HHS Public Access

Author manuscript

Genet Epidemiol. Author manuscript; available in PMC 2024 February 01.

Published in final edited form as:

Genet Epidemiol. 2023 February ; 47(1): 105–118. doi:10.1002/gepi.22508.

Multivariate analysis of a missense variant in *CREBRF* reveals associations with measures of adiposity in people of Polynesian ancestries

Jerry Z. Zhang[1,*], **Lacey W. Heinsberg**[2,*,^], **Mohanraj Krishnan**[2], **Nicola L. Hawley**[3], **Tanya J. Major**[4], **Jenna C. Carlson**[1,2], **Jennie Harré Hindmarsh**[5], **Huti Watson**[5], **Muhammad Qasim**[5], **Lisa K. Stamp**[6], **Nicola Dalbeth**[7], **Rinki Murphy**[7,8], **Guangyun Sun**[9], **Hong Cheng**[9], **Take Naseri**[10], **Muagututi'a S. Reupena**[11], **Erin E. Kershaw**[12], **Ranjan Deka**[9], **Stephen T. McGarvey**[13,14], **Ryan L. Minster**[2], **Tony R. Merriman**[4,8,15], **Daniel E. Weeks**[1,2,^]

[1]Department of Biostatistics, School of Public Health, University of Pittsburgh, Pittsburgh, PA

[2]Department of Human Genetics, School of Public Health, University of Pittsburgh, Pittsburgh, PA

[3]Department of Chronic Disease Epidemiology, Yale University School of Public Health, New Haven, CT

[4]Department of Biochemistry, University of Otago, Dunedin, New Zealand

[5]Ngāti Porou Hauora Charitable Trust, Te Puia Springs, Tairāwhiti, New Zealand

[^]**Corresponding Authors:** Lacey W. Heinsberg, PhD, RN, Postdoctoral Associate, Department of Human Genetics, School of Public Health, University of Pittsburgh, Public Health 3102A, 130 De Soto Street, Pittsburgh, PA 15261 USA, law145@pitt.edu, Daniel E. Weeks, PhD, Professor of Human Genetics and Biostatistics, Department of Human Genetics, University of Pittsburgh, Public Health 3119, 130 DeSoto Street, Pittsburgh, PA 15261 USA, Phone: (412) 624-5388, weeks@pitt.edu.

[*]Equal first author contributions

Author Contributions: DEW and JZZ proposed the analysis plan. JZZ developed the analysis code, analyzed the data, and JZZ and LWH wrote the paper with guidance from DEW, as well as from JCC and RLM. MK and LWH assisted in the analyses of the data. LWH cleaned, annotated, and refined the code for public sharing on GitHub. TRM, ND, LKS, TJM, JHH, and RM contributed to collection of the Aotearoa New Zealand data. GS, HC, and RD generated the genetic marker data for the Samoa cohort. MK and TJM generated the genetic marker data for the Aotearoa New Zealand cohort. STM is the principal investigator of the Samoa cohort parent study and with DEW, was responsible for acquisition of the study funding as well as, with NLH, TN, and MSR, overseeing the 2010 Samoa data collection. All authors contributed to data interpretation, critically revised the manuscript, and approved the final version. All authors agree to be accountable for all aspects of the work.

Conflict of Interest Statement: Nicola Dalbeth has received consulting fees from AstraZeneca, Dyve Biosciences, Horizon, Selecta, Arthrosi, JW Pharmaceutical Corporation, PK Med, PTC Therapeutics, and Protalix, outside of the submitted work. Tanya J Major has received speaker fees from Education in Nutrition, outside of the submitted work. The remaining authors declare that they have no conflicts of interest. The funders had no role in the design of the study; collection, analyses, or interpretation of data; writing of the manuscript; or decision to publish the results.

Ethical Statement: Research protocols and the informed consent were approved both by the Brown University Institutional Review Board (IRB) and the Health Research Committee of the Samoan Ministry of Health. Data analysis activities at the University of Pittsburgh were reviewed by their IRB and were determined to be exempt (IRB #CR20050288) based on their receipt of only deidentified data. The study was also approved by the New Zealand 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 research was conducted with understanding and written consent of research participants.

Code Availability Statement: To more fully document the details of our analyses, annotated analysis code, along with an example synthetic data set mirroring the statistical properties of the Samoan data set, is available in the GitHub repository https://github.com/lwheinsberg/mvCREBRF.

Related Abstracts: An early version of this work was presented at the American Society of Human Genetics meeting: Zhang JZ, Carlson JC, Hawley NL, Sun G, Cheng H, Naseri T, Reupena MS, Deka R, McGarvey ST, Minster RL, Weeks DE. (2019). A multivariate Bayesian genetic association analysis of a CREBRF variant and adiposity-related phenotypes. Paper presented at the 69th Annual Meeting of The American Society of Human Genetics. Houston, Texas.

[6]Department of Medicine, University of Otago, Christchurch, New Zealand

[7]Department of Medicine, University of Auckland, Auckland, New Zealand

[8]Maurice Wilkins Centre for Molecular Biodiscovery, Auckland, New Zealand

[9]Department of Environmental and Public Health Sciences, College of Medicine, University of Cincinnati, Cincinnati, OH

[10]Ministry of Health, Government of Samoa, Apia, Samoa

[11]Lutia i Puava ae Mapu i Fagalele, Apia, Samoa

[12]Division of Endocrinology and Metabolism, Department of Medicine, University of Pittsburgh, Pittsburgh, PA

[13]International Health Institute, Department of Epidemiology, School of Public Health, Brown University, Providence, RI

[14]Department of Anthropology, Brown University, Providence, RI

[15]Division of Clinical Immunology and Rheumatology, University of Alabama Birmingham, Birmingham, AL

Abstract

The minor allele of rs373863828, a missense variant in CREB3 Regulatory Factor, is associated with several cardiometabolic phenotypes in Polynesian peoples. To better understand the variant, we tested the association of rs373863828 with a panel of correlated phenotypes (body mass index [BMI], weight, height, HDL cholesterol, triglycerides, and total cholesterol) using multivariate Bayesian association and network analyses in a Samoa cohort (n=1,632), Aotearoa New Zealand cohort (n=1,419), and combined cohort (n=2,976). An expanded set of phenotypes (adding estimated fat and fat-free mass, abdominal circumference, hip circumference, and abdominal-hip ratio) was tested in the Samoa cohort (n=1,496). In the Samoa cohort, we observed significant associations ($\log_{10}$ Bayes Factor≥5.0) between rs373863828 and the overall phenotype panel (8.81), weight (8.30) and BMI (6.42). In the Aotearoa New Zealand cohort, we observed suggestive associations ($1.5<\log_{10}BF<5$) between rs373863828 and the overall phenotype panel (4.60), weight (3.27), and BMI (1.80). In the combined cohort, we observed concordant signals with larger $\log_{10}$BFs. In the Samoa-specific expanded phenotype analyses, we also observed significant associations between rs373863828 and fat mass (5.65), abdominal circumference (5.34), and hip circumference (5.09). Bayesian networks provided evidence for a direct association of rs373863828 with weight and indirect associations with height and BMI.

Keywords

Bayesian association analyses; Bayesian network analyses; Pacific Islander; Māori; Samoan

1. INTRODUCTION

A missense variant in the CREB3 Regulatory Factor (*CREBRF*) gene, rs373863828, has been associated with higher body mass index (BMI) in Samoan adults, with each copy

of the minor allele associated with a 1.4 kg/m^2 increase in BMI (Minster et al., 2016). This association has been replicated in a Saipanese and Guamanian (Micronesian) cohort, a Tongan cohort, a Māori and Pacific Island (Polynesian) cohort from Aotearoa New Zealand, and a Native Hawaiian cohort (Hanson et al., 2019; Krishnan et al., 2018; Lin et al., 2020; Naka et al., 2017). The rs373863828 variant has also been associated with greater average body fat percentage, abdominal circumference, hip circumference, and height (Carlson et al., 2020; Lin et al., 2020; Metcalfe et al., 2020; Minster et al., 2016) as well as favorable lipid profiles (Minster et al., 2016; Ohashi et al., 2018) and a decreased odds of diabetes (Minster et al., 2016). Little is understood about these complex relationships. While this allele is exceedingly rare in populations of non-Pacific ancestry (Genome Aggregation Database minor allele frequency [MAF] < 3×10^{-4}), it is common in Samoan and other Polynesian populations (MAF = 0.10 to 0.26) (Gudmundsson et al., 2021; Karczewski et al., 2020, 2021). Given the disproportionate burden of global obesity and related cardiovascular/inflammatory diseases (e.g., dyslipidemia, gout, chronic kidney disease) observed in Polynesian groups, an improved understanding of rs373863828 is still needed.

