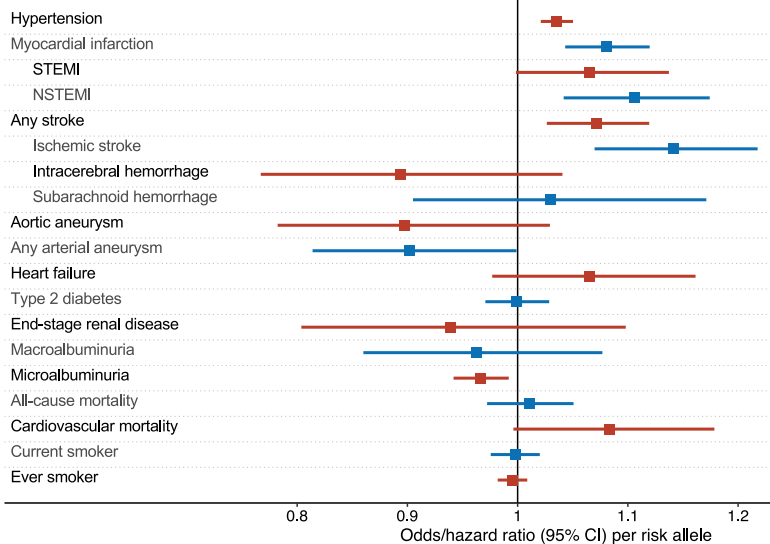
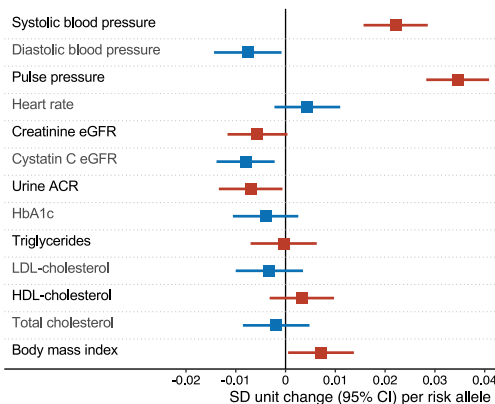


Article

# Atherosclerotic Aortic Calcification-Associated Polymorphism in *HDAC9* and Associations with Mortality, Cardiovascular Disease, and Kidney Disease

- UK Biobank, n = 335,146 middle-aged adults of European descent
- *HDAC9* variant rs57301765, chr7:19013110 (GRCh38.p12)
- Associations between abdominal aortic calcification-raising allele and vascular smooth muscle cell-driven outcomes



Johan Ärnlöv,  
Douglas F. Dluzen,  
Christoph Nowak

christoph.nowak@ki.se

**HIGHLIGHTS**

AAC-associated variation in *HDAC9* was associated with hypertension, STEMI, and stroke

A suggestive protective association with kidney function was not significant

*HDAC9* could be explored as a therapeutic target for arterial calcific disease

Ärnlöv et al., iScience 23, 101253  
July 24, 2020 © 2020 The Author(s).  
<https://doi.org/10.1016/j.isci.2020.101253>



## Article

Atherosclerotic Aortic Calcification-Associated Polymorphism in *HDAC9* and Associations with Mortality, Cardiovascular Disease, and Kidney DiseaseJohan Ärnlöv,<sup>1,2</sup> Douglas F. Dluzen,<sup>3</sup> and Christoph Nowak<sup>1,4,\*</sup>

## SUMMARY

**Histone deacetylase 9 (HDAC9) has recently been demonstrated as a key regulator of vascular smooth muscle cell (VSMC) phenotype and is associated with abdominal aortic calcification, myocardial infarction, and ischemic stroke. It is uncertain whether HDAC9 is also implicated in other VSMC-driven diseases. Our objective was to assess associations between abdominal aortic calcification-associated genetic variation in HDAC9 and VSMC-associated phenotypes. In this prospective population study of 335,146 adults enrolled in the UK Biobank, the abdominal aortic calcification-associated risk allele of a genetic variant in HDAC9 was associated with increased risk of systolic hypertension, non-ST segment elevation myocardial infarction, and ischemic stroke. There was a suggestive protective association with kidney disease outcomes that did not reach experiment-wise significance. These genetic results lend further support for HDAC9 as a potential therapeutic target for arterial stenotic and calcific disease.**

## INTRODUCTION

In a recent translational genomics study, Malhotra et al. (2019) identified the first genetic risk locus for abdominal aortic calcification (AAC) in *HDAC9* on chromosome 7, which encodes histone deacetylase 9 (HDAC9), a co-regulator of gene transcription. In genome-wide association study (GWAS) and *in vitro* studies, the authors demonstrated HDAC9 as a key regulator of vascular smooth muscle cell (VSMC) phenotype and calcification. Mice deficient in HDAC9 had reduced vascular calcification, suggesting HDAC9 as a possible drug target for vascular calcific disease (Malhotra et al., 2019). A separate study found that HDAC9 inhibition in mice reduced arterial neointimal formation and improved stenotic disease (Lino Cardenas et al., 2019). These findings suggest HDAC9 as a potential drug target for VSMC-associated cardiovascular disease.

Arterial calcification is a risk factor for cardiovascular disease and mortality (Bastos Goncalves et al., 2012; Criqui et al., 2014) and is associated with hypertension, chronic kidney disease, and end-stage renal disease (Vervloet and Cozzolino, 2017). The results obtained by Malhotra et al. (2019) demonstrate HDAC9's role in altering VSMC phenotype, and associations with myocardial infarction, ischemic stroke, and pulse pressure have been reported (Traylor et al., 2012; Bellenguez et al., 2012; Malik et al., 2016, 2017; Nikpay et al., 2015; Kato et al., 2015). However, it is uncertain whether HDAC9 is also implicated in other VSMC-driven diseases. Investigating these associations can further delineate the potential role of HDAC9 as a treatment target in arterial calcific and stenotic disease. Here, we analyzed a population sample of 335,146 adults to assess the role of *HDAC9* in mortality and VSMC-driven cardiovascular and renal pathology.

