

TITLE

CREBRF missense variant rs373863828 has both direct and indirect effects on type 2 diabetes and fasting glucose in Polynesians living in Samoa and Aotearoa New Zealand.

SHORT RUNNING TITLE

CREBRF and type 2 diabetes

AUTHORS

Emily M. Russell, PhD¹, Jenna C. Carlson, PhD^{1,2}, Mohanraj Krishnan, PhD¹, Nicola L. Hawley, PhD³, Guangyun Sun, PhD⁴, Hong Cheng, BS⁴, Take Naseri, MBBS, MPH⁵, Muagututi'a Sefuiva Reupena, MS⁶, Satupa'itea Viali, MBChB, MPH⁷, John Tuitele, MBBS⁸, Tanya J. Major, PhD⁹, Iva Miljkovic, MD, PhD¹⁰, Tony R. Merriman, PhD^{9,11}, Ranjan Deka, PhD⁴, Daniel E. Weeks, PhD^{1,2}, Stephen T. McGarvey, PhD^{12,13}, and Ryan L. Minster, PhD¹

AUTHORS AFFILIATIONS

1. Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
2. Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
3. Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, Connecticut, USA
4. Department of Environmental and Public Health Sciences, College of Medicine, University of Cincinnati, Cincinnati, Ohio, USA
5. Ministry of Health, Government of Samoa, Apia, Samoa
6. Lutia i Puava 'ae Mapu i Fagalele, Apia, Samoa
7. School of Medicine, National University of Samoa, Apia, Samoa
8. Department of Public Health, Lyndon B. Johnson Tropical Medical Center, Faga'alu, American Samoa, USA
9. Department of Biochemistry, University of Otago, Dunedin, New Zealand
10. Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
11. Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, Alabama, USA
12. International Health Institute and Department of Epidemiology, School of Public Health, Brown University, Providence, Rhode Island, USA
13. Department of Anthropology, Brown University, Providence, Rhode Island, USA

CORRESPONDING AUTHOR

Ryan L. Minster
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh
Pittsburgh, PA 15261
E-mail: ryan.minster@pitt.edu
Primary Phone: 412-624-6928

Abstract

Objective

The minor allele of rs373863828 in *CREBRF* is associated with higher BMI, lower fasting glucose, and lower odds of type 2 diabetes. We examined the associations between BMI and rs373863828 on type 2 diabetes and fasting glucose with a large sample of adult Polynesians from Samoa, American Samoa and Aotearoa New Zealand and estimated direct and indirect (via BMI) effects of rs373863828 on type 2 diabetes and fasting glucose.

Research Design and Methods

We regressed type 2 diabetes and fasting glucose on BMI and rs373863828 stratified by obesity, regressed type 2 diabetes and fasting glucose on BMI stratified by rs373863828 genotype, and assessed the effects of rs373863828 on type 2 diabetes and fasting glucose with path analysis.

Results

Association of BMI with fasting glucose was greater in those without obesity than in those with obesity. We did not observe evidence of differences by genotype. In the path analysis, the minor allele has direct negative and indirect positive effects on type 2 diabetes risk and fasting glucose, with the indirect effect mediated through a *direct* positive effect on BMI.

Conclusions

There may be a stronger effect of BMI on fasting glucose in Polynesians without obesity than in those with obesity. Carrying the rs373863828 minor allele does not decouple higher BMI from odds of type 2 diabetes. Given the current cost of genotyping compared

Russell *et al.* *CREBRF* and type 2 diabetes

to the accessibility of measuring BMI, including rs373863828 as a clinical predictor of type 2 diabetes may not be indicated.

Keywords (up to 10)

CREBRF, type 2 diabetes, Samoa, Polynesia, BMI, genetics

Abbreviations

cAge	mean-centered scaled age
cAge ²	mean-centered scaled age squared
<i>CREBRF</i>	cAMP-responsive element binding protein 3 regulatory factor
lnBMI	natural logarithm of BMI
lnFG	natural logarithm of fasting blood glucose
MAF	minor allele frequency
OR	odds ratio
RMSEA	root mean square error of approximation

Main Document

High BMI is a risk factor for type 2 diabetes and elevated fasting glucose levels. While type 2 diabetes and obesity are highly correlated, there is little overlap in genetic variants associated with each phenotype [1]. One genetic variant associated with both, the minor allele A of a missense variant in *CREBRF* (rs373863828, c.1370G>A, p.R457Q) is simultaneously associated with higher BMI (1.36 kg/m² per copy of the A allele) and, contrary to expectations, lower odds of type 2 diabetes (0.586 OR) and lower fasting glucose (−1.65 mg/dL per copy of the A allele) in Samoans and other Pacific Island populations [2–6]. Alleles with such discrepant effects have been termed “favorable adiposity” alleles [7, 8]. Favorable adiposity alleles like those of rs373863828 in *CREBRF* and variants in *ADAMTS9*, *GRB14/COBLL1*, and *TCF7L2* [1, 9] could reveal avenues for the study of the variation in biological pathways underlying the relationship between BMI and type 2 diabetes. The biological mechanisms of *CREBRF* and its role in development of obesity and type 2 diabetes remain largely unknown.

The association of rs373863828 with BMI, obesity, type 2 diabetes, and fasting glucose was first identified in Samoans from Samoa and American Samoa [2] and has since been replicated in other Pacific populations[3–6]. rs373863828 is common in Samoans (minor allele frequency [MAF] = 0.259)[2], Tongans (MAF = 0.150)[5], Pukapukans (MAF = 0.243), Niueans (MAF = 0.096), Cook Island Māori (MAF = 0.195), New Zealand Māori (MAF = 0.174)[3], Native Hawaiians (MAF = 0.128)[6], and Pacific Island people from Guam and Saipan (MAF = 0.042)[4] but is very rare in non-Pacific populations (MAF < 0.001) [10, 11].

