Target genes, variants, tissues and transcriptional pathways influencing human serum urate levels

Adrienne Tin, Jonathan Marten, Victoria Halperin Kuhns,
Yong Li, Matthias Wuttke, Holger Kirsten, *et al.*

SUPPLEMENTARY INFORMATION

Table of Contents

Supplementary Note

Ancestry-specific meta-analyses of serum urate

In addition to the trans-ethnic meta-analysis, ancestry-specific GWAS meta-analyses were also conducted. Genome-wide significant loci for each ancestry group are summarized in **Supplementary Tables 5-8**. Five of 123 EA-specific significant loci were ancestry-specific, i.e. not contained within the 183 significant trans-ethnic loci (**Supplementary Table 5**). Among 14 significant loci in AA, five were ancestry-specific (**Supplementary Table 6**), and among the 46 significant loci in EAS, seven were ancestry-specific (**Supplementary Table 7**). None of the six genome-wide significant loci in SA was ancestry-specific (**Supplementary Table 8**). The ancestry-specific findings contained several plausible candidates, including *NIPAL1* in EA, a risk locus previously identified for renal under-excretion gout[1], and rs334 at *HBB* in AA, encoding the sickle cell allele. This finding is consistent with higher urate levels and gout incidence in sickle cell disease patients as compared to the general population[2].

Association between urate-associated SNPs and gene expression in *trans*

We investigated whether any trans-ethnic index SNPs or their proxies were reproducibly associated with gene expression in *trans* in whole blood or peripheral blood mononuclear cell in large eQTL studies (Methods). A total of 17 inter-chromosomal trans-eQTLs based on five index SNPs were identified (**Supplementary Table 21**). The urate-increasing allele of the index SNP rs174594 at *FADS2* on chromosome 11 was associated with higher expression of *LDLR* on chromosome 19 in two independent studies. Higher *LDLR* expression under inflammatory conditions was reported to promote transformation of macrophages into foam cells leading to increased vascular calcification.[3] The urate-increasing allele of the intronic SNP rs12530084 in *RREB1* on chromosome 6 was associated with higher expression of *SYT11* on chromosome 1 in two independent studies. *SYT11* is involved in calcium-dependent regulation of membrane trafficking in synaptic transmission, and both *RREB1* and *SYT11* are predominantly expressed in brain cerebellum. In addition, *SYT11* maps into one of the novel urate-associated loci identified in this study. The associated trans-eQTLs were enriched for several pathways as summarized in the GO, KEGG and Reactome databases (**Supplementary Table 22**), as well as for “cardiovascular disease” based on the Human Disease Ontology database.[4]

Phenotype generation and quality control of study-specific results

We developed an automated analysis workflow to collect and integrate results from 74 GWAS of serum urate from five ancestry groups participating in the CKDGen Consortium. We used a distributive model for study-specific GWAS, with meta-analyses conducted centrally. An analysis plan was circulated to all participating studies accompanied by custom shell and R scripts for phenotype generation (https://github.com/genepi-freiburg/ckdgen-pheno). Study-specific GWAS were conducted after centralized review and approval of the phenotype summary statistics.

Genotype imputation was conducted by each study. Haplotypes were estimated using MACH[5], ShapeIT[6], Eagle[7], or Beagle[8]. Imputation of genotypes was conducted using reference panels from the Haplotype Reference Consortium (HRC) version 1.1[9], 1000 Genomes Project (1000G) phase 3 v5 ALL, or the 1000G phase 1 v3 ALL[10] and ImputeV2[11], minimac3[12], PBWT[13], the Sanger[14], or the Michigan Imputation Server[12]. The imputed genetic dosages were annotated using NCBI b37 (hg19). Each study provided an imputation quality for each variant: ImputeV2 info score, the MACH/minimac RSQ or the SNPTest info score.

Study-specific GWAS results were checked using GWAtoolbox[15], including *P*-value inflation, allele frequency distribution, imputation quality, and completeness of genotypes. Custom scripts were used to compare imputed allele frequencies to those of ancestry-matched reference panels and to visualize variant positions. In addition, quality metrics, including the genomic control factor[16], were compared across studies.

DEPICT analysis details for functional enrichment

DEPICT performs gene set enrichment analysis by testing whether genes in 14,461 reconstituted gene sets were enriched for urate-associated SNPs. These reconstituted gene sets were generated based on similarity analysis from gene expression of 77,840 samples, manually curated gene-sets, molecular pathways from protein-protein interaction screening, and gene sets from mouse gene knock-out studies. Additionally, gene expression levels of from 37,427 human microarray samples are compiled in DEPICT and the tissue and cell types were mapped to 209 MeSH first level terms including physiological systems, tissues and cells. Enrichment of tissues and cell types was conducted by assessing whether urate-associated genes are highly expressed in the tissue and cell types.

All variants with urate association *P*-values <$1x10^{-5}$ from the trans-ethnic meta-analysis results were used as input. Independent SNPs were identified using PLINK 1.9[17] clump command within 500 kb flanking regions and r^2>0.1 in the 1000 Genomes phase1 version 3 data excluding the MHC region (chr6:25–35 Mb). False discovery rates (FDRs) were computed using 500 repetitions, and *P*-values were computed using 5,000 permutations from 500 null GWAS sets adjusting for gene length.

Affinity propagation clustering (APC)[18], implemented in the R package 'APCluster'[19], was applied to all reconstituted gene sets with FDR-corrected enrichment *P*-value <0.01 to cluster gene sets containing similar combinations of genes. Similarity between sets was assessed by the Z-scores of the top ten genes assigned to each gene set. The algorithm identifies a single reconstituted gene set from each cluster as an exemplar (meta gene set) that best represents the reconstituted gene sets within that cluster. A correlation matrix was calculated from the Z-scores of the top ten genes within the meta gene set. Correlations >0.2 were visualized as edges in a network in Cytoscape (http://cytoscape.org).

Stratified LD score regression for functional enrichment

Stratified LD score regression[20] estimates the SNP heritability of urate contributed by the SNPs linked to histone marks in each cell type. The enrichment of a category is defined as the proportion of SNP heritability in that cell type divided by the proportion of SNPs in the same cell type. Here, we assessed urate heritability enrichment in 10 cell types via stratified LD score regression with the EA-specific urate meta-analysis results as the input to match the ancestry of the LD score estimates. The 10 cell types were collapsed from 220 cell type-specific annotations for four histone marks: H3K4me1, H3K4me3, H3K9ac, and H3K27ac. Analyses were also carried out using trans-ethnic meta-analysis summary statistics as input and results were similar (data not shown).

Quality control of the UKBB sample used as to compute an LD reference panel for statistical fine-mapping

Fine-mapping based on summary statistics relies on LD estimates from an ancestry-matched reference panel whose sample size should scale with that of the GWAS[21]. Given the large sample size in the urate meta-analysis, we estimated LD from genotypes of 15,000 randomly

selected UKBB participants (Application ID 2027, Dataset ID 8974). As reported previously[22], we first removed individuals who withdrew consent, those with mismatched sex between self-reported and genetic, with genotype call rates <95%, outliers of variant heterozygosity or along the first two principal components from a principal component analysis seeded with the HapMap phase 3 release 2 populations. We retained only one member of each pair of individuals with pair-wise identity-by-descent statistic ≥0.1875. Altogether, 13,558 individuals with 16,969,363 SNPs were selected as the LD reference panel for fine-mapping.

Workflow for statistical fine-mapping

First, we combined neighboring loci with correlated index SNPs (r^2≥0.2) in the EA meta-analysis into discrete regions. Next, we performed stepwise model selection using GCTA (cojo-slct option) to identify independent index SNPs in each region. For regions with more than one independent index SNPs, we performed conditional analysis (GCTA cojo-cond option) to obtain conditional beta and standard errors. Approximate Bayes factors (ABF) were calculated using the Wakefield's formula[23], as implemented in the R package 'gtx' version 2.0.1 (https://github.com/tobyjohnson/gtx) using the conditional betas and standard errors for regions with multiple independent SNPs and the betas and standard errors of the original EA meta-analysis for regions with a single independent index SNP. The prior standard deviation was calculated as 0.061 based on formula (8) of the original publication of the Wakefield's formula [23] and the 95% interval of the SNP effect sizes in the EA meta-analysis. The posterior probability for a variant being the driver of the association signal was calculated as the ABF of the variant divided by the sum of the ABF in the region. The 99% credible sets of a region is derived by summing the posterior probabilities in descending order until the cumulative posterior probability was >99%. We prioritized variants in credible sets containing ≤5 SNPs or SNPs with posterior probabilities >0.5.

We compared the list of prioritized variants with results using the conditional mode in FINEMAP, an alternative approach for statistical fine-mapping.[24] We used the same prior standard deviation and EA meta-analysis summary statistics as input as those for the Wakefield method. The correlations between SNPs were based on the same UKBB reference panel as used in GCTA model selection and conditional analysis.

Trans-eQTL annotation

We performed *trans*-eQTL annotation for the 183 trans-ethnic index SNPs and their proxies (r^2 >0.8 in both the 1000 Genomes phase 3 European and East Asian reference panels) in eQTL studies with >1,000 individuals: the Framingham Heart Study[25], Westra et al.[26], Zeller et al.[27], Fehrmann et al.[28], and the LIFE Heart[29] and LIFE-Adult[30] studies. Westra *et al*. used whole blood from 5,257 individuals from the Framingham Heart Study and Affymetrix Human Exon 1.0ST expression-chip technology. Westra *et al*. is an eQTL-meta-analysis of nine datasets that used whole blood and Illumina HumanRef-8v2.0, Illumina HT12v3, or Illumina HT12v4 expression-chip technology. *Trans*-eQTL analysis of Westra *et al*. was restricted to previously reported GWAS SNPs. Fehrmann *et al*.[28] used whole blood from 1,469 individuals and Illumina HumanRef-8v2.0 and Illumina HT12v3 expression-chip technology. *Trans*-eQTL analysis of Fehrmann *et al*. was performed genome-wide. Fehrmann *et al*. was not regarded as an independent study in respect to Westra *et al.* as it was included in the meta-analysis of the latter. With respect to the other eQTL-studies, Fehrmann *et al*. is an independent study. Zeller *et al*[27]. used monocytes from 1,469 individuals and Illumina HT12v3 technology. LIFE Heart and LIFE Adult comprised an update of a previous eQTL-study[31] combining 4,285 peripheral blood mononuclear cell-samples from LIFE-Heart[321] and 2,360 whole-blood samples from LIFE-Adult[33] totaling 6,645 samples. Measurement and analysis of the combined data was done as previously described,[31] with the difference that FDR of the *trans*-eQTL from the LIFE Heart and LIFE Adult combined dataset was calculated after applying a hierarchical testing procedure based on Benjamini and Bogomolov[34] to avoid excess of false positives on the gene level. To account for differences between the LIFE Heart and LIFE Adult studies, a study identifier was included as a binary covariate in the regression model.

