Discordant association of the *CREBRF rs373863828* minor allele with increased body mass index and protection from type 2 diabetes in Māori and Pacific (Polynesian) people living in Aotearoa New Zealand

Mohanraj Krishnan[1,2^], Tanya J Major[3^], Ruth K Topless[3], Ofa Dewes[4], Lennex Yu[5], John MD Thompson[1,6], Lesley McCowan[1], Janak de Zoysa[7], Lisa K Stamp[8], Nicola Dalbeth[2], Jennie Harré Hindmarsh[9], Nuku Rapana[10], Ranjan Deka[11], Winston W H Eng[12], Daniel E Weeks[12,13], Ryan L Minster[13], Stephen T McGarvey[14], Satupa'itea Viali[15], Phillip Wilcox[16], David Grattan[4,17], Peter R Shepherd[2,4], Andrew N Shelling[1], Rinki Murphy[2,4#], Tony R Merriman[3,4#*]

1, Department of Obstetrics and Gynaecology, University of Auckland, Auckland, New Zealand; 2, Department of Medicine, Faculty of Medical and Health Sciences, University of Auckland University of Auckland, Auckland, New Zealand; 3, Department of Biochemistry, University of Otago, Dunedin New Zealand; 4, Maurice Wilkins Centre for Molecular Biodiscovery, New Zealand; 5, Department of Anatomy and Medical Imaging, University of Auckland, Auckland, New Zealand; 6, Department of Paediatrics, Child and Youth Health, University of Auckland, Auckland, New Zealand; 7, Renal Services, Waitemata District Health Board, Auckland, New Zealand; 8, Department of Medicine, University of Otago Christchurch, Christchurch, New Zealand; 9, Ngāti Porou Hauora Charitable Trust, Te Puia Springs, Tairāwhiti East Coast, New Zealand; 10, Pukapuka Community of New Zealand Inc, 23 Canning Crescent, Mangere, Auckland, New Zealand; 11, Department of Environmental Health, University of Cincinnati College of Medicine, Cincinnati, Ohio; 12, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, USA; 13, Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, USA; 14, International Health Institute, Department of Epidemiology, Brown University School of Public Health, and Department of Anthropology, Brown University, Providence, USA; 15, Samoa National Health Service, Apia, Samoa; 16, Department of Mathematics and Statistics, University of Otago, Dunedin, New Zealand; 17, Department of Anatomy, University of Otago, Dunedin, New Zealand.

^These authors contributed equally

#These authors contributed equally

*Corresponding author (tony.merriman@otago.ac.nz)

Abstract

The minor allele of *CREBRF rs373863828* associates with increased body mass index (BMI) and reduced risk of type 2 diabetes (T2D) in the Samoan population of Samoa and American Samoa. Our aim was to test *rs373863828* for association with BMI and odds of T2D, gout and chronic kidney disease (CKD) in Māori and Pacific (Polynesian) people living in Aotearoa New Zealand in 2,286 adults. Association analyses were performed by linear and logistic regression with BMI, log-transformed BMI, waist circumference, T2D, gout, CKD, and serum and urate lipid measures. Analyses were adjusted for age, sex, the first four genome-wide principal components, and (when appropriate) BMI and T2D. For the minor allele of *rs373863828* the effect size for log-transformed BMI was 0.038 (P=4.8x10^{-6}) and for T2D was OR=0.59 (P=1.9x10^{-6}). The minor allele conferred a dominant effect for both BMI and T2D. Although there was weak evidence for association with increased urate (P=0.026), there was no evidence for association with gout (P=0.98) or CKD (P=0.59). Our results replicated, with very similar effect sizes, association of the minor allele of *rs373863828* with higher BMI but lower odds of T2D among New Zealand Polynesian adults as in Samoan adults living in Samoa and American Samoa.

Introduction

A missense variant (*rs373863828*, p.(Arg457Gln)) in *CREBRF* (encoding CREB3 regulatory factor) has been associated with body mass index (BMI) in the Samoan population residing in Samoa and American Samoa (1). The minor allele (c.1370A p.(457Gln)) associated with a 1.36kg/m^2 higher BMI (or approximately 4kg in body weight, assuming a height of 1.7m) per copy, with an ~1.3 fold greater risk of obesity (1; 2). In a small sample of individuals from the Kingdom of Tonga (n=171) the minor allele associated with a greater BMI of 3.1kg/m^2 (3). This effect size is considerably larger than any other currently known common obesity gene variant observed, such as the *FTO* (e.g *rs9939609*) minor allele which increases BMI by 0.39kg/m^2 (~1.3kg in body weight) in people of European ancestry (4). The same *FTO* alleles associated with BMI in Europeans are not associated with BMI in Oceanic populations (5; 6). The high minor allele frequency (MAF) of the *rs373863828* missense variant among Samoans in Samoa and American Samoa (MAF = 0.26) and in the Kingdom of Tonga (MAF = 0.15) (1; 3), compared to an exceedingly rare frequency in other populations in the Genome Aggregation Database (gnomad.broadinstitute.org; MAF=$5.3x10^{-5}$ in East Asian, $3.3x10^{-5}$ in South Asian, $4.0x10^{-5}$ in European, absent in ~12K African individuals) (7), supports the hypothesis that *rs373863828* is an important risk factor for obesity unique to the Samoan and Tongan populations and possibly other Polynesian populations.

Unlike *FTO* and many other obesity risk variants, the BMI-increasing allele of *rs373863828* associates with lower odds of type 2 diabetes (T2D) (OR = 0.59-0.74 after adjustment for BMI) (1). This is contrary to the established association between T2D and increased BMI. Recently, several genetic variants have been associated with higher BMI and lower risk of T2D, hypertension and coronary artery disease in European populations (8-10). These have been termed 'favourable adiposity' alleles due to their association with a higher subcutaneous to visceral adipose tissue ratio, lower waist-hip ratio in women, and potential to protect from metabolic disease by increasing the capacity of subcutaneous tissue to store excess calories as lipids (11). However, aside from fasting

glucose, the CREBRF variant does not associate with other features of insulin resistance such as hypertension, lipids or homeostatic model assessment (HOMA) of insulin resistance (1).

In Aotearoa New Zealand, obesity and T2D are both highly prevalent in Māori and Pacific people (12). These conditions are also strongly associated with other prevalent metabolic-based conditions in Māori and Pacific people, specifically gout and chronic kidney disease (13-15) as complications of T2D and hypertension. Moreover increased BMI is causal of increased urate and risk of gout [reviewed in (16)]. Polynesian populations include those from West Polynesia (originating from Samoa, Tonga, Niue and Tokelau) and East Polynesia (Aotearoa New Zealand Māori and Cook Island Māori). Given the presence of different pathogenic allele frequencies and different linkage disequilibrium structure between East and West Polynesian populations (17; 18), understanding the genetic variation in the *CREBRF* gene in other Polynesian population groups besides those of Samoan and Tongan ancestry may provide novel insights into its association with increasing BMI, yet apparent reduction in risk of T2D.

In the present study, we tested the association of the *CREBRF rs373863828* variant with BMI, waist circumference, T2D, gout, chronic kidney disease (CKD) and relevant biochemical measures in people of Polynesian ancestry living in Aotearoa New Zealand. We reasoned that if the association of the minor allele with higher BMI but lower risk of T2D was due to imparting favourable adiposity or muscle expansion in non-central locations, then the minor allele would not associate with waist circumference once adjusted for BMI. In addition, based on association of *FTO* genotype with variance in BMI in Europeans (19) we tested the association of *rs373863828* with variance in log-transformed BMI to detect possible underlying genetic and/or environmental influences in phenotypic variability.

Methods

Study population

Individuals, primarily from the Auckland, Waikato, and Christchurch regions of Aotearoa New Zealand, not known to be first-degree relatives, and aged 16 years and older, were recruited as participants of the "Genetics of Gout, Diabetes and Kidney Disease in Aotearoa New Zealand" case-control studies (20). A separate Māori sample set from the *rohe* (area) of the Ngāti Porou *iwi* (tribe) of the Tairāwhiti region on the East Coast of the North Island of New Zealand was also included in the Aotearoa New Zealand Māori analysis. This sample-set was recruited in collaboration with Ngati Porou Hauora (Health Service) Charitable Trust. A unique Pukapuka Island sample set was recruited in collaboration with the Pukapuka Community of New Zealand Inc. in Mangere, South Auckland.

Information obtained at recruitment for each study included age (years), sex, height (cm), weight (kg) and waist circumference (cm) measured by trained assessors. BMI was calculated by dividing weight by the square of height in metres. Participants were also asked the ancestry of each of their grandparents. Gout was ascertained by the 1977 American Rheumatism Association preliminary classification criteria (21). T2D was ascertained by physician-diagnosis and/or patient reports and/or use of glucose lowering therapy. Blood samples were collected and biochemical measurements were performed at the Southern Community Laboratories (www.sclabs.co.nz). Estimated glomerular filtration rates (eGFR) were derived from participants' serum creatinine, age and sex using the Chronic Kidney Disease Epidemiology Collaboration equation (22). Stage 4 and 5 CKD was defined by eGFR <30mL/min. Obesity was defined as BMI >32 kg/m^2 (23). Ethical approval was given by the NZ Multi-Region Ethics Committee (MEC/05/10/130; MEC/10/09/092; MEC/11/04/036) and the Northern Y Region Health Research Ethics Committee (NPHCT study; NTY07/07/074). All participants provided written informed consent for the collection of samples and subsequent analysis.