## RESULTS

## Associations in the UK Biobank

Among 335,146 participants, the AAC-associated single nucleotide polymorphism rs57301765 was genotyped in 332,425 persons (99.19%, 53.7% women, age  $56.9 \pm 8.0$  years, 10.1% smokers). We ascertained genotyping quality as detailed in the Supplemental Information (Figures S1–S4, Tables S1 and S2). Maximum follow-up was 10.0 years (mean  $7.0 \pm 1.0$ , median 7.1, interquartile range, 6.4–7.8 years). Figures 1 and 2 display associations between the AAC-associated allele and outcomes. At experiment-wise significance, we found associations between rs57301765 and systolic blood pressure ( $p = 1.71 \times 10^{-11}$ ), pulse pressure ( $p = 5.76 \times 10^{-27}$ ), and hypertension ( $p = 1.26 \times 10^{-6}$ ). The associations with myocardial infarction ( $p = 1.72 \times 10^{-5}$ ; driven by non-ST segment elevation myocardial infarction [NSTEMI],  $p = 0.001$ ) and ischemic stroke ( $p = 6.29 \times 10^{-5}$ ) were not significant at

<sup>1</sup>Department of Neurobiology, Care Sciences and Society (NVS), Family Medicine and Primary Care Unit, Karolinska Institutet, Alfred Nobels Allé 23, Huddinge 14183, Sweden

<sup>2</sup>School of Health and Social Studies, Dalarna University, Falun 79188, Sweden

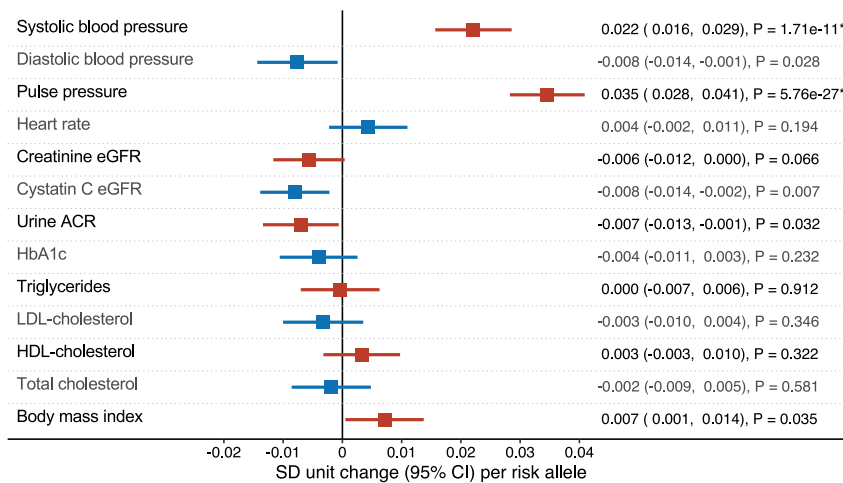
<sup>3</sup>Department of Biology, Morgan State University, Baltimore, MD 21251, USA

<sup>4</sup>Lead Contact

\*Correspondence: christoph.nowak@ki.se

<https://doi.org/10.1016/j.isci.2020.101253>





**Figure 1. Forest Plot of Per-Risk-Allele Associations with Continuous Outcomes**

Associations represent the change in standard deviation unit per added AAC-raising risk allele. Horizontal lines denote 95% confidence intervals. Bonferroni significance is marked by asterisks.

the strict Bonferroni level (Figures 1 and 2). There was no clear association between *HDAC9* and aneurysmal disease, but there was a nominally significant protective association with diastolic blood pressure ( $p = 0.028$ ). We found suggestive directionally consistent protective associations between the AAC-raising variant and kidney disease outcomes that reached nominal (but not experiment-wise) significance: microalbuminuria ( $p = 0.010$ ), albumin-to-creatinine ratio (ACR,  $p = 0.032$ ), and estimated glomerular filtration rate based on cystatin C levels (cystatin C-eGFR,  $p = 0.007$ ).

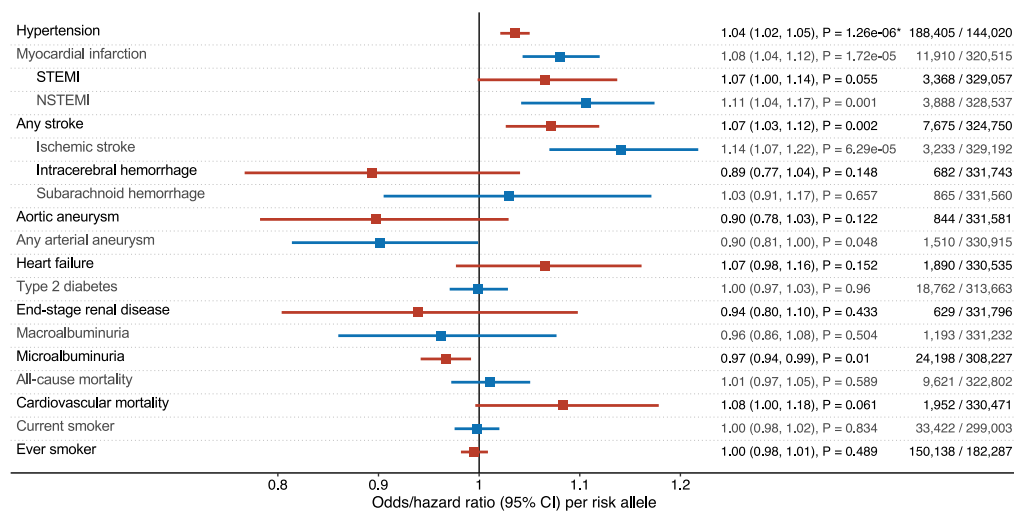
### Associations in Previously Reported GWAS Meta-Analyses

Results from previously reported GWAS meta-analyses (Table S2) confirm our findings in the UK Biobank sample in that the AAC risk allele was associated with raised pulse pressure ( $p = 4.00 \times 10^{-6}$ ) and hypertension ( $p = 0.005$ ). There was no association with diastolic blood pressure ( $p = 0.728$ ) or aortic valve calcification ( $p = 0.767$ ). We found a nominally significant protective effect on ACR risk ( $p = 0.037$ , Table S2). The non-significant ( $p = 0.152$ ) risk-increasing association with heart failure risk in our sample corresponded to a nominal effect in the same direction ( $p = 0.015$ ) in the HERMES heart failure consortium.