The prevalence of obesity in Polynesia is among the highest in the world, and temporal increases in obesity-related disorders such as type 2 diabetes represent a growing public health problem in the region [12]. From 1978 to 2013, the prevalence of obesity (≥ 30 kg/m²) in men rose from 28% to 53% and in women, from 45% to 78% [13]. Similarly, from 1980 to 2014, the age-standardized prevalence of type 2 diabetes of Samoan men and women increased from 6.1% and 9.0%, respectively, to 22.7% and 26.6%, respectively [14].

Previous work has focused on the direct associations of rs373863828 with BMI, type 2 diabetes, and fasting glucose [2–5]. Here we examine, more extensively, the relationships among type 2 diabetes, fasting glucose, BMI, and rs373863828 in three samples of adult Samoans ($n = 2,861$, $n = 1,083$ and $n = 1,013$) living in Samoa and American Samoa and a sample of adult Polynesians ($n = 1,270$) of several ancestries living in Aotearoa New Zealand. This combined study ($n = 7,127$) is the largest one thus far examining the relationship between rs373863828, type 2 diabetes, and BMI [2, 3]. This is also the first time that path analysis has been used to explore the direct and indirect effects of rs373863828 on type 2 diabetes and fasting glucose. Studies such as this may inform the utility of including a screen for rs373863828 in clinical practice.

Methods

Participants and Phenotypes

For this study, we worked with three samples of Samoan participants from Samoa and American Samoa [2, 15–20] from a 1990–95 longitudinal study of cardiometabolic disease, a 2002–03 family-based cross-sectional study of adiposity-related trait genetics, and a 2010 population-based study of cardiometabolic disease genetics; and a sample of participants of Māori and Pacific (Polynesian) ancestry from the Genetics of Gout, Diabetes and Kidney Disease in Aotearoa New Zealand study [3].

Adult participants in the longitudinal 1990–95 study were recruited from Samoa and American Samoa. They reported Samoan ancestry and no previous diagnosis of diabetes, heart disease, or hypertension at baseline [18–20]. Whole blood was collected after an overnight fast for the fasting glucose measurement. Fasting glucose ≥ 126 mg/dl at baseline was used to define type 2 diabetes [2, 18]. The final sample was comprised of 1,013 individuals with full phenotypic and genotypic data (Table 1a).

Participants of the cross-sectional 2002–03 family study reported Samoan ancestry and were recruited from Samoa and American Samoa as members of families of probands selected from the longitudinal 1990–1995 study [2, 16, 17]. Whole blood was collected after a 10 hour fast for the fasting glucose measurement. Fasting glucose ≥ 126 mg/dl and/or a previous diabetes diagnosis was used to define type 2 diabetes [2, 16]. The final sample was comprised of 1,083 adults with full phenotypic and genotypic data (Table 1b).

Participants in the 2010 Samoan sample were drawn from a cross-sectional population-based study of adults of reported Samoan ancestry recruited in Samoa. Details of the sample population and recruitment have been described elsewhere [2, 15]. In brief, whole blood samples were collected following a 10 hour fast for the fasting glucose measurement. Type 2 diabetes was defined as fasting glucose ≥ 126 mg/dl and/or currently taking medication to treat diabetes [2]. The final sample was comprised of 2,861 participants with full phenotypic and genotypic data (Table 1c).

The sample from Aotearoa New Zealand ($n=2,124$) was recruited from 2006–2017 [21]. Type 2 diabetes was ascertained by physician-diagnosis and/or participant reports and/or use of glucose lowering therapy [3]. Ancestry of participants was based on both self-reported NZ Māori and/or Pacific ancestries of their grandparents and clustering of genome-wide principal components resulting in 2,124 adults of Polynesian ancestry. Fasting glucose was not measured in this study (Table 1d).

In all studies, height and weight were measured in duplicate and were used to calculate BMI. Compared to people of European ancestries, people of Polynesian ancestries have a greater ratio of lean mass to fat mass at higher BMIs [22, 23]. Therefore, when classifying participants into those with and without obesity, we used a threshold of BMI > 32 kg/m², which is more appropriate in this population than the standard WHO cutoff of ≥ 30 kg/m² [22].

Table 1: Sample demographics information stratified by sex and diabetes status

a. 1990–95 Samoan and American Samoan sample

Russell *et al.* CREBRF and type 2 diabetes

Overall (n = 1,013)			Males (n = 477, 47.1%)		Females (n = 536, 52.9%)	
	Diabetes (n = 63, 6.2%)	No diabetes (n = 950, 93.8%)	Diabetes (n = 34, 7.1%)	No diabetes (n = 443, 92.9%)	Diabetes (n = 29, 5.4%)	No diabetes (n = 507, 94.6%)
age (mean years (s.d))	46.0 (9.4)	38.8 (9.4)	46.3 (8.8)	38.8 (9.3)	45.7 (10.1)	38.8 (9.6)
BMI (mean kg/m ² (s.d.))	35.7 (8.0)	31.8 (6.1)	34.8 (6.2)	30.6 (5.7)	36.7 (9.7)	32.8 (6.3)
Overall (n = 1,013)			Males (n = 477, 47.1%)		Females (n = 536, 52.9%)	
	Diabetes (n = 63, 6.2%)	No diabetes (n = 950, 93.8%)	Diabetes (n = 34, 7.1%)	No diabetes (n = 443, 92.9%)	Diabetes (n = 29, 5.4%)	No diabetes (n = 507, 94.6%)
fasting glucose (mean mg/dL (s.d.))	201.8 (68.8)	88.4 (12.1)	198.8 (55.3)	88.6 (12.6)	205.4 (82.7)	88.1 (11.6)
b. 2002–03 Samoan and American Samoan sample						
Overall (n = 1,083)			Males (n = 499, 46.1%)		Females (n = 584, 53.9%)	
	Diabetes (n = 161, 14.9%)	No diabetes (n = 922, 85.1%)	Diabetes (n = 74, 14.8%)	No diabetes (n = 425, 85.2%)	Diabetes (n = 87, 14.9%)	No diabetes (n = 497, 85.1%)
age (mean years (s.d))	54.8 (12.3)	40.7 (16.1)	56.0 (12.0)	39.5 (15.8)	53.7 (12.4)	41.6 (16.4)
BMI (mean kg/m ² (s.d.))	36.4 (8.8)	32.8 (7.7)	35.4 (9.7)	30.4 (6.1)	37.1 (7.8)	34.8 (8.3)
Overall (n = 1,077)			Males (n = 497, 46.1%)		Females (n = 580, 53.9%)	
	Diabetes (n = 160, 14.9%)	No diabetes (n = 917, 85.1%)	Diabetes (n = 74, 14.9%)	No diabetes (n = 423, 85.1%)	Diabetes (n = 86, 14.8%)	No diabetes (n = 494, 85.2%)
fasting glucose (mean mg/dL (s.d.))	191.5 (84.2)	88.3 (13.9)	191.9 (92.9)	88.5 (13.3)	191.1 (76.5)	88.1 (14.4)
c. 2010 Samoan sample						
Overall (n = 2,861)			Males (n = 1,149, 40.2%)		Females (n = 1,712, 59.8%)	
	Diabetes (n = 478, 16.7%)	No diabetes (n = 2,383, 83.3%)	Diabetes (n = 185, 16.1%)	No diabetes (n = 964, 83.9%)	Diabetes (n = 293, 17.1%)	No diabetes (n = 1,419, 82.9%)
age (mean years (s.d))	52.5 (8.0)	43.7 (11.2)	53.0 (7.6)	44.3 (11.4)	52.2 (8.3)	43.3 (11.0)
BMI (mean kg/m ² (s.d.))	35.1 (6.3)	33.1 (6.7)	34.0 (6.0)	30.9 (5.8)	34.6 (6.8)	35.9 (6.4)
Overall (n = 2,844)			Males (n = 1,143, 40.2%)		Females (n = 1,701, 59.8%)	