Participants who self-reported any Polynesian ancestry amongst their grandparents were separated into sample-sets based on self-reported Pacific nation of ancestry. Those participants who also reported non-Polynesian ancestry (predominantly European or Chinese) were grouped according to

their Polynesian ancestry. This resulted in seven sample-sets; Aotearoa New Zealand Māori (*n*=1,296, including 270 people from the Ngāti Porou Hauora Charitable Trust study), Cook Island Māori (*n*=205), Samoan (*n*=387), Tongan (*n*=181), Niuean (*n*=47), Pukapukan (*n*=75) and an 'Other' Polynesian group (*n*=271), which included individuals of Tahitian (*n*=3), Tokelauan (*n*=6) and Tuvaluan (*n*=5) ancestry, along with individuals who self-reported grandparental ancestry from more than one Pacific nation (*n*=257). Pukapuka is part of the Cook Islands situated 1,140 km north-west of Rarotonga, the main island of the Cook Islands (East Polynesia), and ~720 km north-east of Samoa (Apia), geographically locating it within West Polynesia. These analysis groups were further refined based on clustering of genome-wide principal component vectors one to four (details of calculation below), resulting in the exclusion of 182 people who clustered outside of their self-reported ancestry group. The final groups used in all analyses were; Aotearoa New Zealand Māori (*n* = 1,154), Cook Island Māori (*n* = 197), Samoan (*n* = 378), Tongan (*n* = 175), Niuean (*n* = 47), Pukapukan (*n* = 70) and the 'Other' Polynesian group (*n* = 265). Baseline characteristics for the final groupings are presented in Table 1.

Whole genome Illumina Infinium CoreExome

The Illumina Infinium CoreExome v24 bead chip platform was used to genotype participants for ~500,000 variants across the whole-genome. Genotyping was performed at the University of Queensland (Centre for Clinical Genomics) for the Genetics of Gout, Diabetes and Kidney Disease in Aotearoa cohorts and at AgResearch (Invermay Agricultural Centre) for the Ngāti Porou Hauora Charitable Trust cohort. Bead chip genotyping batches were auto-clustered using GenomeStudio v2011.1 software (Illumina, San Diego). The Illumina GenomeStudio best practice guidelines and quality control protocols of Guo *et al.* (2014) were applied (24; 25). The genotyping batches were then merged and relevant quality control steps repeated in the full dataset.

Determination of principal components

Whole-genome principal component analysis vectors were calculated using a subset of 2,858 ancestry informative markers (as identified by Illumina) extracted from the CoreExome whole-genome

genotypes. The SmartPCA (EIGENSOFT v6.0.1) (26)) program was used, with an output of 10 eigenvectors, no outlier removal, and no population size limit. Individuals of non-Polynesian ancestry were included, and the first four vectors plotted against each other to view the clustering of ancestral groupings (Asian, European, Eastern Polynesian, and Western Polynesian). The first four vectors, that explained 97.1% of the proportion of variance with the first ten vectors, were chosen for inclusion as covariates in the linear regression to account for population stratification and cryptic relatedness. Clustering by principal component vectors is presented in Figure S1.

CREBRF rs373863828 genotyping

It was required to directly genotype *rs373863828* because this variant was not present on the CoreExome genome-wide genotyping platform and we were unable to impute the region owing to the unavailability of Māori and Pacific reference haplotypes. A custom designed TaqMan™ probe-set (Applied Biosystems, Foster City, CA) was created for *rs373863828* using a Python script (snp_design; DOI:10.5281/zenodo.56250) to annotate the human genome build 37 reference sequence (ftp://ftp.ensembl.org/pub/grch37) with *rs373863828* and any surrounding SNPs (obtained from the NCBI dbSNP build 147 common SNP list; ftp://ftp.ncbi.nlm.nih.gov/snp). Forward Primer: CAAGAGAGGATGCTGAGACCAT; Reverse Primer: ACCATGATGTAAGCCATTTTTCTGATACA; Probe 1 (VIC): TGAGTGGAACC**G**AGATAC Probe 2 (FAM): AGTGGAACC**A**AGATAC. Genotyping was performed using the LightCycler™ 480 Real-Time Polymerase Chain Reaction System (Roche Applied Science, Indianapolis, IN) in 384 well plates. Genotyping of 25% of the sample set as technical replicates demonstrated 100% concordance.

Association testing

Analyses were performed using the R statistical software within RStudio v0.99.902 (https://www.rstudio.com). A multivariable linear regression model was used to test for association between the *rs373863828* minor allele (c.1370A p.(457Gln)) and continuous variables (log-transformed BMI, untransformed BMI, waist circumference, lipid measures and serum urate), with the ß-coefficient representing the estimated effect of each copy of the *rs373863828* minor allele. For

binary outcomes (obesity, T2D, gout and CKD), a multivariable binomial logistic regression model was used in a similar manner, with the allelic odds ratio (OR) representing the estimated effect of each copy of the *rs373863828* minor allele. Each Polynesian population sample-set was analysed separately, and the effects combined using an inverse-variance-weighted fixed effect meta-analysis. Heterogeneity between sample-sets was also assessed during the meta-analysis using Cochran's heterogeneity (Q) statistic. For the BMI, waist circumference and T2D association analyses $P < 0.05$ was set as the significance value, given the prior probability of detecting association (1). For the other phenotypes (lipids, urate, gout, CKD) $P<0.05$ was also set as the significance value, however with cautious interpretation.

Power

Based on estimates from the Minster *et al*. (1) study and $\alpha = 0.05$ the power to detect an effect size of 1.36 kg/m^2 per minor allele was > 80% in the combined Māori and Pacific Island sample-set for a minor allele frequency of 0.15 or greater (Figure S2). The power to detect a moderate protective effect for T2D (OR=0.59) of the minor allele was >80% for a minor allele frequency of 0.10 or greater (Figure S2).

Testing for association of rs373863828 with variance in log-transformed BMI

Association of a genetic variant with variance in phenotype can detect a locus interacting in a non-additive way without prior knowledge of the interacting factor (environmental, intrinsic, genetic). Testing for association of *rs373863828* with variance in log-transformed BMI was performed as previously described (27). The variable used as a measure of variance was produced from residuals obtained from cohort-specific analyses regressing age, age^2 and age-by-sex. An independent ranked inverse normal transformation of absolute residuals generated z-scores, with squared z-scores (z^2) being the variance variable. To account for the influence of *rs373863828* mean-effect on the variance the mean log-transformed BMI (per genotype) was subtracted from the log-transformed BMI of each participant and the z-scores re-calculated. Linear models associating *CREBRF* genotype with both the unadjusted and mean-effect adjusted variance z-scores were performed (Equation 1). This analysis

was also done on the *rs373863828* genotype data of Minster *et al* (1), with age, age^2, age-by-sex and polity as the adjusting variables in the z-score calculation steps. The Aotearoa New Zealand sample sets along with the two sample-sets of Minster *et al*. (1) were combined by an inverse-variance weighted fixed-effect meta-analysis.

Equation 1: A) unadjusted variance analysis. B) mean-effect adjusted variance analysis.

A.

$$\log(\text{BMI}) \sim \text{adjusting variables} + \text{residuals}$$

$$\text{unadjusted } z^2 = (\text{ranked inverse-normal transformed residuals})^2$$

$$z^2 \sim rs373863828$$

B.

$$\log(\text{BMI} - \text{mean}(\text{BMI per } rs373863828 \text{ genotype})) \sim \text{adjusting variables} + \text{residuals}$$

$$\text{adjusted } z^2 = (\text{ranked inverse-normal transformed residuals})^2$$

$$z^2 \sim rs373863828$$

Results

Prevalence of rs373863828

The allele and genotype frequencies of *rs373863828* in each Polynesian sample-set are presented in Table 1. The relative frequencies of the minor (c.1370A) allele differed among the Polynesian groups, with the Samoan (MAF=0.236) and Pukapukan (MAF=0.243) groups exhibiting the highest frequency the Niuean (MAF=0.096) group exhibiting the lowest frequency. Significant differences in allele frequency were observed between the Samoan and New Zealand Māori, Tongan and Niuean groups, and the Niuean and the Cook Island Māori and Pukapukan groups (Figure 1).