### DISCUSSION

In a population sample of >330,000 adults, we used a genetic variant to query the potential effects of targeting *HDAC9* for the treatment of VSMC-associated diseases. We discovered associations between the AAC-raising variant and increased systolic hypertension and raised risk of NSTEMI that confirm previously reported effects on cardiovascular outcomes. A suggestive protective association with kidney disease phenotypes did not reach experiment-wise significance but was directionally consistent across renal endpoints. We confirmed associations with ischemic stroke previously reported in independent samples. There were no clear associations with metabolic cardiovascular risk factors such as plasma lipid and glucose levels, or mortality.

The stronger association between *HDAC9* and NSTEMI, compared with ST segment elevation myocardial infarction (STEMI), could suggest *HDAC9* acting on myocardial infarction risk primarily through non-atherosclerotic mechanisms, as a higher proportion of NSTEMI can be attributed to non-atherosclerotic causes (Zipes et al., 2019). Whether arterial calcification is a particular risk factor for NSTEMI has not been assessed before. The causes of myocardial oxygen demand versus consumption mismatch in NSTEMI include stenotic disease and systemic pathologies such as hypertension (Zipes et al., 2019). Our results suggest a role of *HDAC9* in myocardial infarction through non-coronary plaque-related, possibly stenotic, hypertensive, or calcific aortic pathology. However, the differential associations with STEMI and NSTEMI in our study need to be interpreted with great caution as (1) the clinical diagnoses in the UK Biobank have not been validated for accuracy, (2) our study may be underpowered (e.g., the effect size of 1.07 and  $p$ -value of 0.055 for STEMI suggests that larger samples may detect an effect that we missed because of lack of power), and (3) although non-atherosclerotic flow-limiting causes (such as systemic arterial hypertension) are more common for NSTEMI, both STEMI and NSTEMI share



**Figure 2. Forest Plot of Per-Risk-Allele Associations with Binary Outcomes**

Effects per added AAC-raising risk allele are shown as odds ratios or hazard ratios (for all-cause mortality and cardiovascular mortality). Horizontal lines denote 95% confidence intervals. Bonferroni significance is marked by asterisks.

atherosclerosis as the predominant cause. Hence, our discovery of an apparently stronger effect of AAC-raising genetic variation in *HDAC9* on NSTEMI versus STEMI risk points to an interesting distinction that has received little attention in previous research. Yet, the limitations of our study do not allow firm conclusions, and additional studies exploring a potential differential role of *HDAC9*-driven AAC in different cardiovascular pathologies are needed.

The strong association with systolic hypertension confirms *HDAC9* as a likely key regulator of VSMC-related cardiovascular diseases (Malhotra et al., 2019; Lino Cardenas et al., 2019). The suggestive protective effect on kidney disease due to AAC-associated genetic variation in our comparatively healthy sample has not been reported before and requires independent replication. Although we found consistent effect sizes in the protective direction across renal phenotypes, none of the associations reached experiment-wise significance. Our sample comprising comparatively healthy, middle-aged adults was likely underpowered to detect potential effects on renal outcomes (the proportion of participants with eGFR below 60 was 2.2%, and 58.2% had normal kidney function with an eGFR above 90).

We found no association between *HDAC9* and several conventional cardiovascular risk factors (apart from blood pressure traits), which supports *HDAC9* as a possible add-on drug target for cardiovascular disease acting through different mechanisms than established therapies. Whether pharmacological targeting of *HDAC9* could enhance existing blood pressure treatments or act as an alternative treatment remains uncertain and needs to be addressed in future experimental studies.

### Limitations of the Study

Limitations include exclusive European ancestry, limited power for some outcomes (which may have missed true effects), reliance on a single genetic variant, and the caveats of extrapolating from genetic effects to clinical reality. The clinical phenotypes in the UK Biobank have not been validated for clinical accuracy (leaving doubts about the distinction between NSTEMI and STEMI based on hospital admissions records), and there is evidence for a “healthy selection bias” in the UK Biobank. Our epidemiologic study provides limited mechanistic insights, and cardiovascular risk groups such as elderly persons or those with chronic kidney disease were likely underrepresented. Strengths include the large, contemporary sample representative of the UK population, consistent case ascertainment, exclusion of genotyping errors, and confirmation in independent GWAS data.

### Conclusion

In a community sample of >330,000 adults, we found confirmatory evidence in support of *HDAC9* as a key regulator and potential therapeutic target for VSMC-driven cardiovascular disease, with suggestive evidence of a reverse effect on kidney disease.

## Resource Availability

### Lead Contact

Dr. Christoph Nowak, M.D., Ph.D., Dipl.-Psych., Department of Neurobiology, Care Sciences and Society (NVS), Karolinska Institutet, Alfred Nobels Allé 23, SE-14183 Huddinge, Sweden, email: [christoph.nowak@ki.se](mailto:christoph.nowak@ki.se), phone: +46-739806535, ORCID iD: 0000-0001-8435-3978.

### Materials Availability

We used existing data as described below. No new data or materials were generated in our study.

### Data and Code Availability

Researchers can apply to use the UK Biobank data for health-related research in the public interest (<http://www.ukbiobank.ac.uk/register-apply/>). After ethical approval, the UK Biobank releases de-identified data to approved researchers for specific research projects. All the data used obtained through look-ups of GWAS repositories are in the public domain and available via the webpages provided in the main text and [Table S2](#). The analysis code can be obtained from the corresponding author (C.N.) on request.

## METHODS

All methods can be found in the accompanying [Transparent Methods supplemental file](#).

## SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.isci.2020.101253>.

## ACKNOWLEDGMENTS

We thank all contributors and participants of the UK Biobank study. Summary data have been kindly contributed by the HERMES consortium for heart failure (<http://www.hermesconsortium.org/>) and by the CKDGen consortium (<https://ckdgen.imbi.uni-freiburg.de/>) for kidney disease phenotypes.

