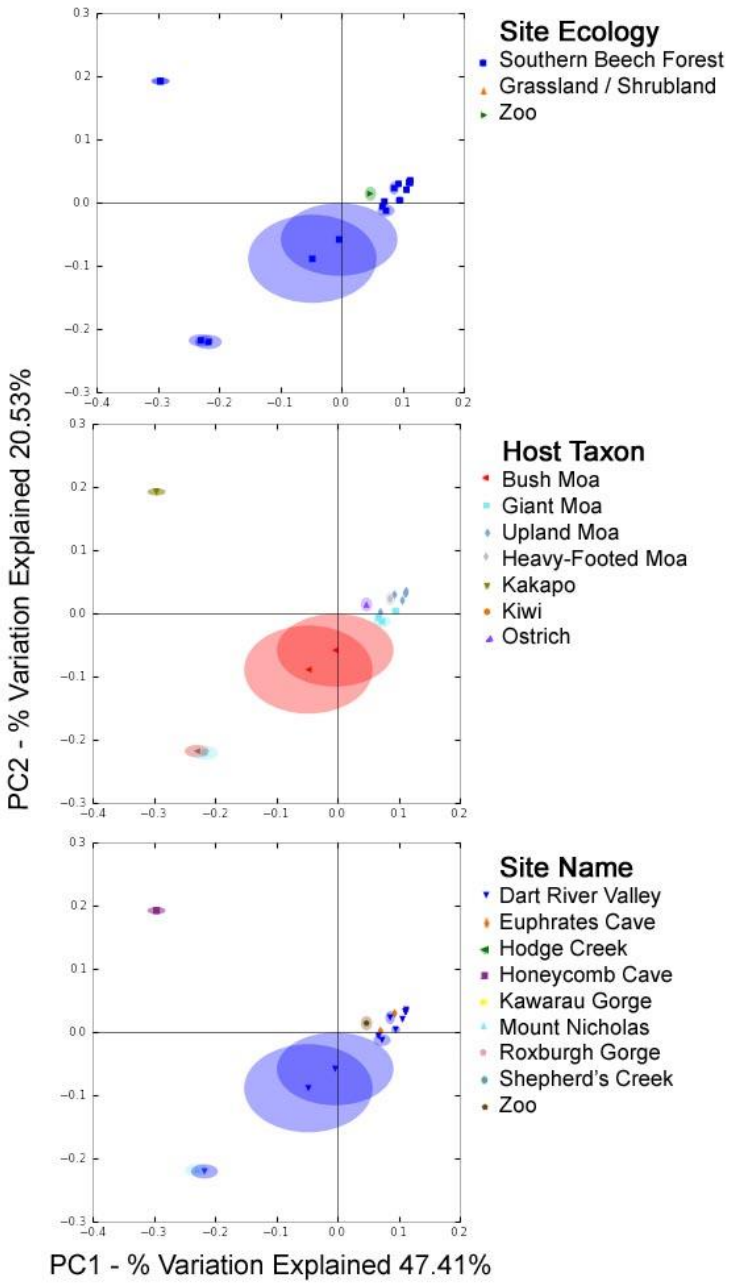


Fig. S13. Bayesian phylogeny of available Cortinariaceae (Fungi: phylum Basidiomycota: class Agaricoycetes: order Agaricales: family Cortinariaceae) sequences of the 18S barcoding region, inclusive of sequenced OTUs identified as members of this taxon group. *Lepiota cristata* (Agaricomycetes: family Agaricaceae) used as outgroup. Bayesian posterior probabilities shown.

