



Published in final edited form as:

*Diabetologia*. 2019 September ; 62(9): 1647–1652. doi:10.1007/s00125-019-4932-z.

## Association of *CREBRF* variants with obesity and diabetes in Pacific Islanders from Guam and Saipan

Robert L. Hanson<sup>1</sup>, Saied Safabakhsh<sup>2</sup>, Jeffrey M. Curtis<sup>1</sup>, Wen-Chi Hsueh<sup>1</sup>, Lois I. Jones<sup>1</sup>, Tanisha F. Aflague<sup>2</sup>, Jenny Duenas Sarmiento<sup>2</sup>, Satish Kumar<sup>3,4</sup>, Nicholas B. Blackburn<sup>3,4</sup>, Joanne E. Curran<sup>3,4</sup>, Darin Mahkee<sup>1</sup>, Leslie J. Baier<sup>1</sup>, William C. Knowler<sup>1</sup>, Robert G. Nelson<sup>1</sup>

<sup>1</sup>Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, 1550 E. Indian School Road, Phoenix, AZ 85014, USA

<sup>2</sup>Micronesian Institute for Disease Prevention and Research, Sinajana, Guam

<sup>3</sup>South Texas Diabetes and Obesity Institute, University of Texas Rio Grande Valley School of Medicine, Brownsville, TX, USA

<sup>4</sup>Department of Human Genetics, University of Texas Rio Grande Valley School of Medicine, Brownsville, TX, USA

### Abstract

**Aims hypothesis**—Variants in *CREBRF* (rs12513649 and rs373863828) have been strongly associated with increased BMI and decreased risk of type 2 diabetes in Polynesian populations; the A allele at rs373863828 is common in Polynesians but rare in most other global populations. The aim of the present study was to assess the association of *CREBRF* variants with obesity and diabetes in Pacific Islander (largely Marianas and Micronesian) populations from Guam and Saipan.

**Methods**—*CREBRF* rs12513649 and rs373863828 were genotyped in 2022 participants in a community-based cross-sectional study designed to identify determinants of diabetes and end-stage renal disease (ESRD). Associations were analysed with adjustment for age, sex, ESRD and the first four genetic principal components from a genome-wide association study (to account for

---

**Correspondence to:** Robert L. Hanson, Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, 1550 E. Indian School Road, Phoenix, AZ 85014, USA, rhanson@phx.niddk.nih.gov.

**Contribution statement** RLH contributed to study conception and design, data acquisition, analysis and interpretation of data, and drafting the manuscript. SS and RGN contributed to study conception and design, data acquisition, analysis and interpretation of data, and revising the draft for intellectual content. JMC and W-CH contributed to data acquisition, analysis and interpretation of data and revising the draft for intellectual content. WCK contributed to study conception and design, interpretation of data and revising the draft for intellectual content. LIJ, TFA, JDS, SK, NBB, JEC, DM and LJB contributed to data acquisition, interpretation of data, and revising the draft for intellectual content. All authors read and approved the final manuscript. RLH is the guarantor of the integrity of the work.

**Data availability** Data for consenting individuals will be made available through the Database of Genotype and Phenotype (<https://www.ncbi.nlm.nih.gov/gap>) pending Institutional Review Board approval.

**Duality of interest** The authors declare that there is no duality of interest associated with this manuscript.

**Publisher's Disclaimer:** This Author Accepted Manuscript is a PDF file of a an unedited peer-reviewed manuscript that has been accepted for publication but has not been copyedited or corrected. The official version of record that is published in the journal is kept up to date and so may therefore differ from this version.

population stratification); a genomic control procedure was used to account for residual stratification.

**Results**—The G allele at rs12513649 had an overall frequency of 7.7%, which varied from 2.2% to 20.7% across different Marianas and Micronesian populations; overall frequency of the A allele at rs373863828 was 4.2% (range: 1.1–5.4%). The G allele at rs12513649 was associated with higher BMI ( $\beta=1.55$  kg/m<sup>2</sup> per copy;  $p=0.0026$ ) as was the A allele at rs373863828 ( $\beta=1.48$  kg/m<sup>2</sup>,  $p=0.033$ ). The same alleles were associated with lower risk of diabetes (OR per copy: 0.63 [ $p=0.0063$ ] and 0.49 [ $p=0.0022$ ], respectively). Meta-analyses combining the current results with previous results in Polynesians showed a strong association between the A allele at rs373863828 and BMI ( $\beta=1.38$  kg/m<sup>2</sup>;  $p=2.5\times 10^{-29}$ ) and diabetes (OR=0.65,  $p=1.5\times 10^{-13}$ ).

