A stop-gain variant in *BTNL9* is associated with atherogenic lipid profiles

Jenna C. Carlson,[1,2,18,19,*] Mohanraj Krishnan,[2] Samantha L. Rosenthal,[2,3] Emily M. Russell,[2] Jerry Z. Zhang,[1] Nicola L. Hawley,[4] Jaye Moors,[5] Hong Cheng,[6] Nicola Dalbeth,[7] Janak R. de Zoysa,[7] Huti Watson,[8] Muhammad Qasim,[8] Rinki Murphy,[7,9] Take Naseri,[10] Muagututi'a Sefuiva Reupena,[11] Satupa'itea Viali,[12] Lisa K. Stamp,[13] John Tuitele,[14] Erin E. Kershaw,[15] Ranjan Deka,[6] Stephen T. McGarvey,[16,17] Tony R. Merriman,[5,9] Daniel E. Weeks,[1,2] and Ryan L. Minster[2]

Summary

Current understanding of lipid genetics has come mainly from studies in European-ancestry populations; limited effort has focused on Polynesian populations, whose unique population history and high prevalence of dyslipidemia may provide insight into the biological foundations of variation in lipid levels. Here, we performed an association study to fine map a suggestive association on 5q35 with high-density lipoprotein cholesterol (HDL-C) seen in Micronesian and Polynesian populations. Fine-mapping analyses in a cohort of 2,851 Samoan adults highlighted an association between a stop-gain variant (rs200884524; c.652C>T, p.R218*; posterior probability = 0.9987) in *BTNL9* and both lower HDL-C and greater triglycerides (TGs). Meta-analysis across this and several other cohorts of Polynesian ancestry from Samoa, American Samoa, and Aotearoa New Zealand confirmed the presence of this association ($\beta_{HDL-C} = -1.60$ mg/dL, $p_{HDL-C} = 7.63 \times 10^{-10}$; $\beta_{TG} = 12.00$ mg/dL, $p_{TG} = 3.82 \times 10^{-7}$). While this variant appears to be Polynesian specific, there is also evidence of association from other multiancestry analyses in this region. This work provides evidence of a previously unexplored contributor to the genetic architecture of lipid levels and underscores the importance of genetic analyses in understudied populations.

Introduction

Atherogenic lipid profiles—increased total cholesterol (TC), low-density lipoprotein cholesterol (LDL-C), and triglycerides (TGs) as well as decreased high-density lipoprotein cholesterol (HDL-C)—are well-documented and heritable risk factors for cardiovascular disease (CVD) worldwide. While behavioral modifications and medication have been successful in improving lipid profiles, CVD is still the leading cause of death worldwide, particularly among people of Polynesian and Pacific Island ancestry.[1–5]

The examination of the genetic underpinnings of lipid variation through genome-wide association studies (GWASs) has identified numerous genetic associations. These discoveries have furthered the understanding of CVD and potential therapeutic targets; however, this research has primarily come from studies in European-ancestry populations. Recent efforts to diversify research on this topic have highlighted the importance of including diverse populations for gene discovery, yielding novel associations, improved fine mapping, and better polygenic risk scores.[6] Despite this, limited effort has focused on Polynesian populations, whose unique population history including genetic drift from founder effects, small population sizes, and population bottlenecks may provide insight into the biological foundations of variation in lipid levels, which would not only benefit Polynesian individuals but also those from other populations.[7–9]

One region of interest is 5q35, which has previously been associated with HDL-C levels in Micronesian and Samoan populations.[10,11] The causal variant at this locus, and the biological mechanism underlying this association with an atherogenic lipid profile, is unknown.

The aim of this study was to fine map this association signal in a cohort of 2,851 Samoan adults and replicate this association in several independent Polynesian cohorts from Samoa, American Samoa, and Aotearoa New Zealand. We identified a strong candidate causal variant at this locus, rs200884524—a stop-gain variant in

[1]Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA, USA; [2]Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA, USA; [3]Center for Craniofacial and Dental Genetics, Department of Oral Biology, University of Pittsburgh, Pittsburgh, PA, USA; [4]Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, CT, USA; [5]Department of Biochemistry, University of Otago, Dunedin, New Zealand; [6]Department of Environmental Health, College of Medicine, University of Cincinnati, Cincinnati, OH, USA; [7]Department of Medicine, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand; [8]Ngāti Porou Hauora Charitable Trust, Te Puia Springs, Tairāwhiti East Coast, New Zealand; [9]Maurice Wilkins Centre for Molecular Biodiscovery, University of Auckland, Auckland, New Zealand; [10]Ministry of Health, Government of Samoa, Apia, Samoa; [11]Lutia i Puava ae Mapu i Fagalele, Apia, Samoa; [12]School of Medicine, National University of Samoa, Apia, Samoa; [13]Department of Medicine, University of Otago Christchurch, Christchurch, New Zealand; [14]Department of Public Health, Government of American Samoa, Pago Pago, American Samoa; [15]Division of Endocrinology, Department of Medicine, University of Pittsburgh, Pittsburgh, PA, USA; [16]International Health Institute, Department of Epidemiology, Brown University, Providence, RI, USA; [17]Department of Anthropology, Brown University, Providence, RI, USA
[18]Twitter: @jenccarlson
[19]Lead contact
*Correspondence: jnc35@pitt.edu
https://doi.org/10.1016/j.xhgg.2022.100155

BTNL9—that is associated with lower HDL-C and higher TG concentrations.

Subjects and methods

An overview of the analytical methods is presented in Figure S1.

Discovery

We performed association testing between variants on 5q35 and lipid levels in a discovery cohort comprising 2,851 Samoan adults (at least 18 years old) drawn from a population-based sample recruited from Samoa in 2010 (Table 1), data for which are available from dbGaP (accession number: phs000914.v1.p1). The sample selection, data collection methods, and phenotyping, including the laboratory assays for serum lipid and lipoprotein levels, have been previously described.[12,13] Briefly, serum lipid levels (TC, HDL-C, LDL-C, and TG) were derived from fasting whole-blood samples collected after a minimum 10-h overnight fast. LDL-C levels were estimated using the Friedewald equation,[14] which does not accurately estimate LDL-C in individuals with TG ≥ 400 mg/dL, and, therefore, those individuals are missing LDL-C values. Analyses were performed without consideration of hypolipidemic medication, as it was not measured. However, individuals who reported taking medication for heart disease were excluded from analyses, as previous sensitivity analyses identified an association between self-reported use of heart disease medication and TC and LDL-C.[10]

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 the Samoan language.

Genotyping was performed using Genome-Wide Human SNP 6.0 arrays (Affymetrix), and extensive quality control was conducted via a pipeline developed by Laurie et al.[15] Additional details for sample genotyping and genotype quality control are described in Minster et al.[13]

A Samoan-specific genotype reference panel for imputation was created using the whole-genome sequence (WGS) data from the Trans-Omics for Precision Medicine (TOPMed) program from the National Institutes of Health's Heart, Lung, and Blood Institute (NHLBI). The reference panel is comprised of high-quality (i.e., passing all quality control [QC] filters and with a minimum depth of 10), bi-allelic markers from the freeze 8 call subset of the 1,284 individuals from the Samoan Adiposity Study. The reference panel was phased using Eagle v.2.4.1.[16] The Samoan-specific reference panel was then used to impute genotype data for the remaining Samoan participants in the discovery cohort who were not in the reference panel. Briefly, genotyped variant coordinates were converted to the hg38 genome build using LiftOver (https://genome.ucsc.edu/cgi-bin/hgLiftOver). Then, data were aligned to the reference panel using Genotype Harmonizer,[17] using the mafAlign option to align to the minor allele when LD alignment failed and the minor allele frequency (MAF) was $\leq 30\%$ in both the input and reference sets. The resulting variants were then phased with Eagle v.2.4.1,[16] and genome-wide imputation was performed using Minimac4.[18] Imputed genotype dosages from variants passing QC ($r^2 > 0.3$) were combined with genotyped variants for association testing.