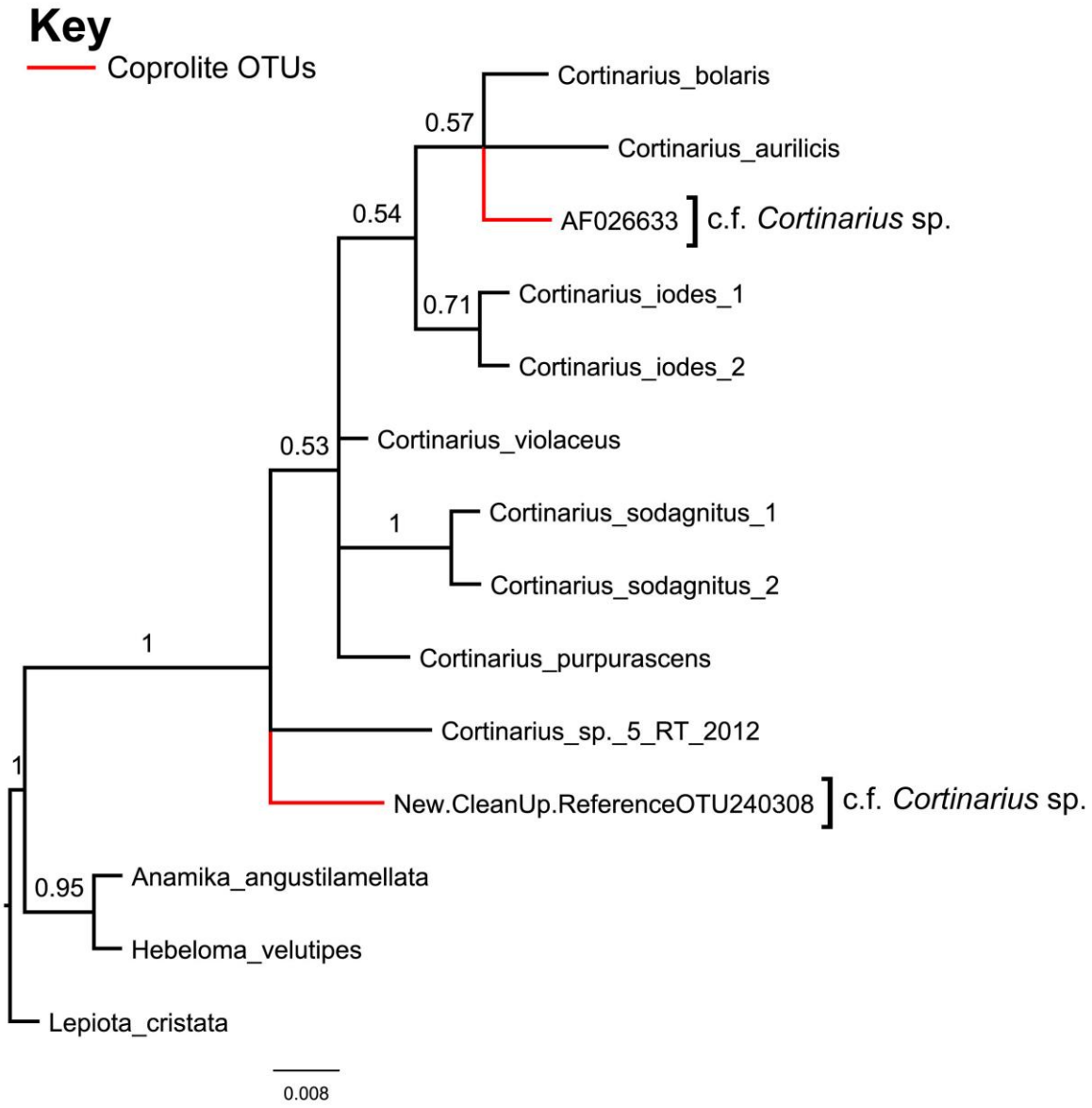


Fig. S14. Bayesian phylogeny of available Inocybaceae (Fungi: phylum Basidiomycota: class Agaricomycetes: order Agaricales: family Inocybaceae) sequences of the 18S barcoding region, inclusive of sequenced OTUs identified as members of this taxon group. *Lepiota cristata* (Agaricomycetes: family Agaricaceae) used as outgroup. Bayesian posterior probabilities shown.

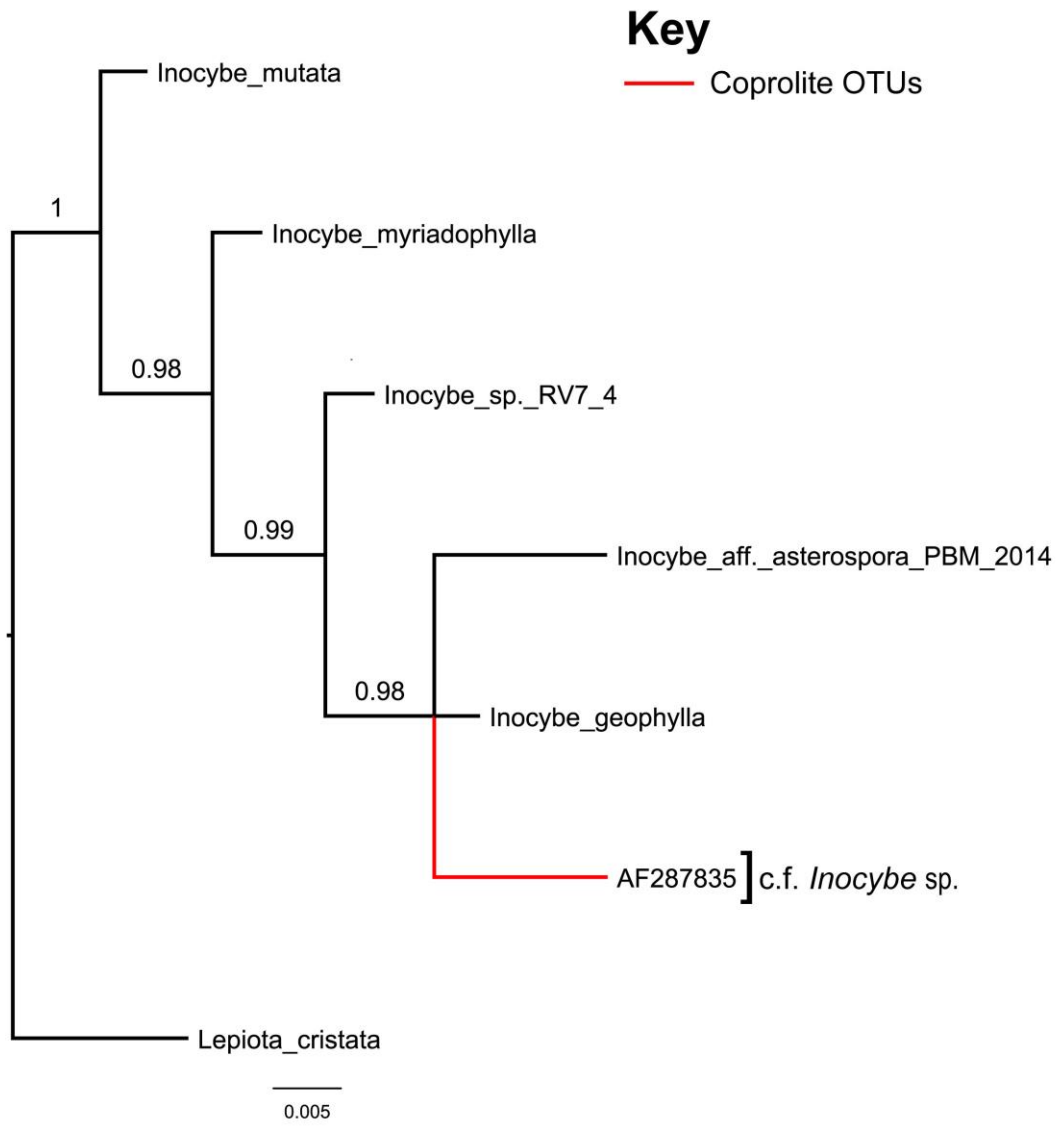


Fig. S15. Bayesian phylogeny of available Physalacriaceae (Fungi: phylum Basidiomycota: class Agaricoycetes: order Agaricales: family Physalacriaceae) sequences of the 18S barcoding region, inclusive of sequenced OTUs identified as members of this taxon group. *Lepiota cristata* (Agaricomycetes: family Agaricaceae) used as outgroup. Bayesian posterior probabilities shown.

