natureresearch

Corresponding Author: ___Sekar Kathiresan___

Date: ___Jun 9, 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.

    > All samples available for analysis (after quality control) were used in our analysis. Sample size was determined based on genetic data available from UK Biobank

2. Data exclusions

    Describe any data exclusions.

    > Non-European ancestry individuals, or those who did not pass quality control

3. Replication

    Describe whether the experimental findings were reliably reproduced.

    > Replication was performed using data from one of 2 sources: 1) CARDIoGRAM exome consortium CAD summary statistics, or 2) CARDIoGRAMplusC4D imputed CAD summary statistics

4. Randomization

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

    > N/A, no randomization performed for GWAS analysis

5. Blinding

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

    > No blinding was performed for our GWAS analysis

    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.

> SNPTEST, R statistical software programs, and Graphpad Prism

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.

> Data for GWAS are publicly available, either from UK Biobank with application, or at: http://www.cardiogramplusc4d.org/

9. Antibodies

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

> Commercial antibodies; the application on human samples have been validated by manufactures

10. Eukaryotic cell lines

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

> Human Aortic Endothelial Cells (HAEC) and Human Coronary Artery Smooth Muscle Cells (HCASMC) were purchased from Lifeline Cell Technology. HL60 cell line was purchased from Sigma-Aldrich. HEK-293 and THP-1 cell line was purchased from ATCC.

b. Describe the method of cell line authentication used.

> Cell line specificity was confirmed with tissue-specific markers: HAEC were von Willebrand Factor positive and smooth muscle a- actin negative, HCASMC were von Willebrand Factor negative and smooth muscle a-actin positive. Extensive characterization of HL60, HEK-293, and THP-1 cells have been performed and provided as Certificate of Analysis by Sigma-Aldrich and ATCC.

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

> The cell types above were confirmed to be mycoplasma negative.

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.

> N/A

▶ 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.

> N/A, no animals in this study

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.

> We used a population based biobank to perform a GWAS for coronary artery disease. Individuals were stratified based on CAD status using EHR data (interview/diagnosis/procedure codes). Relevant descriptive statistics appear in supplementary table 1 (age, gender, lipid medication use). Individuals were genotyped using the UK Biobank array, and imputed using a 1000G/UK10K reference panel. Association analysis was performed with the covariates of age, gender, specific chip related statistics using linear mixed modeling