Replication of Newly Identified Genetic Associations Between Abdominal Aortic Aneurysm and *SMYD2*, *LINC00540*, *PCIF1/MMP9/ZNF335*, and *ERG*

Weihong Tang [a,*], Athanasios Saratzis [b], Jack Pattee [c], Jacqueline Smith [d], Nathan Pankratz [e], Olivia C. Leavy [f,g], Weihua Guan [c], Frank Dudbridge [g], James S. Pankow [a], George D. Kitas [h], Pamela L. Lutsey [a], Matthew J. Bown [b,**]

[a] Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN, USA
[b] Department of Cardiovascular Sciences and National Institute of Health Research (NIHR), Leicester Biomedical Research Centre, University of Leicester, Leicester, UK
[c] Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN, USA
[d] Dudley Group of Hospitals NHS Trust, Dudley, UK
[e] Department of Laboratory Medicine and Pathology, School of Medicine, University of Minnesota, Minneapolis, MN, USA
[f] Department of Non-communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK
[g] Department of Health Sciences, University of Leicester, Leicester, UK
[h] College of Life and Environmental Sciences, University of Birmingham, Birmingham, UK

WHAT THIS PAPER ADDS

A recently published genome wide association study of abdominal aortic aneurysms (AAA) identified four new loci for AAA: *SMYD2* (top single nucleotide polymorphism [SNP] rs1795061), *LINC00540* (rs9316871), *PCIF1/MMP9/ZNF335* (rs3827066), and *ERG* (rs2836411). Based on data from a US based prospective cohort study and a Greece based case control study, associations between rs9316871 and rs2836411 and AAA risk were replicated in a meta-analysis of the two independent cohorts, providing further support for the importance of these loci in the aetiology of AAA. Findings from the study contribute to improved understanding of the pathophysiology of AAA and may aid in the management of patients with this condition.

Objective: A recently published genome wide association study of abdominal aortic aneurysms (AAA), based on pooled case control data of European ancestry, identified four new loci for AAA: *SMYD2* (top single nucleotide polymorphism [SNP] rs1795061), *LINC00540* (rs9316871), *PCIF1/MMP9/ZNF335* (rs3827066), and *ERG* (rs2836411). Of the four, rs1795061 and rs2836411 showed significant heterogeneity across studies and the *p* value for rs9316871 did not reach the genome wide significance threshold until discovery and replication data were pooled together in that study. The objective of this study was to replicate these newly identified genetic associations for AAA in a US based prospective cohort study, the Atherosclerosis Risk in Communities (ARIC) Study, and a Greece based case control study.

Methods: ARIC identified 408 clinically diagnosed AAAs among 8 962 individuals of European ancestry during a median of 22 years of follow up. The Greek case control study included 341 AAAs of European ancestry recruited in a tertiary referral centre and 292 geographically and ethnically matched controls recruited from the same institution. A Cox proportional hazards model was used to analyse the ARIC data and logistic regression to analyse the Greek data.

Results: In ARIC, rs9316871 and rs3827066 were significantly associated with AAA risk (HR [p] was 0.77 [.004] and 1.22 [.03], respectively), rs2836411 was associated at borderline significance (1.13 [.08]), whereas rs1795061 was not associated (p = .55). In the Greek case control study, rs1795061 and rs2836411 were significantly associated with AAA (OR [p] was 1.66 [< .001] and 1.29 [.04], respectively), whereas rs9316871 was not (p = .81). Genotyping of rs3827066 did not succeed. In the meta-analysis of the two studies, the association for rs9316871and rs2836411 was statistically significant and consistent between the two studies: *p* = .02 and .007, respectively.

Conclusions: Associations between rs9316871and rs2836411 and AAA risk were replicated in the meta-analysis of the two independent cohorts, providing further support for the importance of these loci in the aetiology of AAA.

* Corresponding author. Division of Epidemiology & Community Health, School of Public Health, University of Minnesota, 1300 South 2nd Street, Suite 300, Minneapolis, MN 55454, USA.
** Corresponding author. Cardiovascular Sciences, University of Leicester, University Road, Leicester, LE2 7LX, UK.
E-mail addresses: tang0097@umn.edu (Weihong Tang); mjb42@leicester.ac.uk (Matthew J. Bown).