Of note, anthropometric and lipid phenotypes are highly correlated, making them problematic for statistical analyses, particularly in genetic association studies. Specifically, the web of correlation structures both between phenotypes and between genotypes and phenotypes can be difficult to untangle through marginal analyses (i.e., statistical modeling of a single phenotype/outcome) alone. Even more problematic for rs373863828, specifically, are the paradoxical associations observed. As highlighted above, the minor allele is simultaneously a risk factor for obesity but protective against diabetes and unfavorable lipid profiles (Minster et al., 2016). Traditional marginal analyses considering a single phenotype simply cannot support a full understanding of the phenotypic complexity of rs373863828.

In contrast, multivariate Bayesian approaches can provide more information about the statistical dependencies within a *complete disease system* as it considers many phenotypes simultaneously through a more realistic complex “web” of relationships (Lewis & Ward, 2013). By considering the correlation structure between variables, multivariate approaches focus on data-driven structure discovery and can provide novel insights into both previously established and unknown relationships, including the likelihood of direct and indirect effects amongst variables of interest (Lewis & Ward, 2013). This information provides great value for understanding biology and disease processes, which can lead to the design of better disease control and prevention programs in future public health translation.

As such, the purpose of this study was to characterize the associations of rs373863828 with a panel of correlated anthropometric phenotypes and a set of lipid profile measurements through a powerful Bayesian multivariate framework in cohorts of individuals of Polynesian descent from Samoa and Aotearoa New Zealand.

2. MATERIALS AND METHODS

2.1 Design

This was a secondary analysis of cross-sectional, observational data collected from individuals of Polynesian descent from Samoa and Aotearoa New Zealand.

2.2 Cohorts

The Samoa cohort consisted of 3,102 individuals of Samoan ancestry recruited in 2010 from 33 villages located across the two main islands of the Independent State of Samoa – 'Upolu and Savai'i. Details of sample selection, genome-wide genotype and phenotype data collection, and data quality checking methods have been previously reported (N. L. Hawley et al., 2014; Minster et al., 2016). The overall objective of the original study was to understand the genetic architecture and behavioral/environmental moderators of adiposity and related phenotypes among adults in Samoa (N. L. Hawley et al., 2014) which led to the discovery of rs373863828 (Minster et al., 2016).

For this secondary data analysis, a maximally unrelated set of 1,829 individuals was selected from the Samoan cohort using PRIMUS software (Staples et al., 2014) based on a second-cousin kinship threshold to remove any potential confounding effects of kinship on the analysis. An additional 169 individuals were removed due to missingness within their phenotype panels (as the statistical approaches taken in this study require complete data for all phenotypes of interest) and 28 individuals were removed following data screening for normality and outliers (described below). The final sample size for primary multivariate analyses was 1,632 individuals (Table S1). Following identical filtering steps, the final sample size for the expanded phenotype analysis (described below) was 1,496 participants (Table S1).

The Aotearoa New Zealand cohort consisted of 2,335 individuals of Polynesian ancestry relating to nine Island Nation groups: New Zealand Māori, Cook Island Māori, Samoan, Tongan, Pukapukan, Niuean, Tahitian, Tokelauan, and Tuvaluan. This cohort was recruited across the North and South Islands of Aotearoa New Zealand, including 270 (predominantly Māori) participants recruited from the Tairāwhiti region (East Coast, North Island) in collaboration with the Ngāti Porou Hauora Charitable Trust. Details of sample selection, genome-wide genotype and phenotype data collection, and data quality checking have been previously described (Krishnan et al., 2018). Participants were recruited to better understand the genetics of gout, diabetes, and kidney disease (Krishnan et al., 2018; Phipps-Green et al., 2016).

Following removal of 395 individuals due to missingness within their phenotype panels, 490 individuals due to relatedness (using the same second-cousin kinship threshold approach as above), and 31 individuals due to data screening for normality and outliers (described below), a subset of 1,419 individuals remained for multivariate analysis in the Aotearoa New Zealand cohort (Table S1).

2.3 Phenotypes of Interest

A common set of six anthropometric and lipid profile phenotypes composed of BMI (kg/m^2), height (m), weight (kg), HDL cholesterol (HDL-C, mg/dL), triglycerides (TG, mg/dL), and total cholesterol (Chol., mg/dL) were used within both the Samoa cohort and the Aotearoa New Zealand cohort. Cohorts were first analyzed separately, followed by a mega analysis of the combined Samoa and Aotearoa New Zealand data using the common subset of traits.

LDL cholesterol data estimated with the Friedewald formula (Friedewald et al., 1972) were also available. However, this marker was ultimately excluded from the analysis because (1) Friedewald-derived LDL cholesterol is linearly dependent on measured HDL, TG, and total cholesterol values and (2) the Friedewald formula is invalid for high levels of TG (Martin et al., 2013). Therefore, inferring LDL cholesterol in this manner would have resulted in a substantial sample size reduction.

While hypolipidemic medication use was not collected for the Samoa cohort, individuals who self-reported use of heart disease medication (n=17) were excluded based on prior sensitivity analyses which revealed significant associations with cholesterol levels (Minster et al., 2016). In contrast, for the Aotearoa New Zealand cohort, heart disease medication information was not collected, but statin and diuretic medication usage was collected. To assess the impact of medication usage on the multivariate approach, we performed a sensitivity analysis using the subset of individuals from the Aotearoa New Zealand cohort who were not on statin and diuretic medication.

In addition to the phenotypes above, additional anthropometric phenotypes – fat mass (FM, kg), fat-free mass (FFM, kg), abdominal circumference (Abd C, cm), hip circumference (Hip C, cm), and abdominal-hip ratio (AHR), were available only in the Samoa cohort and were used in an expanded Samoa cohort analysis. Fat mass and fat-free mass were estimated from age, sex, height, weight, and bioelectrical impedance resistance using equations derived from direct body composition studies of Polynesian peoples living in Aotearoa New Zealand using duel-energy X-ray absorptiometry (DXA) scans as described elsewhere (Keighley et al., 2006; Swinburn et al., 1999). Of note, fat mass estimated in this matter has nearly perfect correlation with DXA-derived fat mass (r^2=0.97) in a subset of individuals from the Samoa cohort (n=425), demonstrating the validity of these equations in our sample (Heinsberg et al., Submitted).

2.4 Statistical Analysis

We used R statistical software as the framework for data preparation, analysis routine calling, and result reporting (Team, 2018). Within both cohorts, we adjusted phenotypes for age and sex (validated by genotyping) using ordinary linear regression models. In the Aotearoa New Zealand cohort, we also adjusted phenotypes using each individual's first four principal components derived from genome-wide genotype data to correct for potential confounding effects of population structure between different Polynesian population sub-groups (Krishnan et al., 2018). Principal components adjustment was not necessary in the Samoa cohort as all participants self-reported having four Samoan grandparents (N.

L. Hawley et al., 2014); homogeneous Samoan ancestry was analytically confirmed via principal components analysis of genome-wide single nucleotide polymorphism data (Minster et al., 2016).

