Journal of Human Genetics (2021) 66:111–121
https://doi.org/10.1038/s10038-020-0816-9

ARTICLE

Genome-wide association studies in Samoans give insight into the genetic architecture of fasting serum lipid levels

Jenna C. Carlson[1,2] · Daniel E. Weeks[1,2] · Nicola L. Hawley[3] · Guangyun Sun[4] · Hong Cheng[4] · Take Naseri[5] · Muagututi'a Sefuiva Reupena[6] · John Tuitele[7] · Ranjan Deka[4] · Stephen T. McGarvey[8,9] · Ryan L. Minster[1]

Received: 9 March 2020 / Revised: 16 July 2020 / Accepted: 20 July 2020 / Published online: 5 August 2020

Abstract

The current understanding of the genetic architecture of lipids has largely come from genome-wide association studies (GWAS). To date, few GWAS have examined the genetic architecture of lipids in Polynesians, and none have in Samoans, whose unique population history, including many population bottlenecks, may provide insight into the biological foundations of variation in lipid levels. Here we performed a GWAS of four fasting serum lipid levels: total cholesterol (TC), high-density lipoprotein (HDL), low-density lipoprotein (LDL), and triglycerides (TG) in a sample of 2849 Samoans, with validation genotyping for associations in a replication cohort comprising 1798 Samoans and American Samoans. We identified multiple genome-wide significant associations ($P < 5 \times 10^{-8}$) previously seen in other populations—*APOA1* with TG, *CETP* with HDL, and *APOE* with TC and LDL—and several suggestive associations ($P < 1 \times 10^{-5}$), including an association of variants downstream of *MGAT1* and *RAB21* with HDL. However, we observed different association signals for variants near *APOE* than what has been previously reported in non-Polynesian populations. The association with several known lipid loci combined with the newly identified associations with variants near *MGAT1* and *RAB21* suggest that while some of the genetic architecture of lipids is shared between Samoans and other populations, part of the genetic architecture may be Polynesian-specific.

Introduction

The Samoan Islands, comprising both the U.S. Territory of American Samoa (American Samoa) and the Independent State of Samoa (Samoa), have experienced a rise in prevalence of cardiovascular disease and other non-communicable diseases in the last 30 years partly due to economic modernization, rapid urbanization, and lifestyle changes such as increased caloric intake and sedentary behavior [1–3]. A 2010 population-based survey in Samoa, which gathered the "discovery cohort" studied further here, found that many Samoans are at elevated risk of

These authors jointly supervised this work: Daniel E. Weeks, Ranjan Deka, Stephen T. McGarvey, Ryan L. Minster

Joint Senior Authors: Daniel E. Weeks, Ranjan Deka, Stephen T. McGarvey and Ryan L. Minster

Supplementary information The online version of this article (https://doi.org/10.1038/s10038-020-0816-9) contains supplementary material, which is available to authorized users.

✉ Ryan L. Minster
ryan.minster@pitt.edu

1 Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA

2 Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA

3 Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, CT, USA

4 Department of Environmental Health, College of Medicine, University of Cincinnati, Cincinnati, OH, USA

5 Ministry of Health, Government of Samoa, Apia, Samoa

6 Lutia i Puava ae Mapu i Fagalele, Apia, Samoa

7 Department of Public Health, Government of American Samoa, Pago Pago, AS, USA

8 International Health Institute and Department of Epidemiology, School of Public Health, Brown University, Providence, RI, USA

9 Department of Anthropology, Brown University, Providence, RI, USA

Table 1 Demographic, anthropometric, and blood biochemistry statistics of the genotyped discovery and replication cohorts

	2010 Discovery Cohort				2002–2003 Replication Sample Set				1994–1995 Replication Sample Set				2002–2003 Replication Sample Set				1994–1995 Replication Sample Set			
	Samoa, $n = 2849$				Samoa, $n = 490$				Samoa, $n = 468$				American Samoa, $n = 592$				American Samoa, $n = 248$			
	Women		Men		Women		Men		Women		Men		Women		Men		Women		Men	
	$n = 1703$		$n = 1146$		$n = 245$		$n = 245$		$n = 243$		$n = 225$		$n = 337$		$n = 255$		$n = 137$		$n = 111$	
age (years)	44.8	(11.1)	45.6	(11.1)	44.0	(17.0)	40.7	(16.3)	43.3	(8.6)	43.2	(8.9)	43.0	(16.0)	43.2	(16.6)	43.3	(9.1)	45.5	(10.6)
BMI (kg/m^2)	34.8	(6.7)	31.3	(5.9)	33.2	(7.7)	28.8	(5.4)	31.9	(5.7)	28.8	(5.0)	36.5	(8.4)	33.4	(7.6)	37.4	(6.9)	34.9	(6.0)
total cholesterol (mg/dL)	199.3	(36.1)	200.3	(38.7)	202.3	(35.9)	195.5	(40.6)	199.1	(37.2)	197.4	(35.0)	187.1	(38.6)	189.6	(37.8)	197.8	(31.2)	194.1	(31.0)
HDL (mg/dL)	46.5	(10.8)	43.7	(11.2)	47.1	(10.3)	46.3	(11.2)	43.4	(11.0)	42.5	(11.4)	40.9	(8.8)	40.0	(8.8)	34.8	(7.6)	32.8	(8.3)
LDL (mg/dL)	130.0	(32.6)	129.6	(35.3)	130.3	(32.8)	128.6	(37.5)	140.2	(34.8)	133.5	(33.0)	118.5	(33.7)	118.9	(35.1)	136.6	(32.4)	136.0	(37.5)
triglycerides (mg/dL)	114.9	(80.5)	139.3	(113.1)	110.9	(58.9)	120.1	(91.3)	85.6	(72.8)	102.8	(97.2)	130.8	(78.1)	200.2	(206.5)	118.9	(63.4)	168.2	(109.0)

cardiovascular disease based on known risk factors—increased total cholesterol (TC), low-density lipoprotein cholesterol (LDL), and triglycerides (TG), as well as decreased high-density lipoprotein (HDL) [4–6]—with 47% of the 2938 adult Samoans examined in the study having elevated TC (≥5.2 mmol/L), 88% of men and 91% of women having elevated LDL (>2.59 mmol/L), and 43% of women having low HDL (<1.29 mmol/L) [1].

Our current understanding of the genetic component of serum lipid level variation has been largely due to genome-wide association studies (GWAS) [7]. The Global Lipids Genetics Consortium found strong evidence for 157 loci associated with one or more of these traits using a sample of 188,577 individuals of European, East Asian, South Asian, and African ancestry [8]. However, few GWAS of serum lipid levels have been conducted in Pacific Islanders [9, 10] and to our knowledge only one has included a small number of Polynesians [11]. Previous studies have estimated the heritability of serum lipid levels in Samoans, ranging from 16% for HDL to 23% for TG, and have identified genetic susceptibility loci via linkage analysis [12], warranting further study of the genetic architecture of serum lipid levels in Samoans. Samoans are a genetically-isolated founder population, with unique evolutionary history, making them particularly useful in genomic studies [13, 14]. Thus, genomic studies of serum lipid levels could reveal novel lipid-altering loci specific to Pacific Islander populations, as well as highlight susceptibility loci shared with global populations.

