Common coding variant in *SERPINA1* increases the risk for large artery stroke

Rainer Malik[a,1], Therese Dau[b,1], Maria Gonik[a,2], Anirudh Sivakumar[c], Daniel J. Deredge[c], Evgeniia V. Edeleva[d], Jessica Götzfried[b], Sander W. van der Laan[e], Gerard Pasterkamp[e], Nathalie Beaufort[a], Susana Seixas[f,g], Steve Bevan[h,3], Lisa F. Lincz[i], Elizabeth G. Holliday[i], Annette I. Burgess[j], Kristiina Rannikmäe[k], Jens Minnerup[l,m], Jennifer Kriebel[n], Melanie Waldenberger[n], Martina Müller-Nurasyid[o], Peter Lichtner[p,q], Danish Saleheen[r], International Stroke Genetics Consortium[4], Peter M. Rothwell[j], Christopher Levi[i], John Attia[i], Cathie L. M. Sudlow[k], Dieter Braun[d], Hugh S. Markus[h], Patrick L. Wintrode[c], Klaus Berger[l], Dieter E. Jenne[b], and Martin Dichgans[a,s,5]

[a]Institute for Stroke and Dementia Research, Klinikum der Universität München, Munich 81377, Germany; [b]Comprehensive Pneumology Center, Institute of Lung Biology and Disease, University Hospital, Ludwig Maximilians University and Helmholtz Zentrum München, German Center for Lung Research, Munich 81377, Germany; [c]Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, Baltimore, MD 21201; [d]Systems Biophysics, Physics Department, Nanosystems Initiative Munich and Center for NanoScience, Ludwig-Maximilians-Universität Munich, Munich 80799, Germany; [e]Laboratory of Experimental Cardiology, Division of Heart and Lungs, University Medical Center Utrecht, 3584 CG, Utrecht, The Netherlands; [f]Instituto de Investigação e Inovação em Saúde, Universidade do Porto, 4200-135 Porto, Portugal; [g]Institute of Molecular Pathology and Immunology, University of Porto, 4200-135 Porto, Portugal; [h]Department of Clinical Neurosciences, University of Cambridge, Cambridge CB2 0QQ, United Kingdom; [i]Public Health Research Program, Hunter Medical Research Institute, Newcastle, 2305, Australia; [j]Stroke Prevention Research Unit, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford OX3 9DU, United Kingdom; [k]Division of Clinical Neurosciences, University of Edinburgh, Edinburgh EH16 4SB, United Kingdom; [l]Institute of Epidemiology and Social Medicine, University of Münster, Muenster 48149, Germany; [m]Department of Neurology, University of Münster, Muenster 48149, Germany; [n]Research Unit Molecular Epidemiology, Helmholtz Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt, Neuherberg 85764, Germany; [o]Institut für Genetische Epidemiologie, Helmholtz Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt, Neuherberg 85764, Germany; [p]Institut für Humangenetik, Helmholtz Zentrum München, Munich 85764, Germany; [q]Institut für Humangenetik, Technische Universität München, Munich 85764, Germany; [r]Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104; and [s]Munich Cluster for Systems Neurology, Munich 81377, Germany

Edited by Costantino Iadecola, Weill Cornell Medical College, New York, NY, and accepted by Editorial Board Member Carl F. Nathan February 8, 2017 (received for review September 30, 2016)

Large artery atherosclerotic stroke (LAS) shows substantial heritability not explained by previous genome-wide association studies. Here, we explore the role of coding variation in LAS by analyzing variants on the HumanExome BeadChip in a total of 3,127 cases and 9,778 controls from Europe, Australia, and South Asia. We report on a nonsynonymous single-nucleotide variant in serpin family A member 1 (*SERPINA1*) encoding alpha-1 antitrypsin [AAT; p.V213A; $P = 5.99E-9$, odds ratio (OR) = 1.22] and confirm histone deacetylase 9 (*HDAC9*) as a major risk gene for LAS with an association in the 3′-UTR (rs2023938; $P = 7.76E-7$, OR = 1.28). Using quantitative microscale thermophoresis, we show that M1 (A213) exhibits an almost twofold lower dissociation constant with its primary target human neutrophil elastase (NE) in lipoprotein-containing plasma, but not in lipid-free plasma. Hydrogen/deuterium exchange combined with mass spectrometry further revealed a significant difference in the global flexibility of the two variants. The observed stronger interaction with lipoproteins in plasma and reduced global flexibility of the Val-213 variant most likely improve its local availability and reduce the extent of proteolytic inactivation by other proteases in atherosclerotic plaques. Our results indicate that the interplay between AAT, NE, and lipoprotein particles is modulated by the gate region around position 213 in AAT, far away from the unaltered reactive center loop (357–360). Collectively, our findings point to a functionally relevant balance between lipoproteins, proteases, and AAT in atherosclerosis.