2.4.1 mvBIMBAM—The association of rs373863828 with the panel of phenotypes was performed using the Bayesian multivariate *mvBIMBAM* framework (Stephens, 2013, 2019). In this framework, a global null model representing no association between phenotypes and genotype is compared with an exhaustive combination of alternative models, in which all different combinations of phenotypes are associated with the genotype. For the alternative models, the methodology splits phenotypes into all possible partitions of *U*, *D*, and *I*, each representing 'unassociated', 'directly', and 'indirectly' associated. Both directly and indirectly associated phenotypes are associated with genotype, but indirectly associated phenotypes are conditionally independent of the genotype given the presence of a directly associated phenotype in the model. The evidence against the null hypothesis is the sum of Bayes factors (BF) ($\log_{10}$ scale) of all partitions weighted by a diffuse prior (Shim et al., 2015; Stephens, 2013, 2019). Strong evidence of association is defined as $\log_{10}$ BF > 5; suggestive evidence is defined as $1.5 < \log_{10}$ BF < 5; and negligible evidence is defined as $\log_{10}$ BF < 1.5. Marginal posterior probabilities of association (MPPA) are calculated by summing the marginal posterior probabilities of direct and indirect association.

The sensitivity of the Bayesian multivariate *mvBIMBAM* framework to outlier values and non-normality necessitated the normalization of phenotypes (Stephens, 2013, 2019). Residualized phenotypes were order quantile-normalized using the *OrderNorm* function from the R package *bestNormalize* (Peterson & Cavanaugh, 2019). We removed observations in violation of multivariate normality at an $\alpha = 0.01$ level based on Mahalanobis distance-based test statistics following a χ^2_d null distribution corresponding to a d-dimensional multivariate phenotype panel.

For the mega analysis, we combined the cohort-specific residuals and quantile normalized them jointly. As justified above, the Samoa cohort was adjusted for age and sex, and the Aotearoa New Zealand cohort was adjusted for age, sex, and the first four principal components estimated based on genome-wide genotype data. The residualized values were then quantile normalized jointly once again. This normalization procedure allowed for the preservation of relative ranks of observations across the two cohorts while ensuring multivariate normality.

We calculated 95% confidence intervals for BFs and association category posterior probabilities (i.e., those for unassociated, direct association, indirect association category) via bootstrapping. This was done by sampling with replacement and repeating the *mvBIMBAM analyses* 1,000 times. We reported final multi-model point estimates as the median of the bootstrap distribution, as well as the 95% confidence intervals.

In a *post hoc* exploratory analysis, we performed sex-specific *mvBIMBAM* analyses.

2.4.2 Bayesian Network Analyses—We further explored the relationships between rs373863828 and the phenotypes in the quantile normalized datasets with Bayesian networks

learned with the R package *bnlearn* (Nagarajan et al., 2014; Scutari, 2010). A constrained learning algorithm based on conditional independence testing (semi-interleaved HITON-PC) based on Scutari et al (Scutari, 2015) was used to infer association and structure within the network (Aliferis et al., 2010; Scutari, 2015, 2017). Phenotypes and the rs373863828 were modeled as nodes with edges representing associations between nodes. We restricted the analyses so that rs373863828 could have only outgoing edges connecting to phenotypes.

The strength and directionalities of the edges of the Bayesian networks were inferred through a bootstrapped process resulting in networks that varied slightly between runs. As such, representative networks were plotted, but the quantitative strength (E_s) and direction (E_d) of each edge that summarized results across the total number of bootstrapped realizations was labeled on each plot. Edge strength is a measure of confidence of that edge while fixing the rest of the network structure and is defined as the empirical frequency a specific edge is observed over a set of networks learned from bootstrapped samples (i.e., the number of times the edge was present out of the total number of bootstrapped realizations). An edge was included in the network graph if its strength was larger than a significance threshold learned from the bootstrapped samples. Edge direction represents the probability of the edge's direction conditional on the edge's presence within the network (i.e., the number of times the edge traveled in a specific direction out of the total number of bootstrapped realizations in which it was present in either direction). While Markov networks are undirected and may be cyclic, Bayesian networks are directed/acyclic and are designed to support causal interpretation via a set of potentially unverifiable assumptions (Briganti et al., 2022; Nagarajan et al., 2014). Despite this, arrow direction depicts statistical dependencies, not molecular dependencies. Paired with the cross-sectional observational nature of the data, arc direction in this study should be interpreted with caution (Lewis & Ward, 2013).

For complete analytical details, refer to documented example analysis code in the GitHub repository https://github.com/lwheinsberg/mvCREBRF.

3. RESULTS

3.1 mvBIMBAM

Using the Bayesian multivariate *mvBIMBAM* framework, we found that in the Samoa cohort the association with rs373863828, while taking the multivariate correlation structures between phenotypes (Figures S1–S2) into account, was strong for the overall phenotype panel, along with the individual phenotypes weight and BMI ($\log_{10}$ BF overall = 8.81, weight = 8.30, BMI = 6.42, Table 1). There was suggestive evidence for an association with height ($\log_{10}$ BF = 2.02, Table 1) and negligible evidence for an association with HDL-C, TG, and total cholesterol ($\log_{10}$ BF < 1.5, Table 1). The evidence of association was weaker in the Aotearoa New Zealand cohort with suggestive associations between rs373863828 and the overall phenotype panel, weight, and BMI ($\log_{10}$ BF overall = 4.60, weight = 3.27, BMI = 1.80, Table 1), and negligible evidence for an association with height, HDL-C, TG, and total cholesterol ($\log_{10}$ BF < 1.5, Table 1).

When we performed a sensitivity analysis using the subset of Aotearoa New Zealand individuals who were not on statin or diuretic medication, a similar pattern of association between rs373863828 and the phenotype panel was observed, with the strongest effects noted with weight and BMI ($\log_{10}$ BF weight = 1.46, BMI = 0.90, Table S2), although all Bayes factors were below the cut-off for suggestive association ($\log_{10}$ BF < 1.5).

Within the mega analysis of the combined cohort, the evidence of association of rs373863828 with the overall phenotype panel, weight, and BMI was even greater than among the Samoa or Aotearoa New Zealand cohorts alone ($\log_{10}$ BF overall = 12.37, weight = 11.86, BMI = 8.06, Table 1). The association with height remained suggestive ($\log_{10}$ BF = 3.16), and evidence for all other associations remained negligible ($\log_{10}$ BF < 1.5, Table 1).

Within the expanded anthropometric panel in the Samoa cohort, there was strong evidence for an association between rs373863828 and the overall phenotype panel, weight, BMI, fat mass, abdominal circumference, and hip circumference ($\log_{10}$ BF overall = 8.05, weight = 7.26, BMI = 5.75, fat mass = 5.65, abdominal circumference = 5.34, hip circumference = 5.09, Table 2). There was suggestive evidence for an association with height ($\log_{10}$ BF = 1.58) and fat-free mass ($\log_{10}$ BF = 1.85) and negligible evidence for associations with abdominal-hip ratio, HDL-C, TG, or total cholesterol ($\log_{10}$ BF < 1.5, Table 2). Sex-specific *mvBIMBAM* results are presented in Tables S3–S4.

3.2 Bayesian Network Analyses

Bayesian networks were trained for the Samoa cohort (Figure 1), the Aotearoa New Zealand cohort (Figure 2), the combined cohort (Figure S3), and the expanded phenotype panel analysis in the Samoa cohort (Figure 3). As stated above, the strength and directionalities of the edges of the Bayesian networks were inferred through a bootstrapped process resulting in networks that varied slightly from run to run, thus representative networks are presented with quantitative edge strength:direction values that summarize results across the total number of bootstrapped realizations. In this approach, edges with high strength and strong directionality are more likely to appear in any single realization of a network, while edges with weaker directionalities may change direction across different realizations.

