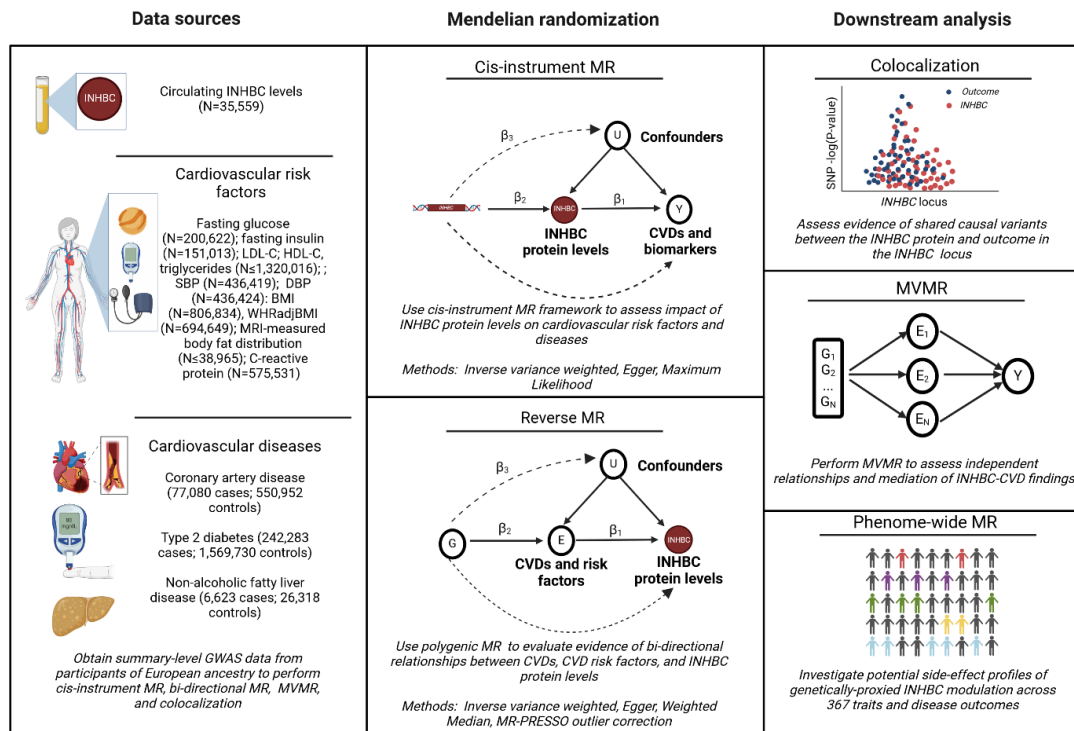


Bidirectional Mendelian randomization highlights causal relationships between circulating INHBC and multiple cardiometabolic diseases and traits.

Nellie Y. Loh, Daniel B. Rosoff, Rebecca Richmond, Raymond Noordam, George Davey Smith, David Ray, Fredrik Karpe, Falk W. Lohoff, and Constantinos Christodoulides.

[Supplemental Information](#)

