Corresponding Author: Thomas Hoffmann and Neil Risch

Date: 11/16/2017

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.

   We used existing cohort data. No statistical methods were used to predetermine sample sizes.

2. Data exclusions

   Describe any data exclusions.

   The original GERA cohort QC has been described in Kvale et. al.. 2015; briefly, individuals with DQC score<0.82 and CR<97% were excluded, followed by per package removal of (1) CR<90%, (2) variance ratio <31, (3) sex frequency difference>0.15, (4) overall CR<0.60, (5) poor genotype discordance at an array-specific threshold (>208/851 EUR array, >23/61 EAS array, >8/12 AFR array, >26/71 LAT array). Here, we included all RPGEH GERA cohort individuals who had lipid measurements. We further excluded genotyped variants with poor call rates (CR<90%) and analyzed SNPs that had a count of at least 20. After imputation, we excluded poorly imputed variants with r_info^2<0.8, and a count of at least 20. Finally, we excluded GLGC SNPs with <80,000 subjects.

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   Two additional cohorts were used for replication.

4. Randomization

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

   Participants randomly genotyped.

5. Blinding

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

   Investigators blinded.

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.

Shape-it v2.r727, Impute2 v2.3.0, Bolt-LMM v2.1, Metasoft v2.0, ImpG v1.0.1, Eagle v2.3, Minimac3 v2.0.1, HaploReg v4.1, Annovar v2015June17, GTeX v6, Gear v0.7.7, PLINK v1.9, GENCODE v13, R v3.2.3, PC-Relate v1, GCTA v1.24.7.

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).

No antibodies.

10. Eukaryotic cell lines

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

No cell lines used.

b. Describe the method of cell line authentication used.

No cell lines used.

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

No cell lines used.

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.

No cell lines 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.

No animals used.

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 GERA cohort of 94,674 is 80.9% non-Hispanic white, 8.2% Latino, 7.2% East Asian, 3.1% African American, and 0.5% South Asian, was 58.6% women, and had an average age of 55.4 years. All lipid measurements were untreated.