Association analysis with adiposity measures

Association of the minor allele of *rs373863828* (c.1370A p.(457Gln)) with log-transformed BMI was detected in the New Zealand Māori (ß=0.038, $p=3.7x10^{-3}$), Pukapukan (ß=0.090, $p=0.015$) and 'Other' Polynesian (ß=0.069, $p=4.3x10^{-3}$) sample sets, but not for the Samoan (ß=0.030, $p=0.11$), Tongan (ß=0.016, $p=0.55$), Cook Island Māori (ß=0.042, $p=0.17$) and Niuean (ß=-0.070, $p=0.20$) sample sets (Table S1). A fixed-effect meta-analysis of the Polynesian samples showed significant association of *rs373863828* with log-transformed BMI (ß=0.038, $p=4.8x10^{-6}$) (Tables 2; Figure 2). Association analysis with untransformed BMI revealed similar results (Table 2, S2; Figure 2). One copy of the minor allele was sufficient to confer the effect (Table 2; ß=1.80 for the heterozygote group and ß=1.49 for the minor allele homozygote group compared to the major allele homozygotes in untransformed BMI analysis). In the combined group there was association with higher odds of obesity (Table 2; OR=1.33, $P=8x10^{-4}$ for >32 kg/m^2 and OR=1.54, $P=1x10^{-5}$ for >40 kg/m^2). There was no evidence for sex-specific effects (Table 2).

Significant associations were found between *rs373863828* and increased waist circumference in the New Zealand Māori (ß=2.36 cm per minor allele, $P=0.029$), Pukapukan (ß=8.75, $P=4.2x10^{-3}$) and other (ß=6.98, $P=5.0x10^{-4}$) Polynesian sample sets (Tables 2, S3). A fixed effect meta-analysis showed evidence of association between *rs373863828* with increased waist circumference in the full

Polynesian sample-set (ß=2.98 cm per minor allele, P=1.3x10^{-5}) (Figure 2**)**. Adjustment of the waist circumference analysis by BMI abrogated the association with waist circumference (Table 2; ß lowered from 2.98 to 0.66, P=0.092).

Association analysis with T2D

There was association of the minor allele of *rs373863828* with reduced odds of T2D in the New Zealand Māori (OR=0.57, P=3.0x10^{-4}), Samoan (OR=0.45, P=5.6x10^{-3}) and other Polynesian (OR=0.31, P=5.0x10^{-4}) sample sets (Tables 2, S4; Figure 4). A fixed-effect meta-analysis of the various sample sets revealed significant association with reduced odds of T2D (OR=0.65, P=3,4x10^{-5}) that was strengthened after adjustment by BMI (OR=0.59, P=1.9x10^{-6}) (Table 2). Similar to the observation for BMI, one copy of the minor allele was sufficient to confer the effect (Table 2: OR=0.55 for the heterozygote group compared to the major allele homozygote group).

Association analysis with gout, CKD and other metabolic parameters

There was no evidence for association with gout (OR=1.10, P=0.34) (Table 2, Figure S3). However there was weak evidence for association with CKD (OR=0.72, P=0.030) although not after adjustment by T2D and BMI (OR=0.91, P=0.59) (Table 2, Figure S3). Excluding individuals with T2D from the CKD analysis did not provide evidence for association with CKD in a single analysis of all samples pooled and adjusted for the first 4 PCA vectors, age, sex and BMI (OR=1.07 [0.65~1.69], P=0.79).

We also tested for association with serum urate, cholesterol, triglycerides, HDL-cholesterol and LDL-cholesterol (Table 2; Fig S4). After adjusting for BMI there was weak evidence for association with serum urate levels, with the minor allele of *rs373863828* associated with an increase in serum urate (ß=0.012 mmol/L, P=0.026), however this was not significant after a conservative Bonferroni correction for the number of additional phenotypes tested (excluding BMI, T2D and waist circumference there was seven, $P_{corrected}$ = 0.18).

Association analysis of rs373863828 with variance in phenotype

The *rs373863828* variant demonstrated no effect on log-transformed BMI variance at the *CREBRF* locus for any of the New Zealand sample-sets nor the two Samoan cohorts in the study of Minster *et al.* (1). Fixed effect meta-analysis of the New Zealand (*n*=2282), 1990s Samoan (*n*=1020) and discovery Samoan (*n*=1876) cohorts showed no evidence of association of *rs373863828* with variance in log-transformed BMI in either the unadjusted (ß=-0.053, *p*=0.15) or adjusted models (ß=-0.047, *p*=0.15) (Figure S5).

Discussion

The prevalence of the minor allele of the *CREBRF* p.Arg457Gln missense variant (*rs373863828*, c.1370G>A) is highest in populations of Polynesian ancestry (Figure 1; 0.10-0.25), there is a prevalence of <0.06 in some Melanesian and Micronesian populations (3) and it is not detectable in other major population groups. We have now confirmed the presence of this *CREBRF* variant in both East (Aotearoa New Zealand and Cook Island Māori) and non-Samoan West Polynesians. Moreover we replicated, in a non-Pacific Island environment, the association of the minor allele at *rs373863828* with higher BMI (1.38 kg/m^2 per minor allele), higher waist circumference (2.98 cm) and with reduced odds of T2D (OR = 0.59). These results are very similar to that of Samoans living in Samoa and American Samoa, with each copy of the minor allele associated with a 1.36 kg/m^2 increase in BMI and an OR for T2D of 0.59 (1) suggesting that the main effect of the *rs373863828* minor allele is relatively impervious to environment. Consistent also with the Samoan data (1), adjustment by BMI strengthened the association with T2D in the Aotearoa New Zealand sample set (Table 2; OR=0.65 to 0.59; Minster *et al.* discovery OR=0.64 to 0.59; Minster *et al.* replication OR=0.83 to 0.74). However, adjustment by BMI removed the association with waist circumference (Table 2), indicating that the waist circumference association was driven by overall body mass distribution rather than central adiposity.

Precisely how these genetic epidemiological findings relate to the actual CREBRF-mediated molecular pathogenesis of obesity and T2D is unclear in the absence of detailed knowledge of the molecular pathways involving CREBRF and in the absence of genetic association data with detailed body composition measures as outcome. Most population genetic variants associated with generalised obesity also associate with insulin resistance, hypertension, dyslipidemia and T2D compatible with the degree of adiposity. However, there are some genetic variants which are associated with higher BMI and percentage body fat, but lower T2D along with lower insulin resistance, hypertension, circulating triglycerides, and LDL-cholesterol. These variants are also known as ‘favourable adiposity’ variants (11) due to higher subcutaneous-to-visceral adipose tissue that suggests preferential fat storage away from visceral organs. The association of the minor allele of *rs373863828*

with a higher BMI but reduced odds of T2D is not entirely compatible with 'favourable adiposity', due to the lack of association with lower insulin resistance, hypertension or lipids (Table 2 and (1)). Cellular bioenergetics models show that the *rs373863828* minor allele promotes lipid and triglyceride storage at a reduced energy cost in the adipocyte suggesting that the presence and metabolic activity of *CREBRF* in fat is important (1), but detailed clinical studies are required to clarify whether visceral and subcutaneous body fat storage depots are altered among carriers of this variant. In a systematic analysis investigating inter-tissue gene co-expression of metabolic healthy obese and metabolic unhealthy obese individuals, inflammatory biomarkers IL-6 and IL1B were found to exhibit co-expression differences related to the metabolic state (28), hence cytokines should be tested for a relationship with *rs373863828*.

There was weak evidence for association of *rs373863828* with CKD, with the minor allele conferring protection. Based on the non-significant association after adjustment by T2D this effect is likely to be mediated by the effect of the *rs373863828* minor allele on T2D. The minor allele of *rs373863828* was weakly associated with an increase in serum urate levels. This is unlikely to be via the established causal effect of increased BMI on increased serum urate (reviewed in ref (16)) because the effect was maintained after adjusting for BMI. However there was no significant association with gout.

It is notable that, from what is understood about the physiological role of CREBRF, there is no obvious role in regulation of appetite, which is seen at *FTO-IRX3* and other loci regulating BMI in Europeans. However, *CREBRF* is widely expressed (www.gtex.org), including throughout the brain and the *CREBRF* gene is known to regulate the CREB3/Luman protein, which is localised to the endoplasmic reticulum and plays an important role in axonal regeneration (29). Interestingly, the CREB3/Luman protein was identified through its association with herpes simplex virus-related host cell factor 1, which has led to the hypothesis that Luman may play a role in the viral emergence from latency (30). It will be important to explore the relationship between *CREBRF-rs373863828* expression in the hypothalamic nuclei and the CREB3/Luman role in the intra–axonal translation and

retrograde trafficking to promote neuronal survival in response to viral stimuli and how this may relate to BMI and T2D.

This study has confirmed in people of Māori and Pacific (Polynesian) ancestry living in New Zealand that the presence of each additional minor *CREBRF rs373863828* allele is associated with 1.36 kg/m^2 higher BMI (equivalent to 4kg/allele for an individual 1.7m in height), yet a decreased odds for T2D (OR=0.59). The association with 3 cm larger waist circumference is a manifestation of its association with overall BMI, since once adjusted for BMI, there is no association with central obesity. While the prevalence of both obesity and T2D is increased amongst New Zealand Māori and those of Pacific (Polynesian) ancestry, compared to New Zealand Europeans, our study confirms that population-specific genetic variation underpins some of the inter-individual heterogeneity observed in the discordant manifestation of obesity and T2D. This study supports the need to conduct comprehensive gene-phenotype studies in populations currently under-represented in genomic studies, in which different genetically segregating pathways linking obesity and T2D clearly exist. Such studies are not only important for these populations *per se*, but are also important in illuminating the molecular biology of the pathogenesis of metabolic disease in the wider human population and have the potential to lead to novel clinical interventions.