Funding: The UK Biobank was supported by the Medical Research Council, the Wellcome Trust, the UK Department of Health, the British Heart Foundation, Cancer Research UK, the US National Institute for Health Research, the Scottish Government, the North West Development Agency, Diabetes UK, and the Welsh Government (grants are listed here <https://www.ukbiobank.ac.uk/wp-content/uploads/2018/10/Funding-UK-Biobank-summary.pdf>). C.N. was supported by a European Foundation for the Study of Diabetes EFSD/Lilly Young Investigator Research Award and grants from Karolinska Institutet (Loo and Hans Osterman Foundation; Foundation for Geriatric Diseases at Karolinska Institutet). J.Ä. was supported by the Swedish Research Council (Vetenskapsrådet 2012–2215). D.F.D. was supported by NIMHD RCMI@Morgan #U54MD013367-8281. The funding sources had no role in any aspect of the study or the writing and interpretation of the manuscript.

## AUTHOR CONTRIBUTIONS

The corresponding author (C.N.) attests that all authors meet the authorship criteria and that all actual or perceived conflicts of interests have been declared. C.N., with assistance from J.Ä., conceived of, and planned, the study. C.N., D.F.D., and J.Ä. developed the methods. C.N. wrote the first draft of manuscript. C.N. led, J.Ä. verified, and D.F.D. critically revised, data analysis and results. C.N. is the guarantor of the study. J.Ä. and C.N. obtained funding for the study. C.N., D.F.D., and J.Ä. critically revised the manuscript.

## DECLARATION OF INTERESTS

The authors declare no support from any organization for the submitted work, no financial relationships with any organizations that might have an interest in the submitted work in the previous three years, and no other relationships or activities that could appear to have influenced the submitted work.

Received: February 10, 2020

Revised: April 20, 2020

Accepted: June 4, 2020

Published: July 24, 2020

## REFERENCES

- Bastos Goncalves, F., Voute, M.T., Hoeks, S.E., Chonchol, M.B., Boersma, E.E., Stolker, R.J., and Verhagen, H.J. (2012). Calcification of the abdominal aorta as an independent predictor of cardiovascular events: a meta-analysis. *Heart* 98, 988–994.
- Bellenguez, C., Bevan, S., Gschwendtner, A., Spencer, C.C., Burgess, A.I., Pirinen, M., Jackson, C.A., Traylor, M., Strange, A., Su, Z., et al. (2012). Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. *Nat. Genet.* 44, 328–333.
- Criqui, M.H., Denenberg, J.O., McClelland, R.L., Allison, M.A., Ix, J.H., Guerci, A., Cohoon, K.P., Srikanthan, P., Watson, K.E., and Wong, N.D. (2014). Abdominal aortic calcium, coronary artery calcium, and cardiovascular morbidity and mortality in the Multi-Ethnic Study of Atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* 34, 1574–1579.
- Kato, N., Loh, M., Takeuchi, F., Verweij, N., Wang, X., Zhang, W., Kelly, T.N., Saleheen, D., Lehne, B., Leach, I.M., et al. (2015). Trans-ancestry genome-wide association study identifies 12 genetic loci influencing blood pressure and implicates a role for DNA methylation. *Nat. Genet.* 47, 1282–1293.
- Lino Cardenas, C.L., Kessinger, C.W., Chou, E.L., Chou, E.L., Ghoshhajra, B., Yeri, A.S., Das, S., Weintraub, N.L., Malhotra, R., Jaffer, F.A., et al. (2019). HDAC9 complex inhibition improves smooth muscle-dependent stenotic vascular disease. *JCI Insight* 4, 124706.
- Malhotra, R., Mauer, A.C., Lino Cardenas, C.L., Guo, X., Yao, J., Zhang, X., Wunderer, F., Smith, A.V., Wong, Q., Pechlivanis, S., et al. (2019). HDAC9 is implicated in atherosclerotic aortic calcification and affects vascular smooth muscle cell phenotype. *Nat. Genet.* 51, 1580–1587.
- Malik, R., Traylor, M., Pulit, S.L., Bevan, S., Hopewell, J.C., Holliday, E.G., Zhao, W., Abrantes, P., Amouyel, P., Attia, J.R., et al. (2016). Low-frequency and common genetic variation in ischemic stroke: the METASTROKE collaboration. *Neurology* 86, 1217–1226.
- Malik, R., Dau, T., Gonik, M., Sivakumar, A., Deredge, D.J., Edeleva, E.V., Götzfried, J., van der Laan, S.W., Pasterkamp, G., Beaufort, N., et al. (2017). Common coding variant in SERPINA1 increases the risk for large artery stroke. *Proc. Natl. Acad. Sci. U S A* 114, 3613–3618.
- Nikpay, M., Goel, A., Won, H.H., Hall, L.M., Willenborg, C., Kanoni, S., Saleheen, D., Kyriakou, T., Nelson, C.P., Hopewell, J.C., et al. (2015). A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat. Genet.* 47, 1121–1130.
- Traylor, M., Farrall, M., Holliday, E.G., Sudlow, C., Hopewell, J.C., Cheng, Y.C., Fornage, M., Ikram, M.A., Malik, R., Bevan, S., et al. (2012). Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol.* 11, 951–962.
- Vervloet, M., and Cozzolino, M. (2017). Vascular calcification in chronic kidney disease: different bricks in the wall? *Kidney Int.* 91, 808–817.
- Zipes, A.P., Libby, P., Bonow, R.O., Mann, D.L., and Tomaselli, G.F. (2019). *Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine*, Eleventh Edition (Elsevier).

iScience, Volume 23

## Supplemental Information

### **Atherosclerotic Aortic Calcification-Associated Polymorphism in *HDAC9* and Associations with Mortality, Cardiovascular Disease, and Kidney Disease**

**Johan Ärnlöv, Douglas F. Dluzen, and Christoph Nowak**

## Transparent Methods

### Study rationale

Large population studies are valuable for querying the predicted effects of therapeutic targeting of proteins. Genetic variants that mimic manipulation of the target protein act over the entire lifespan and can serve as proxies to assess expected long-term effects. As genetic variants are randomly allocated at conception, any non-random differences between carriers and non-carriers of an exposure-associated variant should be attributable to different genotypes. Carriers of a variant known to influence an outcome through a specific protein can therefore provide insights into the possible effects of manipulating the protein pharmacologically (Plenge et al., 2013, Nguyen et al., 2019). Important caveats include ignoring biological canalization, sensitive periods, selection pressure, threshold effects, and off-target effects. Although altering protein activity in adult life can have different consequences than genetically determined activity, querying epidemiologic studies has proven useful in predicting drug effects (Schmidt et al., 2019).

