Corresponding Author: Torben Hansen

Date: Nov 14, 2017

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work we publish. This form is published with all life science papers and is intended to promote consistency and transparency in reporting. All life sciences submissions use this form; while some list items might not apply to an individual manuscript, all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

   Describe how sample size was determined.

   Under an additive effect model the sample size of ~3,000 is large enough to have the power (>80%) to detect associations with a single genetic variant with MAF=2% with modest to high effect sizes (0.4 s.d.) comparable to the known genetic variants associated with lipid levels in Europeans.

2. Data exclusions

   Describe any data exclusions.

   In the association analyses of the Greenlandic cohorts, we excluded individuals for whom we did not have quality filtered Metabochip data, because such data was needed to be able to correct for admixture and relatedness.

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   The variant was found through exome sequencing and validated by genotyping and with RNA sequencing. Since the identified variant is extremely rare, if at all present, in other populations we were not able to perform replication of the association signal for this exact variant in other populations. Instead, we performed a burden test for other loss-of-function variants in the same gene in other populations. This test supported our initial finding and conclusion.

4. Randomization

   Describe how samples/organisms/participants were allocated into experimental groups.

   NA

5. Blinding

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   NA

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or the Methods section if additional space is needed).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly.
☒	☐	A statement indicating how many times each experiment was replicated
☐	☒	The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒	The test results (e.g. *p* values) given as exact values whenever possible and with confidence intervals noted
☐	☒	A summary of the descriptive statistics, including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☒	☐	Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

As described in the online methods, we used the publicly available software package GEMMA to perform all the association tests in the Greenlandic cohorts. To perform the burden test, we used the standard implementation of logistic regression in the statistical software R. Software used for RNA seq analysis included: Trimmomatic, FastQC, Salmon v0.8.2 and R.

For all studies, we encourage code deposition in a community repository (e.g. GitHub). Authors must make computer code available to editors and reviewers upon request. The *Nature Methods* guidance for providing algorithms and software for publication may be useful for any submission.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

No restrictions

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

NA

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

NA

b. Describe the method of cell line authentication used.

NA

c. Report whether the cell lines were tested for mycoplasma contamination.

NA

d. If any of the cell lines used in the paper are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

NA

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

NA

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

The Greenlandic samples are from two different cohorts: B99 (N=1,401) and Inuit Health in Transition (IHIT) (N=3,115), which were collected in Greenland as a part of general population health surveys conducted in 1999-2001 and 2005-2010, respectively. Two hundred and ninety-five individuals overlapped between the two cohorts and were assigned to the B99 dataset. Clinical characteristics of the participants with genotype data are shown in Supplementary Table 6.