Here we report the results of a GWAS of fasting TC, LDL, HDL, and TG in up to 2,849 individuals from independent Samoa followed by replication in up to 1798 individuals from independent Samoa and American Samoa, as part of ongoing GWAS of cardiometabolic disease and adiposity-related traits in the Samoan Islands [1]. We identified multiple genome-wide significant associations previously seen in other populations—*APOA1* with TG, *CETP* with HDL, and *APOE* with TC and LDL—and several suggestive associations, including an association between variants downstream of *MGAT1* and *RAB21* with HDL.

Methods

Discovery cohort and genotyping

The discovery cohort data are available from dbGaP (accession number: phs000914.v1.p1). The discovery cohort of 2849 individuals is drawn from a population-based sample recruited from Samoa in 2010 (Table 1). The sample selection, data collection methods, and phenotyping, including the laboratory assays for serum lipid and

lipoprotein levels, have been previously reported [1, 13]. Briefly, adults aged 24.5 to <65 years old who reported Samoan ancestry (based on having four Samoan grandparents) were recruited from 33 villages from the islands of 'Upolu and Savai'i of Samoa through a population-based sample strategy [1]. Consenting participants completed interviews targeting lifestyle factors related to cardiometabolic health (health history, socio-economic position, dietary intake, and physical activity) and anthropometric measurements (height, weight, blood pressure, body composition), and gave fasting whole blood samples for biochemical and genetic assays. A description of the prevalence of non-communicable diseases and associated risk factors is provided in Hawley et al. [1]. Serum lipid levels were derived from fasting whole blood samples collected after a minimum 10-h overnight fast. Genotyping was performed using Genome-Wide Human SNP 6.0 arrays (Affymetrix). Extensive quality control was conducted on the basis of a pipeline developed by Laurie et al. [15]. Additional details for sample genotyping and genotype quality control are described in Minster et al. [13]. This study was approved by the institutional review board of Brown University and the Health Research Committee of the Samoa Ministry of Health. All participants gave written informed consent via consent forms in Samoan language. Imputation was not performed in this study because prior experience with this population using extant imputation panels such as the Phase 3 1000 Genomes panel showed that the resulting imputed genotypes did not correlate well with observed genotypes [13]. Moreover, extant imputation panels do not contain any haplotypes from individuals of known Polynesian or Pacific Island descent [16–18]. The extent of genetic similarity between currently available reference panels and these Samoan study samples is unknown. We believe it would be inappropriate to use imputed genotypes based on continental reference panels for these samples as they are drawn from a genetically homogeneous Samoan population. Future reference panels that include Polynesian haplotypes are needed to leverage unobserved genetic information in genetically isolated Polynesian populations like those in this study sample.

Replication cohort and genotyping

The replication cohort of 1798 individuals contains two sample sets recruited from Samoa and American Samoa (Table 1). Although there is substantial economic variation across the two polities, with American Samoans generally having a higher standard of living, the Samoans from both territories form a single socio-cultural unit with frequent exchange of mates; genetically, they represent a single homogenous population [3, 19]. The first sample set, referred to as the 1990–95 replication sample set, contains 716 unrelated individuals derived from a longitudinal study of adiposity and cardiovascular disease risk factors among adults from American Samoa and Samoa (Table 1). Detailed descriptions of the sampling and recruitment have been reported previously [20–22]. Briefly, participants were recruited from 46 villages and worksites in American Samoa in 1990 and 9 villages in Samoa (Western Samoa, at the time of recruitment) in 1991 and followed up 4 years later in 1994 and 1995, respectively. All participants were, at baseline, free of self-reported history of heart disease, hypertension, or diabetes. This study was approved by the institutional review boards of the Miriam Hospital and the American Samoa Department of Health, and the Health Research Committee of the Samoa Ministry of Health. All participants gave written informed consent via consent forms in Samoan language.

The second sample set, referred to as the 2002–2003 replication sample set, contains 1082 individuals from American Samoa and Samoa and was drawn from an extended family-based genetic linkage analysis of cardiometabolic traits (Table 1) [12, 23–26]. Probands and relatives were unselected for obesity or related phenotypes. The recruitment process and criteria used for inclusion in this study have been described in detail previously [24, 26]. This study was approved by the institutional review boards of Brown University and the American Samoa Department of Health and the Health Research Committee of the Samoa Ministry of Health. All participants gave written informed consent via consent forms in Samoan language. Imputation was not performed in these studies for the same rationale as the discovery cohort above, but also because genome-wide marker data was not available for the samples in these studies.

In both replication sample sets, blood samples were collected in the morning after a minimum of 10 h fasting, from which serum lipid levels were derived using assay methods published previously [12, 20]. Genotyping of variants selected for validation in the replication cohort (described below) was performed using custom-designed TaqMan OpenArray Real-Time PCR assays (Applied Biosystems). SNPs that could not be genotyped using OpenArray assays were genotyped individually using TaqMan SNP Genotyping assays (Applied Biosystems). Eight variants could not be genotyped due to technical difficulties.

Statistical analyses

Prior to association analyses, residuals were generated for all four lipid traits. First, traits were transformed to normality with the Box–Cox power transformation; secondly, model selection was performed using step-wise linear regression with initial model covariates previously associated with serum lipid levels: age, age^2, sex, log-

transformed BMI, fasting glucose, smoking status, farming status (as a measure of physical activity), and interactions between age, age^2, and sex. The final TC model adjusted for age, age^2, sex, age × sex, and age^2 × sex; the final LDL and TG models adjusted for age, age^2, sex, and age^2 × sex; the final HDL model adjusted for age and sex.

Preliminary associations were performed, and variants were selected for validation without consideration of hypolipidemic medication use, as it was not measured. However, participants did self-report use of heart disease medication. Sensitivity analysis revealed that this self-reported use of medication to treat heart disease was significantly associated with TC and LDL (results not shown); individuals reporting such medication use ($n = 17$) were excluded from analyses. The prioritization of variants for validation genotyping was updated using these analyses, but only after available resources were fully expended. Unfortunately, not all variants that should have been prioritized for validation genotyping were successfully genotyped. All results presented are those of the corrected analyses, removing the individuals with heart disease medication use.

Additional sensitivity analysis was performed for TG by excluding one outlying observation (i.e., TG >4 standard deviations above mean); results did not change qualitatively, and, since the recorded value was within the range of plausible values for TG, the individual was retained for presented analyses.