**Conclusions interpretation**—These results confirm the associations of *CREBRF* variants with higher BMI and lower risk of diabetes and, importantly, they suggest that these variants contribute to the risk of obesity and diabetes in Oceanic populations.

### Keywords

*CREBRF*; Genetics; Micronesians; Marianas; Obesity; Type 2 diabetes mellitus

## Introduction

*CREBRF* (encoding cAMP response element binding protein 3 [CREB3] regulatory factor) was recently identified as an obesity gene in a genome-wide association study (GWAS) in Samoans, a Polynesian population [1]. The discovery variant, rs12513649, was strongly associated with BMI. In Samoans, rs12513649 was in virtually perfect concordance with rs373863828 (G>A), at which the A allele (frequency=25.9%), which is associated with higher BMI and decreased risk of diabetes, codes for an arginine to glutamine amino acid substitution in *CREBRF* (p.Arg457Gln or R457Q). A study in Polynesians from New Zealand replicated these associations with BMI and diabetes, with frequencies of the A allele ranging from 9.6% to 23.4% [2]. The BMI association with these variants was also replicated in Tongans, another Polynesian population with an A allele frequency of 15% [3]. A small study found frequency of the A allele to be 0–6% in other Oceanic populations (Melanesians and Micronesians) but its association with BMI was not statistically significant [3]. In the present study, we examined the associations of rs12513649 and rs373863828 with BMI and diabetes in Pacific Islanders, predominantly Marianas and Micronesians, from Guam and Saipan.

## Methods

### Participants

The study included 2022 participants from the Micronesian extension of the Family Investigation of Nephropathy and Diabetes, designed to identify genetic determinants of diabetes and kidney disease [4]. Participants were 18 years old and were recruited in 2008 and 2009 from dialysis centres, referrals of family members and general advertising. Participants were predominantly Chamorros, indigenous inhabitants of the Marianas islands, but also included Filipinos and several other Micronesian populations. Fasting plasma

glucose was measured in participants without known diabetes. Participants were defined as having diabetes if they had fasting plasma glucose  $\geq 7.0$  mmol/l, HbA<sub>1c</sub>  $\geq 6.5\%$  (44 mmol/mol) or a previous diagnosis of the condition [5]. Fasting serum insulin was measured and measures of insulin resistance and beta cell function were calculated according to the homeostatic model (HOMA-IR and HOMA-B) [6]. Participants reported their maximum lifetime body weight and height at this weight and the corresponding BMI was calculated. Characteristics of participants are shown in electronic supplementary material (ESM) Table 1.

## Genotypes

**Index *CREBRF* variants**—rs12513649 and rs373863828 were typed individually by KASP on Demand according to manufacturer's instructions (LGC, Teddington, England).

**Additional *CREBRF* variants**—We analysed 54 additional variants within 200 kb of the transcription start/stop sites for *CREBRF*. We generated the genotypes as part of a GWAS using the Illumina OmniExpress CoreExome array (Illumina, San Diego, CA, USA).

**Genomic control variants**—The GWAS-derived genotypes were used for a statistical genomic control procedure (see below). These variants include 269,234 SNPs with minor allele frequency  $>1\%$ , call rates  $>95\%$ , discordance rates  $<2.5\%$  in 100 duplicate specimens, and genotypic distributions consistent with the Hardy-Weinberg equilibrium ( $p > 0.0001$ ). In addition, whole genome sequence data (BGI, Shenzhen, Guangdong, China) were obtained for a subset of 40 full-heritage Chamorros.