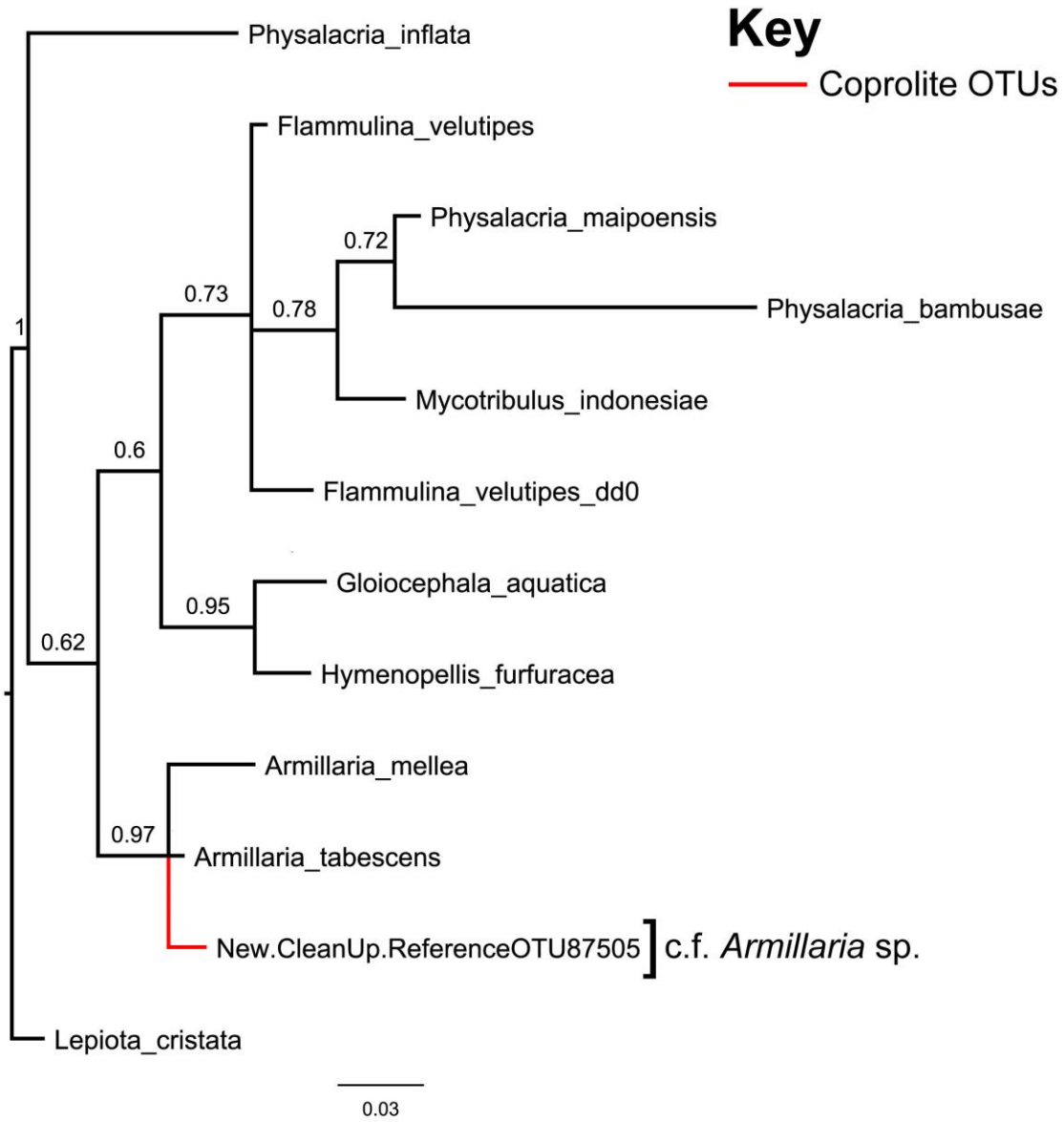


Fig. S16. Bayesian phylogeny of available Eimeriidae (Alveolata: phylum Apicomplexa: class Conoidasida: order: Eucoccidiorida: suborder Eimeriorina: family Eimeriidae) sequences of the 18S barcoding region, inclusive of sequenced OTUs identified as members of this taxon group. *Calyptospora funduli* (Eimeriorina: family Calyptosporiidae) used as outgroup. Bayesian posterior probabilities shown.