INTRODUCTION

Significant genetic influence has been demonstrated for abdominal aortic aneurysms (AAA), with heritability of up to 77%.[1,2] Recently, Jones et al. reported a large genome wide association study (GWAS) of AAA that pooled data from several case control studies of populations of European ancestry (EA).[3] This study identified four new genes/loci: *SMYD2* (top single nucleotide polymorphism [SNP] rs1795061), *LINC00540* (rs9316871), *PCIF1/MMP9/ZNF335* (rs3827066), and *ERG* (rs2836411). Of the four, rs1795061 and rs2836411 showed significant heterogeneity across studies and the *p* value for rs9316871 did not reach the genome wide significance threshold until discovery and replication data were pooled together. It is important to replicate these new findings in independent populations and in prospective data. This study sought to replicate the newly identified associations of these SNPs with AAA in two independent studies that were not part of the study by Jones *et al.*: a prospective cohort study of European Americans with 24 years of follow up in the US, the Atherosclerosis Risk in Communities (ARIC) Study, and a case control study using a white Caucasian population from Greece.

MATERIALS AND METHODS

Study populations and genotyping

The ARIC Study recruited a population based cohort of 15 792 persons (73% EAs) aged 45–64 years in 1987–1989 and followed them through to 2011 for clinically diagnosed AAAs, identified by searching hospitalisation and death records as well as Medicare data.[4] Clinically diagnosed AAAs were defined as *ICD*-9-CM codes of 441.3 (ruptured AAA) or 441.4 (AAA without mention of rupture), or procedure codes of 38.44 (AAA resection and replacement) or 39.71 (AAA endovascular repair), or the following cause of death codes: *ICD-9* 441.3 or 441.4 or *ICD-10* code I71.3 (ruptured AAA) or I71.4 (AAA without mention of rupture).[4] Details of AAA ascertainment and risk factor measurement in ARIC have been described elsewhere.[4] Participants who reported AAA surgery prior to baseline (*n* = 11 without information on ICD codes) and uncertain AAA status during follow up (*n* = 30) were excluded.

ARIC extracted genomic DNA from blood and conducted genotyping with the Affymetrix Genome-Wide Human SNP array 6.0. To expand the number of genetic markers beyond the genotype array for association analyses,[5] ARIC conducted race specific imputation of variant dosages to the 1000 Genomes Project Phase I version 3 reference panel. Before imputation, individuals were removed for being first degree relatives, genetic outliers, or not matching existing genotype data. Detailed information on GWAS genotyping, quality control, and imputation procedures has been provided elsewhere.[6] Principal components (PCs) based on the GWAS data were generated by EIGENSTRAT[7] to reflect the population structure or genetic ancestry of the ARIC participants. Of 11 447 European Americans who were at risk

of AAA at baseline in 1987–1989, 8 962 EAs had the GWAS array genotyping and imputation data. Among them, 408 individuals were diagnosed with AAA from 1987 to 1989 through to 2011, with a median of 22 years of follow up for the cohort.

At the end of follow up for clinically diagnosed AAAs, ARIC conducted an ultrasound exam in 2011–2013 among its surviving participants (*n* = 5 911) and identified an additional 75 asymptomatic AAAs. Ultrasound detected AAAs were not included in this study because doing so would require a separate analysis of these AAAs with adjustment for the potential selection bias caused by differential attrition of the cohort members who had the ultrasound, which is underpowered.

The Greek case control study prospectively recruited 418 consecutive AAA patients between July 2009 and December 2012 in a tertiary referral vascular surgery centre in northern Greece and a group of 447 geographically matched controls recruited within the same time interval.[8] An AAA was defined as a maximum antero-posterior diameter of the infrarenal aorta >3.0 cm. The AAA patients, who were referred to a vascular surgeon (on an outpatient basis) for consideration of surgical treatment, either had an incidental diagnosis of AAA while undergoing investigations for other conditions or were referred via a local informal screening program. Although rupture was not an exclusion criterion, none of the AAA patients presented with rupture (given the logistical difficulties with consent). All controls were recruited from the same surgical unit and had cross sectional imaging of their abdomen within four weeks of recruitment to exclude an AAA. Cases and controls were all Caucasians born in northern Greece. Patients with inflammatory aneurysms or connective tissue disorders associated with aneurysm were excluded. A full medical and surgical history was recorded at baseline, including anthropometric assessments and a list of medications. Of the 418 AAAs and 447 controls, 341 AAAs and 292 controls had DNA samples available for genotyping for the current study.