### Study population

The UK Biobank (<http://www.ukbiobank.ac.uk>) recruited 2006-2010 altogether 502,655 persons aged 40-69 years from the UK population. Informed consent was followed by standardized health assessments and follow-up across national health registers. We included all participants with genotyping data included in genetic principle component analysis (excluding related individuals) and assigned to "white British" ethnicity. Participants who had withdrawn consent by August 2019 were excluded. Maximum follow-up was until 31 March 2017. We selected the following VSMC-associated outcomes (requiring >500 cases): blood and pulse pressure, heart rate, hypertension, myocardial infarction, stroke, aortic/arterial aneurysm, heart failure, estimated glomerular filtration rate (eGFR) by serum creatinine or cystatin C, urine albumin-to-creatinine ratio (ACR), end-stage renal disease, all-cause and cardiovascular mortality, glycated hemoglobin (HbA1c), diabetes, serum lipids, body mass index, and smoking history. Disease endpoints were defined based on nurse interview, hospital episode diagnoses (ICD-9/10), and medication records (**Table S1**).

### Representativeness of the sample

UK Biobank participants are more likely to be female, older, healthier, and less exposed to socioeconomic deprivation than the general UK population (Fry et al., 2017). The limited number of persons with genotyping data who reported non-European ancestry required us to exclude these participants from the analysis. The Generalizability of our results to non-European ethnicities is therefore uncertain.

### Statistical analysis

We selected the top single nucleotide polymorphism (SNP) rs57301765 (hg38 chr7:19,052,733) in HDAC9, whose minor A-allele was associated with worse AAC (Malhotra et al., 2019). Out of six reported AAC-associated SNPs in HDAC9 (Malhotra et al., 2019), only rs57301765 was directly genotyped in the UK Biobank.

We queried public GWAS data using PhenoScanner (<http://www.phenoscanter.medschl.cam.ac.uk/>). We also performed look-ups in the HERMES heart failure (<http://www.hermesconsortium.org/mou>) and CKDGen chronic kidney disease (<https://ckdgen.imbi.uni-freiburg.de/>) consortia to identify genetic associations in addition to the previously reported associations with ischemic stroke (METASTROKE) and myocardial infarction (CARDIoGRAMplusC4D) (Malik et al., 2016, Malik et al., 2017, Nikpay et al., 2015).

We used logistic and linear regression adjusted for age, sex, genotyping array, and the top ten genetic principle components to assess associations between the AAC-associated allele and outcomes. Adjusted Cox regression was used for mortality, and proportional hazards and linearity ascertained in Schoenfeld and Martingale residual plots. Skewed distributions of HbA1c and ACR prompted natural log-transformation to normality. We used Bonferroni (experiment-wise) correction for multiple comparisons adjusting for a total number of phenotypes in the UK Biobank of about 7,500 ( $P < 5 \times 10^{-6}$ ). Analyses were carried out in R v.3.3.

### Ethical approval

The UK Biobank study was approved by the North West Multi-Centre Research Ethics Committee. The current study (UK Biobank project 42176) complies with the Declaration of Helsinki and ethical approval was granted by the Swedish Ethical Review Authority (Dnr. 2019-02328). Data from all participants who withdrew consent were excluded from the study. The purpose of this study was approved by the UK Biobank organization in accordance with participant consent as representing health-related research in the public interest.



### Patient and public involvement

The purpose of this study was approved by the UK Biobank organization in accordance with participant consent as representing health-related research in the public interest. We did not directly involve patients or the general public in this study.

### Data sharing information

Researchers can apply to use the UK Biobank data for health-related research in the public interest (<http://www.ukbiobank.ac.uk/register-apply/>). After ethical approval, the UK Biobank releases de-identified data to approved researchers for specific research projects. All the data used obtained through look-ups of GWAS repositories are in the public domain and available through the webpage provided in the main text and **Table S2**. The analysis code can be obtained from the corresponding author (C.N.) on request.

### Ascertainment of genotyping quality

Concerns about the genotyping quality of a specific genetic variant in the UK Biobank recently led to a very prominent example of a retracted publication in *Nature Medicine* (Wei and Nielsen, 2019). In a series of analyses published independently by three different research groups and published on the preprint server *bioRxiv*, it was demonstrated that poor genotyping performance on the UK Biobank arrays and technical error plausibly led to erroneous results that could not be replicated in other samples (Gudbjartsson et al., 2019, Maier et al., 2019, Tanigawa and Rivas, 2019). In order to ascertain the genotyping quality of the variant rs57301765 (hg38 chr7:19,052,733) in the UK Biobank, we followed the procedure by Yosuke Tanigawa and Manuel Rivas (2019). The variant rs57301765 was genotyped on both the UK Biobank Affymetrix array (n = 95 batches) and the UKBiLIEVE array (n = 11 batches). None of the batches failed genotyping for this variant. The overall genotyping missingness rate of 0.81% in our sample of 335,146 White British individuals was acceptably low. There were no systematic differences between the arrays and no outlying genotyping batches were apparent with regard to genotyping rate, which ranged from 98.08% - 99.69% (median 99.25%), and 98.20% - 99.30% (median 98.68%) for the Affymetrix and UKBiLIEVE arrays, respectively. Cluster plots and additional genotyping performance data were obtained using the *ScatterShot* application by Mark McCarthy's group (<http://mccarthy.well.ox.ac.uk/ScatterShotWebApp/ui/scattershot>). We selected the sample of 409,634 White British individuals in *ScatterShot*. **Figure S1** shows cluster plots of genotype alternatives for women and men, demonstrating acceptable separation of genotype classes.