To improve stringency of results, we only report inter-chromosomal *trans*-eQTLs showing gene expression association *P*-values $<5x10^{-8}$ in at least two of the above mentioned independent samples. To characterize whether the identified *trans*-associated eQTL transcripts were enriched in any biological pathways, we conducted gene enrichment analysis using DOSE, an R/Bioconductor package for disease ontology semantic and enrichment analysis[35], as well as packages ReactomePA and clusterProfiler for enrichment analysis using the Human Disease Ontology database, GO, KEGG, and Reactom[4,36-39]. The background included all 19,327 protein-coding genes from Ensembl/Havana reported in Ensembl release 91.

Site-Directed Mutagenesis Primers

PAGE purified primers:

HNF1A-A98V-Forward: CCCTGAGGAGGCGGTCCACCAGAAAGCCG;

HNF1A-A98V-Reverse: CGGCTTTCTGGTGGACCGCCTCCTCAGGG;

HNF4A-T139I-Forward: GACCGGATCAGCATTCGAAGGTCAAGC;

HNF4A-T139I-Reverse: GCTTGACCTTCGAATGCTGATCCGGTC.

Extended Acknowledgements

The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS.

Study-specific acknowledgements and funding sources

AGES This study has been funded by NIA contract N01-AG-12100 and HHSN271201200022C with contributions from NEI, NIDCD and NHLBI, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). The study is approved by the Icelandic National Bioethics Committee, VSN: 00-063. The researchers are indebted to the participants for their willingness to participate in the study.

Amish We thank our Amish research volunteers for their long-standing partnership in research, and the research staff at the Amish Research Clinic for their hard work and dedication. The Amish contribution was supported by NIH grants R01 AG18728, R01 HL088119, U01 GM074518, U01 HL072515, U01 HL084756, and NIH K12RR023250, and P30 DK072488. Additional support was provided by the University of Maryland General Clinical Research Center, grant M01 RR 16500, the Baltimore Veterans Administration Medical Center Geriatrics Research, and the Paul Beeson Physician Faculty Scholars in Aging Program.

ARIC The Atherosclerosis Risk in Communities study has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services (contract numbers HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700004I and HHSN268201700005I), R01HL087641, R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. The work of AK was supported by a Heisenberg Professorship (KO 3598/3-1, 3598/5-1) as well as CRC 1140 project number 246781735 and CRC 992 of the

German Research Foundation. Adrienne Tin is supported by NIAMS grant R01AR073178-01A1.

ASPS, ASPS-Fam The authors thank the staff and the participants for their valuable contributions. We thank Birgit Reinhart for her long-term administrative commitment, Elfi Hofer for the technical assistance at creating the DNA bank, Ing. Johann Semmler and Anita Harb for DNA sequencing and DNA analyses by TaqMan assays and Irmgard Poelzl for supervising the quality management processes after ISO9001 at the biobanking and DNA analyses.The research reported in this article was funded by the Austrian Science Fond (FWF) grant number P20545-P05, P13180 and PI904 as well as by the Austrian National Bank (OeNB) Anniversary Fund grant number P15435 and the Austrain Federal Ministry of Science, Research and Economy under the aegis of the EU Joint Programme-Neurodegenerative Disease Research (JPND)-www.jpnd.eu. The Medical University of Graz supports the databank of the ASPS.

BBJ The BioBank Japan project is supported by the Ministry of Education, Culture, Sports, Sciences and Technology of Japanese government and the Japan Agency for Medical Research and Development.

BioMe The Mount Sinai IPM Biobank Program is supported by The Andrea and Charles Bronfman Philanthropies. Ruth loos is funded by R01DK110113, U01HG007417, R01DK101855, R01DK107786.

CHNS The China Health and Nutrition Survey (CHNS) was supported by the China National Institute of Nutrition and Food Safety; the China Center for Disease Control; the National Institutes of Health (R01HD30880, R01HL108427, and R01DK104371); the Fogarty International Center of the National Institutes of Health; the China-Japan Friendship Hospital; the Chinese Ministry of Health; and the Carolina Population Center (R24 HD050924). Cassandra N. Spracklen was supported by the American Heart Association Postdoctoral Fellowship (17POST3650016).

CHRIS Full acknowledgements for the CHRIS study are reported here: http://translational-medicine.biomedcentral.com/articles/10.1186/s12967-015-0704-9#Declarations. The CHRIS study was funded by the Department of Innovation, Research, and University of the Autonomous Province of Bolzano-South Tyrol.

CHS This CHS research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, HHSN268201800001C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, R01HL120393, R01HL085251, and U01HL130114 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org.

The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

	The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.
Cilento	We thank the populations of Cilento for their participation in the study. This work was supported by grants from the Italian Ministry of Universities and Research and CNR (Interomics Flagship Project, PON03PE_00060_7), the Assessorato Ricerca Regione Campania, the Fondazione con il SUD (2011-PDR-13), and the Istituto Banco di Napoli - Fondazione to MC.
CoLaus	The CoLaus|PsyCoLaus study was and is supported by research grants from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne, and the Swiss National Science Foundation (grants 3200B0–105993, 3200B0-118308, 33CSCO-122661, 33CS30-139468 and 33CS30-148401).
CROATIA-Korcula, CROATIA-Split, CROATIA-Vis	10001 Dalmatians: The Croatian Biobank (CROATIA) The CROATIA-Vis, CROATIA-Korcula and CROATIA-Split were funded by grants from the Medical Research Council (UK), from the Republic of Croatia Ministry of Science, Education and Sports (108-1080315-0302; 216-1080315-0302) and the Croatian Science Foundation (8875); and the CROATIA-Korčula genotyping was funded by the European Union framework program 6 project EUROSPAN (LSHGCT2006018947). We would like to acknowledge the staff of several institutions in Croatia that supported the field work, including Zagreb Medical Schools, the Institute for Anthropological Research in Zagreb, the recruitment team from the Croatian Centre for Global Health, University of Split and all the study participants. We are grateful to the Helmholtz Zentrum Munchen (Munich, Germany), AROS Applied Biotechnology, (Aarhus, Denmark) and the Edinburgh Clinical Research facility, University of Edinburgh (Edinburgh, United Kingdom) for SNP array genotyping. Genetic analyses were supported by the MRC HGU “QTL in Health and Disease” core programme.
CZECH POST-MONICA	The study was supported by research grant 15-27109A provided by the Health Research Agency of the Ministry of Health, Czech Republic; Krka, tovarna zdravil, d.d., Novo mesto, Slovenia; Servier s.r.o., Czech Republic.
DECODE	The study was funded by deCODE Genetics/Amgen inc. We thank the study subjects for their valuable participation and our colleagues, who contributed to data collection, sample handling, and genotyping.
DIACORE	Cohort recruiting and management was funded by the KfH Stiftung Präventivmedizin e.V. (CAB). Genome-wide genotyping was funded the Else Kröner-Fresenius-Stiftung (2012_A147), the KfH Stiftung Präventivmedizin and the University Hospital Regensburg. Data analysis was funded by the Else Kröner-Fresenius Stiftung (2012_A147), by the Deutsche Forschungsgemeinschaft with grant DFG BO 3815/1-4 and CRC 1350/C6.
EGCUT	The EGCUT studies were financed by Estonian Government (grants IUT20-60 and IUT24-6) and by European Commission through the European Regional Development Fund in the frame of grant Estonian Center of Genomics/Roadmap II (project No. 2014-2020.4.01.16-0125) and grant GENTRANSMED (Project No. 2014-2020.4.01.15-0012) and through H2020 grant no 692145 (ePerMed).
ERF	The Erasmus Rucphen Family (ERF) has received funding from the Centre for Medical Systems Biology (CMSB) and Netherlands

Consortium for Systems Biology (NCSB), both within the framework of the Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research (NWO). ERF study is also a part of EUROSPAN (European Special Populations Research Network) (FP6 STRP grant number 018947 (LSHG-CT-2006-01947)); European Network of Genomic and Genetic Epidemiology (ENGAGE) from the European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F4-2007-201413; "Quality of Life and Management of the Living Resources" of 5th Framework Programme (no. QLG2-CT-2002-01254); FP7 project EUROHEADPAIN (nr 602633), the Internationale Stichting Alzheimer Onderzoek (ISAO); the Hersenstichting Nederland (HSN); and the JNPD under the project PERADES (grant number 733051021, Defining Genetic, Polygenic and Environmental Risk for Alzheimer's Disease using multiple powerful cohorts, focused Epigenetics and Stem cell metabolomics). We are grateful to all study participants and their relatives, general practitioners and neurologists for their contributions and to P. Veraart for her help in genealogy, J. Vergeer for the supervision of the laboratory work, and P. Snijders M.D. for his help in data collection of the data.J.L., C.M.v.D., and A.D. have used exchange grants from the Personalized pREvention of Chronic DIseases consortium (PRECeDI). A.D. is supported by a Veni grant (2015) from ZonMw.

FHS The Framingham Heart Study is supported by HHSN268201500001.

FINCAVAS The Finnish Cardiovascular Study (FINCAVAS) has been financially supported by the Competitive Research Funding of the Tampere University Hospital (Grant 9M048 and 9N035), the Finnish Cultural Foundation, the Finnish Foundation for Cardiovascular Research, the Emil Aaltonen Foundation, Finland, the Tampere Tuberculosis Foundation, and EU Horizon 2020 (grant 755320 for TAXINOMISIS). The authors thank the staff of the Department of Clinical Physiology for collecting the exercise test data.

GCKD The GCKD study was funded by the German Ministry of Research and Education (Bundesminsterium für Bildung und Forschung, BMBF) and by the Foundation KfH Stiftung Präventivmedizin. Unregistered grants to support the study were provided by Bayer, Fresenius Medical Care and Amgen. Genotyping was supported by Bayer Pharma AG. The work of MW was supported by DFG CRC 1140 project number 246781735 and the Else Kroener Fresenius Forschungskolleg NAKSYS. The work of YL was supported by DFG KO 3598/4-1.

GS:SFHS Generation Scotland received core support from the Chief Scientist Office of the Scottish Government Health Directorates [CZD/16/6] and the Scottish Funding Council [HR03006]. We are grateful to all the families who took part, the general practitioners and the Scottish School of Primary Care for their help in recruiting them, and the whole Generation Scotland team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, healthcare assistants and nurses. Genotyping of the GS:SFHS samples was carried out by the Genetics Core Laboratory at the Wellcome Trust Clinical Research Facility, Edinburgh, Scotland and was funded by the Medical Research Council UK and the Wellcome Trust (Wellcome Trust Strategic Award "STratifying Resilience and Depression Longitudinally" (STRADL) Reference 104036/Z/14/Z).

GSK Funding Source: Max-Planck Society, German Federal Ministry of Education and Research (BMBF) in the framework of the National Genome Research Network (NGFN), Foerderkennzeichen 01GS0481.

HANDLS The authors would thank all study participants and the Healthy Aging in Neighborhoods of Diversity across the Life Span (HANDLS) study medical staff for their contributions. The HANDLS study was supported by the Intramural Research Program of the NIH, National Institute on Aging and the National Center on Minority Health and Health Disparities (project # Z01-AG000513 and human subjects protocol number 09-AG-N248). Data analyses for the HANDLS study utilized the high-performance computational resources of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD. (http://biowulf.nih.gov; http://hpc.nih.gov).