Population genetics-based studies



In vitro studies

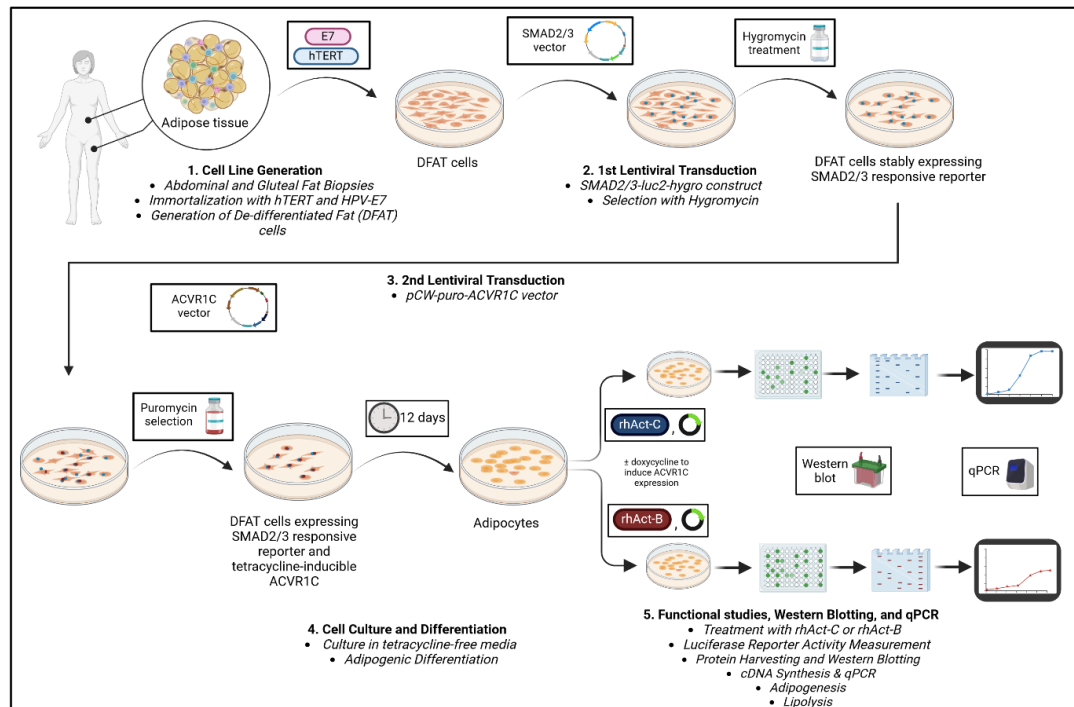


Fig. S1. Study overview. (Created with BioRender.com)

