Journal Pre-proof

Loss of sucrase-isomaltase function increases acetate levels and improves metabolic health in Greenlandic cohorts

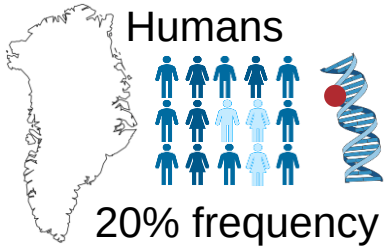
Mette K. Andersen, Line Skotte, Emil Jørsboe, Ryan Polito, Frederik F. Stæger, Peter Aldiss, Kristian Hanghøj, Ryan K. Waples, Cindy G. Santander, Niels Grarup, Inger K. Dahl-Petersen, Lars J. Diaz, Maria Overvad, Ninna K. Senftleber, Bolette Søborg, Christina V.L. Larsen, Clara Lemoine, Oluf Pedersen, Bjarke Feenstra, Peter Bjerregaard, Mads Melbye, Marit E. Jørgensen, Nils J. Færgeman, Anders Koch, Thomas Moritz, Matthew P. Gillum, Ida Moltke, Torben Hansen, Anders Albrechtsen

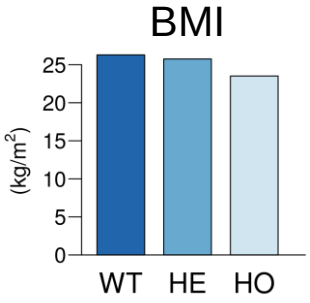
PII: S0016-5085(21)04065-8
DOI: https://doi.org/10.1053/j.gastro.2021.12.236
Reference: YGAST 64770

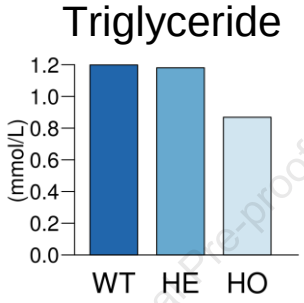
To appear in: *Gastroenterology*
Accepted Date: 2 December 2021

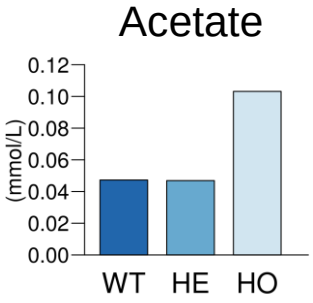
Please cite this article as: Andersen MK, Skotte L, Jørsboe E, Polito R, Stæger FF, Aldiss P, Hanghøj K, Waples RK, Santander CG, Grarup N, Dahl-Petersen IK, Diaz LJ, Overvad M, Senftleber NK, Søborg B, Larsen CVL, Lemoine C, Pedersen O, Feenstra B, Bjerregaard P, Melbye M, Jørgensen ME, Færgeman NJ, Koch A, Moritz T, Gillum MP, Moltke I, Hansen T, Albrechtsen A, Loss of sucrase-isomaltase function increases acetate levels and improves metabolic health in Greenlandic cohorts, *Gastroenterology* (2022), doi: https://doi.org/10.1053/j.gastro.2021.12.236.

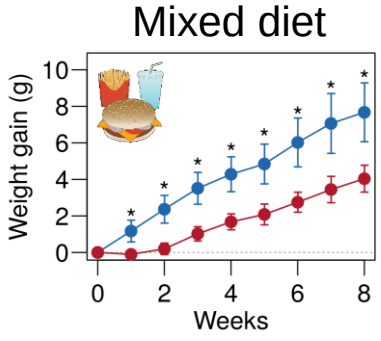
SI Loss-of-function variant improves metabolic health

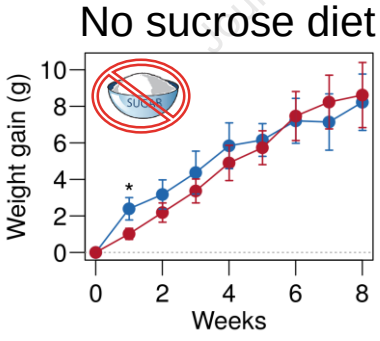


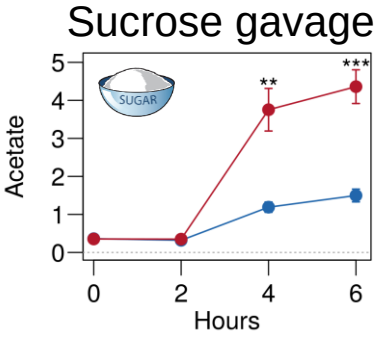


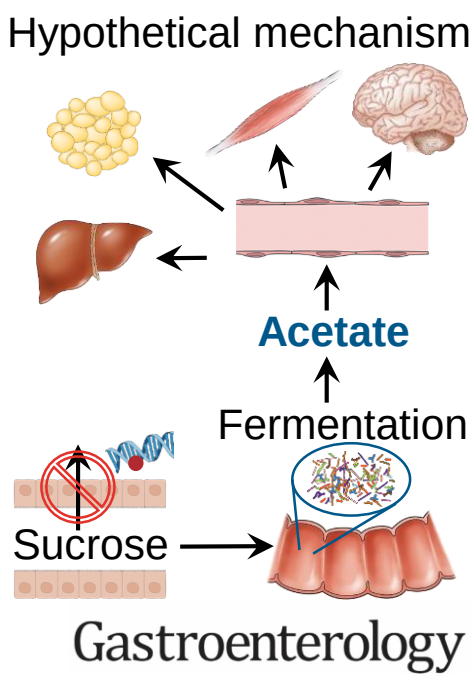












Loss of sucrase-isomaltase function increases acetate levels and improves metabolic health in Greenlandic cohorts

Short title: Sucrase-isomaltase and metabolic health

Mette K. Andersen[1‡], Line Skotte[2‡], Emil Jørsboe[1,3‡], Ryan Polito[1], Frederik F. Stæger[3], Peter Aldiss[1], Kristian Hanghøj[3], Ryan K. Waples[3], Cindy G. Santander[3], Niels Grarup[1], Inger K. Dahl-Petersen[4,5], Lars J. Diaz[5], Maria Overvad[5], Ninna K. Senftleber[3,5], Bolette Søborg[2], Christina V. L. Larsen[4,6], Clara Lemoine[1], Oluf Pedersen[1], Bjarke Feenstra[2], Peter Bjerregaard[4], Mads Melbye[2,7,8], Marit E. Jørgensen[4,5,6], Nils J. Færgeman[9], Anders Koch[2,6,10], Thomas Moritz[1], Matthew P. Gillum[1], Ida Moltke[3]*, Torben Hansen[1,11]*, Anders Albrechtsen[3]*.

[1] Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[2] Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark

[3] Section for Computational and RNA Biology, Department of Biology, University of Copenhagen, Copenhagen, Denmark

[4] National Institute of Public Health, University of Southern Denmark, Copenhagen, Denmark

[5] Steno Diabetes Center Copenhagen, Gentofte, Denmark

[6] Greenland Centre for Health Research, University of Greenland, Nuuk, Greenland

[7] Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[8] Department of Medicine, Stanford University School of Medicine, Stanford, California, USA

[9] Department of Biochemistry and Molecular Biology, Villum Center for Bioanalytical Sciences, University of Southern Denmark, Odense, Denmark

[10] Department of Infectious Diseases, Rigshospitalet University Hospital, Copenhagen, Denmark

[11] Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark

‡Mette K. Andersen, Line Skotte, and Emil Jørsboe contributed equally to this work.

*joint senior and corresponding authors:

- Ida Moltke (ida@binf.ku.dk; tel: +45 35330239)

Ole Maaløes Vej 5, DK-2200 Copenhagen N, Denmark

- Torben Hansen (torben.hansen@sund.ku.dk; tel: +45 35337129)

Blegdamsvej 3B, DK-2200 Copenhagen N, Denmark

- Anders Albrechtsen (albrecht@binf.ku.dk; tel: +45 35330246)

Ole Maaløes Vej 5, DK-2200 Copenhagen N, Denmark

Word count, abstract (including key words): 260; main text (including references and figure and table legends): 6,907; tables: 2; Figures: 5.

Grant support

This project was funded by the Danish Council for Independent Research (DFF-4090-00244, Sapere Aude grant DFF-11-120909 and DFF-4181-00383), the Steno Diabetes Center Copenhagen (www.steno.dk), the Lundbeck Foundation (R215-2015-4174) and the Novo Nordisk Foundation (NNF0064142, NNF20OC0061343, NNF15OC0017918, NNF16OC0019986, NNF17SH0027192, NNF17OC0028136 and NNFCC0018486). The Greenlandic health surveys (IHIT and B99) were supported by Karen Elise Jensen's Foundation, the Department of Health in Greenland, NunaFonden, Medical Research Council of Denmark, Medical Research Council of Greenland, and the Commission for Scientific Research in Greenland. The population-based study referred to as cohort II was supported by grants from the Danish Council for Independent Research, The Greenlandic Ministry of Education, Church, Culture and Gender Equality, the Maersk Foundation (Fonden til Lægevidenskabens Fremme), and the Aase and Ejnar Danielsens Foundation. Samples were handled and stored in the Danish National Biobank, which is supported by the Novo Nordisk Foundation. IM was supported by a Danish National Research Foundation Award (DNRF 143). BF was supported by the Oak Foundation. LS was supported by the Carlsberg Foundation. MKA was supported by a research grant from the Danish Diabetes Academy supported by the Novo Nordisk Foundation.

Abbreviations

Congenital sucrase-isomaltase deficiency (CSID), diet containing 17 kcal% sucrose (17S), diet containing no sucrose (NS), high fat diet containing 12.6 kcal% sucrose (HFS), high-fat diet containing no sucrose

(HFNS), homeostasis model assessment of insulin resistance (HOMA-IR), irritable bowel syndrome (IBS), maltase-glucoamylase (MGAM), oral glucose tolerance test (OGTT), odds ratios (OR), short chain fatty acid (SCFA), sucrase-isomaltase (SI), sucrase-isomaltase gene (*SI*), sucrase-isomaltase homozygous knock-out (Sis-KO), sucrase-isomaltase homozygous wild-type (Sis-WT).

Competing interests

The authors declare no competing interests.

Author contributions

AA, TH, and IM conceived and headed the project, and AA, IM, and MKA conceptualized the project. RP, PA, CL, and MPG designed and headed the mice studies. TH, OP, and NG designed the experimental setup for generation of genotype data. IM and AA set up the framework for association testing, and EJ and LS performed the statistical analyses. FFS, KH, RKW, CGS, BF, and TM contributed to data processing, analysis, and interpretation. MEJ, PB, CVLL, and IKDP collected and managed the cohort I study samples, AK, MM, BF, BS, and LS collected and managed the cohort II study samples. LJD, MO, NKS, and MEJ curated the register data on cardiovascular disease events. NJF supervised the data interpretation. MKA and IM wrote the manuscript, with input from all other authors. All authors approved the final version of the manuscript.

Data availability

The Greenlandic Metabochip-genotype data are deposited in the European Genome-phenome Archive (https://www.ebi.ac.uk/ega/home) under the accessions EGAS00001002641.

Abstract

Background & Aims The sucrase-isomaltase (*SI*) c.273_274delAG loss-of-function variant is common in Arctic populations and causes congenital sucrase-isomaltase deficiency, an inability to breakdown and absorb sucrose and isomaltose. Children with this condition experience gastrointestinal symptoms, when dietary sucrose is introduced. Here we aimed to describe the health of adults with sucrase-isomaltase deficiency.

Methods Association between c.273_274delAG and phenotypes related to metabolic health was assessed in two cohorts of Greenlandic adults (*N*=4,922 and *N*=1,629). A sucrase-isomaltase knock-out (Sis-KO) mouse model was used to further elucidate the findings.

Results homozygous carriers of the variant had a markedly healthier metabolic profile, than the remaining population, including lower BMI (β (SE), -2.0 kg/m^2 (0.5), P=3.1x10^{-5}), body weight (-4.8 kg (1.4), P=5.1x10^{-4}), fat percentage (-3.3% (1.0), P=3.7x10^{-4}), fasting triglyceride (-0.27 mmol/L (0.07), P=2.3x10^{-6}), and remnant cholesterol (-0.11 mmol/L (0.03), P=4.2x10^{-5}). Further analyses suggested that this was likely mediated partly by higher circulating levels of acetate observed in homozygous carriers (0.056 mmol/L (0.002), P=2.1x10^{-26}), and partly by reduced sucrose uptake, but not lower caloric intake. These findings were verified in Sis-KO mice, which compared to wild-type mice were leaner on a sucrose-containing diet, despite similar caloric intake, had significantly higher plasma acetate levels in response to a sucrose gavage, and had lower plasma glucose level in response to a sucrose-tolerance test.

Conclusions These results suggest that sucrase-isomaltase constitutes a promising drug target for improvement of metabolic health, and that the health benefits are mediated by reduced dietary sucrose uptake and possibly also by higher levels of circulating acetate.

Keywords: sucrase-isomaltase, genetics, loss-of-function, metabolic health, drug target

Introduction

To prevent or delay age-related conditions like type 2 diabetes and cardiovascular disease, it is vital to sustain metabolic health. Metabolic health is determined by genetic factors and health behavior, including dietary habits. Hence, understanding how different dietary components are metabolized and utilized may identify pathways important for sustaining or improving metabolic health. For most people, carbohydrates constitute the primary dietary component[1,2]. Carbohydrates are mainly ingested as starch and sugars, and in a Westernized diet, the most abundant dietary sugar is sucrose. The health effects of the increased carbohydrate, and in particular sugar consumption, are heavily debated[3,4].

When ingested, carbohydrates in the form of starch and sugar need to be broken down to monosaccharides, in order to move across the intestinal epithelium, and be taken up by the body. This carbohydrate digestion is initiated by α-amylases in the mouth, and is finalized in the small intestine by the α-glucosidases, maltase-glucoamylase (MGAM) and sucrase-isomaltase (SI)[5–7]. These digestive enzymes are among the targets of the anti-diabetic α-glucosidase inhibitor drugs acarbose, voglibose, and miglitol. These drugs target a combination of enzymes, and are thought to reduce the degradation of starch and sugars, thereby reducing the amount of glucose absorbed to the blood. Acarbose mainly inhibits α-amylases and partly maltase and sucrase[8–10], whereas miglitol and voglibose bind all four α-glucosidase subunits, but have no or very limited affinity for α-amylases[10–12].

Naturally occurring genetic variation that disrupts the function of MGAM and SI can help indicate the effect of specifically targeting these enzymes. Deficiency of both MGAM and SI has been linked to maldigestion and severe gastrointestinal symptoms in children[13,14]. Thus, congenital sucrase-isomaltase deficiency (CSID) is associated with a range of symptoms in children, including diarrhea, abdominal pain, and bloating[15–17], yet gastrointestinal and metabolic-health status in adults has not been reported. CSID is rare in most parts of the world, except in Arctic populations, where the condition has an estimated prevalence of up to 10%[15]. Recently, the c.273-274delAG frameshift variant in the sucrase-isomaltase gene (*SI*) encoding SI, was identified in a Canadian CSID patient[18]. This variant is predicted to result in complete loss of SI function[18]. Hence, homozygous carriers of the variant represent human SI knock-outs, which facilitates assessment of health-related implications of targeted SI inhibition. Importantly, the variant has an estimated allele frequency of 39% in the Greenlandic population[19]. Thus, it is possible to assess the effect of being a homozygous carrier of this variant in Greenlanders. We, therefore, aimed to thoroughly assess how *SI* knock-out affects metabolic, gastrointestinal, and cardiovascular health in 6,551 Greenlandic adults, by assessing two cohorts with complimentary phenotypes. Additionally, to gain further mechanistic insights, we monitored food intake, body weight,

and body composition for 8 weeks in sucrase-isomaltase knock-out (Sis-KO) mice on two different diet regimens.

Methods

Ethics statement

All participants gave written informed consent, and the study was approved by the Scientific Ethics Committee in Greenland (cohort I: project 2011–13 (ref. no. 2011–056978), project 2013–13 (ref.no. 2013–090702), and project 2012-16/17 (ref.no. 2017-12997); cohort II: project 2013-17), and was conducted in accordance with the Declaration of Helsinki, second revision.

Study populations

Cohort I comprised Greenlanders living in Greenland, from population surveys during 1999-2001 (B99, *N*=1,401) and 2005-2010 (IHIT, *N*=3,115), as well as Greenlanders living in Denmark collected as part of the B99 survey (BBH, *N*=547)[20,21]. Cohort II was collected in 2013 as a population-based sample of Greenlanders (*N*=1,629)[22]. Basic clinical data for cohort I and cohort II are presented in Supplementary Table 1.

Assays and measurements

Cohort I

Anthropometrics, concentrations of fasting serum lipids, plasma apolipoproteins AI and B as well as levels of fatty acids were measured, and BMI, fat percentage, lean mass, LDL, VLDL, and remnant-cholesterol calculated as previously described in detail[23,24]. All IHIT participants above 18 years, and B99 participants above 35 years, underwent an oral glucose tolerance test (OGTT), and serum insulin, plasma glucose, serum c-peptide, and HbA1c were measured, and homeostasis model assessment of insulin resistance (HOMA-IR) calculated[24]. Type 2 diabetes was defined based on the WHO 1999 criteria, and controls were defined as normal glucose tolerant based on the OGTT data.

Daily intake of macronutrients, selected types of carbohydrates and fat, as well as total energy was calculated based on data from food frequency questionnaires and published food tables[25,26]. Information on the participants overall health and gastrointestinal health was obtained from questionnaires, and was

analyzed with a case-control design. We classified cardiovascular disease events based on data from registries (Supplementary Table 2 and Supplementary Table 3).

Cohort II

Height and weight were measured, and BMI calculated. Samples for measurement of serum metabolites, characterized with a high-throughput NMR metabolomics platform[27,28], and plasma samples for measurement of alkaline phosphatase, albumin, aspartate aminotransferase, and bilirubin were collected at a clinical visit, without prior fasting.

Genotyping

The *SI* c.273-274delAG variant was genotyped using the KASP Genotyping Assay (LGC Genomics) in 4,922 and 1,629 individuals from cohorts I and II, respectively. The genotyping call rate was 99.4% in both cohorts, and there were no mismatches in 357 individuals genotyped in duplicate in cohort I.

