An *LDLR* missense variant poses high risk of familial hypercholesterolemia in 30% of Greenlanders and offers potential of early cardiovascular disease intervention

Emil Jørsboe,[1,2,13,15] Mette K. Andersen,[2,13] Line Skotte,[3] Frederik F. Stæger,[1] Nils J. Færgeman,[4] Kristian Hanghøj,[1] Cindy G. Santander,[1] Ninna K. Senftleber,[1,5] Lars J. Diaz,[5] Maria Overvad,[5] Ryan K. Waples,[1] Frank Geller,[3] Peter Bjerregaard,[6] Mads Melbye,[7,8,9,10] Christina V.L. Larsen,[6,11] Bjarke Feenstra,[3] Anders Koch,[3,11] Marit E. Jørgensen,[5,6,11,14,*] Niels Grarup,[2,14,*] Ida Moltke,[1,14,*] Anders Albrechtsen,[1,14,*] and Torben Hansen[2,12,14,*]

Summary

The common Arctic-specific *LDLR* p.G137S variant was recently shown to be associated with elevated lipid levels. Motivated by this, we aimed to investigate the effect of p.G137S on metabolic health and cardiovascular disease risk among Greenlanders to quantify its impact on the population. In a population-based Greenlandic cohort (n = 5,063), we tested for associations between the p.G137S variant and metabolic health traits as well as cardiovascular disease risk based on registry data. In addition, we explored the variant's impact on plasma NMR measured lipoprotein concentration and composition in another Greenlandic cohort (n = 1,629); 29.5% of the individuals in the cohort carried at least one copy of the p.G137S risk allele. Furthermore, 25.4% of the heterozygous and 54.7% of the homozygous carriers had high levels (>4.9 mmol/L) of serum LDL cholesterol, which is above the diagnostic level for familial hypercholesterolemia (FH). Moreover, p.G137S was associated with an overall atherosclerotic lipid profile, and increased risk of ischemic heart disease (HR [95% CI], 1.51 [1.18–1.92], p = 0.00096), peripheral artery disease (1.69 [1.01–2.82], p = 0.046), and coronary operations (1.78 [1.21–2.62], p = 0.0035). Due to its high frequency and large effect sizes, p.G137S has a marked population-level impact, increasing the risk of FH and cardiovascular disease for up to 30% of the Greenlandic population. Thus, p.G137S is a potential marker for early intervention in Arctic populations.

Introduction

Cardiovascular disease (CVD) is the number one cause of death in many populations worldwide.[1] Also among Greenlanders, the CVD burden is high, and the CVD prevalence is likely to increase in the future, due to increasing life expectancy and changing lifestyle.[2,3] Therefore, additional focus on prevention and treatment of CVD is essential, also in the Greenlandic population.

CVD is tightly linked to alterations in the lipid homeostasis, including elevated levels of low-density lipoprotein (LDL) cholesterol.[4–7] Alterations in the lipid homeostasis is associated with lifestyle factors, such as unhealthy diet, physical inactivity, smoking, and alcohol intake, as well as genetic predisposition. The genetic predisposition can be monogenic, where a single variant has a large effect on lipid levels causing familial hypercholesterolemia (FH),[8] or polygenic where a large number of common var-

iants each contributes with a small increase in the risk of dyslipidemia. The monogenic forms of dyslipidemia are associated with increased risk of CVD, especially ischemic heart disease.[9] On average, individuals with FH die 19 years earlier than the general population from a CVD event.[10,11] Importantly, monogenic FH mutations seem to impose a 2-fold higher risk of CVD, compared with higher LDL cholesterol caused by polygenic variants, likely because monogenic variants manifest earlier in life than polygenic variants, leading to a greater cumulative LDL cholesterol exposure.[12,13] Early treatment with cholesterol-lowering drugs has been shown to improve health outcomes for FH patients, and current guidelines recommend cholesterol-lowering treatment beginning from the age of 8 to 10 years.[9,14,15] Hence, it is of great importance to identify and treat individuals carrying FH causing mutations early in life to reduce morbidity and mortality. The most common cause of FH is mutations in the gene (*LDLR*) encoding

[1]Bioinformatics Centre, Department of Biology, University of Copenhagen, Copenhagen, Denmark; [2]Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; [3]Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark; [4]Villum Center for Bioanalytical Sciences, Department of Biochemistry and Molecular Biology, University of Southern Denmark, Denmark; [5]Steno Diabetes Center Copenhagen, Gentofte, Denmark; [6]National Institute of Public Health, University of Southern Denmark, Copenhagen, Denmark; [7]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark; [8]Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA; [9]K.G.Jebsen Center for Genetic Epidemiology, Norwegian University of Science and Technology (NTNU), Norway; [10]Center for Fertility and Health, Norwegian Institute of Public Health, Norway; [11]Greenland Centre for Health Research, University of Greenland, Nuuk, Greenland; [12]Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark
[13]These authors contributed equally
[14]These authors contributed equally
[15]Lead contact
*Correspondence: maej@peqqik.gl (M.E.J.), niels.grarup@sund.ku.dk (N.G.), ida@binf.ku.dk (I.M.), albrecht@binf.ku.dk (A.A.), torben.hansen@sund.ku.dk (T.H.)
https://doi.org/10.1016/j.xhgg.2022.100118

the LDL receptor (LDLR). This cell surface receptor mediates the uptake of LDL-cholesterol particles from the blood, primarily into the liver. The hepatic LDL uptake suppresses de novo synthesis of cholesterol, further reducing the concentration of circulating cholesterol.[16]

In Arctic populations, a common *LDLR* missense variant (p.G137S; rs730882082) has been identified. This variant is located in exon 4, which encodes the ligand binding domain of the receptor, and has been predicted to be deleterious,[17] and to reduce LDLR ligand binding by around 60%.[18] The p.G137S variant has been shown to be associated with markedly elevated levels of LDL cholesterol, total cholesterol, and apolipoprotein B,[18] yet it is not known if the variant is associated with increased risk of CVD. Motivated by this, we aimed to investigate the influence of the common Arctic-specific *LDLR* p.G137S missense variant on the metabolic and cardiovascular health in Greenlanders.

Materials and methods

Study cohorts
The cohort used in all analyses, except for the NMR analysis, is composed of Greenlanders living in Greenland from The Population Study in Greenland 1999 (B99; n = 1,401) and The Inuit Health in Transition Study (IHIT; n = 3,115), as well as Greenlanders living in Denmark (BBH; n = 547), all collected as a part of general population health surveys of the Greenlandic population, which took place in 1999–2001 and 2005–2010.[19,20] Altogether, 295 individuals participated in both B99 and IHIT, and were treated as a separate survey when adjusting for survey in the statistical analyses.