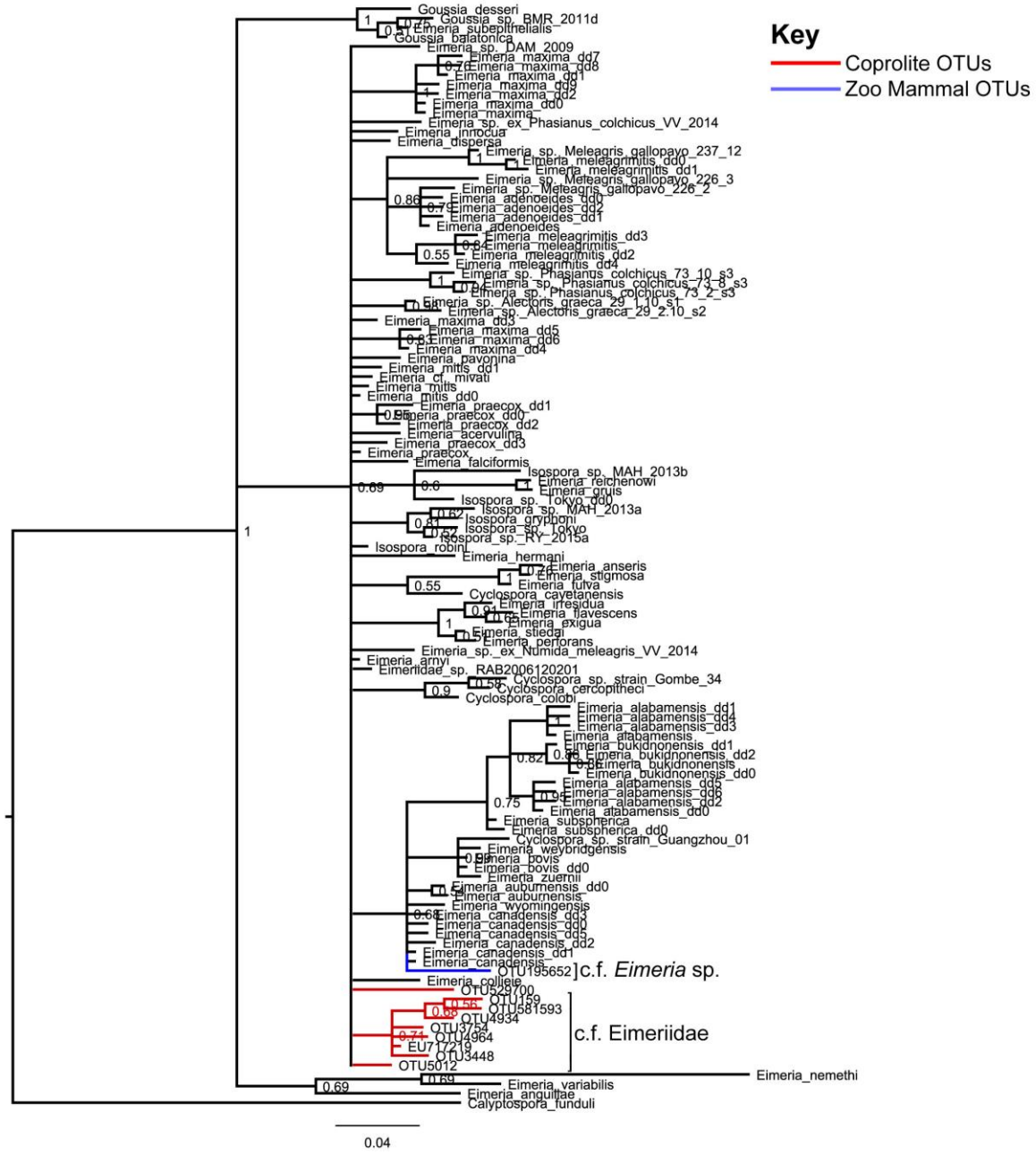


Fig. S17. Bayesian phylogeny of available Sarcocystidae (Alveolata: phylum Apicomplexa: class Conoidasida: order: Eucoccidiorida: suborder Eimeriorina: subfamily Eimeriorina: family Sarcocystidae) sequences of the 18S barcoding region, inclusive of sequenced OTUs identified as members of this taxon group. *Isospora* sp. (Eimeriorina: family Eimeriidae) used as outgroup. Bayesian posterior probabilities shown.

Fig. S18. Bayesian phylogeny of available Ascaridida (Nematoda: class Chromadorea: order: Acaridida) sequences of the 18S barcoding region, inclusive of sequenced OTUs identified as members of this taxon group. Guinea worm *Dracunculus medinensis* (Chromadorea: order Camallanida) used as outgroup. Bayesian posterior probabilities shown.

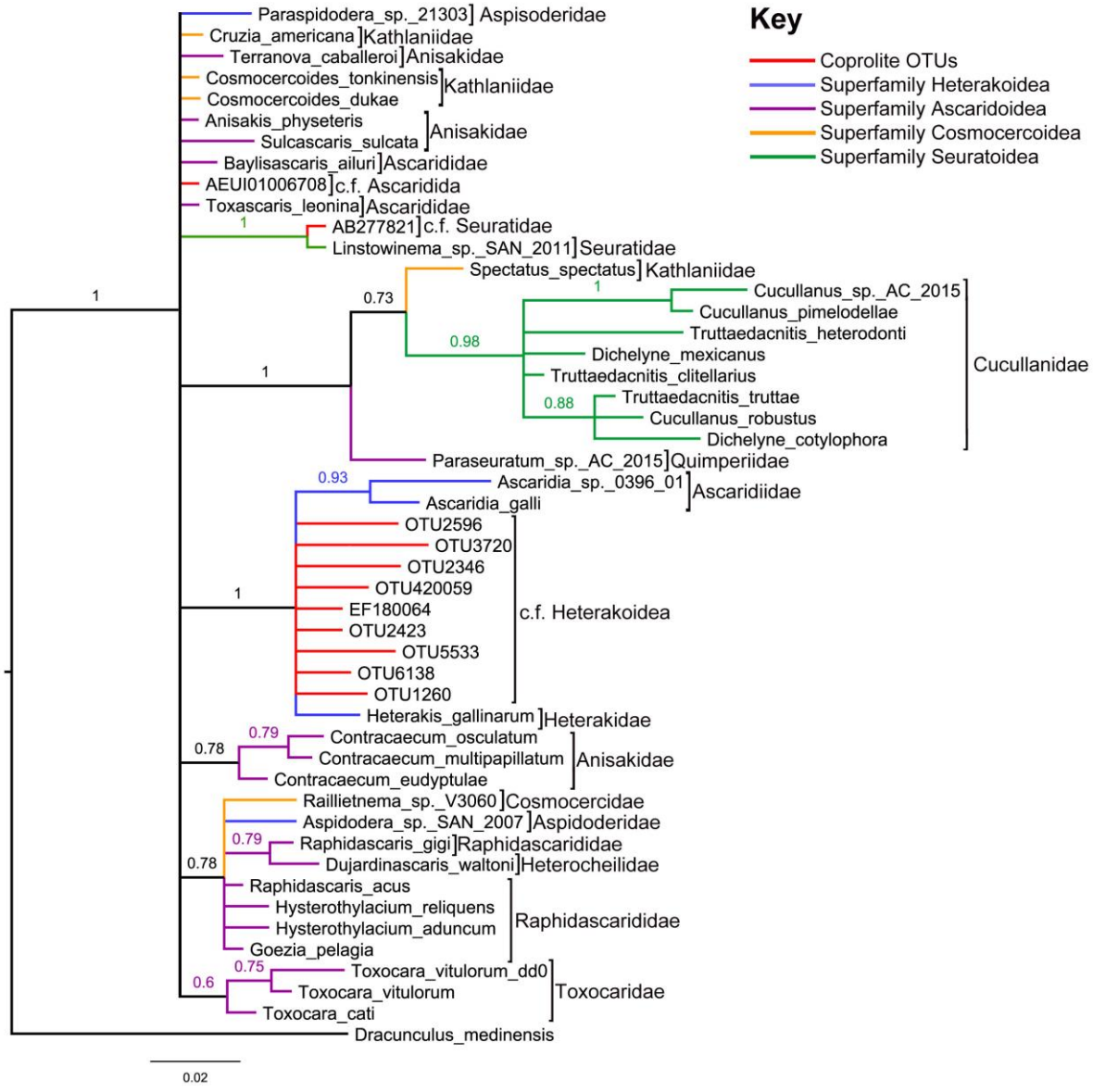


Fig. S19. Bayesian phylogeny of available Pronocephaloidea (Platyhelminthes: class Trematoda: order Echinostomida: superfamily Pronocephaloidea) sequences of the 18S barcoding region, inclusive of sequenced OTUs identified as members of this taxon group. Sheep liver fluke *Fasciola hepatica* (Echinostomida: superfamily Echinostomatoidea) used as outgroup. Bayesian posterior probabilities shown.