genetics | ischemic stroke | large artery stroke | antitrypsin | variation

Stroke is the leading cause of long-term disability and the second most common cause of death worldwide (1, 2). About a quarter of ischemic stroke cases are caused by large artery atherosclerotic stroke (LAS) (3, 4). Atherosclerosis is a chronic inflammatory condition that involves a number of well-characterized steps. Initial stages include the deposition of lipids in vascular endothelial cells, whereas more advanced stages are characterized by fibrotic changes with formation of a fibrotic cap and, eventually, plaque rupture (5). LAS exhibits the highest heritability of all stroke subtypes, with estimates ranging from 40.3 to 66.6% (6, 7). This fact is reflected by recent genome-wide association studies that found common variants for LAS at multiple genomic loci (8–10). The lead SNPs from these regions all reside within intergenic (4, 7, 11) or

Significance

Common single-amino acid variations of proteins are traditionally regarded as functionally neutral polymorphisms because these substitutions are mostly located outside functionally relevant surfaces. In this study, we present an example of a functionally relevant coding sequence variation, which, as we show here, confers risk for large artery atherosclerotic stroke. The single-residue variation M1(A213V) in serpin family A member 1 (*SERPINA1*) [encoding alpha-1 antitrypsin (AAT)] is situated outside the protease-reactive inhibitory loop and is found in a β-turn on the protein surface. We show that the Ala-to-Val exchange in the gate region of AAT alters its functional dynamics toward neutrophil elastase in the presence of complex lipid-containing plasma and also affects the overall structural flexibility of the protein.

Author contributions: R.M., T.D., D.E.J., and M.D. designed research; R.M., T.D., M.G., E.V.E., G.P., N.B., S.B., L.F.L., E.G.H., A.I.B., D.S., P.M.R., C.L., J.A., C.L.M.S., D.B., H.S.M., P.L.W., and K.B. performed research; A.S., D.J.D., E.V.E., J.G., S.W.v.d.L., G.P., N.B., S.B., L.F.L., E.G.H., A.I.B., J.M., J.K., M.W., M.M.-N., P.L., D.S., I.S.G.C., P.M.R., C.L., J.A., C.L.M.S., D.B., H.S.M., P.L.W., K.B., D.E.J., and M.D. contributed new reagents/analytic tools; R.M., T.D., M.G., S.W.v.d.L., S.S., E.G.H., K.R., D.S., P.M.R., C.L., and M.D. analyzed data; and R.M., T.D., G.P., S.S., D.E.J., and M.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. C.I. is a Guest Editor invited by the Editorial Board.

Freely available online through the PNAS open access option.

See Commentary on page 3555.

[1]R.M. and T.D. contributed equally to this work.

[2]Present address: Biomax Informatics AG, Planegg 82152, Germany.

[3]Present address: School of Life Sciences, University of Lincoln LN6 7TS, United Kingdom.

[4]Members of the International Stroke Genetics Consortium: Daniel Woo, Stephanie Debette, Jane Maguire, John W. Cole, Jennifer Majersik, Jordi Jimenez-Conde, Jin-Moo Lee, Natalia Rost, Guillaume Pare, Christina Jern, Arne G. Lindgren, and Israel Fernandez Cardenas.

[5]To whom correspondence should be addressed. Email: martin.dichgans@med.uni-muenchen.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1616301114/-/DCSupplemental.