For the NMR analysis, we used data from a separate Greenlandic cohort collected in 2013 as part of a population-based sample (n = 1,629), as previously described.[21]

Genetic data
The *LDLR* p.G137S variant (rs730882082) was genotyped using the KASP Genotyping Assay (LGC Genomics). The genotyping call rate was 99.42%, and 0 mismatches were observed in 361 samples genotyped in duplicate. We used previously published genome-wide genotype data from the same individuals to obtain estimates of a genetic similarity matrix needed in the association analyses. This genome-wide genotype data was generated using the Illumina MetaboChip (Illumina, San Diego, CA[22]) and consisted of data from 4,674 of the Greenlandic participants (IHIT, 2,791; B99, 1336; BBH, 547) and 115,182 SNPs after quality control. Both the genotyping procedure and quality control for the SNP chip dataset have previously been described in detail.[23] For the *LDLR* p.G137S variant genotyping and the MetaboChip data there was an overlap of 4,653 individuals. We based our analyses of metabolic phenotypes on this dataset.

For the cohort with NMR data, the genome-wide genotype data were generated using the Illumina OmniExpressExome SNP chip (Illumina). After quality control, this cohort comprised 1,570 participants and genotypes of 643,734 SNPs.[21]

Analysis of metabolic phenotypes
Since the Greenlandic study cohorts contain admixed individuals with both Inuit and European ancestry, and numerous close

relatives,[24] we used a linear mixed model to perform the association analyses. This type of model takes admixture and relatedness into account by including them as a random effect. Specifically, we used the linear mixed model software tool GEMMA (v0.95alpha)[25] for quantitative traits and GMMAT[26] for binary traits. To perform the analyses we first estimated a genetic similarity matrix for all participants with both genotype and phenotype data available from quality controlled genome-wide SNP chip data. The estimation was performed by applying GEMMA to standardized genotypes from the SNPs with minor allele frequency (MAF) of at least 5% and missingness of maximum 1%. Association tests were then performed using a score test in GEMMA and a Wald test in GMMAT and effect sizes and standard errors were estimated using a restricted maximum likelihood approach. For the quantitative traits we used a rank-based inverse normal transformation prior to analysis and we reported effect size estimates in standard deviations β_{SD}. This transformation was done independently for men and women. To get effect size estimates in the measured units, we also performed analyses of non-transformed trait values. In all tests of metabolic phenotypes we assumed an additive model and included sex, age, and survey (IHIT, B99, BBH) as covariates.

Analysis of CVD data
We used the R-package survival[27] for Cox regression analyses, adjusting for sex, age, survey, and the top 10 genetic principal components (PCs), based on the genome-wide genotype data. The top 10 PCs were included to correct for population structure. Follow-up time was calculated as years lived from birth to an event, death of other causes, emigration, or end of follow-up (December 31, 2016). A second analysis was done in the same way, but with follow-up time calculated as years from inclusion in study, and thereby excluding individuals with a CVD event of that type before the study initiation. Also a logistic regression using GMMAT, in the same way as previously described, was done for each type of CVD event defining cases as anyone with a CVD event of that type, to estimate the cross-sectional CVD risk.

Accounting for cholesterol-lowering drug intake
In general, the data presented were not adjusted for intake of cholesterol-lowering drugs, as we do not have systematic information on medication across all the surveys. However, in the IHIT survey (n = 3,115) we do have systematic information from the questionnaires. We have therefore assessed the effect of cholesterol-lowering drugs by also analyzing LDL cholesterol, total cholesterol, and CVD outcomes in the IHIT survey only where we exclude those who take cholesterol-lowering drugs (n = 141).

More detailed information on the methods used is available in the supplementary material.

Results

In the Greenlandic population, which has both Inuit and European ancestry, we estimated that the *LDLR* missense variant (p.G137S) had a MAF of 15.8%, and that the minor allele was carried by 29.5% of the population. Within the Inuit ancestry component of the population, we estimated an MAF of 22.7% and of 0.0% in the European ancestry component. In line with this, we estimated the variant to be extremely rare (MAF <0.005%) in non-Arctic populations (Table S1). Even though p.G137S is a missense

Table 1. Analyses of association between the p.G137S variant and circulating lipids and cardiovascular health markers

Trait	N	β_{SD}(SE)	β	p
Lipid profile				
Fs LDL cholesterol (mmol/L)	4,077	0.66 (0.031)	0.75	$7.3 \cdot 10^{-88}$
Fs total cholesterol (mmol/L)	4,636	0.55 (0.029)	0.69	$5.0 \cdot 10^{-73}$
Fs HDL cholesterol (mmol/L)	4,631	−0.11 (0.030)	−0.058	0.00041
Fs triglyceride (mmol/L)	4,105	0.043 (0.033)	0.025	0.19
Apolipoprotein B (g/L)	1,232	0.51 (0.057)	0.12	$1.7 \cdot 10^{-17}$
Apolipoprotein A1 (g/L)	1,232	−0.13 (0.058)	−0.034	0.032
Fs VLDL cholesterol (mmol/L)	2,086	−0.043 (0.045)	−0.0091	0.34
Remnant cholesterol (mmol/L)	3,938	0.034 (0.034)	0.0055	0.31
Cholesterol-lowering drugs^	2,870	1.54 OR (0.18)	–	0.016
Cardiovascular profile				
Pulse rate (beats per minute)	1,524	0.0023 (0.050)	0.12	0.96
Systolic blood pressure (mm Hg)	4,133	−0.014 (0.030)	−0.12	0.65
Diastolic blood pressure (mm Hg)	4,133	−0.056 (0.031)	−0.61	0.072
Albumin-to-creatinine ratio in urine (mg/L)	2,560	0.076 (0.037)	7.9	0.042
Carotid intima-media thickness (mm)	1,063	0.080 (0.039)	0.013	0.042

∧ = binary trait.
Each analysis was both run with the phenotype quantile transformed to a standard normal distribution for each sex (β_{SD}) and without any transformation (β). The p-values are from the transformed analyses. β, effect size; SD, standard deviation; SE, standard error; Fs, fasting serum.

variant, it might still act as an expression quantitative trait locus; however, we observed no association between p.G137S genotype and expression of *LDLR* RNA in blood (p = 0.22; Figure S1).