Association analyses

Prior to analyzing, quantitative traits were transformed independently for men and women, using a rank-based inverse normal transformation, and effect size estimates were reported in standard deviations (β_{SD}) as well as in non-transformed trait values (β). We applied a linear mixed model, to take admixture and relatedness into account by including them as random effects. We estimated a genetic similarity matrix with GEMMA (v0.95alpha)[29] from SNPs with MAF of minimum 5% and missingness of maximum 1% from previously generated genome-wide genotype data from the Illumina Metabochip (Illumina, San Diego, CA, USA) and Illumina OmniExpressExome chip (Illumina) for cohort I[23] and cohort II[22], respectively. The estimated genetic similarity matrix was used as input for association testing. For quantitative traits, we included sex, age, and survey as covariates, and association tests were performed with GEMMA using a score test, whereas effect sizes and standard errors were estimated using a restricted maximum likelihood approach. For dichotomous traits, association tests were performed with the GMMAT package[30] in R, odds ratios and *P* values were obtained from a logistic mixed model using the Wald test including sex, age, and survey as covariates.

A full model, allowing for separate effects of being heterozygous and homozygous carriers of the c.273_274delAG variant, showed a strong effect on metabolic traits in homozygous carriers, but no effect

in heterozygous carriers (Supplementary Table 4). Hence, we report results generated with a recessive model, unless otherwise stated. For discovery analyses in cohort I and cohort II, *P* values below $7.2x10^{-4}$ and $3.1x10^{-4}$, respectively, corresponding to Bonferroni correction were considered statistically significant. We verified that the linear mixed model was able to account for admixture by performing association analyses for BMI and triglycerides in cohort I, split according to Inuit ancestry proportion (Supplementary Figure 1). Also, we performed a test for each of these traits against common variants on the Metabochip to ensure that the test statistics were not inflated (Supplementary Figure 2).

Analyses of register-based cardiovascular disease data

We applied a Cox regression, adjusted for sex, birth year (as number of years since 1900), survey, and the top 10 principal components, to estimate the number of years lived until the first cardiovascular event, until getting censored, or until the conclusion of the study (December 31, 2016) with the R-package *survival* (https://cran.r-project.org/web/packages/survival/index.html). We allowed individuals to have their first event counted in each type of event analysed. For information about selection analysis and estimation of allele frequencies in ancestral population components and in other populations see Supplementary Methods.

Sucrase-isomaltase knock-out mice

The mice experiments adhered to the Animal Research: Reporting of In Vivo Experiments guidelines, and were approved by the Animal Experiments Inspectorate. Heterozygous breeding pairs of C57BL/6NJ-Sisem1(IMPC)J mice were obtained from The Jackson Laboratory. Litters were weaned at 7-8 weeks and separated into new cages by sex. Unless specifically stated, all mice were kept in individually ventilated cages (IVC) (Scanbur). Groups were matched by littermate. The facility was humidity controlled and temperature was 23ºC; the light cycle was from 6:00-18:00.

Diets

All diets were ordered from Research Diets Inc., and matched as much as possible for macronutrients and ingredient composition. For the choice diet experiment, wild-type (Sis-WT, n=9) and knock-out (Sis-KO, n=13) littermate mice between the ages of 8-29 weeks were separated by sex. Males (n=6 Sis-WT, n=7 Sis-KO) were individually caged, and females (n=3 Sis-WT, n=6 Sis-KO) were group caged in IVC. To

ensure sucrose intake, the mice had ad-libitum access to high-fat 12.6 kcal% sucrose (HFS) diet (Research Diets #D12331), low-fat 17 kcal% sucrose (17S) (Research Diets #D12450H) diet, and low-fat no-sucrose (NS) diet (Research Diets #D12450K) for 8 weeks (Supplementary Table 5 and Supplementary Table 6). For the HFNS diet experiment, Sis-WT (n=6) and Sis-KO (n=6) littermate mice between the ages of 5-11 weeks were placed into a mixture of group (Sis-WT n=4, Sis-KO n=4) and individual (Sis-WT n=2, Sis-KO n=2) caging according to how they arrived due to lack of room in the animal housing units to individually house all mice (Supplementary Table 7). Mice had ad-libitum access to high-fat no-sucrose (HFNS) diet (Research Diets #D0806014B) for 8 weeks (Supplementary Table 6).

Sucrose gavage, tolerance test and plasma measurements

Mice were given an oral gavage of sucrose (3g/Kg body weight) following an overnight fast. For the sucrose tolerance test, blood was taken from the tail vein of Sis-WT (n=7) and Sis-KO (n=7) mice and blood glucose was determined by glucometer (Roche) at 0, 15, 30, 60 and 120 minutes. To quantify plasma acetate and conversion of sucrose to short-chain fatty acids 75ul of blood was collected from Sis-WT (n=5) and Sis-KO (n=7) by retro orbital bleed. This was performed on two separate occasions due to the maximum sampling volume and recovery times for a mouse (i.e. bleed one for 0 and 2h time points and bleed two for 4h and 6h time points).

Measurements

Food intake was calculated as weekly intake, by weighing the amount of each diet given at the beginning of each week, and at the same time 7-days later after a thorough search of the cages. An average per mouse was calculated for multi-caged mice.

Individual weights were measured at baseline and at the end of each week following placement of mice on diets. Fat and lean mass were measured using a Minispec LF90II low frequency NMR system (Bruker) in the case of the HFNS experiment or an EchoMRI™-500 for mice in the choice diet experiment. Mice were awake during procedure and immobilized using a plunger system. The Minispec system was applied to measure total lean mass, fat mass, and free fluid. Body fat fraction was calculated as a percentage of total mass determined from the sum of fat mass, lean mass, and free fluid analysed by the system in Microsoft Excel (Office 2009). Liver tryglycerides were determined using the glycerine phosphate oxidase peroxidase (GPO-PAP) colorimetric assay (Randox #TR210) according to manufacturer's

instructions. Plasma levels of acetate was measured by LC-MS (for additional information see Supplementary Methods).

Statistical analysis

To test for differences in weight gain, fat percentage, and lean mass gain at each of the eight weeks separately, we used a linear model adjusted for sex. Confidence intervals were estimated using a profile likelihood approach. For the sucrose gavage experiments, sex was not included in the model, as all mice were female.

Results

Frequency of c.273_274delAG in Greenlanders and in other populations

In cohort I and cohort II, the frequency of the *SI* c.273_274delAG variant was 14.2% (95% CI: 13.5-15.1) and 14.1% (12.8-15.3), and the number of homozygous carriers was 99 and 34, respectively (Supplementary Table 1). The Greenlandic population is admixed, and we estimated the Inuit ancestry-specific allele frequency in cohort I to be 20.0% (19.0-21.1). We also estimated the frequency of the variant in populations from across the world, using publicly available datasets, and found it to be close to zero in non-Arctic populations, except in Siberians (Supplementary Table 8). Despite the higher frequency of the variant among Greenlanders, and in particular Inuit, we observed no signatures of selection at the locus (Supplementary Figure 3).

Anthropometric and metabolic traits

In cohort I, homozygous carriers of the c.273_274delAG variant had a healthier metabolic profile, and results from a full model showed that these effects were mainly recessive (Figure 1 and Supplementary Table 4). Specifically, with a recessive model we found that homozygous carriers had markedly lower BMI (β (SE), -2.0 kg/m^2 (0.5), P=3.1x10^{-5}), smaller waist and hip circumference (-4.9 cm (1.3), P=1.8x10^{-4} and -3.3 cm (0.9), P=2.3x10^{-4}), and lower weight (-4.8 kg (1.4), P=5.1x10^{-4}). Homozygous carriers also had less body fat (subcutaneous adipose tissue, -0.70 cm (0.17), P=5.8x10^{-7}; subcutaneous adipose tissue to visceral adipose tissue ratio, -0.08 (0.03), P=3.8x10^{-6}; fat percentage, -3.3% (1.0), P=3.7x10^{-4}), and a healthier lipid profile (triglyceride, -0.27 mmol/L (0.07), P=2.3x10^{-6}; remnant

cholesterol, -0.11 mmol/L (0.03), P=4.2x10^{-5}; VLDL-cholesterol, -0.13 mmol/L (0.04), P=6.0x10^{-4}; Supplementary Table 9). We observed no association between the variant and risk of type 2 diabetes or traits related to glucose homeostasis (Supplementary Table 9).

In the smaller cohort II, we replicated the association with lower BMI and lower weight, with comparable effect sizes, whereas the association with lower level of triglyceride was non-significant, however, with a comparable effect size (Supplementary Table 10). Moreover, from markers of liver health, we observed significantly lower levels of alkaline phosphatase among homozygous carriers (-15.41 U/l (4.20), P=9.8x10^{-6}) (Supplementary Table 10).

Additional markers of metabolic health

To further understand the impact of the variant, we tested for associations with circulating metabolic markers measured by NMR spectroscopy, available for cohort II. Interestingly, we observed markedly higher levels of circulating acetate in homozygous carriers (β (SE), 0.056 mmol/L (0.002), P=2.1x10^{-26}; Figure 1 and Supplementary Table 11), but no significant associations with markers of glycolysis, ketone bodies, or amino acids when adjusting for multiple testing (Supplementary Table 11).

With respect to lipoproteins, the variant had the strongest impact on HDL metabolism, with significantly higher concentrations of very large HDL particles (β_{SD} (SE), 0.621 SD (0.167), P=2.1x10^{-4}) (Figure 2 and Supplementary Table 12), and significantly higher content of free cholesterol, cholesterol esters, total cholesterol, and total lipids (P<2.5x10^{-4} for all), as well as a nominally higher content of phospholipids (P=4.4x10^{-4}) in these particles (Supplementary Table 12).

From the NMR measurements in cohort II, we also assessed the fatty acid composition in serum. Relative to the total amount of fatty acids, we found significantly higher levels of PUFA (0.704 SD (0.172), P=4.7x10^{-5}), total omega-6 fatty acids (0.883 SD (0.166), P=1.2x10^{-7}), and linoleic acid (0.956 SD (0.163), P=5.8x10^{-9}) in homozygous carriers, as well as lower levels of MUFA (-0.822 SD (0.169), P=1.2x10^{-6}) (Supplementary Table 13). For comparison, we assessed the fatty acid composition in erythrocyte membranes in cohort I, and validated the association with higher levels of omega-6 fatty acids (0.253 SD (0.107), P=0.018) and linoleic acid (0.371 SD (0.102), P=2.6x10^{-4}) in homozygous carriers. Additionally, we observed significantly lower levels of oleic acid (-0.450 SD (0.125), P=3.2 x10^{-4}; Supplementary Table 14).

Gastrointestinal and cardiovascular health

In questionnaire-based data from cohort I, we observed no significant associations with neither gastrointestinal symptoms nor overall health perception (Table 1). With respect to cardiovascular disease events, queried from register-based data from cohort I, effect estimates indicated a lower risk of ischemic heart disease and heart failure in homozygous carriers, however, this risk reduction was statistically non-significant (Figure 3 and Supplementary Table 15).

Dietary composition

In cohort I, the daily intake of added sugar, i.e. sucrose, was significantly lower among homozygous carriers (β (SE), -28.55 g/day (7.92), P=2.8x10^{-7}), whereas we found no significant differences in intake of protein, fat, including the specific fat categories monounsaturated fatty acids (MUFA), polyunsaturated fatty acids (PUFA), and saturated fat, or carbohydrates, including fiber, whole grain, refined grain, and fruit. In line with the non-significant differences for the majority of these dietary components, there was no difference in total daily energy intake (Table 2).

Analyses of factors potentially mediating the association between c.273_274delAG and metabolic health

In cohort I, we tested whether the lower intake of added sugar among homozygous carriers of c.273_274delAG, could explain their healthier metabolic phenotype, but the associations with anthropometric and metabolic traits remained when adjusting for intake of added sugar (Supplementary Table 16). Next, we tested in cohort II, whether serum acetate levels might mediate the associations, and found that associations were attenuated for BMI (P=0.077), weight (P=0.160), and alkaline phosphatase (P=0.018), when adjusting for acetate level indicating that serum acetate might mediate these associations (Supplementary Table 10).

Characterization of sucrase-isomaltase knock-out mice

To further investigate our findings, we studied sucrase-isomaltase knock-out (Sis-KO) mice. Mimicking a Westernized human diet, Sis-KO and wild type (Sis-WT) mice were fed a choice diet with ad libitum access to high-fat 12.6% sucrose (HFS) diet, low-fat 17% sucrose (17S) diet, and low-fat no-sucrose (NS) diet for 8 weeks. There was no difference in the overall calorie intake between Sis-KO and Sis-WT mice (Figure 4A), but the choice of diet differed slightly, with the Sis-KO mice having a lower intake of

sucrose in form of the 17S diet, and a higher intake of the NS diet (Supplementary Figure 4). Even though the caloric intake was similar, Sis-KO had significantly lower weight gain (week 8, β=-3.06 g, *P*=0.029), and lower body fat percentage (week 8, β=-10.2%, *P*=0.0013), but similar lean mass gain (week 8, β=0.62 g, P=0.252) compared to Sis-WT mice (Figure 4B-D and Supplementary Table 17). To investigate the effect of sucrose in the diet, we repeated the experiment where Sis-KO and Sis-WT mice had ad libitum access to a high-fat no-sucrose (HFNS) diet for 8 weeks. Again, the total calorie intake over 8 weeks was similar in Sis-KO and Sis-WT mice (Figure 4E), but without the sucrose in the diet, we observed no differences in weight gain, fat fraction, or lean mass gain between Sis-KO and Sis-WT mice (Figure 4F-G and Supplementary Table 18). Before sacrificing the mice, we measured hepatic triglyceride levels, and observed around 20% lower levels in the Sis-KO mice on the Choice diet, but no difference in mice on HFNS diet. This difference was in the same direction and of the same magnitude as observed in serum in the Greenlanders, but was non-significant (Supplementary Figure 5).

To further explore the mechanism underlying the observed phenotype, Sis-KO and Sis-WT mice were gavaged with 3 g/kg sucrose after a 12-hour fast. At 4 and 6 hours post-gavage, plasma acetate levels in the Sis-KO mice had increased to a level several times higher than the level in Sis-WT mice (4 hours, $P=0.0037$; 6 hours, $P=4.0\times10^{-4}$; Figure 5 and Supplementary Table 19). A separate gavage experiment showed that the Sis-KO mice took up less sugar in response to a 120 minute sucrose-tolerance test, which resulted in lower levels of plasma glucose (15 minutes, $P=2.8\times10^{-4}$; 30 minutes, $P=9.1\times10^{-5}$; Figure 5 and Supplementary Table 20), compared to Sis-WT mice.

Discussion

We assessed the impact of the *SI* c.273-274delAG loss-of-function variant in Greenlandic adults, and *Sis* knock-out in mice. In humans, the c.273-274delAG variant was only observed in Arctic and Siberian populations, and its frequency was estimated to 20% in the Inuit ancestry component of the Greenlanders. Given that we found no signatures of positive selection, this high frequency among Inuit, compared to other populations, has likely been possible due to lack of negative selection pressure combined with strong genetic drift, which is a particularly powerful process affecting small isolated populations like the Greenlandic. Interestingly, in adults we found that genetic loss of SI function was associated with a substantially healthier metabolic profile, with lower BMI, body weight, and fat percent, as well as a favorable lipid profile. Importantly, we replicated the associations with lower BMI and body weight in an independent cohort of Greenlanders. In this other cohort, we also found that homozygous carriers had

markedly higher levels of circulating acetate, which was likely only detectable due to the lack of fasting in these participants. Notably, the effect of the naturally occurring specific loss of SI function, seemed to be greater on weight and levels of triglycerides, compared to drug induced unspecific inhibition of α-glucosidases by acarbose, voglibose, or miglitol[31–37], and the impact of loss of SI function on triglyceride levels was equal to the reported effect of the lipid lowering drug statins[38,39]. Moreover, altered HDL-metabolism among homozygous carriers, suggested increased health promoting removal of cholesterol from extrahepatic tissues. We also observed lower serum concentrations of alkaline phosphatase among homozygous carriers. This might be a consequence of the lower degree of adiposity[40], but could potentially be an indication of a healthier liver function. Even though SI affects the ability to metabolize sugar, we observed no effect on glucose homeostasis in homozygous carriers of the c.273-274delAG variant. However, a difference in measures from the OGTT is not necessarily expected, as loss of SI function should not affect the uptake of glucose. In response to intake of food containing sucrose or isomaltose, a lower uptake of sugar could be expected, but was not apparent from HbA_{1c} measures. This could be due to compensatory mechanisms of higher hepatic gluconeogenesis to sustain blood glucose levels, which is in line with observations from previous studies of MGAM-KO mice[41]. When testing for other effects of the variant, we did not find any significant associations with overall self-reported well-being or risk of cardiovascular disease, which could be due to the limited number of events in our analyses. A follow-up study with a larger sample size and longer follow-up is necessary to determine the potential cardio-protective effects of the SI loss-of-function.

To elucidate the mechanism underlying the healthier metabolic profile associated with loss of SI function, we firstly investigated intake of selected dietary components in the Greenlanders. These analyses suggested that compared to other Greenlanders, the homozygous carriers of the c.273-274delAG variant did not have a significantly different intake of total energy or intake of any specific dietary component, except for added sugar. Hence, the observed differences in lipid levels could not be explained by differences in the composition of dietary fat. Moreover, conditional analyses showed that the lower intake of added sugar did not explain the healthier metabolic profile. Secondly, we performed analyses conditional on acetate levels. Interestingly, adjusting for acetate attenuated the observed associations with a healthier metabolic profile, suggesting that higher levels of circulating acetate could be part of the functional link, between the lack of SI function and improved metabolic health. Thirdly, we performed several mice experiments. In line with the observations in the humans, Sis-KO mice on a diet mimicking a Westernized diet, had a slightly lower intake of sucrose, but a similar total energy intake as the Sis-WT mice. Yet, compared to the Sis-WT mice, the Sis-KO mice gained significantly less weight and had lower fat percentage gain, as well as lower liver triglyceride levels. These findings indicated that the healthier

metabolic profile linked to loss of SI function, in both humans and mice, is likely caused in part by altered intestinal sucrose uptake, rather than altered amounts of total energy intake. In the mice, a sucrose tolerance test clearly demonstrated, that with loss of SI function, sucrose uptake was diminished, indicated by significantly lower levels of plasma glucose. Also, we found that the healthier metabolic profile, associated with loss of SI function, was dependent on presence of dietary sucrose, as Sis-KO mice on a HFNS diet displayed a body composition similar to Sis-WT mice. We, therefore, hypothesize that the metabolic health promoting effect was mediated by increased colonic bacterial fermentation of undigested carbohydrates, particularly sucrose and isomaltose, escaping small intestinal digestion due to the loss of SI function. Increased bacterial fermentation of these carbohydrates may also explain the markedly higher circulating levels of the short chain fatty acid (SCFA) acetate, we observed in humans with loss of SI function. This hypothesis was strongly supported by induction of significantly higher levels of plasma acetate in Sis-KO mice after a sucrose gavage. With the available data, it is not possible to exclude the possibility that other processes, including ketogenesis, contributed to the higher levels of acetate in humans and mice with loss-of SI function. However, it seems unlikely that a 12-hour fast and 6 hours of gavage experiment, could induce increased acetate production by ketogenesis, as a much longer fast of 48 hours did not induce higher acetate levels in previous mice studies[42]. Moreover, the level of ketone bodies in the Greenlanders did not differ between homozygous carriers and the rest of the study population, which indicated that ketogenesis was not increased among homozygous carriers of the variant. Also, our hypothesis of increased gut bacterial acetate production in response to loss of SI function, is supported by previous human studies showing that acarbose treatment is associated with higher fecal concentration of starch and starch-fermenting bacteria, as well as higher levels of SCFAs in feces and circulation[43–45]. In line with this, a common *SI* missense variant (rs9290264), estimated to reduce the SI enzymatic activity by 35%, has been associated with lower abundance of the gut bacterial genus *Parabacteroides*[46], which has been associated with changes in body weight and fat mass[47].

