Contribution of rare and common genetic variants to plasma lipid levels and carotid stiffness and geometry – a substudy of the Paris Prospective Study 3 Short title: Genetic variants and carotid stiffness

Carole Proust, Jean-Philippe Empana, Pierre Boutouyrie, Maureen Alivon, Pascal Challande, Nicolas Danchin, Guillaume Escriou, Ulrike Esslinger, Stéphane Laurent, Zhenlin Li, et al.

Contribution of rare and common genetic variants to plasma lipid levels and carotid stiffness and geometry – a substudy of the Paris Prospective Study 3

Short title: Genetic variants and carotid stiffness

Carole Proust[1] #, MsC #, Jean-Philippe Empana[2] #, MD, PhD, Pierre Boutouyrie[2] #, MD, PhD, Maureen Alivon[2], PhD, Pascal Challande[3], MD, PhD, Nicolas Danchin[4], MD, Guillaume Escriou[1,2], MsC, Ulrike Esslinger[1], MsC, Stéphane Laurent[2], MD, PhD, Zhenlin Li[5], PhD, Bruno Pannier[6], MD, Veronique Regnault[7], PhD, Frederique Thomas[6], PhD, Xavier Jouven[2] ##, MD, PhD, François Cambien[1] ##, MD, Patrick Lacolley[7] ##, MD, PhD

[#] First co-authors, [##] Equal contribution

[1]Inserm, UMR_S 1166; Sorbonne Universités, UPMC Univ Paris 06, Paris, France

[2]Inserm, UMR_S 970; Sorbonne Paris Cité, Paris, France

[3]Sorbonne Universités, UPMC Univ Paris 06; CNRS UMR 7190, Paris, France

[4]Department of Cardiology, European Hospital of Georges Pompidou; Université Paris Descartes, Paris, France

[5] CNRS, UMR 8256; Sorbonne Universités, UPMC Univ Paris 06, Paris, France

[6]Centre d'Investigations Préventives et Cliniques (IPC Center), Paris, France

[7]Inserm, UMR_S 1116; Université de Lorraine, Nancy, France

Correspondence: Patrick Lacolley, UMR_S 1116, Faculté de Médecine, 9 avenue de la forêt de Haye, CS 50184, 54505 Vandoeuvre-les-Nancy, France. Tel: 33 3 83683621, Fax: 33 3 83683639; mail: patrick.lacolley@inserm.fr

Word count: total count: 6632 words; abstract: 249 words

Journal subject codes: 109, 193, 97

Summary

Background: We assess the contribution of common and rare putatively functional genetic variants (most of them coding) present on the *Illumina exome Beadchip* to the variability of plasma lipids and stiffness of the common carotid artery.

Methods and results: Measurements were obtained from 2283 men and 1398 women, and after filtering and exclusion of monomorphic variants, 32827 common (minor allele frequency >0.01) and 68770 rare variants were analyzed. A large fraction of the heritability of plasma lipids is attributable to variants present on the array, especially for Triglycerides (fraction of variance attributable to measured genotypes: V(G)/Vp=31.4%, $P<3.1\times10^{-11}$) and HDLc (V(G)/Vp=26.4%, $P<4.2\times10^{-12}$). Plasma lipids were associated with common variants located in known candidate genes but no implication of rare variants could be established, however this study had limited power to detect an effect of rare variants at the gene level. Gene-sets for plasma lipids, blood pressure and coronary artery disease were defined on the basis of recent meta-analyses of genome wide association studies (GWAS). We observed a strong association between the plasma lipids gene-set and plasma lipid variables but none of the 3 GWAS gene-sets was associated with the carotid parameters. Significant V(G)/Vp ratios were observed for external (14.5%, $P<2.7\times10^{-5}$) and internal diameter (13.4%, $P<4.3\times10^{-4}$), stiffness (12.5%, $P< 8.0\times10^{-4}$), intima-media thickness (10.6%, $P<7.9\times10^{-4}$) and wall cross sectional area (13.2%, $P<2.4\times10^{-5}$). A significant association was observed between the common rs2903692 polymorphism of the *CLEC16A* gene and the internal diameter ($P<4.3\times10^{-7}$).

Conclusion: These results suggest an involvement of *CLEC16A*, a gene that has been reported to be associated with immune disorders, in the modulation of carotid vasodilatation.

Key Words: exome array; vascular stiffness; heritability; *CLEC16A* gene; plasma lipids

Introduction

Increased central arterial stiffening is a hallmark of the ageing process and the consequence of many disease states such as diabetes, atherosclerosis, and chronic renal disease. Aortic stiffness has independent predictive value for total and cardiovascular (CV) mortality, coronary morbidity and mortality, and fatal

stroke in patients with essential hypertension, end-stage renal failure, or diabetes mellitus.[1-4] Arterial stiffness is defined by a reduction in arterial distensibility. We previously showed that aortic stiffness (carotid-femoral pulse wave velocity) remained a significant predictor of coronary events in hypertensive patients after adjustment for classical CV risk factors.[5] The growing prevalence and associated risk of arterial stiffness provide a major incentive to increase our understanding of the underlying molecular, cellular and genetic causes and the resulting physiological impact of this condition.

Since the initial study showing an association between a common polymorphism of the angiotensin II type 1 receptor gene (1166 A/C) and arterial stiffness parameters,[6] genome-wide linkage studies have identified chromosomal regions associated with arterial stiffness, without relying on any prior biological hypothesis.[7] In addition, gene expression profiling studies have identified novel transcriptional biomarkers of arterial stiffness in humans[8] and variants in genes encoding molecules involved in cytoskeletal organization, vascular smooth muscle cell (VSMC) differentiation or the contractile state of the cell have been suggested to contribute to CV stiffness.[9]

The first genome-wide association study (GWAS) of arterial stiffness was performed using a 100K panel of common single-nucleotide polymorphisms (SNPs).[10] The most recent GWAS with a larger sample size and using more informative arrays have identified a number of new candidate genes, some of them being potentially involved in the pathophysiology of arterial stiffness, in particular collagen, type IV (*COL4A1*),[11] the B-cell CLL/lymphoma 11B (*BCL11B*) gene desert, a transcriptional repressor of various genes that may be relevant to aortic stiffness[12] and adrenomedullin (*ADM*), a potent vasodilator and angiogenic peptide.[13] The recent development of whole-exome arrays targeting mainly coding regions of genes offers great potential for the identification of rare or common variants with important phenotypic effects.