Metabolic phenotypes

We first assessed how the p.G137S variant affected a range of metabolic traits in a cohort of 4,653 Greenlanders. We confirmed a strong association between the A-allele of the p.G137S variant and elevated concentrations of circulating LDL cholesterol (β, 0.75 mmol/L, p = $7.3 \cdot 10^{-88}$, Table 1 and Figure 1A). The effect was additive for LDL cholesterol and the associated metabolic phenotypes, and similar across age groups (Figure S2), as well as degree of Inuit ancestry (Figure S3). As shown in Figure 1B, we observed that 8.0%, 25.4%, and 54.7% of non-, heterozygous, and homozygous p.G137S A-allele carriers, respectively, had high LDL cholesterol levels (>4.9 mmol/L), which is the diagnostic threshold for FH.[9] We also validated the association between p.G137S and levels of total cholesterol (β, 0.69 mmol/L, p = $5.0 \cdot 10^{-73}$, Table 1) and apolipoprotein B (β, 0.12 g/L, p = $1.7 \cdot 10^{-17}$, Table 1). Furthermore, we observed a novel association between p.G137S and lower high-density lipoprotein (HDL) cholesterol levels (β, −0.058 mmol/L, p = 0.00041, Table 1), and an increased frequency of the use of cholesterol-lowering drugs among risk A-allele carriers (odds ratio [OR] [95% confidence interval (CI)], 1.54 [1.08–2.20], p = 0.016, Table 1). To account for the effect of cholesterol-lowering

drugs in analyses of LDL cholesterol and total cholesterol, these analyses were repeated in a subset of individuals where individuals taking cholesterol-lowering drugs were excluded. For both LDL cholesterol (n = 2,625, β, 0.78 mmol/L, p = $4.2 \cdot 10^{-69}$) and total cholesterol (n = 2,644, β, 0.74 mmol/L, p = $2.4 \cdot 10^{-52}$) this resulted in higher effect estimates.

With respect to basic clinical markers of cardiovascular health, we observed nominally significantly higher levels of urinary albumin-to-creatinine ratio as well as carotid intima-media thickness (Table 1). We found no associations to phenotypes related to body composition (Table S2). Our tests were adjusted for different ancestries through a mixed model and we observed no indication of inflation of the test statistics based on p-value QQ plots for all genotyped variants in the genome (Figure S4). Moreover, the QQ plot for LDL cholesterol clearly indicated that there were no other Arctic-specific large-impact variants, which could possibly affect the association between p.G137S and LDL cholesterol levels.

Variance explained by the p.G137S variant

Given the large effect size estimates obtained for some of the lipid traits, we estimated the amount of phenotypic variance explained by the variant for these traits. Based on partial R^2 from a linear model that included PCs, sex, age, and survey, the p.G137S variant explained 11.8% of the variance for LDL cholesterol. Restricting the analyses to 696 individuals with more than 95% Inuit ancestry,
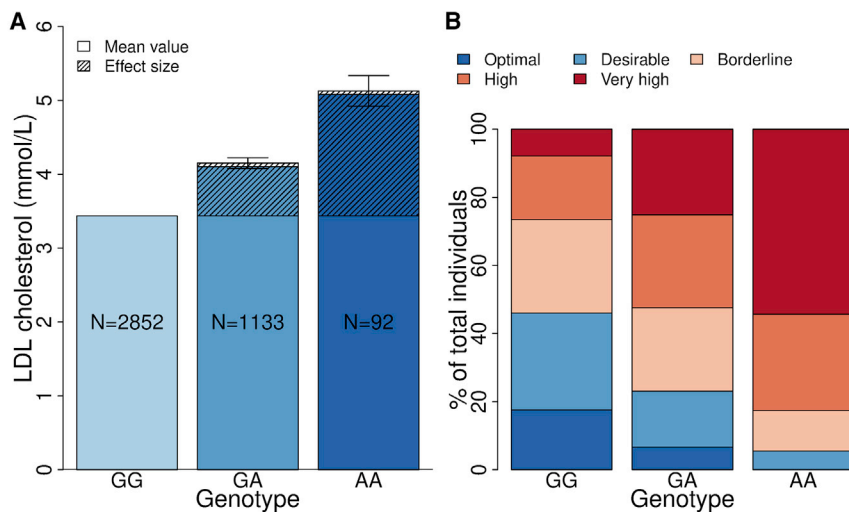
Figure 1. Effect of p.G137S variant on circulating lipid levels

(A) Mean levels of LDL cholesterol stratified per genotype, as well as estimated effect sizes from GEMMA according to a full model (hatched area) with 95% confidence intervals. The number of participants in each group is written inside the bars.

(B) For each genotype group, the proportion of different LDL cholesterol level categories is shown. The LDL cholesterol level categories are according to US guidelines[28]) with <2.6 mmol/L being Optimal, 2.6–3.3 mmol/L being Desirable, 3.4–4.0 mmol/L being Borderline high, 4.1–4.8 mmol/L being High, and ≥4.9 mmol/L being Very high.

the variance explained increased to 16.8% (Figure S5B). We also compared the impact of the variant with that of the main known clinical risk factors for elevated LDL cholesterol, namely BMI, sex, smoking, age, and waist-hip ratio. We found that the p.G137S variant explained more variance of LDL cholesterol than any of these risk factors both when including all risk factors in the model and when estimating partial R^2 from a model, which only included a single risk factor and the PCs (Figure 2A).

To enable comparisons with other variants, we also estimated the variance explained using a less flexible, but commonly used approach based on summary statistics.[31] This led to similar, but slightly lower estimates of 10.4% variance explained in all individuals and 14.9% in individuals with more than 95% Inuit ancestry only (Figure S5A). Compared with the p.G137S variant, the variants identified in a recent large European genome-wide association study[29] had much lower effect sizes and explained much less of the observed variance in LDL cholesterol levels (Figures 2B and 2C).

Analyses of NMR-based lipoprotein profile and metabolic phenotypes

To explore the effect of the p.G137S variant further, we tested for associations with NMR-based measurements of lipoprotein particles and metabolic markers in up to 1,520 Greenlanders from our separate Greenlandic cohort. We found a similar effect on overall lipid levels as in the main Greenlandic cohort (Figure S6); however, we also found that variant carriers had a significantly smaller diameter of LDL particles (β_{SD} [SE], −0.22 [0.05], p = $1.7 \cdot 10^{-6}$), and higher concentrations of all three subclasses of LDL particles (Large, 0.64 [0.04], p = $2.6 \cdot 10^{-48}$; Medium, 0.63 [0.04], p = $1.8 \cdot 10^{-46}$; Small 0.62 [0.04], p = $1.6 \cdot 10^{-45}$), intermediate-density lipoprotein (IDL) particles (0.62 [0.04], p = $4.7 \cdot 10^{-45}$), and the smallest subclasses of very low-density lipoprotein (VLDL) particles (Small, 0.18 [0.05], $1.6 \cdot 10^{-4}$; Very small, 0.49 [0.04], p = $6.6 \cdot 10^{-29}$; Figure 3).