HYPERGENES HYPERGENES project (FP7-HEALTH-F4-2007-201550) and InterOmics (PB05 MIUR-CNR Italian Flagship Project).

INGI-CARL The project was approved by the local administration of Carlantino, the Health Service of Foggia Province, Italy, and ethical committee of the IRCCS Burlo-Garofolo of Trieste. We would like to thank the people of the Friuli Venezia Giulia Region and of Carlantino for the everlasting support.

INGI-FVG Project co-financed by the European Regional Development Fund under the Regional Operational Programme of Friuli Venezia Giulia - Objective "Regional Competitiveness and Employment" 2007/2013, Telethon Foundation (GGP09037), Fondo Trieste (2008), Regione FVG (L.26.2008), and Italian Ministry of Health (RC16/06, ART. 13 D.LGS 297/99) (to PG). We would like to thank the people of the Friuli Venezia Giulia Region and of Carlantino for the everlasting support.

INGI-VBI The research was supported by funds from Compagnia di San Paolo, Torino, Italy; Fondazione Cariplo, Italy and Ministry of Health, Ricerca Finalizzata 2011 and CCM 2010 to DT. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank the inhabitants of the VB that made this study possible, the local administrations, the Tortona and Genova archdiocese and the ASL-22, Novi Ligure (Al) for support. We also thank Clara Camaschella for data collection supervision and organization of the clinical data collection, Fiammetta Viganò for technical help, Corrado Masciullo and Massimiliano Cocca for building the analysis platform.

JHS The Jackson Heart Study (JHS) is supported and conducted in collaboration with Jackson State University (HHSN268201800013I), Tougaloo College (HHSN268201800014I), the Mississippi State Department of Health (HHSN268201800015I/HHSN26800001) and the University of Mississippi Medical Center (HHSN268201800010I, HHSN268201800011I and HHSN268201800012I) contracts from the National Heart, Lung, and Blood Institute (NHLBI) and the National Institute for Minority Health and Health Disparities (NIMHD). The authors also wish to thank the staffs and participants of the JHS. James G. Wilson is supported by U54GM115428 from the National Institute of General Medical Sciences. Laura M. Raffield is supported by T32 HL129982.

KORA The KORA research platform (KORA, Cooperative Health Research in the Region of Augsburg) was initiated and financed by the Helmholtz Zentrum München - German Research Center for

Environmental Health, which is funded by the German Federal Ministry of Education and Research and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ. Data analysis was funded by the Else Kröner-Fresenius Stiftung (2012_A147) and by the Deutsche Forschungsgemeinschaft with grant DFG BO 3815/1-4. The work of IMH was supported by DFG CRC 1350/C6 and DFG BO 3815/4-1.

LIFE-Adult, LIFE-Heart, LIFE-Child LIFE-Adult, LIFE-Heart and LIFE-Child are funded by the Leipzig Research Center for Civilization Diseases (LIFE). LIFE is an organizational unit affiliated to the Medical Faculty of the University of Leipzig. LIFE is funded by means of the European Union, by the European Regional Development Fund (ERDF) and by funds of the Free State of Saxony within the framework of the excellence initiative.

Lifelines The Lifelines Cohort Study, and generation and management of GWAS genotype data for the Lifelines Cohort Study is supported by the Netherlands Organization of Scientific Research NWO (grant 175.010.2007.006), the Economic Structure Enhancing Fund (FES) of the Dutch government, the Ministry of Economic Affairs, the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the Northern Netherlands Collaboration of Provinces (SNN), the Province of Groningen, University Medical Center Groningen, the University of Groningen, Dutch Kidney Foundation and Dutch Diabetes Research Foundation. The authors wish to acknowledge the services of the Lifelines Cohort Study, the contributing research centers delivering data to Lifelines, and all the study participants.

LOLIPOP The LOLIPOP study is supported by the National Institute for Health Research (NIHR) Comprehensive Biomedical Research Centre Imperial College Healthcare NHS Trust, the British Heart Foundation (SP/04/002), the Medical Research Council (G0601966, G0700931), the Wellcome Trust (084723/Z/08/Z, 090532 & 098381) the NIHR (RP-PG-0407-10371), the NIHR Official Development Assistance (ODA, award 16/136/68), the European Union FP7 (EpiMigrant, 279143) and H2020 programs (iHealth-T2D, 643774). We acknowledge support of the MRC-PHE Centre for Environment and Health, and the NIHR Health Protection Research Unit on Health Impact of Environmental Hazards. The work was carried out in part at the NIHR/Wellcome Trust Imperial Clinical Research Facility. The views expressed are those of the author(s) and not necessarily those of the Imperial College Healthcare NHS Trust, the NHS, the NIHR or the Department of Health. We thank the participants and research staff who made the study possible. JC is supported by the Singapore Ministry of Health's National Medical Research Council under its Singapore Translational Research Investigator (STaR) Award (NMRC/STaR/0028/2017).

LURIC LURIC was supported by the 7th Framework Program of the EU (AtheroRemo, grant agreement number 201668 and RiskyCAD, grant agreement number 305739). The work of W.M. and M.E.K. is supported as part of the Competence Cluster of Nutrition and Cardiovascular Health (nutriCARD) which is funded by the German Ministry of Educatio and Research (grant agreement number 01EA1411A). The authors thank the LURIC study team who were involved in patient recruitment as well as sample and data handling, in addition to the laboratory staff at the Ludwigshafen General

Hospital and the universities of Freiburg, Ulm and Heidelberg, Germany.

MICROS We owe a debt of gratitude to all participants, all primary care practitioners, and the personnel of the Hospital of Silandro (Department of Laboratory Medicine) for their participation and collaboration in the research project. The study was supported by the Ministry of Health and Department of Educational Assistance, University and Research of the Autonomous Province of Bolzano, the South Tyrolean Sparkasse Foundation, and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947).

MVP This work was supported by the MVP-VA Award #I01BX003360 (PI A.M.H), and Adriana M. Hung (A.M.H) was supported by VA Award #I01BX003360. Additional MVP acknowledgements are included at the end of this section.

MyCode (Geisinger) We acknowledge the participants, staff, and our colleagues associated with the Geisinger MyCode Community Health Initiative. We also thank the staff of the PACDC of Geisinger for assistance with the phenotypic data, and the staff of the Biomedical & Translational Informatics and Kidney Health Research Institute.

NEO The authors of the NEO study thank all individuals who participated in the Netherlands Epidemiology in Obesity study, all participating general practitioners for inviting eligible participants and all research nurses for collection of the data. We thank the NEO study group, Pat van Beelen, Petra Noordijk and Ingeborg de Jonge for the coordination, lab and data management of the NEO study. The genotyping in the NEO study was supported by the Centre National de Génotypage (Paris, France), headed by Jean-Francois Deleuze. The NEO study is supported by the participating Departments, the Division and the Board of Directors of the Leiden University Medical Center, and by the Leiden University, Research Profile Area Vascular and Regenerative Medicine. Dennis Mook-Kanamori is supported by Dutch Science Organization (ZonMW-VENI Grant 916.14.023).

NESDA Funding was obtained from the Netherlands Organization for Scientific Research (Geestkracht program grant 10-000-1002); the Center for Medical Systems Biology (CSMB, NWO Genomics), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL), VU University's Institutes for Health and Care Research (EMGO+) and Neuroscience Campus Amsterdam, University Medical Center Groningen, Leiden University Medical Center, National Institutes of Health (NIH, R01D0042157-01A, MH081802, Grand Opportunity grants 1RC2 MH089951 and 1RC2 MH089995). Part of the genotyping and analyses were funded by the Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health.Computing was supported by BiG Grid, the Dutch e-Science Grid, which is financially supported by NWO.

OGP The study was supported by grant from the Italian Ministry of Education, University and Research (MIUR) n°: 5571/DSPAR/2002. We express our gratitude to all the study participants for their contributions and to the municipal administrations for their economic and logistic support.

ORCADES The Orkney Complex Disease Study (ORCADES) was supported by the Chief Scientist Office of the Scottish Government (CZB/4/276, CZB/4/710), a Royal Society URF to J.F.W., the MRC Human

Genetics Unit quinquennial programme “QTL in Health and Disease”, Arthritis Research UK and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh.

PREVEND The Prevention of Renal and Vascular Endstage Disease Study (PREVEND) genetics is supported by the Dutch Kidney Foundation (Grant E033), the EU project grant GENECURE (FP-6 LSHM CT 2006 037697), the National Institutes of Health (grant LM010098), the Netherlands organisation for health research and development (NWO VENI grant 916.761.70), and the Dutch Inter University Cardiology Institute Netherlands (ICIN). Niek Verweij was supported by NWO VENI grant 016.186.125.

QIMR adolescent We acknowledge funding by Australian National Health and Medical Research Council (NHMRC) grants 241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 552485, 552498 and Australian Research Council grants A7960034, A79906588, A79801419, DP0770096, DP0212016, DP0343921 for building and maintaining the adolescent twin family resource through which samples were collected.

QIMR adult We acknowledge the contributions of many staff in the Genetic Epidemiology Unit, Queensland Institute of Medical Research, in interviewing study participants, sample processing and DNA extraction, and data management. Funding for aspects of this work was provided by the Australian National Health and Medical Research Council (241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 552485, 552498), the Australian Research Council (A7960034, A79906588, A79801419, DP0770096, DP0212016, DP0343921), the EU 5th Framework Programme GenomEUtwin Project (QLG2-CT-2002-01254), and the U.S. National Institutes of Health (AA07535, AA10248, AA11998, AA13320, AA13321, AA13326, AA14041, AA17688, DA12854, MH66206). G.W.M. was supported by National Health and Medical Research Council (NHMRC) Fellowship Schemes.

RS The Rotterdam Study has been approved by the Medical Ethics Committee of the Erasmus MC (registration number MEC 02.1015) and by the Dutch Ministry of Health, Welfare and Sport (Population Screening Act WBO, license number 1071272-159521-PG). The Rotterdam Study has been entered into the Netherlands National Trial Register (NTR; www.trialregister.nl) and into the WHO International Clinical Trials Registry Platform (ICTRP; www.who.int/ictrp/network/primary/en/) under shared catalogue number NTR6831. All participants provided written informed consent to participate in the study and to have their information obtained from treating physicians. The generation and management of GWAS genotype data for the Rotterdam Study (RS I, RS II, RS III) was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands. The GWAS datasets are supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012), the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA),

project nr. 050-060-810. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters, MSc, and Carolina Medina-Gomez, MSc, for their help in creating the GWAS database, and Karol Estrada, PhD, Yurii Aulchenko, PhD, and Carolina Medina-Gomez, MSc, for the creation and analysis of imputed data. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists.

SHIP/SHIP-Trend SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network 'Greifswald Approach to Individualized Medicine (GANI_MED)' funded by the Federal Ministry of Education and Research (grant 03IS2061A). Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthineers, Erlangen, Germany and the Federal State of Mecklenburg- West Pomerania. The University of Greifswald is a member of the Caché Campus program of the InterSystems GmbH. Variant annotation was supported by software resources provided via the Caché Campus program of the InterSystems GmbH to Alexander Teumer.

