

Supplementary Information for

Subsistence practices, past biodiversity, and anthropogenic impacts revealed by New Zealand-wide ancient DNA survey

Frederik V. Seersholm, Theresa L. Cole, Alicia Grealy, Nicolas J. Rawlence, Karen Greig, Michael Knapp, Michael Stat, Anders J. Hansen, Luke J. Easton, Lara Shepherd, Alan J. D. Tennyson, R. Paul Scofield, Richard Walter & Michael Bunce

Frederik Valeur Seersholm Email: frederikseersholm@gmail.com

Michael Bunce Email: michael.bunce@curtin.edu.au

This PDF file includes:

Supplementary text Figs. S1 to S6 Tables S1 to S10 Captions for additional data table S11 References for SI

Other supplementary materials for this manuscript include the following:

Additional data table S11

Supplementary Information Text

1. Contamination and introduced species

To monitor laboratory contamination, at least two negative controls were included in each batch of sample preparations (see Table S5). Extraction and PCR-blanks represent notemplate controls for the extraction and PCR setup protocols, respectively. Grinding blanks were prepared by running the ball mill with 15 mL of ultrapure water in the grinding pod and concentrating the water to 500 uL using an Amicon®Ultra-4 Centrifugal Filter (Millipore).

In the controls we identify a low level of background contamination from the three well-known contaminants *Homo sapiens*, Phasianinae (presumably *Gallus gallus*) and *Sus* (1) (see Table S7). Accordingly, reads assigned to these three species in the bulk-bone samples were excluded from downstream analyses. In addition, we identified one critical contamination from red kangaroo (*Marcopus rufus*) in a single sample from Cobden cave. As ancient DNA from this species has been studied previously in our laboratory, and since only 58 contaminant reads are identified out of 37,386 (0.16%), it is likely that the DNA represents low-level background contamination. Furthermore, red kangaroo has never been present in New Zealand prior to European arrival, and was not amongst the species introduced by Europeans (e.g. wallabies). To investigate whether contamination occurred at the bone grinding, DNA extraction, or PCR amplification stages, the remaining bones from Cobden cave were ground and extracted (Cob_A_NZ4), two new extracts were made from the original bone powder (Cob_B_NZ4, Cob_C_NZ4), and a second amplification was carried out from the original extract (Cob_NZ1). As none of these yielded any DNA from *M. rufus*, and since no DNA of *M. rufus* was amplified in any of our blanks or other test samples, we conclude that this is a product of a single contamination incidence. The *M. rufus* DNA is most likely a result of a contaminated tagged primer or a combination of post-amplification contamination and tag-jumping. Furthermore, most samples were prepared in batches from very different bulk-bone projects (e.g., geographically distinct sites) to better detect cross contamination between DNA extracts. Reassuringly, we do not detect species endemic to New Zealand in data from other projects or vice versa.

Additionally, a handful of introduced species are identified among the different samples analyzed. These fall into two categories: (i) common contaminants and (ii) other introduced species. The former group consists of the domesticates cattle (*Bos*), sheep (*Ovis*), goat (*Capra*), and cat (*Felis sylvestris*), and are by far the most commonly identified introduced species in the dataset, detected at 12 midden sites, seven paleontological sites and one mixed site. However, as these are well-known laboratory contaminants, it is challenging to determine whether this DNA is endogenous. Furthermore, sheep DNA has previously been shown to leak into moa deposits at the Hukanui pool site (2) and, like sheep, both goat and cattle produce large volumes of urine and fecal material, that could leak into both midden and cave deposits. Moreover, the notion that cattle and sheep DNA could stem from background contamination or DNA leaching is reinforced by the fact that these species, along with human and pig, are the only taxa identified in the samples from Lake Poukawa, where no endogenous DNA could be amplified. However, it is also possible that early historic-period deposition by Māori or Europeans at some sites can have contaminated bulk bone samples with historic introduced species. The cultural layers derived from pre-contact Māori activity typically represent a thin horizon and light sediment layer perturbation could cause overlying post-European layers to contaminate pre-historic layers with remains from introduced species. Lastly, while the detection of dog (*Canis lupus familiaris*) is expected in archaeological sites, we cannot rule out the possibility that the dog DNA could stem from laboratory contamination as dog is a common contaminant.