Whether increased circulating levels of acetate in fact is beneficial has been debated. Some rodent studies have indicated that acetate is linked to increased lipogenesis and possibly induces components of the metabolic syndrome[48,49]. However, a range of studies show that increased levels of circulating acetate, obtained by direct administration or by increased microbial production induced by diet, are linked to lower body weight and lower levels of plasma cholesterol in most studies of humans and rodents[50]. Whether acetate is beneficial or harmful could depend on the site of acetate catabolism. It has been shown, that activated hepatic acetate uptake and catabolism can induce de novo lipogenesis, and thereby hepatic lipid accumulation[49]. The fact that we observed markedly higher levels of acetate in circulation, could indicate that acetate bypasses hepatic catabolism, and thereby reaches systemic circulation, where it

might induce beneficial signaling pathways in other tissues, including liver, brain, muscle, and adipose tissue. In humans, colonic infusion of acetate has been shown to increase fat oxidation and inhibit lipolysis resulting in lower levels of circulating free fatty acids, and lower flux of fatty acids to the liver[51,52]. These effects likely result in reduced hepatic synthesis of triglycerides[53], in line with our observation of lower levels of fasting serum triglycerides in Greenlanders homozygous for the variant, and lower levels of liver triglycerides in Sis-KO mice. Further contributing to improved metabolic health, acetate has been shown to increase resting energy expenditure and affect appetite regulation in humans[52], and to induce adipogenesis in mice[54]. The latter, indicating a healthier expansion of the lipid storage capacity in the adipose tissue.

Interestingly, both increasing the level of circulating acetate and targeting the α-glucosidases, including SI, are intensely studied as ways to improve metabolic health and to induce weight loss[36,50,55–58]. And given the markedly healthier metabolic profile among homozygous carriers of the *SI* loss-of-function variant, it seems relevant to consider specifically targeting SI with a drug to improve metabolic health. SI is a promising drug target, as the enzyme expression is highly specific to the small intestine, which may be favorable compared to targets in the central nervous system affecting appetite regulation, where more undesired effects are expected[59]. Notably, our study constitutes a particularly good first step towards such a consideration since homozygous carriers of the *SI* loss-of-function variant have great predictive value for benefits and side effects of targeting SI with a drug to improve metabolic health[60]. Support from naturally occurring human knock-outs has even been estimated to double the success rate of drug development[61]. In terms of side effects, it seems particularly relevant to consider possible gastrointestinal problems, as CSID is associated with severe gastrointestinal symptoms in children. We were unable to show any differences in self-reported digestive problems in the adult homozygous c.273-274delAG carriers compared to the rest of the Greenlandic study population. This discrepancy between adults and children, may be due to the maturation and growth of the small intestine, increasing the capacity to absorb luminal fluid with increasing age[16], and to dietary adaptation caused by symptoms in childhood. It has been shown that a common *SI* missense variant (rs9290264) was associated with increased risk of irritable bowel syndrome (IBS)[46], and that IBS patients show increased prevalence of rare *SI* variants[62]. However, for rs9290264 this conclusion was supported neither by analysis of 452,264 individuals from the UK biobank (http://geneatlas.roslin.ed.ac.uk/), nor by analyses of 117,050 individuals from the FinnGen study (http://r3.finngen.fi/), including 6,041 and 2,727 cases with IBS, respectively. This aspect should be addressed further in large studies with careful IBS phenotyping to verify whether inhibition of SI will result in unwanted side effects.

In conclusion, our data indicated that lack of SI function in human adults, and in mice, seems to be specifically linked to altered uptake, and metabolism of dietary components, which result in a healthier metabolic phenotype, likely mediated by decreased intestinal sucrose absorption and possibly also by increased levels of circulating acetate. Our data also indicated that isolated targeting of SI may refine the effects already reported for other α-glucosidase inhibitors, and thus that SI is a potential treatment target to improve metabolic health.

References

1. Galgani J, Ravussin E. Energy metabolism, fuel selection and body weight regulation. Int J Obes 2008;32:S109-19.

2. Azaïs-Braesco V, Sluik D, Maillot M, et al. A review of total & added sugar intakes and dietary sources in Europe. Nutr J 2017;16:1–15.

3. Rippe JM, Angelopoulos TJ. Sugars and health controversies: What does the science say? Adv Nutr 2015;6:493S-503S.

4. Bier DM. Dietary Sugars: Not as Sour as They Are Made Out to Be. 2020;95:1–12.

5. Naim HY, Roth J, Sterchi EE, et al. Sucrase-isomaltase deficiency in humans. Different mutations disrupt intracellular transport, processing, and function of an intestinal brush border enzyme. J Clin Invest 1988;82:667–679.

6. Naim HY, Sterchi EE, Lentze MJ. Structure, biosynthesis, and glycosylation of human small intestinal maltase-glucoamylase. J Biol Chem 1988;263:19709–19717.

7. **Gericke B, Amiri M,** Naim HY. The multiple roles of sucrase-isomaltase in the intestinal physiology. Mol Cell Pediatr 2016;3:2.

8. **Lee BH, Eskandari R**, Jones K, et al. Modulation of starch digestion for slow glucose release through "Toggling" of activities of mucosal α-glucosidases. J Biol Chem 2012;287:31929–31938.

9. Yoon SH, Robyt JF. Study of the inhibition of four alpha amylases by acarbose and its 4 IV-α-maltohexaosyl and 4IV-α-maltododecaosyl analogues. Carbohydr Res 2003;338:1969–1980.

10. Mohan S, Eskandari R, Pinto BM. Naturally occurring sulfonium-ion glucosidase inhibitors and their derivatives: a promising class of potential antidiabetic agents. Acc Chem Res 2014;47:211–25.

11. Lembcke B, Fölsch UR, Creutzfeldt W. Effect of 1-desoxynojirimycin derivatives on small intestinal disaccharidase activities and on active transport in vitro. Digestion 1985;31:120–127.

12. Chen X, Zheng Y, Shen Y. Voglibose (Basen®, AO-128), One of the Most Important α-Glucosidase Inhibitors. Curr Med Chem 2009;13:109–116.

13. Nichols BL, Avery SE, Karnsakul W, et al. Congenital maltase-glucoamylase deficiency associated with lactase and sucrase deficiencies. J Pediatr Gastroenterol Nutr 2002;35:573–579.

14. Lebenthal E, Khin-Maung-U, Zheng BY, et al. Small intestinal glucoamylase deficiency and starch malabsorption: A newly recognized alpha-glucosidase deficiency in children. J Pediatr 1994;124:541–546.

15. Treem WR. Congenital sucrase-isomaltase deficiency. J Pediatr Gastroenterol Nutr 1995;21:1–14.

16. Treem WR. Clinical aspects and treatment of congenital sucrase-isomaltase deficiency. J Pediatr Gastroenterol Nutr 2012;55 Suppl 2:S7-13.

17. Antonowicz I, Lloyd-Still JD, Khaw KT, et al. Congenital sucrase-isomaltase deficiency. Observations over a period of 6 years. Pediatrics 1972;49:847–53.

18. Marcadier JL, Boland M, Scott CR, et al. Congenital sucrase-isomaltase deficiency: Identification of a common Inuit founder mutation. CMAJ 2015;187:102–107.

19. Malyarchuk BA, Derenko M V., Denisova GA. The frequency of inactive sucrase-isomaltase variant in indigenous populations of Northeast Asia. Russ J Genet 2017;53:1052–1054.

20. Bjerregaard P, Curtis T, Borch-Johnsen K, et al. Inuit health in Greenland: a population survey of life style

and disease in Greenland and among Inuit living in Denmark. Int J Circumpolar Health 2003;62 Suppl 1:3–79.

21. Bjerregaard P. *Inuit Health in Transition Greenland survey 2005-2010 Population sample and survey methods*. 2011. Available at: http://www.si-folkesundhed.dk/upload/inuit_health_in_transition_greenland_methods_5_2nd_revision.pdf.

22. Skotte L, Koch A, Yakimov V, et al. CPT1AMissense Mutation Associated With Fatty Acid Metabolism and Reduced Height in Greenlanders. Circ Cardiovasc Genet 2017;10:e001618.

23. **Andersen MK, Jørsboe E**, Sandholt CH, et al. Identification of Novel Genetic Determinants of Erythrocyte Membrane Fatty Acid Composition among Greenlanders Zeggini E, ed. PLoS Genet 2016;12:e1006119.

24. Andersen MK, Jørsboe E, Skotte L, et al. 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 2020;16:e1008544.

25. Jeppesen C, Bjerregaard P. Consumption of traditional food and adherence to nutrition recommendations in Greenland. Scand J Public Health 2012;40:475–481.

26. Jeppesen C, Jørgensen ME, Bjerregaard P. Assessment of consumption of marine food in Greenland by a food frequency questionnaire and biomarkers. Int J Circumpolar Health 2012;71:18361.

27. **Soininen P, Kangas AJ**, Würtz P, et al. High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. Analyst 2009;134:1781.

28. **Soininen P, Kangas AJ**, Würtz P, et al. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. Circ Cardiovasc Genet 2015;8:192–206.

29. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet 2012;44:821–4.

30. **Chen H, Wang C**, Conomos MP, et al. Control for Population Structure and Relatedness for Binary Traits in Genetic Association Studies via Logistic Mixed Models. Am J Hum Genet 2016;98:653–666.

31. Baron AD, Eckel RH, Schmeiser L, et al. The effect of short-term alpha-glucosidase inhibition on carbohydrate and lipid metabolism in type II (noninsulin-dependent) diabetics. Metabolism 1987;36:409–15.

32. Kado S, Murakami T, Aoki A, et al. Effect of acarbose on postprandial lipid metabolism in type 2 diabetes mellitus. Diabetes Res Clin Pract 1998;41:49–55.

33. Ogawa S, Takaeuchi K, Ito S. Acarbose lowers serum triglyceride and postprandial chylomicron levels in type 2 diabetes. Diabetes, Obes Metab 2004;6:384–390.

34. Narita T, Yokoyama H, Yamashita R, et al. Comparisons of the effects of 12-week administration of miglitol and voglibose on the responses of plasma incretins after a mixed meal in Japanese type 2 diabetic patients. Diabetes Obes Metab 2012;14:283–7.

35. Shimabukuro M, Higa M, Yamakawa K, et al. Miglitol, α-glycosidase inhibitor, reduces visceral fat accumulation and cardiovascular risk factors in subjects with the metabolic syndrome: A randomized comparable study. Int J Cardiol 2013;167:2108–2113.

36. Sugihara H, Nagao M, Harada T, et al. Comparison of three α-glucosidase inhibitors for glycemic control and bodyweight reduction in Japanese patients with obese type 2 diabetes. J Diabetes Investig 2014;5:206–212.

37. Domecq JP, Prutsky G, Leppin A, et al. Drugs commonly associated with weight change: A systematic review and meta-analysis. J Clin Endocrinol Metab 2015;100:363–370.

38. Stein EA, Lane M, Laskarzewski P. Comparison of statins in hypertriglyceridemia. Am J Cardiol 1998;81:66B-69B.

39. Manoria PC, Chopra HK, Parashar SK, et al. The nuances of atherogenic dyslipidemia in diabetes: Focus on triglycerides and current management strategies. Indian Heart J 2013;65:683–690.

40. **Johansen MJ, Gade J**, Stender S, et al. The Effect of Overweight and Obesity on Liver Biochemical Markers in Children and Adolescents. J Clin Endocrinol Metab 2020;105:dgz010.

41. Nichols BL, Quezada-Calvillo R, Robayo-Torres CC, et al. Mucosal maltase-glucoamylase plays a crucial role in starch digestion and prandial glucose homeostasis of mice 1-3. J Nutr 2009;139:684–690.

42. Sakakibara I, Fujino T, Ishii M, et al. Fasting-Induced Hypothermia and Reduced Energy Production in Mice Lacking Acetyl-CoA Synthetase 2. Cell Metab 2009;9:191–202.

43. Holt PR, Atillasoy E, Lindenbaum J, et al. Effects of acarbose on fecal nutrients, colonic pH, and short-chain fatty acids and rectal proliferative indices. Metabolism 1996;45:1179–1187.

44. Weaver GA, Tangel CT, Krause JA, et al. Acarbose enhances human colonic butyrate production. J Nutr 1997;127:717–23.

45. Wolever TMS, Chiasson J-L. Acarbose raises serum butyrate in human subjects withimpaired glucose tolerance. Br J Nutr 2000;84:57–61.

46. **Henström M, Diekmann L**, Bonfiglio F, et al. Functional variants in the sucrase-isomaltase gene associate with increased risk of irritable bowel syndrome. Gut 2018;67:263–270.

47. **Lecomte V, Kaakoush NO**, Maloney CA, et al. Changes in gut microbiota in rats fed a high fat diet correlate with obesity-associated metabolic parameters. PLoS One 2015;10:e0126931.

48. Perry RJ, Peng L, Barry NA, et al. Acetate mediates a microbiome-brain-β-cell axis to promote metabolic syndrome. Nature 2016;534:213–7.

49. Zhao S, Jang C, Liu J, et al. Dietary fructose feeds hepatic lipogenesis via microbiota-derived acetate. Nature 2020;579:586–591.

50. Hernández MAG, Canfora EE, Jocken JWE, et al. The short-chain fatty acid acetate in body weight control and insulin sensitivity. Nutrients 2019;11:1943.

51. Beek CM van der, Canfora EE, Lenaerts K, et al. Distal, not proximal, colonic acetate infusions promote fat oxidation and improve metabolic markers in overweight/obese men. Clin Sci (Lond) 2016;130:2073–2082.

52. Canfora EE, Beek CM Van Der, Jocken JWE, et al. Colonic infusions of short-chain fatty acid mixtures promote energy metabolism in overweight/obese men: A randomized crossover trial. Sci Rep 2017;7:2360.

53. Alves-Bezerra M, Cohen DE. Triglyceride metabolism in the liver. Compr Physiol 2018;8:1–22.

54. Hong YH, Nishimura Y, Hishikawa D, et al. Acetate and propionate short chain fatty acids stimulate adipogenesis via GPCR43. Endocrinology 2005;146:5092–5099.

55. Canfora EE, Blaak EE. Acetate: A diet-derived key metabolite in energy metabolism: Good or bad in context of obesity and glucose homeostasis? Curr Opin Clin Nutr Metab Care 2017;20:477–483.

56. **Depommier C, Everard A**, Druart C, et al. Supplementation with Akkermansia muciniphila in overweight and obese human volunteers: a proof-of-concept exploratory study. Nat Med 2019;25:1096–1103.

57. Zhang X, Li G, Wu D, et al. Emerging strategies for the activity assay and inhibitor screening of alpha-glucosidase. In: *Food and Function*.Vol 11. Royal Society of Chemistry; 2020:66–82.

58. Scott LJ, Spencer CM. Miglitol: A review of its therapeutic potential in type 2 diabetes mellitus. Drugs 2000;59:521–549.

59. Sugimoto S, Nakajima H, Kosaka K, et al. Review: Miglitol has potential as a therapeutic drug against obesity. Nutr Metab 2015;12:51.

60. Minikel EV, Karczewski KJ, Martin HC, et al. Evaluating drug targets through human loss-of-function genetic variation. Nature 2020;581:459–464.

61. Nelson MR, Tipney H, Painter JL, et al. The support of human genetic evidence for approved drug indications. Nat Genet 2015;47:856–860.

62. **Garcia-Etxebarria K, Zheng T**, Bonfiglio F, et al. Increased Prevalence of Rare Sucrase-isomaltase Pathogenic Variants in Irritable Bowel Syndrome Patients. Clin Gastroenterol Hepatol 2018;16:1673–1676.

Author names in bold designate shared co-first authorship.

Figure legends

Figure 1. Effect of the *SI* c.273_274delAG variant on selected metabolic phenotypes according to a full model. Raw mean values stratified by genotype (top) and untransformed effect sizes with 95% confidence intervals (bottom), for **A)** BMI, **B)** Fasting serum triglyceride, **C)** Fasting serum remnant cholesterol, and **D)** Serum acetate.

Figure 2. Effect of the c.273_274delAG variant on concentration of lipoprotein particles according to a recessive model. Effect estimates plotted as quantile transformed values, and error bars as 95% confidence intervals. HDL, High-density lipoprotein; IDL, Intermediate-density lipoprotein; L, large; LDL, Low Density Lipoprotein; M, medium; S, small; VLDL, Very low-density lipoprotein; XL, very large; XXL, extremely large; XS, very small.

Figure 3. Association between the c.273_274delAG variant and cardiovascular disease events.

Effects were estimated with a recessive model as hazard ratios (95% confidence intervals) based on register data from up to 4,551 individuals from cohort I (WT/HE/HO, 3355/1000/96). The numbers of individuals with an event according to genotype (WT/HE/HO) was 507/155/13 for any event, 242/64/3 for ischemic heart disease, 261/93/8 for cerebrovascular disease, and 128/30/2 for heart failure. Results for peripheral artery disease and coronary operations as not shown due to non-finite confidence intervals.