SKIPOGH The SKIPOGH study is supported by a grant from the Swiss national science foundation (FN33CM30-124087).

Sorbs This work was supported by grants from the German Research Foundation (Project number 209933838; SFB-1052 ''Obesity mechanisms'' A01, B03, SPP 1629 TO 718/2- 1), from the German Diabetes Association, from the DHFD (Diabetes Hilfs- und Forschungsfonds Deutschland) and from IFB Adiposity Diseases (AD2-060E, AD2-06E95, AD2-06E99). IFB Adiposity Diseases is supported by the Federal Ministry of Education and Research (BMBF), Germany (FKZ: 01EO1501)

UKBB This research has been conducted using the UK Biobank Resource under Application Numbers 20272, 7439 and 19655.

ULSAM JÄ was supported by The Swedish Research Council, Swedish Heart-Lung Foundation, Dalarna University and Uppsala University. The ULSAM study was supported by Wellcome Trust grants WT098017, WT064890, WT090532, Uppsala University, Uppsala University Hospital, the Swedish Research Council and the Swedish Heart-Lung Foundation.

Vanderbilt The dataset(s) used for the analyses described were obtained from Vanderbilt University Medical Center's BioVU which is supported by numerous sources: institutional funding, private agencies, and federal grants. These include the NIH funded Shared Instrumentation Grant S10RR025141; and CTSA grants UL1TR002243, UL1TR000445, and UL1RR024975. Genomic data are also supported by investigator-led projects that include U01HG004798, R01NS032830,

RC2GM092618, P50GM115305, U01HG006378, U19HL065962, R01HD074711; and additional funding sources listed at https://victr.vanderbilt.edu/pub/biovu/.

VIKING The Viking Health Study – Shetland (VIKING) was supported by the MRC Human Genetics Unit quinquennial programme grant "QTL in Health and Disease". DNA extractions and genotyping were performed at the Edinburgh Clinical Research Facility, University of Edinburgh. We would like to acknowledge the invaluable contributions of the research nurses in Shetland, the administrative team in Edinburgh and the people of Shetland.

YFS The Young Finns Study has been financially supported by the Academy of Finland: grants 286284, 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Competitive State Research Financing of the Expert Responsibility area of Kuopio, Tampere and Turku University Hospitals (grant X51001); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation for Cardiovascular Research ; Finnish Cultural Foundation; The Sigrid Juselius Foundation; Tampere Tuberculosis Foundation; Emil Aaltonen Foundation; Yrjö Jahnsson Foundation; Signe and Ane Gyllenberg Foundation; Diabetes Research Foundation of Finnish Diabetes Association; and EU Horizon 2020 (grant 755320 for TAXINOMISIS); and European Research Council (grant 742927 for MULTIEPIGEN project). We thank the teams that collected data at all measurement time points; the persons who participated as both children and adults in these longitudinal studies; and biostatisticians Irina Lisinen, Johanna Ikonen, Noora Kartiosuo, Ville Aalto, and Jarno Kankaanranta for data management and statistical advice.

Experimental study Functional experiments in the Woodward laboratory were supported by grants from the NIDDK: (A) R01DK114091, and (B) the Baltimore PKD Research and Clinical Core Center 5P30DK090868. We also want to thanks Douglas D. Ross of the University of Maryland for the generous gift of the *ABCG2* promoter luciferase construct.

Groups with Consortia Authorship

German Chronic Kidney Disease (GCKD) – Author Group

University of Erlangen-Nürnberg: Kai-Uwe Eckardt, Heike Meiselbach, Markus P. Schneider, Mario Schiffer, Thomas Dienemann, Hans-Ulrich Prokosch, Barbara Bärthlein, Andreas Beck, Detlef Kraska, André Reis, Arif B. Ekici, Susanne Avendaño, Dinah Becker-Grosspitsch, Ulrike Alberth-Schmidt, Birgit Hausknecht, Anke Weigel;
University of Freiburg: Gerd Walz, Anna Köttgen, Ulla T. Schultheiß, Fruzsina Kotsis, Simone Meder, Erna Mitsch, Ursula Reinhard;
RWTH Aachen University: Jürgen Floege, Georg Schlieper, Turgay Saritas;
Charité, University Medicine Berlin: Elke Schaeffner, Seema Baid-Agrawal, Kerstin Theisen;
Hannover Medical School: Hermann Haller, Jan Menne;
University of Heidelberg: Martin Zeier, Claudia Sommerer, Rebecca Woitke;
University of Jena: Gunter Wolf, Martin Busch, Rainer Paul;
Ludwig-Maximilians University of München: Thomas Sitter;
University of Würzburg: Christoph Wanner, Vera Krane, Antje Börner-Klein, Britta Bauer;
Medical University of Innsbruck, Division of Genetic Epidemiology: Florian Kronenberg, Julia Raschenberger, Barbara Kollerits, Lukas Forer, Sebastian Schönherr, Hansi Weissensteiner;
University of Regensburg, Institute of Functional Genomics: Peter Oefner, Wolfram Gronwald, Helena Zacharias;
Department of Medical Biometry, Informatics and Epidemiology (IMBIE), University Hospital of Bonn: Matthias Schmid, Jennifer Nadal.

LifeLines Cohort Study – Author Group

Behrooz Z Alizadeh (1), H Marike Boezen (1), Lude Franke (2), Pim van der Harst (3), Gerjan Navis (4), Marianne Rots (5), Morris Swertz (2), Bruce HR Wolffenbuttel (6), Cisca Wijmenga (2)
(1) Department of Epidemiology, University of Groningen, University Medical Center Groningen, The Netherlands
(2) Department of Genetics, University of Groningen, University Medical Center Groningen, The Netherlands
(3) Department of Cardiology, University of Groningen, University Medical Center Groningen, The Netherlands
(4) Department of Internal Medicine, Division of Nephrology, University of Groningen, University Medical Center Groningen, The Netherlands
(5) Department of Pathology and Medical Biology, University of Groningen, University Medical Center Groningen, The Netherlands
(6) Department of Endocrinology, University of Groningen, University Medical Center Groningen, The Netherlands

Million Veteran Program: Consortium Acknowledgement

MVP Executive Committee
- Co-Chair: J. Michael Gaziano, M.D., M.P.H.

- Co-Chair: Rachel Ramoni, D.M.D., Sc.D.
- Jim Breeling, M.D. (ex-officio)
- Kyong-Mi Chang, M.D.
- Grant Huang, Ph.D.
- Sumitra Muralidhar, Ph.D.
- Christopher J. O'Donnell, M.D., M.P.H.
- Philip S. Tsao, Ph.D.

MVP Program Office
- Sumitra Muralidhar, Ph.D.
- Jennifer Moser, Ph.D.

MVP Recruitment/Enrollment
- Recruitment/Enrollment Director/Deputy Director, Boston – Stacey B. Whitbourne, Ph.D.; Jessica V. Brewer, M.P.H.
- MVP Coordinating Centers
    - o Clinical Epidemiology Research Center (CERC), West Haven – John Concato, M.D., M.P.H.
    - o Cooperative Studies Program Clinical Research Pharmacy Coordinating Center, Albuquerque - Stuart Warren, J.D., Pharm D.; Dean P. Argyres, M.S.
    - o Genomics Coordinating Center, Palo Alto – Philip S. Tsao, Ph.D.
    - o Massachusetts Veterans Epidemiology Research Information Center (MAVERIC), Boston - J. Michael Gaziano, M.D., M.P.H.
    - o MVP Information Center, Canandaigua – Brady Stephens, M.S.
- Core Biorepository, Boston – Mary T. Brophy M.D., M.P.H.; Donald E. Humphries, Ph.D.
- MVP Informatics, Boston – Nhan Do, M.D.; Shahpoor Shayan
- Data Operations/Analytics, Boston – Xuan-Mai T. Nguyen, Ph.D.

MVP Science
- Genomics - Christopher J. O'Donnell, M.D., M.P.H.; Saiju Pyarajan Ph.D.; Philip S. Tsao, Ph.D.
- Phenomics - Kelly Cho, M.P.H, Ph.D.
- Data and Computational Sciences – Saiju Pyarajan, Ph.D.
- Statistical Genetics – Elizabeth Hauser, Ph.D.; Yan Sun, Ph.D.; Hongyu Zhao, Ph.D.

MVP Local Site Investigators
- Atlanta VA Medical Center (Peter Wilson)
- Bay Pines VA Healthcare System (Rachel McArdle)
- Birmingham VA Medical Center (Louis Dellitalia)
- Cincinnati VA Medical Center (John Harley)
- Clement J. Zablocki VA Medical Center (Jeffrey Whittle)
- Durham VA Medical Center (Jean Beckham)
- Edith Nourse Rogers Memorial Veterans Hospital (John Wells)
- Edward Hines, Jr. VA Medical Center (Salvador Gutierrez)
- Fayetteville VA Medical Center (Gretchen Gibson)
- VA Health Care Upstate New York (Laurence Kaminsky)
- New Mexico VA Health Care System (Gerardo Villareal)
- VA Boston Healthcare System (Scott Kinlay)
- VA Western New York Healthcare System (Junzhe Xu)
- Ralph H. Johnson VA Medical Center (Mark Hamner)
- Wm. Jennings Bryan Dorn VA Medical Center (Kathlyn Sue Haddock)
- VA North Texas Health Care System (Sujata Bhushan)
- Hampton VA Medical Center (Pran Iruvanti)
- Hunter Holmes McGuire VA Medical Center (Michael Godschalk)
- Iowa City VA Health Care System (Zuhair Ballas)