We performed preliminary association testing between variants in a 1-Mb region around the sentinel SNP from the preliminary GWAS[10] (hg38 chr5:180414895-181414895) and the four lipid levels (TC, HDL-C, LDL-C, and TG) in the discovery cohort using linear mixed modeling with inverse-normally transformed traits, marginally rescaled variance (to restore it to the original variance before the transformation), and additive genotype coding as implemented in the *GENESIS* R package[19,20] adjusting for the following fixed effects: three principal components of ancestry (generated through PC-AiR,[21] which together accounted for >98% variation), age, age^2, sex, age × sex interaction, and age^2 × sex interaction. Age was mean centered for all analyses to avoid multicollinearity issues. Relatedness was measured through an empirical kinship matrix generated with *GENESIS* and was modeled through a random effect. Then, to estimate the effect size of rs200884524 for use in the meta-analysis across cohorts, we used a linear mixed-modeling strategy on the original lipid levels (i.e., not transformed), adjusting for the same fixed and random effects as implemented in the *lmekin* function in the *coxme* R package.[22] These models were also repeated for several cardiometabolic traits as described in Minster et al.[13] to examine the potential pleiotropic effects of rs200884524.

As a sensitivity analysis to check for robustness against deviations from normality, the association tests with rs200884524 were repeated using lipid-level residuals that were adjusted for three principal components of ancestry, age, age^2, sex, and age × sex and age^2 × sex interactions and were then inverse-normally transformed[23] using *RankNorm* function in the *RNOmni* R package.[24] To preserve the interpretation of the effect estimates, the results of the untransformed trait analyses are presented. The results of the sensitivity analysis were similar to the models in which lipid levels were not transformed (Table S1).

To detect a potential secondary signal, conditional analyses were run in the 5q35 region in the discovery cohort. rs200884524, the lead SNP in the region, was modeled as a fixed-effect covariate; the rest of the model parameters stayed the same.

Bayesian fine mapping of the discovery cohort results was performed using PAINTOR v.2.1.[25] Annotations of variant impact (high, moderate, low/modifier) as indicated by VEP v.91.1[26] were used to inform the fine mapping.

Evidence of selection at this locus was assessed using nS_L (number of segregation sites by length) analysis in 419 unrelated Samoans from the discovery cohort with no evidence of non-Oceanic admixture.[27] WGS was completed by TOPMed, and a MAF of >0.01 was used. Haplotype phasing was completed using Eagle v.2.4.1.[16] nS_L was calculated with *selscan* in R.[28]

Replication

Samoan/American Samoan replication cohorts

Two cohorts consisting of a total of 1,466 Samoan and American Samoan adults (at least 18 years old) were used for replication. The first was a longitudinal study of adiposity and CVD risk factors among adults from American Samoa and Samoa recruited between 1994 and 1995, during which lipid levels were measured from fasting serum samples. LDL-C levels were estimated using the Friedewald equation,[14] which does not accurately estimate LDL-C in individuals with TG ≥ 400 mg/dL, and, therefore, those individuals are missing LDL-C values. Detailed descriptions of the sampling, recruitment, and phenotyping have been reported previously.[29–31] The second study included adults from American Samoa and Samoa, recruited in 2002–2003, and was drawn from an extended family-based genetic linkage study of cardiometabolic traits.[32–34] Probands and relatives were unselected for obesity or

Table 1. Characteristics of discovery and replication cohorts

	Samoan discovery		Samoan/American Samoan replication n = 1,466								Aotearoa New Zealand replication n = 1,810					
			1994–1995 recruitment (n = 557)				2002–2003 recruitment (n = 909)									
	n = 2,851		Samoa		American Samoa		Samoa		American Samoa		East Polynesia (n = 1,109)		West Polynesia (n = 603)		East/West Polynesia (n = 98)	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
	n = 1,705	n = 1,146	n = 202	n = 184	n = 96	n = 75	n = 195	n = 185	n = 308	n = 221	n = 456	n = 653	n = 182	n = 421	n = 38	n = 60
Age (years)																
Mean	44.77	45.56	43.5	43.08	42.53	45.4	45.05	42.01	42.96	43.24	51.23	52.72	44.05	44.46	37.66	39.48
SD	11.09	11.29	8.49	8.81	7.57	10.5	16.79	15.72	16.03	16.42	15.48	14.32	15.71	13.99	13.77	16.15
TC (mg/dL)																
Mean	199.23	200.27	199.35	196.79	198.68	194.57	203.87	195.86	187.39	188.71	191.41	191.55	180.42	190.12	178.9	193.74
SD	36.13	38.66	36.88	35.07	30.63	30.46	35.92	38.05	38.97	37.17	42.36	45.54	45.41	43.61	43.01	40.53
N Miss	–	–	–	–	–	–	–	–	–	–	1	–	1	–	–	–
HDL-C (mg/dL)																
Mean	46.48	43.7	43.87	41.38	34.75	32.39	47.46	45.7	42	38.54	49.27	42.36	47.8	41.42	52.66	41.46
SD	10.82	11.18	10.62	11.02	7.66	8.22	10.37	11.48	8.25	8.86	15.08	12.52	12.79	12.32	14.18	12.17
N Miss	–	–	–	3	–	4	1	–	–	–	1	–	1	1	–	–
LDL-C (mg/dL) [a]																
Mean	130.17	130.2	139.12	137.45	141.34	134.57	134.65	127.27	118.85	117.48	107.67	110.29	100.6	113.79	95.85	119.37
SD	32.43	34.38	33.22	32.37	30.85	35.91	32.39	36.57	34.58	33.57	34.82	40.94	37.18	39.16	35.7	32.63
N Miss	10	25	1	5	–	9	1	3	4	19	34	76	9	37	2	8
TG (mg/dL)																
Mean	114.97	139.16	85.01	101.42	115.2	172.83	109.38	119.26	131.28	191.64	180.88	216.8	163.07	201.47	145.98	208.49
SD	80.47	112.7	76.41	95.75	52.17	121.92	54.4	94.9	75.68	171.68	120.22	153.45	117.71	133.27	94.24	137.91

Mean (SD) are given for age, total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), and triglycerides (TGs) for each cohort. When applicable, the number of participants missing the measurement (N miss) is given.
[a]LDL-C was estimated using the Friedewald equation so individuals with TG ≥400 mg/dL have missing LDL-C values.

related phenotypes, and all individuals self-reported Samoan ancestry. The recruitment process, criteria used for inclusion in this study, and phenotyping have been described in detail previously.[35,36] To account for temporal differences, each of the two studies was modeled separately—Samoa/American Samoa 1994–1995 (n = 557) and Samoa/American Samoa 2002–2003 (n = 909). Participants in these two cohorts were genotyped using the Infinium Global Screening Array-24 v.3.0 BeadChip (Illumina, CA, USA) with custom content that included rs200884524. These studies were approved by the institutional review boards of Brown University and the American Samoa Department of Health, as well as the Health Research Committee of the Samoa Ministry of Health. All participants gave written informed consent.