## Statistics

Missing genotypes were imputed from phased data of all typed variants with BEAGLE (<http://faculty.washington.edu/browning/beagle/beagle.html>; version 3.3.2, accessed 30 May 2013) [7]. Statistical analyses were performed with SAS (SAS Institute, Cary, NC, USA). Associations between genotype and continuous variables (BMI, height, weight, HOMA) were assessed by linear regression models in which phenotype was regressed against number of risk alleles. Covariates included age, sex, end-stage renal disease (ESRD) and the first four principal components from the GWAS, which were included to account for potential population stratification (ESM Fig.1). To account for residual population stratification, a genomic control procedure was applied across all GWAS markers with a separate inflation parameter for each trait [8]. Association with diabetes was analysed by logistic regression with the same covariates (ESRD was included in the analyses presented to account for the study design; results were similar if it was not included.) A fixed-effects meta-analysis for associations with BMI and diabetes was performed to combine results for the present study with those reported in other Oceanic populations [1-3]. Effect sizes were weighted by the inverse of their variance, and heterogeneity was assessed using Cochran's  $Q$  statistic and the  $I^2$  measure [9].

## Results

### Allele frequencies

Frequencies of alleles at rs373863828 and rs12513649 and their haplotypes are shown in Table 1. There was strong linkage disequilibrium between the A allele at rs373863828 and the G allele at rs12513649, as indicated by high  $D'$  values. Since allele frequencies differed, allelic concordance was moderate, with an overall  $r^2$  of 0.48.

### Associations with BMI and diabetes

Associations of rs373863828 and rs12513649 with BMI and diabetes are shown in Fig. 1a-d. The G allele at rs12513649 was associated with higher BMI ( $\beta=1.55$  kg/m<sup>2</sup> per copy;  $p=0.0026$ ) as was the A allele at rs373863828 ( $\beta=1.48$  kg/m<sup>2</sup>;  $p=0.033$ ). These obesity-associated alleles were associated with lower prevalence of diabetes: the OR per copy of the G allele at rs12513649 was 0.63 ( $p=0.0063$ ), while the OR for the A allele at rs373863828 was 0.49 ( $p=0.0022$ ). To further account for potential population stratification, we additionally adjusted for self-reported primary ethnicity with similar results (ESM Table 2, which also shows results according to primary ethnicity). There was a modest association of the A allele at rs373863828 with taller height ( $p=0.0097$ , ESM Fig. 2); however, both rs12513649 and rs373863828 were significantly associated with weight after adjustment for height (ESM Fig. 3). Higher BMI was strongly associated with higher prevalence of diabetes (ESM Fig. 4), and there was no significant difference in this relationship when rs373863828 A allele carriers were compared with non-carriers ( $p=0.81$  for interaction).

By design, a large proportion of individuals had ESRD. The analyses were repeated in the 1682 individuals without ESRD to exclude undue influence from ESRD. The associations of rs12513649 and rs373863828 with BMI and diabetes remained significant after this exclusion (ESM Table 3). Similarly, associations with BMI remained significant when analysed in the 1065 individuals who did not have diabetes or ESRD (ESM Table 3). In the 987 non-diabetic individuals who did not have ESRD, HOMA-B was significantly higher in carriers of the A allele at rs373863828 (by 20% per copy;  $p=0.031$ ), but HOMA-IR was not ( $p=0.12$ , ESM Fig. 5); the association with HOMA-B was attenuated with adjustment for BMI ( $p=0.22$ , ESM Fig. 5).

### Associations with additional *CREBRF* variants

In the analysis of all 56 GWAS variants mapping within 200 kb of the transcription start/stop site, rs12513649 achieved the strongest association with BMI (ESM Fig. 6, ESM Table 4), and rs373863828 had the strongest association with diabetes (ESM Fig. 7, ESM Table 4). In models where the effects of rs12513649 and rs373863828 were each conditioned on the effects of the other, rs12513649 remained significantly associated with BMI ( $p=0.033$ , data not shown) while rs373863828 did not ( $p=0.97$ , ESM Fig. 6), and rs373863828 was suggestively, but non-significantly, associated with diabetes ( $p=0.080$ , data not shown) while rs12513649 was not ( $p=0.45$ , ESM Fig. 7). In addition, other than rs373863828, missense variants in *CREBRF* were not detected in whole-genome sequence data from 40 Chamorros.