Acknowledgements

The authors sincerely thank participants for generously donating their time and information to this study. The authors would like to thank Jill Drake, Jordyn de Kwant, Roddi Laurence, Christopher Franklin, Meaghan House, Nancy Aupouri, Ria Akuhata, Carol Ford and Gabrielle Sexton for recruitment, and the Health Research Council of New Zealand (Grant #'s 08/075, 10/548, 11/1075, 14/527) and the Maurice Wilkins Centre for funding the New Zealand component, and the National Institute of Health for funding the Samoa and American Samoa component (Grant # R01-HL093093). The authors declare no conflict of interest relevant to the material presented herein. MK generated and analyzed data, drafted the manuscript and approved the final version; TJM, RKT, LY, JMDT, WWHE, DEW, RLM, analyzed data, critically revised the manuscript and approved the final version; OD, JZ, LKS, ND, JHH, NR, STM, SV contributed data or to the acquisition of data, critically revised the manuscript and approved the final version; LM, PW, DG, ANS contributed to the interpretation of data, critically revised the manuscript and approved the final version; PRS, RM, TRM contributed to study design, data collection, data interpretation and analysis, critically revised the manuscript and approved the final version. TRM is the guarantor of the manuscript.

References

1. Minster RL, Hawley NL, Su CT, Sun G, Kershaw EE, Cheng H, Buhule OD, Lin J, Reupena MS, Viali S, Tuitele J, Naseri T, Urban Z, Deka R, Weeks DE, McGarvey ST. A thrifty variant in CREBRF strongly influences body mass index in Samoans. Nat Genet 2016;48:1049-1054

2. Loos RJ. CREBRF variant increases obesity risk and protects against diabetes in Samoans. Nat Genet 2016;48:976-978

3. Naka I, Furusawa T, Kimura R, Natsuhara K, Yamauchi T, Nakazawa M, Ataka Y, Ishida T, Inaoka T, Matsumura Y. A missense variant, rs373863828-A (p. Arg457Gln), of CREBRF and body mass index in Oceanic populations. J Hum Genet 2017 Apr 13. doi: 10.1038/jhg.2017.44. [Epub ahead of print]

4. Loos RJ, Yeo GS. The bigger picture of FTO: the first GWAS-identified obesity gene. Nat Rev Endocrinol 2014;10:51-61

5. Karns R, Viali S, Tuitele J, Sun G, Cheng H, Weeks DE, McGarvey ST, Deka R. Common variants in FTO are not significantly associated with obesity-related phenotypes among Samoans of Polynesia. Ann Hum Genet 2012;76:17-24

6. Ohashi J, Naka I, Kimura R, Natsuhara K, Yamauchi T, Furusawa T, Nakazawa M, Ataka Y, Patarapotikul J, Nuchnoi P, Tokunaga K, Ishida T, Inaoka T, Matsumura Y, Ohtsuka R. FTO polymorphisms in oceanic populations. J Hum Genet 2007;52:1031-1035

7. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, Kosmicki JA, Duncan LE, Estrada K, Zhao F, Zou J, Pierce-Hoffman E, Berghout J, Cooper DN, Deflaux N, DePristo M, Do R, Flannick J, Fromer M, Gauthier L, Goldstein J, Gupta N, Howrigan D, Kiezun A, Kurki MI, Moonshine AL, Natarajan P, Orozco L, Peloso GM, Poplin R, Rivas MA, Ruano-Rubio V, Rose SA, Ruderfer DM, Shakir K, Stenson PD, Stevens C, Thomas BP, Tiao G, Tusie-Luna MT, Weisburd B, Won HH, Yu D, Altshuler DM, Ardissino D, Boehnke M, Danesh J, Donnelly S, Elosua R, Florez JC, Gabriel SB, Getz G, Glatt SJ, Hultman CM, Kathiresan S, Laakso M, McCarroll S, McCarthy MI, McGovern D, McPherson R, Neale BM, Palotie A, Purcell SM, Saleheen D, Scharf JM, Sklar P, Sullivan PF, Tuomilehto J, Tsuang MT, Watkins HC, Wilson JG, Daly MJ, MacArthur DG, Exome Aggregation Consortium. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285-291

8. Kilpelainen TO, Zillikens MC, Stancakova A, Finucane FM, Ried JS, Langenberg C, Zhang W, Beckmann JS, Luan J, Vandenput L, Styrkarsdottir U, Zhou Y, Smith AV, Zhao JH, Amin N, Vedantam S, Shin SY, Haritunians T, Fu M, Feitosa MF, Kumari M, Halldorsson BV, Tikkanen E, Mangino M, Hayward C, Song C, Arnold AM, Aulchenko YS, Oostra BA, Campbell H, Cupples LA, Davis KE, Doring A, Eiriksdottir G, Estrada K, Fernandez-Real JM, Garcia M, Gieger C, Glazer NL, Guiducci C, Hofman A, Humphries SE, Isomaa B, Jacobs LC, Jula A, Karasik D, Karlsson MK, Khaw KT, Kim LJ, Kivimaki M, Klopp N, Kuhnel B, Kuusisto J, Liu Y, Ljunggren O, Lorentzon M, Luben RN, McKnight B, Mellstrom D, Mitchell BD, Mooser V, Moreno JM, Mannisto S, O'Connell JR, Pascoe L, Peltonen L, Peral B, Perola M, Psaty BM, Salomaa V, Savage DB, Semple RK, Skaric-Juric T, Sigurdsson G, Song KS, Spector TD, Syvanen AC, Talmud PJ, Thorleifsson G, Thorsteinsdottir U, Uitterlinden AG, van Duijn CM, Vidal-Puig A, Wild SH, Wright AF, Clegg DJ, Schadt E, Wilson JF, Rudan I, Ripatti S, Borecki IB, Shuldiner AR, Ingelsson E, Jansson JO, Kaplan RC, Gudnason V, Harris TB, Groop L, Kiel DP, Rivadeneira F, Walker M, Barroso I, Vollenweider P, Waeber G, Chambers JC, Kooner JS, Soranzo N, Hirschhorn JN, Stefansson K, Wichmann HE, Ohlsson C, O'Rahilly S, Wareham NJ, Speliotes EK, Fox CS, Laakso M, Loos RJ. Genetic variation near IRS1 associates with reduced adiposity and an impaired metabolic profile. Nat Genet 2011;43:753-760

9. Yaghootkar H, Scott RA, White CC, Zhang W, Speliotes E, Munroe PB, Ehret GB, Bis JC, Fox CS, Walker M, Borecki IB, Knowles JW, Yerges-Armstrong L, Ohlsson C, Perry JR, Chambers JC, Kooner JS, Franceschini N, Langenberg C, Hivert MF, Dastani Z, Richards JB, Semple RK, Frayling

TM. Genetic evidence for a normal-weight "metabolically obese" phenotype linking insulin resistance, hypertension, coronary artery disease, and type 2 diabetes. Diabetes 2014;63:4369-4377

10. Scott RA, Fall T, Pasko D, Barker A, Sharp SJ, Arriola L, Balkau B, Barricarte A, Barroso I, Boeing H, Clavel-Chapelon F, Crowe FL, Dekker JM, Fagherazzi G, Ferrannini E, Forouhi NG, Franks PW, Gavrila D, Giedraitis V, Grioni S, Groop LC, Kaaks R, Key TJ, Kuhn T, Lotta LA, Nilsson PM, Overvad K, Palli D, Panico S, Quiros JR, Rolandsson O, Roswall N, Sacerdote C, Sala N, Sanchez MJ, Schulze MB, Siddiq A, Slimani N, Sluijs I, Spijkerman AM, Tjonneland A, Tumino R, van der AD, Yaghootkar H, Group RS, Consortium EP-I, McCarthy MI, Semple RK, Riboli E, Walker M, Ingelsson E, Frayling TM, Savage DB, Langenberg C, Wareham NJ. Common genetic variants highlight the role of insulin resistance and body fat distribution in type 2 diabetes, independent of obesity. Diabetes 2014;63:4378-4387

11. Yaghootkar H, Lotta LA, Tyrrell J, Smit RA, Jones SE, Donnelly L, Beaumont R, Campbell A, Tuke MA, Hayward C, Ruth KS, Padmanabhan S, Jukema JW, Palmer CC, Hattersley A, Freathy RM, Langenberg C, Wareham NJ, Wood AR, Murray A, Weedon MN, Sattar N, Pearson E, Scott RA, Frayling TM. Genetic Evidence for a Link Between Favorable Adiposity and Lower Risk of Type 2 Diabetes, Hypertension, and Heart Disease. Diabetes 2016;65:2448-2460

12. New Zealand Ministry of Health. The Health of New Zealand Children 2014/15: Key Findings of the NZ Health Survey. Wellington: Ministry of Health 2015.