Taxa from the second group, on the other hand, most likely represent endogenous DNA from bones that have made their way into the deposits from the surface by postdepositional reworking of the sediment or post-European deposition. This group contains common brushtail possum (*Trichosurus vulpecula*), rabbit (*Oryctolagus cuniculus*), hare (*Lepus europaeus*), ship rat (*Rattus rattus*), and European goldfinch (*Carduelis carduelis*), and these species are detected much less frequently than group 1 (six paleontological and no archaeological sites). All of these species are found in New Zealand today and the detection of these corresponds well to the morphological record, where modern day species are identified occasionally in the upper layers of bone deposits (3-6).

2. Contributions and limitations of morphology and BBM

Both morphological analysis and BBM are challenged by similar inherent limitations, and each methodology cannot easily stand alone. In both methods, failure to detect a taxon does not necessarily translate to absence from the deposit, but can be explained by a number of factors: In BBM, the ability to detect species may be affected by primer specificity, variable DNA preservation and the availability of relevant reference sequences, whereas, for morphological identification, taphonomic factors and the level of taxonomic expertise available could cause species to be missed, or incorrectly assigned. In our data, this is illustrated by the detection of species that are frequently missed by morphological approaches, such as whales, frogs and passerine birds, along with the absence of commonly detected species such as the New Zealand owlet nightjar, which does not have a relevant reference sequence available. Furthermore, owing to the different number of bones sampled for each approach (hundreds for BBM and thousands for morphology), the number of taxa identified with BBM is generally lower than that of morphology. Hence, here, BBM offers a conservative estimate of past biodiversity, providing a lower threshold of the past species richness in New Zealand which can be built on in future research. In particular, the emphasis on spatial patterns in the current study, warrants an examination of undiagnostic bones from a single site with a long depositional history, which could expand on the data presented here by improving our understanding of the exact temporal dynamics of biodiversity turnover in New Zealand. Ideally, BBM will become more widely applied in palaeontological and archaeological practice (51), and seen as a powerful complement to traditional analyses.

3. Palimpsest deposits.

The midden deposits investigated here are examples of cumulative palimpsests (53), in which the lack of fine stratigraphic differentiation along with the mixing of materials prevent temporal precision. Midden deposits are a result of subsistence practices which are formed when leftover bones are discarded. This is a complex process affected not only by the hunting episode itself, but by transportation issues, butchering methods and cooking practices. Additionally, post-depositional factors such as reworking of sediments by rodents and commensals along with taphonomic processes variably affecting different parts of the assemblage further blurs the original pattern. As such, individual hunting episodes are not discernible within the midden sequence, rather the deposits provide information on subsistence practices over longer time spans. However, this also highlights the advantage of cumulative palimpsest midden assemblages: with the repetition of similar hunting events at the same location over long time spans, large deposits are formed which are more easily analyzed and provide strong evidence for general trends at one location.

4. Materials and Methods

4.1 Subsampling and extraction. Bone fragments were subsampled into bulk bone pools of approximately 100 (batch: NZ1) or 50 bones each (batches: NZ2-5; *SI Appendix*, Table S9) of roughly equal size. For larger bones, a small fragment was broken off and added to the subsample. Depending on the size of each bulk bone sample, between one and five bulk bone pools were subsampled for each site. All subsamples were ground using a Retsch PM200 Planetary Ball Mill at 400 rpm until pulverized. Extractions were carried out with approximately 100 mg of bone powder dissolved in digestion buffer (0.25 mg Proteinase K + 1 mL 0.5 m EDTA) and incubated with rotation over night at 55°C. Next, samples were centrifuged, and the supernatant was concentrated to 50 µL in a MWCO 30,000 Vivaspin 500

column (Sigma-Aldrich). Lastly, the DNA was purified using the MinElute PCR Purification Kit (Qiagen) following the manufacturer's instructions, except for the use of a modified binding buffer optimized for ancient DNA(46) (40% Isopropanol, 0.05% µL Tween 20, 90 nm NaAc and 5M GuanHydCh in Ultrapure water).