As the  $P$ -value for Hardy-Weinberg equilibrium (HWE) implied deviations from HWE ( $P_{women} = 2.9 \times 10^{-8}$ ,  $P_{men} = 8.2 \times 10^{-8}$ ), we assessed whether poor genotyping quality across batches or arrays could explain the deviation from HWE. As in the publication from Manuel Rivas' group (Tanigawa and Rivas, 2019), we extracted genotyping rate, batch and array identifiers, heterozygosity rate, and HWE  $P$ -values for each batch. As illustrated in **Figure S2**, we found no association between  $-\log_{10}(\text{HWE } P\text{-value})$  and either array (linear regression  $P = 0.141$ ), or batch ( $P = 0.491$ ).

The distributions of  $-\log_{10}(\text{HWE } P\text{-value})$  did not differ between arrays (Kolmogorov-Smirnov test,  $P = 0.412$ ), or between batches with genotyping rates above or below the median ( $P = 0.911$ ) (**Figure S3**, **Figure S4**). Finally, the minor allele frequency in our sample of 0.1532 was comparable to the minor allele frequency of 0.1696 in the European non-Finnish reference sample reported on the gnomAD browser (<http://gnomad.broadinstitute.org/>).

Taken together, these data demonstrate good genotyping performance for rs57301765 in the UK Biobank. Technical error or inadequate quality control were not apparent and are unlikely to explain the possible deviation from Hardy-Weinberg equilibrium.

## Supplemental References

Plenge, R.M., Scolnick, E.M., and Altshuler, D. (2013). Validating therapeutic targets through human genetics. *Nat. Rev. Drug. Discov.* *12*, 581-594.

Nguyen, P.A., Born, D.A., Deaton, A.M., Nioi, P., and Ward, L.D (2019). Phenotypes associated with genes encoding drug targets are predictive of clinical trial side effects. *Nat. Commun.* *10*, 1579.

Schmidt, A.F., Holmes, M.V., Preiss, D., et al (2019). Phenome-wide association analysis of LDL-cholesterol lowering genetic variants in PCSK9. *BMC Cardiovasc. Disord.* *19*, 240.

Fry, A., Littlejohns, T.J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., Collins, R., and Allen, N.E. (2017). Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am. J. Epidemiol.* *186*, 1026-1034.

Malhotra, R., Mauer, A.C., Lino Cardenas, C.L., et al (2019). HDAC9 is implicated in atherosclerotic aortic calcification and affects vascular smooth muscle cell phenotype. *Nat. Genet.* *51*, 1580-1587.

Malik, R., Traylor, M., Pulit, S.L., et al (2016). Low-frequency and common genetic variation in ischemic stroke: The METASTROKE collaboration. *Neurology* *86*, 1217-1226.

Malik, R., Dau, T., Gonik, M., et al (2017). Common coding variant in SERPINA1 increases the risk for large artery stroke. *Proc. Natl. Acad. Sci. U S A.* *114*, 3613-3618.

Nikpay, M., Goel, A., Won, H.H., et al (2015). A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat. Genet.* *47*, 1121-1130.

Wei, X., and Nielsen, R. (2019). Retraction Note: CCR5- $\Delta$ 32 is deleterious in the homozygous state in humans. *Nat. Med.* *25*, 1796.

Gudbjartsson, D., Sulem, P., Stefansson, K., Mars, N., Karjalainen, J., Ripatti, S., Palotie, A., and Daly, M. (2019). CCR5-del32 is not deleterious in the homozygous state in humans. *bioRxiv*. 2019:788117. <https://doi.org/10.1101/788117>. Accessed December 1, 2019.

Maier, R., Akbari, A., Wei, X., Patterson, N., Nielsen, R., and Reich, D. (2019). No statistical evidence for an effect of CCR5- $\Delta$ 32 on lifespan in the UK Biobank cohort. *bioRxiv*. 2019:787986. <https://doi.org/10.1101/787986>. Accessed December 1, 2019.

Tanigawa, Y., and Rivas, M.A. (2019). Reported CCR5- $\Delta$ 32 deviation from Hardy-Weinberg equilibrium is explained by poor genotyping of rs62625034. *bioRxiv*. 2019:791517. <https://doi.org/10.1101/791517>. Accessed December 1, 2019.

**Table S1. Outcome definitions in the UK Biobank sample, Related to Figure 1 and Figure 2.**

Variable	UK Biobank data field and/or study-specific definition	Additional comments
<b>Systolic blood pressure</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=4080">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=4080</a>	Automated reading, two measures taken a few moments apart. OMRON Healthcare BP device.
<b>Diastolic blood pressure</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=4079">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=4079</a>	Automated reading, two measures taken a few moments apart. OMRON Healthcare BP device.
<b>Hypertension</b>	Self-reported diagnosis of hypertension, current blood pressure medication, physician diagnosis of hypertension, SBP >140 mmHg or DBP >90mmHg	
<b>Heart rate</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=95">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=95</a>	(during blood pressure assessment)
<b>Pulse pressure</b>	SBP minus DBP	
<b>Stroke</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=42006">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=42006</a>	Centrally algorithmically defined by UK Biobank
<b>Ischemic stroke</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=42008">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=42008</a>	Centrally algorithmically defined by UK Biobank
<b>Intracerebral hemorrhage</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=42010">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=42010</a>	Centrally algorithmically defined by UK Biobank
<b>Subarachnoid hemorrhage</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=42012">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=42012</a>	Centrally algorithmically defined by UK Biobank
<b>Myocardial infarction</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=42000">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=42000</a>	Centrally algorithmically defined by UK Biobank
<b>STEMI</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=42002">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=42002</a>	Centrally algorithmically defined by UK Biobank
<b>NSTEMI</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=42004">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=42004</a>	Centrally algorithmically defined by UK Biobank
<b>Aortic aneurysm</b>	ICD-9: 441, ICD-10: I71	
<b>Any arterial aneurysm</b>	ICD-9: 441, 442, 414.1, 417.1, 437.3; ICD-10: I71, I72, I79.0, I25.3, I25.4, I28.1, I67.1	
<b>Creatinine eGFR</b>	Serum creatinine measured by enzymatic assay (Beckman Coulter, UK) on a Beckman Coulter AU5800 analyzer.	CKD-EPI equation adjusted for sex, ethnicity and age
<b>Cystatin C eGFR</b>	Serum cystatin C measured by immuno-turbidimetric assay (Siemens, Germany) on a Siemens ADVIA 1800 analyzer	CKD-EPI equation adjusted for sex and age
<b>Urine albumin to creatinine ratio (ACR)</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=30510">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=30510</a> <a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=30500">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=30500</a>	Urine microalbumin measured by immuno-turbidimetric assay (Randox Bioscience, UK) on an AU5400 analyzer. Values below analytical range 6.7-200 mg/L set to lower limit of detection divided by two (3.35 mg/L). Urine creatinine measured by enzymatic assay (Beckman Coulter) on an AU5400 analyzer within analytical range 88-44,200 micromol/L.
<b>Macroalbuminuria</b>	Urine albumin to creatinine ratio > 300mg/g	
<b>Microalbuminuria</b>	Urine albumin to creatinine ratio 30-300mg/g	
<b>End-stage renal disease</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=42026">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=42026</a>	Centrally algorithmically defined by UK Biobank