In these particles, the content of total lipids, phospholipids, cholesterol, cholesterol esters, free cholesterol, and triglycerides were all significantly higher in p.G137S carriers (Figure S6). In line with the higher levels of LDL cholesterol, the p.G137S carriers also had higher levels of apolipoprotein B (0.50 [0.05], $2.9 \cdot 10^{-29}$; Figure 3). Moreover, p.G137S was significantly associated with higher concentration of several cholesterol subtypes, and with a pattern of HDL particles, with significantly higher concentrations of HDL_3 cholesterol (0.20 [0.05], p = $1.3 \cdot 10^{-5}$), but significantly lower concentrations of large and medium HDL particles (Large, −0.18 [0.05], p = $1.3 \cdot 10^{-5}$; Medium −0.28 [0.05], p = $5.7 \cdot 10^{-9}$; Figure 3). With respect to phospholipids and glycerides, the p.G137S carriers had significantly higher concentrations of total phosphoglycerides (0.19 [0.05], p = $3.0 \cdot 10^{-5}$), phosphatidylcholine and other cholines (0.19 [0.05], p = $3.8 \cdot 10^{-5}$), sphingomyelins (0.45 [0.04], p = $5.7 \cdot 10^{-24}$), and total cholines (0.26 [0.04], p = $4.7 \cdot 10^{-9}$; Figure 3), as well as altered levels of some classes of fatty acids. We observed no effect on the concentration of amino acids, or markers of glucose metabolism (Figure S7).

Analysis of CVD outcomes

We also analyzed if the p.G137S variant had an effect on CVD outcomes in a dataset comprising the 4,565 individuals, for whom we had both genetic and CVD data. For the p.G137S carriers, we observed a significantly higher risk of ischemic heart disease (hazard ratio [HR] [95% CI], 1.51 [1.18–1.92], p = 0.00096), peripheral artery disease (1.69 [1.01–2.82], p = 0.046), and coronary operations (1.78 [1.21–2.62], p = 0.0035; Figure 4 and Table S3) when assessing years from birth to event. For these three types of CVD outcomes, heterozygous and homozygous carriers had lower survival rates as a function of years lived compared with non-carriers (Figure 4, Table S3, Figures S8, and S9). A QQ plot of the test statistics from testing all genotyped variants in the genome showed no indications of
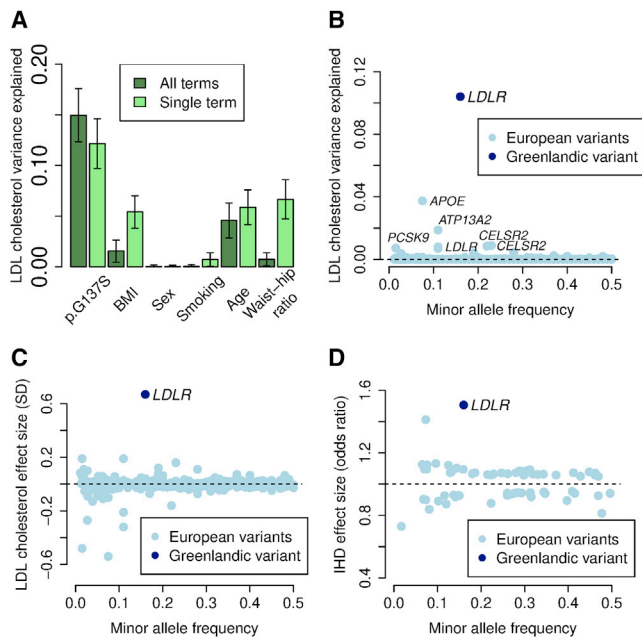
Figure 2. Effect sizes and explained variance of common variants associated with LDL cholesterol and ischemic heart disease in Europeans and the *LDLR* p.G137S variant in Greenlanders

(A) Comparison of LDL cholesterol variance explained by the p.G137S variant, BMI, sex, smoking, age, and waist-hip ratio, respectively, in a model including all terms (All terms) as well as a model including the terms one at a time (Single term) with 95% confidence intervals.

(B) Estimated LDL cholesterol variance (using quantile transformed effect sizes) explained by *LDLR* p.G137S in Greenlanders, and the genetic variants identified to be associated with LDL cholesterol levels in Europeans[29].

(C) LDL cholesterol effect sizes (SD, standard deviation) for *LDLR* p.G137S in Greenlanders, and the genetic variants identified to be associated with LDL cholesterol levels in Europeans.[29]

(D) The odds ratio for ischemic heart disease for *LDLR* p.G137S variant in Greenlanders, and genetic variants identified to be associated with increased risk of ischemic heart disease in Europeans.[30] Genetic variants with a minor allele frequency <0.01, have been omitted from the plot.

inflation of the test statistics in the Cox regression analyses (Figure S10). We observed similar HRs when assessing years since inclusion, instead of years lived (Table S4), and when analyzing CVD risk cross-sectionally (Table S5). Accounting for use of cholesterol-lowering drugs, in analyses of a subset where individuals taking cholesterol-lowering drugs were excluded, resulted in higher effect estimates across all assessed CVD outcomes, and the association with ischemic heart disease remained significant even with the lower sample size (Table S6). Remarkably, the observed effect of the p.G137S variant on ischemic heart disease, when doing a cross-sectional analysis in Greenlanders was much greater than the effect of any common variant identified in Europeans[30] (Figure 2D). Moreover, these associations remained when adjusting for BMI, smoking, and blood pressure (data not shown). On the contrary, we observed no association with cerebrovascular disease, or a combined measure of the different types of CVD events (Figure 4 and Table S3).

Discussion

In this study, we investigated the impact of the common Arctic-specific *LDLR* p.G137S missense variant on specific lipids as well as metabolic and cardiovascular health in the Greenlandic population.

First, we replicated the previously reported associations with elevated circulating concentrations of LDL cholesterol, total cholesterol, and apolipoprotein B. Notably, our results showed that the p.G137S variant had an even larger impact on the lipid profile of Greenlanders than previously reported,[18] even when we did not exclude individuals treated with cholesterol-lowering drugs, which resulted in a lower effect estimate. For LDL cholesterol, we estimated an effect size of 0.75 mmol/L (0.66 SD) per p.G137S allele, whereas the previously reported estimate was 0.54 mmol/L.[18] Furthermore, we found an association with HDL cholesterol not previously reported. A likely explanation for these differences is that the previous analyses were performed using a multivariate linear model adjusting only for age, sex, BMI, and geographic location, whereas we used a linear mixed model accounting additionally for population structure and relatedness. In the admixed Greenlandic population, this model enabled a more accurate estimation of effect sizes.