Aotearoa New Zealand replication cohorts

A total of 1,810 adults (at least 16 years old) of Polynesian (NZ Maori and Pacific) ancestry were recruited from Aotearoa New Zealand as part of a study of risk factors for gout, type 2 diabetes, and kidney disease. Lipid parameters (VLDL and TG) have previously been associated with gout in an Aotearoa NZ Polynesian cohort.[37] Participants were divided into three replication cohorts based on the self-reported ancestry of their grandparents (East Polynesia [Aotearoa New Zealand Māori and Cook Island Māori] n = 1,109; West Polynesia [Samoa, Tonga, Pukapuka, and Niue], n = 603; and Mixed East/West Polynesia, n = 98). A separate Māori sample set from the rohe (area) of the Ngāti Porou iwi (tribe) of the Tairāwhiti (East Coast of the North Island of New Zealand) region was also included in the East Polynesia group. This sample set was recruited in collaboration with the Ngāti Porou Hauora (Health Service) Charitable Trust. The details of sampling, recruitment, phenotyping, and creation of principal components of ancestry have been previously reported.[38,39] Briefly, lipid levels were measured from serum samples at Southern Community Laboratories (Dunedin, NZ). LDL-C levels were estimated using the Friedewald equation,[14] which does not accurately estimate LDL-C in individuals with TG ≥400 mg/dL, and, therefore, those individuals are missing LDL-C values. Genotyping of rs200884524 was carried out using the TaqMan SNP genotyping assay technology (Applied Biosystems, Foster City, CA, USA) using a LightCycler 480 Real-Time Polymerase Chain Reaction (PCR) system (Roche, Indianapolis, IN, USA). This study was approved by the New

Zealand Multi-Region Ethics Committee and the Northern Y Region Health Research Ethics Committee. All participants gave written informed consent.

Replication analyses

Association testing for rs200884524, the sentinel SNP in the 5q35 region as identified in the discovery cohort, and the lipid levels (TC, HDL-C, LDL-C, and TG) was performed in each of the five replication cohorts using linear mixed modeling, adjusting for the following fixed effects: principal components (PCs) of ancestry (four PCs generated through PC-AiR[21] were used for the Samoan/American Samoan replication cohorts, three PCs for the Aotearoa New Zealand replication cohorts), polity (an indicator of Samoa/American Samoa in the two respective cohorts), age, age^2, sex, age × sex interaction, and age^2 × sex interaction. Age was mean centered within each cohort for all analyses to avoid multicollinearity issues. Relatedness was measured through empirical kinship matrices and modeled as a random effect using the *lmekin* function in the *coxme* R package.[22] For the Samoan/American Samoan replication cohorts, the empirical kinship matrix was generated with *GENESIS*;[19,20] for the Aotearoa New Zealand replication cohorts, the empirical kinship matrix was calculated in PLINK v.1.9[40,41] as described previously.[38]

Meta-analysis

The effects of rs200884524 on the four lipid traits from the discovery and 5 replication cohort analyses were combined using an inverse-variance fixed-effect meta-analysis as implemented in the *rmeta* R package.[42] Additionally, meta-analyses were conducted separately across the two Samoan/American Samoan replication cohorts and the three Aotearoa New Zealand replication cohorts. Heterogeneity was assessed with Cochran's Q statistic.

Results

Descriptions of age, TC, HDL-C, LDL-C, and TG levels in the discovery and replication cohorts are given in Table 1. Generally, the cohorts were very similar (with the exception of TG, which varies more widely across cohorts).

The association analyses in the discovery cohort in the 1-Mb region on 5q35 (Figures S2–S5) highlighted an association between a stop-gain variant (rs200884524; c.652C>T, p.R218*) in the butyrophilin-like 9 (*BTNL9*) gene with HDL-C (p = 4.05×10^{-8}; Figure S2). Conditional analyses on rs200884524 showed some evidence of a secondary signal approximately 0.449 Mb upstream from the lead variant in the region (rs71680280; CA>C, conditional p = 6.52×10^{-6}, linkage disequilibrium [LD] with rs200884524 $r^2 = 0.035$, Figure S2). Bayesian fine mapping pointed to rs200884524 with high posterior probability (PP) of causality (PP = 0.9987); all other variants in the region had PP <0.10. rs200884524 did not show evidence of association with any other cardiometabolic phenotype measured (Table S1).

The estimated effects of rs200884524 on the four lipid levels for each cohort are given in Table 2. rs200884524 was significantly associated with lower average HDL-C levels overall (meta-analysis β = −1.60 mg/dL, p = 7.63×10^{-10}; Figure 1), with no evidence of heterogeneity of effect across cohorts (p = 0.65). Similarly, rs200884524

was significantly associated with higher average TG levels overall (meta-analysis β = 12.00 mg/dL, p = 3.82×10^{-7}, heterogeneity p = 0.75). Sensitivity analyses using inverse-normal-transformed phenotypes showed similar results (Table S2). rs200884524 accounts for 0.94% and 0.68% of the variation in HDL-C and TG, respectively, in the discovery cohort.

rs200884524 has MAF <0.0001 in gnomAD v.3.1.1 (accessed February 16, 2022); of the 14 alleles seen in gnomAD, 7 are from East Asians, with no observed homozygotes. However, we observed a MAF of 0.223 in the Samoan discovery cohort with 163 observed homozygotes. The MAFs of rs200884524 ranged from 0.049 to 0.233 in the Samoan/American Samoan and Aotearoa New Zealand replication cohorts, with higher frequencies observed among the Western Polynesian participants (Table 2). Consistent with these disparate allele frequencies, rs200884524 also showed evidence of positive selection in the discovery cohort (nS_L score of 1.79, 98.9 percentile in the Samoan genome).

Discussion

We found strong evidence of association between a stop-gain variant in *BTNL9*, rs200884524, and atherogenic serum lipid profiles (lower HDL-C, higher TG) across several independent cohorts of Polynesian individuals. Interestingly, this variant appears to be Polynesian specific—common in those of Western Polynesian ancestry (which includes Samoa), rarer in those of Eastern Polynesian ancestry, and absent from other non-Polynesian populations. When compared with effect sizes of other index variants from GWASs of HDL-C,[6] the effect size of rs200884524 was moderate to large, which is similar to other findings in this population.[13,43,44]

Moreover, the identification of rs200884524 was achieved using a population-specific genotype imputation panel, highlighting the importance of such panels when working with isolated populations. Previous work (using the discovery cohort from these results) observed a suggestive association between this region and HDL-C[10] but was limited to SNP array data only rather than the dense data from imputed genotypes. The sparse genotypes in this region did not facilitate accurate fine mapping. In fact, we previously nominated *MGAT1* as the causal gene for this locus as the stop-gain variant in *BTNL9* was not genotyped. Once the dense information from imputed genotypes was available, it was easy to identify the stop-gain variant rs200884524 as a variant of interest, which led us to include it with the custom content selected for genotyping in the Samoan/American Samoan replication cohorts. These results highlight the necessity of population-specific imputation reference panels and the fruitfulness of gene discovery in isolated populations.