In the present report we investigated for the first time the possible influence of variants present on the *Illumina exome beadchip* on phenotypes related to arterial pathophysiology, i.e. plasma lipids and parameters defining the geometry and distensibility of the carotid artery, in a population-based study.

Material and Methods

Population

The design and objectives of the Paris Prospective Study 3 (PPS3) have been previously described.[14] It is an ongoing observational prospective study evaluating the possible implication of numerous vascular health parameters in CV disease in healthy men and women. Briefly, the PPS3 cohort consists of individuals working or in early retirement, and their families. The participants are recruited on a voluntary basis in one of the largest French preventive health centers, the Centre d'Investigations Preventives et Cliniques (IPC Center) that is subsidized by "the National Health Insurance System for wage earners" (Securite Sociale-CNAMTS). Female subjects represent ~1/3 of the participants and the age of inclusion ranges from 50 to 75 years. The study protocol was approved by the Ethics Committee of the Cochin Hospital (Paris) and, all the participants have signed an informed consent form. Our study is registered in the international trial registry (NCT00741728). Here, we report a substudy in which we have selected 4056 index subjects. After exclusion of samples with > 1% missing genotypes (n=96), outliers and genetically related individuals (n=279, see ***data preprocessing***), 3681 subjects were kept in the analyses.

Measurement of lipid variables

The standard health check-up included a complete clinical examination, coupled with standard biological tests performed after overnight fasting. Total cholesterol and triglycerides were measured using standardized enzymatic methods (automat HITACHI 917); high density lipoprotein cholesterol (HDLc) was measured by direct enzymatic assay with cyclo-dextrin; low density lipoprotein cholesterol (LDLc) was estimated using the Friedwald formula in individuals with plasma triglycerides below 450 mg/dL.

Assessment of carotid geometry and distensibility

Carotid properties were determined using a high resolution echotracking system (Artlab, Esaote, Maastricht, the Netherlands). Briefly, a 10 MHz 128 transducer linear array probe was positioned on the carotid area. Measurements were performed on a 4 cm segment of the right common carotid artery, 1 cm proximal to the bifurcation/sinus throughout the cardiac cycle for 6 seconds. A longitudinal section showing clear interfaces for blood/intima and media/adventitia was obtained. The system allows real time

radiofrequency signal analysis with operator-independent determination of external diameter (Dext), internal diameter (Dint) and intima-media thickness (IMT) on 128 lines throughout the cardiac cycle. Distension was measured on 14 lines at high pulsed radiofrequency (600 Hz). The axial resolution was 34 µm for diameter, 17 µm for IMT and 1.7 µm for distension.[15] Central pressure was estimated from the distension waveform according to van Bortel et al.[16] The distensibility coefficient (DC), representing the elastic properties of the artery as a hollow structure, was calculated as dLCSA/(LCSA×PPC) where LCSA is the lumen cross sectional area and PPC, the central pulse pressure. Carotid stiffness (Cstif) was calculated as $DC^{-0.5}$ and circumferential wall stress as diastolic blood pressure (DBP)×Dint/2×IMT. WCSA is wall cross sectional area.

Genotyping

The *Illumina HumanExome-12v1.1 BeadChip* targets 242,901 markers and was designed to genotype variants with strong predicted functional impact, such as variants altering protein sequences and reported GWAS hits (see http://genome.sph.umich.edu/wiki/Exome_Chip_Design for a description of the array). We used the standard *Infinium® HD Assay Ultra* protocol according to the manufacturer instructions (See ***Appendix I.1*** for details).

Data preprocessing

Genotype calling was done with *Illumina GenomeStudio* software and quality control was performed with the *PLINK* software[17] and in the R environment.[18] Markers with genotyping success rates < 99% (n=2065) and not in Hardy-Weinberg equilibrium ($P<10^{-5}$, n=4325) were not investigated further. Outlier samples identified using the "neighbour" function implemented in *PLINK* were excluded. We used the Genome-wide complex trait analysis software *GCTA*[19] to detect cryptic relatedness among study participants and removed one individual from each pair of related samples (using a genetic relatedness matrix (GRM)-cutoff = 0.1 which corresponds to a moderate level of relatedness). After exclusion of monomorphic variants (n=134914), the final data set included 3681 individuals and 101597 polymorphic markers.

Statistical analysis

Statistical analysis was performed on common and rare variants separately, a threshold of 0.01 for the minor allele frequency (MAF) was adopted to separate rare from common variants.

All phenotypes were log-transformed before analysis and association analyses were adjusted for gender, age, body mass index (BMI) and body surface area (BSA). Adjustment on both BMI and BSA was performed because associations with the 2 body size measurements were dissimilar for some arterial parameters. The correlation between BMI and BSA was 0.625 and the correlations of both measurements with gender were 0.212 and 0.681 respectively. We also tested polynomial adjustments but the results were almost unchanged (not reported).

Associations between phenotypes and common variants (variant-level analysis) assuming an additive mode of inheritance were explored using a mixed-linear model (MLM) implemented in the *GCTA* software[19] accounting for the remaining genetic relatedness among participants and covariates (age, gender, BMI and BSA). For each phenotype investigated, we estimated the fraction of the phenotypic variance attributable to measured genotypes (V(G)/Vp) using the MLM approach implemented in *GCTA*[20] and also tested a possible heterogeneity of these associations according to gender. Note that in our analysis we used the leave-one-chromosome-out analysis implemented in *GCTA* software to avoid adjusting on variants that are in LD with the variant tested.