The networks from both cohorts and the combined cohort suggested a direct association between rs373863828 and weight, indirect associations between rs373863828 and both height and BMI through weight, and a web of associations between lipid and other anthropometric traits (Figures 1–2, Figure S3). These results aligned with the evidence from *mvBIMBAM,* which suggested the association between rs373863828 and weight was more likely to be direct than indirect (probability 90% direct, 10% indirect in the combined cohort, Table 1). The inferred Bayesian networks were also consistent with *mvBIMBAM* results in the Samoa cohort, which suggested greater likelihood of indirect associations between rs373863828 and both height (probability 42% direct, 56% indirect, Table 1) and BMI (probability 46% direct, 54% indirect, Table 1). This was not the case in the Aotearoa New Zealand cohort, however, which resulted in more likely direct than indirect effects between rs373863828 and height (probability 62% direct, 29% indirect, Table 1) and BMI (probability 65% direct, 29% indirect, Table 1). Of note, the edge metrics for many of the relationships across all networks presented were 1:1 (i.e., consistent presence/direction in

100% of bootstrapped realizations), indicating very high statistical confidence in the strength and directions of effect (Figures 1–3, Figure S3).

In the Bayesian network analysis of the Samoa cohort with the expanded phenotype panel, an edge between rs373863828 and total cholesterol was also inferred with an edge strength of 0.68 (present in 68% of the bootstrapped realizations, Figure 3). The *mvBIMBAM* results suggested a similar probability of direct association with total cholesterol (84%), though the evidence of association was negligible ($\log_{10}$ BF = 0.98, Table 2). This rs373863828-cholesterol edge was not observed in Bayesian network analyses of the two cohorts with the reduced phenotype panel (Figures 1–2) nor in the analysis of the combined cohort (Figure S3).

4. DISCUSSION

When correlation structures were considered, multivariate Bayesian analyses provided strong evidence of the pleiotropic effects of rs373863828 including associations with weight, BMI, fat mass, abdominal circumference, and hip circumference (Tables 1, 2). Of note, rs373863828 was first discovered in a genome-wide association study of BMI (Minster et al., 2016). Most variants that associate with increased BMI are typically associated with increased weight, but not height. Because rs373863828 is associated with increased weight (Minster et al., 2016), height (Carlson et al., 2020), and BMI (Minster et al., 2016) in marginal analyses, the Bayesian approach taken here is uniquely positioned to offer probabilities of direct vs. indirect effects of rs373863828.

To that end, the strongest and most persistent direct association presented here was observed between rs373863828 and weight using both *mvBIMBAM* ($\log_{10}$ BF = 3.3 to 11.9, Tables 1–2) and *bnlearn* (Figures 1–3, Figure S3). Specifically, in the combined cohort there was strong evidence ($\log_{10}$ BF = 11.86) for an 90.25% probability of a direct effect of rs373863828 on weight. These results aligned with the relationships learned from the Bayesian networks in which 100% of bootstrapped realizations found a direct association between rs373863828 and weight across all cohorts.

The signal for association between rs373863828 and height was more varied with suggestive evidence of association in the Samoa and combined cohorts ($\log_{10}$ BF = 1.58 to 3.16, Tables 1–2) but negligible evidence of association in the Aotearoa New Zealand cohort ($\log_{10}$ BF = 0.87). In contrast to weight, in the combined cohort, the association between rs373863828 and height was equally as likely to be direct or indirect in *mvBIMBAM* results, while it was clearly favored as an indirect relationship in *bnlearn* results (i.e., 100% of bootstrapped realizations suggested an indirect association between variant and height, through weight). Considered together, these results suggest that, despite the variant's trait-specific marginal associations with both weight and height, any difference in BMI due to rs373863828 genotype is most attributable to the variant's effect on weight.

In the expanded phenotype panel analyses in the Samoa cohort, we observed strong evidence of association between rs373863828 and hip and abdominal circumference phenotypes ($\log_{10}$ BF abdominal circumference = 5.34, hip circumference = 5.09, Table 2), both

of which are highly correlated with weight. Further, we observed stronger evidence of association between rs373863828 and estimated fat mass compared to fat-free mass ($\log_{10}$ BF of 5.65 vs 1.85, respectively, Table 2). Upon exploring this finding further in the sex-stratified analyses (Table S4), we observed that this association had strong support only in females, which is consistent with prior observations (N. Hawley et al., In Press). While the sex-stratified analyses should be interpreted with caution due to the reduced sample sizes, other potentially interesting sex-specific results included stronger associations between the variant and weight, BMI, abdominal circumference, and hip circumference in females (particularly in the Samoan cohort) and stronger associations between the variant and height in Samoan males (Tables S3–S4).

Despite the many strengths of this study, including the large sample size, complementary Bayesian analytical approaches, and focus on a historically understudied group, there are some limitations that should be acknowledged. Given that the analytical approach applied here only accepts quantitative, normally distributed traits, we were unable to include some important phenotypes that have been associated with rs373863828, including type 2 diabetes diagnosis (Minster et al., 2016). While we considered analyzing a related phenotype such as fasting glucose, the addition was complicated by a substantial reduction in sample size based on the exclusion of participants by diabetes diagnosis (Russell et al., 2022) and heterogeneity of medication use and adherence (LaMonica et al., 2022). Therefore, as with all studies, there were unmeasured/unaccounted for factors which could impact the conclusions (e.g., arc direction, direct vs. indirect determinations) (Stephens, 2013, 2019). To reduce this burden, and provide more complete/causal conclusions, future directions of this work include the application of multivariate methods that allow other variable types (e.g., binary, categorical, or nonparametric) (Hackinger & Zeggini, 2017). Relatedly, we examined total cholesterol, which included both detrimental LDL cholesterol and favorable HDL cholesterol. While it would be ideal to examine these values separately, the use of LDL cholesterol would have reduced our sample size substantially as described in the methods. Given the association between the rs373863828 minor allele and decreased fasting glucose (Russell et al., 2022), enhanced early insulin release (Burden et al., 2021), and favorable lipid profiles (Minster et al., 2016; Ohashi et al., 2018), it will be important to examine these more analytically complex traits in the future.

In conclusion, we used two complementary Bayesian approaches to explore how rs373863828 was associated with a panel of correlated adiposity-related phenotypes when considered in a multivariate context. There was strong evidence for an association between rs373863828 and weight, BMI, fat mass, abdominal circumference, and hip circumference (Tables 1, 2), providing additional evidence of rs373863828's pleiotropic effects. Most notably, both multivariate approaches highlighted a strong direct effect only on weight (Table 1, Figures 1–3, Figure S3), suggesting that rs373863828-specific interventions focused directly on reducing body weight might be most effective at improving other measures of adiposity and lipid profiles. In future work, it will be interesting to further examine these patterns in the context of more precise measurements of body composition and energy metabolism, as well as glucose levels, insulin release/sensitivity, and detailed lipid panels.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We would like to thank the participants from both Samoa and Aotearoa New Zealand and all the field organizers who have made this research possible. In addition, we would like to thank the Samoan Ministry of Health, the Samoa Bureau of Statistics, and the Ministry of Women, Community, and Social Development for their continued support. We would specifically like to thank Vaimoana Lupematisila and Melania Selu for their continued engagement with the Samoan cohort. Gabrielle Sexton, Roddi Laurence, Fiona Taylor, Jordyn de Kwant, Grace Muyoma, Jill Drake, and Chris Franklin are thanked for their engagement and recruitment of Aotearoa New Zealand participants and Ria Akuhata, Nancy Aupouri and Carol Ford of the Ngāti Porou Hauora Charitable Trust for their engagement and recruitment of participants in Tairāwhiti. Finally, thank you to the anonymous peer reviewers who provided excellent feedback on this paper as it improved the quality and clarity of our work.

Funding:

This work was funded by National Institute of Health (grants R01HL093093, R01HL133040, TL1TR001858, and K99HD107030) and the Health Research Council of New Zealand (grants 08/075, 10/548, 11/1075, 14/527).

Data Availability Statement:

Sample data from the Samoa cohort data are available from dbGAP, accession number: phs000914.v1.p1. Sample data from the Aotearoa New Zealand cohort are not publicly available owing to consent restrictions, but can be requested from author TRM (https://orcid.org/0000-0003-0844-8726) under an appropriate arrangement.

References

Aliferis C, Statnikov A, Tsamardinos I, Mani S, & Koutsoukos X (2010). Local Causal and Markov Blanket Induction for Causal Discovery and Feature Selection for Classification Part I: Algorithms and Empirical Evaluati. Journal of Machine Learning Research, 11(7), 171–234.