<b>Heart failure</b>	ICD-9: 428, ICD-10: I50, I11.0, I13.0, I13.2	
<b>Type 2 diabetes</b>	Self-reported diagnosis of diabetes, current diabetes medication, physician diagnosis of type 2 diabetes, HbA1c > 6.5%	
<b>All-cause mortality</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=40007">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=40007</a>	
<b>Cardiovascular mortality</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=40001">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=40001</a>	Underlying primary cause of death ICD-10 I00-I99
<b>HbA1c</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=30750">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=30750</a>	Measured by HPLC analysis on a Bio-Rad VARIANT II Turbo
<b>Body mass index</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=21001">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=21001</a>	Calculated from height and weight measured during baseline assessment center visit.
<b>Current smoker</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=20116">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=20116</a>	Answer "Current smoker"
<b>Ever smoker</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=20116">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=20116</a>	Answers "Current smoker" + "Previous smoker"
<b>Serum cholesterol</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=30690">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=30690</a>	Enzymatic assay (Beckman Coulter, UK); AU5800 Analyzer
<b>Serum HDL-cholesterol</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=30760">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=30760</a>	Enzymatic Immuno-inhibition assay (Beckman Coulter); AU5800 analyzer
<b>Serum LDL-cholesterol</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=30780">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=30780</a>	Enzymatic selective protection assay (Beckman Coulter); AU5800 Analyzer
<b>Serum triglycerides</b>	<a href="http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=30870">http://biobank.ctsu.ox.ac.uk/showcase/field.cgi?id=30870</a>	Enzymatic assay (Beckman Coulter); AU5800 Analyzer

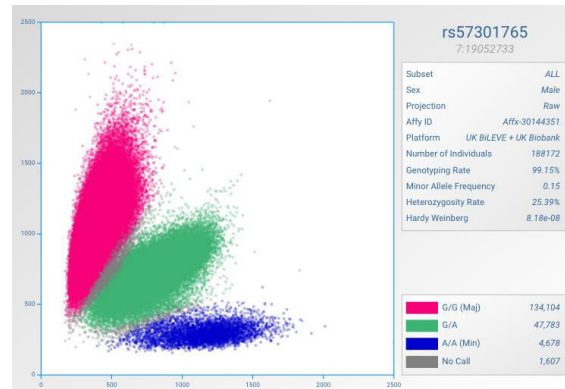
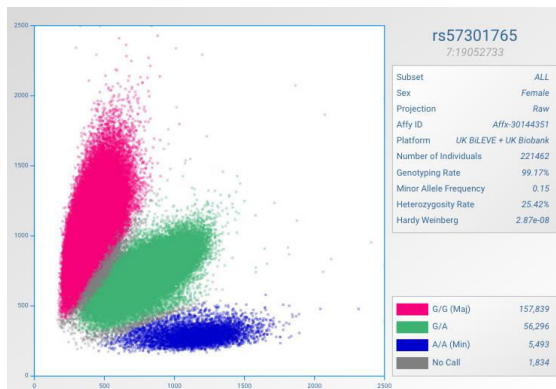
**Table S2. Publicly available summary genome-wide association study results, Related to Figure 1 and Figure 2.** Obtained as search results in PhenoScanner, or manual search of GWAS results for heart failure (HERMES Consortium) and kidney disease endpoints (CKDGen Consortium). We selected rs2107595 (chr7:19,009,765) as proxy if rs57301765 was not available (linkage disequilibrium in the European reference sample,  $R^2$  0.965,  $D'$  0.986, <https://ldlink.nci.nih.gov/>).

RSID	Hg38	A1	A2	Proxy	Trait	PMID/Source	Ethn	Beta	SE	P	N
rs57301765	chr7:19013110	A	G	rs2107595	Pulse pressure	26390057	Mixed	0.307	0.046	4.0E-11	320,251
rs57301765	chr7:19013110	A	G	rs2107595	Pulse pressure	28739976	EUR	0.291	0.063	4.0E-06	378,379
rs57301765	chr7:19013110	A	G	rs2107595	SBP	21909115	EUR	NR	NR	0.005	69,395
rs57301765	chr7:19013110	A	G	rs2107595	SBP	21909115	EUR	NR	NR	0.005	69,395
rs57301765	chr7:19013110	A	G	rs2107595	Aortic valve calcification	23388002	Mixed	NR	NR	0.728	6,942
rs57301765	chr7:19013110	A	G	rs2107595	DBP	21909115	EUR	NR	NR	0.767	69,395
rs57301765	chr7:19013110	A	G	-	Heart failure	HERMES*	EUR	0.026	0.011	0.015	956,557
rs57301765	chr7:19013110	A	G	-	eGFRcreat	CKDGen**	EUR	1e-04	5e-04	0.817	525,153
rs57301765	chr7:19013110	A	G	-	CKD	CKDGen**	EUR	0.013	0.013	0.322	402,682
rs57301765	chr7:19013110	A	G	-	eGFRcreat	CKDGen**	Mixed	-3e-04	4e-04	0.493	721,594
rs57301765	chr7:19013110	A	G	-	CKD	CKDGen**	Mixed	0.014	0.010	0.147	579,035
rs57301765	chr7:19013110	A	G	-	Microalbuminuria	CKDGen**	Mixed	-0.013	0.009	0.171	336,215
rs57301765	chr7:19013110	A	G	-	Urine ACR	CKDGen**	EUR	-0.005	0.003	0.047	533,701
rs57301765	chr7:19013110	A	G	-	Urine ACR	CKDGen**	Mixed	-0.006	0.003	0.037	549,054