13. Winnard D, Wright C, Jackson G, Gow P, Kerr A, McLachlan A, Orr-Walker B, Dalbeth N. Gout, diabetes and cardiovascular disease in the Aotearoa New Zealand adult population: co-prevalence and implications for clinical practice. NZ Med J 2012;126:53-64

14. Winnard D, Wright C, Taylor WJ, Jackson G, Te Karu L, Gow PJ, Arroll B, Thornley S, Gribben B, Dalbeth N. National prevalence of gout derived from administrative health data in Aotearoa New Zealand. Rheumatol 2012;51:901-909

15. Hollis-Moffatt JE, Phipps-Green AJ, Chapman B, Jones GT, van Rij A, Gow PJ, Harrison AA, Highton J, Jones PB, Montgomery GW, Stamp LK, Dalbeth N, Merriman TR. The renal urate transporter SLC17A1 locus: confirmation of association with gout. Arthritis Res Ther 2012;14:R92

16. Robinson PC, Choi HK, Do R, Merriman TR: Insight into rheumatological cause and effect through the use of Mendelian randomization. Nat Rev Rheumatol 2016;12:486-496.

17. Friedlaender JS, Friedlaender FR, Reed FA, Kidd KK, Kidd JR, Chambers GK, Lea RA, Loo JH, Koki G, Hodgson JA, Merriwether DA, Weber JL. The genetic structure of Pacific Islanders. PLoS Genet 2008;4:e19

18. Terrell JE. Social network analysis of the genetic structure of Pacific Islanders. Ann Hum Genet 2010;74:211-232

19. Yang J, Loos RJ, Powell JE, Medland SE, Speliotes EK, Chasman DI, Rose LM, Thorleifsson G, Steinthorsdottir V, Magi R, Waite L, Smith AV, Yerges-Armstrong LM, Monda KL, Hadley D, Mahajan A, Li G, Kapur K, Vitart V, Huffman JE, Wang SR, Palmer C, Esko T, Fischer K, Zhao JH, Demirkan A, Isaacs A, Feitosa MF, Luan J, Heard-Costa NL, White C, Jackson AU, Preuss M, Ziegler A, Eriksson J, Kutalik Z, Frau F, Nolte IM, Van Vliet-Ostaptchouk JV, Hottenga JJ, Jacobs KB, Verweij N, Goel A, Medina-Gomez C, Estrada K, Bragg-Gresham JL, Sanna S, Sidore C, Tyrer J, Teumer A, Prokopenko I, Mangino M, Lindgren CM, Assimes TL, Shuldiner AR, Hui J, Beilby JP, McArdle WL, Hall P, Haritunians T, Zgaga L, Kolcic I, Polasek O, Zemunik T, Oostra BA, Junttila MJ, Gronberg H, Schreiber S, Peters A, Hicks AA, Stephens J, Foad NS, Laitinen J, Pouta A, Kaakinen M, Willemsen G, Vink JM, Wild SH, Navis G, Asselbergs FW, Homuth G, John U, Iribarren C, Harris T, Launer L, Gudnason V, O'Connell JR, Boerwinkle E, Cadby G, Palmer LJ, James AL, Musk AW, Ingelsson E, Psaty BM, Beckmann JS, Waeber G, Vollenweider P, Hayward C, Wright AF, Rudan I, Groop LC, Metspalu A, Khaw KT, van Duijn CM, Borecki IB, Province MA, Wareham NJ, Tardif JC, Huikuri HV, Cupples LA, Atwood LD, Fox CS, Boehnke M, Collins FS, Mohlke KL, Erdmann J, Schunkert H, Hengstenberg C, Stark K, Lorentzon M, Ohlsson C, Cusi D, Staessen JA, Van der Klauw MM, Pramstaller PP, Kathiresan S, Jolley JD, Ripatti S, Jarvelin MR, de

Geus EJ, Boomsma DI, Penninx B, Wilson JF, Campbell H, Chanock SJ, van der Harst P, Hamsten A, Watkins H, Hofman A, Witteman JC, Zillikens MC, Uitterlinden AG, Rivadeneira F, Zillikens MC, Kiemeney LA, Vermeulen SH, Abecasis GR, Schlessinger D, Schipf S, Stumvoll M, Tonjes A, Spector TD, North KE, Lettre G, McCarthy MI, Berndt SI, Heath AC, Madden PA, Nyholt DR, Montgomery GW, Martin NG, McKnight B, Strachan DP, Hill WG, Snieder H, Ridker PM, Thorsteinsdottir U, Stefansson K, Frayling TM, Hirschhorn JN, Goddard ME, Visscher PM. FTO genotype is associated with phenotypic variability of body mass index. Nature 2012;490:267-272

20. Phipps-Green AJ, Merriman ME, Topless R, Altaf S, Montgomery GW, Franklin C, Jones GT, van Rij AM, White D, Stamp LK, Dalbeth N, Merriman TR: Twenty-eight loci that influence serum urate levels: analysis of association with gout. Ann Rheum Dis 2016;75:124-130

21. Wallace SL, Robinson H, Masi AT, Decker JL, McCarty DJ, Yu TF. Preliminary criteria for the classification of the acute arthritis of primary gout. Arthritis Rheum 1977;20:895-900

22. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, 3rd, Feldman HI, Kusek JW, Eggers P, Van Lente F, Greene T, Coresh J. CKD EPI: A new equation to estimate glomerular filtration rate. Ann Int Med 2009;150:604-612

23. Swinburn BA, Ley SJ, Carmichael HE, Plank LD. Body size and composition in Polynesians. Int J Obes Relat Metab Disord 1999;23:1178-1183

24. Guo Y, He J, Zhao S, Wu H, Zhong X, Sheng Q, Samuels DC, Shyr Y, Long J. Illumina human exome genotyping array clustering and quality control. Nat Protocols 2014;9:2643-2662

25. Illumina: Infinium Genotyping Data Analysis. A Guide for Analyzing Infinium Genotyping Data Using the GenomeStudio Genotyping Module. Tech Note 2014:1-10

26. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904-909

27. Topless RK, Flynn TJ, Cadzow M, Stamp LK, Dalbeth N, Black MA, Merriman TR. Association of SLC2A9 genotype with phenotypic variability of serum urate in pre-menopausal women. Front Genet 2015;6:313

28. Kogelman LJ, Fu J, Franke L, Greve JW, Hofker M, Rensen SS, Kadarmideen HN. Inter-tissue gene co-expression networks between metabolically healthy and unhealthy obese individuals. PloS One 2016;11:e0167519

29. Ying Z, Misra V, Verge VM. Sensing nerve injury at the axonal ER: activated Luman/CREB3 serves as a novel axonally synthesized retrograde regeneration signal. Proc Natl Acad Sic U S A 2014;111:16142-16147

30. Liang G, Audas TE, Li Y, Cockram GP, Dean JD, Martyn AC, Kokame K, Lu R. Luman/CREB3 induces transcription of the endoplasmic reticulum (ER) stress response protein Herp through an ER stress response element. Mol Cell Biol 2006;26:7999-8010