4.2 Barcoding. All barcoding primers were fused to Illumina sequencing adaptors and tagged by a unique combination of 6-8 bp indexes on each primer. For metabarcoding assays, 25 µL PCR reactions were setup with 1 μ L of template DNA, 1X buffer (ThermoFisher), 2mM MgCl₂ 0.25mM dNTPs, 1U AmpliTaq Gold[®] DNA Polymerase, 0.6 µL 5X SYBR green, 0.4 mg/ml BSA and 0.4 μ M of primers. DNA was amplified on an Applied Biosystems StepOnePlus™ Real-Time PCR System with the following cycling conditions: an initial denaturation step at 95°C for 10 min, followed by 50 cycles of 95°C for 30 sec, 57/54°C for 30 sec (*SI Appendix*, Table S2) and 72°C for 45 sec; and a final elongation step of 72°C for 10 min. For the kākāpō specific barcoding assay, PCR reactions where setup as described above with one of the primer sets described in Bergner et al. (Table S2). This region was chosen because it had the fewest low complexity strings and length variable regions of the 4 assays described by Bergner et al. PCR reactions were setup in four replicates for each extract under the following cycling conditions: an initial denaturation step at 95° C for 5 min, followed by 50 cycles of 95°C for 30 sec, 52/58°C for 30 sec (*SI Appendix*, Table S2) and 72°C for 2 min; and a final elongation step of 72°C for 10 min. After PCR amplification, all reactions (metabarcoding and kākāpō specific assays) were pooled in equimolar concentrations based on end point fluorescence intensity, and LabChip GX Touch HT (PerkinElmer) quantifications. Lastly, amplicons were sequenced unidirectionally on the Illumina MiSeq platform using the 300-cycle MiSeq Reagent Nano Kit v2 (Illumina) with custom sequencing primers.

4.3 Sequence analysis. Sequences were filtered using a custom-made pipeline based on OBItools(56) (http://www. grenoble.prabi.fr/trac/OBITools; *SI Appendix*, Table S10). Firstly, samples were demultiplexed and adapter and primer sequences were removed using ngsfilter, only retaining reads with a 100% match to both tags and primers. Next, samples were dereplicated (obiuniq), discarding reads shorter than 80 bp, or reads represented by fewer than 5 replicates (obigrep -l 80 -p "count>=5"). To account for PCR and sequencing errors, the data was denoised using the obiclean program (obiclean -r 0.2 -d 2 -H) and Sumaclust, collapsing clusters at 95% and 93% with abundance thresholds of 50% and 1%, respectively, in a sample-wise manner. Lastly, chimeric sequences were removed using vsearch (vsearch - -uchime_denovo)(57).

4.4 Taxonomic assignments. Raw taxonomic assignments were achieved using the script blast getLCA.py (https://github.com/frederikseersholm/blast getLCA) where reads are assigned to the taxonomic node of the lowest common ancestor of the best hits to the database based on percent similarity. Briefly, filtered and denoised sequences were queried against the NCBI nt database $(ftp://ftp.ncbi.nlm.nih.gov/blast/db/nt*gz)$, downloaded 31st of August 2017(58)) and a local database of the reference sequences for *Leiopelma markhami*, *Nestor meridionalis, Circus teauteensis* and *Aquila moorei* using the megablast algorithm (59). Next, blast-files were parsed using the script blast_getLCA.py assigning each read to the taxonomic node of the lowest common ancestor of the best hits to the database based on percent similarity. To ensure that species level identifications were only called based on high confidence alignments, species level assignments were made only on hits with above 98% similarity to the reference, with assignments between 95% and 98% called at the genus level (7). Lastly, after this raw taxonomic assignment, each taxonomic node was examined and correlated to records of species present in New Zealand and the availability of reference sequences of similar species known from New Zealand (NCBI genbank). In cases of lower taxonomic nodes only represented by a single New Zealand species, the assignment was upgraded to species level, e.g., *Falco* was reassigned to *Falco novaeseelandiae* as this is the

only falcon present in New Zealand. Likewise, in cases of species level assignments where similar species within the same genus from New Zealand were absent from the reference database, the assignment was downgraded to genus level. e.g., *Cyanoramphus novaezelandiae* was reassigned to *Cyanoramphus*, as other *Cyanoramphus* species from New Zealand are absent from the database.