Interestingly, the p.G137S effect size for LDL cholesterol (0.66 SD per allele) was much larger than the effect sizes reported for any common variant (MAF >1%) identified in Europeans of maximum 0.19 SD,[29] and it was also larger than the effect size for predicted loss-of-function and missense variants in the UK Biobank with allele frequencies below 0.1% (0.42 SD), and similar to predicted loss-of-function and missense variants with allele frequencies below 0.001% (0.65 SD).[32]

The large effect estimate, combined with the fact that p.G137S is carried by 29.5% of the Greenlandic population, means that the variant has a large impact on this population. This large impact on the population level was particularly evident by the fact that 25.4% of the heterozygous and 54.7% of the homozygous carriers had LDL cholesterol levels above the diagnostic level for FH and by the fact that the explained variance for LDL cholesterol in Greenlanders was estimated to be around 12%, and more than 16% when only looking at unadmixed Inuit. This estimated amount of variance explained by the *LDLR* p.G137S variant is more than that of any identified common variant in Europeans, and also more than that of clinical risk factors like BMI, sex, smoking, age, or waist-hip ratio explained. For comparison, estimates of the amount of phenotypic variance explained by genetics in large-scale European studies using polygenic risk scores range from 10%[13] to 19.9%.[33] Hence, the p.G137S variant alone has an impact on the Greenlandic population comparable to an entire polygenic risk score in Europeans.

The *LDLR* p.G137S variant is an example of a genetic variant with a combination of having a large effect size and being common, similar to the common variants
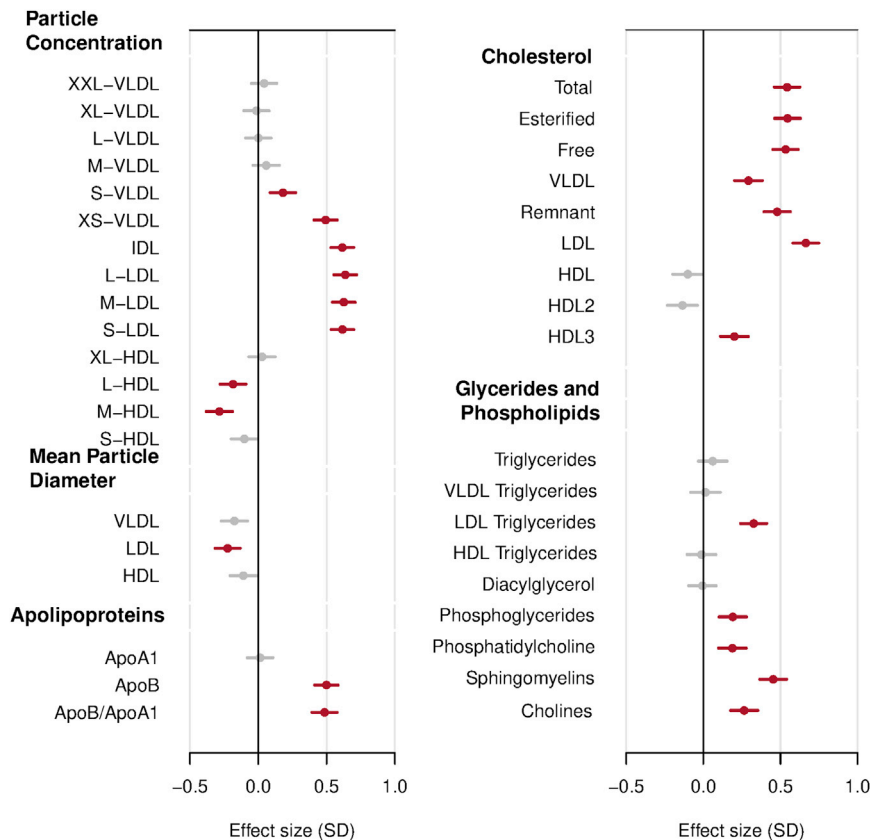
Figure 3. Effect of the p.G137S variant on concentration of lipoprotein particles and cholesterol species

Lipoprotein particle concentration and metabolic measures in Greenlanders with NMR phenotypes. Estimates of the effect size of the p.G137S variant are shown with their 95% confidence interval, furthermore the confidence intervals are colored red if their p-value is below the Bonferroni correction of $3.3 \cdot 10^{-4}$, and gray otherwise. Phenotype values were quantile transformed to a standard normal distribution. XXL, extremely large; XL, very large; L, large; M, medium; S, small; XS, very small; VLDL, very low-density lipoprotein; IDL, intermediate-density lipoprotein; LDL, low-density lipoprotein; HDL, high-density lipoprotein; Apo, apolipoprotein.

with large effects on lipid levels, BMI, or risk of type 2 diabetes identified in small isolated populations of Greek, Greenlandic, Samoan, or Pima Indian individuals, respectively.[34–39] Variants like this are unusual in large outbred populations like the European, but can be found in small isolated populations, where genetic drift has a large impact on allele frequencies.[40,41] Importantly, the *LDLR* variant has an additive effect for LDL cholesterol and the associated metabolic phenotypes, and thereby potentially a bigger impact on the general population health as compared with variants with a recessive effect.[37] Hence, this variant further highlights the importance of carrying out genetic studies in smaller populations like the Greenlandic.

When we looked further into the effects of the p.G137S variant on the lipid-related measures, from the NMR data, we observed higher concentrations of atherosclerotic lipoproteins, namely large, medium, and small LDL, IDL, and the smallest subclasses of VLDL particles, as well as apolipoprotein B. This is in line with reduced function of the LDLR and with findings in FH children.[42] Further contributing to the atherosclerotic profile in p.G137S carriers, we also observed significantly smaller LDL particles[43] and higher concentrations of both phosphatidylcholines and sphingomyelins, the latter class including ceramides. Specific subtypes of these lipid species have previously been linked to increased CVD risk, and these associations seem to be independent of other atherosclerotic risk factors, including age, BMI, smoking, and triglycerides, as well as LDL and total cholesterol.[44] Extended lipidomics profiling assessing specific phosphatidylcholines and ceramides further elucidating these possibly causal relationships would be of great interest. With respect to HDL cholesterol, we observed higher concentrations of HDL₃, but lower concentrations of large and medium HDL particles, which might indicate reduced flux of cholesterol through the reverse cholesterol transport system, and thereby reduced hepatic cholesterol uptake. The observed lipoprotein profile in the carriers of the p.G137S carriers was overall very similar to the profile reported for children with FH with different causative mutations; however, with larger effect sizes in the children.[42] This difference could be due to the lack of exclusion of individuals treated with cholesterol-lowering medication in our study, or to the functional effect of the p.G137S variant, approximately reducing the ligand binding to the LDLR by 60%, as compared with the FH mutations, which in general have greater functional impact.[18,45]

Regarding the CVD impact of the p.G137S variant, we found a significantly increased risk of ischemic heart disease, coronary operations, and peripheral artery disease, and observed similar effect estimates when limiting the analyses to events after inclusion or when analyzing CVD risk cross-sectionally. Importantly, these findings are in line with previous results for patients with FH, who are known to have a higher risk of CVD, especially ischemic heart disease[9] but risk of cerebrovascular disease on level with the general population.[13,46] Generally, our findings were consistent with a recent Mendelian randomization study where elevated LDL cholesterol levels were shown to be causally linked to increased risk of ischemic heart disease but not ischemic stroke.[47]