Figure 4. Energy intake and body composition of Sis-KO and Sis-WT mice.

Mean total energy intake is indicated by the horizontal lines for mice with ad libitum access to **A)** a choice diet of high-fat 12.6% sucrose (HFS), low-fat 17% sucrose (17S), and low-fat no-sucrose (NS) diet (Sis-WT, n=9 (single/multi-cage, 6/3); Sis-KO, n=13 (7/6)), and **E)** a high-fat no-sucrose (HFNS) diet (Sis-WT, n=6 (2/4); Sis-KO, n=6 (2/4)). Closed circles indicate single-caged mice, and open circles indicate multi-caged mice. Mean weekly weight gain, body fat fraction, and lean mass gain are indicated by circles, and standard errors are indicated by the error bars for **B-D)** Sis-WT (n=9) and Sis-KO (n=13) mice on the choice diet, and **F-H)** Sis-WT (n=6) and Sis-KO (n=6) mice on the HFNS diet. The asterisks indicate level of significance (*, $P<0.05$; **, $P<0.01$; and ***, $P<0.001$).

Figure 5. Blood plasma acetate and glucose levels following a sucrose gavage in Sis-KO and Sis-WT mice.

Mean values of **A**) plasma acetate levels (mmol/L), and **B**) plasma glucose levels (mmol/L) following a sucrose gavage. The points indicate mean values with error bars of standard errors. The asterisks indicate level of significance (*, $P<0.05$; **, $P<0.01$; and ***, $P<0.001$).

Tables

Table 1. Association between *SI* c.273_274delAG and gastrointestinal and overall health according to a recessive model

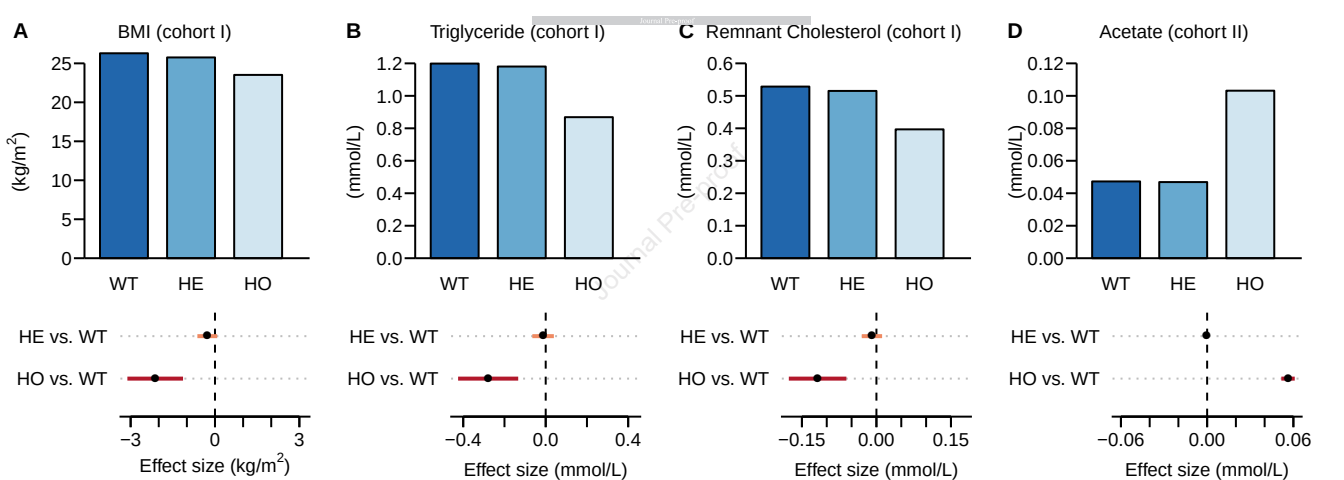
Trait	***N* (with/without condition)**	**OR (95% CI)**	***P* value**
Digestive problems	784/3058	1.58 (0.95-2.65)	0.081
Stomach pain	770/3051	0.83 (0.46-1.49)	0.530
Poor health	149/3762	0.52 (0.12-2.22)	0.380

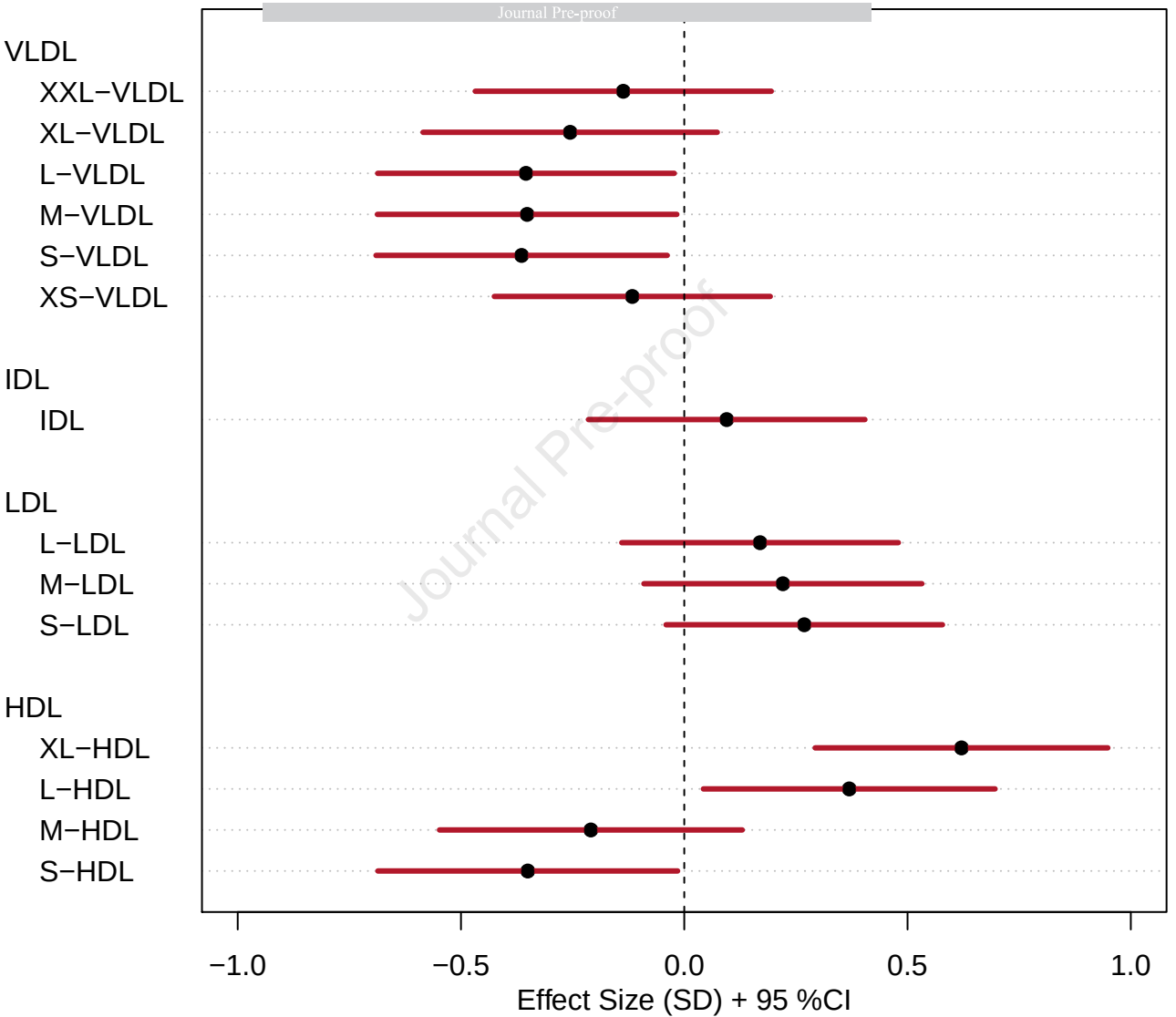
Data were questionnaire based and obtained from up to 3,911 individuals from cohort I. Effect sizes were estimated as odds ratios (OR) with 95% confidence intervals (95% CI).

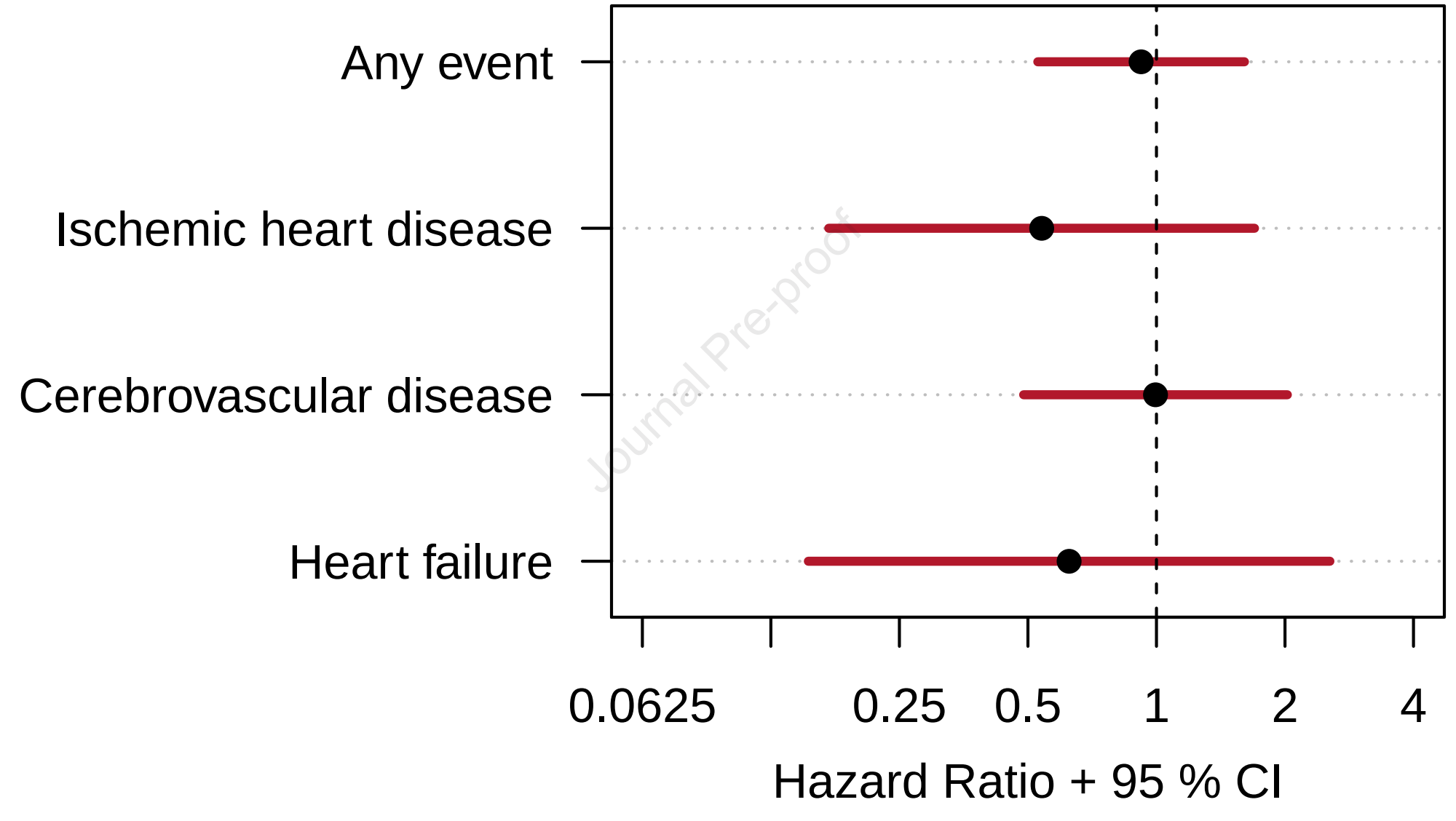
Table 2. Association between *SI* c.273_274delAG and questionnaire based diet information according to a recessive model

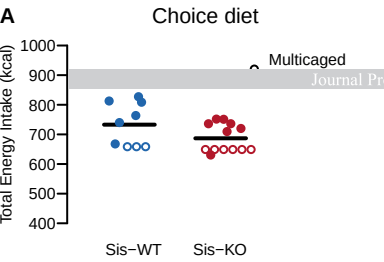
Trait	**β_{SD} (SE)**	**β (SE)**	***P* value**
Total energy (kJ/day)	-0.06 (0.13)	-146.69 (378.54)	0.634
Macronutrients			
Carbohydrate (g/day)	-0.24 (0.12)	-19.77 (12.43)	0.048
Protein (g/day)	0.07 (0.13)	2.92 (6.11)	0.603
Fat (g/day)	0.08 (0.13)	2.66 (4.26)	0.542
Fat components			
MUFA (g/day)	0.06 (0.13)	0.78 (2.38)	0.624
PUFA (g/day)	0.07 (0.13)	0.54 (1.08)	0.565
Saturated fat (g/day)	0.15 (0.13)	1.91 (1.53)	0.232
Carbohydrates			
Added sugar (g/day)	-0.65 (0.13)	-28.55 (7.92)	2.8×10^{-7}
Fruit (g/day)	-0.02 (0.13)	-6.36 (21.07)	0.874
Fiber (g/day)	0.10 (0.13)	1.39 (1.36)	0.419
Whole grain (g/day)	0.22 (0.13)	24.00 (15.29)	0.089
Refined grain (g/day)	0.05 (0.13)	3.00 (8.46)	0.710

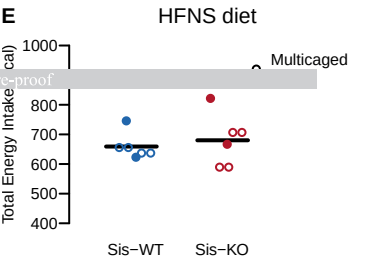
Results were obtained with a linear mixed model for 2,469 individuals from cohort I. Effect sizes are shown as quantile transformed (β_{SD}), or untransformed (β), and *P* values were calculated based on the quantile transformed values using the score test in GEMMA, including only individuals with a realistic energy intake.

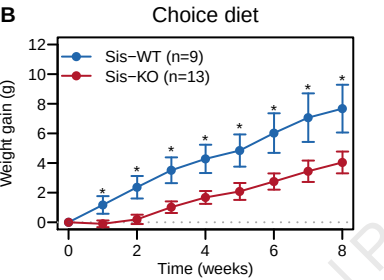

A
BMI (cohort I)
(kg/m2)
25
20
15
10
5
0
WT
HE
HO
B
Triglyceride (cohort I)
(mmol/L)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
WT
HE
HO
C
Remnant Cholesterol (cohort I)
(mmol/L)
0.6
0.5
0.4
0.3
0.2
0.1
0.0
WT
HE
HO
D
Acetate (cohort II)
(mmol/L)
0.12
0.10
0.08
0.06
0.04
0.02
0.00
WT
HE
HO
HE vs. WT
HO vs. WT
−3
0
3
Effect size (kg/m2)
HE vs. WT
HO vs. WT
−0.4
0.0
0.4
Effect size (mmol/L)
HE vs. WT
HO vs. WT
−0.15
0.00
0.15
Effect size (mmol/L)
HE vs. WT
HO vs. WT
−0.06
0.00
0.06
Effect size (mmol/L)