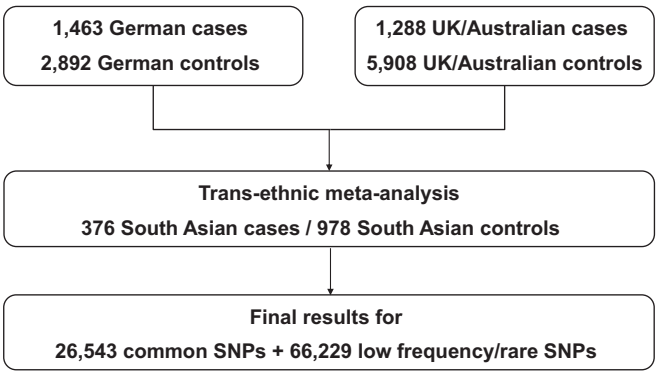

Fig. 1. Flow chart of the exome chip analysis. Samples from Germany, the United Kingdom, and Australia were pooled with samples from South Asia for the transethnic metaanalysis.

intronic (12) regions, and most of them are situated within a regulatory sequence marked by DNase I hypersensitivity sites.

Whole-exome (13) and whole-genome (14) sequencing efforts have identified multiple common, low-frequency, and rare variants that have not yet been examined for association with LAS. Conceivably, these variants might account for some of the missing heritability of LAS. To search for novel variants and genes implicated in atherosclerotic stroke, we assembled the largest cohort of LAS cases to date (3,127 cases and 9,778 controls from Germany, the United Kingdom, Australia, and Pakistan).

In the current study, we found two exome-wide significant variants, one in the established LAS risk gene histone deacetylase 9 (*HDAC9*) and one in serpin family A member 1 (*SERPINA1*). The main target of the inhibitor alpha-1 antitrypsin (AAT), encoded by *SERPINA1*, is neutrophil elastase (NE). AAT and NE are both involved in inflammation, and an imbalance between AAT and NE has previously been discussed as a possible mechanism in atherosclerotic plaque and aneurysm formation (15–18).

The rate-determining step of the inhibitory reaction between ATT and NE is the reversible formation of a noncovalent docking intermediate (encounter complex), which subsequently progresses to the covalent tetrahedral complex (16, 19, 20). The common M1 variants of AAT, the minor A213 allele and the prevailing V213 allele (rs6647), have previously been characterized as normal, functionally equivalent plasma isoforms (20) with very similar plasma levels and association rate constants for the purified isoforms (16, 19, 20). However, adjacent to the loop containing residue 213 is a hydrophobic groove (16), which, together with the polymorphic side chain, may differentially interact with endogenous hydrophobic components of plasma. In this study, we functionally characterized the interaction between AAT and NE in human plasma using microscale thermophoresis, and provide evidence for differential behavior of the two major M1(V213) and M1(A213) alleles toward lipoproteins.

Results

Common Variants in *SERPINA1* and *HDAC9* Associate with LAS. Characteristics of the case and control samples, including details on quality control (QC), are presented in Figs. S1 and S2, Table S1, and *SI Materials and Methods*. The overall strategy for the transethnic, exome-wide association study is shown in Fig. 1. We found two common variants to be associated with LAS on an exome-wide level ($P < 1.88E\text{-}6$ with Bonferroni correction for all common variants studied; Fig. 2 *A–C* and Table 1). The first variant, exm1124208 (rs6647) in *SERPINA1*, encoding M1(A213) in AAT showed a minor allele frequency (MAF) of 17.8% in Caucasian controls and 17.1% in South Asian controls [odds ratio (OR) 95% confidence interval [CI_95] = 1.22 [1.13–1.31], $P = 5.99E\text{-}9$ in the transethnic meta-analysis] (Fig. 2 *D* and *E*). PolyPhen2 (score = 0.0), PROVEAN (score = 1.11), and SIFT (score = 0.54) predict this variant to be likely benign and tolerated. The association between LAS and rs6647 remained significant ($P = 7.01E\text{-}9$) when removing carriers of the low-frequency Z and S alleles (98 cases and 262 controls) that have previously been shown to be associated with lower plasma levels of AAT. The second variant, rs2023938 in *HDAC9* (MAF = 9.1% and 10.8% in Caucasians and South Asian controls, respectively; OR [CI_95] = 1.28 [1.16–1.40, $P = 7.76E\text{-}7$]), is in the 3′-UTR of *HDAC9* and in high linkage disequilibrium (LD) with previously published risk variants for LAS in the 3′ region of HDAC9 [rs11984041 (11): $r^2 = 1$; rs2107595 (4): $r^2 = 0.53$].