Association between lipid residuals and autosomal genotypes of 659,492 SNPs with minor allele frequency ≥0.05 and Hardy-Weinberg Equilibrium test P value $\geq 5 \times 10^{-5}$ was assessed using linear mixed modeling in GenABEL, including previously-derived empirical kinship estimates to adjust for subject relatedness [13, 27]. The association between X chromosome genotypes and the lipid phenotypes were calculated in GenABEL, without adjustment using the empirical kinship estimates. Genomic inflation due to population stratification and cryptic relatedness was assessed by estimating λ_{GC} using the lower 90% of the P value distribution [28]. GWAS P values in the discovery cohort (P_D) were compared with a threshold for genome-wide significance of $P_D < 5 \times 10^{-8}$ and a suggestive association threshold of $P_D < 1 \times 10^{-5}$. Statistical power to detect signals at these thresholds was calculated using the Genetic Power Calculator [29].

Gene-set enrichment analysis with MAGENTA was also performed to identify any biological pathways enriched for discovery association signals [30]. Briefly, gene scores were obtained from the most significant P value among SNPs located within each gene using the association results from each lipid GWAS. Genes scores were adjusted for confounding factors including gene size, number of variants, and linkage disequilibrium-related properties by using stepwise multiple linear regression. The 95th percentile of all gene scores was used as the enrichment cutoff for each trait [31]. Gene-set enrichment P values were obtained for highly ranked gene scores. Gene sets were obtained from Gene Ontology (April 2010), pathway information from the Ingenuity (June 2008) and KEGG (June 2010), and biological processes and molecular function from PANTHER (January 2010).

For each of the lipid traits, the INRICH program [32] was used to test for enrichment of known genes (as constructed from Teslovich et al. [33] and Willer et al. [8]) INRICH tests if more known genes are contained in associated intervals than expected by chance, using permutation based on 1 million replicates to generate experiment-wide empirical P values. For each lipid trait, we defined the associated intervals as 100 kb intervals centered on the most significant SNP within association peaks with $P_D < 1 \times 10^{-4}$.

We selected 21 regions demonstrating at least suggestive association for association validation in the replication cohort. Subsequent analyses—which excluded 17 participants who were taking heart disease medication—revealed additional 10 regions should have selected for validation. These regions were not followed-up because resources were not available for genotyping; six additional regions from the initial 21 failed to genotype and resources were not available to reattempt the genotyping. The variant from each locus with smallest P value across the four lipid scans (defined as 1 Mb windows surrounding the peak SNP) or a proxy SNP in high linkage disequilibrium with the lowest-P value SNP was selected as representative of the locus for replication genotyping.

Statistical association was measured in the 1990–1995 and 2002–2003 replication sample sets independently, and results were combined using meta-analysis (see below). Association analyses for both sample sets were performed using GenABEL [34] in R [35], using the same regression models as in the discovery cohort but additionally adjusting for polity (American Samoa or Samoa); the 2002–2003 sample set was additionally adjusted using expected kinship, as derived from familial pedigree information [36].

Prior to meta-analysis, quality control was performed using EasyQC to check for strand and allele frequency consistency [37]. P value-based meta-analysis using sample sizes as weights was performed using METAL [38] to generate two P values: one for the meta-analysis of the two replication cohorts (P_R) and one for the replication cohorts and discovery sample together (P_{DR}). Resulting meta-analysis signals were evaluated based on genome-wide significance and suggestive thresholds (as described above) and by the contribution of the replication sample to the signal. Effect directions for meta-analysis results of peak SNPs were qualitatively compared with those of previously reported lead SNPs.

For ease of reference, any locus identified here with a corresponding signal within 1 mega base pairs (Mb) in a prior lipid study is referred to by the previously prescribed locus name; [8, 10] for loci not previously associated with lipid traits, the symbol of the gene nearest the peak SNP in the locus or the hyphen-separated symbols of the nearest two genes is used as the locus label.

Results

The demographic, anthropometric, and biochemical characteristics of the 2849 participants composing the discovery cohort for this GWAS of serum lipids levels and the 1798 participants composing the replication cohort are presented in Table 1. A detailed description of the discovery cohort and its trends compared with the historical sample sets making up the replication cohort has been previously reported [1]. Briefly, the average age was similar for all cohorts; average BMI was higher among women compared with men and in American Samoa compared with Samoa; average BMI for men and women in Samoa was higher in more recent studies; average lipid levels are largely similar across cohorts with minor exceptions.

We assessed 659,492 unique genome-wide markers for association with 4 traits—TC, HDL, LDL, and TG—in up to 2849 Samoans in the discovery cohort. Relatedness within the discovery cohort was well controlled using the empirical kinship coefficients; λ_{GC} ranged between 1.03 and 1.07 for the four lipid traits (Figs. S2, S4, S6, and S8 in S1 Appendix). We observed 38 genome-wide suggestive or significant associations across 31 loci from the four GWAS (Fig. 1; Figs. S1, S3, S5, and S7 in S1 Appendix; Tables S1, S7, S10, and S13 in S1 Appendix).

Genome-wide significant association was observed in the discovery cohort between all four traits and markers near *APOE*: TC and rs4420638 ($P_D = 2.67 \times 10^{-16}$, Fig. 2e); HDL and rs4420638 ($P_D = 9.07 \times 10^{-9}$, Fig. 2f); LDL and rs1160985 ($P_D = 2.61 \times 10^{-20}$; Fig. 2g); and TG and rs4420638 ($P_D = 7.44 \times 10^{-10}$, Fig. 2h). In addition, HDL was associated with markers near *CETP* (rs289708, $P_D = 1.19 \times 10^{-11}$), and TG, with *APOA1* (rs6589566, $P_D = 3.98 \times 10^{-18}$, Fig. 2c). Suggestive associations were observed between lipid levels and markers at an additional 28 loci, including the *MGAT1* and *RAB21* loci and HDL (Fig. 2a, d), and *APOA1* with TC (rs3741298, $P_D = 1.63 \times 10^{-7}$, Fig. 2b). We had 80% power at $\alpha = 1 \times 10^{-5}$ and $\alpha = 5 \times 10^{-8}$ to detect SNPs that account for 1.0% and 1.5%, respectively, of the residual variance in a phenotype.

Gene-set enrichment analysis with MAGENTA highlighted, at a <5% false-discovery rate (FDR), several lipid homeostasis pathways and gene ontologies for HDL and TG (Tables S8 and S14 in S1 Appendix). Four gene sets were below the FDR for both HDL and TG: HDL particle remodeling, reverse cholesterol transport, cholesterol efflux, and phospholipid efflux. An additional 12 gene sets were implicated for HDL and three gene sets for TG. The HDL particle remodeling and reverse cholesterol transport gene sets had significant enrichment for TC (Table S3 in S1 Appendix), and a single gene set was implicated with LDL, the amylase pathway (Table S11 in S1 Appendix). All four traits had significant enrichment for known TC, HDL, LDL, and TG loci using the INRICH method (Tables S5, S9, S12, and S15 in S1 Appendix).