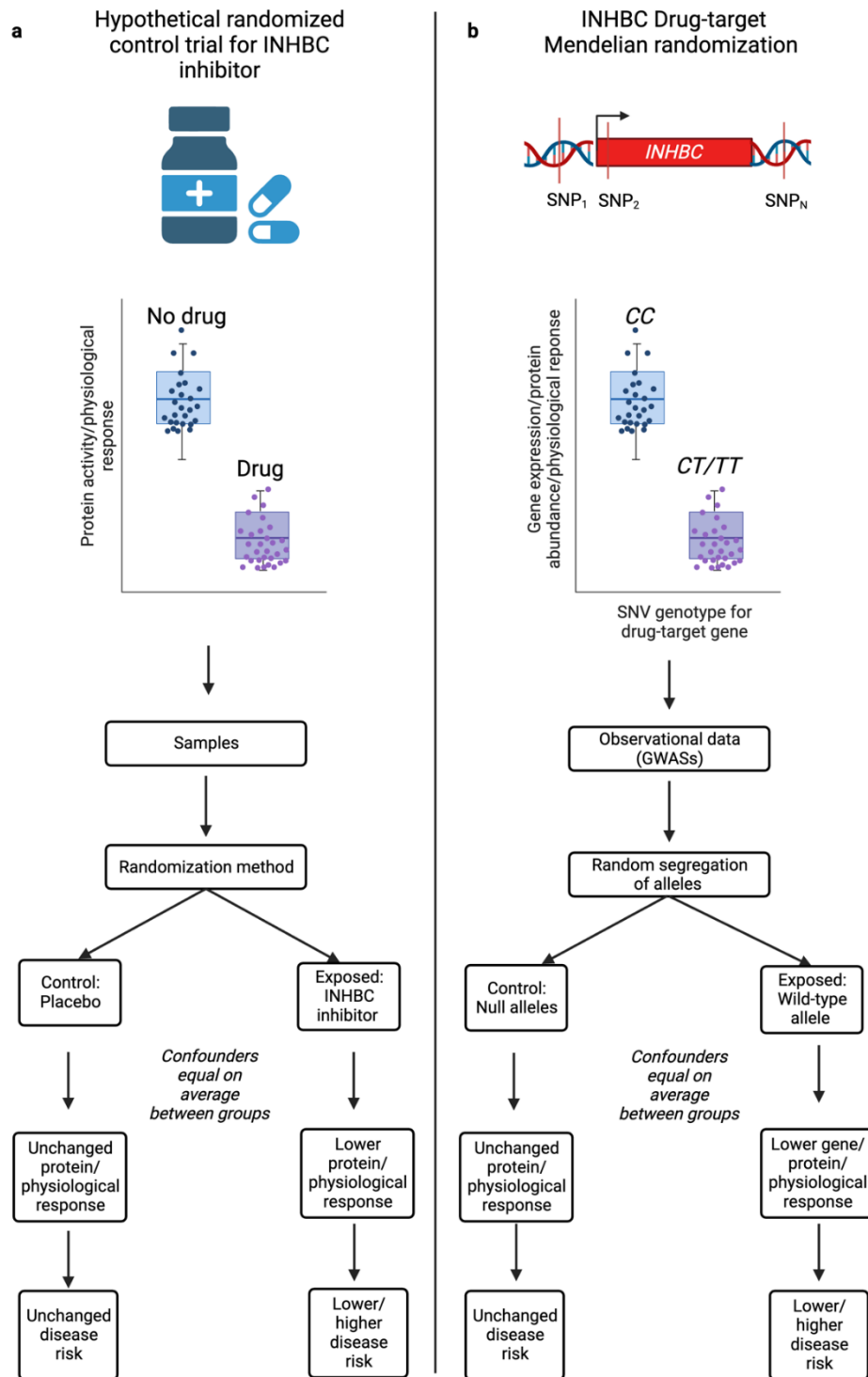


Fig. S2. Comparison of drug-target Mendelian randomization (DTMR) with randomized control trials (RCTs).

The outcome from a RCT for a hypothetical INHBC inhibitor **(a)** compared with drug-target MR analysis assessing the impact of genetically-proxied associations of circulating INHBC protein levels, using single-nucleotide polymorphisms (SNPs) located within or near the genomic loci of the drug target (here the *INHBC* locus: GRCh38, chromosome 12: 57,434,784-57,452,062) associated with the downstream physiological response (e.g., INHBC protein levels) **(b)**. The drug-target instrument SNPs may influence the downstream effects, for example, by impacting transcription regulators. In this diagram, the hypothetical T allele in the SNP used as a drug-target instrument for INHBC is associated with INHBC protein levels. Therefore, individuals with one or more copies of the T allele, i.e., genotypes CT or TT, have analogies to individuals randomized to a RCT treatment arm given the therapeutic under evaluation, while individuals that are homozygous for the C allele have analogies to those randomized to the control arm of a RCT. (Created with BioRender.com)