In the Greek case control study, DNA samples were extracted from peripheral venous blood and stored at 4 °C until analysis. The candidate SNPs were identified using LightSNip (TIB MOLBIOL GmbH, Berlin, Germany) on the Roche LightCycler 480 system (Roche Diagnostics Corp, Indianapolis, IN, USA) based on real time polymerase chain reaction and melting curve analysis. The LightCycler 480 instrument was programmed according to the instructions for use for each LightSNip. DNA samples of both AAAs and controls were included in each plate. Clinical data were not available during genotyping. Genotyping for rs3827066 failed because of insufficient quality and quantity of DNA samples. Sixty-four duplicate samples were included in the genotyping, yielding an overall concordance rate of 98.6% between the duplicates. Three AAAs were excluded from all analyses because of questionable allele calls or missing data for more than one of the SNPs. After further exclusion of missingness for each individual SNP, the final sample size was 333 AAAs and 292 controls for rs1795061, 336 and 292 for rs9316871, and 323 and 291 for rs2836411.

In both studies, all participants provided informed consent, and the institutional review board or ethics committee at each institute approved the study. None of the data on the associations between AAA and the four genetic markers in ARIC or the Greek study have been published previously.

Statistical analysis

In ARIC, Cox proportional hazards regression was used to test the association of SNP allele dosages with AAA, with time to event calculated from baseline to the time of first AAA event, date of death, loss to follow up, or through December 31, 2011, whichever came first. HR and 95% CI were computed with adjustment for age, sex, field centre (i.e., study site), and the first five PCs for population stratification to reduce potential confounding by these characteristics.

In the Greek case control study, logistic regression was performed to calculate OR and 95% CI to assess the association between the allele dosage of each SNP and AAA status, with adjustment for age and sex.

In both studies, the at risk allele was chosen to be consistent with the report by Jones et al.[3]

Meta-analysis

To reduce the influence of variation of individual studies and increase statistical precision, a sample size weighted meta-analysis was used to pool p values and direction of genetic effects from the two studies for rs1795061, rs9316871, and rs2836411. The p value based approach was chosen because different study designs were used which yielded different effect estimates (HR vs. OR) in the two studies. The number of AAAs in each study was used as the weight. An I^2 was estimated in a heterogeneity test to assess the difference of test statistics and the direction of association between the two studies. Furthermore, the data from the two studies were pooled with those from Jones et al. to evaluate whether adding the ARIC and Greek data strengthened the overall GWAS associations of Jones et al. The meta-analyses were conducted using the METAL package.[9]

Power analysis

The power was estimated to replicate the associations for the three SNPs reported by Jones et al. in the combined ARIC and Greek samples. At an α of 0.05, there was acceptable power to replicate the reported associations in the pooled ARIC and Greek samples: rs1795061 (MAF 0.34, OR 1.13, power 61%), rs9316871 (0.20, 0.86, 62%), and rs2836411 (0.37, 1.13, 60%).

RESULTS

Table 1 presents the baseline characteristics of the two study populations.

In ARIC, rs2836411 and rs9316871 were genotyped while rs1795061 and rs3827066 were imputed. Imputation quality, reflected by the ratio of observed to expected variance of the dosage statistic, was excellent for these two SNPs (0.98 and 0.99, respectively). In the whole study population, a test of genotype frequencies against Hardy–Weinberg equilibrium was marginally significant for rs2836411 ($p = .05$) and not significant for the other three SNPs. Minor allele frequencies (MAF) for the four SNPs agreed with those in Jones et al.[3] After adjustment for age, sex, field centre, and the first five PCs, rs9316871 and rs3827066 were significantly associated with AAA (HR [95% CI] 0.77

Table 1. Baseline characteristics of patients with and without abdominal aortic aneurysm (AAA) in Atherosclerosis Risk in Communities (ARIC) cohort study (1987–2011) and Greek case cohort study