There is evidence that rs200884524 plays a role in nonsense-mediated decay (NMD) efficacy; the variant is

Table 2. rs200884524 association results for untransformed lipid levels: Total cholesterol (TC), HDL cholesterol (HDL-C), LDL cholesterol (LDL-C), and triglycerides (TG)

Cohort	TC β (95% CI) mg/dL p	HDL-C β (95% CI) mg/dL p	LDL-C β (95% CI) mg/dL p	TG β (95% CI) mg/dL p
Samoan Discovery (n = 2,851; MAF = 0.223)	−1.65 (−3.91, 0.61)	−1.84 (−2.54, −1.14)[a]	−0.93 (−2.98, 1.12)	13.48 (7.49, 19.47)[a]
	0.15	2.74×10^{-7a}	0.37	1.04×10^{-5a}
Samoan/American Samoan replication cohorts				
1994–1995 Samoan/American Samoan (n = 557; MAF = 0.233)	0.68 (−3.91, 5.26)	−1.98 (−3.35, −0.60)	0.85 (−3.64, 5.34)	15.60 (3.67, 27.53)
2002–2003 Samoa/American Samoan (n = 909; MAF = 0.202)	0.76 (−3.27, 4.79)	−0.83 (−1.99, 0.33)	0.71 (−3.09, 4.50)	4.65 (−7.88, 17.17)
Meta-analysis of 2 Samoan/American Samoan replication cohorts	0.72 (−2.30, 3.75)	−1.31 (−2.19, −0.42)[a]	0.77 (−2.13, 3.66)	10.39 (1.75, 19.03)[a]
Meta-analysis effect p value	0.64	3.86×10^{-3a}	0.60	0.018[a]
Heterogeneity p value	0.98	0.21	0.96	0.21
Aotearoa New Zealand replication cohorts				
Eastern Polynesian (n = 1,109; MAF = 0.049)	−1.40 (−9.57, 6.76)	−0.58 (−3.16, 1.99)	−1.63 (−9.10, 5.84)	15.52 (−10.85, 41.88)
Western Polynesian (n = 603; MAF = 0.216)	−1.21 (−6.98, 4.56)	−1.54 (−3.21, 0.14)	0.06 (−5.27, 5.40)	4.90 (−11.97, 21.76)
Mixed Eastern/Western Polynesian (n = 98; MAF = 0.133)	0.48 (−15.27, 16.24)	−2.90 (−7.71, 1.91)	4.88 (−8.05, 17.80)	12.41 (−35.63, 60.44)
Meta-analysis of 3 Aotearoa New Zealand replication cohorts	−1.13 (−5.65, 3.39)	−1.38 (−2.73, −0.04)[a]	0.04 (−4.08, 4.15)	8.34 (−5.29, 21.96)
Meta-analysis effect p value	0.62	0.044[a]	0.99	0.23
Heterogeneity p value	0.98	0.68	0.69	0.79
Meta-analysis (all cohorts)	−0.85 (−2.53, 0.83)	−1.60 (−2.11, −1.09)[a]	−0.31 (−1.86, 1.24)	12 (7.37, 16.63)[a]
Meta-analysis effect p value	0.32	7.63×10^{-10a}	0.70	3.82×10^{-7a}
Heterogeneity p value	0.90	0.65	0.89	0.75

For individual cohorts, βs, 95% confidence intervals (CIs), and p values obtained from linear mixed models adjusting for fixed effects of principal components of ancestry, polity (Samoa/American Samoa), age, age^2, sex, age × sex interaction, and age^2 × sex interaction. Age was mean centered for all analyses to avoid multicollinearity issues. Relatedness was measured through an empirical kinship matrix and was modeled through a random effect. For meta-analyses, βs, 95% CIs, and p values were obtained from inverse-variance fixed-effect meta-analyses. Heterogeneity p values are from Cochran's Q test.
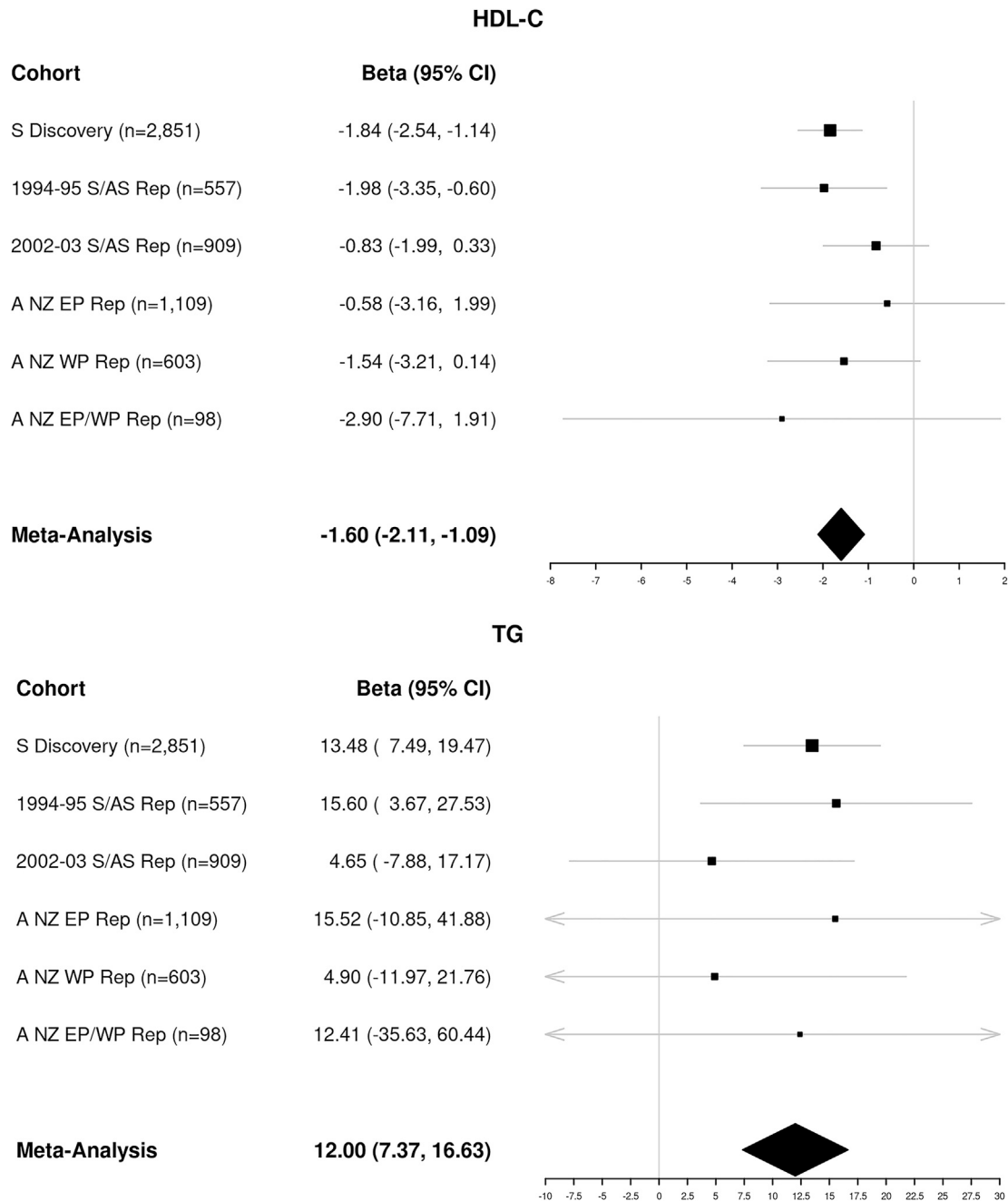[a]Results with p <0.05.

located in exon 4 of 11 in *BTNL9* and has strong mRNA degradation predicted (NMDetective score 0.6[45,46]). While *BTNL9* is not constrained/predicted to be loss-of-function intolerant (predicted LoF observed/expected = 0.94, gnomAD v.2.1.1), the role of the BTNL9 protein in human disease is still unknown.

Despite the T allele of rs200884524 having a very low frequency in other populations, the association between variants in this region and lipid levels does not appear to be Polynesian specific. Results from studies of the UKBiobank and the Global Lipids Genetics Consortium demonstrate evidence of association in this region as well. Specifically, an intron variant in nearby *BTNL3* (∼61 kb away from rs200884524) was significantly associated with HDL-C in a GWAS of ∼459,000 individuals of European descent (rs138354839, $p_{HDL-C} = 9.70 \times 10^{-13}$).[47] Additionally, variants in and upstream of *BTNL8* (∼128–169 kb away from rs200884524) were associated with total and HDL-C in a large, multiancestry meta-analysis (rs138692142, $p_{HDL-C} = 8.09 \times 10^{-17}$; rs188238483 $p_{TC} = 3.39 \times 10^{-9}$).[6] Each of these variants is rare in the Samoan discovery cohort presented here (MAF = 0.009, 0.0003, and 0.00003, respectively). The positions of these variants in relationship to the stop gain are depicted in Figure 2.