Briganti G, Scutari M, & McNally RJ (2022). A tutorial on bayesian networks for psychopathology researchers. Psychological Methods. 10.1037/met0000479

Burden HJ, Adams S, Kulatea B, Wright-McNaughton M, Sword D, Ormsbee JJ, Watene-O'Sullivan C, Merriman TR, Knopp JL, Chase JG, Krebs JD, Hall RM, Plank LD, Murphy R, Shepherd PR, & Merry TL (2021). The CREBRF diabetes-protective rs373863828-A allele is associated with enhanced early insulin release in men of Māori and Pacific ancestry. Diabetologia, 64(12), 2779–2789. 10.1007/s00125-021-05552-x [PubMed: 34417843]

Carlson JC, Rosenthal SL, Russell EM, Hawley NL, Sun G, Cheng H, Naseri T, Reupena MS, Tuitele J, Deka R, McGarvey ST, Weeks DE, & Minster RL (2020). A missense variant in CREBRF is associated with taller stature in Samoans. American Journal of Human Biology, 32(6), e23414. 10.1002/ajhb.23414 [PubMed: 32190945]

Friedewald WT, Levy RI, & Fredrickson DS (1972). Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clinical Chemistry, 18(6), 499–502. [PubMed: 4337382]

Gudmundsson S, Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, Solomonson M, Watts NA, Rhodes D, Singer-Berk M, England EM, Seaby EG, … MacArthur DG (2021). Addendum: The mutational constraint spectrum quantified from variation in 141,456 humans. Nature, 597(7874), E3–E4. 10.1038/s41586-021-03758-y [PubMed: 34373650]

Hackinger S, & Zeggini E (2017). Statistical methods to detect pleiotropy in human complex traits. Open Biology, 7(11). 10.1098/rsob.170125

Hanson RL, Safabakhsh S, Curtis JM, Hsueh W-C, Jones LI, Aflague TF, Duenas Sarmiento J, Kumar S, Blackburn NB, Curran JE, Mahkee D, Baier LJ, Knowler WC, & Nelson RG (2019). Association of CREBRF variants with obesity and diabetes in Pacific Islanders from Guam and Saipan. Diabetologia, 62(9), 1647–1652. 10.1007/s00125-019-4932-z [PubMed: 31280340]

Hawley N, Duckham R, Carlson J, Naseri R, Reupena M, Lameko V, Pomer A, Wetzel A, Selu M, Lupematisila V, Unasa F, Vesi L, Fatu R, Unasa S, Faasalele-Savusa K, Rivara A, Russell E, Delany J, Viali S, … McGarvey S (n.d.). The protective effect of rs373863828 on type 2 diabetes does not operate through a body composition pathway in adult Samoans. Obesity, In press.

Hawley NL, Minster RL, Weeks DE, Viali S, Reupena MS, Sun G, Cheng H, Deka R, & Mcgarvey ST (2014). Prevalence of adiposity and associated cardiometabolic risk factors in the samoan genome-wide association study. American Journal of Human Biology, 26(4), 491–501. 10.1002/ajhb.22553 [PubMed: 24799123]

Heinsberg L, Hawley N, Duckham R, Pomer A, Rivara A, Naseri T, Reupena M, Weeks D, McGarvey S, & Minster R (2022). Validity of anthropometric equation-based estimators of fat mass in Samoan adults. American Journal of Human Biology, Submitted.

Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, Solomonson M, Watts NA, Rhodes D, Singer-Berk M, England EM, Seaby EG, Kosmicki JA, … MacArthur DG (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. Nature, 581(7809), 434–443. 10.1038/s41586-020-2308-7 [PubMed: 32461654]

Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, Solomonson M, Watts NA, Rhodes D, Singer-Berk M, England EM, Seaby EG, Kosmicki JA, … MacArthur DG (2021). Author Correction: The mutational constraint spectrum quantified from variation in 141,456 humans. In Nature (Vol. 590, Issue 7846, p. E53). 10.1038/s41586-020-03174-8 [PubMed: 33536625]

Keighley ED, McGarvey ST, Turituri P, & Viali S (2006). Farming and adiposity in Samoan adults. American Journal of Human Biology : The Official Journal of the Human Biology Council, 18(1), 112–122. 10.1002/ajhb.20469 [PubMed: 16378333]

Krishnan M, Major TJ, Topless RK, Dewes O, Yu L, Thompson JMD, McCowan L, de Zoysa J, Stamp LK, Dalbeth N, Harré Hindmarsh J, Rapana N, Deka R, Eng WWH, Weeks DE, Minster RL, McGarvey ST, Viali S, Naseri T, … Merriman TR (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. 10.1007/s00125-018-4623-1 [PubMed: 29721634]

LaMonica LC, McGarvey ST, Rivara AC, Sweetman CA, Naseri T, Reupena MS, Kadiamada H, Kocher E, Rojas-Carroll A, DeLany JP, & Hawley NL (2022). Cascades of diabetes and hypertension care in Samoa: Identifying gaps in the diagnosis, treatment, and control continuum - a cross-sectional study. The Lancet Regional Health. Western Pacific, 18, 100313. 10.1016/j.lanwpc.2021.100313 [PubMed: 35024652]

Lewis FI, & Ward MP (2013). Improving epidemiologic data analyses through multivariate regression modelling. Emerging Themes in Epidemiology, 10(1), 4. 10.1186/1742-7622-10-4 [PubMed: 23683753]

Lin M, Caberto C, Wan P, Li Y, Lum-Jones A, Tiirikainen M, Pooler L, Nakamura B, Sheng X, Porcel J, Lim U, Setiawan VW, Le Marchand L, Wilkens LR, Haiman CA, Cheng I, & Chiang CWK (2020). Population-specific reference panels are crucial for genetic analyses: an example of the CREBRF locus in Native Hawaiians. Human Molecular Genetics, 29(13), 2275–2284. 10.1093/hmg/ddaa083 [PubMed: 32491157]

Martin SS, Blaha MJ, Elshazly MB, Brinton EA, Toth PP, McEvoy JW, Joshi PH, Kulkarni KR, Mize PD, Kwiterovich PO, DeFilippis AP, Blumenthal RS, & Jones SR (2013). Friedewald-Estimated Versus Directly Measured Low-Density Lipoprotein Cholesterol and Treatment Implications. Journal of the American College of Cardiology, 62(8), 732–739. 10.1016/j.jacc.2013.01.079 [PubMed: 23524048]

Metcalfe LK, Krishnan M, Turner N, Yaghootkar H, Merry TL, Dewes O, Hindmarsh JH, De Zoysa J, Dalbeth N, Stamp LK, Merriman TR, Smith G, Shepherd P, & Murphy R (2020). The Māori and

Pacific specific CREBRF variant and adult height. International Journal of Obesity (2005), 44(3), 748–752. 10.1038/s41366-019-0437-6 [PubMed: 31543511]

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 (2016). A thrifty variant in CREBRF strongly influences body mass index in Samoans. Nature Genetics, 48(9), 1049–1054. 10.1038/ng.3620 [PubMed: 27455349]

Nagarajan R, Scutari M, & Lèbre S (2014). Bayesian Networks in R: with Applications in Systems Biology. Springer New York.

