

# Complete Mitochondrial Genome Sequencing Reveals Novel Haplotypes in a Polynesian Population

Miles Benton<sup>1,3</sup>, Donia Macartney-Coxson<sup>2</sup>, David Eccles<sup>1,2</sup>, Lyn Griffiths<sup>3</sup>, Geoff Chambers<sup>1</sup>, Rod Lea<sup>2,3\*</sup>

**1** School of Biological Science, Victoria University of Wellington, New Zealand, **2** Kenepuru Science Centre, Institute of Environmental Science and Research, Sandringham, New Zealand, **3** Genomics Research Centre, Griffith Health Institute, Griffith University, Queensland, Australia

## Abstract

The high risk of metabolic disease traits in Polynesians may be partly explained by elevated prevalence of genetic variants involved in energy metabolism. The genetics of Polynesian populations has been shaped by island hopping migration events which have possibly favoured thrifty genes. The aim of this study was to sequence the mitochondrial genome in a group of Maoris in an effort to characterise genome variation in this Polynesian population for use in future disease association studies. We sequenced the complete mitochondrial genomes of 20 non-admixed Maori subjects using Affymetrix technology. DNA diversity analyses showed the Maori group exhibited reduced mitochondrial genome diversity compared to other worldwide populations, which is consistent with historical bottleneck and founder effects. Global phylogenetic analysis positioned these Maori subjects specifically within mitochondrial haplogroup - B4a1a1. Interestingly, we identified several novel variants that collectively form new and unique Maori motifs – B4a1a1c, B4a1a1a3 and B4a1a1a5. Compared to ancestral populations we observed an increased frequency of non-synonymous coding variants of several mitochondrial genes in the Maori group, which may be a result of positive selection and/or genetic drift effects. In conclusion, this study reports the first complete mitochondrial genome sequence data for a Maori population. Overall, these new data reveal novel mitochondrial genome signatures in this Polynesian population and enhance the phylogenetic picture of maternal ancestry in Oceania. The increased frequency of several mitochondrial coding variants makes them good candidates for future studies aimed at assessment of metabolic disease risk in Polynesian populations.

**Citation:** Benton M, Macartney-Coxson D, Eccles D, Griffiths L, Chambers G, et al. (2012) Complete Mitochondrial Genome Sequencing Reveals Novel Haplotypes in a Polynesian Population. PLoS ONE 7(4): e35026. doi:10.1371/journal.pone.0035026

**Editor:** Manfred Kayser, Erasmus University Medical Center, The Netherlands

**Received:** September 26, 2011; **Accepted:** March 11, 2012; **Published:** April 13, 2012

**Copyright:** © 2012 Benton et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This research was funded by the Wellington Health Research Council Grant #1387. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: r.lea@griffith.edu.au

## Introduction

Scientific evidence from linguistics, archaeology and genetics indicates that the Maori population of New Zealand (NZ) represents the final link in a long chain of island-hopping voyages by Polynesians, which began in Taiwan and stretched through Melanesia and across the Pacific Islands over a period of 5–6000 years (Figure 1). Around 800 years ago one or more small groups of voyagers arrived in NZ from Tahiti, via the Cook Islands. This event marked the last of the great human migrations and the creation of an isolated founder population. The widespread intermarriage between Maoris and Europeans over the past 200 years (8–10 generations) has introduced substantial European genomic ancestry (~40%) into the contemporary Maori gene pool [1].

Maoris, and Polynesians more generally, are disproportionately affected with certain metabolic disease traits eg. obesity and type 2 diabetes mellitus [2,3,4,5]. Given that these traits are partially influenced by genetic factors it is likely that genes involved in energy metabolism play a role in disease risk [6]. Mitochondrial genes could potentially account for some of the high prevalence of metabolic disease traits in Maoris. Coding variants in mitochondrial genes that exhibit unusually high frequencies in Maori may have been driven to high frequency by positive selection due to periods of feast and famine during the migrations (ie. thrifty genes) [7]. Alternatively, these mitochondrial variants may have simply