Furthermore, there were significant associations between pLoF variants in *BTNL9* and both apolipoprotein A and HDL-C observed in gene-based studies of European-ancestry individuals also from the UKBiobank.[48–50] Both associations are driven by a single variant (rs367635312, $p_{ApoA} = 1.93 \times 10^{-9}$, $p_{HDL-C} = 1.48 \times 10^{-7}$, MAF = 0.0125 in non-Finnish Europeans). This variant is also rare in the Samoan discovery cohort (MAF = 0.00016).

Interestingly, the region also harbors a ∼55-kb common deletion with consequence in *BTNL8* and *BTNL3* (DEL_5_65831 gnomAD SVs v.2.1). This deletion is common in those of Asian, American, and European descent

HDL-C

Cohort	Beta (95% CI)
S Discovery (n=2,851)	-1.84 (-2.54, -1.14)
1994-95 S/AS Rep (n=557)	-1.98 (-3.35, -0.60)
2002-03 S/AS Rep (n=909)	-0.83 (-1.99, 0.33)
A NZ EP Rep (n=1,109)	-0.58 (-3.16, 1.99)
A NZ WP Rep (n=603)	-1.54 (-3.21, 0.14)
A NZ EP/WP Rep (n=98)	-2.90 (-7.71, 1.91)
Meta-Analysis	**-1.60 (-2.11, -1.09)**

TG

Cohort	Beta (95% CI)
S Discovery (n=2,851)	13.48 (7.49, 19.47)
1994-95 S/AS Rep (n=557)	15.60 (3.67, 27.53)
2002-03 S/AS Rep (n=909)	4.65 (-7.88, 17.17)
A NZ EP Rep (n=1,109)	15.52 (-10.85, 41.88)
A NZ WP Rep (n=603)	4.90 (-11.97, 21.76)
A NZ EP/WP Rep (n=98)	12.41 (-35.63, 60.44)
Meta-Analysis	**12.00 (7.37, 16.63)**

Figure 1.   Meta-analysis results of rs200884524 for high-density lipoprotein cholesterol (HDL-C) and triglycerides (TG)
Effect estimates (beta) and 95% confidence intervals (CIs) for HDL-C (mg/dL) and TG (mg/dL) are given for each cohort (abbreviations: Samoan [S], American Samoan [AS], Replication [Rep], Aotearoa New Zealand [An NZ], Eastern Polynesian [EP], Western Polynesian [WP]). Meta analysis results were obtained using inverse-variance fixed-effects meta-analysis. Arrows indicate CIs that have been truncated for plotting.

but rare in African and Oceanic (Papuan and Melanesian) populations.[51] The frequency of this deletion in Polynesians is unknown as Polynesians have differing genetic ancestries from Papuans and Melanesians. Functional analysis has shown that this deletion down-regulates *BTNL9* in lymphoblastoid cell lines.[51] This deletion is in high LD ($r^2 > 0.80$) with 26 SNVs in the UKBiobank.[50,52] The SNV in very high LD with the deletion (rs72494581, LD with deletion 0.97) showed minimal evidence of asso-

ciation with HDL-C in the Pan-UKBiobank ($p = 0.056$ in meta-analysis across all populations, $p = 0.037$ in individuals of European descent).[50,53] Because of the low allele frequency of rs200884524 in non-Polynesian populations, it is unknown if the stop-gain variant is in LD with this deletion. Further work is needed to explore the interplay of this deletion with rs200884524 to examine the causal mechanism and resulting biological impact on lipids.
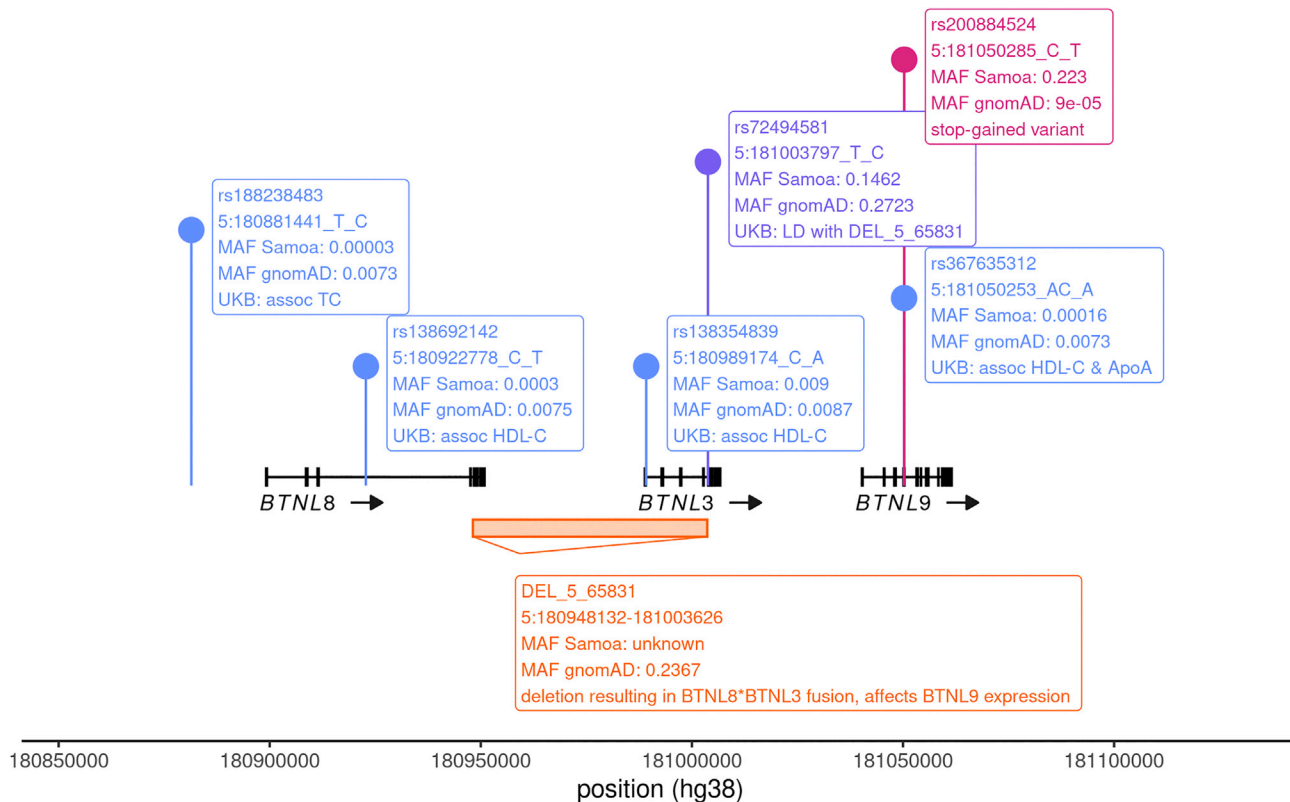
Figure 2. Variants of interest in *BTNL9* region
Variant IDs, position (build hg38), reference and alternate alleles, and minor allele frequency (MAF) in the Samoan discovery cohort and in gnomAD are given for the stop-gained variant in *BTNL9* identified in the Samoan discovery cohort (pink), 4 variants with lipid associations in the UK Biobank (UKB; blue), a deletion upstream of *BTNL9* (orange), and a variant in high linkage disequilibrium (LD) with the deletion in UKB (purple). *BTNL8*, *BTNL3*, and *BTNL9* coordinates for MANE transcripts are plotted for reference, with black bars representing exonic regions. For DEL_5_65831, hg38 coordinates were obtained using LiftOver (https://genome.ucsc.edu/cgi-bin/hgLiftOver) on the positions defined in gnomAD SV v.2.1 (hg19).

