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.

1. Sample size
   Describe how sample size was determined.
   
   Our sample size is ~300,000 individuals with lipids and exome chip genotypes. This sample size is the largest sample that has been analyzed for plasma lipids against a large set of genetic markers.

2. Data exclusions
   Describe any data exclusions.
   
   Individuals were excluded on a study-specific basis based on quality control of the genotypes. Study-specific exclusion can be found in Table S3.

3. Replication
   Describe whether the experimental findings were reliably reproduced.
   
   We tested the new single variant association findings for replication in an independent set of up to 286,268 participants from three studies – Nord-Trøndelag Health Study, (HUNT; max n = 62,168), Michigan Genomics Initiative (MGI; www.michigangenomics.org; max n = 6,411) and the Million Veteran Program18 (MVP; max n = 218,117). Of the novel primary trait associations, 73/73 associations were directionally consistent (Table S10); two SNPs were not available for replication (rs201148465, rs75862065). Furthermore, we were able to replicate the associations of 66/73 (90%) at an alpha of 0.05.

4. Randomization
   Describe how samples/organisms/participants were allocated into experimental groups.
   
   There was no randomization of the participants as this is an observational study.

5. Blinding
   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.
   
   Blinding was not relevant to the study. It is an observational study.

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).

- 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.

7. Software

Describe the software used to analyze the data in this study. Single-variant association statistics and inter-marker linkage disequilibrium information summarized across 1 megabase sliding windows were generated from each cohort using RAREMETALWORKER or RVTESTS13,14 software. Meta-analyses of genetic associations were performed using the R-package rareMETALS (version 6.0).

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.

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 unique materials were used.

9. Antibodies

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

1. A1CF, mouse polyclonal (ab89050, Abcam)
2. β-Actin, mouse monoclonal (A5316, Sigma)
3. APOB, mouse monoclonal (sc-393636, Santa Cruz).

10. Eukaryotic cell lines

- State the source of each eukaryotic cell line used.
- Describe the method of cell line authentication used.
- Report whether the cell lines were tested for mycoplasma contamination.
- 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.

Huh7 and HepG2 human hepatoma cells

Standard authentication was done.

Standard testing for mycoplasma contamination.

No commonly misidentified cell lines were used.
**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.

Jak2 p.Val617Phe MxCre mice were created and reported previously. C57BL/6J mice were used for the A1CF knock-ins.

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.

Covariate-relevant population characteristics of the included studies are available in Table S1 and S2.