Russell *et al.* CREBRF and type 2 diabetes

	Diabetes (n = 461, 16.1%)	No diabetes (n = 2,383, 83.3%)	Diabetes (n = 179, 15.7%)	No diabetes (n = 964, 84.3%)	Diabetes (n = 282, 16.6%)	No diabetes (n = 1,419, 83.4%)
fasting glucose (mean mg/dL (s.d.))	206.1 (81.9)	88.6 (14.0)	203.5 (84.0)	89.6 (14.4)	207.7 (80.7)	88.0 (13.7)
d. 2006–2013 Aotearoa New Zealand sample						
	Overall (n = 2,124)		Males (n = 1,334, 62.8%)		Females (n = 790, 37.2%)	
	Diabetes (n = 557, 26.2%)	No diabetes (n = 1,567, 73.8%)	Diabetes (n = 335, 25.1%)	No diabetes (n = 999, 74.9%)	Diabetes (n = 222, 28.1%)	No diabetes (n = 568, 71.9%)
age (mean years (s.d.))	57.4 (11.4)	45.8 (15.2)	57.3 (11.5)	46.1 (14.5)	57.4 (11.4)	45.3 (16.2)
BMI (mean kg/m²(s.d.))	36.6 (8.2)	33.8 (7.6)	36.4 (7.5)	34.0 (7.1)	36.9 (9.1)	33.4 (8.3)
fasting glucose (mean mg/dL (s.d.))	*	*	*	*	*	*

*Fasting glucose was not measured in the Aotearoa New Zealand sample

Details of the DNA extraction and genotyping for the Samoan, American Samoan and Aotearoa New Zealand samples have been described elsewhere [2, 3]. Subject relatedness in the 2010 sample was measured using an empirical kinship matrix calculated with OpenMendel using genome-wide genotype data [2, 24, 25]. In the 2002–03 sample, expected kinship was derived from familial pedigrees with OpenMendel [24]. Neither genome-wide genotype data nor pedigree information was available for participants in the 1990–95 sample, so subjects were treated as unrelated in statistical models. In the Aotearoa New Zealand sample, a kinship coefficient matrix was calculated from 40,156 independent autosomal markers using PLINK v1.9 [26].

All participants in these studies gave written informed consent. The research in Samoa and American Samoa was reviewed and approved by the institutional review boards of Miriam Hospital, Providence, RI; Brown University; University of Cincinnati; and

University of Pittsburgh. Research in Samoa was also reviewed and approved by the Health Research Committee of the Samoan Ministry of Health. Research in American Samoa was additionally reviewed and approved by the American Samoa Department of Health institutional review board. Ethical approval for the Aotearoa New Zealand study 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).

Statistical Analysis

We explored the impact of BMI and rs373863828 on type 2 diabetes and fasting glucose in three ways: (1) regression analysis to assess the effect of rs373863828 and natural log transformed BMI (lnBMI) on type 2 diabetes and natural log transformed fasting glucose (lnFG) in the two obesity-stratified groups; (2) regression analysis to assess the effect of lnBMI on type 2 diabetes and lnFG within each of the three rs373863828 genotype groups; and (3) path analysis to model the effects of lnBMI and rs373863828 on type 2 diabetes and lnFG.

Except where noted, all statistical analyses were conducted using R (version 3.6.0), and a significance threshold of $\alpha < 0.05$ was used for all tests [27]. Age and age² were scaled and centered (cAge and cAge²), to separate the effects of the sex \times age and sex \times age² interactions from the main effects. The rs373863828 genotype was coded as an additive genetic effect so that it represents the number of copies of the minor allele, A.

Participants who reported a diabetes diagnosis—and thus were potentially being treated for the disease—were excluded from fasting glucose analyses (Table 1). We did not exclude individuals with fasting glucose ≥ 126 mg/dL from the fasting glucose analysis if they did not report a diagnosis because they were presumed to not be part of any intervention affecting their fasting glucose levels.

Each regression model described below was fit using logistic mixed models for type 2 diabetes status and linear mixed models for fasting glucose (Table S1). We adjusted for kinship as a random effect using the *lme4qtl* package [28].

Unless otherwise noted, we performed each analysis in each of the three Samoan/American Samoan samples and the Aotearoa New Zealand sample and then meta-analyzed the results together. In the regression analyses, the samples were meta-analyzed using standard error–weights based on the effect sizes and heterogeneity was tested in METAL [29]. In the path analyses, we pooled the raw data across the samples.