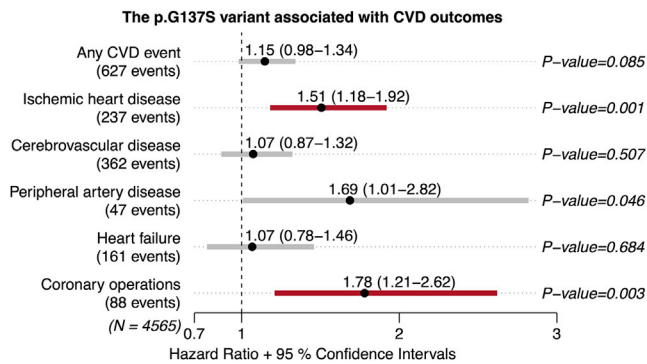In this study, only a low proportion of individuals with very high levels of LDL cholesterol were receiving

The p.G137S variant associated with CVD outcomes

Any CVD event (627 events): 1.15 (0.98–1.34), P–value=0.085
Ischemic heart disease (237 events): 1.51 (1.18–1.92), P–value=0.001
Cerebrovascular disease (362 events): 1.07 (0.87–1.32), P–value=0.507
Peripheral artery disease (47 events): 1.69 (1.01–2.82), P–value=0.046
Heart failure (161 events): 1.07 (0.78–1.46), P–value=0.684
Coronary operations (88 events): 1.78 (1.21–2.62), P–value=0.003

(N = 4565)

Hazard Ratio + 95 % Confidence Intervals

Figure 4. Estimated hazard ratios for the p.G137S variant on CVD
Hazard ratios are shown with their 95% confidence interval and p-values (P) estimated with Cox regression testing for an association between the p.G137S variant and risk of different types of CVD events. The red color denotes those where the p-value is below the Bonferroni correction of 0.00833.

cholesterol-lowering therapy and only around 10% of the homozygous p.G137S carriers were prescribed cholesterol-lowering drugs, even though most of them had LDL cholesterol levels above the 4.9 mmol/L level considered diagnostic for FH.[9] Part of the explanation for the low number is that the individuals are from a population-based survey of mostly healthy adults and because statin guidelines in Greenland were less aggressive at the time of sampling around 15 years ago compared with now. Current treatment guidelines in Greenland are based on international recommendations, and recent quality data from Greenland show a reasonably high proportion of individuals with diabetes on statin treatment, although still with room for improvement.[48] Thus, the current study suggests that there are many currently untreated individuals with very high lipid levels who potentially could benefit from treatment irrespective of other CVD risk factors. Moreover, we observed that the elevated levels of LDL cholesterol for p.G137S carriers was independent of age, indicating that these individuals would benefit from early intervention and treatment, similar to the recommendations for treatment of FH.[9] In addition, studies have shown that monogenic FH variants increase the risk of CVD more relative to polygenic hypercholesterolemia variants; it is hypothesized that this is due to monogenic variants leading to greater cumulative LDL cholesterol exposure throughout life.[13] It has been shown for FH individuals that early statin therapy can help mitigate their risk of developing IHD so that it is almost similar to individuals without FH.[15,49] Our findings, coupled with the fact that the p.G137S variant is common among Greenlanders, indicate that a screening program for the p.G137S variant could be highly useful for early identification of individuals at increased CVD risk, hence, potentially improving preventive care and public health. To this end, it would be valuable to perform an intervention study to investigate if any behavioral factors, such as physical activity, diet, or smoking, could possibly also play a role in such preventive care.

In summary, the common p.G137S variant had a major effect on the circulating lipid profile of Greenlanders, where carriers were characterized by an atherogenic lipid profile including elevated levels of LDL cholesterol, total cholesterol, and apolipoprotein B as well as an increased risk of CVD. Furthermore, the large effect size for LDL cholesterol combined with the high frequency of the variant, mean that the variant has a large population-level impact. The variant is therefore a simple marker that could be used as an early indicator of future risk of elevated lipid levels and CVD. Early detection of individuals likely benefiting from intervention with cholesterol-lowering drugs, could potentially lead to improved CVD prevention in Arctic populations.

Data and code availability

The Greenlandic Metabochip-genotype data and the RNA sequencing data are deposited in the European Genome-phenome Archive (https://www.ebi.ac.uk/ega/home) under the accessions EGAS00001002641 and EGAS00001004127, respectively. The summary statistics from the NMR analyses are available in Table S9.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.xhgg.2022.100118.

Acknowledgments

Declaration of interests

References

1. World Health Organization (2018). Global health estimates 2016: disease burden by cause, age, sex, by country and by region, 2000-2016. https://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.html.

2. Andersen, M.K., and Hansen, T. (2018). Genetics of metabolic traits in Greenlanders: lessons from an isolated population. J. Intern. Med. *284*, 464–477. https://doi.org/10.1111/joim.12814.

3. Tvermosegaard, M., Dahl-Petersen, I.K., Nielsen, N.O., Bjerregaard, P., and Jørgensen, M.E. (2015). Cardiovascular disease susceptibility and resistance in circumpolar inuit populations. Can. J. Cardiol. *31*, 1116–1123. https://doi.org/10.1016/j.cjca.2015.06.002.

4. Scientific Steering Committee on behalf of the Simon Broome Register Group (1991). Risk of fatal coronary heart disease in familial hypercholesterolaemia. Scientific steering committee on behalf of the simon broome register group. BMJ *303*, 893–896.

5. Versmissen, J., Oosterveer, D.M., Yazdanpanah, M., Defesche, J.C., Basart, D.C.G., Liem, A.H., Heeringa, J., Witteman, J.C., Lansberg, P.J., Kastelein, J.J.P., and Sijbrands, E.J.G. (2008). Efficacy of statins in familial hypercholesterolaemia: a long term cohort study. BMJ *337*, a2423. https://doi.org/10.1136/bmj.a2423.

6. Elis, A., Zhou, R., and Stein, E.A. (2011). Effect of lipid-lowering treatment on natural history of heterozygous familial hypercholesterolemia in past three decades. Am. J. Cardiol. *108*, 223–226. https://doi.org/10.1016/j.amjcard.2011.03.027.

7. Raal, F.J. (2013). Lomitapide for homozygous familial hypercholesterolaemia. Lancet *381*, 7–8. https://doi.org/10.1016/s0140-6736(12)61845-5.

8. Sharifi, M., Futema, M., Nair, D., and Humphries, S.E. (2017). Genetic architecture of familial hypercholesterolaemia. Curr. Cardiol. Rep. *19*, 44. https://doi.org/10.1007/s11886-017-0848-8.

9. Youngblom, E., Pariani, M., and Knowles, J.W. (2014). Familial hypercholesterolemia. In GeneReviews®, M.P. Adam, H.H. Ardinger, and R.A. Pangon, et al., eds. (University of Washington).