Validation of peak SNPs was attempted for 21 loci. At loci with multiple associated variants, the most significant variant was chosen as representative of the locus. For some loci, the exclusion of participants using self-reported heart disease medication resulted in a different peak SNP. Thus, for the *APOE* locus rs1160985 was genotyped instead of rs4420638 ($P_D = 2.67 \times 10^{-16}$ for TC, $P_D = 9.07 \times 10^{-9}$ for HDL, and $P_D = 7.44 \times 10^{-10}$ for TG); for the *APOA1* locus rs964184 was genotyped instead of rs3741298 ($P_D = 1.63 \times 10^{-7}$ for TC) or rs6589566 ($P_D = 3.98 \times 10^{-18}$ for TG); for the *MGAT1* locus rs1038143 was genotyped instead of rs249356 ($P_D = 1.06 \times 10^{-6}$ for HDL); for the *APOB* locus rs754523 was genotyped instead of rs1469513 ($P_D = 2.71 \times 10^{-6}$ for LDL).

We successfully genotyped the peak SNP, or a proxy SNP, in the replication cohorts for 15 loci. Two loci (*APOA1* with TG, $P_{DR} = 1.81 \times 10^{-29}$; *APOE* with TC, $P_{DR} = 4.29 \times 10^{-21}$, and LDL, $P_{DR} = 1.53 \times 10^{-27}$) demonstrated genome-wide significant associations in the discovery-replication meta-analysis (Table 2 and Tables S1, S7, S10, and S13 in S1 Appendix). An additional four associations demonstrated evidence of replication with consistent directions of effect and suggestive joint P_{DR} values (*GCKR* with TG, $P_{DR} = 5.62 \times 10^{-8}$; *MGAT1* with HDL, $P_{DR} = 2.91 \times 10^{-7}$; *APOA1* with TC, $P_{DR} = 1.72 \times 10^{-6}$; *RAB21* with HDL, $P_{DR} = 5.92 \times 10^{-7}$). Three associations had suggestive joint P_{DR} values driven by the discovery associations only (*APOB* with LDL, $P_{DR} = 5.81 \times 10^{-6}$; *LIPC* with HDL, $P_{DR} = 9.15 \times 10^{-7}$; *CDH4* with HDL, $P_{DR} = 8.77 \times 10^{-6}$); associations at *APOB* and *CDH4* had consistent directions of effect. Among the remaining loci with at least suggestive association in the discovery sample, but not in the discovery-replication meta-analysis, consistent effect directions were also seen for TC and *APOB* and *ZHX2*; LDL and *ALG10* and *CPNE8* (Table 2 and Tables S1, S7, S10, and S13 in S1 Appendix).

We compared the directions of effect to those previously reported in Willer et al. [8] and Teslovich et al. [33] (as made available by the Center for Statistical Genetics at the University of Michigan: http://csg.sph.umich.edu/willer/public/lipids2013/ and http://csg.sph.umich.edu/willer/public/lipids2010/) for *APOB*, *GCKR*, *APOA1*, *LIPC*,

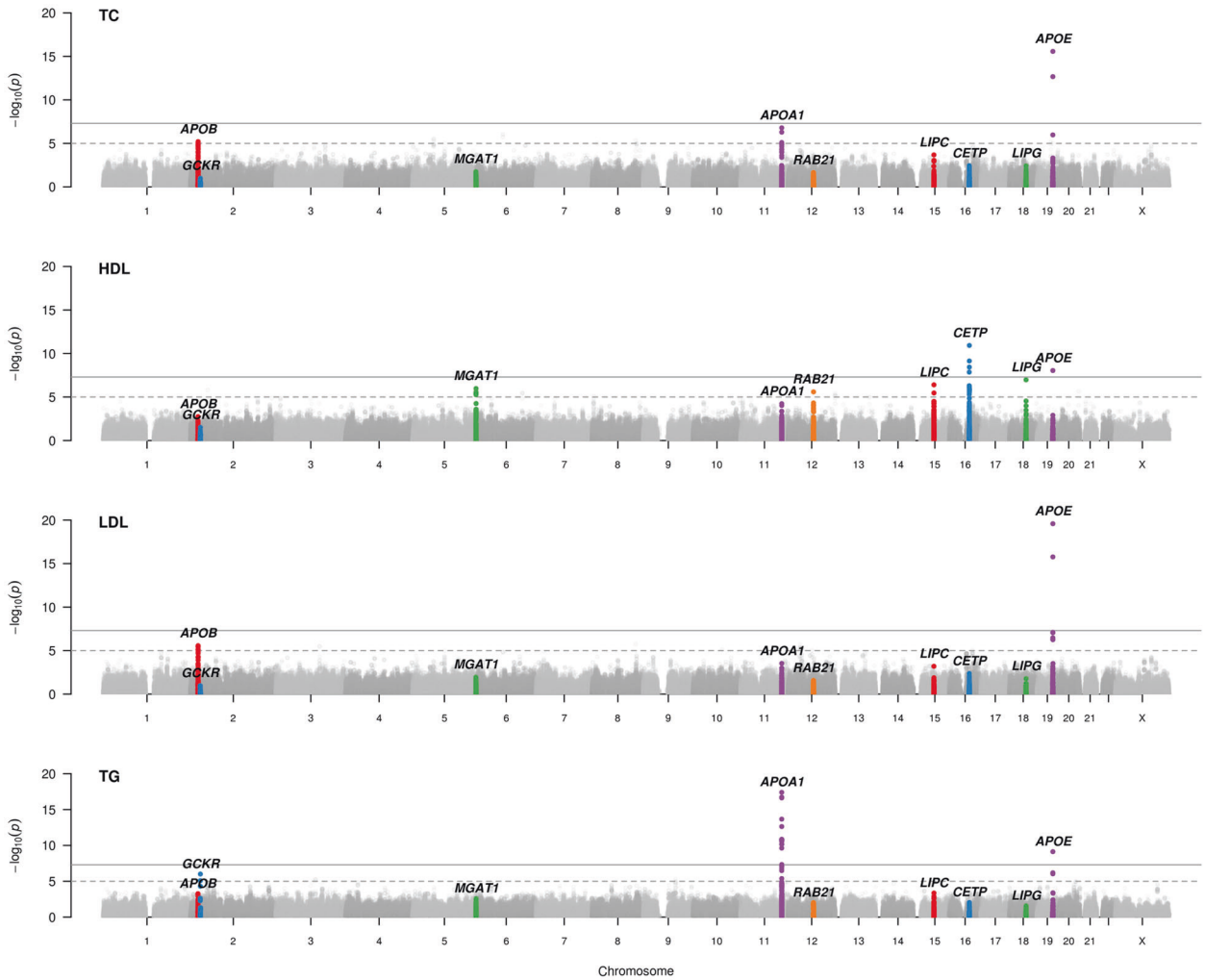

Fig. 1 Manhattan plots for GWAS of four lipid traits in the discovery cohort of 2849 Samoans. The dashed and solid lines denote genome-wide suggestive and genome-wide significant *P* value thresholds ($P < 1 \times 10^{-5}$ and $P < {}^{5} \times 10^{-8}$, respectively). Peaks are labeled with the candidate gene or closest gene in the region if they have at least suggestive association in the discovery cohort for at least one trait and demonstrate evidence of replication or have been previously associated

LIPG, and *APOE* (i.e., genome-wide suggestive loci that have been previously associated with lipid traits). For all variants from established lipid loci, the locus name and associated traits from Teslovich et al. and Willer et al. are given in Table 2. We observed a consistent direction of effect for the representative SNP for all associations (Table 2).