Baseline characteristics	ARIC cohort study			Greek case control study		
	AAA (n = 408)	Non-AAA (n = 8554)	p	AAA (n = 338[a])	Controls (n = 292[b])	p
Age, years	57 ± 5	55 ± 6	<.001	69 ± 8	73 ± 5	.001
Male	295 (72)	3875 (45)	<.001	310 (92)	219 (75)	.006
Body mass index kg/m²	27 ± 4	27 ± 5	.10	28 ± 3	27 ± 3	.73
Smoking status			<.001			.43
Current smoker	206 (50)	1984 (23)		66 (20)	63 (22)	
Former smoker	151 (37)	3037 (35)		192 (57)	149 (51)	
Never smoker	51 (13)	3533 (41)		80 (24)	80 (27)	
Pack-years smoking[c]	35 ± 24	16 ± 21	<.001	NA	NA	NA
Hypertension	149 (37)	2250 (26)	<.001	260 (77)	226 (77)	.87
Peripheral arterial disease	20 (5)	163 (2)	<.001	71 (21)	51 (17)	.27
Diabetes	24 (6)	744 (9)	.05	75 (22)	68 (23)	.78
Hypercholesterolaemia	147 (36)	2224 (26)	<.001	196 (58)	155 (53)	.20

Data are provided as n (%) or mean ± standard deviation (SD).

AAA = abdominal aortic aneurysm; ARIC = Atherosclerosis Risk in Communities; NA = not assessed or available.

[a] n was 333 AAAs for rs1795061, 336 AAAs for rs9316871, and 323 AAAs for rs2836411.

[b] n was 292 controls for both rs1795061 and rs9316871, and 291 controls for rs2836411.

[c] Including non-smokers.

[0.65−0.92] and 1.22 [1.02, 1.46], respectively, per one copy increase in the effect allele dose), and rs2836411 was close to the threshold for significance (HR [95% CI] 1.13 [0.99−1.31], *p* = .08) (Model 1 in Table 2). Notably, the association for rs9316871 suggests a protective effect associated with the minor allele G. The direction of association for the three SNPs was consistent with that of Jones *et al.*[3] rs1795061 was not associated with AAA in ARIC (*p* = .55) (Model 1 in Table 2).

In the Greek case control study, genotype frequencies for rs1795061, rs9316871, and rs2836411 were in accordance with Hardy−Weinberg equilibrium (p > .05) in the controls. MAF for rs1795061 was lower than that in ARIC and Jones *et al.*, whereas MAF for the other two SNPs was similar to that of ARIC and Jones *et al.* rs1795061 and rs2836411 were significantly associated with the risk of AAA (OR [95% CI] 1.66 [1.27−2.17] and 1.29 [1.01−1.65], respectively) (Model 1 in Table 2), rs9316871 was not associated with AAA (p = .81).

Associations of rs1795061, rs9316871, and rs2836411 with AAA were statistically significant in the meta-analysis of the two studies (Model 1 in Table 2). Heterogeneity (*I²* > 0) was observed for rs1795061 and rs9316871, with that for rs1795061 being statistically significant (p < .05). The direction of association for rs1795061 was not consistent between the two studies. Pooling the data of the two studies with those from Jones *et al.* in a second meta-analysis strengthened the statistical significance of the GWAS associations for the four SNPs (Model 1 in Table 2).

In a secondary analysis, further adjustments were made for two important risk factors for AAA, smoking (pack years of smoking in ARIC and smoking status in the Greek study) and cholesterol (total cholesterol in ARIC and binary hypercholesterolaemia in the Greek study). The additional adjustments did not result in appreciable changes in either individual study or the meta-analyses (Model 2 in Table 2). As a note, 12 AAAs and 99 non-AAAs in ARIC were excluded from this analysis because of missing information on the smoking variable.

DISCUSSION

This study replicated the previously reported new associations of AAA risk with rs9316871 and rs3827066 in the US based prospective study, and with rs1795061 and rs2836411 in the Greek case control study. The direction of association with rs9316871 and rs2836411 was consistent between the two studies and the associations were significant in the meta-analysis combining the two studies. The ARIC data showed the association of rs9316871, rs3827066, and rs2836411 with longitudinal AAA risk based on 24 years of follow up. To the best of the present authors' knowledge, there are few data in the literature reporting longitudinal AAA risk associated with genetic variants in a community based study. While the effect size for the replicated variants was modest, their public health implication is not trivial because their at risk alleles are common in the general population.