At the gene and gene-set levels, associations with the lipid and vascular traits were investigated using the Sequence Kernel Association Tests (*SKAT*) implemented in the R package *SKAT* (*R/SKAT*) for both rare and common variants. All tests were adjusted on age, gender, BMI, BSA and first 10 principal components to account for population structure by entering these covariates in the *SKAT* null model. The *SKAT* test allows for both effect increasing and effect decreasing variants, this appears appropriate in the context of this study in which tested variants were collected independently of the phenotypes investigated, nevertheless we also conducted a one-sided collapsing test to account for the possibility that most variants in a gene or a gene-set affect a phenotype in the same direction (see *Appendix I.2.* for details).

For each trait, association P-values were summarized by a Manhattan plot (*R/qqman* package) and the fit of observed to expected association chi-square was plotted as QQ-plots (*R/snpStats* package).

To account for multiple testing, we applied a Bonferroni correction for the number of variants tested in the variant-level analysis of common variants (n=32827) and for the number of genes tested in the gene-level analysis of rare and common variants (n=13453 genes harboring at least 2 variants). These corrections are conservative since they do not account for the linkage disequilibrium (LD) existing among variants in the same gene and nearby genes. To further account for the interdependent vascular phenotypes, we applied a Bonferroni correction for the 3 main phenotype categories: plasma lipids, carotid stiffness and carotid geometry. After these corrections, the significance thresholds for common variants and rare variants were set to 5×10^{-7} and to 1×10^{-6} respectively.

Power of the study. The power of the variant level analysis for various MAF (>0.01) and allele effect size is reported in ***supplementary Table 1.*** We note that given our study size, for a MAF of 0.01, effect size must be at least equal to 0.7 standard deviation (SD) unit to reach a power of 0.8. For a MAF of 0.1, a similar power is attained for an effect size of 0.23 SD unit. For sets of variants, power estimation requires modeling the genetic architecture (number, respective frequencies and effects of variants, LD between variants). The *R/SKAT* package provides an analytic method to compute power for SKAT (See ***Supplementary Methods I.3***). Given the relatively small number of rare variants in our study the power of the gene-level analysis was low for most genes. On the other hand the power of the gene-set level analysis was more appropriate (***Supplementary Table 2***).

Results

3681 individuals (2283 men and 1398 women) contributed to the analysis. The mean values of the relevant characteristics are reported in ***Table 1*** for men and women separately. The results of association analyses were very similar in men and women, as a consequence we report results for both genders combined. Correlations among vascular parameters are shown in ***Table 2*** before (above the diagonal) and after (below the diagonal) adjustment on covariables (age, gender, BMI and BSA).

Heritability of investigated traits attributable to variants on the array

The ratio of the genetic variance to the total variance (V(G)/Vp) of lipid and vascular phenotypes are reported in ***Table 3***; these ratios correspond to the heritability attributable to the 101597 variants available. The ratios for LDLc, HDLc and triglycerides were 10.5% ($P < 0.0081$), 26.4% ($P < 4.2\times10^{-12}$) and 31.4% ($P < 3.1\times10^{-11}$) respectively. The V(G)/Vp ratios for blood pressure related variables (MBP and PPC) were not significant; on the other hand the ratios for Cstif (12.5%, $P < 0.0008$), Dext (14.5%, $P < 2.7\times10^{-5}$), Dint (13.4%, $P < 0.00043$), IMT (10.6%, $P < 0.00078$) and WCSA (13.2%, $P < 2.4\times10^{-5}$) were significantly different from zero; however, the large standard errors (***Table 3***), imply that the heritability estimates are rather imprecise. To better characterize the source of this heritability, genetic effects were investigated at 3 levels: i) the variant level, for common variants (MAF > 0.01), ii) the gene level, for common and rare variants (MAF $\leq$ 0.01) and iii) the gene-set level, with 3 gene-sets, each combining candidate genes identified in GWAS of plasma lipids, coronary artery disease (CAD) and blood pressure, respectively.

Variant level analysis

The variant-level analysis was conducted on the 32827 common variants. Results of the association analysis with the MLM adjusted for age, gender, BMI and BSA are reported for each phenotype as a Manhattan and QQ plot (see ***Appendix, section II***); the most significant associations at each associated locus are shown in ***Table 4*** and a complete report of all associated SNPs is provided in ***supplementary Table 3.*** The highest estimated lambda statistics for all phenotypes was less than 1.05 (see QQ-plots in ***Appendix section II***), suggesting that population stratification if any was well controlled in the MLM model and was unlikely to affect the association results in a significant way.

The QQ and Manhattan plots (***Appendix section II***) reveal the presence of numerous associations with lipid variables. Significant associations were observed for LDLc with variants at the *APOE/TOMM40/APOC1, PCSK9 and SNAPC2* loci, for HDLc, with variants at the *CETP, LPL, ABCA1, SNAPC2* and *LIPC* loci, and for triglycerides with variants at the *GCKR, APOA5, LPL*, and *TRIB1* loci. These associations are consistent with results reported in large scale GWAS of plasma lipids.[22] Note that the exm1417699 variant

(rs116635738) in the *SNAPC2* gene is located 445798 bp from the rs7255436 intronic SNP in the *ANGPTL4* gene, which is a GWAS-identified locus for HDLc.[22] However, rs7255436 was available in our data set but was not associated with any of the lipid variables ($P > 0.05$).

For arterial phenotypes, a significant association ($P < 4.2\times10^{-7}$) was observed between carotid Dint and rs2903692 at the *CLEC16A* locus. The minor allele (MAF = 0.39) of this variant was associated with a reduced Dint value. rs2903692 was also associated with Dext ($P < 4.5\times10^{-6}$) but not with WCSA ($P < 0.04$) and IMT ($P < 0.9$), suggesting that it is related to vasodilation rather than to wall hypertrophy. Two other common variants of *CLEC16A*, rs12708716 and rs12924729, in tight LD with rs2903692 exhibited slightly weaker associations with Dint, $P < 2.2\times10^{-6}$ and 3.4×10^{-6} respectively. No significant association was found with stiffness parameters (Cstif, CWS, WCSA).