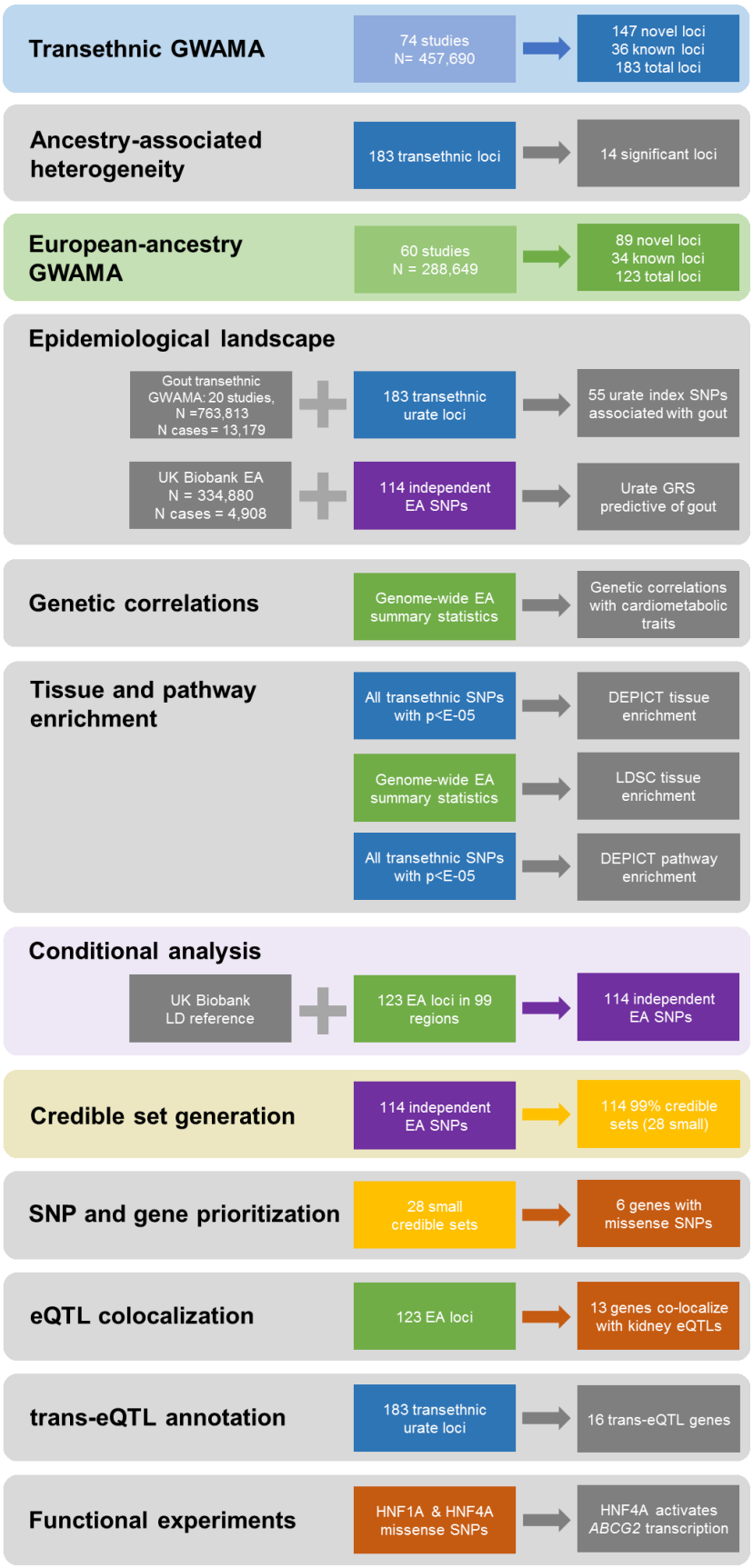
- Jack C. Montgomery VA Medical Center (Malcolm Buford)
- James A. Haley Veterans' Hospital (Stephen Mastorides)
- Louisville VA Medical Center (Jon Klein)
- Manchester VA Medical Center (Nora Ratcliffe)
- Miami VA Health Care System (Hermes Florez)
- Michael E. DeBakey VA Medical Center (Alan Swann)
- Minneapolis VA Health Care System (Maureen Murdoch)
- N. FL/S. GA Veterans Health System (Peruvemba Sriram)
- Northport VA Medical Center (Shing Shing Yeh)
- Overton Brooks VA Medical Center (Ronald Washburn)
- Philadelphia VA Medical Center (Darshana Jhala)
- Phoenix VA Health Care System (Samuel Aguayo)
- Portland VA Medical Center (David Cohen)
- Providence VA Medical Center (Satish Sharma)
- Richard Roudebush VA Medical Center (John Callaghan)
- Salem VA Medical Center (Kris Ann Oursler)
- San Francisco VA Health Care System (Mary Whooley)
- South Texas Veterans Health Care System (Sunil Ahuja)
- Southeast Louisiana Veterans Health Care System (Amparo Gutierrez)

Current as of 19-JUL-2018

- Southern Arizona VA Health Care System (Ronald Schifman)
- Sioux Falls VA Health Care System (Jennifer Greco)
- St. Louis VA Health Care System (Michael Rauchman)
- Syracuse VA Medical Center (Richard Servatius)
- VA Eastern Kansas Health Care System (Mary Oehlert)
- VA Greater Los Angeles Health Care System (Agnes Wallbom)
- VA Loma Linda Healthcare System (Ronald Fernando)
- VA Long Beach Healthcare System (Timothy Morgan)
- VA Maine Healthcare System (Todd Stapley)
- VA New York Harbor Healthcare System (Scott Sherman)
- VA Pacific Islands Health Care System (Gwenevere Anderson)
- VA Palo Alto Health Care System (Philip Tsao)
- VA Pittsburgh Health Care System (Elif Sonel)
- VA Puget Sound Health Care System (Edward Boyko)
- VA Salt Lake City Health Care System (Laurence Meyer)
- VA San Diego Healthcare System (Samir Gupta)
- VA Southern Nevada Healthcare System (Joseph Fayad)
- VA Tennessee Valley Healthcare System (Adriana Hung)
- Washington DC VA Medical Center (Jack Lichy)
- W.G. (Bill) Hefner VA Medical Center (Robin Hurley)
- White River Junction VA Medical Center (Brooks Robey)
- William S. Middleton Memorial Veterans Hospital (Robert Striker)

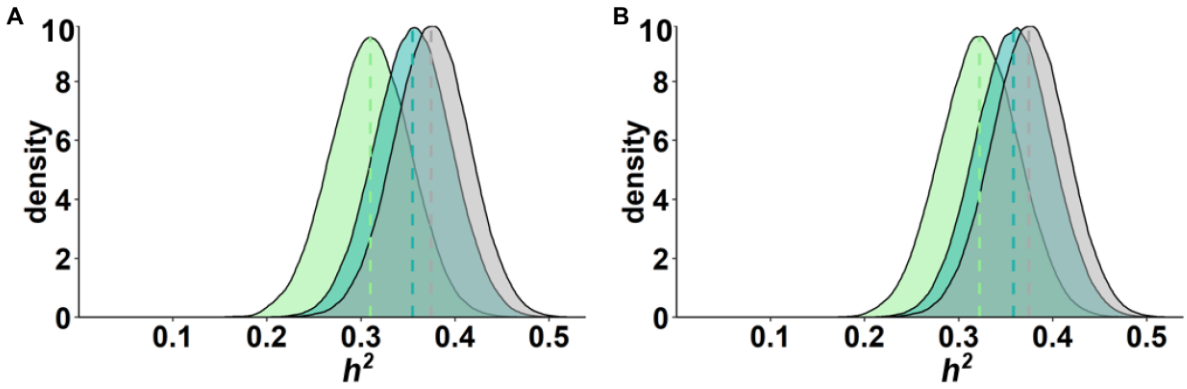
Supplementary Figure 1: Workflow of the project

Overview of the analysis workflow from trans-ethnic meta-analysis, gout risk prediction, genetic correlation, enrichment analyses, and statistical fine-mapping to functional study.



Supplementary Figure 2: Density plots of the heritability of serum urate estimated from the Cooperative Health Research in South Tyrol (CHRIS) study

Heritability of serum urate estimated from the Cooperative Health Research in South Tyrol (CHRIS) study (n=4,373 individuals in 186 pedigrees) using index SNPs from the trans-ethnic meta-analysis (A) and from the European-ancestry meta-analysis (B). Gray: overall heritability; blue: heritability after removing index SNPs in *SLC2A9*, *ABCG2*, and *SLC22A12*; green: heritability after removing all index SNPs in the present study.

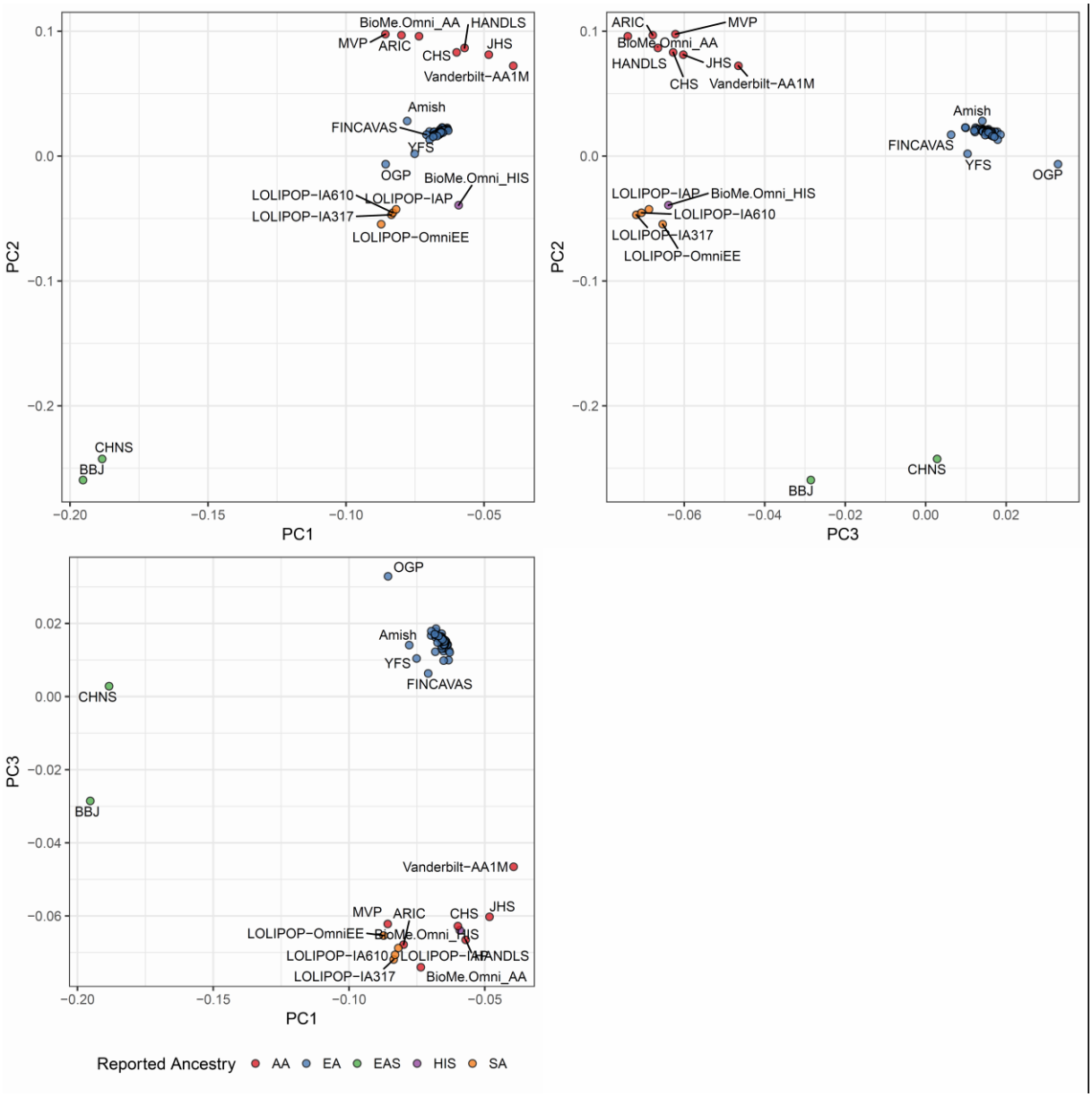


h2 (95% credible interval)	
Overall:	0.37 (0.29-0.45)
Without SNPs in 3 major loci:	0.35 (0.28-0.43)
Without all index SNPs:	0.31 (0.23-0.39)

h2 (95% credible interval)	
Overall:	0.37 (0.30-0.45)
Without SNPs in 3 major loci:	0.36 (0.28-0.44)
Without all index SNPs:	0.32 (0.24-0.40)