The effect allele frequencies (EAFs) in the two samples —discovery and replication—were largely similar for each of the 15 successfully genotyped SNPs (Tables S1, S7, S10, and S13 Appendix). However, many loci had markedly different EAFs between Samoans and other 1000 Genomes populations (Table 2). For example, compared with 1000 Genomes populations, there were higher EAFs in Samoans for rs964184 near *APOA1* (G allele frequency: 0.440 in Samoans vs. <0.277 in 1000 Genomes populations), rs1160985 near *APOE* (C allele frequency: 0.724 in Samoans vs. <0.659 in 1000 Genomes populations), and rs1038143 near *MGAT1* (A allele frequency: 0.309 in Samoans vs. <0.148 in 1000 Genomes populations).

Discussion

In this study, we examined four measures of fasting lipid levels—TC, HDL, LDL, and TG—for associations with 659,492 SNPs from a genome-wide array in a discovery cohort of 2849 Samoans, with follow-up genotyping of significant and suggestive findings in a replication cohort comprising 1798 Samoans from Samoa and American Samoa. Thirty-one loci had at least suggestive evidence of association with one or more lipid traits in the discovery cohort, of which eight have been reported to be associated with lipid levels previously: *APOB, GCKR, MGAT1,*

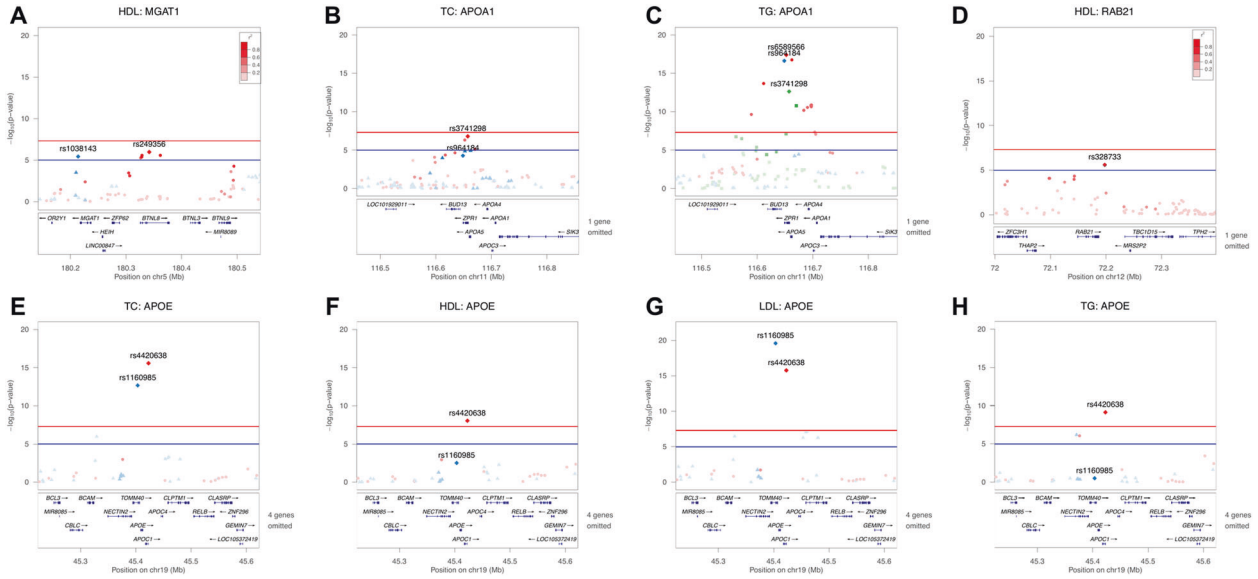

Fig. 2 Regional association plots for selected loci. Regional association plots generated in LocusZoom [49] showing $-\log_{10}$ (*P* values) for SNPs in (**a**) *MGAT1* locus and HDL, (**b**) *APOA1* locus and TC, (**c**) *APOA1* locus and TG, (**d**) *RAB21* locus and HDL, and the *APOE* locus and (**e**) TC, (**f**) HDL, (**g**) LDL, and (**h**) TG. Points are color coded within each plot according to pairwise linkage disequilibrium (r^2) with the labeled SNPs; the saturation of the color of each plotted SNP measures the linkage disequilibrium (r^2) with the labeled SNP sharing the same color

APOA1, *LIPC*, *CETP*, *LIPG*, and *APOE* [8, 10, 33]. Enrichment analyses highlighted known lipid metabolism gene sets and previously associated lipid loci.

We observed a difference in the architecture of the statistical association signals between the four lipid traits and variants near *APOE* (Fig. 2e–h). The peak SNP for TC and LDL was rs1160985, an intronic variant in *TOMM40* upstream of *APOE*; whereas the peak SNP for HDL and TG was rs4420638, an intergenic variant downstream of *APOC1* and *APOE*. rs1160985 demonstrated evidence of replication for TC and LDL but not for HDL and TG, consistent with the discovery findings. Sanna et al. report two independent signals for APOE, led by missense variants rs7412 and rs429358 [39]. Neither rs7412 nor rs429358 was genotyped in these samples so *APOE* isoform could not be determined. However, rs4420638, the sentinel SNP in our analyses of TC, HDL, and TG and in Sanna et al., is in high LD with rs429358. The sentinel SNP from our analysis of LDL was rs1160985, which is not in LD with rs4420638 in Samoans ($r^2 = 0.093$). It appears that this variant may be flagging an additional independent signal than the two reported in Sanna et al., as it is in low LD with rs7412 in all 1000 Genomes population (max $r^2 = 0.23$ in Southern Han Chinese; LDPop accessed 5 May 2020 [40]). While this could support a shared genetic architecture for TC and LDL and for HDL and TG in Samoans, this study was not positioned to adequately capture the association signal present at this locus. Further fine-mapping of this locus, including sequencing or imputation of the 19q13.2 region, is needed to dissect the genetic architecture of *APOE* and lipid levels in Polynesians and would give further insight into how much genetic architecture is shared between Samoans and better-studied populations.

We did not observe suggestive or genome-wide significant association with several loci which have figured prominently in multiple lipid GWAS (e.g., *LPL*, *LDLR*, *CILP2*, *FADS1/2/3*, *ANGPTL3*, *SORT1*, *PPP1R3B*, *MIXIPL*, *HNF4A*, *PCSK9*, *GALNT2*, *HMGCR*) either because we lacked sufficient power to detect their effects, the effects are negligible in Samoans, or the allele frequencies of associated variants are different enough in Samoans to hinder detection. However, it is important to note that this study was not designed to evaluate the effect of known lipids loci in Samoans, nor were previously associated loci examined specifically.