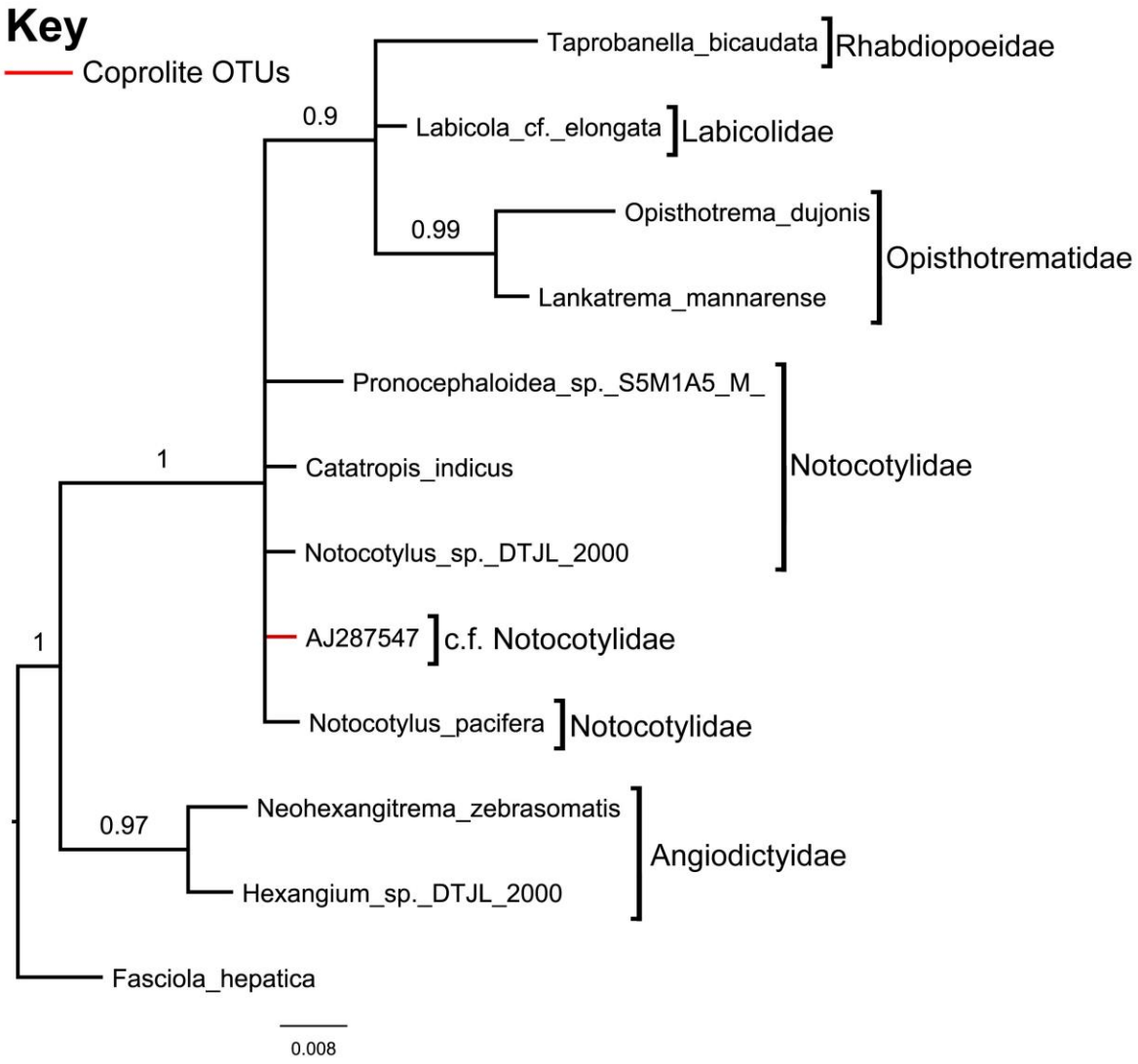
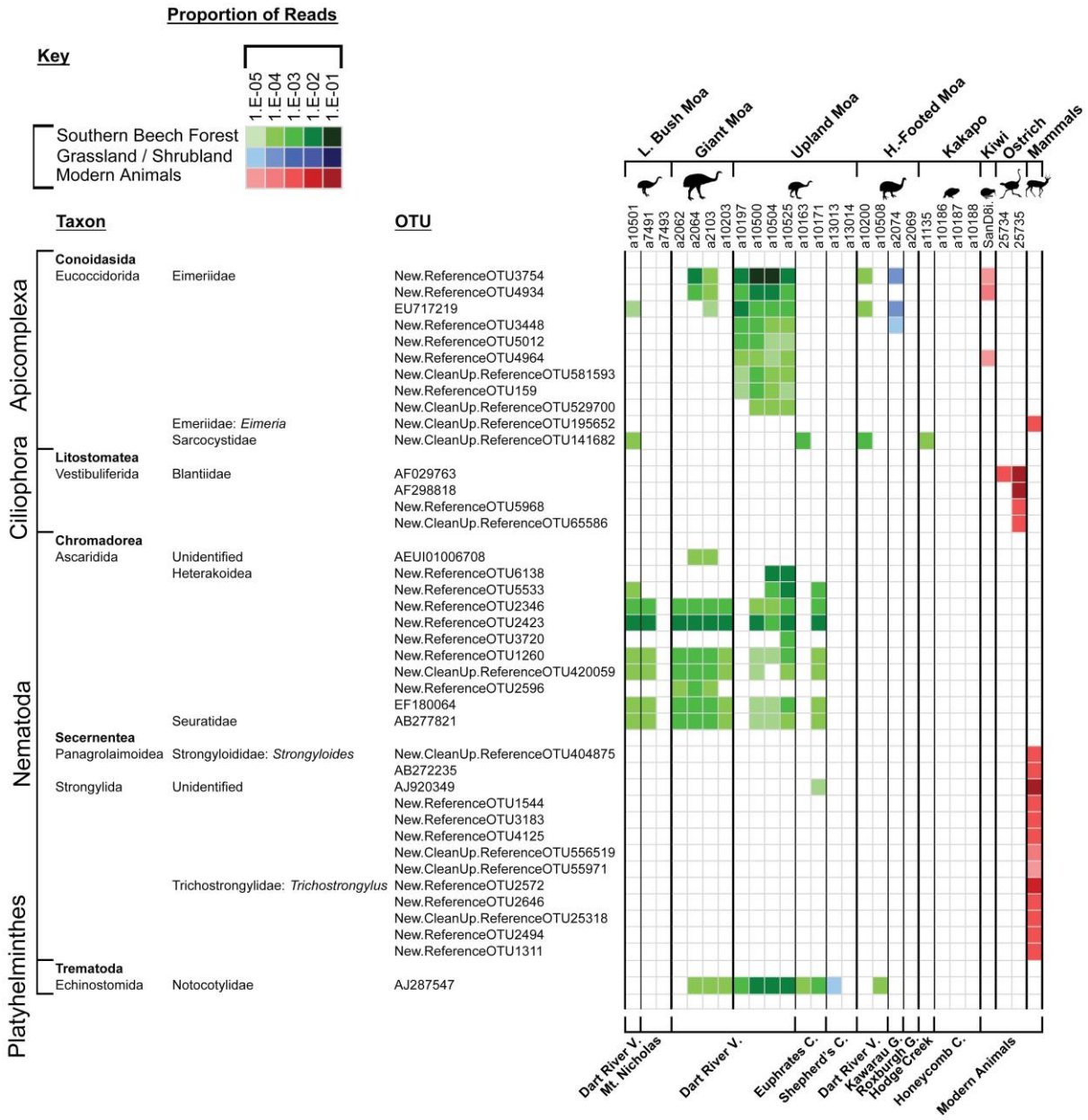