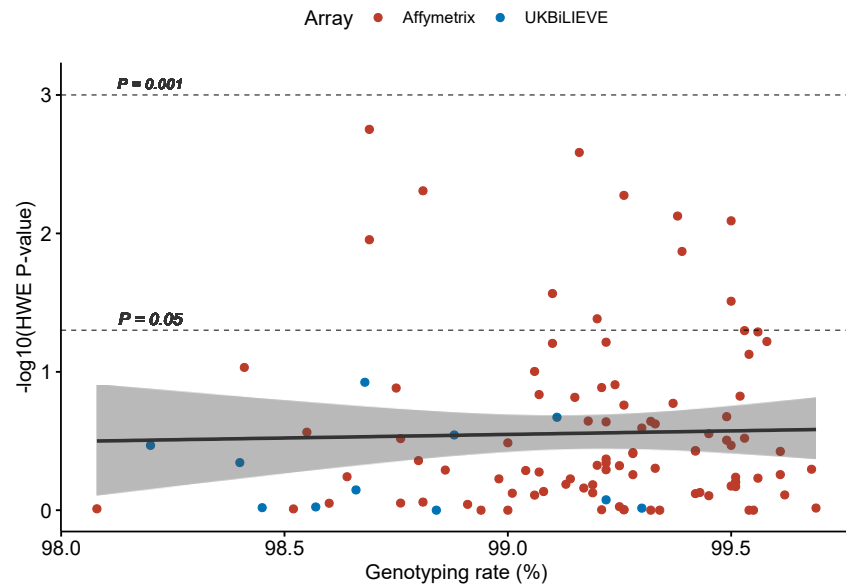
\*<https://www.biorxiv.org/content/10.1101/682013v1>; Data available here: <http://www.broadcvdi.org/>

\*\*Summary data from different studies in CKDGen available here: <https://ckdgen.imbi.uni-freiburg.de/>

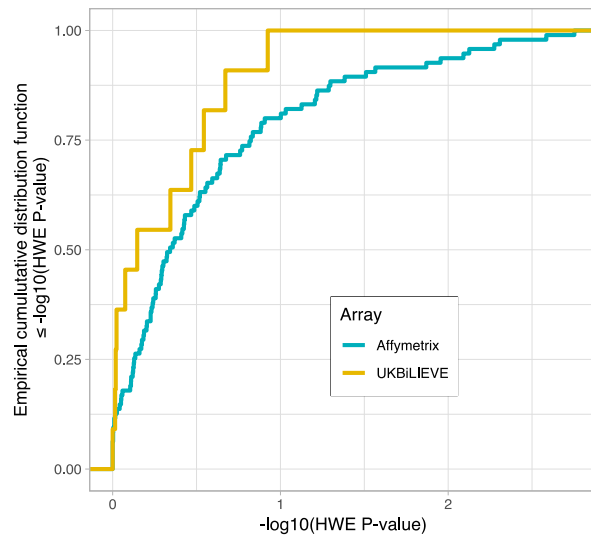
A1: effect allele, A2: other allele, Ethn: Ethnicity, AVC: aortic valve calcification, DBP: diastolic blood pressure, NR: not reported, PP: pulse pressure, SBP: systolic blood pressure, SE: standard error. Linkage disequilibrium between rs57301765 and rs2107595 in EUR reference  $r^2$  0.9649,  $D'$  0.9858



**Figure S1. Cluster plots of genotype classes, Related to Figure 1 and Figure 2.** Genotypes G/G, A/G, and A/A in 409,634 White British participants in the UK Biobank for women (left), and men (right) are shown as provided by *ScatterShot*.

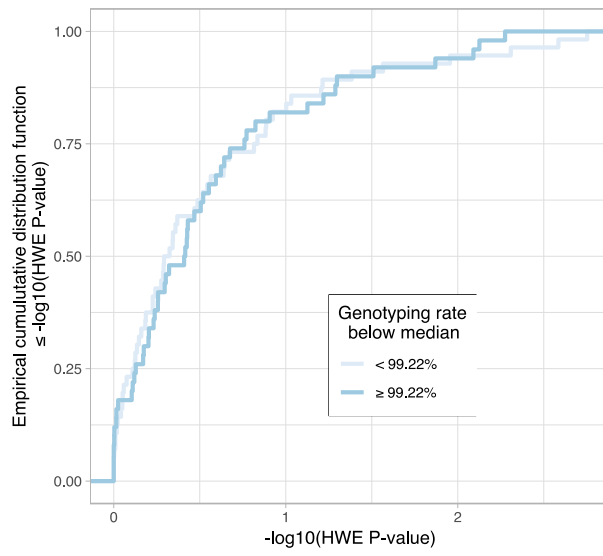


**Figure S2. Scatterplot of  $-\log_{10}(\text{HWE P-value})$  against genotyping rate, Related to Figure 1 and Figure 2.** Each dot represents one of the genotyping batches of the Affymetrix (red), or UKBiLIEVE arrays (blue). The line drawn through the dots represents the ordinary least squares best fit line with associated 95% confidence interval (shaded area).



**Figure S3. Cumulative  $-\log_{10}(\text{HWE P-value})$  distribution functions for batches of the Affymetrix and UKBiLIEVE array, Related to Figure 1 and Figure 2.**





**Figure S4. Cumulative  $-\log_{10}(\text{HWE P-value})$  distribution functions for batches with genotyping rates above or below the median rate, Related to Figure 1 and Figure 2.**