## Meta-analyses

Meta-analyses combining results of associations with rs373863828 in the present study with those from previous publications are shown in Fig. 1e-f. The A allele was associated with higher BMI (combined  $\beta=1.38$  kg/m<sup>2</sup> per allele;  $p=2.5\times 10^{-29}$ ) and there was no evidence for heterogeneity across studies ( $I^2=0\%$ ,  $p=0.58$ ). The A allele was also associated with lower risk of diabetes, and there was little evidence for heterogeneity ( $I^2=31\%$ ,  $p=0.23$ ). Although none of the studies investigating the association between rs373863828 and diabetes individually achieved genome-wide statistical significance, the combined effect was significant (OR 0.65,  $p=1.5\times 10^{-13}$ ).

## Discussion

*CREBRF* variants are associated with increased BMI in Polynesian populations [1-3, 10]. The frequency of the A allele at rs373863828 in Polynesians is 9.6–25.9%, whereas this allele is exceedingly rare in most other global populations (e.g. the frequency was 0.004% in the Exome Aggregation Project; <https://www.ncbi.nlm.nih.gov/snp/>, accessed 3 November 2018). In Polynesians, the A allele was associated with higher BMI but lower risk of diabetes [1,2]. In the present study, we extend these findings and show that the A allele is present at low frequencies, between 1.1% and 5.4%, in Marianas and Micronesian populations (Chamorros, Chuukese, Carolinians, Palauans), in whom it is also associated with higher BMI and lower risk of diabetes.

The association with BMI was initially identified with rs12513649, which is in nearly perfect concordance with rs373863828 in Polynesian populations [1]. In the current sample, the concordance is less strong, which, theoretically, allows statistical analyses to separate effects of each SNP. However, results of the present conditional analyses are equivocal and larger sample sizes are required to confidently identify independent effects. Functional studies in murine adipocytes found that the A allele at rs373863828 results in increased lipid accumulation and decreased cellular energy expenditure [1]. The functionality of rs12513649 is not clear, but scores derived from the Genome-Wide Annotation of Variants program (GWAVA; [https://www.sanger.ac.uk/sanger/StatGen\\_Gwava](https://www.sanger.ac.uk/sanger/StatGen_Gwava), accessed 5 April 2019), which assesses potential functionality from a variety of genomic annotation features, including epigenetic marks and conservation, are 0.27–0.41; this is only modestly suggestive that it is functional [11]. Given this, and the fact that rs12513649 is not associated with BMI in GWAS from East Asian populations [12], in whom the G allele has a frequency of 6.3% (<https://www.ncbi.nlm.nih.gov/snp/>, accessed 3 November 2018), it seems likely that rs373863828 is the functional variant. The present study did not find significant evidence for additional variants in *CREBRF* that influence BMI or diabetes risk.

The mechanisms behind the ‘discordant’ associations, in which the allele associated with higher BMI is associated with lower risk for diabetes, require further investigation. The present findings suggest that the A allele at rs373863828 does not modify the effect of BMI on diabetes risk but that it separately lowers diabetes risk while increasing BMI. The association of the A allele with higher HOMA-B suggests that it may result in greater capacity for compensatory insulin secretion, but more detailed physiological studies are required. Regardless of these mechanisms, the study findings demonstrate that genetic

variants in *CREBRF* are important determinants of obesity and diabetes risk in Oceanic populations.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors thank the participants who volunteered for these studies and the staff who helped to conduct them. Particular thanks are due to J. Loebel and V. Ossowski, Phoenix Epidemiology and Clinical Research Branch, for help with laboratory work. This work was presented in part at the 78th Scientific Sessions of the American Diabetes Association, Orlando, FL, USA, 22–26 June 2018.

**Funding** This work was supported by the intramural research programme of the National Institute of Diabetes and Digestive and Kidney Diseases. The genotyping was conducted in part in facilities constructed under the support of NIH grant C06 RR020547.

## Abbreviations

|             |                               |
|-------------|-------------------------------|
| <b>ESRD</b> | End-stage renal disease       |
| <b>GWAS</b> | Genome-wide association study |