10. Mundal, L., Sarancic, M., Ose, L., Iversen, P.O., Borgan, J., Veierod, M.B., Leren, T.P., and Retterstol, K. (2014). Mortality among patients with familial hypercholesterolemia: a registry-based study in Norway, 1992-2010. J. Am. Heart Assoc. *3*, e001236. https://doi.org/10.1161/jaha.114.001236.

11. Krogh, H.W., Mundal, L., Holven, K.B., and Retterstøl, K. (2016). Patients with familial hypercholesterolaemia are characterized by presence of cardiovascular disease at the time of death. Eur. Heart J. *37*, 1398–1405. https://doi.org/10.1093/eurheartj/ehv602.

12. Trinder, M., Li, X., DeCastro, M.L., Cermakova, L., Sadananda, S., Jackson, L.M., Azizi, H., Mancini, G.J., Francis, G.A., Frohlich, J., and Brunham, L.R. (2019). Risk of premature atherosclerotic disease in patients with monogenic versus polygenic familial hypercholesterolemia. J. Am. Coll. Cardiol. *74*, 512–522. https://doi.org/10.1016/j.jacc.2019.05.043.

13. Trinder, M., Francis, G.A., and Brunham, L.R. (2020). Association of monogenic vs polygenic hypercholesterolemia with risk of atherosclerotic cardiovascular disease. JAMA Cardiol *5*, 390–399. https://doi.org/10.1001/jamacardio.2019.5954.

14. Wiegman, A., Gidding, S.S., Watts, G.F., Chapman, M.J., Ginsberg, H.N., Cuchel, M., Ose, L., Averna, M., Boileau, C., Boren, J., et al. (2015). Familial hypercholesterolaemia in children and adolescents: gaining decades of life by optimizing detection and treatment. Eur. Heart J. *36*, 2425–2437. https://doi.org/10.1093/eurheartj/ehv157.

15. Vuorio, A., Docherty, K.F., Humphries, S.E., Kuoppala, J., and Kovanen, P.T. (2013). Statin treatment of children with familial hypercholesterolemia–trying to balance incomplete evidence of long-term safety and clinical accountability: are we approaching a consensus? Atherosclerosis *226*, 315–320. https://doi.org/10.1016/j.atherosclerosis.2012.10.032.

16. Brown, M.S., and Goldstein, J.L. (1974). Familial hypercholesterolemia: defective binding of lipoproteins to cultured fibroblasts associated with impaired regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase activity. Proc. Natl. Acad. Sci. U S A *71*, 788–792. https://doi.org/10.1073/pnas.71.3.788.

17. Thormaehlen, A.S., Schuberth, C., Won, H.H., Blattmann, P., Joggerst-Thomalla, B., Theiss, S., Asselta, R., Duga, S., Merlini, P.A., Ardissino, D., et al. (2015). Systematic cell-based phenotyping of missense alleles empowers rare variant association studies: a case for LDLR and myocardial infarction. PLoS Genet. *11*, e1004855. https://doi.org/10.1371/journal.pgen.1004855.

18. Dubé, J.B., Wang, J., Cao, H., McIntyre, A.D., Johansen, C.T., Hopkins, S.E., Stringer, R., Hosseinzadeh, S., Kennedy, B.A., Ban, M.R., et al. (2015). Common low-density lipoprotein receptor p.G116S variant has a large effect on plasma low-density lipoprotein cholesterol in circumpolar inuit populations. Circ. Cardiovasc. Genet. *8*, 100–105. https://doi.org/10.1161/circgenetics.114.000646.

19. Bjerregaard, P. (2011). Inuit Health in Transition: Greenland Survey 2005-2010: Population Sample and Survey Methods (Statens Institut for Folkesundhed).

20. Bjerregaard, P., Curtis, T., Borch-Johnsen, K., Mulvad, G., Becker, U., Andersen, S., and Backer, V. (2003). Inuit health in Greenland: a population survey of life style and disease in Greenland and among Inuit living in Denmark. Int. J. Circumpolar Health *62*, 3–79. https://doi.org/10.3402/ijch.v62i0.18212.

21. Skotte, L., Koch, A., Yakimov, V., Zhou, S., Soborg, B., Andersson, M., Michelsen, S.W., Navne, J.E., Mistry, J.M., Dion, P.A., et al. (2017). *CPT1A* missense mutation associated with fatty acid metabolism and reduced height in Greenlanders. Circ. Cardiovasc. Genet. *10*. https://doi.org/10.1161/circgenetics.116.001618.

22. Voight, B.F., Kang, H.M., Ding, J., Palmer, C.D., Sidore, C., Chines, P.S., Burtt, N.P., Fuchsberger, C., Li, Y., Erdmann, J., et al. (2012). The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. PLoS Genet. *8*, e1002793. https://doi.org/10.1371/journal.pgen.1002793.

23. Andersen, M.K., Jorsboe, E., Sandholt, C.H., Grarup, N., Jorgensen, M.E., Færgeman, N.J., Bjerregaard, P., Pedersen, O., Moltke,

I., Hansen, T., and Albrechtsen, A. (2016). Identification of novel genetic determinants of erythrocyte membrane fatty acid composition among Greenlanders. PLoS Genet. *12*, e1006119. https://doi.org/10.1371/journal.pgen.1006119.

24. Moltke, I., Fumagalli, M., Korneliussen, T., Crawford, J., Bjerregaard, P., Jorgensen, M., Grarup, N., Gullov, H., Linneberg, A., Pedersen, O., et al. (2015). Uncovering the genetic history of the present-day Greenlandic population. Am. J. Hum. Genet. *96*, 54–69. https://doi.org/10.1016/j.ajhg.2014.11.012.

25. Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. *44*, 821–824. https://doi.org/10.1038/ng.2310.

26. Chen, H., Wang, C., Conomos, M., Stilp, A., Li, Z., Sofer, T., Szpiro, A., Chen, W., Brehm, J., Celedon, J., et al. (2016). Control for population structure and relatedness for binary traits in genetic association studies via logistic mixed models. Am. J. Hum. Genet. *98*, 653–666. https://doi.org/10.1016/j.ajhg.2016.02.012.

27. Therneau, T.M., and Grambsch, P.M. (2000). Modeling Survival Data: Extending the Cox Model.

28. Author anonymous. (2002). II. Rationale for intervention. Circulation *106*, 3163–3223.

29. Liu, D.J., Peloso, G.M., Yu, H., Butterworth, A.S., Wang, X., Mahajan, A., Saleheen, D., Emdin, C., Alam, D., Alves, A.C., et al. (2017). Exome-wide association study of plasma lipids in >300,000 individuals. Nat. Genet. *49*, 1758–1766. https://doi.org/10.1038/ng.3977.

30. van der Harst, P., and Verweij, N. (2018). Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. Circ. Res. *122*, 433–443. https://doi.org/10.1161/circresaha.117.312086.