Gene level analysis

The gene-level analysis was conducted with *SKAT* for genes harboring at least two variants (rare or common) present in our data set (n=13453 genes). QQ-plots and tables reporting gene-phenotype associations with P-values $<10^{-4}$ are provided in the ***Appendix section III***. Although this analysis was more specifically focused on the analysis of rare variants, the results of the joint analysis of all and common variants are also reported in ***supplementary Table 4***. The gene-level analysis of common variants revealed significant associations ($P < 10^{-6}$) implicating lipid variables but all associated loci (*APOE*, *TOMM40*, *LPL*, *APOA5*, *GCKR*, *CETP*, *LPL* and *SNAPC2*) had already been identified in the variant-level analysis. No significant gene-level effect of common variants on vascular variables was observed. The results of the *SKAT* and collapsing tests for rare variants are reported in ***supplementary Table 4***. Overall for both lipid and cardiovascular phenotypes, the gene-level analysis did not reveal any association reaching the significance threshold of 10^{-6}. It must be considered however that due to the large number of monomorphic variants, relatively few rare variants were available for most investigated genes (see ***supplementary Table 4***) and as a consequence the study had limited power to detect an effect of rare variants at the gene level. To try overcome the lack of statistical power resulting from the separate investigation of each gene, we

extended the analysis to sets of genes whose genetic variability is known to affect lipid and cardiovascular phenotypes.

Gene-set level analysis

We defined candidate gene-sets based on the most recent GWAS reports for plasma lipids,[22] coronary artery disease[23] and blood pressure/hypertension[24,25]. All variants available in our study and located within or close to the sequence of the genes in each gene-set were identified and analyzed with *SKAT* to assess their contribution to the variability of the traits investigated. The genes included in the sets and the number of variants are reported in the ***Appendix section IV*** and the results of the *SKAT* analysis are shown in ***supplementary Tables 5-7*** for all, common and rare variants respectively. Analysis of the blood pressure gene-set did not reveal any significant association with the lipid or the vascular phenotypes. As expected, the plasma lipid gene-set was associated with the plasma lipid variables, but this association was only attributable to the common variants. No association of this gene-set with the vascular phenotypes was observed. Associations with the CAD gene-set were significant for LDLc and triglycerides, but again these associations were observed for the common variants only.

To assess whether common variants significantly associated with plasma lipids in the variant level analysis could account for the associations observed with the lipid set, we excluded 38 common variants associated with plasma lipids in the variant-level analysis ($P<0.0001$) and performed the *SKAT* analysis again using the reduced plasma lipid gene-set. Associations of common variants in this gene-set with LDLc ($P < 9.5\times10^{-5}$), HDLc ($P < 2.3\times10^{-5}$) and triglycerides ($P < 1.4\times10^{-4}$) were reduced but remained significant, indicating that common variants with weak effects present in the lipid gene-set but non-identified in the variant level analysis contribute to plasma lipid levels. In ***supplementary Tables 5-7***, we also report the results of the gene-set level analysis limited to variants with a high putative severity score as defined by the Combined-Annotation-Dependent Depletion (CADD) score. However, this analysis did not reveal any association with the rare variants in the 3 gene sets.

Discussion

The exome array with its enrichment in rare variants provides a unique tool to investigate the contribution of genetic variants affecting protein structure to various quantitative traits and diseases in the human population. Compared to current genome-wide arrays, it provides a much reduced genome coverage of < 10% of variants with a MAF > 1% (based on a $r^2 > 0.8$). The great advantage of the exome array is therefore not related to its genome-wide tagging ability but to its enrichment in putatively functional variants and therefore to the increased possibility that an associated variant is causal in comparison to a random marker.

Rare versus common variants. To define rare variants, it has been proposed to adopt a threshold of 1/sqrt(2*n) where n is the study sample size. In our study this value was 0.0116, so we adopted a threshold for rare variants of 0.01. For the gene and set-level analyses we used the default parameters in *R/SKAT* which up-weight rare variants relative to common ones. However we also performed the analyses without differentially weighting rare and common variants (see ***supplementary methods I.2***), the results were very similar.

Plasma lipids. The heritability of plasma lipids is close to 50%. The last meta-analysis of GWAS of plasma lipids conducted in 188,578 European ancestry individuals and 7,898 non-European ancestry individuals has confirmed 97 previously reported loci and identified 62 new ones.[22] In our much smaller study mainly focused on coding variants, we identified only a small subset of these loci, but the V(G)/Vp ratios for the lipid phenotypes indicate that the variants present on the exome array may contribute to > 50% of the heritability of HDLc and triglycerides.

Blood pressure. In the Framingham Heart Study, heritability of mean arterial pressure, and pulse pressure were respectively 0.33 and 0.50.[26] In sharp contrast in our study, the fraction of phenotypic variance of MBP and PPC attributable to variants on the genome array was non-significant. In addition variants extracted from the blood pressure gene-set genotyped in our study were not associated with MBP or PPC. Most variants on the exome array were chosen because they are coding and affect protein sequence. We

may therefore hypothesize that variants contributing to MBP and PPC do not belong to this category and are more likely to be regulatory.

Arterial stiffness. In the Framingham Heart Study, heritability of carotid-femoral pulse wave velocity (a measure of arterial stiffness) was 0.40.[26] In the PPS3, the V(G)/Vp ratio for carotid stiffness (Cstif) was 12.5% ($P < 0.0008$) . However the association analysis of Cstif at the variant, gene or gene-set level did not reveal any significant association, suggesting that the variants affecting heritability of carotid stiffness may exert a too weak individual effect to be detectable in our study. It must also be considered that PWV is directly measured whereas Cstif is calculated from measures of diameter, distension, and pulse pressure, raising the possibility that cumulated measurement error may explain our negative result.