## References

1. Minster RL, Hawley NL, Su CT, et al. (2016) A thrifty variant in *CREBRF* strongly influences body mass index in Samoans. *Nat Genet* 48:1049–1054. [PubMed: 27455349]
2. Krishnan M, Major TJ, Topless RK, et al. (2018) Discordant association of the *CREBRF* rs373863828 A allele with increased BMI and protection from type 2 diabetes in Maori and Pacific (Polynesian) people living in Aotearoa/New Zealand. *Diabetologia* 61:1603–1613. [PubMed: 29721634]
3. Naka I, Furusawa T, Kimura R, et al. (2017) A missense variant, rs373863828-A (p.Arg457Gln), of *CREBRF* and body mass index in Oceanic populations. *J Hum Genet* 62:847–849. [PubMed: 28405013]
4. Knowler WC, Coresh J, Elston RC, et al. (2005) The Family Investigation of Nephropathy and Diabetes (FIND): design and methods. *J Diabetes Complications* 19:1–9. [PubMed: 15642484]
5. American Diabetes Association. (2010) Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care* 33 (Suppl 1):S62–S69. [PubMed: 20042775]
6. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC (1985) Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 28:412–419. [PubMed: 3899825]
7. Browning SR, Browning BL (2007) Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet* 81:1084–1097. [PubMed: 17924348]
8. Devlin B, Bacanu SA, Roeder K (2004) Genomic control to the extreme. *Nat Genet* 36:1129–1130. [PubMed: 15514657]
9. Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. *Brit Med J* 327:557–560. [PubMed: 12958120]
10. Berry SD, Walker CG, Ly K, et al. (2018) Widespread prevalence of a *CREBRF* variant amongst Maori and Pacific children is associated with weight and height in early childhood. *Int J Obes* 42:603–607.
11. Ritchie GRS, Dunham I, Zeggini E, Flicek P (2014) Functional annotation of noncoding sequence variants. *Nature Methods* 11:294–296. [PubMed: 24487584]

12. Wen W, Zheng W, Okada Y, et al. (2014) Meta-analysis of genome-wide association studies in East Asian-ancestry populations identifies four new loci for body mass index. *Hum Molec Genet* 23:5492–5504. [PubMed: 24861553]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



## Research in context

### What is already known about this subject?

- Genetic variants in *CREBRF* (the G allele at rs12513649 and the A allele at rs373863828] are strongly associated with higher BMI in Polynesian populations
- The alleles associated with higher BMI have also been associated with lower risk of type 2 diabetes
- The A allele at rs373863828 is common in Polynesian populations, in whom it is associated with higher BMI, but it is very rare in most global populations; it is present at low frequency in some non-Polynesian Oceanic populations

### What is the key question?

- What are the frequencies of variants at rs12513649 and rs373863828 in Pacific Islander populations, largely Marianas and Micronesian, from Guam and Saipan, and what are the associations of these alleles with BMI and diabetes?

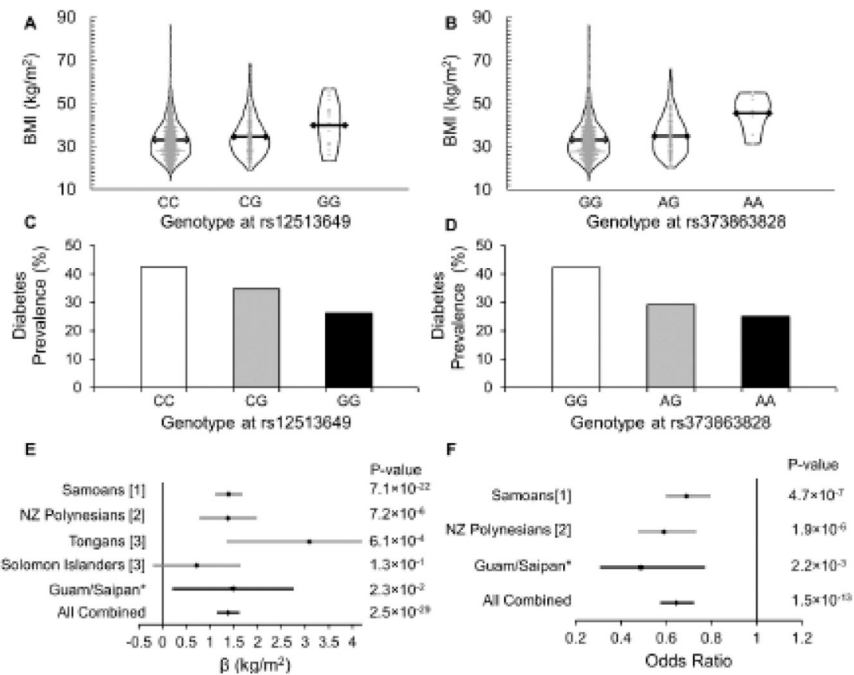
### What are the new findings?