All regression models below were adjusted for the following fixed-effect covariates: cAge, cAge², sex, sex \times cAge, sex \times cAge². The analyses in the 2002–03 and 1990–95 Samoan/American Samoan sample sets were also adjusted for polity (Samoa or American Samoa) to account for different economic and cultural contexts. For each regression model, we compared the effect sizes and the overlap of confidence intervals, a conservative measure of heterogeneity, to assess the whether the effects were different per stratification [30].

Association of lnBMI and rs373863828 with type 2 diabetes and lnFG stratified by obesity: We stratified the samples by the Polynesian obesity cutoff (32 kg/m²). We examined if the effects of lnBMI and rs373863828 on type 2 diabetes differed for individuals with and without obesity by regressing type 2 diabetes on lnBMI and rs373863828 in the obesity-stratified groups. We also examined whether the effects of lnBMI and rs373863828 on fasting glucose differed by regressing lnFG on lnBMI and rs373863828 in the obesity-stratified groups.

Association of lnBMI with type 2 diabetes and lnFG stratified by rs373863828 genotype: We stratified the samples by rs373863828 genotype (GG, GA, and AA). We examined if the effect of lnBMI on type 2 diabetes differed depending on genotype group by regressing type 2 diabetes on lnBMI in each of the groups. We also examined whether the effect of lnBMI on lnFG differed by regressing lnFG on BMI in each of the genotype groups.

Modeling the effects of lnBMI and rs373863828 on type 2 diabetes and lnFG with path analysis: The relationships among sex, age, rs373863828, lnBMI, and type 2 diabetes were modeled with path analysis using the *lavaan* package in R [31]. Interaction terms were not included in the path model, and therefore variables were not centered. In the path analysis model, we regressed lnBMI on sex, age, and rs373863828, and type 2 diabetes on lnBMI, sex, age, and rs373863828. The indirect effect of rs373863828 on

Russell *et al.* CREBRF and type 2 diabetes

type 2 diabetes as mediated by lnBMI was calculated by multiplying the path coefficient of the direct effect of rs373863828 on lnBMI by the path coefficient of the direct effect of lnBMI on type 2 diabetes. Fasting glucose was modeled using path analysis by substituting lnFG for type 2 diabetes in the above model. Relatedness was not accounted for in these models because the methods for accounting for relatedness do not exist, and this may increase type I error.

In path analysis, the absolute value of an effect size of around 0.1 is commonly considered to be small, around 0.3 is medium, and 0.5 is large [32]. The direct paths of sex to type 2 diabetes and sex to lnFG had effect sizes less than 0.1 and were not part of the direct or indirect path of rs373863828 on type 2 diabetes/lnFG so they were trimmed from the models. For both type 2 diabetes and lnFG, two models were examined: with and without a dummy variable for study (2010 Samoa, 2002–2003 Samoa, 2002–2003 American Samoa, 1990–95 Samoa, 1990–95 American Samoa or 2006–2013 Aotearoa New Zealand) to assess for study sample differences.

Results

Association of lnBMI and rs373863828 with type 2 diabetes and lnFG stratified by obesity:

In the regression of type 2 diabetes on lnBMI and rs373863828, higher lnBMI was associated with higher type 2 diabetes odds in participants without and with obesity (OR = 6.27, $p = 0.003$; and OR = 4.41, $p = 3.68 \times 10^{-6}$, respectively). lnBMI was also associated with lnFG in participants without and with obesity ($\beta = 0.191$ (1.21 mg/dL higher fasting glucose with each higher unit of lnBMI), $p = 4.11 \times 10^{-6}$ and $\beta = 0.093$ (1.10 mg/dL higher fasting glucose), $p = 0.036$, respectively) (Figures 1 and S1 and Table 2). There was evidence of heterogeneity ($p < 0.01$) between studies in the subset with obesity in the type 2 diabetes analysis ($p = 0.082$) and the subset without obesity in the lnFG analysis ($p = 0.063$) (Figure S1 and Table S2).

On the other hand, rs373863828 was associated with lower type 2 diabetes odds in participants, both without and with obesity (OR = 0.58, $p = 9.70 \times 10^{-6}$; and OR = 0.73, $p = 2.82 \times 10^{-5}$, respectively). rs373863828 was also associated with lower lnFG in participants both without and with obesity ($\beta = -0.030$ (-0.97 mg/dL fasting glucose with each copy of the missense allele), $p = 8.17 \times 10^{-5}$ and $\beta = -0.033$ (-0.97 mg/dL fasting glucose), $p = 1.22 \times 10^{-4}$, respectively) (Figure 1 and S1 and Table 2).

Table 2: Effect of lnBMI and rs373863828 on type 2 diabetes and fasting glucose in the meta-analysis, stratified by obesity status (a) and genotype (b)

a.

Meta-analysis lnBMI				Meta-analysis rs373863828		
Dependent variable: type 2 diabetes						
Stratum*	n	OR (95% CI)	<i>p</i>	n	OR (95% CI)	<i>p</i>
without obesity	3,211 (393 cases)	6.27 (1.90, 20.75)	0.003	3,211	0.58 (0.45, 0.74)	9.70 × 10^{−6}
with obesity	3,870 (866 cases)	4.41 (2.35, 8.25)	3.68 × 10^{−6}	3,870	0.73 (0.63, 0.85)	2.82 × 10^{−5}

Meta-analysis lnBMI				Meta-analysis rs373863828		
Dependent variable: natural log transformed fasting glucose						
Stratum*	n	β (s.e.)	p	n	β (s.e.)	p
without obesity	2,180	0.191 (0.041)	4.11 × 10 ⁻⁶	2,180	-0.030 (0.008)	8.17 × 10 ⁻⁵
with obesity	2,352	0.093 (0.044)	0.036	2,352	-0.033 (0.009)	1.22 × 10 ⁻⁴

*without obesity: BMI ≤ 32 kg/m²; with obesity: BMI > 32 kg/m²

b.