Naka I, Furusawa T, Kimura R, Natsuhara K, Yamauchi T, Nakazawa M, Ataka Y, Ishida T, Inaoka T, Matsumura Y, Ohtsuka R, & Ohashi J (2017). A missense variant, rs373863828-A (p.Arg457Gln), of CREBRF and body mass index in Oceanic populations. Journal of Human Genetics, 62(9), 847–849. 10.1038/jhg.2017.44 [PubMed: 28405013]

Ohashi J, Naka I, Furusawa T, Kimura R, Natsuhara K, Yamauchi T, Nakazawa M, Ishida T, Inaoka T, Matsumura Y, & Ohtsuka R (2018). Association study of CREBRF missense variant (rs373863828:G > A; p.Arg457Gln) with levels of serum lipid profile in the Pacific populations. Annals of Human Biology, 45(3), 215–219. 10.1080/03014460.2018.1461928 [PubMed: 29877158]

Peterson R, & Cavanaugh J (2019). Ordered quantile normalization: a semiparametric transformation built for the cross-validation era. Journal of Applied Statistics, 1–16. [PubMed: 35707602]

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 (2016). Twenty-eight loci that influence serum urate levels: analysis of association with gout. Annals of the Rheumatic Diseases, 75(1), 124–130. 10.1136/annrheumdis-2014-205877 [PubMed: 25187157]

Russell EM, Carlson JC, Krishnan M, Hawley NL, Sun G, Cheng H, Naseri T, Reupena MS, Viali S, Tuitele J, Major TJ, Miljkovic I, Merriman TR, Deka R, Weeks DE, McGarvey ST, & Minster RL (2022). CREBRF missense variant rs373863828 has both direct and indirect effects on type 2 diabetes and fasting glucose in Polynesian peoples living in Samoa and Aotearoa New Zealand. BMJ Open Diabetes Research & Care, 10(1). 10.1136/bmjdrc-2021-002275

Scutari M (2010). Learning Bayesian Networks with the bnlearn R Package. Journal of Statistical Software, 35(3 SE-Articles), 1–22. 10.18637/jss.v035.i03 [PubMed: 21603108]

Scutari M (2015). Bayesian Network Constraint-Based Structure Learning Algorithms: Parallel and Optimised Implementations in the bnlearn R Package. ArXiv. https://arxiv.org/abs/1406.7648

Scutari M (2017). Bayesian Network Constraint-Based Structure Learning Algorithms: Parallel and Optimized Implementations in the bnlearn R Package. Journal of Statistical Software, 77(2 SE-Articles), 1–20. 10.18637/jss.v077.i02

Shim H, Chasman DI, Smith JD, Mora S, Ridker PM, Nickerson DA, Krauss RM, & Stephens M (2015). A multivariate genome-wide association analysis of 10 LDL subfractions, and their response to statin treatment, in 1868 Caucasians. PloS One, 10(4), e0120758. 10.1371/journal.pone.0120758 [PubMed: 25898129]

Staples J, Qiao D, Cho MH, Silverman EK, Nickerson DA, & Below JE (2014). PRIMUS: rapid reconstruction of pedigrees from genome-wide estimates of identity by descent. American Journal of Human Genetics, 95(5), 553–564. 10.1016/j.ajhg.2014.10.005 [PubMed: 25439724]

Stephens M (2013). A unified framework for association analysis with multiple related phenotypes. PloS One, 8(7), e65245. 10.1371/journal.pone.0065245 [PubMed: 23861737]

Stephens M (2019). Correction: A Unified Framework for Association Analysis with Multiple Related Phenotypes. In PloS one (Vol. 14, Issue 3, p. e0213951). 10.1371/journal.pone.0213951 [PubMed: 30889189]

Swinburn BA, Ley SJ, Carmichael HE, & Plank LD (1999). Body size and composition in Polynesians. International Journal of Obesity and Related Metabolic Disorders : Journal of the International Association for the Study of Obesity, 23(11), 1178–1183. 10.1038/sj.ijo.0801053 [PubMed: 10578208]

Team RC (2018). R: A language and environment for statistical computing. (3.5.0). R Foundation for Statistical Computing. https://www.r-project.org/

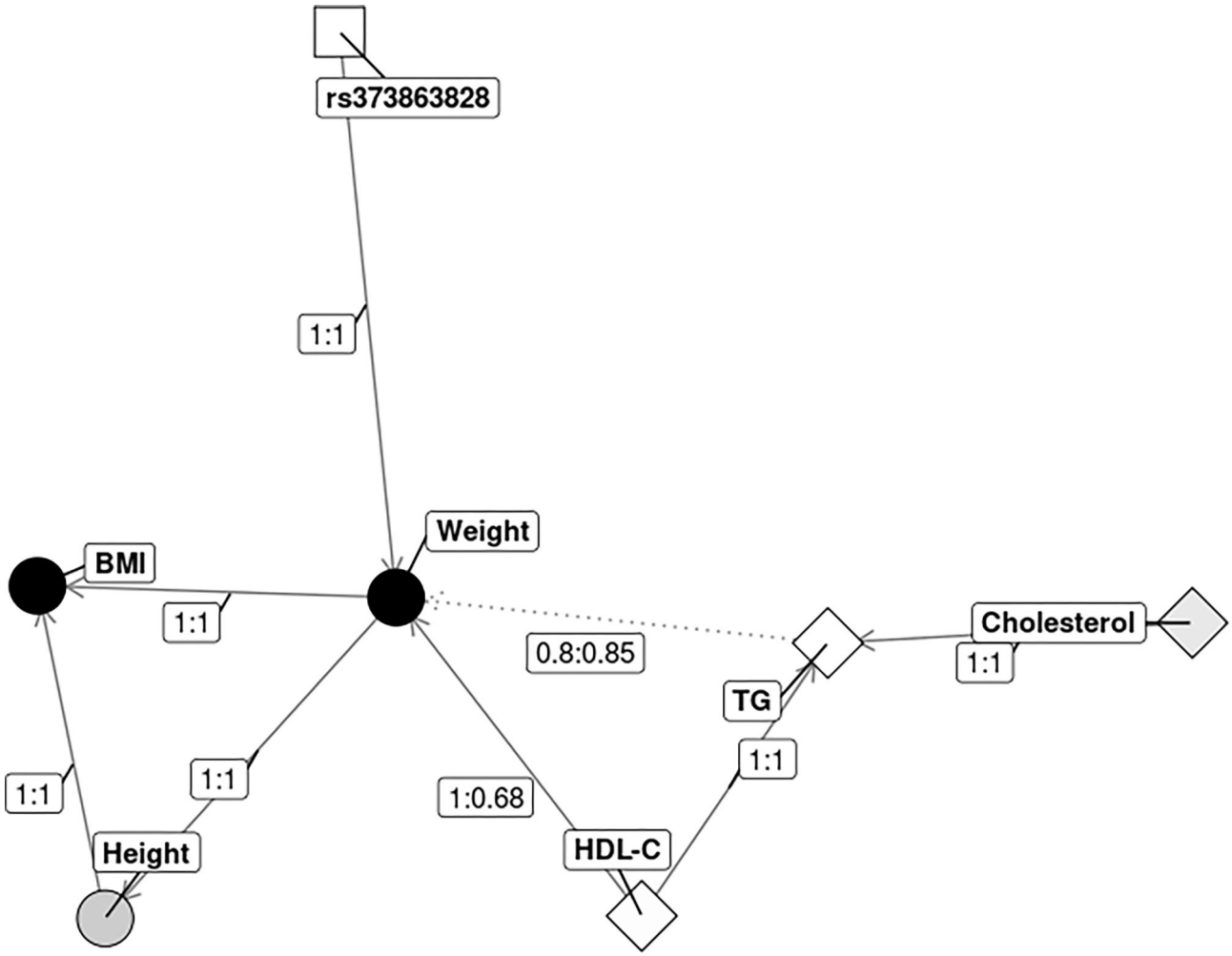

Figure 1.

Samoa Bayesian Network. Nodes are shape coded by type (circle = anthropometric, diamond = lipids, square = the variant). Shading correspond to $\log_{10}$BF 0 (white) to 5 or greater (black). Edge labels represent strength:directionality; edge strength represents the number of times the edge was present out of the total number of bootstrapped realizations (proportion); edge direction represents the number of times the edge traveled in a specific direction out of the total number of bootstrapped realizations in which it was present in either direction (proportion); edges with strength less or equal to 0.90 are dashed. BMI, body mass index; Chol, cholesterol; TG, triglycerides; HDL-C, HDL cholesterol.