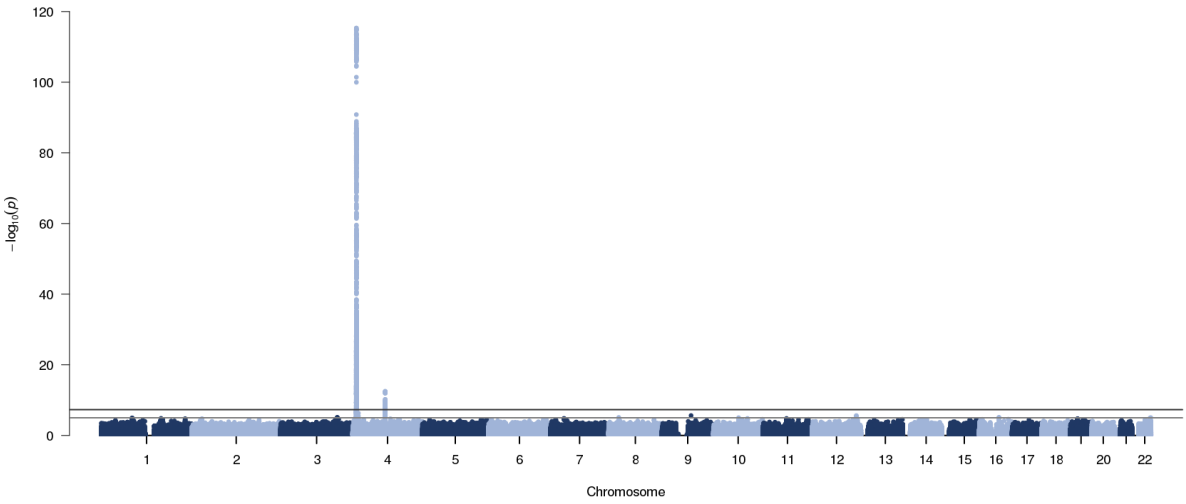
Supplementary Figure 3: Principal components generated using the mean allele frequency difference between studies.

Principal components (PCs) are generated using MR-MEGA from a matrix of mean pairwise allele frequency differences between cohorts (total n=457,690). Color denotes self-reported ancestry for each cohort. Selected outliers are labeled with cohort name. Three PCs were chosen as per author recommendations, and, as shown, are sufficient to separate self-reported ancestry groups.



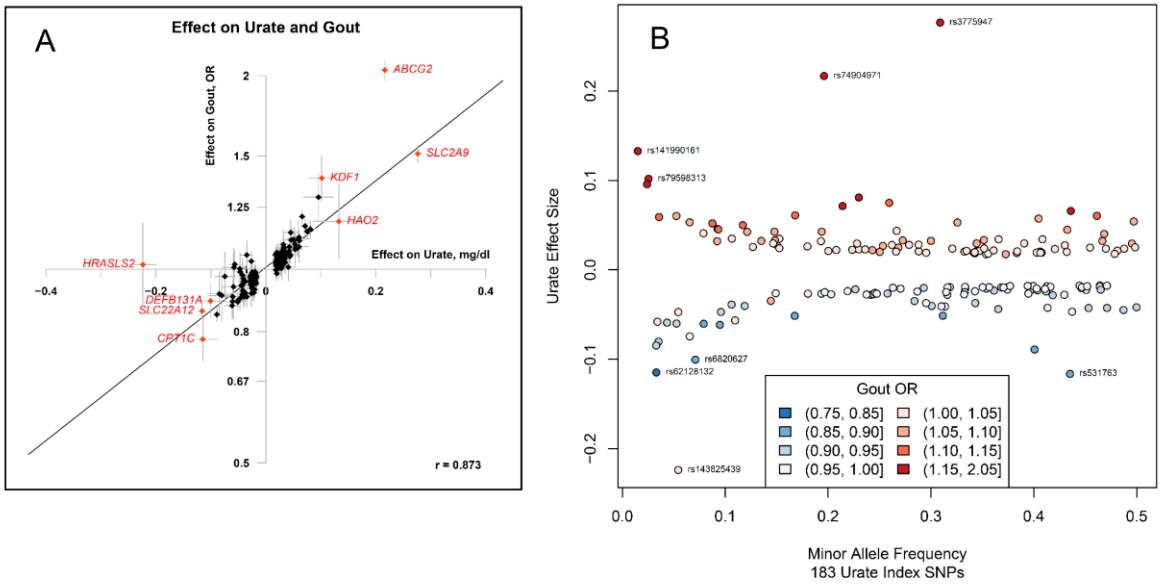
Supplementary Figure 4: Manhattan plot of the *P*-values of the differences in genetic effects on serum urate from trans-ethnic, sex-stratified meta-analyses.

GWAS of 174,111 men and 172,102 women were run separately and the effect sizes compared. *P*-values were obtained from a two-sample two-sided t-test for significant difference of association between men and women using the formula $t = \frac{\beta_M - \beta_F}{\sqrt{SE_M{}^2 + SE_F{}^2}}$, where β_M and β_F were beta coefficients among males and females, respectively, and SE_M and SE_F were the corresponding standard errors.



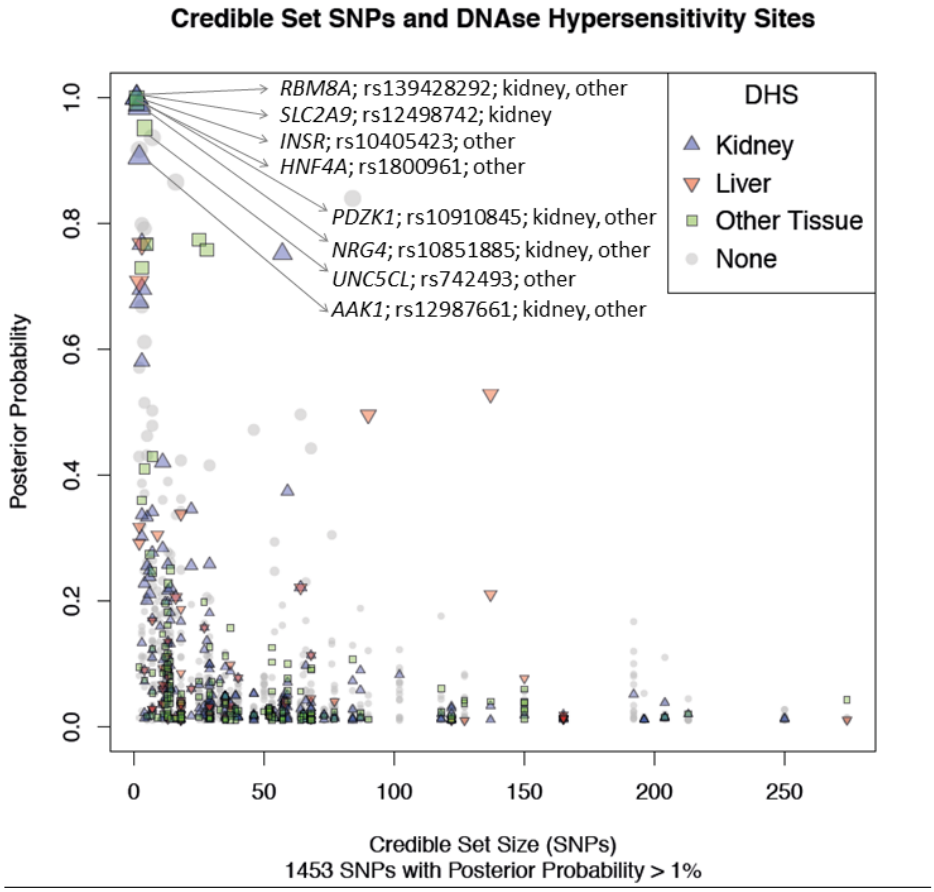
Supplementary Figure 5: Effect size of serum urate loci relating to gout (A) and minor allele frequency (B).

The effect sizes for serum urate and gout are represented by the index SNPs from the trans-ethnic meta-analysis of serum urate (n=457,690) and gout (n=763,813), respectively. The whiskers in (A) represent 95% confidence intervals and r is the Pearson correlation coefficient.



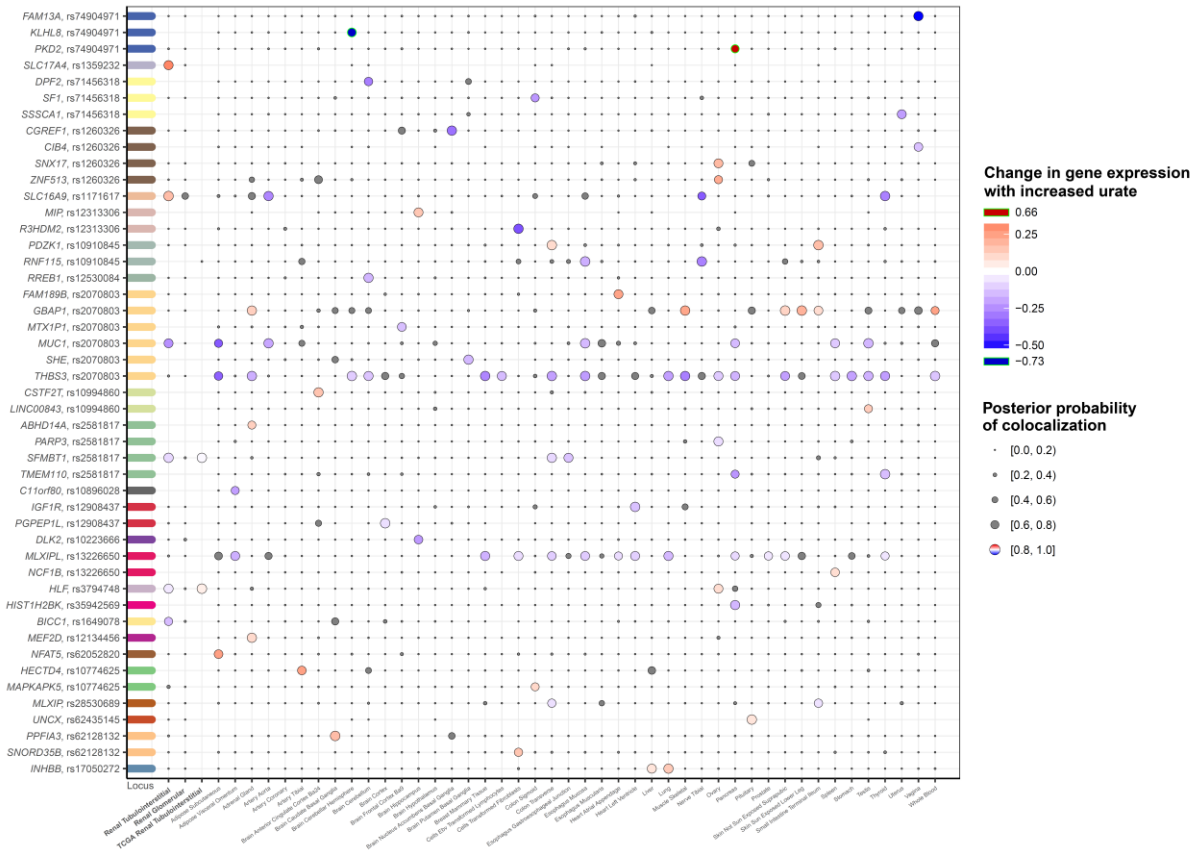
Supplementary Figure 6: Posterior probability of SNPs in 99% credible set with DNAse I hypersensitivity site annotation