Fig. S1. Avifauna structure. Coordination analysis of bird species composition from paleontological (gray trangles) and archaeological sites (red trangles) based on presence/absence. Sites represented by three or fewer bird taxa have been omitted.

Fig. S2. Bird size and abundance comparisons. Only bird taxa identified at species or subspecies level were considered. Estimates of bird weights are the averages of male and female weights from http://nzbirdsonline.org.nz/.

Fig. S3. *Leiopelma* **spp. diversity.** UPGMA consensus tree with 10,000 bootstraps of 12S and 16S barcodes assigned within Leiopelmatidae. *L. markhami* reference sequences were generated for this study while the *L. pakeka* 12S sequence was provided by Luke Easton and Nicolas Rawlence.

Fig. S4. Site locations.

Bulk bone sites where each species was detected Ω

Time range for extinct birds

Time range for extinct herpetofauna

Fig. S5. Extinction times. Extinction times (AD) for extinct species detected in the current study, with information on the number of assemblages in which each species was detected. Sites were binned in Pleistocene, Holocene and pre-historic Māori assemblages (see table S1).

Fig. S6. In silico evaluation of pilot whale primer binding. Primer binding sites for the five whale species detected in this study, along with the two species of pilot whale. For the 16S primer set, the pilot whales have no mismatches to the primer sequences, whereas, for the 12S primer set, there is a single 5' mismatch to the reverse primer unlikely to impact on PCR efficacy.

Table S1. Site descriptions.

Table S2. Primers.

Table S3. Mammal species identified. \land Upgrade (family \Rightarrow species). \lor Downgrade (species -> family).> Inferred (species > species). A: archaeological sites, P: paleontological sites.

***Common laboratory contaminant**

Table S4. Fish and shark species identified. \land Upgrade (family \Rightarrow species). \lor Downgrade (species -> family). > Inferred (species > species). N: North Island sites, S: South Island sites.

*Found in a palaeontological site

Table S5. Bird species identified. Only lowest taxonomic nodes detected are shown. \wedge Upgrade (family -> species). \vee Downgrade (species -> family). E: endemic, I: introduced, N: native, A: archaeological sites, P: paleontological sites.

Family	Taxon	Common Name	NZ status	Conservation status	- P
Sphenodontidae	Sphenodon punctatus	Tuatara	Endemic	Relict	$\overline{\mathbf{3}}$
Scincidae	Oligosoma [^] sp.	Skinks	Native	٠	$\overline{1}$
Leiopelmatidae	Leiopelma archevi	Archey's frog	Endemic	Vulnerable	$\overline{1}$
	Leiopelma hamiltoni	Hamilton's frog	Endemic	Nationally critical	$\overline{1}$
	Leiopelma markhami	Markham's frog	Endemic	Extinct	3

Table S6. Reptiles identified. A: archaeological sites, P: paleontological sites.

Table S7. Taxa detected in control samples. GB: Grinding blank, EB: Extraction blank, PB: PCR blank.

Control	Sample						
type	prep. batch	Homo sapiens	Sus	Phasianinae			
F _{R1}	NZ1	13416	988				
EB ₂	NZ1	5991					
PB.	NZ1	no amplification					
GB	NZ ₂	no amplification					
EВ	NZ ₂	no amplification					
GB	NZ3	61137	n	145			
PB.	NZ3	no amplification					
F _B	NZ4	6418	O	O			
PB	NZ4	11256	ŋ	ŋ			
EВ	NZ5	3999	O	n			
РB	NZ5	2934					

Table S8. Sample descriptions, blanks.

Table S10. Post filtering DNA amplification descriptions

Additional data table S11 (separate file)

Raw taxonomic assignments for all sites studied along with information on reassigned taxa. Numbers in each field of the table designate the number of separate subsamples in which a given taxa was identified in each site.