Meta-analysis lnBMI			
Dependent variable: type 2 diabetes			
Stratum	n	OR (95% CI)	<i>p</i>
GG	4,111 (824 cases)	6.37 (4.03, 10.05)	1.88×10^{-15}
GA	2,547 (389 cases)	7.049 (2.85, 17.47)	2.46×10^{-5}
AA	232 (22 cases)	1.774 (0.19, 16.93)	0.619
Dependent variable: natural log transformed fasting glucose			
Stratum	n	β (s.e.)	<i>p</i>
GG	2,442	0.245 (0.027)	3.12×10^{-20}
GA	1,759	0.214 (0.034)	1.89×10^{-10}
AA	331	0.161 (0.062)	0.001

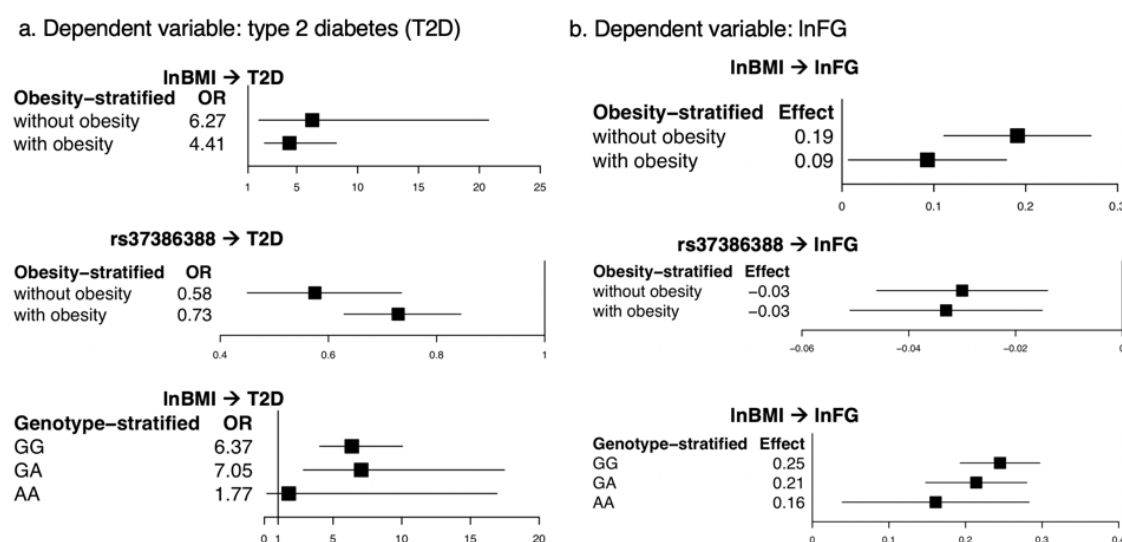
Association of lnBMI with type 2 diabetes and lnFG stratified by rs373863828 genotype:

In the regression of type 2 diabetes on lnBMI in the genotype-stratified groups, higher lnBMI was associated with higher type 2 diabetes odds in both the GG and GA groups (OR = 6.37, $p = 1.88 \times 10^{-15}$; and OR = 7.05, $p = 2.46 \times 10^{-5}$, respectively). lnBMI was not significantly associated with type 2 diabetes in the AA group (OR = 1.774, $p = 0.619$) (Table 2). However, the analysis of the AA group only included participants from the 2010 Samoan sample (type 2 diabetes cases = 22) because the regression models did not

converge with the low numbers of type 2 diabetes cases in the AA group in the 1990–95 and 2002–03 Samoan samples ($n = 2$ and $n = 6$, respectively) and in the Aotearoa New Zealand sample ($n = 16$) (Figure 1 and Table S3). According to a post-hoc power calculation, we had 70.5% power to detect an odds ratio of 1.774 in our AA sample with $\alpha = 0.05$.

In the meta-analysis of lnBMI on lnFG in the genotype-stratified groups, higher lnBMI was associated with higher lnFG in all genotype groups (GG: $\beta = 0.245$ (1.28 mg/dL higher fasting glucose with each unit of lnBMI), $p = 3.12 \times 10^{-20}$; GA: $\beta = 0.214$ (1.24 mg/dL higher fasting glucose), $p = 1.89 \times 10^{-10}$; and AA: $\beta = 0.161$ (1.17 mg/dL higher fasting glucose), $p = 0.001$) (Figure 1 and Table S3). In the lnFG analyses, there was evidence of heterogeneity between studies in the GG group ($p = 0.079$) (Figure S2).

Figure 1: Comparison of the odds ratios and effect sizes in the obesity and genotype-stratified groups.



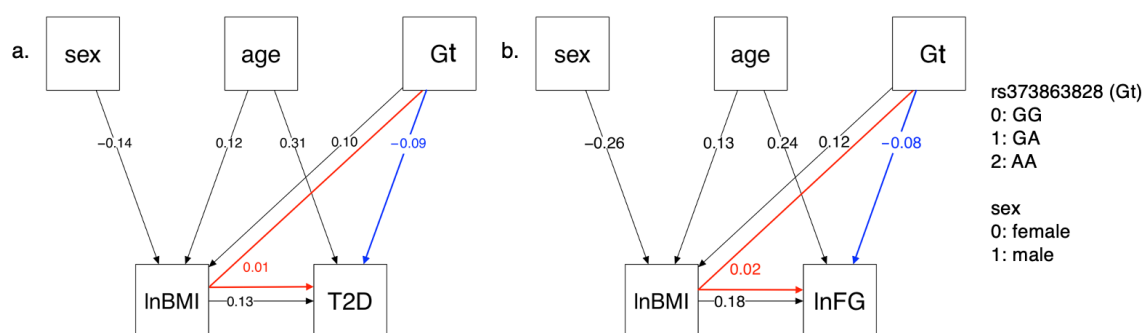
Caption: From the meta-analysis, (a) a comparison of the odds ratios (OR) and 95% confidence intervals in the obesity and genotype-stratified models with type 2 diabetes as

the independent variable and (b) a comparison of the effect sizes and 95% confidence intervals in the obesity and genotype-stratified models with lnFG as the independent variable.