We detected and replicated a suggestive association between HDL and a variant on 5q35.3 (Fig. 2a). While the peak SNP lies within an intron of *BTNL8*, the variant selected for follow-up genotyping is intergenic, downstream of *MGAT1*. Previous association between the *MGAT1* locus and HDL have been reported, including a significant association with a variant in *MGAT1*, rs634501, in a study of Han Chinese and a suggestive association with variants near *MGAT1* in a GWAS in the Micronesian population of Kosrae [10, 41]. Furthermore, in Europeans, variation near *MGAT1* has been associated with BMI, serum fatty acid levels and composition, and glucose response [42–44]. The encoded MGAT enzyme plays a major role in the absorption of dietary fat in the intestine [45]. As this variant has only demonstrated evidence of association in non-European

Table 2 Significant and suggestive loci and replication genotyping

Total cholesterol

Locus	SNP	Chr	BP	EA	OA	Dir	P_D	P_R	P_{DR}	Known gene	Traits	SAM	EAS	SAS	EUR	AMR	AFR
APOB	rs754523	2	21311691	G	A	+++	*6.25E–06*	0.178	1.20E–05	*APOB*	C,H,L,T	0.247	0.265	0.140	0.309	0.306	0.201
PDE4D	rs7711093	5	59593138	G	A	++−	3.01E–06	0.747	5.37E–04			0.506	0.643	0.553	0.856	0.775	0.852
LUCAT1	rs10072084	5	90539203	C	T	+··	*9.48E–06*					0.541	0.559	0.422	0.232	0.429	0.822
FILIP1	rs2951921	6	76165524	T	C	+··	*9.04E–07*					0.073	0.022	0.063	0.015	0.030	0.293
ZHX2	rs7841763	8	123971081	T	C	+++	*4.82E–06*	0.631	1.03E–04			0.043	0.023	0.146	0.102	0.058	0.169
APOA1	rs964184[a]	11	116648917	C	G	+++	5.37E–05	0.009	*1.72E–06*	*APOA1*	C,H,L,T	0.560	0.760	0.771	0.838	0.723	0.779
SIRT2	rs10405150	19	39387919	C	T	+··	*6.34E–06*					0.056	0.147	0.144	0.081	0.117	0.774
ZNF283	rs16976816	19	44339377	G	A	+··	*9.78E–06*					0.977	0.970	0.994	0.987	0.976	0.864
APOE	rs1160985[a]	19	45403412	C	T	+++	**2.13E–13**	**3.36E–09**	**4.29E–21**	*APOE*	C,H,L,T	0.724	0.659	0.590	0.554	0.447	0.378

HDL

Locus	SNP	Chr	BP	EA	OA	Dir	P_D	P_R	P_{DR}	Known gene	Traits	SAM	EAS	SAS	EUR	AMR	AFR
STON1-GTF2A1L	rs6739536	2	48831901	A	G	−··	*1.58E–06*					0.762	0.676	0.837	0.918	0.842	0.609
MGAT1	rs1038143[a]	5	180213878	T	C	−−−	*3.72E–06*	0.016	*2.91E–07*			0.309	0.110	0.137	0.148	0.081	0.013
AKAP7	rs3777486	6	131584648	G	A	−+−	*3.09E–06*	0.808	4.37E–04			0.976	0.909	0.898	0.858	0.793	0.949
CSMD1	rs1626142	8	4345284	T	C	−··	*7.67E–06*					0.612	0.400	0.435	0.298	0.392	0.654
RAB21	rs328733	12	72197574	T	C	−−−	*2.57E–06*	0.036	*5.92E–07*			0.788	0.626	0.700	0.869	0.738	0.611
ZNF10	rs2292029	12	133734113	A	G	−··	*4.05E–06*					0.179	0.112	0.254	0.245	0.156	0.008
HS6ST3	rs16953620	13	97508453	A	G	−··	*8.48E–06*					0.968	0.950	0.949	0.848	0.911	0.836
LIPC	rs10438284	15	58629424	G	A	−+−	*4.00E–07*	0.133	*9.15E–07*	*LIPC*	C,H,T	0.286	0.326	0.173	0.272	0.199	0.047
CETP	rs289708	16	57038162	T	C	−··	**1.19E–11**			*CETP*	C,H,L,T	0.905	0.815	0.800	0.860	0.823	0.597
LIPG	rs16950739	18	47138509	T	C	−··	*1.07E–07*			*LIPG*	C,H	0.019	0.058	0.190	0.057	0.140	0.004
APOE	rs1160985[a]	19	45403412	C	T	−+−	0.003	0.342	0.004	*APOE*	C,H,L,T	0.724	0.659	0.590	0.554	0.447	0.378
CDH4	rs817687	20	59753355	A	G	−−−	*2.31E–06*	0.237	*8.77E–06*			0.986	0.924	0.945	0.967	0.855	0.610

LDL

Locus	SNP	Chr	BP	EA	OA	Dir	P_D	P_R	P_{DR}	Known gene	Traits	SAM	EAS	SAS	EUR	AMR	AFR
APOB	rs754523[a]	2	21311691	G	A	+++	*3.25E–06*	0.158	*5.81E–06*	*APOB*	C,H,L,T	0.247	0.265	0.140	0.309	0.306	0.201
KALRN	rs6789134	3	123942339	G	A	++−	*3.22E–06*	0.925	1.96E–04			0.078	0.161	0.120	0.046	0.118	0.212
ZHX2	rs7841763	8	123971081	T	C	++−	*1.80E–06*	0.557	3.78E–05			0.043	0.023	0.146	0.102	0.058	0.169
SH2D4B	rs10509415	10	82473065	A	C	+··	*7.96E–06*					0.706	0.510	0.735	0.758	0.808	0.708
ALG10	rs3912355	12	34079616	C	T	+++	*2.12E–06*	0.544	3.97E–05			0.855	0.872	0.651	0.605	0.732	0.863
ALG10B	rs10880642	12	38554152	A	G	+··	*5.56E–06*					0.802	0.715	0.555	0.481	0.532	0.299
CPNE8	rs11169807	12	39244161	C	T	+++	*4.77E–06*	0.418	4.13E–05			0.794	0.644	0.536	0.505	0.408	0.898
LINC02408	rs17104016	12	67969929	A	T	+··	*9.29E–06*					0.574	0.734	0.924	0.853	0.906	0.838

Table 2 (continued)