Carotid geometry. In the PPS3, carotid geometry was assessed via IMT, Dint, Dext and WCSA. All 4 parameters were characterized by a V(G)/Vp ratios differing from zero. In the Framingham Heart Study,[27] adjusted heritability was 0.38 for the common carotid artery IMT ($P<0.001$) and 0.35 for the internal carotid artery IMT ($P<0.001$). In our study the heritability of the common carotid artery IMT attributable to the variants on the exome array was 10.6% ($P<0.0008$). In a meta-analysis of GWAS conducted by the CHARGE consortium in over 40,000 participants of European ancestry, common variants associated with carotid IMT were identified in 4 regions of the genome, near *ZHX2*, *APOC1*, *PINX1* and *SLC17A4*.[28] The strongest association was with the rs445925, located at position 19: 45415640 near the *APOC1* gene. This variant was present in our data-set but no association with carotid IMT could be detected ($P > 0.05$). The other lead-SNPs identified in the meta-analysis of GWAS of carotid IMT were not genotyped in our study, however, no variant at the respective loci was associated with the trait.

In a genome wide linkage analysis in 3300 American Indian participants in the Strong Heart Family Study (SHFS), heritability estimates for carotid artery lumen diameters ranged from 0.29 to 0.45 across field centers. In the PPS3 study, Dext was the carotid parameter showing the largest V(G)/Vp ratio (14.5%, $P < 2.7\times10^{-5}$). In the SHFS, significant evidence for a locus influencing carotid artery lumen diameter on chromosome 7q was reported,[29] the authors proposed *KCND2* as a possible candidate at this locus. In the

PPS3, an intronic variant located at position 7:120323727 within the *KCND2* gene sequence was not associated with carotid diameter and there was no evidence for an association of Dint with variants located within 5MB on both sides of *KCND2* (>100 variants, minimum *P* value = 0.014).

Dint was associated with common variants on the *CLEC16A* sequence and the strongest association was observed for rs2903692 ($P < 4.21\times10^{-7}$). Similar but slightly less significant associations were observed for Dext. rs2903692 is located in intron 22 of *CLEC16A* at position 16:11144926. The 3 intronic variants on the *CLEC16A* sequence were added to the exome array because they have been reported to be associated with several immunity-related diseases,[30] particularly type 1 diabetes (T1D) and Multiple Sclerosis. rs2903692 is located at a distance of 58910 base pairs from rs12708716, the lead SNP associated with T1D and both SNPs are in tight LD ($R^2 = 0.96$). The mechanism of implication of *CLEC16A* in immune diseases is not clearly established[30, 31] and it is possible that other genes in the region are responsible for the association of this locus with immune diseases. It is interesting that Chromosome Conformation Capture (3C) studies have demonstrated physical proximity of the promoter of the nearby *DEXI* gene and intron 19 of *CLEC16A*. This intronic region, where several of the disease-associated variants are located, is highly enriched in transcription binding sites.[32] In the same report, it was shown that alleles of *CLEC16A* that confer protection from type 1 diabetes and multiple sclerosis were associated with increased expression of *DEXI* in 2 human monocyte data-sets leading to the conclusion that *DEXI* may be an unappreciated autoimmune candidate gene. In addition to *DEXI*, 2 nearby genes *SOCS1* and *CIITA* are other possible immunity and inflammation-related candidates. In our study, we observed an association with the diameters but not with the other carotid parameters, including IMT. The internal and external diameters of the common carotid artery are very stable and are usually considered as structural parameters which mostly depend on the extracellular matrix content of the vascular wall. However, they may be affected by physiological or pharmacological changes in vasomotor tone via endothelium-related mechanisms. If the influence of *CLEC16A* on arterial diameter was confirmed, the possible implication of a vasomotor mechanism would require investigation. It should be noted that in the SHFS, no linkage of carotid diameter with markers located in the region encompassing *CLEC16A* on chromosome 16 was observed.[29]

Our results concur with those of other reports[33, 34] in showing that rare coding variants, identified in a healthy population and included on the exome array, exert little if any effect on the traits investigated. This may be related to an insufficient power since the published studies in which this array was used are of limited sample size compared to the most recent meta-analysis of common variants. A much larger sample size may be needed[35] to assess the impact of rare variants on vascular phenotypes.

Conclusion

In this study we show that coding and GWAS-identified variants present on the *Illumina* exome array contribute in a significant way to the heritability of plasma lipids and of parameters assessing the geometry of carotid arteries. On the other hand, they exert no influence on blood pressure and carotid stiffness. Our results replicate a number of previously reported associations with plasma lipids involving common variants, but rare variants analyzed at the gene or gene-set levels did not significantly influence the traits studied here. The only significant association involving vascular parameters concerned carotid internal diameter. The candidate region implicated on chromosome 16 encompasses the *CIITA/DEXI/CLEC16A/SOCS1* genes and is well-known for its associations with several immune disorders. If this association was confirmed, it would suggest an implication of immune-related genes in carotid vasomotricity.

Acknowledgments

We thank Erwan Bozec for providing the technicians with echotracking training, Dr. MF Eprinchard, Dr. JM Kirzin and all the medical and technical staff of the IPC Center, the Centre de Ressources Biologiques de l'Hôpital Européen Georges Pompidou staff (C. de Toma, B. Vedie), and the Platform for Biological Resources (PRB) of the Hôpital Européen Georges Pompidou for the management of the biobank.

The PPS3 is organized under an agreement between INSERM and the IPC Center, and between INSERM and the Biological Research Center at the Européen Georges Pompidou Hospital, Paris, France.

We thank the Caisse Nationale d'Assurance Maladie des Travailleurs Salariés (CNAM-TS, France) and the Caisse Primaire d'Assurance Maladie de Paris (CPAM-P, France) for helping make this study possible. We thank Mary Osborne-Pellegrin for help in manuscript editing.

Funding Sources

The PPS3 Study was funded by grants from The National Research Agency (ANR), the Research Foundation for Hypertension (FRHTA), the Research Institute in Public Health (IRESP) and the Region Ile de France (Domaine d'Intérêt Majeur).