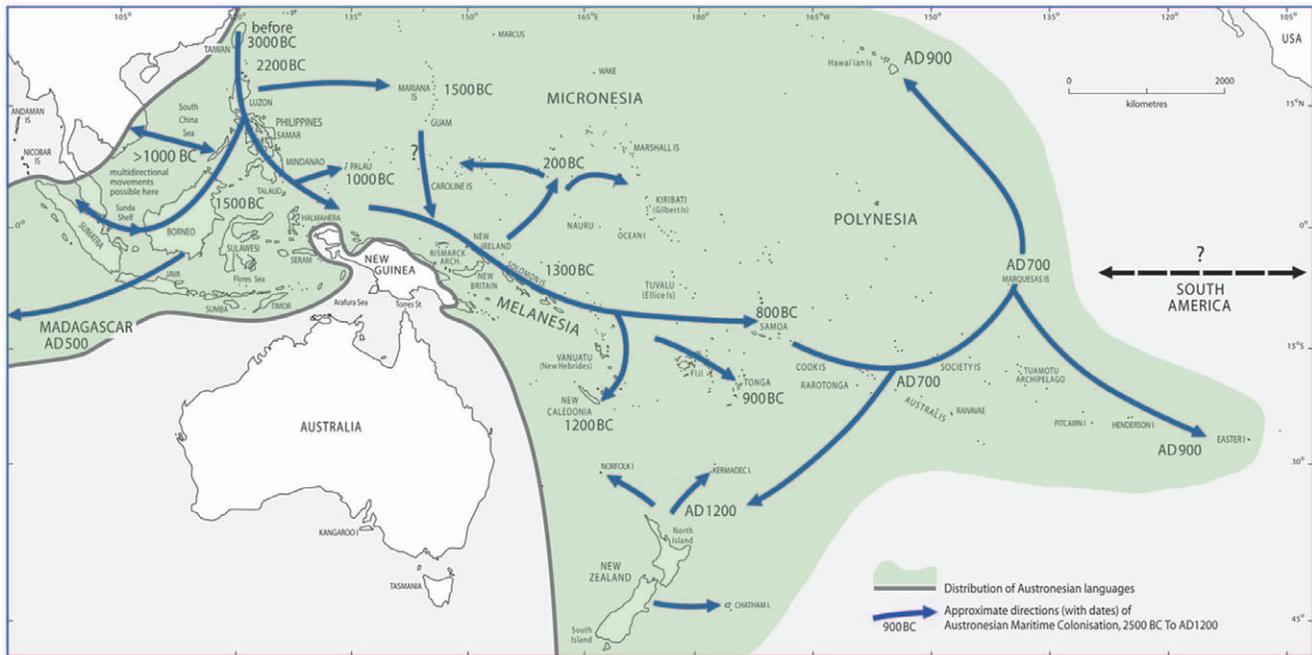
increased in frequency in Maoris via genetic drift as a consequence of repeated founder events and subsequent population bottlenecks.

Complete mitochondrial genome sequence data have been previously investigated to elucidate the evolutionary history among human populations around the world [8]. Studies have also comprehensively investigated mitochondrial genome variation in relation to metabolic syndrome traits [9,10]. However, no complete mitochondrial genome surveys have involved a Maori sample and Polynesians more generally have been under-represented. Given the unusual maternal history of Maoris it is likely that a unique mitochondrial genomic makeup exists in this Polynesian subgroup. In this study we sequenced the entire mitochondrial genome in a group of Maori individuals and performed population genetic analyses to characterise the patterns of genomic variation in this Polynesian population. These new data provide the opportunity to enhance the phylogenetic picture of the mitochondrial genome in the South Pacific region and establish a foundation for future studies of mitochondrial DNA and metabolic disease traits in Polynesian populations.

## Results

### Complete mtDNA sequences and genetic diversity

A summary of mitochondrial (mt) sequence variation for all 20 Maori mtDNA genomes is shown in Table 1. These sequences are



**Figure 1. Map outlining migratory paths of Austronesian speaking populations, including estimated dates.** Adapted from Bellwood *et al.*, (2011) [52].  
doi:10.1371/journal.pone.0035026.g001

the only complete NZ Maori sequences currently available (at time of writing). Previous studies have suggested that there is very limited mtDNA variation in Polynesians in general, and even less in Maori [11,12] and no Maori-specific genetic mtDNA markers have yet been identified.

Sequence variation was identified by comparison against the revised Cambridge Reference Sequence (CRS) [13], which belongs to haplogroup H (commonly found in European peoples). The mt sequence variation identified in the Maori individuals differed from the CRS at 44 variable sites (see Table 1). Of these variant sites, 22 were fixed in all 20 Maori individuals – these are the defining markers of mitochondrial haplogroup B, and its further substructure (haplotypes) such as B4a1a1, to which Maoris belong. There were 12 singleton variants identified and a further 10 variants were shared by two or more individuals and define subclades within the Maori mtDNA phylogeny.

The limited sequence variation was validated by calculation of  $h$  and  $\pi$  diversity statistics in DNAspV5 [14]. Table 2 shows the amount of DNA sequence diversity of 189 complete mtDNA sequences as well as diversity within each specific population. The Maori group were found to exhibit high haplotype diversity ( $h = 0.92$ ), yet diversity was substantially lower than that seen in any of the other three populations (see Table 2). When looking at the nucleotide ( $\pi$ ) diversity it can be seen that Maoris exhibit a value 10-fold lower ( $\pi = 0.00018$ ) compared to that of the other populations. As expected there is no maternal European admixture identified in this group. All mtDNA sequences are clearly Polynesian (Maori) and show the characteristic, and well documented, Polynesian Motif markers [15,16]: 16189, 16217, 16247, and 16261 (see Table 1).

### Phylogenetic Analysis

Phylogenetic analysis of the Maori sequences in the software mtPhy [17] confirmed that all 20 belong to haplogroup B. As expected from previous studies of the Hyper Variable Region

(HVR) in Polynesians [11,12,15,18,19], these Maori sequences all group deep within haplogroup B (for reference see Phylotree [20]). To further investigate the sub-structure of haplogroup B a detailed phylogeny was reconstructed to include 64 complete sequences representing B4a (10 Asian, 14 Taiwanese, 4 Coastal PNG, 16 Pacific Islanders and 20 Maori). Figure 2 illustrates this tree (Asian/Austronesian mt DNA sequences) and shows that all 20 Maori sequences group within B4a1a1, with Pacific Islander and Coastal PNG (Melanesian) mt sequences. These groupings fit well with previous mtDNA work [18,19,21,22,23,24], and complement the hypothetical model of Polynesian origin stemming from Taiwan [25]. Apart from the variants which define haplogroup B, we have identified three novel Polynesian (Maori) haplotypes – until now all documented Polynesian mt haplotypes have been B4a1a1a. Table 3 displays the frequency and specific markers for the haplotypes identified in the 20 complete Maori mt sequences. The most interesting haplotype, B4a1a1a3 (unpublished data), was recently included in an updated build of Phylotree (<http://www.phylotree.org/> [20]). This haplotype was present in 35% ( $n = 7$ ) of the individuals sequenced, and is defined by the variants 1185T and 4769A.