31. Kampstra P. Beanplot: a boxplot alternative for visual comparison of distributors. J Stat Software 2008;28:1-9

Table 1: Baseline characteristics

Characteristics	NZ Māori	Cook Island Māori	Samoan	Tongan	Niuean	Pukapukan	'Other' Polynesian
Number	1,154	197	378	175	47	70	265
GENDER (*n*)							
Male (proportion)	666 (0.577)	117 (0.594)	280 (0.741)	131 (0.749)	36 (0.766)	33 (0.471)	162 (0.611)
AGE (*years*) (SD)	52.25 (14.66)	52.16 (14.28)	45.25 (13.50)	43.61 (14.67)	48.28 (12.72)	44.14 (17.28)	42.38 (15.41)
WEIGHT (*kg*) (SD)	96.97 (23.93)	100.73 (25.28)	106.90 (23.05)	108.06 (21.61)	96.88 (18.78)	95.16 (20.85)	104.63 (28.35)
HEIGHT (*cm*) (SD)	169.57 (9.55)	169.03 (9.99)	173.39 (9.84)	174.18 (8.53)	171.73 (8.16)	165.56 (9.34)	171.59 (8.30)
BMI (kg/m^2) (SD)	33.79 (7.96)	35.37 (8.56)	35.63 (7.84)	35.55 (6.88)	32.77 (5.09)	34.60 (6.61)	35.43 (8.42)
WAIST CIRCUMFERENCE (*cm*) (SD)	108.04 (18.24)	110.75 (15.85)	111.80 (16.28)	112.53 (14.72)	106.06 (12.78)	108.20 (16.07)	110.06 (18.81)
***SUA (*mmol/L*) (SD)**	0.363 (0.11)	0.406 (0.096)	0.395 (0.13)	0.385 (0.13)	0.413 (0.11)	0.405 (0.092)	0.387 (0.11)
eGFR (*mL/min*) (SD)	71.23 (28.42)	69.36 (28.70)	72.49 (27.31)	73.00 (31.76)	70.94 (26.68)	79.79 (22.23)	76.25 (28.25)
[#]TRIGLYCERIDES (*mmol/L*) (SD)	2.25 (1.53)	2.37 (1.96)	2.13 (1.40)	2.21 (1.43)	2.52 (1.92)	2.09 (1.40)	2.10 (1.43)
[#]HDL (*mmol/L*) (SD)	1.16 (0.364)	1.17 (0.393)	1.10 (0.340)	1.09 (0.373)	1.10 (0.332)	1.21 (0.227)	1.13 (0.366)
[#]LDL (*mmol/L*) (SD)	2.80 (1.01)	2.88 (0.99)	2.87 (1.05)	2.82 (1.06)	3.03 (1.21)	2.71 (1.05)	2.76 (0.93)
TYPE 2 DIABETES (*n*) (proportion)	301 (0.273)	63 (0.330)	83 (0.226)	49 (0.290)	10 (0.213)	16 (0.229)	61 (0.234)
GOUT (*n*) (proportion)	456 (0.434)	96 (0.513)	207 (0.570)	92 (0.532)	21 (0.500)	11 (0.164)	100 (0.407)
CHRONIC KIDNEY DISEASE (*n*) (proportion)	115 (0.119)	23 (0.139)	31 (0.095)	23 (0.146)	4 (0.111)	0 (0.000)	24 (0.101)
***rs373863828* GENOTYPE (*n*)**							
G/G (proportion)	772 (0.674)	127 (0.651)	212 (0.568)	120 (0.686)	39 (0.830)	42 (0.600)	168 (0.641)
G/A (proportion)	348 (0.304)	60 (0.308)	146 (0.391)	47 (0.269)	7 (0.149)	22 (0.314)	77 (0.294)
A/A (proportion)	25 (0.022)	8 (0.041)	15 (0.040)	8 (0.046)	1 (0.021)	6 (0.086)	17 (0.065)
HWE *p*-value	0.049	0.79	0.098	0.23	0.34	0.22	0.052
ALLELE (*n*)							
A (proportion)	398 (0.174)	76 (0.195)	176 (0.236)	63 (0.180)	9 (0.096)	34 (0.243)	111 (0.212)

* Serum urate concentrations are reported from individuals not taking urate-lowering therapy. [#] Data were unavailable for lipid-lowering medications.

Table 2: *rs373863828* association with log-transformed BMI, untransformed BMI, metabolic, and lipid traits in the full Polynesian sample-set meta-analysis.

Continuous trait	**n**	**β [95% CI]**	***P***	**Covariables**
Log-transformed BMI – All	2,125	0.038 [0.022-0.055]	$4.8x10^{-6}$	4 PCA, sex, age, T2D
Male	1,335	0.042 [0.023~0.062]	$2.4x10^{-5}$	4 PCA, age, T2D
Female	790	0.032 [0.0014~0.062]	0.040	4 PCA, age, T2D
Untransformed BMI (kg/m^2) - All	2,125	1.38 [0.78~1.98]	$7.3x10^{-6}$	4 PCA, sex, age, T2D
Male	1,335	1.51 [0.78~2.23]	$4.4x10^{-5}$	4 PCA, age, T2D
Female	790	1.17 [0.089~2.26]	0.034	4 PCA, age, T2D
Untransformed BMI (kg/m^2) - All	1890	0.16 [-0.18~0.51]	0.35	4 PCA, sex, age, T2D, Waist
Major allele homozygote	1381	-	-	-
Heterozygote	670	1.80 [1.08~2.53]	$9.7x10^{-7}$	4 PCA, sex, age, T2D
Minor allele homozygote	74	1.49 [-0.32~3.30]	0.11	4 PCA, sex, age, T2D
Waist circumference (cm) - All	1,904	2.98 [1.64~4.33]	$1.4x10^{-5}$	4 PCA, sex, age, T2D
Male	1,221	2.81 [1.23~4.40]	$5.0x10^{-4}$	4 PCA, age, T2D
Female	683	3.29 [0.80~5.78]	$9.7x10^{-3}$	4 PCA, age, T2D
Waist circumference (cm) - All	1,890	0.66 [-0.11~1.42]	0.092	4 PCA, sex, age, T2D, BMI
Total cholesterol (mmol/L)	1,955	0.049 [-0.045~0.14]	0.30	4 PCA, sex, age
Total cholesterol (mmol/L)	1,901	0.029 [-0.067~0.13]	0.55	4 PCA, sex, age, BMI
Triglycerides (mmol/L)	1,957	-0.059 [-0.18~0.064]	0.35	4 PCA, sex, age
Triglycerides (mmol/L)	1,903	-0.11 [-0.24~0.011]	0.074	4 PCA, sex, age, BMI
HDL cholesterol (mmol/L)	1,954	-0.020 [-0.047~0.008]	0.17	4 PCA, sex, age
HDL cholesterol (mmol/L)	1,900	-0.0041 [-0.032~0.024]	0.77	4 PCA, sex, age, BMI
LDL cholesterol (mmol/L)	1,808	0.077 [0.0074~0.16]	0.074	4 PCA, sex, age
LDL cholesterol (mmol/L)	1,754	0.066 [-0.020~0.15]	0.13	4 PCA, sex, age, BMI
Urate (mmol/L)*	1,237	0.016 [0.005~0.026]	$5.0x10^{-3}$	4 PCA, sex, age
Urate (mmol/L)*	1,199	0.012 [0.0014~0.022]	0.026	4 PCA, sex, age, BMI
Dichotomous trait	**n**	**OR [95% CI]**	***P***	**Covariates**
Obesity (>32 kg/m^2)	2,125	1.33 [1.13~1.58]	$8x10^{-4}$	4 PCA, sex, age, T2D
Obesity (>40 kg/m^2)	2,125	1.54 [1.27~1.87]	$1.02x10^{-5}$	4 PCA, sex, age, T2D
Type 2 diabetes - All	2,125	0.65 [0.53~0.80]	$3.42x10^{-5}$	4 PCA, sex, age
Major allele homozygote	1,381	-	-	-
Heterozygote	670	0.55 [0.43~0.71]	$3.2x10^{-6}$	4 PCA, sex, age, BMI
Minor allele homozygote	74	0.72 [0.34~1.53]	0.39	4 PCA, sex, age, BMI
Type 2 diabetes - All	2,125	0.59 [0.47~0.73]	$1.9x10^{-6}$	4 PCA, sex, age, BMI
Male	1,335	0.54 [0.41~0.72]	$1.5x10^{-5}$	4 PCA, age, BMI
Female	779	0.66 [0.46~0.95]	0.026	4 PCA, age, BMI
Gout	2,114	1.10 [0.91~1.32]	0.34	4 PCA, sex, age
Gout	2,009	1.00 [0.81~1.22]	0.98	4 PCA, sex, age, BMI, T2D
Chronic kidney disease	1,849	0.72 [0.53~0.97]	0.030	4 PCA, sex, age
Chronic kidney disease	1,795	0.72 [0.53~0.99]	0.045	4 PCA, sex, age, BMI
Chronic kidney disease	1,810	0.86 [0.63~1.15]	0.36	4 PCA, sex, age, T2D
Chronic kidney disease	1,756	0.91 [0.65~1.28]	0.59	4 PCA, sex, age, BMI, T2D

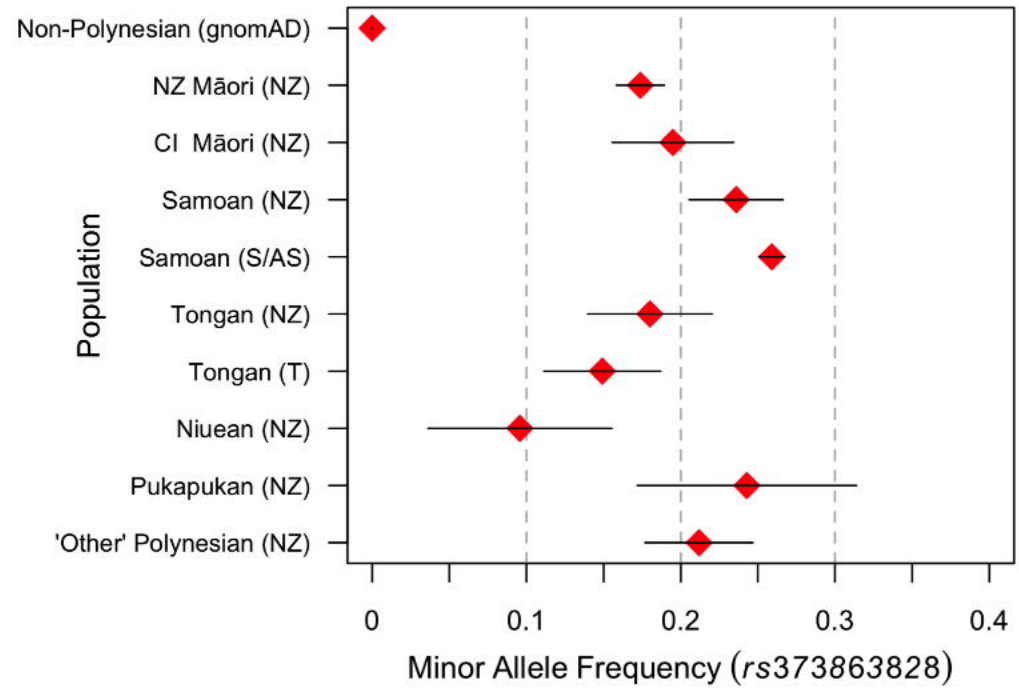
*Analysis excluded individuals taking urate-lowering medication