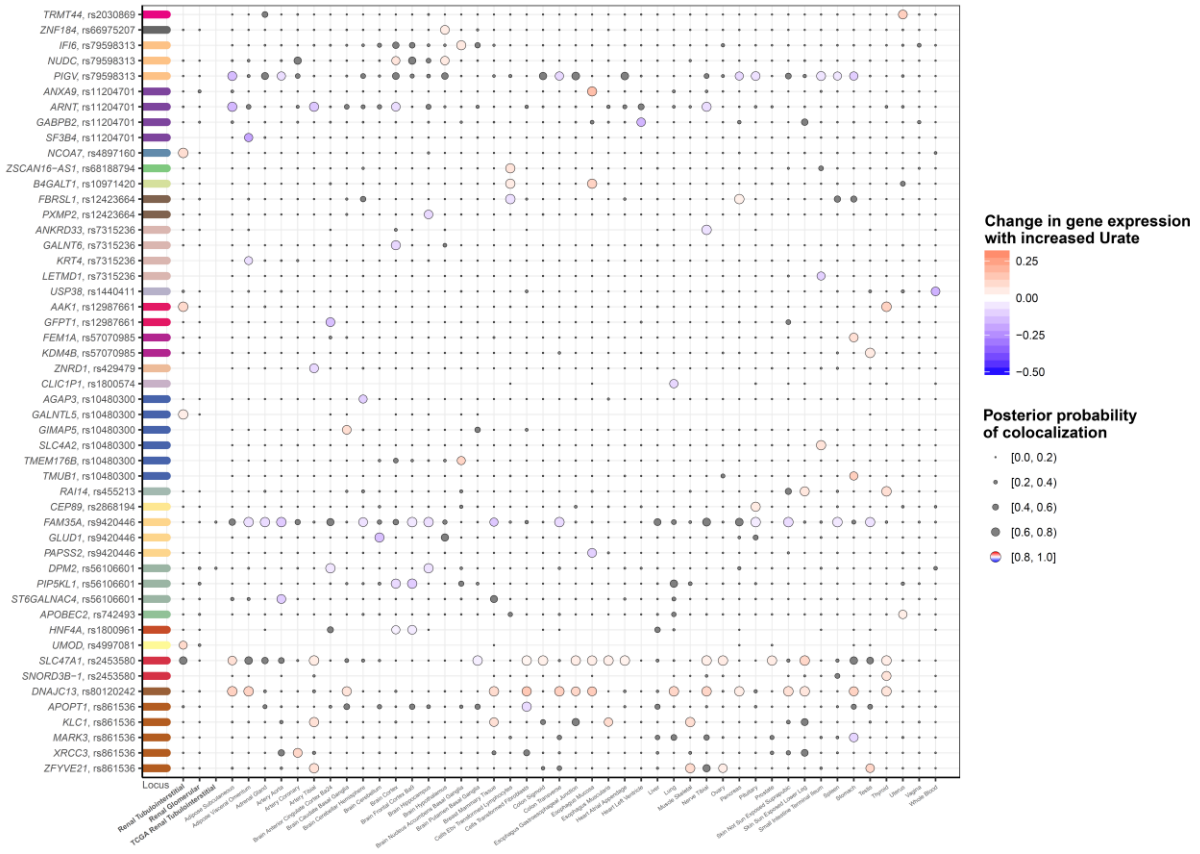
Graph shows credible set size (X-axis) against the posterior probability of association (PPA; Y-axis) for each of 1,453 SNPs with PPA >1% in 114 99% credible sets. SNPs mapping into ENCODE or RoadMap DNAse I hypersensitivity sites are highlighted.

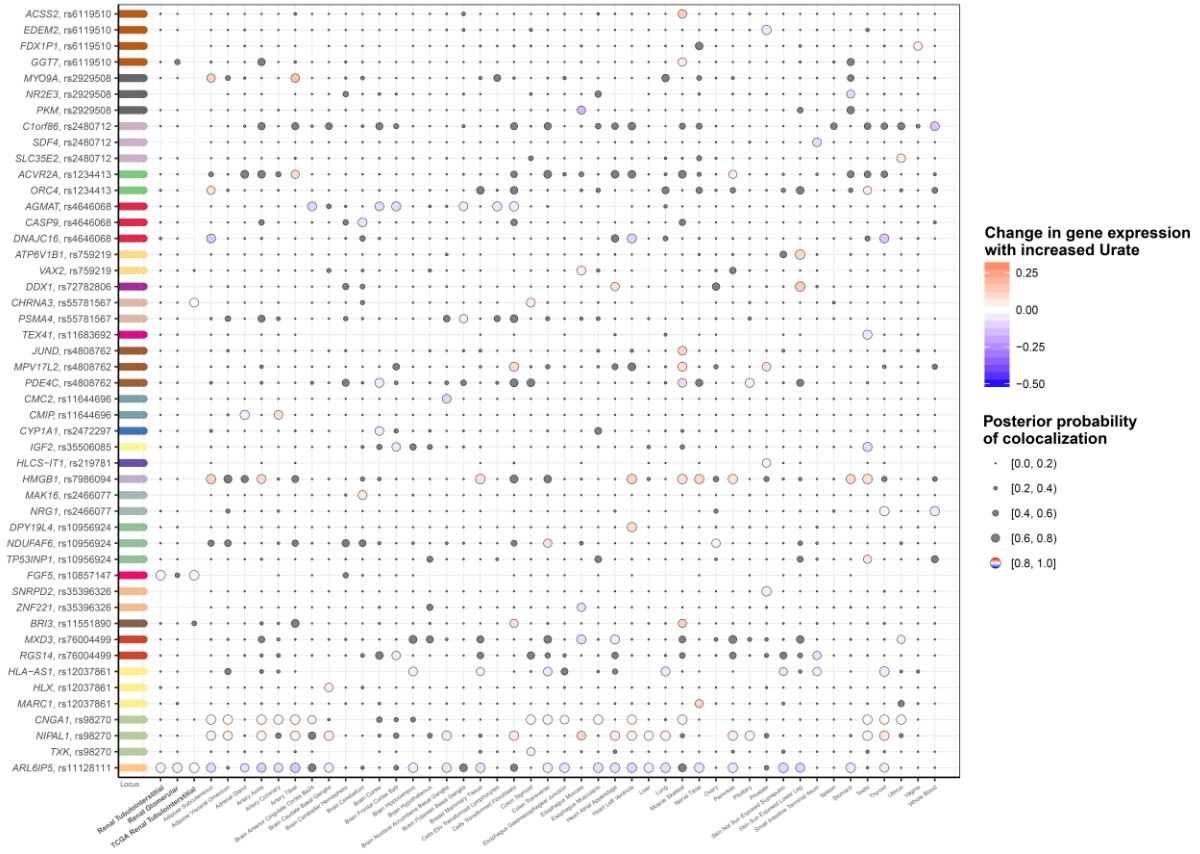


Supplementary Figure 7: Co-localization of urate-association signals with gene expression in *cis* across 47 tissues

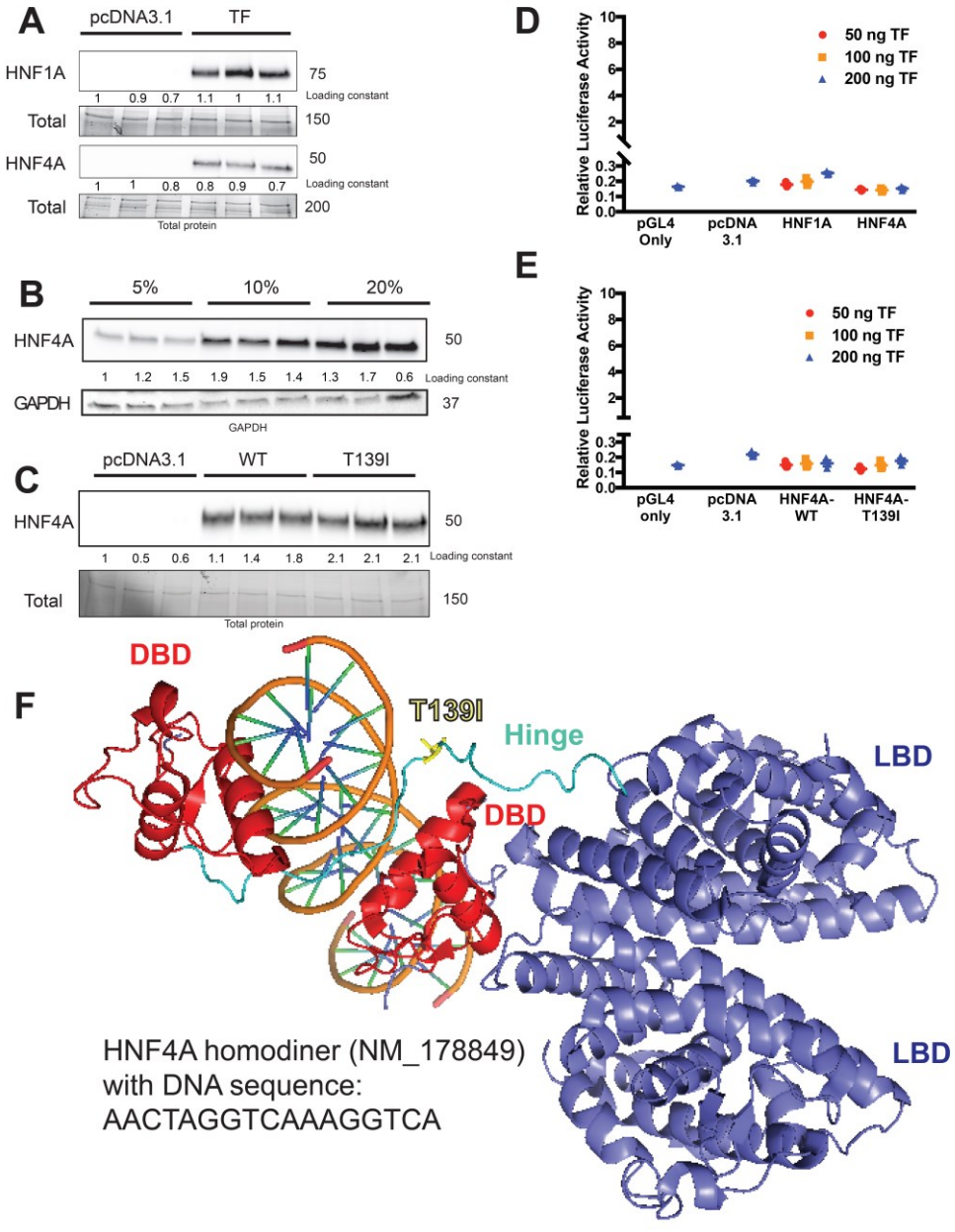
The plot shows genes whose expression levels in any tested tissues co-localized with the urate association signal with a posterior probability of ≥80%. Genes and their respective index SNP are shown on the Y-axis. The color adjacent to the SNP identifier on the Y-axis groups co-localization signals from one locus and helps to rapidly identify loci in which the urate association signal co-localizes with the expression of more than one gene. Co-localization across tissues (X-axis) is illustrated as dots, where the size of the dots indicates the posterior probability of the co-localization. The change in urate levels is color-coded relative to the change in gene expression with a color gradient as indicated in the legend, or gray in case of a posterior probability <0.80. Kidney tissues are labeled in bold.