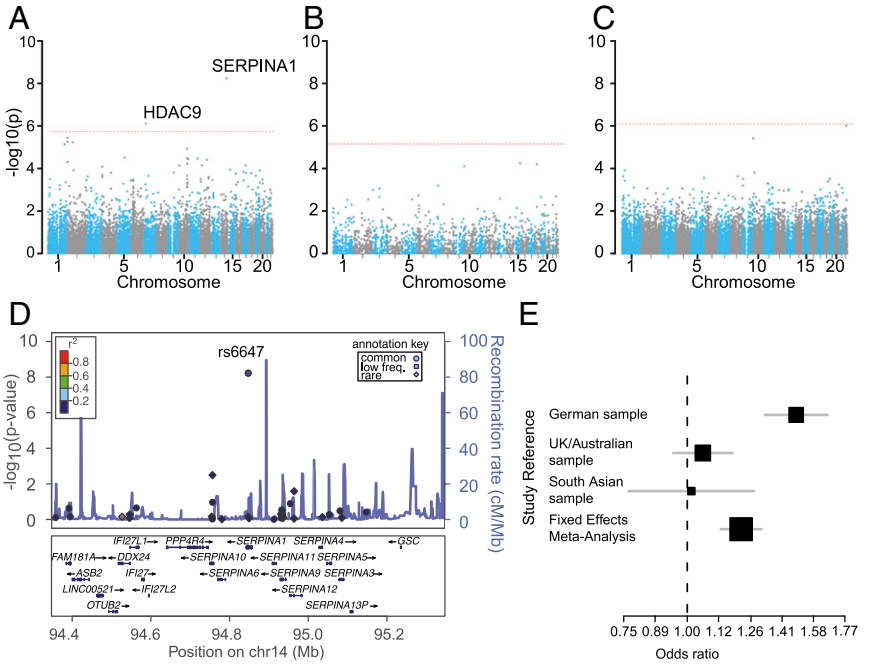

Fig. 2. Main association results. Shown are the results for common (*A*, MAF > 5%), low-frequency (*B*, 1% < MAF < 5%), and rare (*C*, MAF < 1%) variants. Genomic position is plotted on the *x* axis, and the $-\log_{10}$ of the *P* value is displayed on the *y* axis. The threshold for exome-wide significance after Bonferroni correction for the number of variants studied is displayed as a dashed red line. (*D*) LocusZoom plot [$-\log_{10}$ (p-value)] for the *SERPINA1* region and LAS. Shown is the region for the top signal ± 500 kb. Common, low-frequency, and rare variants are displayed as filled circles (●), squares (■), and diamonds (♦), respectively. Blue peaks represent estimated recombination rates. (*E*) Forest plot of associations with rs6647 in individual samples and in the resulting fixed-effects metaanalysis. Sample sizes are reflected by the size of each square. Gray bars show the [CI_95] of the point estimate.