Little is known about the function of *BTNL9* and the mechanism by which it may impact serum lipids; however, it has been previously associated with cardiomyopathy,[54–56] pre-eclampsia,[57] and various cancers.[58,59] *BTNL9* has a cell- and tissue-specific expression pattern—it is most highly expressed in human B cells of adipose tissue, lung, thymus, spleen, and heart.[60] Moreover, BTNL9 is localized to the plasma membrane and binds to immune cell surfaces.[61] BTNL9 belongs to the butyrophilin (BTN) family of membrane proteins, which is part of a superfamily of immunoglobulin (Ig) receptors.[62] Several studies suggest a role of BTN proteins in inflammatory and immunological functions,[61,63–72] supported by the shared extracellular characteristics and homology of the BTN proteins with the B7 family (which includes ligands and receptors for T cell activity).[62] *BTNL9* is member of a BTN subfamily along with *BTNL3* and *BTNL8*, which all diverged from a common ancestor.[62] While *BTNL3* and *BTNL8* are primate specific, *BTNL9* is conserved across several mammalian species; however, the protein structure differs across species. Human BTNL9 has two substitution mutations that result in the loss of the IgC domain that is present in other mammals.[62] In mice, the receptor to Btnl9 was shown to be expressed in immune cells, supporting the hypothe-

sized immune function.[62,73] Future functional work is needed to identify the biological role of BTNL9 in lipid and lipoprotein metabolism, immune function, and, ultimately, cardiovascular (and potentially cancer) risk.

Additional work is necessary to characterize the relationship between variation in *BTNL9* and lipid levels, especially analyzing the impact of the nearby deletion in *BTNL8/BTNL3*, as well as any potential interaction between immune cells and lipid transport pathways. However, these findings fine map the suggestive association in 5q35 to a stop-gain variant in *BTNL9*, providing evidence of a new contributor to the genetic architecture of lipids. This work highlights the importance of measuring association in non-European populations as it can provide insight not only into population-specific findings and benefits but also into cross-population associations, which may lead to therapeutic targets able to benefit multiple population groups.

Data and code availability

The discovery cohort data used for this study are available through dbGaP (accession numbers: phs000914.v1.p1 and phs000972.v5.p1). The Samoan/American Samoan and Aotearoa New Zealand replication cohorts' data have not

been deposited in a public repository because participants did not give consent for data sharing. Code used for data analysis will be made available upon request.

References

1. Rehuher, D., Hishinuma, E.S., Goebert, D.A., and Palafox, N.A. (2021). A historical and contemporary review of the contextualization and social determinants of Health of Micronesian migrants in the United States. Hawaii J. Health Soc. Welf. *80* (9 Suppl 1), 88–101.

2. Mau, M.K., Sinclair, K., Saito, E.P., Baumhofer, K.N., and Kaholokula, J.K. (2009). Cardiometabolic Health disparities in native hawaiians and other pacific islanders. Epidemiol. Rev. *31*, 113–129.

3. Grey, C., Wells, S., Riddell, T., Pylypchuk, R., Marshall, R., Drury, P., Elley, R., Ameratunga, S., Gentles, D., Erick-Peletiy, S., et al. (2010). A comparative analysis of cardiovascular disease risk profiles of five Pacific ethnic groups assessed in New Zealand primary care practice: PREDICT CVD-13. N. Z. Med. J. *123*, 41–52.

4. Cameron, V.A., Faatoese, A.F., Gillies, M.W., Robertson, P.J., Huria, T.M., Doughty, R.N., Whalley, G.A., Richards, M.A., Troughton, R.W., Tikao-Mason, K.N., et al. (2012). A cohort study comparing cardiovascular risk factors in rural Māori, urban Māori and non-Māori communities in New Zealand. BMJ Open *2*, e000799.

5. Sun, H., Lin, M., Russell, E.M., Minster, R.L., Chan, T.F., Dinh, B.L., Naseri, T., Reupena, M.S., Lum-Jones, A., et al.; Samoan Obesity, Lifestyle, and Genetic Adaptations OLaGA Study Group (2021). The impact of global and local Polynesian genetic ancestry on complex traits in Native Hawaiians. PLoS Genet. *17*, e1009273.

6. Graham, S.E., Clarke, S.L., Wu, K.H.H., Kanoni, S., Zajac, G.J.M., Ramdas, S., Surakka, I., Ntalla, I., Vedantam, S., Winkler, T.W., et al. (2021). The power of genetic diversity in genome-wide association studies of lipids. Nature *600*, 675–679.

7. Kristiansson, K., Naukkarinen, J., and Peltonen, L. (2008). Isolated populations and complex disease gene identification. Genome Biol. *9*, 109.

8. Friedlaender, J.S., Friedlaender, F.R., Reed, F.A., Kidd, K.K., Kidd, J.R., Chambers, G.K., Lea, R.A., Loo, J.H., Koki, G., Hodgson, J.A., et al. (2008). The genetic structure of pacific islanders. PLoS Genet. *4*, e19.

9. Tsai, H.J., Sun, G., Smelser, D., Viali, S., Tufa, J., Jin, L., Weeks, D.E., McGarvey, S.T., and Deka, R. (2004). Distribution of genome-wide linkage disequilibrium based on microsatellite loci in the Samoan population. Hum. Genom. *1*, 327–334.

10. Carlson, J.C., Weeks, D.E., Hawley, N.L., Sun, G., Cheng, H., Naseri, T., Reupena, M.S., Tuitele, J., Deka, R., McGarvey, S.T., and Minster, R.L. (2021). Genome-wide association studies in Samoans give insight into the genetic architecture of fasting serum lipid levels. J. Hum. Genet. *66*, 111–121.

11. Lowe, J.K., Maller, J.B., Pe'er, I., Neale, B.M., Salit, J., Kenny, E.E., Shea, J.L., Burkhardt, R., Smith, J.G., Ji, W., et al. (2009). Genome-wide association studies in an isolated founder population from the Pacific Island of Kosrae. Gibson G. PLoS Genet. *5*, e1000365.

12. Hawley, N.L., Minster, R.L., Weeks, D.E., Viali, S., Reupena, M.S., Sun, G., Cheng, H., Deka, R., and Mcgarvey, S.T. (2014). Prevalence of adiposity and associated cardiometabolic risk factors in the Samoan genome-wide association study. Am. J. Hum. Biol. *26*, 491–501.

13. Minster, R.L., Hawley, N.L., Su, C.T., Sun, G., Kershaw, E.E., Cheng, H., Buhule, O.D., Lin, J., Reupena, M.S., Viali, S., et al. (2016). A thrifty variant in CREBRF strongly influences body mass index in Samoans. Nat. Genet. *48*, 1049–1054.

14. Friedewald, W.T., Levy, R.I., and Fredrickson, D.S. (1972). Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin. Chem. *18*, 499–502.

15. Laurie, C.C., Doheny, K.F., Mirel, D.B., Pugh, E.W., Bierut, L.J., Bhangale, T., Boehm, F., Caporaso, N.E., Cornelis, M.C., Edenberg, H.J., et al. (2010). Quality control and quality assurance in genotypic data for genome-wide association studies. Genet. Epidemiol. *34*, 591–602.

16. Loh, P.R., Danecek, P., Palamara, P.F., Fuchsberger, C., A Reshef, Y., K Finucane, H., Schoenherr, S., Forer, L., McCarthy, S., Abecasis, G.R., et al. (2016). Reference-based phasing using the haplotype reference Consortium panel. Nat. Genet. *48*, 1443–1448.