- The frequency of the A allele at rs373863828 was 4.2% overall, and it ranged from 1.1% to 5.4% across different Marianas and Micronesian populations; the frequency of the G allele at rs12513649 was 7.7% (22–20.7%)
- The A allele at rs373863828 and the G allele at rs12513649 were associated with higher BMI and with lower risk of type 2 diabetes in the study population
- The magnitude of the associations observed in the present study was comparable to that previously reported in Polynesian populations

### How might this impact on clinical practice in the foreseeable future?

- Understanding the mechanisms by which *CREBRF* variants confer risk for obesity and diabetes may lead to new therapies for these conditions



**Fig. 1.**

(a) ‘Bean’ plot of the distribution of BMI according to genotypes at rs12513649. Grey horizontal lines represent individual values, with the length of the line corresponding to the number of observations at each level. The symmetrical plots represent the density at each BMI level, estimated using the kernel density estimation function in SAS. Black horizontal bars represent the mean value.  $\beta=1.55 \text{ kg/m}^2$  per copy of the G allele;  $p=0.0026$ . (b) ‘Bean’ plot of the distribution of BMI according to genotypes at rs373863828.  $\beta=1.48 \text{ kg/m}^2$  per copy of the A allele;  $p=0.033$ . (c) Prevalence of diabetes according to genotype at rs12513649. OR 0.63 per copy of the G allele;  $p=0.0063$ . (d) Prevalence of diabetes according to genotype at rs373863828. OR 0.49 per copy of the A allele;  $p=0.0022$ . (e) Meta-analysis of the association of the A allele at rs373863828 with BMI, including data from the present study (labeled ‘Guam/Saipan’). Data are presented as the regression coefficient ( $\beta$ ,  $\text{kg/m}^2$  per copy of the A allele) with 95% CI. (f) Meta-analysis of the association of the A allele at rs373863828 with diabetes. Data are presented as OR per copy of the A allele with 95% CI. NZ, New Zealand

Allele and haplotype frequencies at rs12513649 and rs373863828 in *CREBRF* among Pacific Islanders from Guam and Saipan

**Table 1**

| Population       | n    | rs12513649 frequency (%) |     | rs373863828 frequency (%) | Haplotype frequency (%) <sup>a</sup> |     |      |      | Association measures |                |
|------------------|------|--------------------------|-----|---------------------------|--------------------------------------|-----|------|------|----------------------|----------------|
|                  |      | G                        | A   |                           | G/A                                  | G/G | C/A  | C/G  | D'                   | r <sup>2</sup> |
| Chamorros        | 928  | 8.0                      | 5.4 | 5.2                       | 2.8                                  | 0.2 | 91.7 | 0.94 | 0.59                 |                |
| Chuukese         | 174  | 6.3                      | 4.9 | 4.9                       | 1.4                                  | 0.0 | 93.7 | 1.00 | 0.76                 |                |
| Palauans         | 46   | 20.7                     | 2.2 | 2.2                       | 18.5                                 | 0.0 | 79.3 | 1.00 | 0.09                 |                |
| Carolinians      | 45   | 2.2                      | 1.1 | 1.1                       | 1.1                                  | 0.0 | 97.8 | 1.00 | 0.49                 |                |
| Filipinos        | 249  | 4.0                      | 0.2 | 0.2                       | 3.8                                  | 0.0 | 96.0 | 1.00 | 0.05                 |                |
| Others           | 580  | 8.8                      | 4.0 | 3.9                       | 4.9                                  | 0.1 | 91.1 | 0.98 | 0.41                 |                |
| All participants | 2022 | 7.7                      | 4.2 | 4.0                       | 3.7                                  | 0.2 | 92.1 | 0.96 | 0.48                 |                |

Frequencies were calculated for individuals who were full-heritage Chamorro, Chuukese, Palauan, Carolinian and Filipino by self-report. 'Others' include those from other populations or mixed heritage

<sup>a</sup>Haplotypes are defined by the alleles at rs12513649 and rs373863828 with the allele at rs12513649 listed first

D', linkage disequilibrium coefficient normalised by allele frequencies; r<sup>2</sup> allelic concordance expressed as the square of correlation coefficient