This PPS3 substudy was supported by the Agence Nationale de la Recherche (ANR-09-GENO-010), the Region Lorraine and the Communauté Urbaine du Grand Nancy.

Disclosures

None.

References

1. Benetos A, Adamopoulos C, Bureau JM, Temmar M, Labat C, Bean K et al. Determinants of accelerated progression of arterial stiffness in normotensive subjects and in treated hypertensive subjects over a 6-year period. Circulation. 2002; 105:1202-1207.

2. Laurent S, Boutouyrie P, Lacolley P. Structural and genetic bases of arterial stiffness. Hypertension. 2005; 45:1050-1055.

3. Laurent S, Cockcroft J, Van Bortel L, Boutouyrie P, Giannattasio C, Hayoz D et al. Expert consensus document on arterial stiffness: methodological issues and clinical applications. Eur Heart J. 2006; 27:2588-2605.

4. Zieman SJ, Melenovsky V, Kass DA. Mechanisms, pathophysiology, and therapy of arterial stiffness. Arterioscler Thromb Vasc Biol. 2005; 25:932-943.

5. Boutouyrie P, Tropeano AI, Asmar R, Gautier I, Benetos A, Lacolley P et al. Aortic stiffness is an independent predictor of primary coronary events in hypertensive patients: a longitudinal study. Hypertension. 2002; 39:10-15.

6. Benetos A, Topouchian J, Ricard S, Gautier S, Bonnardeaux A, Asmar R et al. Influence of angiotensin II type 1 receptor polymorphism on aortic stiffness in never-treated hypertensive patients. Hypertension. 1995; 26:44-47.

7. Lacolley P, Challande P, Osborne-Pellegrin M, Regnault V. Genetics and pathophysiology of arterial stiffness. Cardiovasc Res. 2009; 81:637-648.

8. Fassot C, Briet M, Rostagno P, Barbry P, Perret C, Laude D et al. Accelerated arterial stiffening and gene expression profile of the aorta in patients with coronary artery disease. J Hypertens. 2008; 26:747-757.

9. Galmiche G, Labat C, Mericskay M, Aissa KA, Blanc J, Retailleau K et al. Inactivation of serum response factor contributes to decrease vascular muscular tone and arterial stiffness in mice. Circ Res. 2013; 112:1035-1045.

10. Levy D, Larson MG, Benjamin EJ, Newton-Cheh C, Wang TJ, Hwang SJ et al. Framingham Heart Study 100K Project: genome-wide associations for blood pressure and arterial stiffness. BMC Med Genet. 2007; 8 Suppl 1:S3.

11. Tarasov KV, Sanna S, Scuteri A, Strait JB, Orru M, Parsa A et al. COL4A1 is associated with arterial stiffness by genome-wide association scan. Circ Cardiovasc Genet. 2009; 2:151-158.

12. Mitchell GF, Verwoert GC, Tarasov KV, Isaacs A, Smith AV, Yasmin et al. Common genetic variation in the 3'-BCL11B gene desert is associated with carotid-femoral pulse wave velocity and excess cardiovascular disease risk: the AortaGen Consortium. Circ Cardiovasc Genet. 2012; 5:81-90.

13. Beygui F, Wild PS, Zeller T, Germain M, Castagne R, Lackner KJ et al. Adrenomedullin and Arterial Stiffness: An Integrative Approach Combining Monocyte ADM Expression, Plasma MR-proADM and Genome-Wide Association Study. Circ Cardiovasc Genet. 2014;7:634-641.

14. Empana JP, Bean K, Guibout C, Thomas F, Bingham A, Pannier B et al. Paris Prospective Study III: a study of novel heart rate parameters, baroreflex sensitivity and risk of sudden death. Eur J Epidemiol. 2011; 26:887-892.

15. Meinders JM, Brands PJ, Willigers JM, Kornet L, Hoeks AP. Assessment of the spatial homogeneity of artery dimension parameters with high frame rate 2-D B-mode. Ultrasound Med Biol. 2001; 27:785-794.

16. Van Bortel LM, Balkestein EJ, van der Heijden-Spek JJ, Vanmolkot FH, Staessen JA, Kragten JA et al. Non-invasive assessment of local arterial pulse pressure: comparison of applanation tonometry and echo-tracking. J Hypertens. 2001; 19:1037-1044.

17. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559-575.

18. Team TRC. R: A language and environment for statistical computing [Internet]. R Foundation for Statistical Comptuing. 2013; Available from: http://www.R-project.org/.

19. Yang J, Lee SH, Goddard ME, Visscher PM. Genome-wide complex trait analysis (GCTA): methods, data analyses, and interpretations. Methods Mol Biol. 2013; 1019:215-236.

20. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR et al. Common SNPs explain a large proportion of the heritability for human height. Nat Genet. 2010; 42:565-569.

21. Ionita-Laza I, Lee S, Makarov V, Buxbaum JD, Lin X. Sequence kernel association tests for the combined effect of rare and common variants. Am J Hum Genet. 2013; 92:841-853.

22. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S et al. Discovery and refinement of loci associated with lipid levels. Nat Genet. 2013; 45:1274-1283.

23. Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR et al. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet. 2013; 45:25-33.

24. Tragante V, Barnes MR, Ganesh SK, Lanktree MB, Guo W, Franceschini N et al. Gene-centric meta-analysis in 87,736 individuals of European ancestry identifies multiple blood-pressure-related loci. Am J Hum Genet. 2014; 94:349-360.

25. Munroe PB, Barnes MR, Caulfield MJ. Advances in blood pressure genomics. Circ Res. 2013; 112:1365-1379.

26. Mitchell GF, DeStefano AL, Larson MG, Benjamin EJ, Chen MH, Vasan RS et al. Heritability and a genome-wide linkage scan for arterial stiffness, wave reflection, and mean arterial pressure: the Framingham Heart Study. Circulation. 2005; 112:194-199.