### Novel mt DNA sequence variants in Maori

This study has identified six novel (undocumented) mtDNA variants in the Maori sample: five within protein coding regions and one in the control region (see Table 4). An extended database search of mtDB [26] and Mitomap [27] and for these variants returned no hits, thus these positions are deemed to be novel. Two variants result in amino acid sequence changes; 9255T (Pro→Ser) and 15014C (Phe→Leu). Apart from variant 3909T, the novel variants were only noted in individuals and are not present in the wider population and are thus probably sporadic, rather than ancestral.

**Table 1.** Variation across 20 complete Maori mtDNA sequences.

Position	HVRII					12S rRNA					16S rRNA					ND1 ND2					COI					ATP6				
	73	146	151	263	750	1185	1438	1692	1806	2416	2706	3909	4769!	5465	6261	6719	6782	6905	7028	8860	8865	9123	9145							
CHS*	A	T	C	A	A	C	A	A	T	T	A	C	A	T	G	T	T	A	C	A	G	G	G							
mt3	G	C	.	G	G	T	G	.	.	G	.	G	.	C	.	C	.	.	T	G	.	A	.							
mt6	G	C	.	G	G	.	G	.	.	G	.	G	.	G	.	C	.	.	T	G	.	A	.							
mt8	G	C	.	G	G	T	G	.	.	G	.	G	.	C	.	C	.	.	T	G	.	A	.							
mt9	G	C	.	G	G	.	G	.	.	G	T	G	.	C	.	C	.	.	T	G	.	A	.							
mt10	G	C	T	G	G	.	G	G	.	C	G	.	G	.	C	.	.	.	T	G	.	A	.							
mt11	G	C	.	G	G	.	G	.	.	G	.	G	.	C	.	C	.	.	T	G	A	A	.							
mt12	G	C	.	G	G	.	G	.	.	G	.	G	.	C	.	C	.	.	T	G	A	A	.							
mt13	G	C	.	G	G	T	G	.	.	G	.	.	C	.	C	.	.	.	T	G	.	A	.							
mt14	G	C	.	G	G	T	G	.	.	G	.	.	C	.	C	.	.	.	T	G	.	A	.							
mt16	G	C	.	G	G	.	G	.	.	G	.	G	.	C	.	C	.	.	T	G	.	A	.							
mt18	G	C	.	G	G	.	G	.	.	G	.	G	.	C	A	C	.	G	T	G	.	A	A							
mt19	G	C	.	G	G	.	G	.	.	G	T	G	.	C	A	C	.	.	T	G	.	A	.							
mt21	G	C	.	G	G	.	G	.	.	G	T	G	.	C	.	C	.	.	T	G	.	A	.							
mt23	G	C	.	G	G	.	G	.	.	G	.	G	.	C	.	C	.	.	T	G	.	A	.							
mt24	G	C	.	G	G	.	G	.	.	G	.	G	.	C	.	C	.	.	T	G	.	A	.							
mt25	G	C	.	G	G	T	G	.	.	G	.	G	.	C	.	C	.	.	T	G	.	A	.							
mt26	G	C	.	G	G	.	G	.	.	G	.	G	.	C	.	C	C	.	T	G	.	A	.							
mt28	G	C	.	G	G	T	G	.	.	G	.	.	C	.	C	.	.	.	T	G	.	A	.							
mt29	G	C	.	G	G	T	G	.	C	C	G	.	.	C	.	C	.	.	T	G	.	A	.							
mt30	G	C	T	G	G	.	G	G	.	G	.	G	.	C	.	C	.	.	T	G	.	A	.							
n (variant)	20	20	2	20	20	7	20	2	2	20	3	13	20	2	20	2	1	1	20	20	2	20	1							
Amino-acid change											syn	syn	syn	syn	A120T	syn	syn	syn	syn	T112A	syn	syn	A207T							
Conservation Index											8	28	51	97	100	100	100	100	72	92	100	100								
Protein Position											201	100	332	120	272	293	334	375	112	113	199	207								



**Table 2.** Estimated haplotype ( $h$ ) and nucleotide ( $\pi$ ) diversity.

Population	$N_{ind}$	$N_{haplo}$	$h$	$\pi$
European	67	64	0.999	0.00145
Chinese	52	52	1.000	0.00186
Melanesian	50	48	0.998	0.00172
Maori	20	11	0.916	0.00018
All populations	189	174	0.999	0.00174

$N_{ind}$  (number of sequences),  $N_{haplo}$  (number of haplotypes),  $h$  (haplotype diversity),  $\pi$  (nucleotide diversity).