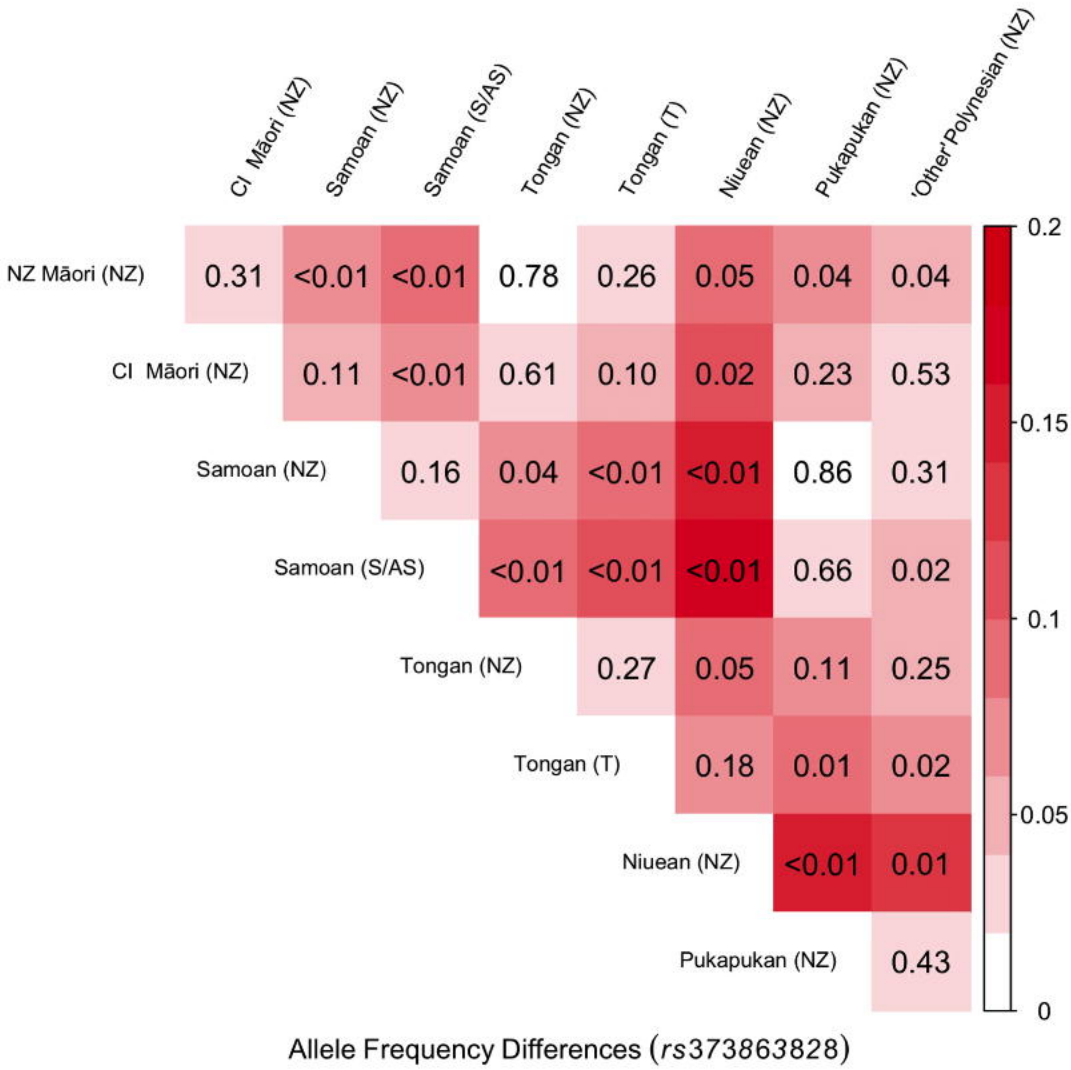
Figure 1: *rs373863828* allele frequencies in various Māori and Pacific ancestral groups (top) and comparison of allele frequencies between the Māori and Pacific ancestral groups (bottom). The *P* values in the bottom figure are derived from a difference in proportions parametric z-test. NZ – New Zealand. S / AS – Samoa / American Samoa (1). T – Kingdom of Tonga (3). Non-Polynesian includes all populations in the Genome Aggregation Database (gnomAD; gnomad.broadinstitute.org).

Figure 2: Forest plot of fixed effect meta-analysis for *rs373863828* with log-transformed BMI (A), untransformed BMI (B), and waist circumference (C). Association adjusted for age, sex, first four PCA vectors, and T2D.

Figure 3: Beanplots of BMI versus *rs373863828* genotype in men and women. A solid line shows the average for each group and the dotted line the overall average. Plots were generated using the R beanplot package (31).

Figure 4: Forest plot of fixed effect meta-analysis for *rs373863828* with T2D. Association adjusted for age, sex, first four PCA vectors and T2D (OR=0.59 [0.74~0.73]), $P=1.9\text{x}10^{-6}$).

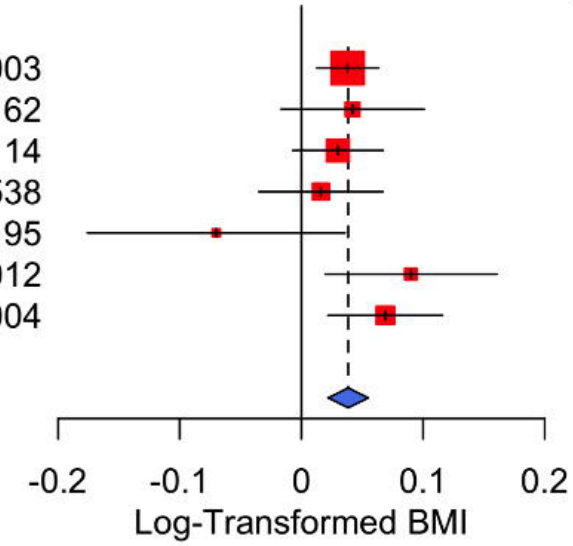

A
Non-Polynesian (gnomAD)
NZ Māori (NZ)
CI Māori (NZ)
Samoan (NZ)
Samoan (S/AS)
Tongan (NZ)
Tongan (T)
Niuean (NZ)
Pukapukan (NZ)
'Other' Polynesian (NZ)
Population
0
0.1
0.2
0.3
0.4
Minor Allele Frequency (rs373863828)


B
CI Māori (NZ)
Samoan (NZ)
Samoan (S/AS)
Tongan (NZ)
Tongan (T)
Niuean (NZ)
Pukapukan (NZ)
'Other' Polynesian (NZ)
NZ Māori (NZ) 0.31 <0.01 <0.01 0.78 0.26 0.05 0.04 0.04
CI Māori (NZ) 0.11 <0.01 0.61 0.10 0.02 0.23 0.53
Samoan (NZ) 0.16 0.04 <0.01 <0.01 0.86 0.31
Samoan (S/AS) <0.01 <0.01 <0.01 0.66 0.02
Tongan (NZ) 0.27 0.05 0.11 0.25
Tongan (T) 0.18 0.01 0.02
Niuean (NZ) <0.01 0.01
Pukapukan (NZ) 0.43
0.2
0.15
0.1
0.05
0
Allele Frequency Differences (rs373863828)

A

Cohort	Beta	[95% CI]	P
NZ Māori	0.038	[0.013; 0.063]	0.003
CI Māori	0.042	[-0.017; 0.101]	0.162
Samoan	0.030	[-0.007; 0.067]	0.114
Tongan	0.016	[-0.035; 0.067]	0.538
Niuean	-0.070	[-0.176; 0.036]	0.195
Pukapukan	0.090	[0.019; 0.161]	0.012
'Other' Polynesian	0.069	[0.022; 0.116]	0.004
Overall Effect	**0.038**	**[0.022; 0.055]**	

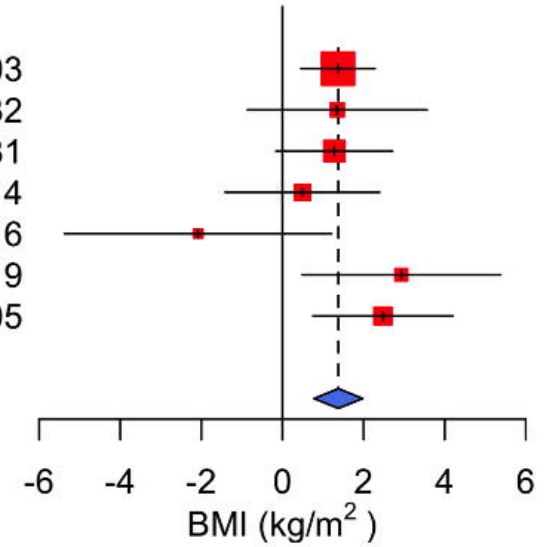
Heterogeneity: $\chi^2_6 = 8.66$ (p = 1.9e-01)
Test for overall effect: $z = 4.57$ (p = 4.8e-06)