Table 1. Top association signals from the exome chip analysis

dbSNP ID no.	Chromosome:position	Minor/major allele	MAF	Functional annotation	Gene	*P* value	OR [CI_95]
Common variants							
rs6647	**14:94847415**	**G/A**	**0.18**	**p.Val237Ala**	***SERPINA1***	**5.99E-9**	**1.22 [1.13–1.31]**
rs2023938	**7:19036775**	**G/A**	**0.09**	**3′-UTR**	***HDAC9***	**7.76E-7**	**1.28 [1.16–1.40]**
rs11553746	2:272203	T/C	0.35	p.Thr95Ile	*ACP1*	3.67E-6	1.14 [1.07–1.21]
rs11681377	2:65394714	T/C	0.47	Intergenic	*Intergenic*	5.91E-6	1.15 [1.08–1.22]
rs2290911	2:224919	G/A	0.34	Synonymous	*SH3YL1*	5.65E-6	1.14 [1.07–1.21]
rs10863936	1:212237798	G/A	0.49	Intronic	*DTL*	7.33E-6	1.19 [1.12–1.26]
rs11191447	10:104652323	T/C	0.09	Intronic	*C10orf32-AS3MT*	1.18E-5	1.22 [1.10–1.36]
rs1004467	10:104594507	C/T	0.10	Intronic	*CYP17A1*	3.22E-5	1.20 [1.08–1.32]
rs17368582	11:102738075	C/T	0.13	Synonymous	*MMP12*	3.30E-5	1.20 [1.10–1.31]
rs1172822	19:55819845	T/C	0.36	Intronic	*BRSK1*	3.69E-5	1.19 [1.10–1.28]
rs11191580	10:104906211	C/T	0.09	5′-UTR	*NT5C2*	5.70E-5	1.21 [1.09–1.35]
Low-frequency variants							
rs62019510	15:85657120	G/A	0.03	p.Asn329Ile	*PDE8A*	5.60E-5	1.39 [1.18–1.63]
rs111986709	18:29054294	T/C	0.04	p.Ser771Tyr	*DSG3*	6.08E-5	1.35 [1.17–1.57]
rs2231400	9:135753629	G/A	0.04	p.Ile5Thr	*AK8*	7.77E-5	1.30 [1.12–1.50]
Rare variants							
rs45439291	22:41077613	A/G	0.003	p.Arg191Gln	*MCHR1*	9.99E-7	2.08 [1.34–3.24]
rs114895119	9:139264888	A/T	0.003	p.Glu270Val	*CARD9*	3.81E-6	1.84 [1.41–2.40]

Shown are the results from the single-variant analysis for common, low-frequency, and rare variants with a *P* value <1E-4. *P* values and ORs are displayed from the transethnic metaanalysis of the German, UK/Australian, and South Asian samples. SNP position and functional annotation are given for hg19. Exome-wide significant results are shown in bold. dbSNP, the single nucleotide polymorphism database.

Among 14 variants meeting the criterion of suggestive association ($P < 1E-4$), three are situated within known risk loci for atherosclerotic phenotypes or risk factors for atherosclerosis (Table 1). *MMP12* (21) and *CYP17A1* (22) are associated with LAS and coronary artery disease, respectively, whereas *C10orf32-AS3MT* (23, 24) is an established risk locus for hypertension. None of the low-frequency (MAF < 5%) and rare (MAF < 1%) variants reached exome-wide significance (Fig. 2 *B* and *C*). Gene-based tests revealed no exome-wide significant signals (Table S2).

M1(A213) and Atherosclerotic Plaque Characteristics in Advanced Stages of Disease. To explore associations between the M1 (A213) variant and histological characteristics of atherosclerotic plaques, we analyzed 1,414 carotid endarterectomy samples from patients with advanced atherosclerotic lesions assembled through the Athero-Express study. The M1(A213) variant was nominally associated with a lower macrophage content ($P = 0.03$, beta = −0.19, SE = 0.09). However, when controlling for the assessment of multiple plaque characteristics, including plaque hemorrhage, collagen content, and smooth muscle cell number, the association did not reach statistical significance. The results did not materially change when correcting for antithrombotic medication, lipid-lowering drugs, and smoking status. Also, expression quantitative trait locus (eQTL) analysis revealed no association between M1(A213) and AAT levels in atherosclerotic plaques.

M1(V213) Has a Higher Dissociation Constant Toward NE than M1(A213). To explore potential functional differences between M1(A213) and M1(V213) with respect to their inhibitory behavior, we analyzed the initial interaction between AAT and NE. We characterized the formation of the encounter complex under equilibrium conditions using fluorescently labeled and catalytically inactive NE (S195A variant) in a microscale thermophoresis assay (25). AAT-deficient plasma [1:12 dilution in PBS (vol/vol)] was used as a matrix to include plasma-specific interactions that may influence the encounter complex formation between AAT and NE. Under this condition, the M1(A213) variant [dissociation constant (K_D) = 3,200 ± 500 nM] exhibited an almost twofold lower dissociation constant than the M1(V213) variant (K_D = 6,800 ± 1,000 nM) toward NE (Fig. 3*A*). These allele-specific differences disappeared when AAT-deficient plasma was freed from lipoproteins by ultracentrifugation and dialyzed against a phosphate buffer over a membrane with a 10-kDa cutoff (Fig. 3*B*).