17. Deelen, P., Bonder, M.J., Van Der Velde, K.J., Westra, H.J., Winder, E., Hendriksen, D., Franke, L., and Swertz, M.A. (2014). Genotype harmonizer: automatic strand alignment and format conversion for genotype data integration. BMC Res. Notes *7*, 901.

18. Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A.E., Kwong, A., Vrieze, S.I., Chew, E.Y., Levy, S., McGue, M., et al. (2016). Next-generation genotype imputation service and methods. Nat. Genet. *48*, 1284–1287.

19. Gogarten, S.M., Sofer, T., Chen, H., Yu, C., Brody, J.A., Thornton, T.A., et al. (2019). Genetic association testing using the

GENESIS R/Bioconductor package. Bioinformatics *35*, 5346–5348. https://doi.org/10.1093/bioinformatics/btz567.

20. Conomos, M.P., Gogarten, S.M., Brown, L., Chen, H., Rice, K., Sofer, T., et al. (2019). GENESIS: GENetic EStimation and Inference in Structured samples (GENESIS): statistical methods for analyzing genetic data from samples with population structure and/or relatedness. https://github.com/UW-GAC/GENESIS.

21. Conomos, M.P., Miller, M.B., and Thornton, T.A. (2015). Robust inference of population structure for ancestry prediction and correction of stratification in the presence of relatedness. Genet. Epidemiol. *39*, 276–293.

22. Therneau, T.M. (2018). Coxme: mixed effects cox models. R package version 2.2-10. https://cran.r-project.org/package=coxme.

23. Sofer, T., Zheng, X., Gogarten, S.M., Laurie, C.A., Grinde, K., Shaffer, J.R., Shungin, D., O'Connell, J.R., Durazo-Arvizo, R.A., Raffield, L., et al. (2019). A fully adjusted two-stage procedure for rank-normalization in genetic association studies. Genet. Epidemiol. *43*, 263–275.

24. McCaw, Z. (2019). RNOmni: rank normal transformation omnibus test. R package version 0.7.1. https://cran.r-project.org/package=RNOmni.

25. Kichaev, G., Yang, W.Y., Lindstrom, S., Hormozdiari, F., Eskin, E., Price, A.L., Kraft, P., and Pasaniuc, B. (2014). Integrating functional data to prioritize causal variants in statistical fine-mapping studies. PLoS Genet. *10*, e1004722.

26. McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P., and Cunningham, F. (2016). The ensembl variant effect predictor. Genome Biol. *17*, 122.

27. Harris, D.N., Kessler, M.D., Shetty, A.C., Weeks, D.E., Minster, R.L., Browning, S., Cochrane, E.E., Deka, R., Hawley, N.L., Reupena, M.S., et al. (2020). Evolutionary history of modern Samoans. Proc. Natl. Acad. Sci. USA *117*, 9458–9465.

28. Szpiech, Z.A., and Hernandez, R.D. (2014). Selscan: an efficient multithreaded program to perform EHH-based scans for positive selection. Mol. Biol. Evol. *31*, 2824–2827.

29. Chin-Hong, P.V., and McGarvey, S.T. (1996). Lifestyle incongruity and adult blood pressure in Western Samoa. Psychosom. Med. *58*, 131–137.

30. Galanis, D.J., McGarvey, S.T., Sobal, J., Bausserman, L., and Levinson, P.D. (1995). Relations of body fat and fat distribution to the serum lipid, apolipoprotein and insulin concentrations of Samoan men and women. Int. J. Obes. Relat. Metab. Disord. *19*, 731–738.

31. McGarvey, S.T., Levinson, P.D., Bausser-Man, L., Galanis, D.J., and Hornick, C.A. (1993). Population change in adult obesity and blood lipids in American Samoa from 1976–1978 to 1990. Am. J. Hum. Biol. *5*, 17–30.

32. Åberg, K., Dai, F., Viali, S., Tuitele, J., Sun, G., Indugula, S.R., Deka, R., Weeks, D.E., and McGarvey, S.T. (2009). Suggestive linkage detected for blood pressure related traits on 2q and 22q in the population on the Samoan islands. BMC Med. Genet. *10*, 107.

33. Åberg, K., Dai, F., Sun, G., Keighley, E., Indugula, S.R., Bausser-man, L., Viali, S., Tuitele, J., Deka, R., Weeks, D.E., and McGarvey, S.T. (2008). A genome-wide linkage scan identifies multiple chromosomal regions influencing serum lipid levels in the population on the Samoan islands. J. Lipid Res. *49*, 2169–2178.

34. Åberg, K., Dai, F., Sun, G., Keighley, E.D., Indugula, S.R., Roberts, S.T., Zhang, Q., Smelser, D., Viali, S., Tuitele, J., et al.

(2009). Susceptibility loci for adiposity phenotypes on 8p, 9p, and 16q in American Samoa and Samoa. Obesity *17*, 518–524.

35. Dai, F., Keighley, E.D., Sun, G., Indugula, S.R., Roberts, S.T., Åberg, K., Smelser, D., Tuitele, J., Jin, L., Deka, R., et al. (2007). Genome-wide scan for adiposity-related phenotypes in adults from American Samoa. Int. J. Obes. *31*, 1832–1842.

36. Dai, F., Sun, G., Åberg, K., Keighley, E.D., Indugula, S.R., Roberts, S.T., Smelser, D., Viali, S., Jin, L., Deka, R., et al. (2008). A whole genome linkage scan identifies multiple chromosomal regions influencing adiposity-related traits among Samoans. Ann. Hum. Genet. *72*, 780–792.

37. Rasheed, H., Hsu, A., Dalbeth, N., Stamp, L.K., McCormick, S., and Merriman, T.R. (2014). The relationship of apolipoprotein B and very low density lipoprotein triglyceride with hyperuricemia and gout. Arthritis Res. Ther. *16*, 495.

38. Krishnan, M., Major, T.J., Topless, R.K., Dewes, O., Yu, L., Thompson, J.M.D., McCowan, L., de Zoysa, J., Stamp, L.K., Dalbeth, N., et al. (2018). Discordant association of the CREBRF rs373863828 A allele with increased BMI and protection from type 2 diabetes in Māori and Pacific (Polynesian) people living in Aotearoa/New Zealand. Diabetologia *61*, 1603–1613.

39. Phipps-Green, A.J., Merriman, M.E., Topless, R., Altaf, S., Montgomery, G.W., Franklin, C., Jones, G.T., van Rij, A.M., White, D., Stamp, L.K., et al. (2016). Twenty-eight loci that influence serum urate levels: analysis of association with gout. Ann. Rheum. Dis. *75*, 124–130.

40. Purcell, S.C.C.. PLINK 1.9. www.cog-genomics.org/plink/1.9/.

41. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience *4*.

42. Lumley, T. (2018). rmeta: meta-Analysis. R package version 3.0. https://cran.r-project.org/package=rmeta.

43. Metcalfe, L.K., Krishnan, M., Turner, N., Yaghootkar, H., Merry, T.L., Dewes, O., Hindmarsh, J.H., De Zoysa, J., Dalbeth, N., Stamp, L.K., et al. (2020). The Māori and Pacific specific CREBRF variant and adult height. Int. J. Obes. *44*, 748–752.

44. Carlson, J.C., Rosenthal, S.L., Russell, E.M., Hawley, N.L., Sun, G., Cheng, H., Naseri, T., Reupena, M.S., Tuitele, J., Deka, R., et al. (2020). A missense variant in CREBRF is associated with taller stature in Samoans. Am. J. Hum. Biol. *32*, e23414.