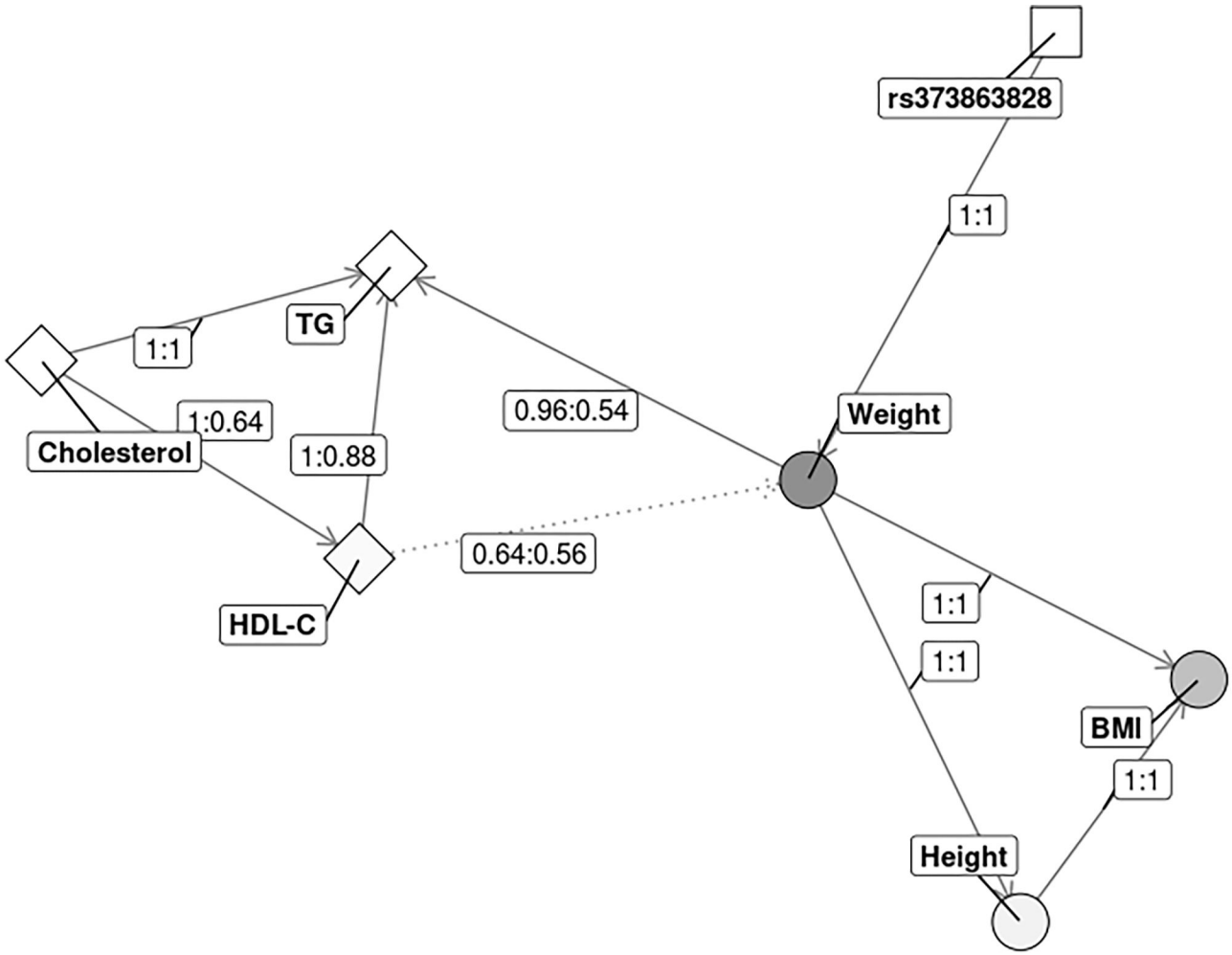

Figure 2.

Aotearoa New Zealand Bayesian Network. Nodes are shape coded by type (circle = anthropometric, diamond = lipids, square = the variant). Shading correspond to $\log_{10}$BF 0 (white) to 5 or greater (black). Edge labels represent strength:directionality; edge strength represents the number of times the edge was present out of the total number of bootstrapped realizations (proportion); edge direction represents the number of times the edge traveled in a specific direction out of the total number of bootstrapped realizations in which it was present in either direction (proportion); edges with strength less or equal to 0.90 are dashed. BMI, body mass index; Chol, cholesterol; TG, triglycerides; HDL-C, HDL cholesterol.

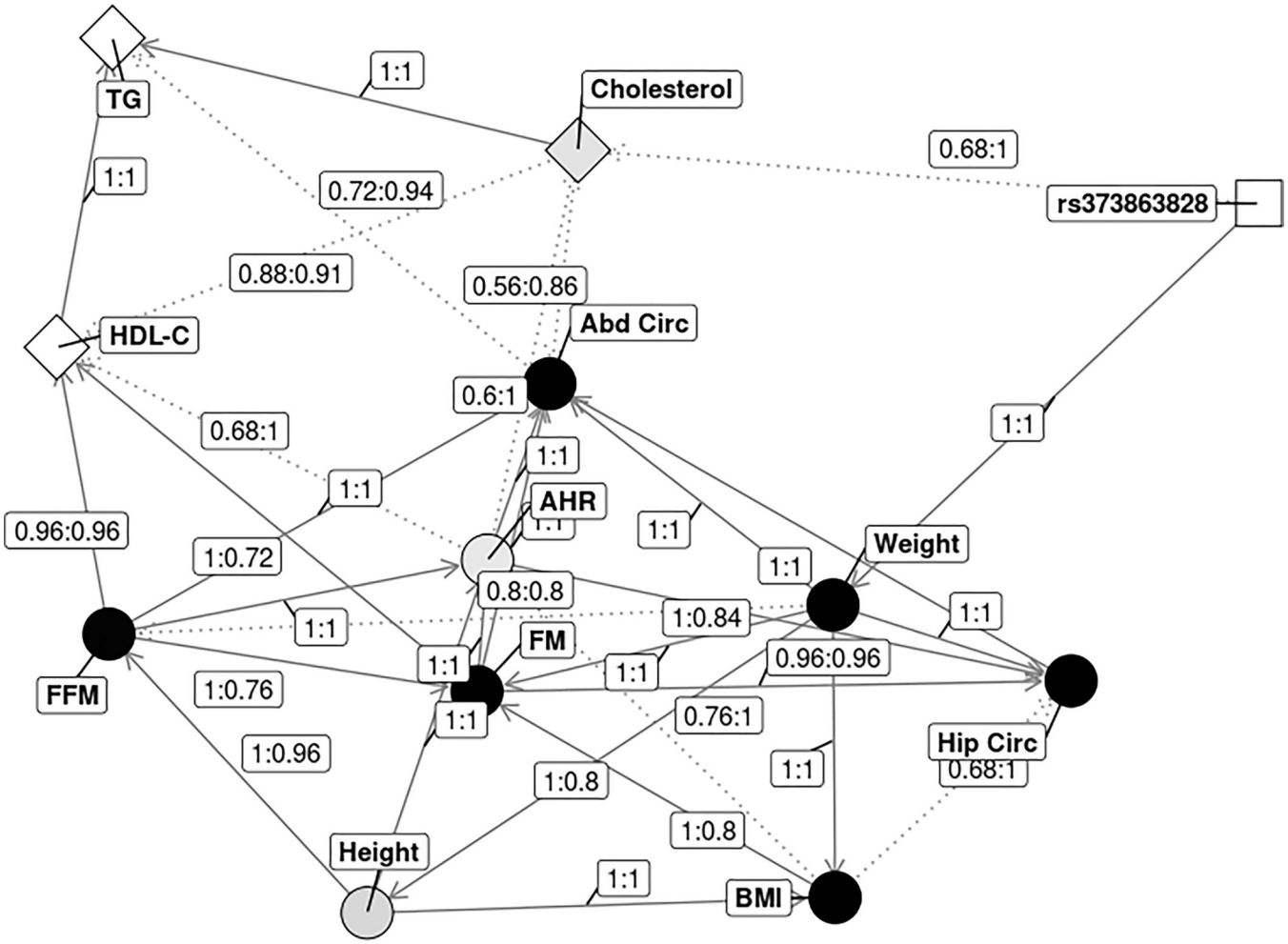

Figure 3.