27. Fox CS, Polak JF, Chazaro I, Cupples A, Wolf PA, D'Agostino RA et al. Genetic and environmental contributions to atherosclerosis phenotypes in men and women: heritability of carotid intima-media thickness in the Framingham Heart Study. Stroke. 2003; 34:397-401.

28. Bis JC, Kavousi M, Franceschini N, Isaacs A, Abecasis GR, Schminke U et al. Meta-analysis of genome-wide association studies from the CHARGE consortium identifies common variants associated with carotid intima media thickness and plaque. Nat Genet. 2013; 43:940-947.

29. Bella JN, Cole SA, Laston S, Almasy L, Comuzzie A, Lee ET et al. Genome-wide linkage analysis of carotid artery lumen diameter: the strong heart family study. Int J Cardiol. 2013; 168:3902-3908.

30. Berge T, Leikfoss IS, Harbo HF. From Identification to Characterization of the Multiple Sclerosis Susceptibility Gene *CLEC16A*. Int J Mol Sci. 2013; 14:4476-4497.

31. Zouk H, D'Hennezel E, Du X, Ounissi-Benkalha H, Piccirillo CA, Polychronakos C. Functional evaluation of the role of C-type lectin domain family 16A at the chromosome 16p13 locus. Clin Exp Immunol. 2014; 175:485-497.

32. Davison LJ, Wallace C, Cooper JD, Cope NF, Wilson NK, Smyth DJ et al. Long-range DNA looping and gene expression analyses identify DEXI as an autoimmune disease candidate gene. Hum Mol Genet. 2012; 21:322-333.

33. Chen MM, Crous-Bou M, Setiawan VW, Prescott J, Olson SH, Wentzensen N et al. Exome-wide association study of endometrial cancer in a multiethnic population. PLoS One. 2014; 9:e97045.

34. Holmen OL, Zhang H, Zhou W, Schmidt E, Hovelson DH, Langhammer A et al. No large-effect low-frequency coding variation found for myocardial infarction. Hum Mol Genet. 2014; 23:4721-4728.

35. Huyghe JR, Jackson AU, Fogarty MP, Buchkovich ML, Stancakova A, Stringham HM et al. Exome array analysis identifies new loci and low-frequency variants influencing insulin processing and secretion. Nat Genet. 2013; 45:197-201.

Table 1. Characteristics of the population

		Mean±SD or Frequency (%)		
		Whole population	**Women N=1398**	**Men N=2283**
Age (years)		59.9±6.4	60.9±6.4	59.3±6.3
Past history of CVD:		64 (2%)	12 (1%)	52 (2%)
Body height (m)		1.69±0.09	1.61±0.06	1.74±0.07
Body weight (kg)		73±14	63±11	79±11
Body mass index (kg/m²)	BMI	25.3±3.7	24.3±4.2	25.9±3.2
Body surface area (m²)	BSA	1.85±0.21	1.67±0.16	1.95±0.16
Systolic blood pressure (mmHg)	SBP	131±16	129±18	133±16
Diastolic blood pressure (mmHg)	DBP	76±10	74±10	77±10
Antihypertensive drugs n (%)		520 (14%)	197 (14%)	323 (14%)
Lipid lowering drugs n (%)		494 (13%)	179 (13%)	315 (14%)
Total Cholesterol (mg/dL)		220±36	229±36	216±35
HDL cholesterol (mg/dL)	HDLc	59±15	67±16	54±12
LDL cholesterol (mg/dL)	LDLc	141±32	143±32	140±32
Triglycerides (mg/dL)	TRIG	103±50	92±41	110±54
Vascular parameters				
Carotid stiffness (m/s)	Cstif	7.0±1.4	7.1±1.4	7.0±1.3
External diastolic diameter (mm)	Dext	7.18±0.72	6.85±0.63	7.38±0.69
Internal diameter (mm)	Dint	5.89±0.65	5.57±0.57	6.09±0.62
Intima media thickness (mm)	IMT	0.64±0.12	0.64±0.11	0.65±0.13
Mean blood pressure (mmHg)	MBP	94±11	92±11	96±11
Central pulse pressure (mmHg)	PPC	45±11	45±12	45±10
Circumferential wall stress (kPa)	CWS	48±12	44±10	50±12
Wall cross sectional area (mm^2)	WCSA	13.3±3.2	12.5±2.7	13.7±3.4

Table 2. Correlations of vascular parameters

	Cstif	**Dext**	**Dint**	**IMT**	**MBP**	**PPC**	**CWS**	**WCSA**
Cstif	1	0.22	0.2	0.13	0.36	0.48	0.09	0.19
Dext	0.15	1	0.95	0.44	0.29	0.25	0.22	0.78
Dint	0.16	0.94	1	0.12	0.26	0.19	0.47	0.54
IMT	0.00039	0.37	0.016	1	0.17	0.25	-0.62	0.9
MBP	0.33	0.21	0.18	0.11	1	0.49	0.47	0.26
PPC	0.41	0.16	0.13	0.13	0.47	1	-0.04	0.29
CWS	0.16	0.21	0.47	-0.65	0.49	0.014	1	-0.33
WCSA	0.068	0.73	0.45	0.9	0.18	0.17	-0.38	1

Cstif, carotid stiffness ; Dext, external diastolic diameter; Dint, internal diameter; IMT, intima media thickness; MBP, mean blood pressure; PPC, central pulse pressure; CWS, Circumferential wall stress; WCSA, Wall cross sectional area. Correlations among vascular parameters are reported without adjustment (above diagonal) and after adjustment for age, gender, BMI and BSA (below diagonal). $P<0.001$ for all correlations with absolute values < 0.05.