doi:10.1371/journal.pone.0035026.t002

### Mitochondrial gene variant frequencies in different subpopulations

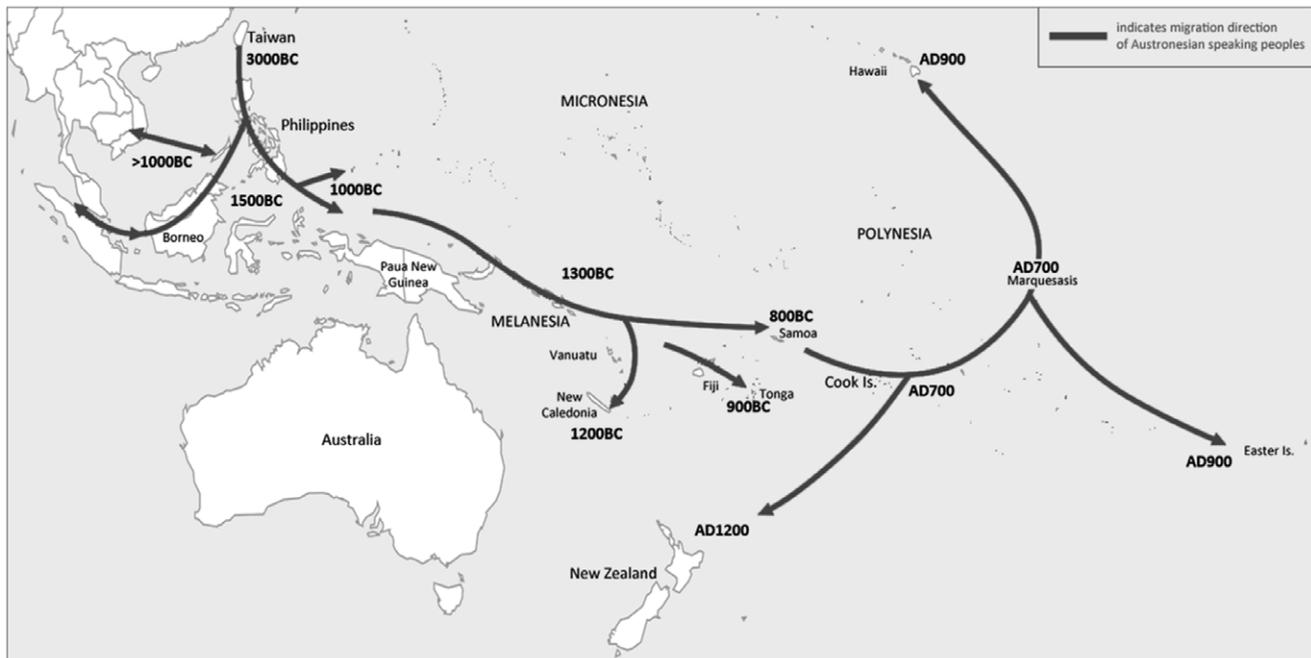
Estimated frequencies of variants within mitochondrial genes were calculated for the NZ Maori as well as for European, Chinese and Melanesian subgroups, selected because they are each putative ancestral contributor populations of NZ Maori. Of the 13 mtDNA genes, the Maori mtDNA sequences contained variable sites in 9 genes, the majority being population specific polymorphisms (haplogroup B defining variants). There were 19 variants spread across these 9 genes, with *COI*, *ATP6*, and *Cyt b* showing the largest number of variants among the ethnic subgroups. The variant frequency differences between these four groups are displayed in Table 5. Of particular interest in terms of metabolic disease risk was the presence of non-synonymous variants in *COI*, *ATP6*, *COIII* and *Cyt b* genes compared to the ancestral subgroups. The most notable population specific polymorphism, variant A15746G in *Cyt b*, was observed in all 20 Maori samples but was absent or rare in all ancestral subgroups (Table 5). The rare variant, 4769A in *ND2*, is also of particular

interest for several reasons; 1) it is a rare polymorphism identified in the CRS [13], yet it is identified at 35% ( $n = 7$ ) in the NZ Maori cohort, and 2) alongside variant 1185T this variant forms a unique Maori haplotype.

### Discussion

This study provides the first complete mitochondrial sequence data for a Polynesian (Maori) population, and as such allows a rare opportunity to enhance the maternal phylogeny in Oceania as well as explore the mitochondrial genome for potential metabolic risk genes in Polynesians. Although sequence alignment of the Maori mt genomes illustrated high concordance with other Polynesian mt sequences, phylogenetic analysis was able to refine the haplotype substructure of Polynesians. Specifically, Maori mt sequences were deemed as belonging to major mt haplogroup B and formed substructures within the B4a1a1 ‘haplotype’. This analysis also confirmed the presence of the 9-bp deletion and characteristic control region variants which have become collectively known as the ‘‘Polynesian motif’’ [15]. Identification of these Polynesian informative sites is consistent with previous mt DNA studies in NZ Maori [11,12].