Samoa cohort Bayesian network with expanded phenotype set. Nodes are shape coded by type (circle = anthropometric, diamond = lipids, square = the variant). Shading correspond to $\log_{10}$BF 0 (white) to 5 or greater (black). Edge labels represent strength:directionality; edge strength represents the number of times the edge was present out of the total number of bootstrapped realizations (proportion); edge direction represents the number of times the edge traveled in a specific direction out of the total number of bootstrapped realizations in which it was present in either direction (proportion); edges with strength less or equal to 0.90 are dashed. BMI, body mass index; Chol, cholesterol; TG, triglycerides; HDL-C, HDL cholesterol; Abd Circ, abdominal circumference; Hip Circ, hip circumference; AHR, abdominal-hip ratio; FM, fat mass; FFM, fat-free mass.

Table 1.

mvBIMBAM $\log_{10}$ Bayes Factors and marginal posterior probabilities of association (MPPA) for all cohorts.

Table 1A. Samoa Cohort, N = 1,632, Overall $\log_{10}$ BF = 8.81 (95% CI 4.16 to 15.28)

	Weight			Height			BMI			HDL-C			TG			Chol.		
	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th
Log_{10} Bayes Factor	**8.30**	3.80	14.36	*2.02*	0.02	5.93	**6.42**	2.61	11.99	0.01	−0.49	1.91	−0.41	−0.50	0.39	0.82	−0.39	3.47
MPPA (%)	100	100	100	100	100	79.47	100	100	100	88.72	99.82	44.28	65.29	94.46	36.41	92.27	100	60.74
Direct (%)	85.27	65.29	96.92	41.54	22.36	82.88	46.47	25.68	84.19	27.10	13.94	70.64	27.14	15.40	81.32	78.82	15.41	99.95
Indirect (%)	14.73	3.08	34.71	55.82	16.40	75.45	53.53	15.75	74.32	55.70	10.98	77.38	32.75	0.88	71.49	12.54	0.03	53.71
Unaffected (%)	0	0	0.02	0.19	0	20.53	0	0	0.09	11.28	0.18	55.72	34.71	5.54	63.59	7.73	0.02	39.26

Table 1B. Aotearoa New Zealand Cohort, N = 1,419, Overall $\log_{10}$ BF = 4.60 (95%CI 1.08 to 10.12)

	Weight			Height			BMI			HDL-C			TG			Chol.		
	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th
Log_{10} Bayes Factor	*3.27*	0.71	7.70	0.87	−0.33	3.57	*1.80*	0.02	5.20	−0.15	−0.43	1.38	−0.26	−0.44	1.05	−0.20	−0.44	1.33
MPPA (%)	100	100	74.00	95.00	100	53.35	99.08	100	61.36	71.87	99.16	43.96	63.32	98.83	44.15	66.18	99.04	41.91
Direct (%)	80.42	54.62	97.88	62.43	24.57	92.74	64.99	37.78	87.54	52.22	21.68	98.45	58.35	25.97	98.82	53.21	21.74	98.79
Indirect (%)	17.54	1.99	32.86	29.45	6.21	51.57	28.74	6.31	53.93	13.03	0.08	49.18	1.76	0	36.98	6.94	0.02	44.35
Unaffected (%)	0.13	0	26.00	5.00	0.01	46.65	0.92	0	38.64	28.13	0.84	56.04	36.68	1.17	55.85	33.82	0.96	58.09

Table 1C. Combined Cohort, N = 2,976, Overall $\log_{10}$ BF = 12.37 (95% CI 6.57 to 20.22)

	Weight			Height			BMI			HDL-C			TG			Chol.		
	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th
Log_{10} Bayes Factor	**11.86**	6.38	19.07	*3.16*	0.57	7.46	**8.06**	3.67	14.70	0.08	−0.58	2.21	−0.45	−0.59	0.84	−0.39	−0.59	0.96
MPPA (%)	100	100	100	100	100	93.90	100	100	100	90.20	100	50.82	44.80	95.46	25.91	68.11	96.05	46.78
Direct (%)	90.25	65.90	98.91	50.75	17.94	85.91	55.51	19.94	90.76	27.87	10.73	96.96	34.41	15.15	94.66	20.00	8.34	67.43
Indirect (%)	9.75	1.09	34.10	49.09	14.06	79.94	44.49	9.24	80.06	53.77	2.53	78.51	2.94	0.01	54.85	48.46	11.11	63.64
Unaffected (%)	0	0	0	0.02	0	6.10	0	0	0.01	9.80	0.11	49.18	55.20	4.54	74.09	31.89	3.95	53.22

Bold log10 Bayes Factors are greater than 5 (strong), italic log10 Bayes Factors are between 1.5 and 5 (suggestive). BF, log10 Bayes Factor; MPPA, marginal posterior probabilities of association; BMI, body mass index; HDL-C, HDL cholesterol; TG, triglycerides; Chol., cholesterol; Est., point estimate for BF, MPPA%, Direct%, Indirect%, or Unaffected% calculated as the median of the bootstrapped resamples; 2.5th and 97.5th percentiles interpreted as limits of 95% confidence intervals calculated via bootstrapping.

Table 2.

mvBIMBAM $\log_{10}$ Bayes Factors and marginal posterior probabilities of association (MPPA) for the Samoa cohort with expanded phenotype set (N = 1,496, Overall log10 BF = 8.05 [95% CI 3.11 to 14.31]).

	Weight			Height			BMI			FFM			FM			AHR		
	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th
$\log_{10}$ Bayes Factors	**7.26**	3.04	12.61	*1.58*	−0.16	4.88	**5.75**	2.07	10.57	*1.85*	−0.09	5.55	**5.65**	2.15	10.37	0.75	−0.38	3.58
MPPA (%)	100	100	99.78	99.76	100	86.10	100	100	99.25	99.84	100	90.69	100	100	99.17	99.02	100	83.17
Direct (%)	73.41	41.15	95.20	37.94	18.45	74.21	42.11	20.27	80.86	44.50	15.24	97.20	45.95	21.17	96.31	33.36	14.09	76.35
Indirect (%)	26.59	4.80	58.70	59.94	23.79	80.71	57.79	19.14	79.73	53.51	2.80	81.99	53.97	3.35	78.83	62.80	21.86	84.67
Unaffected (%)	0	0	0.22	0.24	0	13.90	0	0	0.75	0.16	0	9.31	0	0	0.83	0.99	0	16.83
	Abd C			Hip C			HDL-C			TG			Chol.					
	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th	Est.	2.5th	97.5th			
$\log_{10}$ Bayes Factors	**5.34**	1.67	10.61	**5.09**	1.78	9.98	−0.16	−0.49	1.55	−0.40	−0.49	0.67	0.98	−0.28	3.74			
MPPA (%)	100	100	99.24	100	100	99.04	91.69	99.84	61.54	79.93	96.65	53.71	97.50	100	80.48			
Direct (%)	36.06	15.07	77.93	37.25	15.32	83.89	29.90	13.28	67.31	38.77	14.97	93.17	83.55	21.28	99.95			
Indirect (%)	63.86	22.04	84.93	62.63	16.04	84.68	59.90	10.40	83.26	35.66	0.80	73.92	12.55	0.03	65.04			
Unaffected (%)	0	0	0.76	0	0	0.96	8.31	0.16	38.46	20.07	3.35	46.29	2.50	0	19.52			

Bold $\log_{10}$ Bayes Factors are greater than 5 (strong), italic $\log_{10}$ Bayes Factors are between 1.5 and 5 (suggestive). BF, $\log_{10}$ Bayes Factor; MPPA, marginal posterior probabilities of association; BMI, body mass index; FFM, fat-free mass; FM, fat mass; AHR, abdominal-hip ratio; Abd C, abdominal circumference; Hip C, hip circumference; HDL-C, HDL cholesterol; TG, triglycerides; Chol., cholesterol; Est., point estimate for BF, MPPA%, Direct%, Indirect%, or Unaffected% calculated as the median of the bootstrapped resamples; 2.5th and 97.5th percentiles interpreted as limits of 95% confidence intervals calculated via bootstrapping.