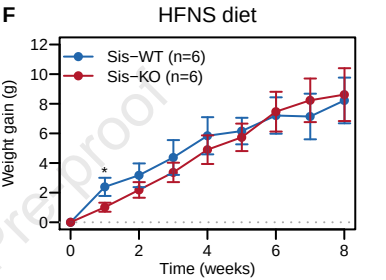


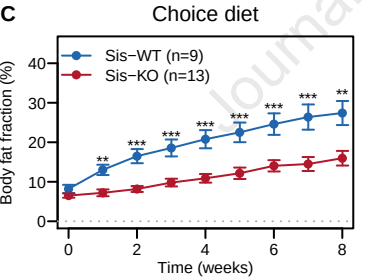


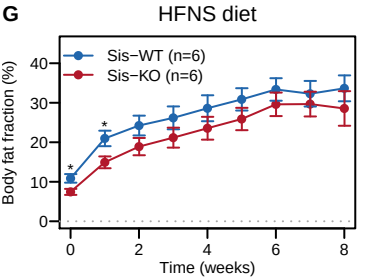


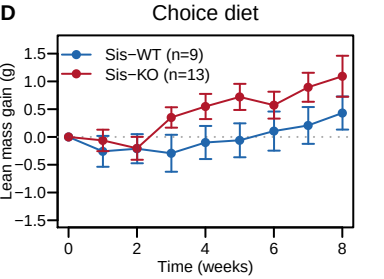


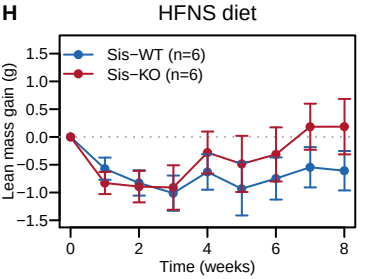


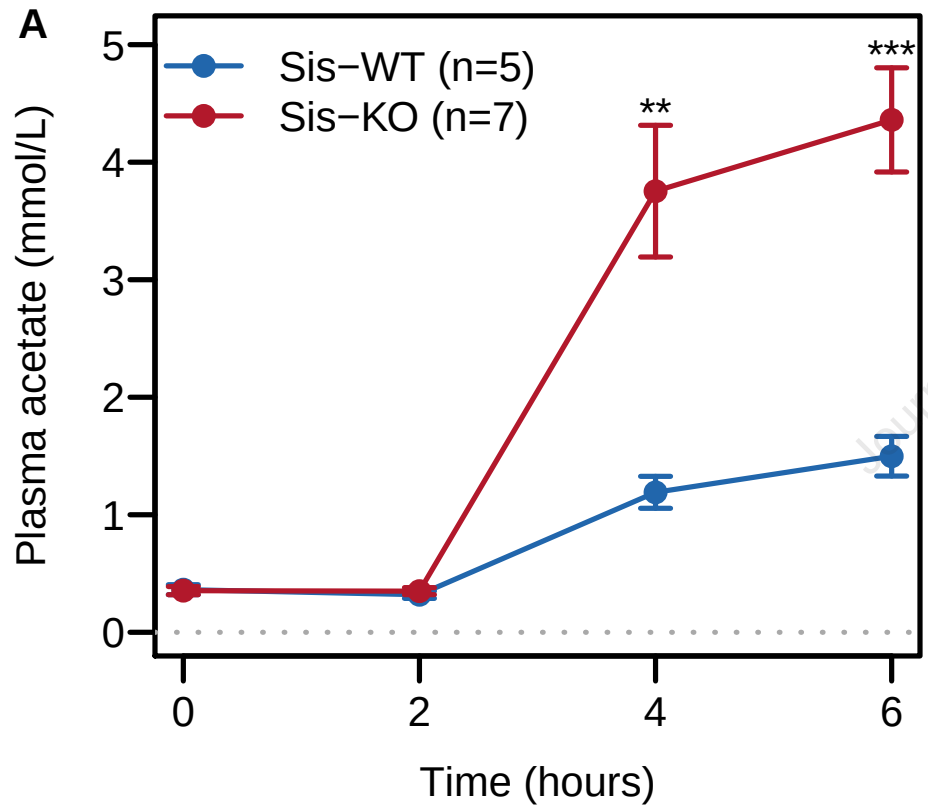


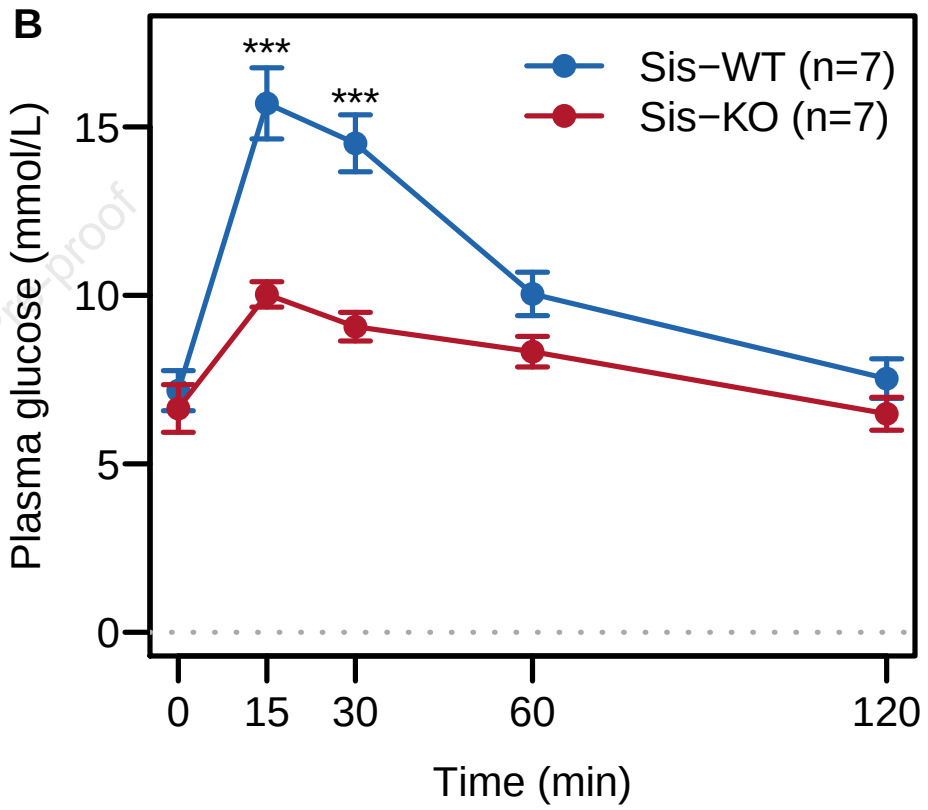












Supplementary Methods

Selection analysis

To assess whether the *SI* variant has been under positive selection we estimated extended haplotype homozygosity[1] and integrated haplotype score[2] at the *SI* c.273_274delAG variant (rs781470490), and across chromosome 3. Estimates were based on analyses of 263 unrelated Greenlanders without European ancestry from cohort I, identified by running an analysis of population structure with ADMIXTURE (v1.3)[3], and RELATEADMIX[4] on Illumina Metabochip SNP array data. To construct a data set for this analysis, we first selected all the 13195 sites from the Illumina Metabochip on chromosome 3 with less than 2% missing data. These data were then used for reference-based phasing and imputation with phased reference data from 40 trio-phased Greenlanders of Inuit descent and 190 individuals of European descent from the CEU and GBR populations from the 1000 Genomes (internationalgenome.org)[5]. The genotype data from cohort I was phased with SHAPEIT (v2.r904)[6] using this reference panel and the HapMap hg19 recombination map. Genetic variants were imputed onto the phased haplotypes with IMPUTE2 (v2.3.2)[7]. We used hapbin (v1.3.0)[8] to calculate EHH and iHS across chromosome 3.

Estimation of allele frequencies in ancestral population components and in other data sets

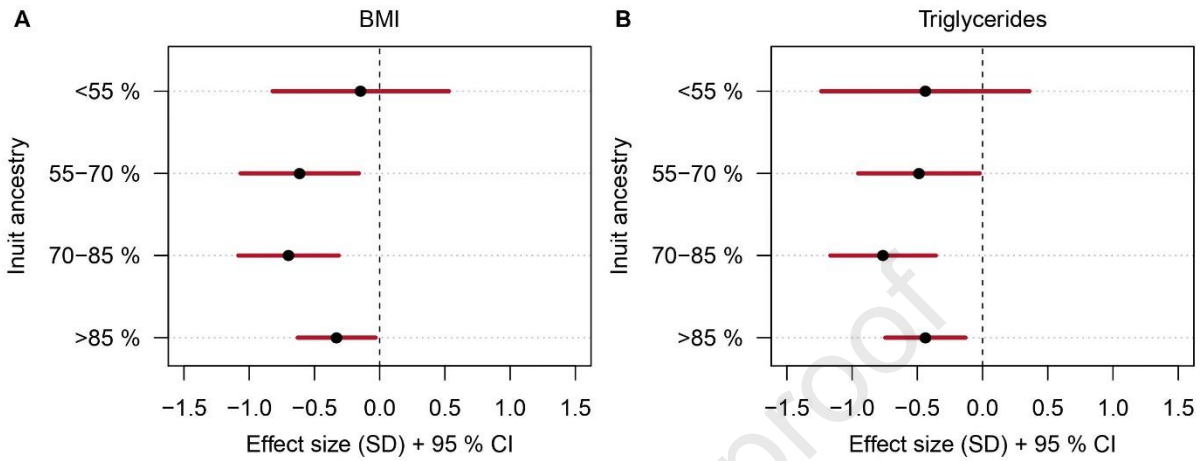
We estimated the *SI* c.273_274delAG frequency separately for the Inuit ancestry component of the admixed Greenlandic population by estimating ancestry proportions[3] for the Greenlandic individuals from cohort I, as well as for 50 Danish individuals, assuming two ancestral populations – Inuit and Europeans. Moreover, we surveyed the allele frequency of c.273-274delAG in a range of available datasets from across the world.

Measurement of plasma acetate in mice

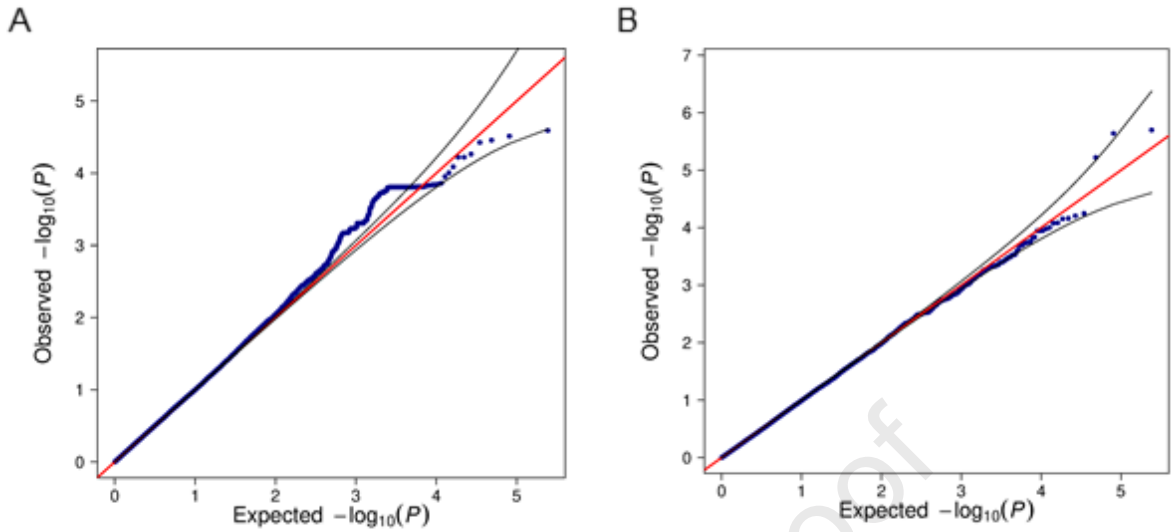
The derivatizing reagent was 200 mM EDC, 120 mM 3-Nitrophenylhydrazine, and pyridine (2% v/v) in 50% acetonitrile. Plasma (10 µL) was mixed with 10µl of stable isotope labelled internal standards (100µM of 13C4-acetate in 50% methanol) and derivatizing reagent (20 µL) and incubated for 1 hour at 40°C. Then, the samples were centrifuged at 14,000 RPM for 10 min at 4°C, and mixed with 40 µL of 0.1% formic acid. Eight different levels of acetate-calibrants were derivatised as the samples. The samples were injected into an Ultrahigh Performance Liquid Chromatography (UHPLC) system (Agilent 1290 Infinity II) connected to a Bruker timsTOF Pro™ instrument (Bruker, Bremen, Germany). Ions were generated in the negative electrospray ionization (ESI) mode. Data acquisition was performed with otofControl version 6.0 and Bruker Compass HyStar version 5.0 (Bruker Daltonics, Bremen, Germany) and data processing was performed with Bruker TASQ 2021b quantitation software. [M-H]- for acetate and internal standard was used as quantifier.

Supplementary Material

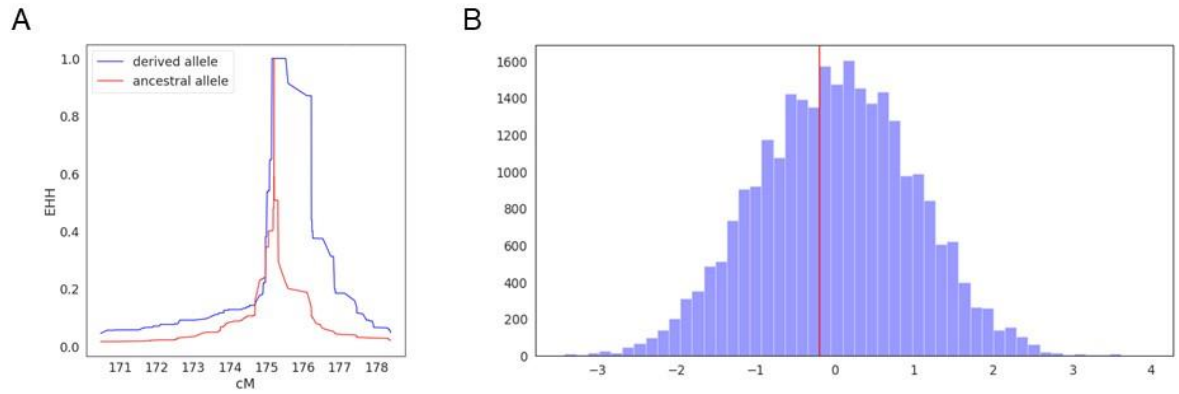
Supplementary Figures



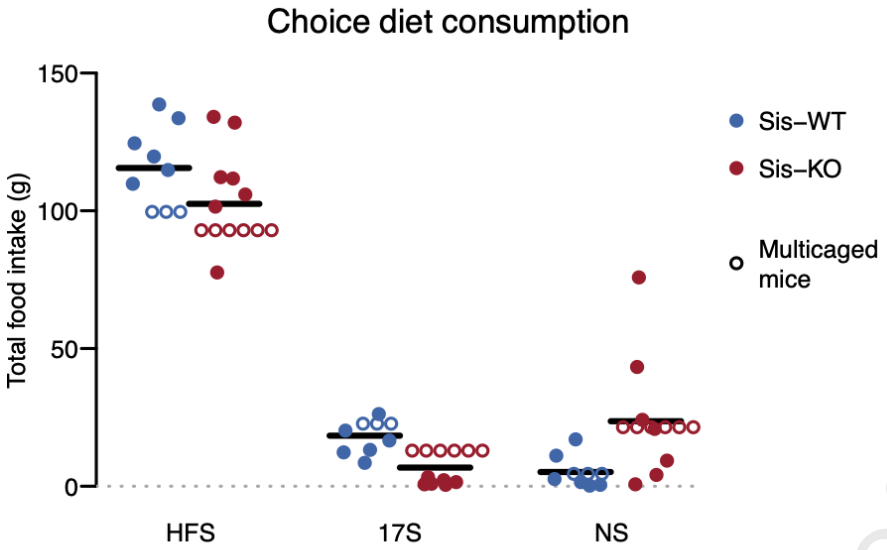
Supplementary Figure 1. Association results for cohort I divided into four subgroups according to Inuit ancestry proportion. Results from analysing association between *SI* c.273_274delAG and **A)** BMI, and **B)** Triglycerides for different subgroups of cohort I using a linear mixed model assuming a recessive effect.



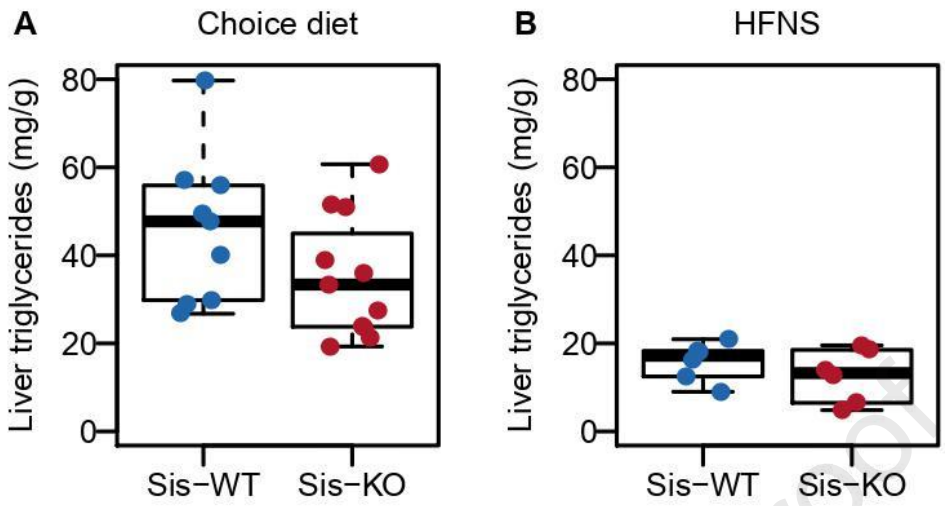
Supplementary Figure 2. QQ-plots of recessive association analyses of **A)** BMI (lambda, 1.01) and **B)** Triglycerides (lambda, 0.97) performed using a linear mixed model.



Supplementary Figure 3. Selection scan of the region on chromosome 3 encompassing *SI*. A) Estimated extended haplotype homozygosity (EHH)[1] decay from the *SI* variant. The blue and the red lines show the decay of EHH for haplotypes carrying the derived and the ancestral allele at the *SI* variant site, respectively. **B)** Estimated normalized integrated haplotype score (iHS)[2] values on chromosome 3. The vertical red line shows the normalized iHS for the *SI* variant (-0.199). As can be seen, the variant is not an outlier in terms of iHS, and thus does not show a signature of recent positive selection.



Supplementary Figure 4. Feeding pattern of Sis-KO and Sis-WT mice on choice diet. Average intake per mouse (Sis-KO, n=13; Sis-WT, n=9), and per cage of high-fat 12.6% sucrose (HFS) diet, low-fat 17% sucrose (17S) diet, and low-fat no-sucrose (NS) diet.



Supplementary Figure 5. Liver triglycerides in Sis-KO and Sis-WT mice. Mean liver triglyceride levels for mice with either ad libitum access to A) a choice diet of high-fat 12.6% sucrose (HFS), low-fat 17% sucrose (17S), and low-fat no-sucrose (NS) diet (Sis-WT, n=9; Sis-KO, n=11), or B) a high-fat no-sucrose (HFNS) diet (Sis-WT, n=6; Sis-KO, n=6).

Supplementary Tables

Supplementary Table 1. Basic clinical characteristics for cohort I and cohort II

	Cohort I	**Cohort II**
N (men/women)	4,639 (1,959/2,680)	1,526 (579/947)
SI c.273_274delAG		
Frequency (95% CI)	14.2% (13.5-15.1)	14.1% (12.8-15.3)
Genotype distribution (WT/HE/HO)	3,418/1,122/99	1,131/361/34
Age (years)	43 (33-54)	32 (25-43)
Height (cm)	162 (156-169)	164 (158-171)
Weight (kg)	67 (58-78)	70 (61-81)
BMI (kg/m^2)	25.3 (22.4-29.0)	25.7 (22.7-29.6)

Data are median (interquartile range) for quantitative traits.

Supplementary Table 2. Classification codes for cardiovascular disease events

	ICD8	ICD10	ICPC2
Ischemic heart disease	41199, 41009, 41099, 41109, 41209, 41299, 41409, 41499	DI21, DI22, DI23, DI24, DI25	K75, K76
Cerebrovascular disease	43100, 43101, 43108, 43109, 43190, 43191, 43198, 43199, 43200, 43201, 43202, 43208, 43209, 43290, 43291, 43292, 43293, 43298, 43299, 43309, 43399, 43409, 43499, 43599, 43509, 43601, 43609, 43690, 43699, 43700, 43701, 43708, 43709, 43790, 43791, 43798, 43799	DI61, DI62, DI63, DI64, DI65, DI66, DI672, DI678, DI693, DI694, DG458, DG459	K89, K90, K91
Peripheral artery disease	44009, 44019, 44020, 44021, 44028, 44029, 44099, 44039, 44408, 44409, 44419, 44420, 44421, 44428, 44442, 44429, 44439, 44440, 44441, 44499, 44490, 44449, 44448, 44443, 44444	DI70, DI739, DI739A, DI739C, DI740, DI740B, DI740D, DI741, DI742, DI743, DI744, DI745, DI748, DI749	K92, K99
Heart failure	78249, 42719, 42710, 42711, 42899, 42709	DI50	K77

International Statistical Classification of Diseases and Related Health Problems- (ICD-) 8, ICD10, and International Classification of Primary Care 2 (ICPC-2) codes were used to classify events.