It has been previously documented that Polynesian and central/eastern Micronesian populations show reduced mtDNA diversity, sharing high frequencies of the single mtDNA haplotype - B4a1a1 [23,28]. We explored the possibility of decreased mt sequence diversity within the Maori population. Both haplotype ( $h$ ) and nucleotide diversity ( $\pi$ ) were shown to be lower in Maori mt genomes compared to putative ancestral populations. Nucleotide diversity exhibited a 10-fold decrease when compared to three ancestral populations. Evidence of such dramatically reduced diversity of the mt genome in Maori is probably due to founder effects during island hopping migrations and is supported by previous studies [11,12]. It is perhaps not surprising that due to



**Figure 2. Phylogenetic reconstruction detailing haplogroup B4a1 in complete mtDNA sequences.** This neighbour-joining tree was created in MEGA4, using the Tamura-Nei substitution model. The sequences used were sourced from Trejaut et al. (2005), and include the 20 complete Maori mtDNA's. Shown in red are the respective tribes of the Taiwanese Aboriginal sequences.

doi:10.1371/journal.pone.0035026.g002

**Table 3.** RHAS Maori mt DNA haplotype markers.

Haplotype	f	1185	4769	14022	16126	16189	16217	16247	16261
B4a1a1c <sup>##</sup>	6	.	G	G	C	C	C	.	T
B4a1a1a <sup>*</sup>	3	.	G	G	.	C	C	G	T
B4a1a1a3 <sup>##</sup>	7	T	.	G	.	C	C	G	T
B4a1a1a5 <sup>##</sup>	4	.	G	G	C	C	C	G	T

\*Previously reported Polynesian mt DNA haplotype.

<sup>##</sup>Novel Polynesian (Maori) haplotypes, numbered following Phylotree nomenclature. B4a1a1a3 was included in a recent update to Phylotree. Variants to the CRS are indicated.

doi:10.1371/journal.pone.0035026.t003

this reduced mt genetic diversity no unique mtDNA haplotypes have so far been discovered within the Maori population. However, our complete mitochondrial genome scan revealed the presence of at least 3 specific sub-haplotypes of haplogroup B in Maori, which are derived from variants 1185T, 4769A, and 16126C. These three variants could have arisen in the seafaring Polynesian ancestors of Maori, or they could have occurred more recently, i.e. since the settlement of NZ. These variants form a unique mt signature within this Maori population, one that is worth exploring further in other NZ Maori populations to determine its generalizability. As there is very little coding region information available for other Polynesian mt sequences, with only 7 complete Polynesian sequences listed in mtDB [26], it is not currently possible to determine whether these ‘signatures’ are unique to Maori. They may in fact be present, but as yet undetected, in the broader Polynesian population. Nevertheless, these new findings provide a more specific mt ancestry informative marker for future genetic studies involving Maori subjects.

Our results also indicate that the protein coding regions within the mitochondrial genome for the populations of Island Southeast Asia, Coastal Melanesia (PNG), Polynesia, and NZ Maori, which are all mt haplogroup B, are heavily conserved and have not changed much over the ~5000 years since the suggested movement from Taiwan. The presence of population specific polymorphisms consistent with those previously identified in haplogroup B was confirmed via comparison across four putative ancestral populations. There was one coding variant (4769A) that is not a haplogroup B defining marker which showed increased frequency in the NZ Maori group, although further work is required to accurately confirm its prevalence in the wider Polynesian community. The lack of coding variation is most likely attributed to genetic drift attributed to the rapid expansion and migration of Austronesian peoples from Taiwan throughout

Oceania in the last ~5000 years [25,29]. Regardless, our findings make these variants good candidates for future genetic association studies of metabolic disease in Maori populations.

Disease association with specific mtDNA variants has been previously noted for several metabolic traits, including; type-2 diabetes (T2D) [30,31,32,33,34], insulin resistance [35,36,37], and BMI/fat mass [10,33,38]. One specific mtDNA variant, 16189C – a fundamental haplogroup B variant, has previously been identified to associate with T2D, insulin resistance and BMI in separate studies [30,33,34,36,38]. Whether the variant itself is the cause of the association, or simply a marker for the larger haplotype/signature or in linkage with other causal variants located in the nuclear genome is yet to be seen.

In conclusion, this study reports the first complete mitochondrial genome sequence data for a Maori population. Overall, these new data reveal unique mitochondrial genome characteristics in this Polynesian population and enhance the phylogenetic picture of maternal ancestry in Oceania. The presence of several newly identified novel variants, as well as the presence of previously identified disease associated variants, offers plausible candidates for future studies aimed at assessment of metabolic disease risk in Polynesian populations.

## Materials and Methods

### Samples

This project is part of the Rakaipaaka Health and Ancestry Study (RHAS) which is aimed at identifying the genetic and environmental determinants of health in the indigenous Maori tribe (iwi) – Ngati Rakaipaaka. Being a DNA-based genetic study involving indigenous Maori participants the RHAS has taken several years to develop in terms of ethical and cultural approval. The RHAS is governed by Te Iwi o Rakaipaaka (TIORI) in

**Table 4.** Novel mtDNA variants observed in 20 Maori individuals.

Gene	Nucleotide Change	Protein Change	No. Individuals	Percentage
16S rRNA	m.1806T>C	NA	1 (20)	5
ND1	m.3909C>T	Syn	3 (20)	15
COXI	m.6782T>C	Syn	1 (20)	5
COXIII	m.9255C>T	p.MT-COXIII:Pro17Ser	1 (20)	5
Cyt b	m.15014T>C	p.MT-Cyt b:Phe90Leu	1 (20)	5
HVRI <sup>*</sup>	m.16295C insA	NA	1 (20)	5