Total cholesterol																
Locus	SNP	Chr	BP	EA	OA	Dir	P_D	P_R	P_{DR}	Known gene	Traits	SAM	EAS	SAS	EUR	AMR	AFR
LINC00922	rs254371	16	65943650	T	C	+··	*9.04E–06*					0.555	0.698	0.650	0.599	0.693	0.884
ZNF283	rs16976816	19	44339377	G	A	+··	*1.78E–06*					0.977	0.970	0.994	0.987	0.976	0.864
APOE	rs1160985	19	45403412	C	T	+++	**2.61E–20**	**5.07E-09**	**1.53E–27**	*APOE*	C,H,L,T	0.724	0.659	0.590	0.554	0.447	0.378
Triglycerides																	
Locus	SNP	Chr	BP	EA	OA	Dir	P_D	P_R	P_{DR}	Known gene	Traits	SAM	EAS	SAS	EUR	AMR	AFR
GCKR	rs780094	2	27741237	T	C	+++	*9.84E–07*	0.01	*5.62E–08*	*GCKR*	C,T	0.334	0.476	0.198	0.411	0.360	0.132
CD200	rs2399416	3	112059213	A	G	++−	*5.12E–06*	0.668	9.29E–04			0.021	0.148	0.140	0.393	0.272	0.101
SPIN1	rs7861888	9	90886340	A	G	+··	*4.24E–06*					0.706	0.719	0.897	0.921	0.842	0.989
APOA1	rs964184[a]	11	116648917	C	G	+++	**2.37E–17**	**8.97E-14**	**1.81E–29**	*APOA1*	C,H,L,T	0.440	0.240	0.229	0.162	0.277	0.221
KIRREL3	rs3018434	11	126805881	G	A	+··	*4.16E–06*					0.916	0.803	0.920	0.866	0.857	0.974
APOE	rs1160985[a]	19	45403412	T	C	+−+	0.312	0.836	0.507	*APOE*	C,H,L,T	0.276	0.341	0.410	0.446	0.553	0.623

$P_D$discovery cohort GWAS p value

P_R replication cohort p value

P_{DR}joint discovery and replication cohorts p value

Known gene Loci observed in Teslovich et al. [33] or Willer et al. [8]

Traits Traits locus associated with in Teslovich et al. [33] or Willer et al. [8]

C TC

H HDL

L LDL

T TG

EA effect allele, *OA* other allele (reference allele), *Dir* direction of the effect in each of the four samples (+ indicates the effect allele is increasing the trait value on the raw scale), *SAM* Samoan effect allele frequency (EAF), *EAS* East Asian EAF, *SAS* South Asian EAF, *EUR* European EAF, *AMR* Admixed American EAF, *AFR* African EAF

Significant P values $< 5 \times 10^{-8}$ are in bold, while suggestive P values between 1×10^{-5} and 5×10^{-8} are in italics

[a]The SNP genotyped in the replication population was not the peak SNP at this locus for this trait

population, it may be tagging another variant or haplotype that is rare or absent in Europeans but is present in Pacific Island or South Asian populations and is itself associated with HDL.

We also detected and replicated a novel suggestive association between HDL and a variant downstream of *RAB21* (Fig. 2d). Unlike the variant downstream of *MGAT1*, we observed similar allele frequencies between Samoans and 1000 Genomes populations (i.e., <20% difference between Samoans and another 1000 Genomes population) in the HDL-associated variant downstream of *RAB21*. This region, 12q21, was previously seen in linkage analysis with both univariate and bivariate scans of TC and LDL [12]. The individuals included in this linkage analysis are also included in our 2002–2003 replication sample set, however, they do not appear to be driving the association signal near *RAB21* (Table S7 in S1 Appendix). Variation near *RAB21* has been previously associated with obesity [46]. *RAB21* belongs to the family of monomeric GTPases involved in control of cellular membrane trafficking and is involved in the targeted trafficking of integrins and the regulation of cell adhesion and migration [47, 48].

This study is limited in drawing conclusions about the genetic architecture of lipids in Samoans, as replication genotyping was unavailable for many loci and due to the lack of genome-wide imputation. Future studies, evaluating the evidence of association between associations seen here in separate cohorts as well fine-mapping loci with genotype imputation (given the availability of a relevant reference panel), are necessary to fully evaluate the genetic architecture of lipids in Samoans.

This is the first GWAS of lipid phenotypes in Samoans, and we observed association with many known lipid loci, which was further supported by the gene-set enrichment analysis highlighting lipid metabolism gene sets. However, the difference in association results near *APOE*, coupled with evidence of Pacific-Islander–specific associations with *MGAT1* and *RAB21* suggest that some, but not all, of the genetic architecture of lipids is shared between Samoans and other populations. Given this evidence of a partially distinct genetic architecture of lipids in Samoans, further investigation and fine-mapping of lipid loci, especially that across multiple ethnicities, is warranted.

Supplementary information is available at the Journal of Human Genetics website.

Data availability

The discovery cohort data are available from dbGaP (accession number: phs000914.v1.p1). The data from the replication cohort (recruited in 1990–1995 and 2002–2003) are not available as participants were not consented for data sharing.

Acknowledgements We would like to thank the Samoan participants of the study, local village authorities, and the many Samoan and other field workers over the years. We acknowledge the Samoan Ministry of Health, the Samoa Bureau of Statistics, and the American Samoan Department of Health for their support of this research. We give particular thanks to two research assistants, Melania Selu and Vaimoana Lupematasila, who contributed to the 2010 recruitment and continue to assist us in our work in Samoa. This work was funded by the National Institute of Health grants R01HL093093 (STM), R01HL133040 (RLM), R01AG09375 (STM), R01HL52611 (MI Kamboh), R01DK59642 (STM), and R01DK55406 (RD). Genotyping was performed in the Core Genotyping Laboratory at the University of Cincinnati, funded by National Institutes of Health grant P30ES006096 (SM Ho).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Hawley NL, Minster RL, Weeks DE, Viali S, Reupena MS, Sun G, et al. Prevalence of adiposity and associated cardiometabolic risk factors in the samoan genome-wide association study. Am J Hum Biol. 2014;26:491–501.
2. Keighley E, McGarvey ST, Quested C, Mccuddin C, Viali S, Maga UA. Nutrition and health in modernizing Samoans: temporal trends and adaptive perspectives. In: Ohtsuka R, Ulijaszek SJ, editors. Health Change in the Asia-Pacific Region: Biocultural and Epidemiological Approaches. Cambridge, NY: Cambridge University Press; 2007. p. 147–91.
3. McGarvey ST. Cardiovascular disease (CVD) risk factors in Samoa and American Samoa, 1990–95. Pac Health Dialog 2001;8:157–62.
4. Kannel WB, Castelli WP, Gordon T, McNamara PM. Serum cholesterol, lipoproteins, and the risk of coronary heart disease. Framingham Study Ann Intern Med 1971;74:1–12.
5. Gordon T, Castelli WP, Hjortland MC, Kannel WB, Dawber TR. High density lipoprotein as a protective factor against coronary heart disease. The Framingham Study. Am J Med 1977;62:707–14.
6. Roeters van Lennep J, Westerveld HT, Erkelens DW, van der Wall EE. Risk factors for coronary heart disease: implications of gender. Cardiovasc Res. 2002;53:538–49.
7. Lange LA, Willer CJ, Rich SS. Recent developments in genome and exome-wide analyses of plasma lipids. Curr Opin Lipido. 2015;26:96–102.
8. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet. 2013;45:1274–83.
9. Burkhardt R, Kenny EE, Lowe JK, Birkeland A, Josowitz R, Noel M, et al. Common SNPs in *HMGCR* in micronesians and whites associated with LDL-cholesterol levels affect alternative splicing of exon13. Arterioscler Thromb Vasc Biol 2008;28:2078–84.
10. Lowe JK, Maller JB, Pe'er I, Neale BM, Salit J, Kenny EE, et al. Genome-wide association studies in an isolated founder population from the Pacific Island of Kosrae. PLoS Genet 2009;5:e1000365.
11. Dumitrescu L, Carty CL, Taylor K, Schumacher FR, Hindorff LA, Ambite JL, et al. Genetic determinants of lipid traits in diverse

