nature research

Corresponding author(s): Declan P O'Regan

Last updated by author(s): May 7, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Cardiovascular magnetic resonance images were segmented in Python 3 using published code which can be found at: https://github.com/baiwenjia/ukbb_cardiac. Bright field images of medaka embryos were processed with NIS-Elements v.4.20.
Data analysis	Data processing and analysis were embedded into different snakemake (v5.1.5) workflows. Genetics analyses used plink v1.9 and v2 as well as flashpca. Data visualisation and processing was done in R, version 3.6.0. Genetic association testing requires Bgenie v1.3. Fractal analysis requires Matlab, MATLAB Compiler Runtime (MCR) 2016b and Ghostscript. Finite element modeling was conducted with Abaqus/Standard, Simulia. the rendering of the medaka inner heart surface was performed with UCSL Chimera 1.11. All analysis code and required configuration files to conduct the analyses can be found at: https://github.com/ImperialCollegeLondon/fractalgenetics/

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The genetic and phenotypic UK Biobank data presented in this work are available to any bona fide researcher upon application to UK Biobank. The research was conducted under access number 40616. The GWAS summary level data used in this study are publicly-available (https://www.ebi.ac.uk/gwas/, http://ldsc.broadinstitute.org/ and http://www.hermesconsortium.org/). Figures 1, 3 and 4 contain raw data.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	24,632 UK Biobank participants, 1,129 healthy volunteers, and 510 patients. We used all samples available from UKBB and no formal power calculation was conducted in advance.
Data exclusions	Participant and genetic marker exclusion criteria were pre-established and included presence of full covariate information (age, sex, height or weight), genotype-derived ancestry estimate overlapping with European ancestry of reference population and CMR images passing defined quality control criteria. 1,604 participants were excluded based on these criteria. Criteria for selection based on ancestry and inclusion criteria of genetic markers for the GWAS are described in detail below. UK Biobank genotypes release version 3 (in GRCh37 coordinates) were used in the genetic association studies, for computing genetic principal components and for the Mendelian randomisation analyses. First, only unrelated or distantly related individuals were selected for further analysis based on the estimated identity of descent (IBD) provided by UK Biobank (via 'ukbgene relkey.enc'). Selection of individuals from families is optimised to retain as many unrelated individuals as possible in the study i.e. in family trios only parents would be considered for analysis. For computation of genetic principal components, genotypes were formated into plink binary format, LD pruned (flag '–indep-pairwise 50kb 1 0.8') and a minor allele frequency (MAF) threshold of 1% applied. The filtered dataset was used with flashpca version 2,74 to compute the first 50 principal components of the UK Biobank genotypes. Population substructures arising due to different ethnic origins of samples were examined by comparing the UK Biobank genotypes to genotypes from the HapMap Phase III study, for four ethnic populations. Download of reference data sets, fusion of UK Biobank and Hapmap data sets and PCA selection was done as in plinkQC. Individuals that clustered with the main cluster of the PC1/PC2 plot, which was also the position of the main cluster of HapMap III individuals of European ancestry were kept for further analyses, as is standard in GWAS analysis to model a well mixed population. The genotypes of the remaining unrelated individuals were filtered for genetic variants that passed a MAF threshold of 0.1% and an imputation INFO threshold of > 0.4. For association analysis this was achieved by providing a variant ID list with variant MAF>0.001to BGENIE. All relevant analysis in UK-Biobank and ancestry in the GitHub repository. After genotyping and phenotyping quality control, the discovery cohort was composed of 18,096 individuals and 14,134,301 genetic variants.
Replication	The UK Biobank released an additional 7,192 CMR images on in December 2018. We applied the same genotyping and phenotyping quality control as described above for the UK Biobank discovery cohort and obtained a replication cohort of 6,536 individuals and 14,069,398 genetic variants. A second replication in 1,129 healthy volunteers was performed from comparable data obtained in the UK Digital Heart Project. To ensure reproducibility of simulation and model results, we provide all analysis code and parameters. Knock-out models to study effect of GWAS associated genes were conducted with a large number of embryos, in two independent experiments (total n=100).
Randomization	Classification into groups was not required as the phenotypes in analyses were quantitative and continuous .
Blinding	Blinding was not required as the phenotypes in analyses were quantitative and continuous .

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Medaka (Oryzias latipes); Unblinded analysis was performed using a sample size sufficient for descriptive analysis. All data were generated on embryos (mean age: 4 days post fertilisation), whose sex is not discernible by external features.
Wild animals	This study did not involve wild animals.
Field-collected samples	The study does not contain samples collected from the field.
Ethics oversight	All fish are maintained in closed stocks at Heidelberg University. Medaka (Oryzias latipes) husbandry (permit number 35–9185.64/BH Wittbrodt) and experiments (permit number 35-9185.81/G-145/15 Wittbrodt) were performed according to local animal welfare standards (Tierschutzgesetz §11, Abs. 1, Nr. 1) and in accordance with European Union animal welfare guidelines. The fish facility is under the supervision of the local representative of the animal welfare agency. All experiments were performed in the embryonic stage before hatching.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Individuals in both the discovery and the replication cohort were phenotyped by cardiovascular magnetic resonance (CMR) imaging and left ventricular volumes were derived from the images. Left ventricular trabecular mass and complexity were measured using fractal dimension as a proxy phenotype for trabecular complexity. In addition, the age at CMR imaging, standing height, weight, and sex of the participants were used in the study. Full details are available on the UKBB data showcase - https://www.ukbiobank.ac.uk/data-showcase/. UK Biobank discovery cohort: 9402/8694 male/female, average age=55, height=1.70m, weight=77kg; UK Biobank replication cohort: 3378/3158 male/female, average age=54, height=1.70m, weight=76kg; UK Digital Heart Project replication cohort: 621/508 male/female, average age=41, height=1.71m, weight=72kg.
Recruitment	For UK Biobank, approximately 500,000 community-dwelling participants aged 40–69 years were recruited across the United Kingdom between 2006 and 2010. Baseline summary characteristics of the cohort can be viewed on the UK Biobank data showcase (http://www.ukbiobank.ac.uk). Since 2014, a subset of participants are being recalled for CMR. All subjects provided written informed consent for participation in the study, which was also approved by the National Research Ethics Service (11/NW/0382). Our study was conducted under terms of access approval number 18545.
Ethics oversight	National Research Ethics Service (11/NW/0382)

Note that full information on the approval of the study protocol must also be provided in the manuscript.