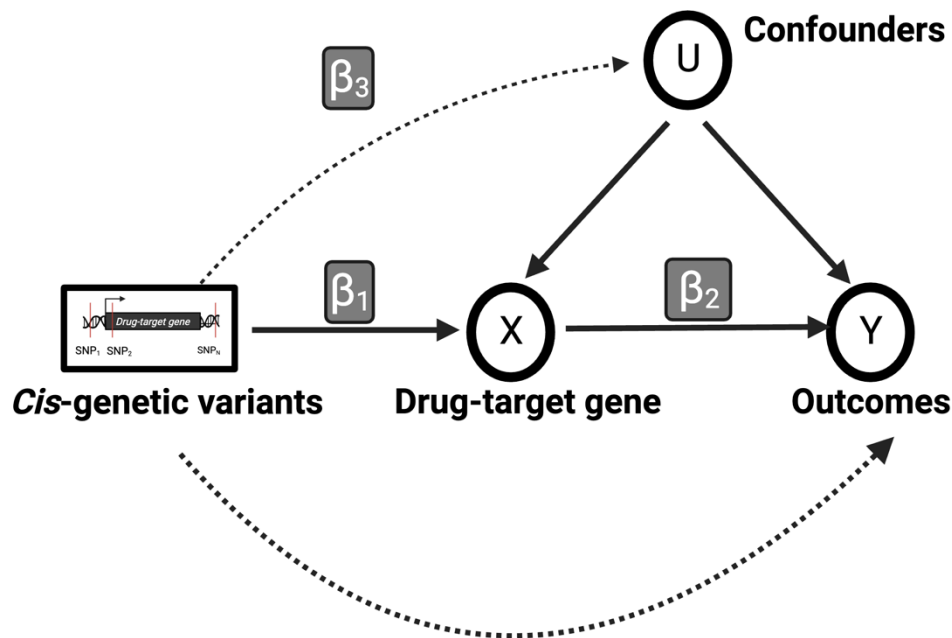
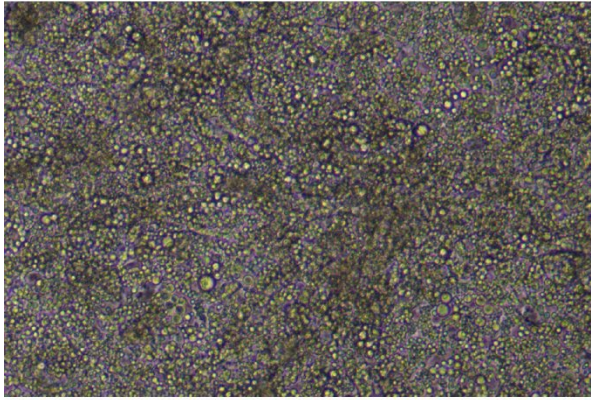


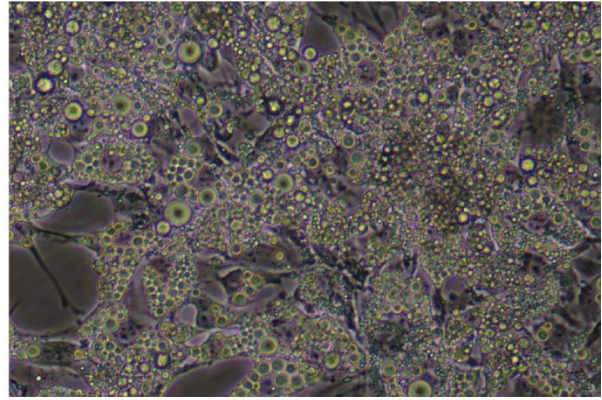
Fig. S3. Drug-target Mendelian randomization (MR) directed acyclic graph.

B_2 is the genetic association of interest estimated by $B_2 = B_1 / B_3$. B_1 and B_3 are the estimated drug-target MR association of the genetic single-nucleotide polymorphisms (SNPs) within the locus ('cis-SNPs') of the drug-target gene (here *INHBC*) on the exposure (here circulating *INHBC* protein levels) and the outcomes. Drug-target MR assumes that the cis-SNPs comprising the exposure instrument for the exposure only impact the outcome(s) of interest via the exposure and not directly, or via confounders (represented by the dotted lines) (1). (Created with BioRender.com)