populations from the population architecture using genomics and epidemiology (PAGE) study. PLoS Genet 2011;7:e1002138.
12. Åberg K, Dai F, Sun G, Keighley E, Indugula SR, Bausserman L, et al. A genome-wide linkage scan identifies multiple chromosomal regions influencing serum lipid levels in the population on the Samoan islands. J Lipid Res. 2008;49:2169–78.
13. Minster RL, Hawley NL, Su C-T, Sun G, Kershaw EE, Cheng H, et al. A thrifty variant in *CREBRF* strongly influences body mass index in Samoans. Nat Genet 2016;48:1049–54.
14. Kristiansson K, Naukkarinen J, Peltonen L. Isolated populations and complex disease gene identification. Genome Biol 2008;9:109.
15. Laurie CC, Doheny KF, Mirel DB, Pugh EW, Bierut LJ, Bhangale T, et al. Quality control and quality assurance in genotypic data for genome-wide association studies. Genet Epidemiol. 2010; 34:591–602.
16. Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. G3 (Bethesda). 2011;1:457–70.
17. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, et al. A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet. 2016;48:1279–83.
18. Das S, Abecasis GR, Browning BL. Genotype imputation from large reference panels. Annu Rev Genomics Hum Genet. 2018;19:73–96.
19. Deka R, Mc Garvey ST, Ferrell RE, Kamboh MI, Yu LM, Aston CE, et al. Genetic characterization of American and Western Samoans. Hum Biol 1994;66:805–22.
20. McGarvey ST, Levinson PD, Bausser-Man L, Galanis DJ, Hornick CA. Population change in adult obesity and blood lipids in American Samoa from 1976–1978 to 1990. Am J Hum Biol 1993;5:17–30.
21. Chin-Hong PV, McGarvey ST. Lifestyle incongruity and adult blood pressure in Western Samoa. Psychosom Med 1996;58: 131–7.
22. Galanis DJ, McGarvey ST, Quested C, Sio B, Afele-Fa'amuli SA. Dietary intake of modernizing Samoans: implications for risk of cardiovascular disease. J Am Diet Assoc. 1999;99:184–90.
23. Åberg K, Dai F, Sun G, Keighley ED, Indugula SR, Roberts ST, et al. Susceptibility Loci for Adiposity Phenotypes on 8p, 9p, and 16q in American Samoa and Samoa. Obesity. 2009;17:518–24.
24. Dai F, Keighley ED, Sun G, Indugula SR, Roberts ST, Åberg K, et al. Genome-wide scan for adiposity-related phenotypes in adults from American Samoa. Int J Obes. 2007;31:1832–42.
25. Åberg K, Dai F, Viali S, Tuitele J, Sun G, Indugula SR, et al. Suggestive linkage detected for blood pressure related traits on 2q and 22q in the population on the Samoan islands. BMC Med Genet. 2009;10:107.
26. Dai F, Sun G, Åberg K, Keighley ED, Indugula SR, Roberts ST, et al. A whole genome linkage scan identifies multiple chromosomal regions influencing adiposity-related traits among Samoans. Ann Hum Genet. 2008;72:780–92.
27. Chen W-M, Abecasis GR. Family-based association tests for genomewide association scans. Am J Hum Genet. 2007; 81:913–26.
28. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999;55:997–1004.
29. Purcell S, Cherny SS, Sham PC. Genetic power calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 2003;19:149–50.
30. Carithers LJ, Ardlie K, Barcus M, Branton PA, Britton A, Buia SA, et al. A novel approach to high-quality postmortem tissue procurement: The GTEx Project. Biopreserv Biobank 2015; 13:311–9.
31. Segrè AV, Groop L, Mootha VK, Daly MJ, Altshuler D. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. PLoS Genet. 2010;6:e1001058.
32. Lee PH, O'Dushlaine C, Thomas B, Purcell SM. INRICH: interval-based enrichment analysis for genome-wide association studies. Bioinformatics. 2012;28:1797–9.
33. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010;466:707–13.
34. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics 2007;23:1294–6.
35. R Core Team. R: A language and environment for statistical computing. Vienna, Austria, Austria: R Foundation for Statistical Computing; 2017.
36. Therneau TM, Sinnwell J kinship2: Pedigree functions. R package version 1.6.4; 2014.
37. Winkler TW, Day FR, Croteau-Chonka DC, Wood AR, Locke AE, Magi R, et al. Quality control and conduct of genome-wide association meta-analyses. Nat Protoc. 2014;9:1192–212.
38. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010;26:2190–1.
39. Sanna S, Li B, Mulas A, Sidore C, Kang HM, Jackson AU, et al. Fine mapping of five loci associated with low-density lipoprotein cholesterol detects variants that double the explained heritability. PLoS Genet. 2011;7:e1002198.
40. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics. 2015;31:3555–7.
41. Wu J, Yin R-X, Zhou Y-G, Zhang Q-H, Wu J-Z, Chen W-X. Association between the *MGAT1* rs634501 polymorphism and serum lipid traits in the Chinese Han and Maonan ethnic groups. Int J Clin Exp Pathol. 2018;11:5923–37.
42. Johansson Å, Marroni F, Hayward C, Franklin CS, Kirichenko AV, Jonasson I, et al. Linkage and genome-wide association analysis of obesity-related phenotypes: association of weight with the *MGAT1* gene. Obesity 2010;18:803–8.
43. Jacobsson JA, Rask-Andersen M, Risérus U, Moschonis G, Koumpitski A, Chrousos GP, et al. Genetic variants near the *MGAT1* gene are associated with body weight, BMI and fatty acid metabolism among adults and children. Int J Obes. 2012;36:119–29.
44. Del-Aguila JL, Beitelshees AL, Cooper-DeHoff RM, Chapman AB, Gums JG, Bailey K, et al. Genome-wide association analyses suggest NELL1 influences adverse metabolic response to HCTZ in African Americans. Pharmacogenomics J. 2014;14:35–40.
45. Yen C-LE, Stone SJ, Cases S, Zhou P, Farese RV. Identification of a gene encoding MGAT1, a monoacylglycerol acyltransferase. Proc Natl Acad Sci. 2002;99:8512–7.
46. Turcot V, Lu Y, Highland HM, Schurmann C, Justice AE, Fine RS, et al. Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. Nat Genet 2018;50:26–41.
47. Pellinen T, Ivaska J Integrin traffic. J Cell Sci. 2006;119(Pt 18):3723–31.
48. Pellinen T, Tuomi S, Arjonen A, Wolf M, Edgren H, Meyer H, et al. Integrin trafficking regulated by Rab21 is necessary for cytokinesis. Dev Cell 2008;15:371–85.
49. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010;26:2336–7.