Table 2. Association between novel abdominal aortic aneurysm (AAA) single nucleotide polymorphism (SNPs) and AAA in Atherosclerosis Risk in Communities (ARIC) cohort study (1987−2011) and Greek case control study

SNP	Gene	A1	A2	ARIC cohort study			Greek case control study			Meta-analysis 1[b]			Meta-analysis 2[c]		
				Minor allele frequency (MAF)	HR (95% CI)[a]	*p*	MAF	OR (95% CI)[d]	*p*	Dir	*p*	*I²*	Dir	*p*	*I²*
Model 1															
rs1795061	SMYD2	T[e]	C	.31	0.96 (0.82−1.12)	.55	.21	1.66 (1.27−2.17)	1.96×10^{-4}	−+	.04	90[f]	−++	1.1×10^{-11}	80[f]
rs9316871	LINC00540	G[e]	A	.23	0.77 (0.65−0.92)	.004	.25	0.97 (0.75−1.25)	.81	−	.02	68	−	3.8×10^{-11}	42
rs3827066	PCIF1-ZNF335-MMP9	T[e]	C	.16	1.22 (1.02−1.46)	.03	NA	NA	NA	NA	NA	NA	++	2.2×10^{-18}	0
rs2836411	ERG	T[e]	C	.34	1.13 (0.99−1.31)	.08	.33	1.29 (1.01−1.65)	.04	++	.007	0	+++	2.6×10^{-10}	0
Model 2															
rs1795061	SMYD2	T[e]	C	.31	0.95 (0.81−1.11)	.52	.21	1.71 (1.30−2.23)	9.6×10^{-5}	−+	.03	91[f]	−++	8.8×10^{-12}	82[f]
rs9316871	LINC00540	G[e]	A	.23	0.75 (0.63−0.90)	.002	.25	0.97 (0.75−1.26)	.83	−	.01	74	−	3.0×10^{-11}	54
rs3827066	PCIF1-ZNF335-MMP9	T[e]	C	.16	1.20 (1.01−1.43)	.04	NA	NA	NA	NA	NA	NA	++	2.8×10^{-18}	0
rs2836411	ERG	T[e]	C	.34	1.12 (0.97−1.30)	.11	.33	1.27 (0.99−1.62)	.06	++	.01	0.1	+++	3.9×10^{-10}	0

AAA = abdominal aortic aneurysm; A1 = effective allele, A2 = alternative allele; Dir = direction; *I²* = heterogeneity test statistics; MAF = minor allele frequency; NA = not available or applicable; SNP = single nucleotide polymorphism; HR = hazard ratio; OR = odds ratio.

[a] HR were modelled per 1 unit increment in the effect allele dose in Cox regression, with covariable adjustment in Model 1 (adjusted for age, sex, field centre, and the first five principal components for population stratification) or Model 2 (Model 1 further adjusted for pack years of smoking and total cholesterol).

[b] Meta-analysis 1: Atherosclerosis Risk in Communities (ARIC) and the Greek case control study (not for rs3827066).

[c] Meta-analysis 2: ARIC, the Greek case control study (not for rs3827066), and Jones *et al.*[3].

[d] OR were modelled per one unit increment in the effect allele dose in logistic regression, with covariable adjustment in Model 1 (adjusted for age and sex) or Model 2 (Model 1 further adjusted for smoking status and binary hypercholesterolemia).

[e] Minor allele.

[f] *p* < .05, other *I²* without this label had *p* > .05.

By extensive look up in expression quantitative trait locus and RNA sequencing data, Jones *et al.* identified functional relevance for three of the variants: rs3827066 with PLTP expression, rs2836411 with ERG expression, and rs9316871 with FGF9 expression.[3] They also searched for these putative AAA risk variants in published GWAS data of other cardiometabolic phenotypes but did not find associations of the new variants with coronary artery disease (CAD), hypertension, lipid traits, or diabetes. However, a newly published GWAS for CAD identified rs3827066 in *PCIF1-ZNF335-MMP9* as a new locus for CAD (OR 1.04).[10] This new GWAS finding further highlights the overlap between genetic risk factors for AAA and other forms of cardiovascular disease.

In Jones *et al.*, the *p* for rs9316871 in *LINC00540* reached the genome wide significance threshold only when the discovery and replication data were pooled together. Therefore, the replication of this locus in ARIC provides further corroboration that it is related to the risk of AAA. The association at this locus showed modest heterogeneity between ARIC and the Greek study, with consistent direction. Unlike the analysis in ARIC, the Greek study did not control for possible population stratification using ancestry informative markers or principal components. However, the threat of confounding by population stratification is probably small, although cannot be completely ruled out, because both cases and controls were Caucasians who were recruited from and born in a small area of northern Greece.