MBdfat5 Abdo d12



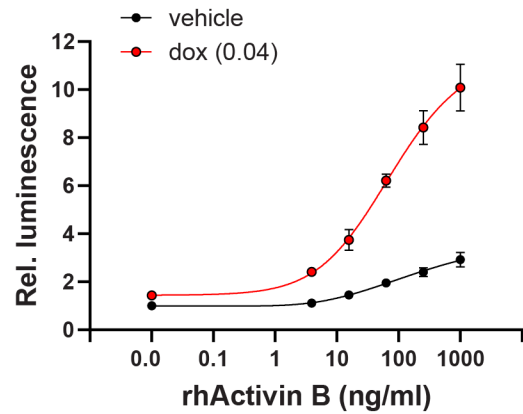
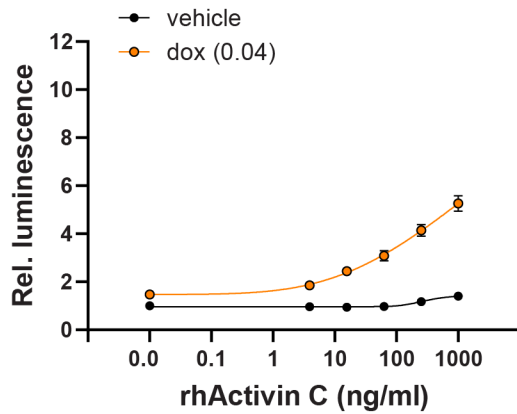
MBdfat5 Glut d12



100μm

Fig. S4. Micrographs of *in vitro* differentiated day 12 MBdfat5 cells.

A Abdo



B Glut

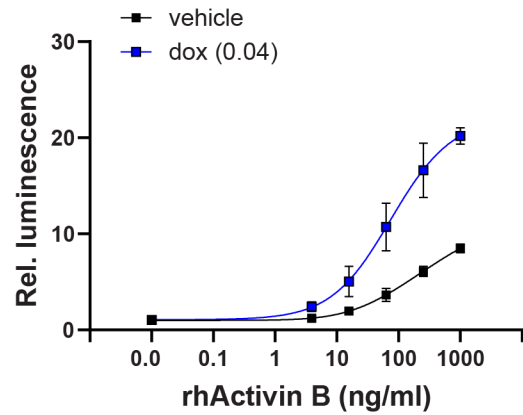
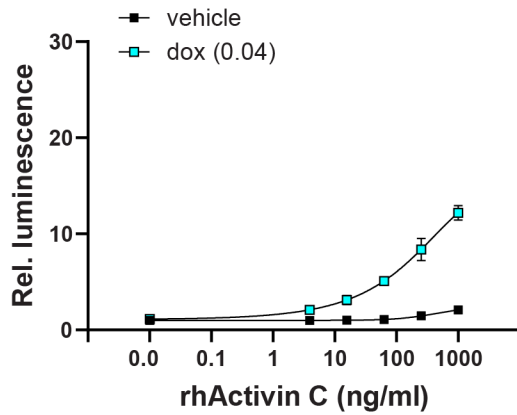


Fig. S5. Comparing the effects of rhAct-C and rhAct-B treatment on SMAD2/3-luciferase reporter activity in DFAT adipocytes co-expressing a dox-inducible ACVR1C vector. Day 10 abdominal (**A**) and gluteal (**B**) adipocytes were treated with 0.04 μ g/ml doxycycline (or vehicle) for 24 hrs, then a further 24 hrs with addition of increasing doses of rhAct-C or rhAct-B ($n = 8$, from 2 independent experiments). Data points are means \pm SD relative luminescence unit (relative to vehicle-treated cells).

Supplemental Methods

Data sources

A study overview is presented in Fig. S1. The MR and other human genetics-based analyses used published, de-identified GWAS data comprised of participants of European ancestry. GWAS for BMI-adjusted waist and hip circumference (WCadjBMI and HCadjBMI) of European-ancestry participants from the UKB were performed on the MRC-IEU UKB GWAS pipeline, using BOLT-LMM (v2.3)(2), and were adjusted for age and sex (and BMI). Details on the data QC and association analysis are found on <https://data.bris.ac.uk/datasets/pnoat8cxo0u52p6ynfaekeigi/MRC%20IEU%20UK%20Biobank%20GWAS%20pipeline%20version%202.pdf>. These studies have existing ethical permissions from their internal review boards and include participant informed consent with rigorous quality control. See Table S1 for a full list of GWAS data used in this study.