Supplementary Table 3. Classification codes for operational procedures related to cardiovascular health

	The Classification of Operations and Treatments	**NOMESCO Classification of Surgical Procedures**
Coronary operations	30359, 30354, 30350, 30245, 30241, 30240, 30200, 30199, 30189, 30179, 30169, 30159, 30149, 30139, 30129, 30120, 30119, 30109, 30099, 30089, 30079, 30069, 30059, 30049, 30039, 30029, 30009, 30019	KFNA, KFNB, KFNC, KFND, KFNE, KFNF, KFNG, KFNH, KFNW96, KFNW98

The Classification of Operations and Treatments and The Nordic Medico-Statistical Committee (NOMESCO) Classification of Surgical Procedures were used to classify events of coronary operations.

Supplementary Table 4. Analyses of association between the *SI* c.273_274delAG variant and metabolic phenotypes applying a full model in cohort I

		Heterozygous effect		Homozygous effect	
Trait	***N***	**β_{SD} (SE)**	***P* value**	**β_{SD} (SE)**	***P* value**
Body composition					
BMI (kg/m^2)	4591	-0.06 (0.04)	0.087	-0.44 (0.10)	$1.2x10^{-5}$
Weight (kg)	4596	-0.05 (0.03)	0.169	-0.36 (0.10)	$2.5x10^{-4}$
Waist (cm)	4559	-0.05 (0.03)	0.127	-0.38 (0.10)	$8.8x10^{-5}$
Hip (cm)	4558	-0.06 (0.04)	0.096	-0.39 (0.10)	$9.1x10^{-5}$
Waist-hip ratio	4557	-0.02 (0.03)	0.529	-0.23 (0.09)	0.011
Fat percentage (%)	2691	-0.02 (0.04)	0.602	-0.41 (0.11)	$3.4x10^{-4}$
SAT (cm)	2663	-0.04 (0.05)	0.322	-0.63 (0.12)	$3.0x10^{-7}$
VAT (cm)	2674	-0.002 (0.04)	0.956	-0.19 (0.12)	0.100
SAT/VAT ratio	2656	-0.07 (0.05)	0.149	-0.59 (0.12)	$1.4x10^{-6}$
Lean mass (kg)	2680	0.02 (0.04)	0.553	-0.19 (0.10)	0.061
Lipid profile					
Fs-Triglyceride (mmol/L)	4091	-0.06 (0.04)	0.085	-0.52 (0.10)	$7.0x10^{-7}$
Fs-Total cholesterol (mmol/L)	4482	-0.04 (0.03)	0.191	0.16 (0.10)	0.094
Fs-LDL-cholesterol (mmol/L)	3924	-0.07 (0.04)	0.053	0.16 (0.10)	0.108
Fs-HDL-cholesterol (mmol/L)	4617	0.03 (0.03)	0.443	0.21 (0.10)	0.027
Fs-VLDL-cholesterol (mmol/L)	2078	-0.001 (0.05)	0.978	-0.46 (0.13)	$6.0x10^{-4}$
Fs-Remnant cholesterol (mmol/L)	3924	-0.07 (0.04)	0.060	-0.48 (0.11)	$1.2x10^{-5}$
Apolipoprotein A1 (g/l)	1233	-0.01 (0.07)	0.905	0.55 (0.25)	0.029
Apolipoprotein B (g/l)	1233	-0.03 (0.07)	0.680	0.11 (0.25)	0.657

Effect sizes and *P* values, calculated based on quantile transformed trait values, assessed with a full model allowing for separate effects of being heterozygous and homozygous carrier of the *SI* c.273_274delAG variant. Fs, fasting serum; SAT, subcutaneous adipose tissues; VAT, visceral adipose tissues.

Supplementary Table 5. Dietary composition of low-fat diets with 17% sucrose (17S) or no-sucrose (NS).

	D12450K, No Sucrose (NS)		**D12450H, 17% Sucrose (17S)**	
	gm%	kcal%	gm%	kcal%
Protein	19.2	20.0	19.2	20.0
Carbohydrate	67.3	70.0	67.3	70.0
Sucrose	0.0	0.0	16.4	17.0
Fat	4.3	10.0	4.3	10.0
Total		100.0		100.0
Kcal/gram	3.85		3.85	
	gm	**kcal**	**gm**	**Kcal**
Casein, 30 Mesh	200	800	200	800
L-Cystein	3	12	3	12
Corn Starch	550	2200	452.2	1808.8
Maltodextrin 10	150	600	75	300
Sucrose	0	0	172.8	691.2
Cellulose BW200	50	0	50	0
Soybean Oil	25	225	25	225
Lard	20	180	20	180
Mineral Mix S10026	10	0	10	0
DiCalcium Phosphate	13	0	13	0
Calcium Carbonate	5.5	0	5.5	0
Potassium Citrate, 1H20	16.5	0	16.5	0
Vitamin Mix V10001	10	40	10	40
Choline Bitartrate	2	0	2	0
FD&C Yellow Dye #5	0	0	0.04	0
FD&C Red Dye #40	0.025	0	0.01	0
FD&C Blue Dye #1	0.025	0	0	0
Total	1055.05	4057	1055.05	4057

Supplementary Table 6. Dietary composition of high-fat diets with (HFS) or without sucrose (HFNS).

D12331, 58 kcal% Fat with Sucrose (HFS)			**D0806014B, 60 kcal% Fat no Sucrose (HFNS)**		
	gm%	**Kcal %**		**gm%**	**kcal%**
Protein	23.0	16	Protein	26.0	20.0
Carbohydrate	34.5	26	Carbohydrate	26.0	20.0
Maltodextrin	17.0	12.2	Maltodextrin	16.2	12.3
Corn Starch	0.0	0	Corn Starch	8.9	6.8
Sucrose	17.5	12.6	Sucrose	0.0	0.0
Fat	35.8	58.0	Fat	35.0	60.0
Soybean Oil	2.5	4.0	Soybean Oil	3.2	5.5
Coconut Oil, Hydrogenated	33.3	54.0	Coconut Oil, Hydrogenated	0.0	0.0
Lard	0.0	0	Lard	31.7	54.4
Total		100.0	Total		100.0
kcal/gm	5.56		kcal/gm	5.2	
Ingredient	**gm**	**kcal**	**Ingredient**	**gm**	**kcal**
Casein	228	912	Casein, 80 Mesh	200	800
DL-Methionine	2	0	L-Cystine	3	12
Maltodextrin 10	170	680	Maltodextrin 10	125	500
Corn Starch	0	0	Corn Starch	68.8	275
Sucrose	175	700	Sucrose	0	0
Soybean Oil	25	225	Soybean Oil	25	225
Coconut Oil, Hydrogenated	333.5	3001	Lard	245	2205
Mineral Mix S10001	40	0	Mineral Mix S10026	10	0
Sodium Bicarbonate	10.5	0	CiCalcium Phosphate	13	0
Potassium Citrate	4	0	Calcium Carbonate	5.5	0
			Potassium Citrate, 1 H20	16.5	0
Vitamin Mix V10001	10	40	Vitamin Mix V10001	10	40
Choline Bitartrate	2	0	Choline Bitartrate	2	0

FD&C Yellow Dye #5	0	0	FD&C Yellow Dye #5	0	0
FD&C Red Dye #40	0.1	0	FD&C Red Dye #40	0	0
FD&C Blue Dye #1	0	0	FD&C Blue Dye #1	0.5	0
			Cellulose, BW200	50	0
			Inulin (orafti HP)	0	0
Total	1000.1	5558	**Total**	773.85	4057

Supplementary Table 7. Housing arrangement for mice in the HFNS diet experiment.

Female Sis-WT			**Female Sis-KO**		
	Mice/Cage	Age		Mice/Cage	Age
Cage 1	2	9	Cage 1	2	8
Cage 2	1	6			
Mean Age		8			8
Male Sis-WT			**Male Sis-KO**		
	Mice/Cage	Age		Mice/Cage	Age
Cage 1	2	11	Cage 1	2	8
Cage 2	1	9	Cage 2	1	5
			Cage 3	1	11
Mean Age		10.3			8

Mean age at baseline, and number of mice per cage for Sis-KO and Sis-WT mice. All mice were kept in IVCs.

Supplementary Table 8. Frequency of the *SI* c.273_274delAG variant in populations from across the world

Region	*N*	Allele Frequency (%)
Americas		
Canadian Inuit[9]	128	17.2
Greenlandic Inuit[10]	18	38.9
Ancient Dorset[11]	16	0
Ancient Saqqaq[12]	2	0
Americas (SGDP)[13]	27	0
Americas (HGDP)[14]	51	0
Latino (1000 Genomes)[5]	347	0
Latino (gnomAD)[15]	6801	0.015
Siberia		
Central Asia Siberia (SGDP)[13]	27	5.6
Central Asia Siberia[16]	205	4.1
Central Asia Siberia (HGDP)[14]	23	0
Unknown		
Other (gnomAD)[15]	1074	0.047
Rest of the World		
HGDP[14]	754	0
SGDP[13]	246	0
1000 Genomes[5]	2157	0
gnomAD[15]	63689	0

The allele frequency of c.273-274delAG was surveyed in a range of datasets from across the world in the public database gnomAD (https://gnomad.broadinstitute.org). We also downloaded and used SAMtools[17], BGT[18], and VCFtools[19] to interrogate modern whole genome data from Human Genetic Diversity Panel (HGDP; ftp://ftp.1000genomes.ebi.ac.uk:/vol1/ftp/data_collections/HGDP/data/), Simons Genome Diversity Panel (SGDP; https://github.com/lh3/sgdp-fermi), and 1000 Genome Project data (http://hgdownload.cse.ucsc.edu/gbdb/hg19/1000Genomes/phase3/). Finally, we used SAMtools to interrogate whole genome data from several ancient genomes from the Americas and the Arctic[11,12,20].

Supplementary Table 9. Association between *SI* c.273_274delAG and quantitative metabolic traits according to a recessive model in cohort I

Trait	*N*	β (SE)	β_{SD} (SE)	*P* value
Body composition				
BMI (kg/m^2)	4591	-2.0 (0.5)	-0.4 (0.1)	$3.1x10^{-5}$
Weight (kg)	4596	-4.8 (1.4)	-0.3 (0.1)	$5.1x10^{-4}$
Waist (cm)	4559	-4.9 (1.3)	-0.4 (0.1)	$1.8x10^{-4}$
Hip (cm)	4558	-3.3 (0.9)	-0.4 (0.1)	$2.3x10^{-4}$
Waist-hip ratio	4557	-0.02 (0.01)	-0.2 (0.1)	0.012
Fat percentage (%)	2691	-3.3 (1.0)	-0.4 (0.1)	$3.7x10^{-4}$
SAT (cm)	2663	-0.70 (0.17)	-0.60 (0.12)	$5.8x10^{-7}$
VAT (cm)	2674	-0.44 (0.26)	-0.19 (0.12)	0.102
SAT/VAT ratio	2656	-0.08 (0.03)	-0.56 (0.12)	$3.8x10^{-6}$
Lean mass (kg)	2680	-1.30 (0.65)	-0.21 (0.10)	0.039
Lipid profile				
Triglyceride (mmol/L)	4091	-0.27 (0.07)	-0.49 (0.10)	$2.3x10^{-6}$
Total cholesterol (mmol/L)	4482	0.22 (0.11)	0.18 (0.09)	0.048
LDL-cholesterol (mmol/L)	3924	0.21 (0.11)	0.21 (0.10)	0.039
HDL-cholesterol (mmol/L)	4617	0.11 (0.05)	0.19 (0.09)	0.037
VLDL-cholesterol (mmol/L)	2078	-0.13 (0.04)	-0.46 (0.13)	$6.0x10^{-4}$
Remnant cholesterol (mmol/L)	3924	-0.11 (0.03)	-0.44 (0.11)	$4.2x10^{-5}$
Apolipoprotein A1 (g/l)	1233	0.18 (0.08)	0.55 (0.25)	0.027
Apolipoprotein B (gl/l)	1233	0.03 (0.06)	0.12 (0.25)	0.634
Glucose homeostasis				
Fp glucose (mmol/L)	3664	0.003 (0.10)	0.05 (0.01)	0.643
2h-p glucose (mmol/L)	3410	-0.01 (0.27)	0.004 (0.11)	0.972
Fs insulin (pmol/l)	3662	-6.53 (4.46)	-0.19 (0.11)	0.088
2h-s insulin (pmol/l)	3410	3.73 (25.41)	-0.10 (0.11)	0.359
Fs C-peptide (pmol/l)	3662	-89.76 (33.18)	-0.33 (0.10)	0.0014
2h-s C-peptide (pmol/l)	3410	-67.02 (113.67)	-0.10 (0.10)	0.308
HbA_{1c} (%)	4589	-0.03 (0.05)	-0.04 (0.08)	0.598
HOMA-IR	3655	-0.29 (0.15)	-0.16 (0.11)	0.135

	N	**OR (95% CI)**		***P* value**
T2D (cases/controls)	320/2619	1.89 (0.93-3.86)		0.081

Results were obtained with a linear mixed model (GEMMA). Effect sizes are shown as quantile transformed (β_{SD}), and untransformed (β). *P* values were calculated based on the quantile transformed values. Lipids were measured in fasting serum samples. For type 2 diabetes (T2D), the effect estimate was the odds ratio (95% CI). F, fasting; p, plasma; s, serum; SAT, subcutaneous adipose tissues; VAT, visceral adipose tissues.

Supplementary Table 10. Association between *SI* c.273_274delAG and quantitative metabolic traits according to a recessive model in cohort II

Trait	***N***	**β (SE)**	**β_{SD} (SE)**	***P* value**	**$N_{acetate\_adj}$**	**β_{SD} (SE)$_{acetate\_adj}$**	**$P_{acetate\_adj}$**
Body composition							
BMI (kg/m^2)	1521	-2.23 (0.85)	-0.42 (0.17)	0.013	1473	-0.35 (0.20)	0.077
Weight (kg)	1521	-5.38 (2.46)	-0.35 (0.17)	0.034	1473	-0.27 (0.19)	0.160
Lipid profile							
Triglyceride (mmol/L)	1467	-0.33 (0.21)	-0.31 (0.17)	0.075	1420	-0.36 (0.20)	0.080
Total cholesterol (mmol/L)	1472	0.18 (0.19)	0.18 (0.16)	0.257	1425	0.02 (0.183)	0.892
LDL-cholesterol (mmol/L)	1466	0.20 (0.24)	0.15 (0.17)	0.373	1419	0.13 (0.20)	0.510
HDL-cholesterol (mmol/L)	1470	0.10 (0.08)	0.21 (0.17)	0.207	1423	0.06 (0.20)	0.748
Markers of liver health							
Alkaline phosphatase (U/l)	1318	-15.41 (4.20)	-0.78 (0.18)	$9.8x10^{-6}$	1273	-0.49 (0.21)	0.018
Albumin (g/l)	998	0.44 (0.55)	0.15 (0.19)	0.436	956	0.19 (0.22)	0.397
ASAT (U/l)	1316	1.53 (3.29)	0.05 (0.18)	0.772	1271	-0.17 (0.21)	0.415
Bilirubin (µmol/l)	1279	-0.16 (1.83)	0.02 (0.19)	0.924	1235	0.09 (0.21)	0.686

Results were obtained with a linear mixed model (GEMMA). Effect sizes and standard error (SE) for analyses of association between the variant and the different phenotypes are shown as quantile transformed (β_{SD}), and untransformed (β). *P* values were calculated based on the quantile transformed values. Additionally, quantile transformed effect size and SE (β_{SD} (SE)$_{acetate\_adj}$), number of individuals ($N_{acetate\_adj}$), and *P* value adjusted for level of serum acetate ($P_{acetate\_adj}$) are shown. Lipids and markers of liver health were measured in plasma samples. ASAT, aspartate aminotransferase; SAT.