M1(A213) Enhances the Structural Flexibility of AAT. To determine the impact of the two variants on AAT structure, we further determined the hydrogen/deuterium exchange rates for M1(V213) and M1(A213). As illustrated in Fig. 4, the exchange rates between the two variants significantly differed for multiple pepsin-generated peptides. In several surface regions, the M1(A213) variant was more susceptible to deuterium uptake by one to three ions per fragment compared with M1(V213). Of note, significant differences in the deuterium uptake were also observed at a distance from residue 213. Specifically, the Ala-213 substitution led to a higher flexibility of the C-terminal end of the reactive center loop, including strand 1 of the C β-sheet (s1c) as well as strand 5 of the A-sheet in the shutter region (s5A) (26) (Fig. 4). Conceivably, the increased dynamics of the Ala-213 variant may reduce exosite interactions with other plasma components, and may thus account for the observed higher affinity for NE in the complex plasma environment.

Discussion

Our study enabled us to detect the common variants with a moderate effect size that are associated with LAS (90% power to detect variants with >5% MAF and OR > 1.2), although the power to detect rare variants was limited. The results obtained for *HDAC9*, *MMP12*, and *CYP17A1* are consistent with previous studies (4, 11, 21, 22) that have shown associations between LAS and common variants in these regions within noncoding DNA. The variant rs2023938, which reached exome-wide significance in the current study, is located in the 3′-UTR of *HDAC9* and is in high LD with rs2107595 and other variants previously shown to be associated with LAS (4, 11). The variant rs2107595 resides in regulatory DNA, and the risk allele of this variant has previously been shown to associate with elevated expression levels of *HDAC9* (27). Genetic ablation of *HDAC9* attenuates atherosclerosis-prone mice (27, 28). Hence, the effects of this locus seem to be mediated by altered *HDAC9* expression levels.

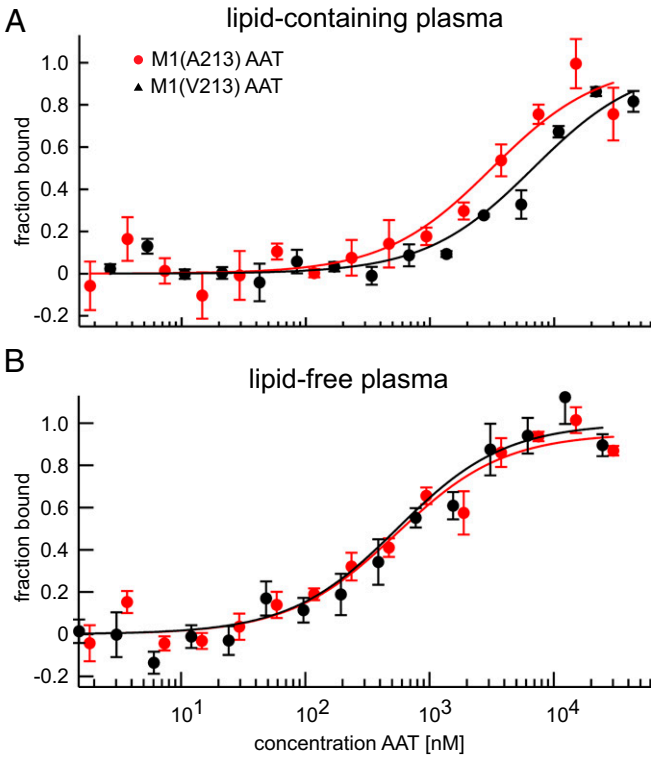

Fig. 3. Functional differences between M1 AAT variants found in lipid-containing, but not lipid-free, plasma. Shown are the fitted binding curves of the microscale thermophoresis assay under equilibrium conditions in lipid-containing (*A*) and lipid-free (*B*) plasma. A fixed concentration of fluorescently labeled and catalytically inactive NE was titrated with the M1(A213) or M1(V213) variant in AAT-deficient plasma. (*A*) In the presence of lipids, the K_D (KD) observed with the M1(V213) variant [KD(A213) = 6,800 ± 1000 nM] was substantially higher than the K_D observed with the M1(A213) variant [KD(V213) = 3,200 ± 500 nM]. (*B*) In the absence of lipids, there was no discernible difference between the two variants [KD(A213) = 540 ± 50 nM; KD(V213) = 550 ± 60 nM]. Fitted binding curves and K_D values (mean ± SD) were derived from global fitting of three measurements (three independent protein expressions). The measurements were performed in 7.5% plasma and with 5 nM labeled NE.