The association between rs1795061 and AAA risk was replicated in the Greek study but not in ARIC (*p* = .55). In fact, the direction of association in ARIC was opposite to that in the Greek study and the GWAS by Jones *et al.* Imputation error was an unlikely explanation because rs1795061 was imputed with high confidence in ARIC (imputation quality score 0.98). The MAF was similar in ARIC and Jones *et al.* (0.31 vs. 0.34) but lower in the Greek study (0.21 among the controls). The association of this SNP with AAA also showed high heterogeneity across studies in Jones *et al.* As rs1795061 is located at an intergenic region near SMYD2 and GWAS is an indirect mapping strategy that relies on variants that are in linkage disequilibrium (LD) with the true causal variants to capture the genetic associations,[11,12] it is possible that the rs1795061-AAA association reflects the effect of an underlying, causal variant that was not directly measured by the GWAS arrays. If so, the extent of LD between rs1795061 and the underlying causal variant may differ across populations. Another interpretation is that different genetic backgrounds and environmental exposures may have modified the effect of this locus in different populations.

In summary, based on the data from two independent populations of European ancestry, the present study consistently replicated newly reported associations for AAA at *LINC00540* and *ERG*, further supporting the importance of these loci in the aetiology of AAA. Improved understanding of AAA aetiology might benefit the management of AAA and suggest new primary prevention or therapeutic strategies. Future studies in diverse populations are needed to further validate these genetic loci as risk factors for AAA as well as to identify the underlying causal variants and mechanisms for validated associations. Furthermore, with the increasing number of genetic loci and variants that might be identified for AAA in future studies, it may be worth evaluating the efficiency of genetic testing in relatives of AAA patients for the identification of individuals at increased risk of developing AAA. Notably, the potential value of using genomic information in cardiovascular risk prediction has recently been demonstrated in the field, for example, on CAD.[13] Future studies with comprehensive assessment of genomic variants in large populations will be needed for evaluation of clinical application of AAA genomic findings.

CONFLICT OF INTEREST

None.

REFERENCES

1 Wahlgren CM, Larsson E, Magnusson PK, Hultgren R, Swedenborg J. Genetic and environmental contributions to abdominal aortic aneurysm development in a twin population. *J Vasc Surg* 2010;**51**:3–7.

2 Joergensen TM, Christensen K, Lindholt JS, Larsen LA, Green A, Houlind K. Editor's choice - high heritability of liability to abdominal aortic aneurysms: a population based twin study. *Eur J Vasc Endovasc Surg* 2016;**52**:41–6.

3 Jones GT, Tromp G, Kuivaniemi H, Gretarsdottir S, Baas AF, Giusti B, et al. Meta-analysis of genome-wide association studies for abdominal aortic aneurysm identifies four new disease-specific risk loci. *Circ Res* 2017;**120**:341–53.

4 Tang W, Yao L, Roetker NS, Alonso A, Lutsey PL, Steenson CC, et al. Lifetime risk and risk factors for abdominal aortic aneurysm in a 24-year prospective study: the ARIC study (Atherosclerosis risk in Communities). *Arterioscler Thromb Vasc Biol* 2016;**36**:2468–77.

5 Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nat Rev Genet* 2010;**11**:499–511.

6 Pankow JS, Tang W, Pankratz N, Guan W, Weng LC, Cushman M, et al. Identification of genetic variants linking protein C and lipoprotein metabolism: the ARIC study (Atherosclerosis risk in Communities). *Arterioscler Thromb Vasc Biol* 2017;**37**:589–97.

7 Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006;**38**:904–9.

8 Saratzis A, Bown MJ, Wild B, Nightingale P, Smith J, Johnson C, et al. Association between seven single nucleotide polymorphisms involved in inflammation and proteolysis and abdominal aortic aneurysm. *J Vasc Surg* 2015;**61**:1120–1128 e1.

9 Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010;**26**:2190–1.

10 van der Harst P, Verweij N. Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. *Circ Res* 2018;**122**:433–43.

11 Kruglyak L. The road to genome-wide association studies. *Nat Rev Genet* 2008;**9**:314–8.

12 Manolio TA, Brooks LD, Collins FS. A HapMap harvest of insights into the genetics of common disease. *J Clin Invest* 2008;**118**:1590–605.

13 Inouye M, Abraham G, Nelson CP, Wood AM, Sweeting MJ, Dudbridge F, et al. Genomic risk prediction of coronary artery disease in 480,000 adults: implications for primary prevention. *J Am Coll Cardiol* 2018;**72**:1883–93.