Path analysis models:

The type 2 diabetes path analysis model included five nodes: lnBMI, sex, rs373863828 genotype (subscript Gt in the following path coefficients), age, and type 2 diabetes (subscript T2D). The path pooled analyses were modeled with and without additional nodes to account for study sample (Figure S2). These models were similar so the simpler models are presented below. In the path pooled analysis displayed in Figure 2a, rs373863828 had a small direct effect on lnBMI ($P_{\ln\text{BMI},\text{Gt}} = 0.10, p < 0.001$). rs373863828 and lnBMI had small direct effects on type 2 diabetes ($P_{\text{T2D},\text{Gt}} = -0.09, p < 0.001$; and $P_{\text{T2D},\ln\text{BMI}} = 0.13, p < 0.001$, respectively). rs373863828 also had a small indirect effect on type 2 diabetes as mediated by lnBMI ($P_{\ln\text{BMI},\text{Gt}} \times P_{\text{T2D},\ln\text{BMI}} = 0.01$). The root mean square error of approximation (RMSEA) was 0.011 (90% CI 0.000–0.035) indicating a good model fit [33].

Figure 2: Type 2 Diabetes and lnFG Path Analysis Models



Caption: In the pooled analysis, rs373863828 genotype (Gt) has both direct (blue) and indirect (red) effects on (a) type 2 diabetes (T2D) and (b) lnFG as mediated by lnBMI. All direct paths shown have $p < 0.001$.

The lnFG path analysis model included five nodes: lnBMI, sex, rs373863828, age, and lnFG. In the path pooled analysis (Figure 2b), rs373863828 had a small direct effect on lnBMI ($P_{\ln\text{BMI},\text{Gt}} = 0.12, p < 0.001$). rs373863828 had a small direct effect on lnFG ($P_{\ln\text{FG},\ln\text{BMI}} = 0.18, p < 0.001$). rs373863828 also had a small indirect effect on lnFG that was mediated by lnBMI ($P_{\ln\text{BMI},\text{Gt}} \times P_{\ln\text{FG},\ln\text{BMI}} = 0.02$). RMSEA was 0.068 (90% CI 0.045–0.094) indicating that the path model was a mediocre fit [33].

Discussion

We see here that the discordant association of rs373863828 with BMI and type 2 diabetes is consistent among individuals with and without obesity and across dietary and physical activity exposures and general nutritional environments from Samoa in 1991 [19] to Aotearoa New Zealand in 2006–13 [3, 34]. While there is evidence of heterogeneity between studies in some of the meta-analyses, this is not surprising, giving the secular trends in fasting glucose and type 2 diabetes. These samples were recruited during different decades and from different countries, and there could be underlying differences in them arising from cohort effects, their different dietary and physical-activity

environments and differences in other exposures. This heterogeneity may limit our power to detect associations.

There may be a stronger effect of BMI on fasting glucose levels in Samoans without obesity than in Samoans with obesity, especially in the more recent samples. This is likely attributable to more rapid increases in hyperglycemia and insulin resistance as adiposity increases from normal to overweight to obesity, relative to later in the pathophysiology process [35]. While higher BMI is associated with higher fasting glucose across all rs373863828 genotypes, the magnitude of effect is less with each copy of the missense allele. However, we do not see a similar pattern in the effect of BMI on type 2 diabetes. Carrying the A allele does not decouple higher BMI from higher odds of type 2 diabetes. Individuals who are GG or GA or AA all have higher odds of type 2 diabetes with higher BMI, but individuals who are AA have a lower risk of type 2 diabetes than those who are GG at any given BMI.

In our exploration of the relationships among BMI, fasting glucose, and genotype via path analysis, we found that the rs373863828 missense variant has both a *direct* protective effect on type 2 diabetes and fasting glucose and an *indirect* opposing risk-increasing effect on type 2 diabetes odds and fasting glucose. This indirect effect is mediated through the direct increasing effect of the rs373863828 missense variant on BMI. Again, the rs373863828 missense allele provides some protection from type 2

diabetes, but higher BMI is associated with higher type 2 diabetes odds regardless of rs373863828 genotype.

Little is known about the biology behind the effects of *CREBRF* on type 2 diabetes and BMI. There is evidence that it has multiple, tissue-specific function. The rs373863828 minor allele is associated with greater bone and lean mass in Samoan infants at four months old, but it is not associated with BMI [36], suggesting that body composition may be involved in the association between rs373863828, type 2 diabetes, and BMI. Another function is indicated by the analog of *CREBRF* in *Drosophila* which is involved in energy homeostasis [37], and along with one of its binding partners, CREBL2, it functions as a metabolic regulator linking nutrient sensor mTORC1 to cellular metabolic response [38]. Finally, there is evidence of a role in adipose differentiation.

Overexpression of *CREBRF* in mouse 3T3-L1 preadipocytes induces the expression of adipogenic markers and results in increased lipid accumulation, and overexpression of *CREBRF*:p.R457Q promotes greater lipid storage while using less energy than wild-type *CREBRF* [2]. Our path analysis with its disparate effects adds to this body of evidence that *CREBRF*'s role in biology is multifarious.

While rs373863828 has disparate effects on BMI and type 2 diabetes, BMI is still an important and readily available clinical predictor of type 2 diabetes risk [39], independent of rs373863828. Currently, given the high cost of genotyping and ease of measuring BMI across contexts, there is limited utility to including genetic factors in risk prediction models for type 2 diabetes [1, 40]. While this work is unlikely to impact clinical practice

in the immediate future, exploring the effect of rs373863828 on BMI and type 2 diabetes adds to current knowledge of the effects of *CREBRF* on BMI and type 2 diabetes.

In summary, we provide evidence that the rs373863828 minor allele has both a direct negative effect and an indirect positive effect on type 2 diabetes and fasting glucose, respectively. We also suggest that there may be a stronger positive association between BMI and type 2 diabetes and fasting glucose in Polynesians without obesity than in Polynesians with obesity. Finally, the A allele of rs373863828 is associated with lower odds of type 2 diabetes, but no matter the genotype, higher BMI is still associated with higher odds of type 2 diabetes.