\*HVRI is non-coding. NA, not applicable; 16S rRNA, 16S ribosomal RNA; ND1, NADH dehydrogenase subunit 1; COXI, cytochrome c oxidase I; COXIII, cytochrome c oxidase III; Cyt b, cytochrome b; HVRI, hyper variable region I.

doi:10.1371/journal.pone.0035026.t004

**Table 5.** mtDNA coding variant frequencies in four human populations.

Gene	Variant	Rare allele frequency*			
		European (n = 101)	Chinese (n = 52)	Melanesian (n = 56)	Maori (n = 20)
ND1	C3909T	0	0	0	0.15
ND2	G4769A	0.01	0	0	0.35
	T5465C	0	0	0.16	1
COI	G6261A	0.01	0	0	0.10
	T6719C	0	0	0.16	1
	T6782C	0	0	0	0.05
	A6905G	0	0	0.04	0.05
	C7028T	0.63	1	0.98	1
ATP6	G8865A	0	0	0.02	0.10
	G9123A	0	0	0.016	1
	G9145A	0	0	0	0.05
COIII	C9255T	0	0	0	0.05
	T9722C	0	0	0	0.05
ND3	T10238C	0.03	0	0.16	1
ND4	G11719A	0.56	1	1	1
ND5	A14022G	0	0	0.14	1
Cyt b	C14766T	0.57	1	1	1
	T15014C	0	0	0	0.05
	A15746G	0	0	0.16	1

\*compared to CRS.

doi:10.1371/journal.pone.0035026.t005

Nuhaka and has received full ethical approval from the Multi-regional ethics committee of New Zealand (MEC/05/12/174). All individuals involved signed a consent form acknowledging they understood the genetic nature of this health research and wished to participate. For this mitochondrial project we selected a subsample of 20 adult individuals who were deemed to be non-admixed (ie. have full Maori ancestry). This was determined by the individual self-reporting that they had four Maori grandparents. Genomic DNA was isolated with the use of commercial kits (FlexiGene – QIAGEN). Polynesian mt DNA ancestry was validated using the previously documented “Polynesian motif” [15] (9-bp deletion plus three control region SNPs), which was found to be present in all 20 DNA (unpublished data).

### mtDNA sequencing of 20 Maori individuals

Complete mitochondrial DNA sequence information was obtained for the 20 Maori individuals using the Mitochip Resequencing Array [39]. The Mitochip Resequencing protocol

### References

- Lea RA, Chambers GK (2007) Pharmacogenetics in Admixed Polynesian Populations. Chapter in Pharmacogenetics and Admixed populations Edited by: Guilherme Suarez-Kurtz Eureka Bioscience Database (ISBN: 1-58706-130-9).
- Ministry of Health (2008) A Portrait of Health: Key Results of the 2006/2007 New Zealand Health Survey. Wellington: Ministry of Health.
- Dowse GK, Zimmet PZ, Finch CF, Collins VR (1991) Decline in Incidence of Epidemic Glucose Intolerance in Nauruans: Implications for the “Thrifty Genotype”. *American Journal of Epidemiology* 133: 1093–1104.
- Zimmet P, Dowse G, Finch C, Serjeantson S, King H (1990) The epidemiology and natural history of midm-lessons from the South Pacific. *Diabetes/Metabolism Reviews* 6: 91–124.
- Zimmet P, Taft P, Guinea A, Guthrie W, Thoma K (1977) The high prevalence of diabetes mellitus on a Central Pacific island. *Diabetologia* 13: 111–115.
- O’Rahilly S (2009) Human genetics illuminates the paths to metabolic disease. *Nature* 462: 307–314.
- Neel JV (1962) Diabetes Mellitus - A thrifty genotype rendered detrimental by progress. *American Journal of Human Genetics* 14: 353–362.
- Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, et al. (2003) Natural selection shaped regional mtDNA variation in humans. *Proceedings of the National Academy of Sciences of the United States of America* 100: 171–176.

(Affymetrix, Santa Clara, CA) laid out by Affymetrix was followed (Affymetrix GeneChip CustomSeq Resequencing Array Protocol version 2.1.), and the chips were run on Affymetrix GeneChip equipment (GeneChip Hybridization Oven, GeneChip Fluidics Station, and GeneChip Scanner 3000). The raw data files were analysed using the GeneChip Sequence Analysis Software 4.1 (GSEQ 4.1). Complete mtDNA sequences were exported from GSEQ4.1 and aligned against the revised Cambridge Reference Sequence [13] (CRS) in MEGA4.1 [40]. All sequence data has been submitted to GenBank (awaiting Accession numbers).

### Sequence analysis and mtDNA diversity

Aligned sequences were exported as FASTA files from MEGA4.1, these were then entered into the program mtPhyl [17], where sequence haplotypes and sequence variation statistics were calculated. The mtPhyl software also reported information regarding changes in amino acids and respective position and conservation of these changes. Mitochondrial DNA diversity calculations were performed in DNAspV5 [14] on groups of sequences from four different ethnic populations; European (n = 101) [8,41,42,43,44], Chinese (n = 52) [42,43,45,46], Melanesian (n = 56) [24,42,47,48], and NZ Maori (n = 20). European, Chinese and Melanesian sequences were obtained from the databases mtDB [26] and PhyloTree [20]. Haplotype (h) and nucleotide ( $\pi$ ) diversity statistics were calculated in each ethnic group, as well as in the total sample (all sequences pooled together).