References

- 1. Leonard JA, et al. (2007) Animal DNA in PCR reagents plagues ancient DNA research. *J Archaeol Sci* 34(9):1361–1366.
- 2. Haile J, et al. (2000) Ancient DNA Chronology within Sediment Deposits : Are Paleobiological Reconstructions Possible and Is DNA Leaching a Factor ? *Mol Biol Evol*. doi:10.1093/molbev/msm016.
- 3. Worthy TH, Zhao JX (2006) A late Pleistocene predator-accumulated avifauna from Kids Cave, West Coast, South Island, New Zealand. *Alcheringa, Spec Issue,* 1(August):389–408.
- 4. Worthy TH, Roscoe D (2003) Takaka Fossil Cave a stratified Late Glacial to Late Holocene deposit from Takaka Hill, New Zealand. *Tuhinga* 14(January):41–60.
- 5. Worthy TH, et al. (2002) A rich Pleistocene-Holocene avifaunal sequence from Te Waka #1: terrestrial fossil vertebrate faunas from inland Hawke's Bay, North Island, New Zealand. *Tuhinga* 13:1–38.
- 6. Worthy TH (1997) Quaternary fossil fauna of South Canterbury, South Island, New Zealand. *J R Soc New Zeal* 27(1):67–162.
- 7. Grealy A, et al. (2016) Tropical ancient DNA from bulk archaeological fish bone reveals the subsistence practices of a historic coastal community in southwest Madagascar. *J Archaeol Sci* 75:82–88.
- 8. Walter R, Jacomb C, Brooks E (2011) Excavations at Cook's Cove, Tolaga Bay, New Zealand. 2(1).
- 9. Fredericksen C (1996) An Archaeological Survey of a Proposed Subdivision, Tokerau Beach, Northland. 1–13.
- 10. Walter R, Jacomb C, Brooks E (2010) Archaeological Excavations at Omaio, Eastern Bay of Plenty. *Whakatane Dist Hist Soc Mem* 11.
- 11. Leach HM, Hamel J (1981) Archaic and Classic Maori Relationships at Long Beach, Otago: the Artefacts and Activity Areas. *New Zeal J Archaeol* 3:109–141.
- 12. Jacomb C (2009) Excavations at 27 Main Rd , Redcliffs. *Rep to New Zeal Hist Places Trust* (July).
- 13. Trotter MM (1975) Archaeological Investigations at Redcliffs, Canterbury, New Zealand. *Rec Cantebury Museum* 9(3):198–220.
- 14. Higham T, Anderson A, Jacomb C (1999) Dating the first New Zealanders: The chronology of Wairau Bar. *Antiquity* 73(280):420–427.
- 15. Jacomb C (2011) Salvage Excavations at the Tokanui River Mouth. *Archaeol New Zeal* 54(3):181+202.
- 16. Cunliffe EA, Brooks E (2016) Prehistoric Whale Bone Technology in Southern New Zealand. *Int J Osteoarchaeol* 26(3):384–396.
- 17. Walter R, Jacomb C, Jennings C (2011) Final Report on Excavations at the Kawatiri Site (K29/8), West Coast, under NZHPT by. *New Zeal Hist Places Trust*:51.
- 18. Trotter MM (1955) First Excavation Of A Moa Hunter Camp Site At Waimataitai Mouth, Katiki. *J Polyn Soc* 80(3):379–385.
- 19. Mcfadgen BG (1987) Beach ridges, breakers and bones : late Holocene geology and archaeology of the Fyffe site, S49/46, Kaikoura Peninsula, New Zealand. *J R Soc New Zeal* 17(4):381–394.
- 20. Trotter MM (1975) Radiocarbon dates for Wairau Bar and Wakanui, South Island. *New Zeal Archaeol Assoc Newsl* 18(2):90–91.
- 21. Weisler MI (2000) Excavations at Shag Point J43-11 North Otago a summary report. *Archaeol New Zeal* 43(3):220–229.
- 22. Trotter MM (1970) Archaeological investigations in the Aviemore area, South Island. *Rec Cantebury Museum* 8(5):439–453.
- 23. Mason GM, Wilkes O (1963) Tumbledown Bay a Banks Peninsula moa-hunter site S94-30. *New Zeal Archaeol Assoc Newsl* 6(2):98–100.