Data Availability

The 2010 sample data from Samoa examined in this study are available from dbGaP (accession number: phs000914.v1.p1). The 1991–95 and 2002–03 sample data from Samoa and American Samoa is not available because participants did not consent to data sharing when enrolled. Sample data from Aotearoa New Zealand is not publicly available owing to consent restrictions but can be requested from author TRM under an appropriate arrangement.

Acknowledgements

We are grateful for the contributions of Melania Selu and Vaimoana Lupematisila to the recruitment of the 2010 Samoan sample. The authors thank the Samoan participants and local village authorities and research assistants over the years. We acknowledge the support of our research collaboration with the Samoa Ministry of Health, Samoa Bureau

Russell *et al.* *CREBRF* and type 2 diabetes

of Statistics, Samoan Ministry of Women, Community and Social Development and the American Samoa Department of Health.

The authors acknowledge the contributions of Nicola Dalbeth, Janak de Zoysa, Rinki Murphy and Lisa Stamp to the recruiting of the Aoteaora New Zealand sample, Jennie Harré Hindmarsh to the recruiting of the Ngati Porou Hauora (NPHCT) sample included within the Aoteaora New Zealand sample, and the contributions of Ruth Topless, Amanda Phipps-Green, Marilyn Merriman and Murray Cadzow for laboratory work and management of the New Zealand datasets.

Funding

Research in Samoa and American Samoa was funded by the National Institute of Health (R01-HL093093, R01-HL133040, R01-AG09375, R01-HL52611, R01-DK55406, P30-ES006096, and R01-DK59642). Research in New Zealand was funded by the Health Research Council of New Zealand (08/075, 10/548, 11/1075, 14/527).

Duality of interest

The authors do not have a conflict of interest to declare.

Contribution statement

EMR and MK performed the analyses with guidance from RLM, DEW, JCC, and IM. EMR, JCC, and MK wrote the manuscript. NLH led the field work data collection in Samoa and phenotype analyses with guidance from STM. GS and HC performed the genotyping of the Samoan samples with guidance from RD. MSR, SV, and JT facilitated

Russell *et al.* *CREBRF* and type 2 diabetes

fieldwork in Samoa and American Samoa. TN contributed to the discussion of the public health implications of the findings. TRM contributed to the data collection in Aotearoa New Zealand. TJM created the kinship matrix for the Aotearoa New Zealand sample. All authors contributed to this work, discussed the results, and critically reviewed and revised the manuscript. RLM is the guarantor of this work.

References

1. Grarup N, Sandholt CH, Hansen T, Pedersen O (2014) Genetic susceptibility to type 2 diabetes and obesity: From genome-wide association studies to rare variants and beyond. *Diabetologia* 57:1528–1541
2. Minster RL, Hawley NL, Su C-T, et al (2016) A thrifty variant in *CREBRF* strongly influences body mass index in Samoans. *Nat Genet* 48(9):1049–1054. <https://doi.org/10.1038/ng.3620>
3. 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 Māori and Pacific (Polynesian) people living in Aotearoa/New Zealand. *Diabetologia* 61(7):1603–1613. <https://doi.org/10.1007/s00125-018-4623-1>
4. Hanson RL, Safabakhsh S, Curtis JM, et al (2019) Association of *CREBRF* variants with obesity and diabetes in Pacific Islanders from Guam and Saipan. *Diabetologia* 1–6. <https://doi.org/10.1007/s00125-019-4932-z>
5. 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(9):847–849. <https://doi.org/10.1038/jhg.2017.44>
6. Lin M, Caberto C, Wan P, et al (2019) Population specific reference panels are crucial for the genetic analyses of Native Hawai‘ians: an example of the *CREBRF* locus. *bioRxiv* 789073. <https://doi.org/10.1101/789073>
7. Yaghootkar H, Scott RA, White CC, et al (2014) Genetic evidence for a normal-weight “metabolically obese” phenotype linking insulin resistance, hypertension, coronary artery

- disease, and type 2 diabetes. *Diabetes* 63(12):4369–4377. <https://doi.org/10.2337/db14-0318>
8. Yaghootkar H, Lotta LA, Tyrrell J, et al (2016) Genetic Evidence for a Link Between Favorable Adiposity and Lower Risk of Type 2 Diabetes, Hypertension, and Heart Disease. *Diabetes* 65(8):2448–2460. <https://doi.org/10.2337/db15-1671>
 9. Fernández-Rhodes L, Howard AG, Graff M, et al (2018) Complex patterns of direct and indirect association between the transcription Factor-7 like 2 gene, body mass index and type 2 diabetes diagnosis in adulthood in the Hispanic Community Health Study/Study of Latinos. *BMC Obes* 5(1). <https://doi.org/10.1186/s40608-018-0200-x>
 10. Taliun D, Harris DN, Kessler MD, et al (2019) Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *bioRxiv* 563866. <https://doi.org/10.1101/563866>
 11. Karczewski KJ, Francioli LC, Tiao G, et al (2019) Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv* 531210. <https://doi.org/10.1101/531210>
 12. Hawley NL, McGarvey ST (2015) Obesity and Diabetes in Pacific Islanders: the Current Burden and the Need for Urgent Action. *Curr Diab Rep* 15(5):29. <https://doi.org/10.1007/s11892-015-0594-5>
 13. Lin S, Naseri T, Linhart C, et al (2017) Trends in diabetes and obesity in Samoa over 35 years, 1978-2013. *Diabet Med* 34(5):654–661. <https://doi.org/10.1111/dme.13197>
 14. Zhou B, Lu Y, Hajifathalian K, et al (2016) Worldwide trends in diabetes since 1980: A pooled analysis of 751 population-based studies with 4.4 million participants. *Lancet* 387(10027):1513–1530. [https://doi.org/10.1016/S0140-6736\(16\)00618-8](https://doi.org/10.1016/S0140-6736(16)00618-8)
 15. Hawley NL, Minster RL, Weeks DE, et al (2014) Prevalence of adiposity and associated