Fig. S20. Chart of parasite distributions. ‘Mammals’ group comprises average of mammals with parasites present only.



Tables

Table S1: Table of site details. Note that coprolite localities in New Zealand are also reviewed in detail in Wood & Wilmshurst (1). Species codes: L = little bush Moa; G = giant moa, U = upland moa, H = heavy-footed moa, K = kakapo.

Site	Site Type	Species Known from Site	Region	Palaeovegetation	Radiocarbon Age Ranges	L	G	U	H	K	References
Dart River Valley (Daley's Flat)	Rock Shelter	*Anomalopteryx didiformis* (little bush moa) *Dinornis robustus* (S.I. giant moa) *Megalapteryx didinus* (upland moa) *Pachyornis elephantopus* (heavy-footed moa)	West Otago	Southern Beech Forest	664 ± 25 - 853 ± 25 (2)	1	4	4	2		(2–4)
Hodge Creek Cave System	Cave	*Strigops habroptilus* (kakapo)	North West Nelson	Southern Beech Forest	725 ± 43 (5)					1	(5, 6)
Honeycomb Cave System	Cave	*Strigops habroptilus* (kakapo)	North West Nelson	Southern Beech Forest	172 ± 20 - 2514 ± 43 (this study) (5)					3	(5, 7, 8)
Euphrates Cave	Cave	*Megalapteryx didinus* (upland moa)	North West Nelson	Subalpine Southern Beech Forest	694 ± 30 - 6368 ± 31 (9)			2			(9, 10)
Shepherd's Creek, Waitaki Valley	Rock Shelter	*Megalapteryx didinus* (upland moa)	North Otago	Dry shrubland, herbfield	Unknown			2			(11)
Mount Nicholas (Possum's Rockshelter)	Rock Shelter	*Anomalopteryx didiformis* (little bush moa) *Dinornis robustus* (S.I. giant moa)	Central Otago	Southern Beech / Podocarp-Broadleaf Forest	1440 ± 30 - 1582 ± 34 (12)	2					(12)
Roxburgh Gorge (Rockshelter B)	Rock Shelter	*Pachyornis elephantopus* (heavy-footed moa)	Central Otago	Dry shrubland, herbfield	Unknown				1		(3, 13, 14)
Kawarau Gorge	Rock Shelter	*Pachyornis elephantopus* (heavy-footed moa)	Central Otago	Dry shrubland, herbfield	Unknown				1		(3, 13, 14)

References (Table S1).

1. Wood JR, Wilmshurst JM (2014) Late Quaternary terrestrial vertebrate coprolites from New Zealand. *Quat Sci Rev* 98(0):33–44.
2. Wood JR, Wilmshurst JM, Rawlence NJ (2011) Radiocarbon-dated faunal remains correlate very large rock avalanche deposit with prehistoric Alpine fault rupture. *N Z J Geol Geophys* 54(4):431–434.
3. Wood JR, et al. (2008) Coprolite deposits reveal the diet and ecology of the extinct New Zealand megaherbivore moa (Aves, Dinornithiformes). *Ice Age Refug Quat Extinctions Issue Quat Evol Palaeoecol* 27(27–28):2593–2602.
4. Wood JR, et al. (2013) Resolving lost herbivore community structure using coprolites of four sympatric moa species (Aves: Dinornithiformes). *Proc Natl Acad Sci* 110(42):16910–16915.
5. Horrocks M, et al. (2008) Plant microfossil analysis of coprolites of the critically endangered kakapo (*Strigops habroptilus*) parrot from New Zealand. *Rev Palaeobot Palynol* 149(3–4):229–245.
6. Worthy TH (1997) Fossil deposits in the Hodges Creek Cave System, on the northern foothills of Mt Arthur, Nelson. South Island, New Zealand. *Notornis* 44:111–108.
7. Worthy TH (1993) *Fossils of Honeycomb Hill* (Museum of New Zealand).
8. Wood JR, Wilmshurst JM, Worthy TH, Holzapfel AS, Cooper A (2012) A Lost Link between a Flightless Parrot and a Parasitic Plant and the Potential Role of Coprolites in Conservation Paleobiology. *Conserv Biol* 26(6):1091–1099.
9. Wood JR, et al. (2012) High-Resolution Coproecology: Using Coprolites to Reconstruct the Habits and Habitats of New Zealand's Extinct Upland Moa (*Megalapteryx didinus*). *PLoS ONE* 7(6):e40025.
10. Rowe P, Millar I, Worthy T (1994) Exploration on Garibaldi Ridge–Euphrates Cave, Kahurangi National Park. *N Z Speleol Bull* 9:271–290.
11. Trotter MM (1970) Archaeological investigations in the Aviemore area, South Island. *Rec Canterb Mus* 8(5):439–453.
12. Wood JR, Wilmshurst JM, Worthy TH, Cooper A (2012) First coprolite evidence for the diet of *Anomalopteryx didiformis*, an extinct forest ratite from New Zealand. *N Z J Ecol* 36(2):164.
13. Wood JR (2008) Pre-settlement Paleoecology of Central Otago's Semi-arid Lowlands, with Emphasis on the Pre-settlement Role of Avian Herbivory in South Island Dryland Ecosystems, New Zealand.
14. Wood JR, Wilmshurst JM (2013) Pollen analysis of coprolites reveals dietary details of heavy-footed moa (*Pachyornis elephantopus*) and coastal moa (*Euryapteryx curtus*) from Central Otago. *N Z J Ecol* 37(1):151–155.