TRMT44, rs2030869
ZNF184, rs66975207
IFI6, rs79598313
NUDC, rs79598313
PIGV, rs79598313
ANXA9, rs11204701
ARNT, rs11204701
GABPB2, rs11204701
SF3B4, rs11204701
NCOA7, rs4897160
ZSCAN16−AS1, rs68188794
B4GALT1, rs10971420
FBRSL1, rs12423664
PXMP2, rs12423664
ANKRD33, rs7315236
GALNT6, rs7315236
KRT4, rs7315236
LETMD1, rs7315236
USP38, rs1440411
AAK1, rs12987661
GFPT1, rs12987661
FEM1A, rs57070985
KDM4B, rs57070985
ZNRD1, rs429479
CLIC1P1, rs1800574
AGAP3, rs10480300
GALNTL5, rs10480300
GIMAP5, rs10480300
SLC4A2, rs10480300
TMEM176B, rs10480300
TMUB1, rs10480300
RAI14, rs455213
CEP89, rs2868194
FAM35A, rs9420446
GLUD1, rs9420446
PAPSS2, rs9420446
DPM2, rs56106601
PIP5KL1, rs56106601
ST6GALNAC4, rs56106601
APOBEC2, rs742493
HNF4A, rs1800961
UMOD, rs4997081
SLC47A1, rs2453580
SNORD3B−1, rs2453580
DNAJC13, rs80120242
APOPT1, rs861536
KLC1, rs861536
MARK3, rs861536
XRCC3, rs861536
ZFYVE21, rs861536
Locus
Renal Tubulointerstitial
Renal Glomerular
TCGA Renal Tubulointerstitial
Change in gene expression with increased Urate
0.25
0.00
−0.25
−0.50
Posterior probability of colocalization
[0.0, 0.2)
[0.2, 0.4)
[0.4, 0.6)
[0.6, 0.8)
[0.8, 1.0]


ACSS2, rs6119510
EDEM2, rs6119510
FDX1P1, rs6119510
GGT7, rs6119510
MYO9A, rs2929508
NR2E3, rs2929508
PKM, rs2929508
C1orf86, rs2480712
SDF4, rs2480712
SLC35E2, rs2480712
ACVR2A, rs1234413
ORC4, rs1234413
AGMAT, rs4646068
CASP9, rs4646068
DNAJC16, rs4646068
ATP6V1B1, rs759219
VAX2, rs759219
DDX1, rs72782806
CHRNA3, rs55781567
PSMA4, rs55781567
TEX41, rs11683692
JUND, rs4808762
MPV17L2, rs4808762
PDE4C, rs4808762
CMC2, rs11644696
CMIP, rs11644696
CYP1A1, rs2472297
IGF2, rs35506085
HLCS−IT1, rs219781
HMGB1, rs7986094
MAK16, rs2466077
NRG1, rs2466077
DPY19L4, rs10956924
NDUFAF6, rs10956924
TP53INP1, rs10956924
FGF5, rs10857147
SNRPD2, rs35396326
ZNF221, rs35396326
BRI3, rs11551890
MXD3, rs76004499
RGS14, rs76004499
HLA−AS1, rs12037861
HLX, rs12037861
MARC1, rs12037861
CNGA1, rs98270
NIPAL1, rs98270
TXK, rs98270
ARL6IP5, rs11128111
Locus
Renal Tubulointerstitial
Renal Glomerular
TCGA Renal Tubulointerstitial
Change in gene expression with increased Urate
0.25
0.00
−0.25
−0.50
Posterior probability of colocalization
[0.0, 0.2)
[0.2, 0.4)
[0.4, 0.6)
[0.6, 0.8)
[0.8, 1.0]

Supplementary Figure 8: Mechanics and mechanisms of *ABCG2* promoter luciferase assay (**A**) Expression of HNF1A and HNF4A in HEK 293 cells used for the luciferase *ABCG2* promoter activation assay demonstrating the successful transfection of the assay cells (n=3 for each treatment group). (**B**) Western blot of HNF4A showing that the resulting dose progression of HNF4A protein reflects the increasing dose of DNA used in the transfection of the HEK 293 cells; with GAPDH as loading control (from the same samples used in Figure 5C; n=3 for each treatment group). (**C**) The mutant HNF4A, T139I, expressed at comparable levels to wild-type in HEK 293 cells (from the same samples used in Figure 5E; n=3 for each treatment group). (**D & E**) Control experiments for the luciferase assay in Figure 5C and E: HEK293 cells transfected with an empty vector show no increase in luciferase activity (± SEM) from the same n=3 independent experiments as in Figure 5C & 5E. (**F**) Entire homodimer HNF4A structural model partially depicted in Figure 5D): location of mutation and relevant domains are identified (PBD 4IQR).



References

1. Nakayama, A. *et al.* GWAS of clinically defined gout and subtypes identifies multiple susceptibility loci that include urate transporter genes. *Ann Rheum Dis* (2016).
2. Bardin, T. *et al.* Lesinurad in combination with allopurinol: a randomised, double-blind, placebo-controlled study in patients with gout with inadequate response to standard of care (the multinational CLEAR 2 study). *Ann Rheum Dis* **76**, 811-820 (2017).
3. Zhang, Y., Ma, K.L., Ruan, X.Z. & Liu, B.C. Dysregulation of the Low-Density Lipoprotein Receptor Pathway Is Involved in Lipid Disorder-Mediated Organ Injury. *Int J Biol Sci* **12**, 569-79 (2016).
4. Schriml, L.M. *et al.* Human Disease Ontology 2018 update: classification, content and workflow expansion. *Nucleic Acids Res* **47**, D955-D962 (2019).
5. Li, Y., Willer, C., Sanna, S. & Abecasis, G. Genotype imputation. *Annu Rev Genomics Hum Genet* **10**, 387-406 (2009).
6. Delaneau, O., Marchini, J. & Zagury, J.F. A linear complexity phasing method for thousands of genomes. *Nat Methods* **9**, 179-81 (2011).
7. Loh, P.R., Palamara, P.F. & Price, A.L. Fast and accurate long-range phasing in a UK Biobank cohort. *Nat Genet* **48**, 811-6 (2016).
8. Browning, B.L. & Browning, S.R. A fast, powerful method for detecting identity by descent. *Am J Hum Genet* **88**, 173-82 (2011).
9. the Haplotype Reference, C. A reference panel of 64,976 haplotypes for genotype imputation. *Nature Genetics* **48**, 1279 (2016).
10. Abecasis, G.R. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56-65 (2012).
11. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* **5**, e1000529 (2009).
12. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat Genet* **48**, 1284-1287 (2016).
13. Durbin, R. Efficient haplotype matching and storage using the positional Burrows-Wheeler transform (PBWT). *Bioinformatics* **30**, 1266-72 (2014).
14. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* **48**, 1279-83 (2016).
15. Fuchsberger, C., Taliun, D., Pramstaller, P.P. & Pattaro, C. GWAtoolbox: an R package for fast quality control and handling of genome-wide association studies meta-analysis data. *Bioinformatics* **28**, 444-5 (2012).
16. Devlin, B., Roeder, K. & Wasserman, L. Genomic control, a new approach to genetic-based association studies. *Theor Popul Biol* **60**, 155-66 (2001).
17. Chang, C.C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
18. Frey, B.J. & Dueck, D. Clustering by passing messages between data points. *Science* **315**, 972-6 (2007).
19. Bodenhofer, U., Kothmeier, A. & Hochreiter, S. APCluster: an R package for affinity propagation clustering. *Bioinformatics* **27**, 2463-4 (2011).
20. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236-41 (2015).

21. Benner, C. *et al.* Prospects of Fine-Mapping Trait-Associated Genomic Regions by Using Summary Statistics from Genome-wide Association Studies. *Am J Hum Genet* **101**, 539-551 (2017).
22. Wuttke, M. *et al.* A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat Genet* **51**, 957-972 (2019).
23. Wakefield, J. A Bayesian measure of the probability of false discovery in genetic epidemiology studies. *Am J Hum Genet* **81**, 208-27 (2007).
24. Benner, C. *et al.* FINEMAP: efficient variable selection using summary data from genome-wide association studies. *Bioinformatics* **32**, 1493-501 (2016).
25. Joehanes, R. *et al.* Integrated genome-wide analysis of expression quantitative trait loci aids interpretation of genomic association studies. *Genome Biol* **18**, 16 (2017).
26. Westra, H.J. *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* **45**, 1238-1243 (2013).
27. Zeller, T. *et al.* Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. *PLoS One* **5**, e10693 (2010).
28. Fehrmann, R.S. *et al.* Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. *PLoS Genet* **7**, e1002197 (2011).
29. Ichida, K. *et al.* Decreased extra-renal urate excretion is a common cause of hyperuricemia. *Nat Commun* **3**, 764 (2012).
30. Qu, L.H., Jiang, H. & Chen, J.H. Effect of uric acid-lowering therapy on blood pressure: systematic review and meta-analysis. *Ann Med* **49**, 142-156 (2017).
31. Kirsten, H. *et al.* Dissecting the genetics of the human transcriptome identifies novel trait-related trans-eQTLs and corroborates the regulatory relevance of non-protein coding locidagger. *Hum Mol Genet* **24**, 4746-63 (2015).
32. Beutner, F. *et al.* Rationale and design of the Leipzig (LIFE) Heart Study: phenotyping and cardiovascular characteristics of patients with coronary artery disease. *PLoS One* **6**, e29070 (2011).
33. Loeffler, M. *et al.* The LIFE-Adult-Study: objectives and design of a population-based cohort study with 10,000 deeply phenotyped adults in Germany. *BMC Public Health* **15**, 691 (2015).
34. Peterson, C.B., Bogomolov, M., Benjamini, Y. & Sabatti, C. Many Phenotypes Without Many False Discoveries: Error Controlling Strategies for Multitrait Association Studies. *Genet Epidemiol* **40**, 45-56 (2016).
35. Yu, G., Wang, L.G., Yan, G.R. & He, Q.Y. DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics* **31**, 608-9 (2015).
36. Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res* **45**, D331-D338 (2017).
37. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25-9 (2000).
38. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**, D353-D361 (2017).
39. Fabregat, A. *et al.* The Reactome Pathway Knowledgebase. *Nucleic Acids Res* **46**, D649-D655 (2018).