Supplementary Table 11. Association between *SI* c.273_274delAG and measures of metabolic markers and lipoproteins measured by NMR in cohort II according to a recessive model

		***SI* c.273_274delAG genotype**				
Trait	***N***	**HO**	**HE+WT**	**β (SE)**	**β_{SD} (SE)**	***P* value**
Mean particle diameter (nm)						
VLDL	1478	36.40 (1.54)	37.03 (1.68)	-0.574 (0.289)	-0.365 (0.173)	0.035
LDL	1515	23.54 (0.11)	23.61 (0.16)	-0.082 (0.027)	-0.559 (0.165)	$7.1x10^{-4}$
HDL	1478	10.13 (0.28)	9.97 (0.28)	0.149 (0.046)	0.551 (0.170)	0.0012
Cholesterol concentration (mmol/L)						
Total	1515	4.74 (1.11)	4.40 (1.09)	0.227 (0.169)	0.211 (0.155)	0.173
Esterified	1509	3.38 (0.80)	3.17 (0.78)	0.136 (0.122)	0.173 (0.156)	0.268
Free	1509	1.35 (0.32)	1.24 (0.32)	0.091 (0.049)	0.295 (0.153)	0.054
VLDL	1526	0.69 (0.30)	0.74 (0.31)	-0.081 (0.050)	-0.286 (0.163)	0.080
Remnant	1515	1.42 (0.50)	1.42 (0.48)	-0.053 (0.076)	-0.114 (0.159)	0.470
LDL	1515	1.74 (0.61)	1.56 (0.60)	0.138 (0.096)	0.251 (0.160)	0.117
HDL	1467	1.58 (0.39)	1.41 (0.38)	0.155 (0.061)	0.413 (0.169)	0.014
HDL_2	1467	1.07 (0.36)	0.91 (0.35)	0.149 (0.056)	0.432 (0.170)	0.011
HDL_3	1509	0.51 (0.04)	0.50 (0.04)	0.005 (0.007)	0.132 (0.161)	0.413
Glycerides and phospholipids concentration (mmol/L)						
Triglycerides	1526	1.18 (0.58)	1.40 (0.73)	-0.225 (0.123)	-0.359 (0.169)	0.034
VLDL triglycerides	1478	0.76 (0.48)	0.96 (0.64)	-0.195 (0.109)	-0.369 (0.171)	0.031
LDL triglycerides	1526	0.17 (0.05)	0.18 (0.07)	-0.008 (0.011)	-0.037 (0.155)	0.810
HDL triglycerides	1526	0.15 (0.05)	0.16 (0.05)	-0.020 (0.009)	-0.402 (0.168)	0.017
Diacylglycerol	1516	0.01 (0.02)	0.02 (0.02)	-0.004 (0.003)	-0.266 (0.161)	0.099
Phosphoglycerides	1509	1.90 (0.37)	1.86 (0.41)	-0.008 (0.066)	0.0039 (0.160)	0.982
Phosphatidylcholine	1509	1.90 (0.35)	1.86 (0.40)	-0.003 (0.062)	0.013 (0.160)	0.937
Sphingomyelins	1509	0.46 (0.10)	0.43 (0.09)	0.029 (0.015)	0.310 (0.158)	0.050
Cholines	1509	2.34 (0.41)	2.25 (0.44)	0.037 (0.068)	0.102 (0.158)	0.519
Apolipoproteins (g/l)						
Apolipoprotein A-I	1467	1.63 (0.23)	1.54 (0.22)	0.073 (0.034)	0.322 (0.162)	0.047
Apolipoprotein B	1515	0.89 (0.24)	0.88 (0.24)	-0.010 (0.038)	-0.026 (0.159)	0.866

Amino acids (mmol/L)						
Alanine	1467	0.42 (0.08)	0.44 (0.08)	-0.028 (0.013)	-0.397 (0.174)	0.022
Glutamine	1509	0.47 (0.06)	0.46 (0.08)	0.022 (0.012)	0.289 (0.170)	0.089
Histidine	1505	0.07 (0.01)	0.07 (0.01)	0.002 (0.002)	0.189 (0.174)	0.278
Glycine	1514	0.27 (0.06)	0.29 (0.06)	-0.023 (0.001)	-0.534 (0.167)	0.001
Isoleucine	1526	0.06 (0.02)	0.06 (0.02)	-0.002 (0.004)	-0.079 (0.174)	0.650
Leucine	1526	0.09 (0.03)	0.09 (0.03)	-0.0002 (0.005)	-0.008 (0.174)	0.964
Valine	1508	0.18 (0.05)	0.18 (0.05)	0.003 (0.008)	0.024 (0.173)	0.890
Phenylalanine	1508	0.08 (0.01)	0.08 (0.01)	-0.006 (0.002)	-0.423 (0.165)	0.010
Tyrosine	1505	0.06 (0.02)	0.06 (0.02)	0.004 (0.003)	0.159 (0.174)	0.361
Glycolysis Related Metabolites (mmol/L)						
Glucose	1515	4.14 (0.68)	4.11 (0.99)	0.039 (0.167)	0.133 (0.170)	0.433
Lactate	1508	1.81 (0.55)	2.17 (0.69)	-0.398 (0.116)	-0.578 (0.169)	$6.4x10^{-4}$
Citrate	1508	0.10 (0.02)	0.10 (0.01)	0.002 (0.002)	0.155 (0.173)	0.370
Glycerol	1470	0.06 (0.03)	0.07 (0.03)	-0.010 (0.005)	-0.445 (0.172)	0.010
Pyruvate	1471	0.09 (0.03)	0.11 (0.04)	-0.016 (0.006)	-0.462 (0.173)	0.008
Ketone bodies (mmol/L)						
Acetoacetate	1478	0.05 (0.02)	0.05 (0.06)	0.018 (0.010)	0.439 (0.171)	0.011
Beta hydroxybutyrate	1478	0.12 (0.03)	0.12 (0.08)	0.010 (0.014)	0.262 (0.171)	0.124
Inflammation						
Glycoprotein acetyls	1515	1.26 (0.16)	1.38 (0.25)	-0.130 (0.043)	-0.554 (0.171)	0.001
Miscellaneous (mmol/L)						
Acetate	1478	0.10 (0.07)	0.05 (0.01)	0.056 (0.002)	1.784 (0.165)	$2.1x10^{-26}$
Creatinine	1498	0.05 (0.01)	0.05 (0.01)	0.0007 (0.002)	0.027 (0.170)	0.876
Albumin	1515	0.09 (0.01)	0.09 (0.01)	-0.0003 (0.001)	-0.023 (0.172)	0.896

Supplementary Table 12. Association between *SI* c.273_274delAG and measures of lipoprotein particle concentration and composition measured by NMR in cohort II according to a recessive model

		***SI* c.273_274delAG genotype**				
Trait	***N***	**HO**	**HE+WT**	**β (SE)**	**β_{SD} (SE)**	***P* value**
Particle concentration (mol/l)						
XXL-VLDL	1526	$1.8x10^{-10}$ ($1.7x10^{-10}$)	$2.2x10^{-10}$ ($2.4x10^{-10}$)	$-2.7x10^{-11}$ ($4.1x10^{-11}$)	-0.137 (0.169)	0.419
XL-VLDL	1526	$7.3x10^{-10}$ ($7.7x10^{-10}$)	$9.9x10^{-10}$ ($1.1x10^{-9}$)	$-2.4x10^{-10}$ ($1.9x10^{-10}$)	-0.256 (0.168)	0.129
L-VLDL	1526	$4.0x10^{-9}$ ($3.8x10^{-9}$)	$5.6x10^{-9}$ ($5.3x10^{-9}$)	$-1.5x10^{-9}$ ($9.0x10^{-10}$)	-0.354 (0.169)	0.037
M-VLDL	1478	$1.5x10^{-8}$ ($9.3x10^{-9}$)	$1.9x10^{-8}$ ($1.2x10^{-8}$)	$-3.7x10^{-9}$ ($2.0x10^{-9}$)	-0.352 (0.171)	0.040
S-VLDL	1478	$2.7x10^{-8}$ ($1.1x10^{-8}$)	$3.1x10^{-8}$ ($1.2x10^{-8}$)	$-4.1x10^{-9}$ ($2.0x10^{-9}$)	-0.364 (0.166)	0.029
XS-VLDL	1515	$3.9x10^{-8}$ ($1.2x10^{-8}$)	$3.9x10^{-8}$ ($1.2x10^{-8}$)	$-1.3x10^{-9}$ ($1.9x10^{-9}$)	-0.116 (0.158)	0.458
IDL	1515	$1.1x10^{-7}$ ($3.3x10^{-8}$)	$1.1x10^{-7}$ ($3.2x10^{-8}$)	$2.8x10^{-9}$ ($5.1x10^{-9}$)	0.095 (0.158)	0.550
L-LDL	1515	$1.9x10^{-7}$ ($5.7x10^{-8}$)	$1.7x10^{-7}$ ($5.6x10^{-8}$)	$8.6x10^{-9}$ ($9.0x10^{-9}$)	0.170 (0.158)	0.283
M-LDL	1526	$1.5x10^{-7}$ ($4.9x10^{-8}$)	$1.4x10^{-7}$ ($5.0x10^{-8}$)	$9.8x10^{-9}$ ($8.0x10^{-9}$)	0.221 (0.159)	0.164
S-LDL	1526	$1.8x10^{-7}$ ($5.3x10^{-8}$)	$1.6x10^{-7}$ ($5.6x10^{-8}$)	$1.3x10^{-8}$ ($8.9x10^{-9}$)	0.269 (0.158)	0.088
XL-HDL	1478	$6.2x10^{-7}$ ($2.2x10^{-7}$)	$4.7x10^{-7}$ ($2.4x10^{-7}$)	$1.4x10^{-7}$ ($3.8x10^{-8}$)	0.621 (0.167)	$2.1x10^{-4}$
L-HDL	1478	$1.4x10^{-6}$ ($6.2x10^{-7}$)	$1.1x10^{-6}$ ($6.3x10^{-7}$)	$2.3x10^{-7}$ ($1.0x10^{-7}$)	0.369 (0.167)	0.026
M-HDL	1515	$1.9x10^{-6}$ ($4.0x10^{-7}$)	$2.0x10^{-6}$ ($4.4x10^{-7}$)	$-8.0x10^{-8}$ ($7.3x10^{-8}$)	-0.209 (0.173)	0.228
S-HDL	1515	$4.7x10^{-6}$ ($5.0x10^{-7}$)	$4.8x10^{-6}$ ($7.4x10^{-7}$)	$-2.0x10^{-7}$ ($1.3x10^{-7}$)	-0.350 (0.171)	0.041
Triglyceride concentration (mmol/L)						
XXL-VLDL	1526	0.028 (0.026)	0.033 (0.037)	-0.0038 (0.0062)	-0.119 (0.169)	0.482
XL-VLDL	1526	0.044 (0.047)	0.061 (0.069)	-0.015 (0.012)	-0.261 (0.168)	0.122
L-VLDL	1526	0.136 (0.128)	0.191 (0.183)	-0.051 (0.031)	-0.348 (0.170)	0.041
M-VLDL	1478	0.252 (0.166)	0.327 (0.225)	-0.069 (0.038)	-0.358 (0.172)	0.037
S-VLDL	1478	0.201 (0.099)	0.241 (0.113)	-0.041 (0.019)	-0.403 (0.169)	0.018
XS-VLDL	1526	0.094 (0.035)	0.105 (0.039)	-0.013 (0.0064)	-0.365 (0.165)	0.027
IDL	1526	0.103 (0.031)	0.108 (0.039)	-0.0083 (0.0063)	-0.180 (0.159)	0.256
L-LDL	1526	0.096 (0.029)	0.097 (0.038)	0.0044 (0.0060)	-0.042 (0.156)	0.785
M-LDL	1526	0.048 (0.015)	0.047 (0.020)	-0.0012 (0.0032)	0.027 (0.156)	0.864
S-LDL	1526	0.031 (0.011)	0.032 (0.014)	-0.0016 (0.0022)	-0.065 (0.158)	0.680

XL-HDL	1526	0.024 (0.010)	0.022 (0.012)	0.0010 (0.0021)	0.139 (0.166)	0.407
L-HDL	1467	0.038 (0.018)	0.035 (0.020)	$-3.5x10^{-5}$ (0.0033)	0.003 (0.167)	0.988
M-HDL	1515	0.038 (0.016)	0.046 (0.016)	-0.0088 (0.0028)	-0.570 (0.172)	$9.7x10^{-4}$
S-HDL	1526	0.047 (0.016)	0.054 (0.018)	-0.0083 (0.0031)	-0.534 (0.171)	0.0018
Free cholesterol concentration (mmol/L)						
XXL-VLDL	1526	0.0027 (0.0027)	0.0034 (0.0039)	-0.00066 (0.00066)	-0.239 (0.169)	0.158
XL-VLDL	1526	0.0075 (0.0077)	0.0096 (0.0109)	-0.0020 (0.0018)	-0.211 (0.167)	0.207
L-VLDL	1526	0.026 (0.027)	0.036 (0.038)	-0.0010 (0.0064)	-0.350 (0.169)	0.039
M-VLDL	1478	0.056 (0.039)	0.072 (0.050)	-0.0162 (0.0085)	-0.363 (0.170)	0.033
S-VLDL	1526	0.076 (0.031)	0.083 (0.032)	-0.0095 (0.0052)	-0.305 (0.164)	0.062
XS-VLDL	1509	0.078 (0.023)	0.074 (0.026)	0.00097 (0.0042)	0.082 (0.161)	0.614
IDL	1515	0.208 (0.065)	0.187 (0.065)	0.015 (0.010)	0.250 (0.161)	0.122
L-LDL	1515	0.257 (0.074)	0.234 (0.073)	0.018 (0.012)	0.261 (0.161)	0.105
M-LDL	1515	0.149 (0.036)	0.137 (0.037)	0.0090 (0.0059)	0.263 (0.158)	0.097
S-LDL	1515	0.091 (0.021)	0.083 (0.023)	0.0061 (0.0037)	0.292 (0.159)	0.066
XL-HDL	1478	0.080 (0.031)	0.058 (0.033)	0.021 (0.0054)	0.633 (0.167)	$1.5x10^{-4}$
L-HDL	1478	0.094 (0.051)	0.070 (0.049)	0.023 (0.0081)	0.453 (0.167)	0.007
M-HDL	1515	0.074 (0.021)	0.074 (0.023)	-0.0018 (0.0038)	-0.084 (0.173)	0.628
S-HDL	1515	0.109 (0.014)	0.114 (0.019)	-0.0054 (0.0033)	-0.369 (0.172)	0.031
Cholesterol esters concentration (mmol/L)						
XXL-VLDL	1526	0.0034 (0.0031)	0.0041 (0.0041)	-0.00077 (0.00069)	-0.197 (0.167)	0.239
XL-VLDL	1478	0.0079 (0.0081)	0.0107 (0.0113)	-0.0027 (0.0019)	-0.253 (0.168)	0.134
L-VLDL	1478	0.0289 (0.0272)	0.0395 (0.0334)	-0.011 (0.0057)	-0.379 (0.169)	0.026
M-VLDL	1526	0.0835 (0.0480)	0.0960 (0.0523)	-0.015 (0.0088)	-0.316 (0.168)	0.061
S-VLDL	1509	0.1377 (0.0587)	0.1415 (0.0553)	-0.0097 (0.0089)	-0.191 (0.161)	0.234
XS-VLDL	1509	0.1782 (0.0606)	0.1739 (0.0531)	-0.0024 (0.0085)	-0.058 (0.159)	0.712
IDL	1515	0.5266 (0.1744)	0.4919 (0.1567)	0.019 (0.025)	0.118 (0.159)	0.459
L-LDL	1515	0.6496 (0.2343)	0.5911 (0.2230)	0.042 (0.036)	0.199 (0.159)	0.210
M-LDL	1515	0.3668 (0.1536)	0.3225 (0.1541)	0.036 (0.025)	0.264 (0.162)	0.102

S-LDL	1526	0.2258 (0.0917)	0.1935 (0.0937)	0.028 (0.015)	0.332 (0.162)	0.040
XL-HDL	1526	0.2351 (0.0722)	0.1810 (0.0847)	0.049 (0.014)	0.604 (0.164)	$2.4x10^{-4}$
L-HDL	1478	0.3363 (0.1606)	0.2645 (0.1576)	0.066 (0.026)	0.424 (0.167)	0.011
M-HDL	1515	0.3261 (0.0805)	0.3306 (0.0806)	-0.0071 (0.014)	-0.105 (0.174)	0.549
S-HDL	1515	0.3213 (0.0758)	0.3103 (0.0921)	0.0082 (0.016)	0.073 (0.172)	0.668
Total cholesterol concentration (mmol/L)						
XXL-VLDL	1526	0.0061 (0.0058)	0.0075 (0.0079)	-0.0014 (0.0013)	-0.216 (0.168)	0.200
XL-VLDL	1526	0.015 (0.016)	0.020 (0.022)	-0.0046 (0.0037)	-0.227 (0.167)	0.176
L-VLDL	1478	0.054 (0.054)	0.076 (0.071)	-0.021 (0.012)	-0.370 (0.169)	0.029
M-VLDL	1478	0.140 (0.086)	0.169 (0.101)	-0.032 (0.017)	-0.338 (0.169)	0.046
S-VLDL	1520	0.214 (0.087)	0.224 (0.084)	-0.019 (0.014)	-0.230 (0.161)	0.153
XS-VLDL	1509	0.256 (0.083)	0.248 (0.076)	-0.0016 (0.012)	-0.029 (0.159)	0.852
IDL	1515	0.734 (0.238)	0.679 (0.218)	0.034 (0.035)	0.151 (0.159)	0.345
L-LDL	1515	0.907 (0.308)	0.825 (0.294)	0.060 (0.047)	0.222 (0.159)	0.165
M-LDL	1515	0.516 (0.189)	0.460 (0.190)	0.045 (0.031)	0.264 (0.161)	0.101
S-LDL	1526	0.317 (0.113)	0.277 (0.116)	0.034 (0.019)	0.326 (0.161)	0.043
XL-HDL	1526	0.315 (0.101)	0.239 (0.117)	0.069 (0.019)	0.608 (0.165)	$2.3x10^{-4}$
L-HDL	1478	0.431 (0.212)	0.335 (0.206)	0.089 (0.034)	0.430 (0.167)	0.001
M-HDL	1515	0.400 (0.101)	0.405 (0.102)	-0.0088 (0.017)	-0.096 (0.173)	0.584
S-HDL	1515	0.431 (0.079)	0.424 (0.100)	0.0030 (0.017)	0.0080 (0.171)	0.961
Phospholipids concentration (mmol/L)						
XXL-VLDL	1526	00046 (0.0046)	0.0055 (0.0065)	-0.00072 (0.0011)	-0.138 (0.169)	0.413
XL-VLDL	1526	0.013 (0.013)	0.017 (0.019)	-0.0037 (0.0033)	-0.223 (0.167)	0.183
L-VLDL	1526	0.043 (0.041)	0.059 (0.056)	-0.016 (0.0095)	-0.358 (0.169)	0.034
M-VLDL	1478	0.101 (0.061)	0.125 (0.077)	-0.024 (0.013)	-0.357 (0.170)	0.037
S-VLDL	1478	0.126 (0.047)	0.139 (0.049)	-0.016 (0.0081)	-0.340 (0.166)	0.040
XS-VLDL	1515	0.144 (0.045)	0.139 (0.048)	0.0020 (0.0077)	0.076 (0.161)	0.640
IDL	1515	0.308 (0.086)	0.288 (0.083)	0.011 (0.013)	0.147 (0.160)	0.359
L-LDL	1515	0.332 (0.087)	0.312 (0.083)	0.0129 (0.013)	0.163 (0.159)	0.305

M-LDL	1515	0.206 (0.050)	0.196 (0.051)	0.0066 (0.0081)	0.150 (0.158)	0.342
S-LDL	1515	0.150 (0.031)	0.141 (0.035)	0.0063 (0.0055)	0.207 (0.156)	0.184
XL-HDL	1478	0.291 (0.121)	0.212 (0.127)	0.074 (0.021)	0.596 (0.169)	$4.4x10^{-4}$
L-HDL	1467	0.413 (0.174)	0.344 (0.180)	0.061 (0.029)	0.336 (0.167)	0.044
M-HDL	1515	0.385 (0.075)	0.392 (0.088)	-0.013 (0.014)	-0.171 (0.172)	0.320
S-HDL	1515	0.558 (0.069)	0.591 (0.096)	-0.037 (0.017)	-0.503 (0.173)	0.004
Total lipids concentration (mmol/L)						
XXL-VLDL	1526	0.039 (0.036)	0.046 (0.051)	-0.0059 (0.0087)	-0.143 (0.169)	0.397
XL-VLDL	1526	0.072 (0.076)	0.097 (0.111)	-0.023 (0.019)	-0.252 (0.168)	0.135
L-VLDL	1526	0.233 (0.222)	0.324 (0.309)	-0.087 (0.052)	-0.356 (0.169)	0.036
M-VLDL	1478	0.493 (0.309)	0.621 (0.398)	-0.125 (0.067)	-0.352 (0.171)	0.040
S-VLDL	1478	0.541 (0.221)	0.605 (0.232)	-0.077 (0.038)	-0.342 (0.166)	0.039
XS-VLDL	1515	0.494 (0.150)	0.491 (0.149)	-0.013 (0.024)	-0.085 (0.158)	0.587
IDL	1515	1.146 (0.343)	1.075 (0.327)	0.037 (0.052)	0.113 (0.158)	0.476
L-LDL	1515	1.337 (0.415)	1.234 (0.403)	0.069 (0.065)	0.191 (0.158)	0.227
M-LDL	1526	0.770 (0.247)	0.701 (0.253)	0.052 (0.041)	0.230 (0.159)	0.149
S-LDL	1526	0.498 (0.149)	0.449 (0.157)	0.039 (0.025)	0.285 (0.158)	0.072
XL-HDL	1478	0.630 (0.220)	0.471 (0.240)	0.146 (0.039)	0.628 (0.167)	$1.7x10^{-4}$
L-HDL	1478	0.882 (0.397)	0.713 (0.400)	0.152 (0.065)	0.381 (0.167)	0.022
M-HDL	1515	0.822 (0.175)	0.841 (0.192)	-0.031 (0.032)	-0.188 (0.173)	0.278
S-HDL	1515	1.035 (0.114)	1.070 (0.166)	-0.041 (0.028)	-0.317 (0.171)	0.064

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.