Table S2. Table of all 23 coprolites and modern faecal samples sequenced in this study.

Specimen	Species	Locality	C^{14} Date	Sequenced Read Count	Depositor DNA	Depositor Haplotype	Fossil Content	DNA Content
A10501 (replicate 1)	*Anomalopteryx didiformis* (little bush moa)	Dart River Valley	664 ± 25 (1)	138532	(2)	C	(2, 3)	(2, 3)
A10501 (replicate 2)	*Anomalopteryx didiformis* (little bush moa)	Dart River Valley	664 ± 25 (1)	139065	(2)	C	(2, 3)	(2, 3)
A7491	*Anomalopteryx didiformis* (little bush moa)	Mount Nicholas	1440 ± 30 (4)	156679	(4)	N/A	(4)	
A7493	*Anomalopteryx didiformis* (little bush moa)	Mount Nicholas		139635	(4)	N/A	(4)	
A10203	*Dinornis robustus* (South Island giant moa)	Dart River Valley		56904	(2)	D3	(2, 3)	(2, 3)
A2062	*Dinornis robustus* (South Island giant moa)	Dart River Valley	721 ± 30 (1)	26102	(6)	D1	(3, 6)	(3, 6)
A2064	*Dinornis robustus* (South Island giant moa)	Dart River Valley	841 ± 30 (1)	88940	(6)	D6	(3, 6)	(3, 6)
A2103	*Dinornis robustus* (South Island giant moa)	Dart River Valley		76220	(6)	D2	(3, 6)	(3, 6)
A10197	*Megalapteryx didinus* (upland moa)	Dart River Valley		193425	(2)	M5	(2, 3)	(2, 3)
A10500 (replicate 1)	*Megalapteryx didinus* (upland moa)	Dart River Valley		90853	(2)	M4	(2, 3)	(2, 3)
A10500 (replicate 2)	*Megalapteryx didinus* (upland moa)	Dart River Valley		133408	(2)	M4	(2, 3)	(2, 3)
A10504	*Megalapteryx didinus* (upland moa)	Dart River Valley		227453	(2)	M1	(2, 3)	(2, 3)
A10525	*Megalapteryx didinus* (upland moa)	Dart River Valley	678 ± 25 (1)	257186	(2)	M6	(2, 3)	(2, 3)
A10163	*Megalapteryx didinus* (upland moa)	Euphrates Cave		174053	(5)	T	(3, 6)	(3, 5)
A10171	*Megalapteryx didinus* (upland moa)	Euphrates Cave		227405	(5)	T	(3, 5)	(3, 5)
A13013	*Megalapteryx didinus* (upland moa)	Shepherd's Creek		143787	This study	M1		
A13014	*Megalapteryx didinus* (upland moa)	Shepherd's Creek		100579	This study	M1		
A10200	*Pachyornis elephantopus* (heavy-footed moa)	Dart River Valley		73537	(2)	P2	(2, 3)	(2, 3)
A10508	*Pachyornis elephantopus* (heavy-footed moa)	Dart River Valley		137043	(2)	P2	(2, 3)	(2, 3)
A2069	*Pachyornis elephantopus* (heavy-footed moa)	Roxburgh Gorge		125775	(6)	P1	(3, 6, 7)	(3, 6)
A2074	*Pachyornis elephantopus* (heavy-footed moa)	Kawarau Gorge		67524	(6)	P3	(3, 6, 7)	(3, 6)
A10186	*Strigops habroptilus* (kakapo)	Honeycomb Cave	172 ± 30 (this study)	52021	This study	N/A		
A10187	*Strigops habroptilus* (kakapo)	Honeycomb Cave	192 ± 30 (this study)	96050	This study	N/A		
A10188	*Strigops habroptilus* (kakapo)	Honeycomb Cave	1020 ± 25 (8)	47517	(8)	N/A	(8)	
A1135	*Strigops habroptilus* (kakapo)	Hodge Creek		90386	This study	N/A		
SanD8i.1	*Apteryx mantelli* (North Island brown kiwi)	San Diego Zoo	Modern	137549	This study	N/A		
25734	*Struthio camelus* (ostrich)	Houston Zoo	Modern	27028	This study	N/A		
25735	*Struthio camelus* (ostrich)	Houston Zoo	Modern	102543	This study	N/A		

References (Table S2)