45. Lindeboom, R.G.H., Vermeulen, M., Lehner, B., and Supek, F. (2019). The impact of nonsense-mediated mRNA decay on genetic disease, gene editing and cancer immunotherapy. Nat. Genet. *51*, 1645–1651.

46. Lindeboom, R.G.H. (2019). NMDetective. https://doi.org/10.6084/m9.figshare.7803398.v1.

47. Richardson, T.G., Sanderson, E., Palmer, T.M., Ala-Korpela, M., Ference, B.A., Davey Smith, G., and Holmes, M.V. (2020). Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: a multivariable Mendelian randomisation analysis. PLoS Med. *17*, e1003062.

48. (2021). Genebass. https://genebass.org/.

49. Karczewski, K.J., Solomonson, M., Chao, K.R., Goodrich, J.K., Tiao, G., Lu, W., Riley-Gillis, B.M., Tsai, E.A., Kim, H.I., Zheng, X., and Rahimov, F. (2021). Systematic single-variant and gene-based association testing of 3,700 phenotypes in 281,850 UK Biobank exomes. https://www.medrxiv.org/content/10.1101/2021.06.19.21259117v3.

50. Neale, B.M. (2018). UK Biobank GWAS round 2. http://www.nealelab.is/uk-biobank/.

51. Aigner, J., Villatoro, S., Rabionet, R., Roquer, J., Jiménez-Conde, J., Martí, E., and Estivill, X. (2013). A common 56-kilobase deletion in a primate-specific segmental duplication creates a novel butyrophilin-like protein. BMC Genet. 14, 61.

52. Collins, R.L., Brand, H., Karczewski, K.J., Zhao, X., Alföldi, J., Francioli, L.C., Khera, A.V., Lowther, C., Gauthier, L.D., Wang, H., et al. (2020). A structural variation reference for medical and population genetics. Nature 581, 444–451.

53. (2020). Pan-UKB team. https://pan.ukbb.broadinstitute.org.

54. Koczor, C.A., Lee, E.K., Torres, R.A., Boyd, A., Vega, J.D., Uppal, K., Yuan, F., Fields, E.J., Samarel, A.M., and Lewis, W. (2013). Detection of differentially methylated gene promoters in failing and nonfailing human left ventricle myocardium using computation analysis. Physiol. Genom. 45, 597–605.

55. Staley, J.R., Blackshaw, J., Kamat, M.A., Ellis, S., Surendran, P., Sun, B.B., Paul, D.S., Freitag, D., Burgess, S., Danesh, J., et al. (2016). PhenoScanner: a database of human genotype–phenotype associations. Bioinformatics 32, 3207–3209.

56. Kamat, M.A., Blackshaw, J.A., Young, R., Surendran, P., Burgess, S., Danesh, J., Butterworth, A.S., and Staley, J.R. (2019). PhenoScanner V2: an expanded tool for searching human genotype–phenotype associations. Bioinformatics 35, 4851–4853.

57. Kaartokallio, T., Cervera, A., Kyllönen, A., Laivuori, K., Kere, J., Laivuori, H.; and FINNPEC Core Investigator Group (2015). Gene expression profiling of pre-eclamptic placentae by RNA sequencing. Sci. Rep. 5, 14107.

58. Lebrero-Fernández, C., Wenzel, U.A., Akeus, P., Wang, Y., Strid, H., Simrén, M., Gustavsson, B., Börjesson, L.G., Cardell, S.L., Öhman, L., et al. (2016). Altered expression of Butyrophilin (BTN) and BTN-like (BTNL) genes in intestinal inflammation and colon cancer. Immun. Inflamm. Dis. 4, 191–200.

59. Hsu, Y.L., Hung, J.Y., Lee, Y.L., Chen, F.W., Chang, K.F., Chang, W.A., Tsai, Y.M., Chong, I.W., and Kuo, P.L. (2017). Identification of novel gene expression signature in lung adenocarcinoma by using next-generation sequencing data and bioinformatics analysis. Oncotarget 8, 104831–104854.

60. Arnett, H.A., Escobar, S.S., and Viney, J.L. (2009). Regulation of costimulation in the era of butyrophilins. Cytokine 46, 370–375.

61. Yamazaki, T., Goya, I., Graf, D., Craig, S., Martin-Orozco, N., and Dong, C. (2010). A butyrophilin family member critically inhibits T cell activation. J. Immunol. 185, 5907–5914.

62. Afrache, H., Gouret, P., Ainouche, S., Pontarotti, P., and Olive, D. (2012). The butyrophilin (BTN) gene family: from milk fat to the regulation of the immune response. Immunogenetics 64, 781–794.

63. Smith, I.A., Knezevic, B.R., Ammann, J.U., Rhodes, D.A., Aw, D., Palmer, D.B., Mather, I.H., and Trowsdale, J. (2010). BTN1A1, the mammary gland butyrophilin, and BTN2A2 are both inhibitors of T cell activation. J Immunol. 184, 3514–3525.

64. Nguyen, T., Liu, X.K., Zhang, Y., and Dong, C. (2006). BTNL2, a butyrophilin-like molecule that functions to inhibit T cell activation. J. Immunol. 176, 7354–7360.

65. Arnett, H.A., Escobar, S.S., Gonzalez-Suarez, E., Budelsky, A.L., Steffen, L.A., Boiani, N., Zhang, M., Siu, G., Brewer, A.W., and Viney, J.L. (2007). BTNL2, a butyrophilin/B7-like molecule, is a negative costimulatory molecule modulated in intestinal inflammation. J. Immunol. 178, 1523–1533.

66. Suzuki, H., Ota, M., Meguro, A., Katsuyama, Y., Kawagoe, T., Ishihara, M., Asukata, Y., Takeuchi, M., Ito, N., Shibuya, E., et al. (2012). Genetic characterization and susceptibility for sarcoidosis in Japanese patients: risk factors of BTNL2 gene polymorphisms and HLA class II alleles. Invest. Ophthalmol. Vis. Sci. 53, 7109–7115.

67. Gang, Q., Bettencourt, C., Machado, P., Hanna, M.G., and Houlden, H. (2014). Sporadic inclusion body myositis: the genetic contributions to the pathogenesis. Orphanet J. Rare Dis. 9, 88.

68. Malcherek, G., Mayr, L., Roda-Navarro, P., Rhodes, D., Miller, N., and Trowsdale, J. (2007). The B7 homolog butyrophilin BTN2A1 is a novel ligand for DC-SIGN. J. Immunol. 179, 3804–3811.

69. Johnson, C.M., Traherne, J.A., Jamieson, S.E., Tremelling, M., Bingham, S., Parkes, M., Blackwell, J.M., and Trowsdale, J. (2007). Analysis of the BTNL2 truncating splice site mutation in tuberculosis, leprosy and Crohn's disease. Tissue Antigens 69, 236–241.

70. Bas, A., Swamy, M., Abeler-Dörner, L., Williams, G., Pang, D.J., Barbee, S.D., and Hayday, A.C. (2011). Butyrophilin-like 1 encodes an enterocyte protein that selectively regulates functional interactions with T lymphocytes. Proc. Natl. Acad. Sci. USA 108, 4376–4381.

71. Martin-Orozco, N., and Dong, C. (2007). Inhibitory costimulation and anti-tumor immunity. Semin. Cancer Biol. 17, 288–298.

72. Driessens, G., Kline, J., and Gajewski, T.F. (2009). Costimulatory and coinhibitory receptors in anti-tumor immunity. Immunol. Rev. 229, 126–144.

73. Abeler-Dörner, L., Swamy, M., Williams, G., Hayday, A.C., and Bas, A. (2012). Butyrophilins: an emerging family of immune regulators. Trends Immunol. 33, 34–41.