Supplementary Table 13. Association between *SI* c.273_274delAG and levels of fatty acids in serum according to a recessive model in cohort II

		***SI* c.273_274delAG genotype**				
Trait	***N***	**HO**	**HE+WT**	**β (SE)**	**β_{SD} (SE)**	***P* value**
Docosahexaenoic acid (22:6 ω-3)	1512	1.318 (0.472)	1.404 (0.614)	-0.112 (0.095)	-0.162 (0.156)	0.299
Total ω-3	1512	4.057 (1.705)	4.218 (1.843)	-0.266 (0.291)	-0.147 (0.157)	0.345
Linoleic acid (*cis*-*cis*-18:2 ω-6)	1512	28.427 (3.913)	25.422 (3.651)	3.275 (0.606)	0.956 (0.163)	$5.8x10^{-9}$
Total ω-6	1512	32.591 (3.739)	29.893 (3.698)	2.967 (0.621)	0.883 (0.166)	$1.2x10^{-7}$
Total SFAs	1518	36.391 (1.833)	36.036 (2.212)	0.438 (0.378)	0.225 (0.171)	0.188
Conjugated linoleic acid	1512	0.122 (0.117)	0.115 (0.100)	0.011 (0.017)	0.108 (0.159)	0.499
MUFA	1518	26.970 (3.497)	29.868 (4.009)	-3.157 (0.677)	-0.822 (0.169)	$1.2x10^{-6}$
PUFA	1518	36.647 (3.753)	34.109 (4.004)	2.674 (0.692)	0.704 (0.172)	$4.7x10^{-5}$
Fatty acid length	1470	17.468 (0.292)	17.648 (0.326)	-0.174 (0.056)	-0.536 (0.172)	0.0018
Total fatty acids	1507	11.394 (2.485)	11.445 (2.827)	-0.268 (0.467)	-0.053 (0.161)	0.739
Estimated degree of unsaturation	1512	1.149 (0.079)	1.135 (0.095)	0.017 (0.160)	0.192 (0.169)	0.254

Supplementary Table 14. Association between *SI* c.273_274delAG and levels of fatty acids in erythrocyte membranes according to a recessive model in cohort I

Trait	*N*	β (SE)	β_{SD} (SE)	*P* value
α-Linolenic acid (18:3 ω-3)	2607	-0.015 (0.015)	-0.115 (0.061)	0.059
Stearidonic acid (18:4 ω-3)	2607	0.011 (0.016)	0.088 (0.056)	0.115
Eicosatetraenoic acid (20:4 ω-3)	2607	-0.014 (0.014)	-0.078 (0.050)	0.118
Eicosapentaenoic acid (20:5 ω-3)	2607	0.386 (0.167)	0.136 (0.096)	0.158
Tetrahomo α-Linolenic acid (22:3 ω-3)	2607	-0.008 (0.020)	-0.015 (0.055)	0.789
Docosapentaenoic acid (22:5 ω-3)	2607	0.058 (0.081)	0.087 (0.119)	0.462
Docosahexaenoic acid (22:6 ω-3)	2607	-0.097 (0.250)	-0.037 (0.115)	0.749
Total ω-3	2607	0.038 (0.468)	0.008 (0.111)	0.946
Linoleic acid (*cis*-*cis*-18:2 ω-6)	2607	0.885 (0.233)	0.371 (0.102)	$2.6x10^{-4}$
γ-linolenic acid (18:3 ω-6)	2607	-0.008 (0.005)	-0.048 (0.030)	0.113
Eicosadienoic acid (20:2 ω-6)	2607	0.006 (0.012)	-0.007 (0.050)	0.886
Dihomo-γ-linolenic acid (20:3 ω-6)	2607	-0.043 (0.057)	-0.087 (0.102)	0.391
Arachidonic acid (20:4 ω-6)	2607	0.336 (0.283)	0.127 (0.110)	0.249
Adrenic acid (22:4 ω-6)	2607	-0.010 (0.048)	0.053 (0.101)	0.600
Osbond acid (22:5 ω-6)	2607	-0.009 (0.026)	-0.011 (0.059)	0.851
Total ω-6	2607	1.161 (0.530)	0.253 (0.107)	0.018
Palmitoleic acid (16:1 ω-7)	2607	-0.094 (0.045)	-0.211 (0.116)	0.067
cis-vaccenic acid (18:1 ω-7)	2607	-0.050 (0.028)	-0.226 (0.110)	0.039
Oleic acid (18:1 ω-9)	2607	-0.809 (0.227)	-0.450 (0.125)	$3.2x10^{-4}$
11-eicosenoic acid (20:1 ω-9)	2607	0.016 (0.024)	-0.030 (0.101)	0.769
Erucic acid (22:1 ω-9)	2607	-0.011 (0.011)	-0.052 (0.046)	0.264
Nervonic acid (24:1 ω-9)	2607	-0.130 (0.156)	-0.093 (0.111)	0.399
Myristic acid (14:0 SFA)	2607	-0.022 (0.019)	-0.097 (0.122)	0.428
Palmitic acid (16:0 SFA)	2607	-0.213 (0.307)	-0.026 (0.121)	0.833
Stearic acid (18:0 SFA)	2607	0.035 (0.142)	0.065 (0.118)	0.582
Arachidic acid (20:0 SFA)	2607	-0.017 (0.017)	-0.131 (0.121)	0.281
Behenic acid (22:0 SFA)	2607	-0.109 (0.052)	-0.239 (0.111)	0.031
Lignoceric acid (24:0 SFA)	2607	-0.079 (0.106)	-0.062 (0.110)	0.575

Levels of fatty acids are reported as the relative level compared to the total amount of fatty acids in each sample.

Supplementary Table 15. Association between *SI* c.273_274delAG and cardiovascular disease events according to a recessive model in cohort I

Trait	***N* (with/without condition)**	**HR (95 % CI)**	***P* value**
Ischemic heart disease	238/4313	0.54 (0.17-1.70)	0.291
Cerebrovascular disease	362/4189	0.99 (0.49-2.03)	0.989
Peripheral artery disease	44/4507	$1.1x10^{-7}$ (0-inf)	0.996
Heart failure	160/4391	0.62 (0.15-2.55)	0.512
Coronary operations	88/4463	$8.7x10^{-8}$ (0-inf)	0.994
Any event	628/3923	0.92 (0.53-1.61)	0.771

Effect sizes were estimated as hazard ratios (HR) with a Cox regression model adjusted for sex, birth year, survey, and the top 10 principal components. The analyses included register data from individuals from cohort I. Inf, infinite.

Supplementary Table 16. Association between *SI* c.273_274delAG and quantitative traits according to a recessive model adjusted for sugar intake in cohort I

Trait	***N***	***P* value**	P_{sugar_adj}
Body composition			
BMI (kg/m^2)	2386	$4.7x10^{-4}$	$1.7x10^{-4}$
Weight (kg)	2386	0.0015	$5.7x10^{-4}$
Waist (cm)	2360	0.0030	0.0014
Hip (cm)	2360	0.0040	0.0022
Waist-hip ratio	2359	0.038	0.026
Fat percentage (%)	2344	0.0018	$8.0x10^{-4}$
SAT (cm)	2323	$3.0x10^{-6}$	$1.1x10^{-6}$
VAT (cm)	2334	0.065	0.047
SAT/VAT ratio	2318	$2.1x10^{-5}$	$1.1x10^{-5}$
Lean mass (kg)	2335	0.0098	0.0039
Lipid profile			
Fs-Triglyceride (mmol/L)	2411	$4.1x10^{-4}$	$7.3x10^{-4}$
Fs-Total cholesterol (mmol/L)	2285	0.195	0.174
Fs-LDL-cholesterol (mmol/L)	2269	0.061	0.052
Fs-HDL-cholesterol (mmol/L)	2411	0.755	0.813
Fs-VLDL-cholesterol (mmol/L)	698	0.012	0.013
Fs-Remnant cholesterol (mmol/L)	2269	0.0090	0.014

P values were estimated based on the score test in GEMMA, with (P_{sugar_adj}) or without (*P* value) daily intake of added sugar (g/day) included as a covariate. The analyses included only individuals with data available on sugar intake, and with a realistic energy intake. Fs, fasting serum; SAT, subcutaneous adipose tissues; VAT, visceral adipose tissues.

Supplementary Table 17. Analyses of weight gain, fat fraction, and lean mass gain in Sis-KO mice over 8 weeks on a choice diet with access to HFS, 17S, and NS diet

	Weight gain (g)		Fat fraction (%)		Lean mass gain (g)	
Week	**β (95%CI)**	***P* value**	**β (95%CI)**	***P* value**	**β (95%CI)**	***P* value**
0			-1.89 (-3.95, 0.17)	0.088		
1	-1.22 (-2.34, -0.10)	0.047	-5.5 (-8.34, -2.65)	0.0012	0.13 (-0.49, 0.74)	0.691
2	-2.05 (-3.48, -0.62)	0.011	-7.94 (-11.2, -4.71)	$1.2x10^{-4}$	-0.06 (-0.69, 0.58)	0.865
3	-2.33 (-3.97, -0.69)	0.012	-8.21 (-12, -4.36)	$5.0x10^{-4}$	0.56 (-0.09, 1.21)	0.108
4	-2.39 (-4.16, -0.63)	0.016	-9.11 (-12.9, -5.33)	$1.5x10^{-4}$	0.56 (-0.12, 1.24)	0.122
5	-2.45 (-4.47, -0.43)	0.028	-9.5 (-14.1, -4.95)	$6.3x10^{-4}$	0.73 (-0.02, 1.47)	0.071
6	-2.92 (-5.22, -0.62)	0.022	-9.67 (-14.5, -4.84)	$9.2x10^{-4}$	0.39 (-0.40, 1.17)	0.348
7	-3.06 (-5.67, -0.45)	0.033	-10.6 (-15.8, -5.41)	$7.6x10^{-4}$	0.70 (-0.15, 1.54)	0.121
8	-3.06 (-5.6, -0.52)	0.029	-10.2 (-15.5, -4.91)	0.0013	0.62 (-0.41, 1.64)	0.252

Supplementary Table 18. Analyses of weight gain, fat fraction, and lean mass gain in Sis-KO mice over 8 weeks on a HFNS diet

	Weight gain (g)		**Fat fraction (%)**		**Lean mass gain (g)**	
Week	**β (95%CI)**	***P* value**	**β (95%CI)**	***P* value**	**β (95%CI)**	***P* value**
0			-2.99 (-5.23, -0.76)	0.028		
1	-1.61 (-2.71, -0.52)	0.018	-6.08 (-11.3, -0.91)	0.047	-0.26 (-0.85, 0.34)	0.419
2	-1.27 (-2.96, 0.43)	0.176	-5.36 (-12.3, 1.61)	0.166	-0.05 (-0.81, 0.71)	0.902
3	-1.57 (-3.35, 0.21)	0.117	-5.98 (-13.1, 1.13)	0.134	0.13 (-0.93, 1.2)	0.815
4	-1.64 (-3.54;0.25)	0.123	-6.31 (-14.1, 1.47)	0.146	0.46 (-0.50, 1.41)	0.373
5	-0.88 (-2.93, 1.18)	0.426	-5.97 (-13.4, 1.48)	0.151	0.68 (-0.46, 1.82)	0.271
6	-0.58 (-2.57, 1.4)	0.580	-5.14 (-11.9, 1.66)	0.173	0.57 (-0.61, 1.75)	0.370
7	0.08 (-2.06, 2.21)	0.944	-4.26 (-11.3, 2.74)	0.264	0.83 (-0.25, 1.92)	0.165
8	-0.65 (-3.53, 2.24)	0.671	-6.72 (-16.3, 2.86)	0.202	0.88 (-0.35, 2.12)	0.195

Supplementary Table 19. Estimated differences between Sis-WT and Sis-KO mice in plasma acetate measured at four time points after sucrose gavage.

Time (hours)	β (95% CI)	*P* value
0	-0.008 (-0.130, 0.113)	0.881
2	0.031 (-0.068, 0.129)	0.504
4	2.563 (1.043, 4.082)	0.0037
6	2.862 (1.637, 4.088)	4.0×10^{-4}

Comparison of plasma acetate levels in Sis-KO and Sis-WT mice at four time points after a sucrose gavage. The β values are the estimated differences, and a positive value indicates that the levels are higher in the Sis-KO mice.

Supplementary Table 20. Plasma glucose following a sucrose gavage (3g/kg)

Time (minutes)	β (95% CI)	*P* value
0	-0.529 (-2.542, 1.485)	0.578
15	-5.671 (-8.114, -3.229)	2.8×10^{-4}
30	-5.443 (-7.503, -3.382)	9.1×10^{-5}
60	-1.714 (-3.432, 0.003)	0.050
120	-1.043 (-2.707, 0.621)	0.197

Comparison of plasma glucose levels in Sis-KO and Sis-WT mice at five time points after a sucrose gavage. The β values are the estimated differences, and a negative value indicates that the levels are lower in the Sis-KO mice.

References

1. Sabeti PC, Reich DE, Higgins JM, et al. Detecting recent positive selection in the human genome from haplotype structure. Nature 2002;419:832–837.
2. **Voight BF**, **Kudaravalli S**, Wen X, et al. A Map of Recent Positive Selection in the Human Genome Hurst L, ed. PLoS Biol 2006;4:e72.
3. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res 2009;19:1655–1664.
4. Moltke I, Albrechtsen A. RelateAdmix: a software tool for estimating relatedness between admixed individuals. Bioinforma Appl 2014;30:1027–1028.
5. Auton A, Abecasis GR, Altshuler DM, et al. A global reference for human genetic variation. Nature 2015;526:68–74.
6. Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. Nat Methods 2013;10:5–6.
7. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. Schork NJ, ed. PLoS Genet 2009;5:e1000529.
8. Maclean CA, Chue Hong NP, Prendergast JG. hapbin: An Efficient Program for Performing Haplotype-Based Scans for Positive Selection in Large Genomic Datasets. Mol Biol Evol 2015;32:3027–9.
9. Marcadier JL, Boland M, Scott CR, et al. Congenital sucrase-isomaltase deficiency: Identification of a common Inuit founder mutation. CMAJ 2015;187:102–107.
10. Pedersen C-ET, Lohmueller KE, Grarup N, et al. The Effect of an Extreme and Prolonged Population Bottleneck on Patterns of Deleterious Variation: Insights From the Greenlandic Inuit. Genetics 2017;205:787–801.
11. Raghavan M, DeGiorgio M, Albrechtsen A, et al. The genetic prehistory of the New World Arctic. Science 2014;345:1255832.
12. **Rasmussen M**, **Li Y**, **Lindgreen S**, et al. Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature 2010;463:757–762.
13. Mallick S, Li H, Lipson M, et al. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. Nature 2016;538:201–206.
14. **Li JZ**, **Absher DM**, Tang H, et al. Worldwide human relationships inferred from genome-wide patterns of variation. Science (80-) 2008;319:1100–1104.
15. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020;581:434–443.
16. Malyarchuk BA, Derenko M V., Denisova GA. The frequency of inactive sucrase-isomaltase variant in indigenous populations of Northeast Asia. Russ J Genet 2017;53:1052–1054.
17. **Li H**, **Handsaker B**, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinforma Appl NOTE 2009;25:2078–2079.
18. Li H. BGT: efficient and flexible genotype query across many samples | Bioinformatics | Oxford Academic. Bioinformatics 2016;32:590–592.
19. **Danecek P**, **Auton A**, Abecasis G, et al. The variant call format and VCFtools. Bioinforma Appl NOTE 2011;27:2156–2158.
20. Posth C, Nakatsuka N, Lazaridis I, et al. Reconstructing the Deep Population History of Central and South America. Cell 2018;175:1185-1197.e22.

Author names in bold designate shared co-first authorship.

What you need to know

Background and context: In Arctic populations the sucrase-isomaltase c.273_274delAG loss-of-function variant causes congenital sucrase-isomaltase deficiency in children, however, the impact of the variant on metabolic health in adults is unknown.

New findings: Among Greenlandic adults, homozygous c.273_274delAG carriers had a markedly healthier metabolic profile than the remaining study population, likely mediated by higher circulating acetate levels and reduced sucrose uptake, but not lower caloric intake.

Limitations: We hypothesize that the healthier metabolic profile observed in homozygous c.273_274delAG carriers was mediated by acetate produced by gut bacteria; however, we lack data to firmly verify this hypothesis.

Impact: Our results suggest that sucrase-isomaltase constitutes a promising drug target for improvement of metabolic health, and in a broader perspective add to the debate about the health effects of sugar consumption.

Lay summary

A sucrase-isomaltase loss-of-function variant was associated with a markedly healthier metabolic profile in Greenlandic adults, suggesting that sucrase-isomaltase constitutes a promising drug target for improvement of metabolic health.