Mendelian Randomization

There are differences in the methodology and interpretation of cis-instrument Mendelian randomization (MR) studies compared to conventional MR studies using polygenic instruments with tens to hundreds of SNPs located throughout the genome (i.e., the MR method used for the reverse MR in this study to assess the impact of cardiometabolic traits on circulating INHBC protein levels),(3-5) and below we provide additional background for the cis-instrument MR methodology used for the forward MR analyses using variants located within or near the INHBC locus to genetically model the impact of circulating INHBC protein levels in cardiometabolic health.

Cis-instrument MR leverages genetic variation within or near the genomic locus of the target of interest to evaluate causal relationships with phenotypic outcomes and provide genetics-based evidence to support for the encoded target (here circulating INHBC protein levels).(1, 3, 6) Cis-instrument MR has been shown to have analogies with randomized control trials and has been used in many applications related to drug target validation and discovery.(7, 8)

We used the cis-instrument MR framework (termed “drug-target MR” when applied to known drug targets like PCSK9 inhibitors) outlined by Schmidt et al.,(3) which describes instrumentation strategies in the application of MR methods to gene targets. This cis-instrument MR instrumentation framework includes implementing different P-value and linkage disequilibrium (LD) R^2 clumping thresholds than the P-value and LD R^2 thresholds used in conventional MR instrumentation strategies for complex, polygenic traits (like low-density lipoprotein levels or systolic blood pressure), and performing cis-instrument MR with estimators that incorporate the pairwise LD correlation matrices constructed using the appropriate population reference panels so as to account for residual correlation between the cis-instruments, and thus adding information to the MR model from SNPs in the locus of the gene(s) of interest that would otherwise not be included (setting P-value and LD R^2 thresholds used in conventional MR instrumentation strategies for complex, polygenic traits). This cis-instrument MR framework has been used extensively in cis-instrument/drug-target MR studies aiming to identify and prioritize new drug targets.(7, 9-14) Specifically, we filtered the variants located within or near the INHBC locus using the 1000 Genomes Project Phase 3 European reference panel,(15) and clumped the variants

at linkage disequilibrium (LD) $R^2 < 0.2$ (250 kb window). This LD R^2 threshold was selected to maximize power while also remaining below LD R^2 thresholds shown by simulation and applied studies to produce unstable MR estimates (LD $R^2 > \sim 0.4$).^(3, 16) We also constructed correlation matrices to account for any SNP-SNP LD (9) in the resulting cis-INHBC instrument and incorporate these matrices into the MR estimates for the inverse variance weighted, Egger, and Maximum Likelihood methods used for the cis-instrument MR analyses. This cis-instrument framework provides a basis from which to optimize instrumentation parameters in cis-instrument MR analyses when using only variants located in/near the locus of the target of interest – here variants within or near the *INHBC* to proxy the impact of circulating INHBC protein levels.

For the reverse Mendelian Randomization (MR), we employed a conventional two-sample MR approach to evaluate the impact of various cardiometabolic traits on circulating levels of the INHBC protein. In this analysis, we utilized polygenic genetic instruments, which consist of single nucleotide polymorphisms (SNPs) distributed across the entire genome, to model the polygenic nature of the exposures. Given the complexity and polygenic nature of these traits, we adhered to the standard two-sample MR clumping parameters for polygenic instrumentation, specifically using a LD R^2 threshold of 0.001 and a physical distance of 10,000 kb. These parameters are essential for effectively capturing the genetic architecture of complex traits, which do not correspond to a single genomic locus encoding the exposure trait while also minimizing potential sources of bias and violations of the instrumental variable assumptions.⁽¹⁾

Multivariate Mendelian Randomization

MVMR analyses were undertaken only if there was evidence for causal effect in the primary MR analyses. For MVMR, we used GWAS summary statistics derived from non-overlapping samples for exposure and outcome data. Exposure data for exposures and the mediator(s) of interest were extracted for MVMR using `mv_extract_exposures_local()`. All MR and MVMR analyses were conducted using the `TwoSampleMR` (v0.5.7) (6) and `MendelianRandomization` (v.0.8.0) (17) packages in R (v4.3.1).