B

Cohort	Beta	[95% CI]	P
NZ Māori	1.370	[0.457; 2.283]	0.003
CI Māori	1.350	[-0.865; 3.565]	0.232
Samoan	1.280	[-0.159; 2.719]	0.081
Tongan	0.490	[-1.413; 2.393]	0.614
Niuean	-2.080	[-5.373; 1.213]	0.216
Pukapukan	2.930	[0.480; 5.380]	0.019
'Other' Polynesian	2.480	[0.755; 4.205]	0.005
Overall Effect	**1.379**	**[0.777; 1.981]**	

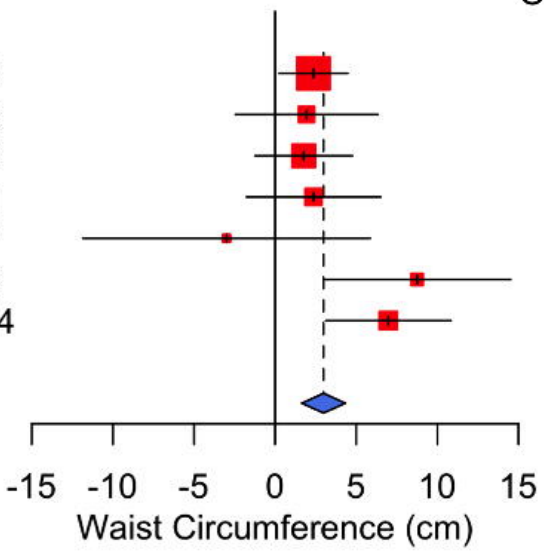
Heterogeneity: $\chi^2_6 = 8.20$ (p = 2.2e-01)
Test for overall effect: $z = 4.49$ (p = 7.2e-06)

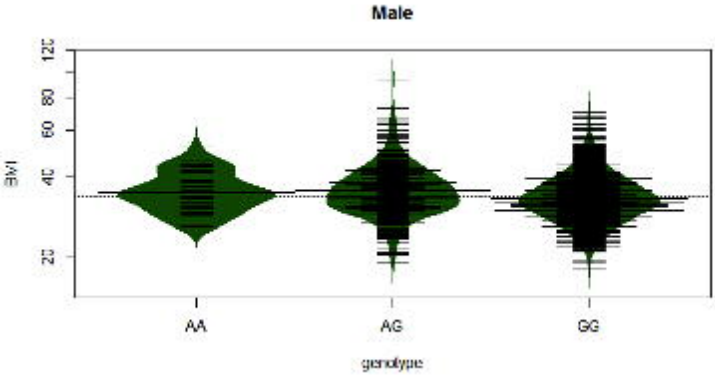


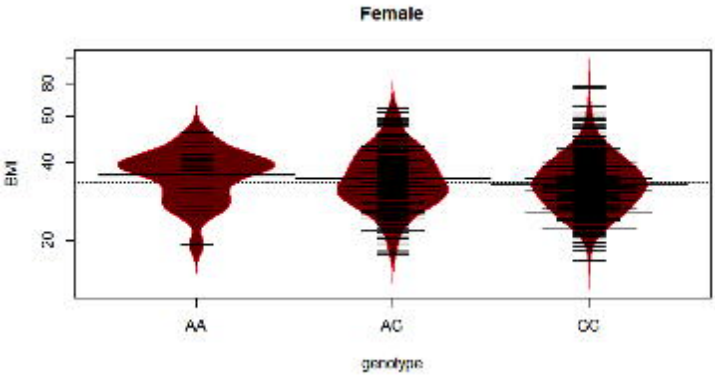
C

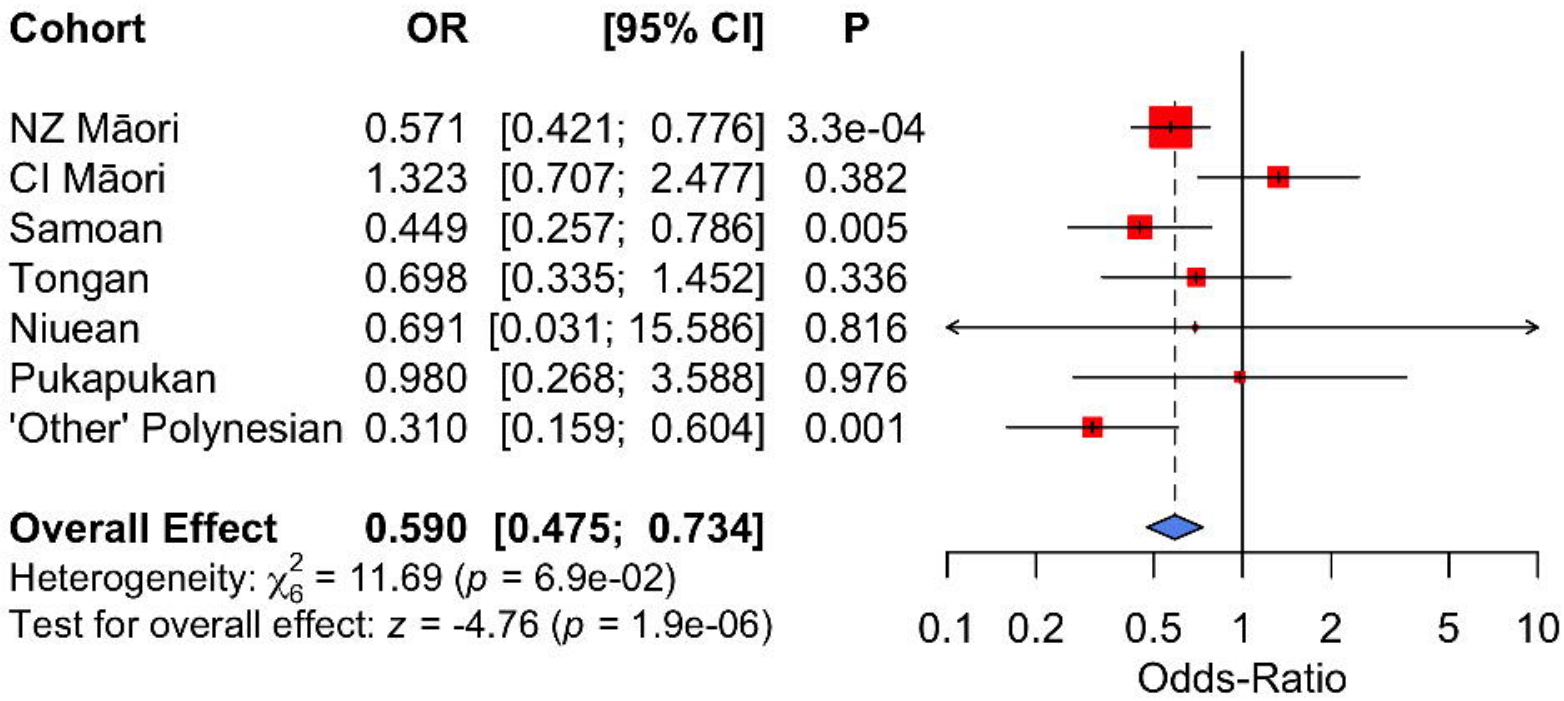
Cohort	Beta	[95% CI]	P
NZ Māori	2.360	[0.243; 4.477]	0.029
CI Māori	1.940	[-2.450; 6.330]	0.386
Samoan	1.760	[-1.239; 4.759]	0.250
Tongan	2.370	[-1.766; 6.506]	0.261
Niuean	-3.000	[-11.859; 5.859]	0.507
Pukapukan	8.750	[2.988; 14.512]	0.003
'Other' Polynesian	6.980	[3.119; 10.841]	4.0e-04
Overall Effect	**2.984**	**[1.641; 4.327]**	

Heterogeneity: $\chi^2_6 = 10.99$ (p = 8.9e-02)
Test for overall effect: $z = 4.36$ (p = 1.3e-05)




Male
BMI
120
80
60
40
20
AA
AG
GG
genotype


Female
BMI
80
60
40
20
AA
AC
CC
genotype


Cohort OR [95% CI] P
NZ Māori 0.571 [0.421; 0.776] 3.3e-04
CI Māori 1.323 [0.707; 2.477] 0.382
Samoan 0.449 [0.257; 0.786] 0.005
Tongan 0.698 [0.335; 1.452] 0.336
Niuean 0.691 [0.031; 15.586] 0.816
Pukapukan 0.980 [0.268; 3.588] 0.976
'Other' Polynesian 0.310 [0.159; 0.604] 0.001
Overall Effect 0.590 [0.475; 0.734]
Heterogeneity: χ^2_6 = 11.69 (p = 6.9e-02)
Test for overall effect: z = -4.76 (p = 1.9e-06)
0.1 0.2 0.5 1 2 5 10
Odds-Ratio