

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but 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.

Of the total ~500,000 subjects from UK Biobank, we analysed 458,577 subjects which passed QC of the genetic data, were of European ancestry, and met our phenotypic data QC requirements for availability of blood pressure data and covariates.

To maximize sample size in the discovery, we recruited an additional 148,890 samples from 23 new cohorts, in addition to the 150,134 samples already existing from the 54 cohorts within the published ICBP-1000G project, giving a total sample size of 299,024 in the ICBP meta-analysis.

Hence a total discovery sample size of $N=757,601$.

Our combined meta-analysis sample size was $N=1,006,863$ after combining with the data from the replication cohorts ($N=220,520$ from MVP and $N=28,742$ from EGCUT).

2. Data exclusions

Describe any data exclusions.

Within UK Biobank, we excluded samples according to both genetic data quality control (QC) and phenotypic data QC. From genetic data QC, we excluded 968 subjects listed as QC outliers for heterozygosity or missingness within the centrally provided UK Biobank sample QC files, and 378 individuals with sex discordance between the phenotypic and genetically inferred sex. We also restricted to subjects of European ancestry, according to both self-reported ethnicity status and ancestry clustering using PCA data. For phenotypic QC, we excluded any subjects with no BP measurements, missing BMI covariate data, pregnant ($N=372$) and those individuals who had withdrawn consent ($N=36$).

Similar sample QC was performed at study level within each of the ICBP and replication cohorts.

3. Replication

Describe whether the experimental findings were reliably reproduced.

Novel loci identified from our 2-stage approach were robustly replicated using independent replication datasets. Novel loci identified from our 1-stage approach met our criteria for internal replication by showing significant support within each of the UKB and ICBP GWAS datasets separately.

4. Randomization

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

N/A for GWAS

5. Blinding

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

N/A for GWAS
(Note data collection of UK Biobank was done centrally, not performed by us)

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 in 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 clear description of 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.

For the Primary GWAS analysis, BOLT-LMM software v2.3 was used for running an association analysis using linear mixed modelling; then METAL software was used for all meta-analyses with a fixed effects inverse variance weighted meta-analysis approach.
We used R software for any general statistical analyses, for secondary analyses (e.g. variance explained analyses, risk score analyses) and for producing plots in the figures.
We used PLINK software for LD calculations of variants.
For the bioinformatics analyses, specific software was used for each different analysis. Each method and the software used is described in the Online Methods, the Supplementary Methods and also summarised in Supplementary Figure 3. For example, the Variant Effect Predictor (VEP) tool is used for variant annotation; DEPICT software is used for enrichment testing, etc.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* [guidance for providing algorithms and software for publication](#) provides further information on this topic.

► 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.

N/A (not labwork)

9. Antibodies

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

N/A

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 are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

N/A

N/A

N/A

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

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.

Summary descriptives of UKB, ICBP and MVP/EGCUT individuals are provided in Sup Tables 1a, 1b and 1c, respectively, showing: blood pressure measurements as the phenotype; age, sex and BMI values as covariates; hypertension status and the use of BP-lowering medication.