31. Shim, H., Chasman, D.I., Smith, J.D., Mora, S., Ridker, P.M., Nickerson, D.A., Krauss, R.M., and Stephens, M. (2015). A multivariate genome-wide association analysis of 10 LDL subfractions, and their response to statin treatment, in 1868 Caucasians. PLoS One *10*, e0120758. https://doi.org/10.1371/journal.pone.0120758.

32. Backman, J.D., Li, A.H., Marcketta, A., Sun, D., Mbatchou, J., Kessler, M.D., Benner, C., Liu, D., Locke, A.E., Balasubramanian, S., et al. (2021). Exome sequencing and analysis of 454,787 UK Biobank participants. Nature *599*, 628–634. https://doi.org/10.1038/s41586-021-04103-z.

33. Surakka, I., Horikoshi, M., Magi, R., Sarin, A.P., Mahajan, A., Lagou, V., Marullo, L., Ferreira, T., Miraglio, B., Timonen, S., et al. (2015). The impact of low-frequency and rare variants on lipid levels. Nat. Genet. *47*, 589–597. https://doi.org/10.1038/ng.3300.

34. Southam, L., Gilly, A., Suveges, D., Farmaki, A.E., Schwartzentruber, J., Tachmazidou, I., Matchan, A., Rayner, N.W., Tsafantakis, E., Karaleftheri, M., et al. (2017). Whole genome sequencing and imputation in isolated populations identify genetic associations with medically-relevant complex traits. Nat. Commun. *8*, 15606. https://doi.org/10.1038/ncomms15606.

35. Andersen, M.K., Jorsboe, E., Skotte, L., Hanghoj, K., Sandholt, C.H., Moltke, I., Grarup, N., Kern, T., Mahendran, Y., Soborg, B., et al. (2020). The derived allele of a novel intergenic variant at chromosome 11 associates with lower body mass index and a favorable metabolic phenotype in Greenlanders. PLoS Genet. *16*, e1008544. https://doi.org/10.1371/journal.pgen.1008544.

36. Grarup, N., Moltke, I., Andersen, M.K., Dalby, M., Vitting-Seerup, K., Kern, T., Mahendran, Y., Jorsboe, E., Larsen, C.V.L., Dahl-Petersen, I.K., et al. (2018). Loss-of-function variants in ADCY3 increase risk of obesity and type 2 diabetes. Nat. Genet. *50*, 172–174. https://doi.org/10.1038/s41588-017-0022-7.

37. Moltke, I., Grarup, N., Jorgensen, M.E., Bjerregaard, P., Treebak, J.T., Fumagalli, M., Korneliussen, T.S., Andersen, M.A., Nielsen, T.S., Krarup, N.T., et al. (2014). A common Greenlandic TBC1D4 variant confers muscle insulin resistance and type 2 diabetes. Nature *512*, 190–193. https://doi.org/10.1038/nature13425.

38. 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. https://doi.org/10.1038/ng.3620.

39. Baier, L.J., Muller, Y.L., Remedi, M.S., Traurig, M., Piaggi, P., Wiessner, G., Huang, K., Stacy, A., Kobes, S., Krakoff, J., et al. (2015). ABCC8 R1420H loss-of-function variant in a southwest American Indian community: association with increased birth weight and doubled risk of type 2 diabetes. Diabetes *64*, 4322–4332. https://doi.org/10.2337/db15-0459.

40. Pedersen, C.-E.T., Lohmueller, K.E., Grarup, N., Bjerregaard, P., Hansen, T., Siegismund, H.R., Moltke, I., and Albrechtsen, A. (2017). The effect of an extreme and prolonged population bottleneck on patterns of deleterious variation: insights from the Greenlandic inuit. Genetics *205*, 787–801. https://doi.org/10.1534/genetics.116.193821.

41. Andersen, M.K., Pedersen, C.E.T., Moltke, I., Hansen, T., Albrechtsen, A., and Grarup, N. (2016). Genetics of type 2 diabetes: the power of isolated populations. Curr. Diab. Rep. *16*, 65. https://doi.org/10.1007/s11892-016-0757-z.

42. Christensen, J.J., Ulven, S.M., Retterstol, K., Narverud, I., Bogsrud, M.P., Henriksen, T., Bollerslev, J., Halvorsen, B., Aukrust, P., and Holven, K.B. (2017). Comprehensive lipid and metabolite profiling of children with and without familial hypercholesterolemia: a cross-sectional study. Atherosclerosis *266*, 48–57. https://doi.org/10.1016/j.atherosclerosis.2017.09.021.

43. Ikezaki, H., Lim, E., Cupples, L.A., Liu, C., Asztalos, B.F., and Schaefer, E.J. (2021). Small dense low-density lipoprotein cholesterol is the most atherogenic lipoprotein parameter in the prospective Framingham offspring study. J. Am. Heart Assoc. *10*, e019140. https://doi.org/10.1161/jaha.120.019140.

44. Ding, M., and Rexrode, K.M.A. (2020). A review of lipidomics of cardiovascular disease highlights the importance of isolating lipoproteins. Metabolites *10*, 163. https://doi.org/10.3390/metabo10040163.

45. Hobbs, H.H., Russell, D.W., Brown, M.S., and Goldstein, J.L. (1990). The LDL receptor locus in familial hypercholesterolemia: mutational analysis of a membrane protein. Annu. Rev. Genet. *24*, 133–170. https://doi.org/10.1146/annurev.ge.24.120190.001025.

46. Huxley, R.R., Hawkins, M., Humphries, S., Karpe, F., and Neil, H. (2003). Risk of fatal stroke in patients with treated familial hypercholesterolemia: a prospective registry study. Stroke *34*, 22–25. https://doi.org/10.1161/01.str.0000047123.14312.3e.

47. Valdes-Marquez, E., Parish, S., Clarke, R., Stari, T., Worrall, B.B., and Hopewell, J.C. (2019). Relative effects of LDL-C

on ischemic stroke and coronary disease: a Mendelian randomization study. Neurology *92*, e1176–e1187. https://doi.org/10.1212/WNL.0000000000007091.

48. Pedersen, M.L. (2019). Diabetes care in the dispersed population of Greenland. A new model based on continued monitoring, analysis and adjustment of initiatives taken. Int. J. Circumpolar Health *78*, 1709257. https://doi.org/10.1080/22423982.2019.1709257.

49. Nordestgaard, B.G., Chapman, M.J., Humphries, S.E., Ginsberg, H.N., Masana, L., Descamps, O.S., Wiklund, O., Hegele, R.A., Raal, F.J., Defesche, J.C., et al. (2013). Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease: consensus statement of the European Atherosclerosis Society. Eur. Heart J. *34*, 3478–3490a. https://doi.org/10.1093/eurheartj/eht273.