1. Wood JR, Wilmshurst JM, Rawlence NJ (2011) Radiocarbon-dated faunal remains correlate very large rock avalanche deposit with prehistoric Alpine fault rupture. *N Z J Geol Geophys* 54(4):431–434.
2. Wood JR, et al. (2013) Resolving lost herbivore community structure using coprolites of four sympatric moa species (Aves: Dinornithiformes). *Proc Natl Acad Sci* 110(42):16910–16915.
3. Wood JR, et al. (2013) A Megafauna's Microfauna: Gastrointestinal Parasites of New Zealand's Extinct Moa (Aves: Dinornithiformes). *PLoS ONE* 8(2):e57315.
4. Wood JR, Wilmshurst JM, Worthy TH, Cooper A (2012) First coprolite evidence for the diet of *Anomalopteryx didiformis*, an extinct forest ratite from New Zealand. *N Z J Ecol* 36(2):164.
5. Wood JR, et al. (2012) High-Resolution Coproecology: Using Coprolites to Reconstruct the Habits and Habitats of New Zealand's Extinct Upland Moa (*Megalapteryx didinus*). *PLoS ONE* 7(6):e40025.
6. Wood JR, et al. (2008) Coprolite deposits reveal the diet and ecology of the extinct New Zealand megaherbivore moa (Aves, Dinornithiformes). *Ice Age Refug Quat Extinctions Issue Quat Evol Palaeoecol* 27(27–28):2593–2602.
7. Wood JR, Wilmshurst JM (2013) Pollen analysis of coprolites reveals dietary details of heavy-footed moa (*Pachyornis elephantopus*) and coastal moa (*Euryapteryx curtus*) from Central Otago. *N Z J Ecol* 37(1):151–155.
8. Wood JR, Wilmshurst JM, Worthy TH, Holzapfel AS, Cooper A (2012) A Lost Link between a Flightless Parrot and a Parasitic Plant and the Potential Role of Coprolites in Conservation Paleobiology. *Conserv Biol* 26(6):1091–1099.

Table S3. Results from Adonis (PERMANOVA) Beta-diversity analyses. All analyses returned significant p-values (not presented). Values represent the estimated proportion of variation between samples explained by each selected variable. Habitat type sorted samples into either “Modern”, “Southern Beech” or “Semi-Arid”.

Variable	All Samples	Birds Only	Coprolites Only	Moa Only
Habitat Type	0.37052	0.24029	0.15215	0.23226
Site Specific	0.53008	0.56722	0.53518	0.50041
Host Genus	0.69482	0.44421	0.34115	0.27814

Table S4. Table of Bayesian phylogenetic trees run in this study, demonstrating taxon group, identified taxa, model selection, and MCMC generations used.

Taxon	Rank	Higher taxon / ecology	Identified coprolite taxa	Outgroup	# OTUs included	Generations	Model Selected	Average S.D. of Split Frequencies
Physalacriaceae	Family	Fungi	*Armillaria* sp.	*Lepiota cristata*	1	25000000	GTR+G	0.007057
Inocybaceae	Family	Fungi	*Inocybe* sp.	*Lepiota cristata*	1	10000000	JC	0.006434
Cortinariacae	Family	Fungi	*Cortinarius* spp.	*Lepiota cristata*	2	25000000	GTR	0.008921
Sarcocystidae	Family	Apicomplexa Parasites	*Sarcocystis* sp.	*Isospora orloxi*	1	25000000	HKY+G	0.007261
Eimeriidae	Family	Apicomplexa Parasites	*Eimeria* sp., Eimeriidae spp.	*Calyptospora funduli*	10	50000000	GTR+G	0.006167
Ascaridida	Order	Nematoda Parasites	Heterakoidea spp., Seuratidae sp., Ascaridida sp.	*Dracunculus medinensis*	11	10000000	HKY+G	0.009734
Pronocephaloidea	Superfamily	Trematoda Parasites	Notocotylidae sp.	*Fasciola hepatica*	1	25000000	SYM+G	0.006288

Dataset Legends

Dataset S1. Modified and formatted, tabulated biom file from all samples, displaying raw read counts (post-filtering). Separate columns are displayed for the original QIIME-based and later MEGAN-based taxonomies. The various mapping categories used for QIIME analyses are also shown. Values in the grey-coloured cells represent total count of filtered reads.

Dataset S2. Results from Kruskal-Wallis Beta-diversity analyses, identifying significant differences in diversity between mapping categories.

Dataset S3. All clustered 18S rRNA sequences found in this study (.txt formatted FASTA file).

Dataset S4. Comparison between plant and parasite taxa identified in all the 23 coprolites analysed in this study, using NGS data (this study), and Sanger-sequencing and fossil (pollen, spores, macrofossils) data (past study). Numbers refer to total number of identified different taxa, DNA sequences or DNA OTUs. Values with *, did not have individual counts per sample available, and these values instead refer to the sum total from all samples from the same moa taxon and collection site.

References (Dataset S4):

1. Wood JR, et al. (2008) Coprolite deposits reveal the diet and ecology of the extinct New Zealand megaherbivore moa (Aves, Dinornithiformes). *Ice Age Refug Quat Extinctions Issue Quat Evol Palaeoecol* 27(27–28):2593–2602.
2. Wood JR, et al. (2013) Resolving lost herbivore community structure using coprolites of four sympatric moa species (Aves: Dinornithiformes). *Proc Natl Acad Sci* 110(42):16910–16915.
3. Wood JR, et al. (2013) A Megafauna's Microfauna: Gastrointestinal Parasites of New Zealand's Extinct Moa (Aves: Dinornithiformes). *PLoS ONE* 8(2):e57315.
4. Wood JR, Wilmshurst JM, Worthy TH, Cooper A (2012) First coprolite evidence for the diet of *Anomalopteryx didiformis*, an extinct forest ratite from New Zealand. *N Z J Ecol* 36(2):164.
5. Wood JR, Wilmshurst JM, Worthy TH, Holzapfel AS, Cooper A (2012) A Lost Link between a Flightless Parrot and a Parasitic Plant and the Potential Role of Coprolites in Conservation Paleobiology. *Conserv Biol* 26(6):1091–1099.
6. Wood JR, et al. (2012) High-Resolution Coproecology: Using Coprolites to Reconstruct the Habits and Habitats of New Zealand's Extinct Upland Moa (*Megalapteryx didinus*). *PLoS ONE* 7(6):e40025.
7. Wood JR, Wilmshurst JM (2013) Pollen analysis of coprolites reveals dietary details of heavy-footed moa (*Pachyornis elephantopus*) and coastal moa (*Euryapteryx curtus*) from Central Otago. *N Z J Ecol* 37(1):151–155.