- 24. Mason GM, Wilkes O (1963) Dashing Rocks, Timaru a preliminary note on excavations – site S111-1. *New Zeal Archaeol Assoc Newsl* 6(2):95–98.
- 25. Harrowfield DL (2013) Archaeological Investigations at Awamoa, North Otago, 1852 – 2013. *Archaeol New Zeal* 56(2):90–99.
- 26. Burrage S (1975) The End of an Era. "Hohou" Closes Down. *Canterbury Museum Archaeol Soc Newsl* 31.
- 27. Jacomb C, Darmody R (2002) Interim Report on Excavations at Watson's Beach (H45/10); An Early Coastal Otago Archaeological Site. *Archaeol New Zeal* 45(1):47– 58.
- 28. Teviotdale D (1932) The Material Culture of The Moa-Hunters in Murihiku. *J Polyn Soc* 41(2):81–120.
- 29. Easton LJ, et al. (2017) Testing species limits of New Zealand's leiopelmatid frogs through morphometric analyses. *Zool J Linn Soc* (January):1–14.
- 30. Yaldwyn JC (1958) Notes on the Environment and Age of the Sub-Fossil Deposits of the Martinborough Caves. *Rec Dom Museum* 3:129–133.
- 31. Worthy TH, Miskelly CM, Ching RA (2002) Taxonomy of North and South Island snipe (Aves: Scolopacidae: Coenocorypha), with analysis of a remarkable collection of snipe bones from Greymouth, New Zealand. *New Zeal J Zool* 29(May 2017):231– 244.
- 32. Worthy TH, Holdaway RN (1996) Quaternary fossil faunas, overlapping taphonomies, and palaeofaunal reconstruction in North Canterbury, South Island, New Zealand. *J R Soc New Zeal* 26(3):275–361.
- 33. Worthy TH, Holdaway RN (1995) Quaternary fossil faunas from caves on Mt Cookson, North Canterbury, South Island, New Zealand. *J R Soc New Zeal* 25(3):333– 370.
- 34. Worthy TH (2004) The Holocene fossil waterfowl fauna of Lake Poukawa, North Island, New Zealand. *Tuhinga* 15:77–120.
- 35. Worthy TH, Holdaway RN (1994) Quaternary fossil faunas from caves in Takaka Valley and on Takaka Hill, northwest Nelson, South Island, New Zealand. *J R Soc New Zeal* 24(3):297–391.
- 36. Worthy TH (2001) A fossil vertebrate fauna accumulated by laughing owls (Sceloglaux albifacies) on the Gouland Downs, Northwest Nelson, South Island. *Notornis* 48(4):225–233.
- 37. Worthy TH, Mildenhall DC (1989) A late Otiran-Holocene paleoenvironment reconstruction based on cave excavations in northwest Nelson, New Zealand. *New Zeal J Geol Geophys* 32(2):243–253.
- 38. Grealy AC, et al. (2015) A critical evaluation of how ancient DNA bulk bone metabarcoding complements traditional morphological analysis of fossil assemblages. *Quat Sci Rev* 128:37–47.
- 39. Riaz T, et al. (2011) ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Res* 39(21):e145.
- 40. Taylor PG (1996) Reproducibility of Ancient DNA Sequences from Extinct Pleistocene Fauna. *Mol Biol Evol* 13(January):283–285.
- 41. Deagle BE, et al. (2007) Studying Seabird Diet through Genetic Analysis of Faeces: A Case Study on Macaroni Penguins (Eudyptes chrysolophus). (9). doi:10.1371/journal.pone.0000831.
- 42. Stat M, et al. (2017) Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. *Sci Rep* 7(1):12240.
- 43. Cooper A (1994) Molecular evolutionary studies of New Zealand birds. *PhD thesis, Sch Biol Sci Victoria Univ Wellington, Wellington, New Zeal*.
- 44. Bergner LM, Dussex N, Jamieson IG, Robertson BC (2016) European Colonization, Not Polynesian Arrival, Impacted Population Size and Genetic Diversity in the Critically Endangered New Zealand Kākāpō. *J Hered*:1–10.