### Phylogeny reconstruction

A consensus neighbour joining tree showing the detail of sub-branching patterns within haplogroup B was reconstructed for a total of 64 complete mitochondrial sequences; all 20 Maori sequences and 44 mitochondrial sequences (Austronesian, Coastal Melanesian and Oceania) obtained from previous studies [25,32,47,49,50]. The phylogeny was constructed in MEGA4.1 [51] using the Tamura-Nei method and a bootstrap of 500 replicates.

### Novel variants and ‘global’ variant frequencies

Observed mtDNA variants in the Maori sequences were searched against known electronic databases (mtDB [26] and mitomap [27]) to identify potential unreported (novel) DNA sequence variants. Identification of possible thrifty genes in Maori involved comparing mitochondrial gene variant frequencies between candidate ancestral populations: European, Chinese (Asian), Melanesian. Sequences were aligned in MEGA4.1 and variant frequencies between population groups were calculated.

### Author Contributions

Conceived and designed the experiments: RL. Performed the experiments: RL MB DM. Analyzed the data: MB DE. Contributed reagents/materials/analysis tools: GC DM. Wrote the paper: MB RL DM LG.

9. Saxena R, de Bakker PIW, Singer K, Mootha V, Burt N, et al. (2006) Comprehensive Association Testing of Common Mitochondrial DNA Variation in Metabolic Disease. *The American Journal of Human Genetics* 79: 54–61.
10. Yang T-L, Guo Y, Shen H, Lei S-F, Liu Y-J, et al. (2011) Genetic Association Study of Common Mitochondrial Variants on Body Fat Mass. *PLoS ONE* 6: e21595.
11. Murray-Mcintosh RP, Scrimshaw BJ, Hatfield PJ, Penny D (1998) Testing migration patterns and estimating founding population size in Polynesia by using human mtDNA sequences. *Proceedings of the National Academy of Sciences of the United States of America* 95: 9047–9052.
12. Whyte ALH, Marshall SJ, Chambers GK (2005) Human evolution in Polynesia. *Human Biology* 77: 157–177.
13. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, et al. (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nature Genetics* 23: 147–147.
14. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.
15. Redd A, Takezake JN, Sherry ST (1995) Evolutionary history of the COII/tRNA lys intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific. *Molecular Biology and Evolution* 12: 604–615.
16. Melton T, Peterson R, Redd AJ, Saha N, Sofro ASM, et al. (1995) Polynesian Genetic Affinities with Southeast Asian Populations as Identified by mtDNA Analysis. *The American Society of Human Genetics* 57: 403–414.
17. Eltsov N, Volodko N (2009) mtPhyl - software tool for human mtDNA analysis and phylogeny reconstruction.
18. Lum JK, Rickards O, Ching C, Cann RL (1994) Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters. *Human Biology* 66: 567(524).
19. Sykes B, Leiboff A, Lowbeer J, Tetzner S, Richards M (1995) The origins of the Polynesians - An interpretation from mitochondrial lineage analysis. *American Journal of Human Genetics* 57: 1463–1475.
20. van Oven M, Kayser M (2009) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Human Mutation* 30: E386–E394.
21. Hertzberg M, Mickelson KN, Serjeantson SW, et al. (1989) An Asian-specific 9-bp deletion of mitochondrial DNA is frequently found in Polynesians. *American Journal of Human Genetics*. pp 504–510.
22. Hagelberg E, Clegg JB (1993) Genetic Polymorphisms in Prehistoric Pacific Islanders Determined by Analysis of Ancient Bone DNA. *Proceedings of the Royal Society of London Series B: Biological Sciences* 252: 163–170.
23. Lum JK, Cann RL (1998) mtDNA and Language Support a Common Origin of Micronesians and Polynesians in Island Southeast Asia. *American Journal of Physical Anthropology* 105: 109–119.
24. Pierson MJ, Martinez-Arias R, Holland BR, Gemmell NJ, Hurles ME, et al. (2006) Deciphering Past Human Population Movements in Oceania: Provably Optimal Trees of 127 mtDNA Genomes. *Society for Molecular Biology and Evolution*.
25. Trejaut JA, Kivisild T, Loo JH, Lee CL, He CL, et al. (2005) Traces of archaic mitochondrial lineages persist in Austronesian-speaking Formosan populations. *PLoS Biology* 3: 1–11.
26. Ingman M, Gyllensten U (2006) mtDB: Human Mitochondrial Genome Database, a resource for population genetics and medical sciences. *Nucleic Acids Research* 34: D749–D751.
27. Ruiz-Pesini E, Lott MT, Procaccio V, Poole J, Brandon MC, Mishmar D, Yi C, Kreuziger J, Baldi P, Wallace CD (2007) An enhanced MITOMAP with a global mtDNA mutational phylogeny. *Nucleic Acids Research* 35 (Database issue): D823–D828.
28. Lum JK, Cann RL (2000) mtDNA Lineage Analyses: Origins and Migrations of Micronesians and Polynesians. *American Journal of Physical Anthropology* 113: 151–168.