Table 3. Fraction of phenotypic variance explained by all variants

Variable	V(G) ±se	V(e) ±se	Vp ±se	V(G)/Vp ±se	P-value
LDLc	0.0059±0.0027	0.0500±0.0028	0.0559 ± 0.0013	0.1048 ± 0.0477	0.0081
TRIG	0.047±0.008	0.102±0.008	0.148±0.004	0.314±0.052	3.13×10^{-11}
HDLc	0.013±0.003	0.037±0.003	0.050±0.001	0.264±0.051	4.25×10^{-12}
Cstif	0.0044±0.0016	0.0304±0.0017	0.0348±0.0008	0.1254±0.0464	0.00080
Dext	0.0010±0.0003	0.0059±0.0003	0.0069±0.0002	0.1448±0.0460	2.70×10^{-5}
Dint	0.0012±0.0004	0.0079±0.0005	0.0091±0.0002	0.1334±0.0470	0.00043
IMT	0.0029±0.0012	0.0239±0.0013	0.0268±0.0006	0.1065±0.0437	0.00078
MBP	0.0000±0.0005	0.0130±0.0006	0.0130±0.0003	0.0000±0.0420	0.5
PPC	0.0009±0.0020	0.0477±0.0023	0.0485±0.0011	0.0176±0.0417	0.35
CWS	0.0013±0.0021	0.0503±0.0024	0.0516±0.0012	0.0258±0.0410	0.26
WCSA	0.0053±0.0018	0.0346±0.0019	0.0398±0.0009	0.1318±0.0441	2.39×10^{-5}

LDLc, LDL cholesterol; TRIG, triglycerides; HDLc, HDL cholesterol; Cstif, carotid stiffness ; Dext, external diastolic diameter; Dint, internal diameter; IMT, intima media thickness; MBP, mean blood pressure; PPC, central pulse pressure; CWS, Circumferential wall stress; WCSA, Wall cross sectional area.

Variance components were estimated with the MLM implemented in *GCTA*. V(G), variance attributable to measured genotypes; Vp, phenotypic variance; V(e), variance not attributable to measured genotypes.

Table 4. Variant-level analysis - significant associations ($P < 5\times10^{-7}$)

	Chr	position	A1	A2	MAF	b	se	P-value	locus
LDLc									
exm1479366	19	45412079	A	G	0.0762	-0.139	0.0104	2.7×10^{-40}	*APOE*
exm62588	1	55505647	A	C	0.0177	-0.121	0.0211	1.2×10^{-8}	*PCSK9*
exm1417699	19	7987428	A	G	0.0622	0.0595	0.0118	4.5×10^{-7}	*SNAPC2*
Triglycerides									
exm181733	2	27730940	A	G	0.458	0.0643	0.00931	4.9×10^{-12}	*GCKR*
rs964184	11	116648917	C	G	0.137	0.0912	0.0136	2.1×10^{-11}	*APOA5*
rs12678919	8	19844222	G	A	0.110	-0.0992	0.0149	3.0×10^{-11}	*LPL*
rs2954029	8	126490972	T	A	0.456	-0.0504	0.00927	5.2×10^{-8}	*TRIB1*
HDLc									
rs247616	16	56989590	A	G	0.310	0.0566	0.00567	1.6×10^{-23}	*CETP*
rs12678919	8	19844222	G	A	0.110	0.0528	0.00837	2.9×10^{-10}	*LPL*
rs1883025	9	107664301	A	G	0.271	-0.0303	0.00587	2.5×10^{-7}	*ABCA1*
exm1417699	19	7987428	A	G	0.0622	-0.0572	0.0112	3.4×10^{-7}	*SNAPC2* [£]
rs261334	15	58726744	G	C	0.213	0.0311	0.00624	6.3×10^{-7}	*LIPC* [$]
Dint									
rs2903692	16	11238783	A	G	0.39	-0.0115	0.00227	4.2×10^{-7}	*CLEC16A*

A1, reference allele (the coded effect allele); A2, the other allele; MAF, frequency of the reference allele; b, SNP effect; se, standard error. A single significant variant is shown at each locus, a more complete list of associated variants up to a P-value of 0.0001 is reported in the ***Appendix supplementary Table 1***.

[£] the exm1417699 variant is located 440745 bp upstream of the *ANGPTL4* gene which is a locus for HDLc identified.

[$] *LIPC* is an obvious candidate, the P-value for rs261334 is just above significance.

Table 5. Gene-set level analysis – results: P values

	Lipids gene set			**Blood Pressure gene set**			**CAD gene set**		
Variants	All (n=902)	Common (n=339)	Rare (n=563)	All (n=563)	Common (n=206)	Rare (n=357)	All (n=552)	Common (n=226)	Rare (n=326)
LDLc	$2.8x10^{-12}$	$1.7x10^{-17}$	0.90	0.27	0.48	0.19	0.00022	$3.1x10^{-05}$	0.42
TRIG	$6.0x10^{-22}$	$9.3x10^{-23}$	0.012	0.073	0.045	0.35	$8.2x10^{-05}$	$1.3x10^{-05}$	0.38
HDLc	$3.7x10^{-38}$	$4.9x10^{-40}$	0.030	0.42	0.15	0.79	0.20	0.23	0.28
Cstif	0.47	0.42	0.52	0.28	0.59	0.14	0.61	0.13	0.97
Dext	0.060	0.34	0.030	0.19	0.26	0.25	0.016	0.0069	0.33
Dint	0.056	0.23	0.052	0.38	0.55	0.27	0.083	0.019	0.61
IMT	0.75	0.79	0.55	0.29	0.50	0.20	0.17	0.25	0.22
MBP	0.86	0.84	0.70	0.032	0.067	0.11	0.58	0.13	0.95
PPC	0.39	0.66	0.20	0.23	0.085	0.63	0.74	0.11	1.0
CWS	0.66	0.86	0.32	0.80	0.93	0.42	0.77	0.43	0.88
WCSA	0.39	0.69	0.19	0.12	0.18	0.19	0.023	0.037	0.12

LDLc, LDL cholesterol; TRIG, triglycerides; HDLc, HDL cholesterol; CAD, coronary artery disease; Cstif, carotid stiffness; Dext, external diastolic diameter; Dint, internal diameter; IMT, intima media thickness; MBP, mean blood pressure; PPC, central pulse pressure; CWS, Circumferential wall stress; WCSA, Wall cross sectional area.