Cell lines

The MBdfat5 abdominal and gluteal de-differentiated fat (DFAT) cells lines were generated as follows: Stromovascular cells (SVCs) from abdominal and gluteal fat biopsies of a healthy female donor were immortalized as previously described (18). DFAT cells were derived from immortalized human SVCs by two rounds of [*in vitro* adipogenesis, selection of lipid-laden cells by trypsinization, centrifugation and ceiling culture, and de-differentiation in growth media] (18). The resulting DFAT daughter cells have high adipogenic potential (Fig. S4).

Cell culture

DFAT cell lines were maintained in tetracycline-free media (DMEM-F12 supplemented with 10% FBS (ThermoFisher Scientific Gibco, #26140079), 2 mM L-glutamine, 0.25 ng/ml fibroblast growth factor, 100 units/ml penicillin and 100 µg/ml streptomycin). For *in vitro* adipogenesis, confluent cells were cultured for 12 days in an adipogenic medium (DMEM-F12 containing 2 mM L-glutamine, 100 units/ml penicillin, 100 µg/ml

streptomycin, 17 μ M pantothenate, 100 nM human insulin, 10 nM 3,3',5-triiodo-L-thyronine, 33 μ M biotin, 10 μ g/ml human transferrin and 1 μ M dexamethasone). For the first 4 days, 250 μ M 3-isobutyl-1-methylxanthine and 4 μ M troglitazone were added to the adipogenic medium.

Lipolysis

Basal and stimulated lipolysis was performed on *in vitro* differentiated DFAT[SMAD2/3-luc2/pCW-ACVR1C] cells in 12-well plates as follows: Cells were differentiated as above for the first seven days, and thereafter cultured in low glucose, low insulin maintenance media (DMEM-F12 with 5mM glucose, L-glutamine, 100 units/ml penicillin, 100 μ g/ml streptomycin, 17 μ M pantothenate, 33 μ M biotin, 10 μ g/ml human transferrin and 50 pM human insulin). *In vitro* differentiated day 10 cells were cultured in the presence of 0.02 μ g/ml doxycycline (to induce *ACVR1C* expression) or vehicle for 24-48 hrs, then a further ~24 hrs with the addition of 25 ng/ml rhAct-C or vehicle, prior to the lipolysis experiments. On the day of lipolysis experiment, cells underwent a 1-hour equilibration period in Krebs Ringer HEPES buffer containing 6mM glucose and 3.5% BSA (KRH), then incubated for 2 hrs in fresh KRH buffer with 50 pM insulin (for basal lipolysis) or KRH buffer containing 50 pM insulin and 10 nM adrenaline (for stimulated lipolysis), with the addition of doxycycline (0.02 μ g/ml) and rhAct-C (25 ng/ml), or vehicle, as indicated. At the end of the 2-hour period, the treatment buffer was collected for glycerol measurements, and cells were rinsed in PBS and harvested in lysis buffer containing 1% IGEPAL CA-630, 150 mM NaCl and 50 mM Tris-HCl pH8.0 for protein. Glycerol concentrations were measured

using the GY105 enzymatic assay (Randox Laboratories Ltd) on an AU480 clinical analyser (Beckman Coulter) and normalised to cellular protein concentration.

RNA purification, first strand cDNA synthesis and quantitative PCR (qPCR)

RNA was purified using the RNeasy Mini kit (Qiagen) and reversed transcribed to cDNA using the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems). qPCR was performed using Taqman assays (ThermoFisher Scientific). qPCR results were normalized to 18S.