29. Tabbada KA, Trejaut J, Loo J-H, Chen Y-M, Lin M, et al. (2010) Philippine Mitochondrial DNA Diversity: A Populated Viaduct between Taiwan and Indonesia? *Molecular Biology and Evolution* 27: 21–31.
30. Bhat A, Koul A, Sharma S, Rai E, Bukhari SA, et al. (2007) The possible role of 10398A and 16189C mtDNA variants in providing susceptibility to T2DM in two North Indian populations: a replicative study. *Human Genetics* 120: 821–826.
31. Chinnery PF, Mowbray C, Patel SK, Elson JL, Sampson M, et al. (2007) Mitochondrial DNA haplogroups and type 2 diabetes: a study of 897 cases and 1010 controls. *Journal of Medical Genetics* 44.
32. Guo LJ, Oshida Y, Fuku N, Takeyasu T, Fujita Y, et al. (2005) Mitochondrial genome polymorphisms associated with type-2 diabetes or obesity. *Mitochondrion* 5: 15–33.
33. Liou CW, Lin TK, Weng HH, Lee CF, Chen TL, et al. (2007) A common mitochondrial DNA variant and increased body mass index as associated factors for development of type 2 diabetes: Additive effects of genetic and environmental factors. *Journal of Clinical Endocrinology and Metabolism* 92: 235–239.
34. Mohlke KL, Jackson AU, Scott LJ, Peck EC, Suh YD, et al. (2005) Mitochondrial polymorphisms and susceptibility to type 2 diabetes-related traits in Finns. *Human Genetics* 118: 245–254.
35. Poulton J, Bednarz AL, Scott-Brown M, Thompson C, Macaulay VA, et al. (2002) The presence of a common mitochondrial DNA variant is associated with fasting insulin levels in Europeans in Auckland. *Diabetic Medicine* 19: 969–971.
36. Poulton J, Brown MS, Cooper A, Marchington DR, Phillips DIW (1998) A common mitochondrial DNA variant is associated with insulin resistance in adult life. *Diabetologia* 41: 54–58.
37. Maechler P, de Andrade PBM (2006) Mitochondrial damages and the regulation of insulin secretion. *Biochemical Society Transactions* 34: 824–827.
38. Kim JH, Park KS, Cho YM, Kang BS, Kim SK, et al. (2002) The prevalence of the mitochondrial DNA 16189 variant in non-diabetic Korean adults and its association with higher fasting glucose and body mass index. *Diabetic Medicine* 19: 681–684.
39. Maitra A, Cohen Y, Gillespie SED, Mambo E, Fukushima N, et al. (2004) The Human MitoChip: a high-throughput sequencing microarray for mitochondrial mutation detection. *Genome Research* 14: 812–819.
40. Kumar S, Nei M, Dudley J, Tamura K (2008) MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in Bioinformatics* 9: 299–306.
41. Achilli A, Rengo C, Magri C, Battaglia V, Olivieri A, et al. (2004) The Molecular Dissection of mtDNA Haplogroup H Confirms That the Franco-Cantabrian Glacial Refuge Was a Major Source for the European Gene Pool. *The American Journal of Human Genetics* 75: 910–918.
42. Ingman M, Kaessmann H, Paabo S, Gyllensten U (2000) Mitochondrial genome variation and the origin of modern humans. *Nature* 408: 708–713.
43. Kivisild T, Shen P, Wall DP, Do B, Sung R, et al. (2006) The Role of Selection in the Evolution of Human Mitochondrial Genomes. *Genetics* 172: 373–387.
44. Moilanen JS, Finnila S, Majamaa K (2003) Lineage-specific selection in human mtDNA: Lack of polymorphisms in a segment of MTND5 gene in haplogroup J. *Molecular Biology and Evolution* 20: 2132–2142.
45. Kong Q-P, Yao Y-G, Sun C, Bandelt H-J, Zhu C-L, et al. (2003) Phylogeny of East Asian Mitochondrial DNA Lineages Inferred from Complete Sequences. *American Journal of Human Genetics* 73: 671–676.
46. Macaulay V, Hill C, Achilli A, Rengo C, Clarke D (2005) Single, Rapid Coastal Settlement of Asia Revealed by Analysis of Complete Mitochondrial Genomes. *Science* 308: 1034–1036.
47. Ingman M, Gyllensten U (2003) Mitochondrial Genome Variation and Evolutionary History of Australian and New Guinean Aborigines. *Genome Research* 13: 1600–1606.
48. Merriwether DA, Hodgson JA, Friedlaender FR, Allaby R, Cerchio S, et al. (2005) Ancient mitochondrial M haplogroups identified in the Southwest Pacific. *Proceedings of the National Academy of Sciences of the United States of America* 102: 13034–13039.
49. Soares P, Rito T, Trejaut J, Mormina M, Hill C, et al. (2011) Ancient Voyaging and Polynesian Origins. *The American Journal of Human Genetics* 88: 239–247.
50. Ingman M, Gyllensten U (2001) Analysis of the Complete Human mtDNA Genome: Methodology and Inferences for Human Evolution. *The American Genetic Association* 92: 454–461.
51. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* 10.1093/molbev/msm092.
52. Bellwood P, Chambers G, Ross M, Hung H-c (2011) Are 'Cultures' Inherited? Multidisciplinary Perspectives on the Origins and Migrations of Austronesian-Speaking Peoples Prior to 1000 BC. In: Roberts BW, Vander Linden M, eds. *Investigating Archaeological Cultures*: Springer New York. pp 321–354.