- cardiometabolic risk factors in the samoan genome-wide association study. *Am J Hum Biol* 26(4):491–501. <https://doi.org/10.1002/ajhb.22553>
16. Dai F, Keighley ED, Sun G, et al (2007) Genome-wide scan for adiposity-related phenotypes in adults from American Samoa. *Int J Obes* 31(12):1832–1842. <https://doi.org/10.1038/sj.ijo.0803675>
 17. Dai F, Sun G, Åberg K, et al (2008) A whole genome linkage scan identifies multiple chromosomal regions influencing adiposity-related traits among Samoans. *Ann Hum Genet* 72(6):780–792. <https://doi.org/10.1111/j.1469-1809.2008.00462.x>
 18. McGarvey ST, Levinson PD, Bausser-Man L, Galanis DJ, Hornick CA (1993) Population change in adult obesity and blood lipids in American Samoa from 1976-1978 to 1990. *Am J Hum Biol* 5(1):17–30. <https://doi.org/10.1002/ajhb.1310050106>
 19. Galanis DJ, McGarvey ST, Quested C, et al (1999) Dietary intake of modernizing Samoans: Implications for risk of cardiovascular disease. *J Am Diet Assoc* 99(2):184–190. [https://doi.org/10.1016/S0002-8223\(99\)00044-9](https://doi.org/10.1016/S0002-8223(99)00044-9)
 20. Chin-Hong P V, McGarvey ST (1996) Lifestyle incongruity and adult blood pressure in Western Samoa. *Psychosom Med* 58(2):131–7
 21. Wrigley R, Phipps-Green AJ, Topless RK, et al (2020) Pleiotropic effect of the ABCG2 gene in gout: involvement in serum urate levels and progression from hyperuricemia to gout. *Arthritis Res Ther* 22(1):45. <https://doi.org/10.1186/s13075-020-2136-z>
 22. Swinburn BA, Ley SJ, Carmichael HE, Plank LD (1999) Body size and composition in Polynesians. *Int J Obes Relat Metab Disord* 23(11):1178–83
 23. Wen JYJ, Rush EC, Plank LD (2010) Assessment of obesity in New Zealand Chinese: a comparative analysis of adults aged 30-39 years from five ethnic groups. *N Z Med J*

123(1327):87–98

24. Lange K, Papp JC, Sinsheimer JS, Sripracha R, Zhou H, Sobel EM (2013) Mendel: The Swiss army knife of genetic analysis programs. *Bioinformatics* 29(12):1568–1570.
<https://doi.org/10.1093/bioinformatics/btt187>
25. Zhou H, Sinsheimer JS, Bates DM, et al (2019) OpenMendel: a cooperative programming project for statistical genetics. *Hum Genet.* <https://doi.org/10.1007/s00439-019-02001-z>
26. Purcell S, Neale B, Todd-Brown K, et al (2007) PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet* 81(3):559–575.
<https://doi.org/10.1086/519795>
27. R Core Team (2017) R: A language and environment for statistical computing.
28. Ziyatdinov A, Vázquez-Santiago M, Brunel H, Martinez-Perez A, Aschard H, Soria JM (2018) lme4qtl: linear mixed models with flexible covariance structure for genetic studies of related individuals. *BMC Bioinformatics* 19(1):68. <https://doi.org/10.1186/s12859-018-2057-x>
29. Willer CJ, Li Y, Abecasis GR (2010) METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinforma Appl NOTE* 26(17):2190–219110.
<https://doi.org/10.1093/bioinformatics/btq340>
30. Schenker N, Gentleman JF (2001) On Judging the Significance of Differences by Examining the Overlap Between Confidence Intervals. *Am Stat* 55(3):182–186.
<https://doi.org/10.1198/000313001317097960>
31. Rosseel Y (2012) lavaan: An R Package for Structural Equation Modeling. *J Stat Softw* 48(2):1–36. <https://doi.org/10.18637/jss.v048.i02>
32. Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd ed. L. Erlbaum

Associates, Hillsdale N.J.

33. Schreiber JB, Stage FK, King J, Nora A, Barlow EA (2006) Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review
34. Phipps-Green AJ, Merriman ME, Topless R, et al (2016) Twenty-eight loci that influence serum urate levels: analysis of association with gout. *Ann Rheum Dis* 75(1):124–130.
<https://doi.org/10.1136/annrheumdis-2014-205877>
35. Kahn SE, Hull RL, Utzschneider KM (2006) Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature* 444(7121):840–846.
<https://doi.org/10.1038/nature05482>
36. Arslanian KJ, Fidow UT, Atanoa T, et al (2020) A missense variant in *CREBRF*, rs373863828, is associated with fat-free mass, not fat mass in Samoan infants. *Int J Obes*.
<https://doi.org/10.1038/s41366-020-00659-4>
37. Tiebe M, Lutz M, De La Garza A, Buechling T, Boutros M, Teleman AA (2015) REPTOR and REPTOR-BP Regulate Organismal Metabolism and Transcription Downstream of TORC1. *Dev Cell* 33(3):272–284.
<https://doi.org/10.1016/j.devcel.2015.03.013>
38. Tiebe M, Lutz M, Senyilmaz Tiebe D, Teleman AA (2019) *Crebl2* regulates cell metabolism in muscle and liver cells. *Sci Rep* 9(1):19869. <https://doi.org/10.1038/s41598-019-56407-w>
39. Tiner J, Hawley NL, Weeks DE, et al (2019) The utility of combining *CREBRF* genotype with BMI to screen Samoans for type 2 diabetes (Abstract/Program 1914/W). American Society of Human Genetics, Houston, Texas
40. Ingelsson E, McCarthy MI (2018) Human Genetics of Obesity and Type 2 Diabetes

Russell *et al.* *CREBRF* and type 2 diabetes

Mellitus: Past, Present, and Future. *Circ. Genomic Precis. Med.* 11:e002090