Statistics

Statistical analyses were performed in GraphPad Prism v10.2.3.

References

1. Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafò MR, et al. Mendelian randomization. *Nature Reviews Methods Primers*. 2022;2:6
2. Loh PR, Tucker G, Bulik-Sullivan BK, Vilhjalmsdottir BJ, Finucane HK, Salem RM, et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet*. 2015;47:284-290
3. Schmidt AF, Finan C, Gordillo-Marañón M, Asselbergs FW, Freitag DF, Patel RS, et al. Genetic drug target validation using Mendelian randomisation. *Nature communications*. 2020;11:3255-3255
4. Gill D, Georgakis MK, Walker VM, Schmidt AF, Gkatzionis A, Freitag DF, et al. Mendelian randomization for studying the effects of perturbing drug targets. *Wellcome Open Res*. 2021;6:16
5. Holmes MV, Richardson TG, Ference BA, Davies NM, Davey Smith G. Integrating genomics with biomarkers and therapeutic targets to invigorate cardiovascular drug development. *Nature Reviews Cardiology*. 2021;18:435-453
6. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018;7
7. Henry A, Gordillo-Marañón M, Finan C, Schmidt AF, Ferreira JP, Karra R, et al. Therapeutic Targets for Heart Failure Identified Using Proteomics and Mendelian Randomization. *Circulation*. 2022;145:1205-1217

8. Wingo TS, Gerasimov ES, Liu Y, Duong DM, Vattathil SM, Lori A, et al. Integrating human brain proteomes with genome-wide association data implicates novel proteins in post-traumatic stress disorder. *Mol Psychiatry*. 2022;27:3075-3084
9. Gordillo-Marañón M, Zwierzyna M, Charoen P, Drenos F, Chopade S, Shah T, et al. Validation of lipid-related therapeutic targets for coronary heart disease prevention using human genetics. *Nature Communications*. 2021;12:6120
10. Schmidt AF, Bourfiss M, Alasiri A, Puyol-Anton E, Chopade S, van Vugt M, et al. Druggable proteins influencing cardiac structure and function: Implications for heart failure therapies and cancer cardiotoxicity. *Science Advances*. 2023;9:eadd4984
11. Prapiadou S, Živković L, Thorand B, George MJ, Laan SWvd, Malik R, et al. Proteogenomic Data Integration Reveals CXCL10 as a Potentially Downstream Causal Mediator for IL-6 Signaling on Atherosclerosis. *Circulation*. 2024;149:669-683
12. Storm CS, Kia DA, Almramhi MM, Bandres-Ciga S, Finan C, Noyce AJ, et al. Finding genetically-supported drug targets for Parkinson's disease using Mendelian randomization of the druggable genome. *Nature Communications*. 2021;12:7342
13. Rosoff DB, Mavromatis LA, Bell AS, Wagner J, Jung J, Marioni RE, et al. Multivariate genome-wide analysis of aging-related traits identifies novel loci and new drug targets for healthy aging. *Nature Aging*. 2023;3:1020-1035
14. Bouras E, Karhunen V, Gill D, Huang J, Haycock PC, Gunter MJ, et al. Circulating inflammatory cytokines and risk of five cancers: a Mendelian randomization analysis. *BMC Med*. 2022;20:3
15. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, et al. A global reference for human genetic variation. *Nature*. 2015;526:68-74
16. Burgess S, Zuber V, Valdes-Marquez E, Sun BB, Hopewell JC. Mendelian randomization with fine-mapped genetic data: Choosing from large numbers of correlated instrumental variables. *Genet Epidemiol*. 2017;41:714-725
17. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. *International Journal of Epidemiology*. 2017;46:1734-1739
18. Loh NY, Minchin JEN, Pinnick KE, Verma M, Todorovic M, Denton N, et al. RSPO3 impacts body fat distribution and regulates adipose cell biology in vitro. *Nat Commun*. 2020;11:2797