The primary finding of this study is an association of the common M1(A213) variant of AAT with LAS. This variant has not previously been identified through regular genome-wide association study (GWAS) approaches, and it constitutes a potential novel risk factor for LAS. No other variant within or near AAT showed a significant association with LAS. Several factors might explain why this association was not identified by previous GWAS approaches: (*i*) variations in the accuracy of phenotyping across studies and, hence, reduced statistical power (29); (*ii*) variations in imputation accuracy; and (*iii*) differences in allele frequencies for M1(A213) within the European superpopulation (14) and different regional compositions of previous GWAS discovery cohorts.

A role of *SERPINA1* in atherosclerosis is further supported by a recent study that found sixfold higher expression levels of *SERPINA1* in human atherosclerotic lesions compared with healthy arteries (30). The target enzyme of AAT, NE, is expressed by macrophages in human atherosclerotic plaques (17), particularly within advanced atherosclerotic lesions (31). AAT might facilitate protection against matrix breakdown by NE and clearance of lipoprotein deposits. However, the issues of whether AAT is protective against atherosclerosis and whether AAT levels correlate with the progression of vascular disease are still debated (16, 32–34).

The observed association between rs6647 and LAS is unlikely to be mediated by differences in AAT levels. First, previous studies have shown similar concentrations of circulating total AAT among carriers of the M1(A213) and M1(V213) alleles (16, 19, 20). Second, we found no differences in the frequency of the Z and S alleles between cases and controls. Third, the association between rs6647 and LAS remained stable when removing carriers of a Z or S allele from our cohorts (35). Finally, eQTL analysis in the Athero-Express data showed no eQTL of M1 (A213).

Previous functional studies found no difference in the inhibitory capacity and association rate constants between the two most frequent M1 AAT variants in Caucasians (20). However, these studies were performed with purified proteins in simple protein- and lipid-free buffer systems. In contrast, the microscale thermophoresis method applied here disclosed a clear difference between the two variants. This method accounts for all interactions of AAT with components of the plasma matrix, which can modify the strength of interactions between AAT and NE, as we recently showed for the AAT-Z variant (36). Under these conditions, the dissociation constant between the enzymatically inactive NE and AAT was found to be lower for M1(V213) than for M1(A213), indicating that components of the plasma matrix interfered more strongly with the former allele. The dimorphic residue at position 213 is located in the loop between strands 3 and 4 of the C sheet (s3C and s4C) and is exposed to the surface. The exposed position and hydrophobic nature of this turn make it suited for hydrophobic interactions with major plasma proteins, particularly lipoprotein particles. Compared with Val-213, the Ala-213 residue displays lower hydrophobicity and side-chain entropy, which reduces the extent of hydrophobic interactions by the s3C-s4C connecting turn of M1(A213).

Apart from the obvious single-residue substitution, further differences of the global structure were noticed in several surface regions at a distance from position 213. The subtle change in global structural flexibility between the two variants evidently does not influence the interaction between AAT and catalytically inactive NE in lipid-free plasma. This finding is consistent with earlier studies that used simple buffer systems (20). However, the lower hydrophobicity of the Ala-213 side chain and the higher surface entropy of the M1(A213) variant most likely reduce its interactions with plasma lipoproteins compared with the M1(V213) variant.

Recent studies have provided evidence for an interaction between AAT and lipoproteins, including both LDL (37) and HDL (38, 39). Binding of AAT to reconstituted HDL reduced the inhibitory capacity of AAT toward NE (39), and a similar effect was seen in our study when lipoproteins were removed from the plasma matrix. Moreover, we noticed that allele-specific differences of the dissociation constant disappeared, supporting our hypothesis of an altered interaction of M1(V213) AAT with lipoproteins in the complex plasma environment.

Better binding to lipoproteins could improve the local availability of the functional Val-213 inhibitor in atherosclerotic plaques. Moreover, we suggest that the more dynamic flexible state of the Ala-213 variant increases the likelihood of proteolytic inactivation by other proteases, particularly metalloproteases (MMP12), in atherosclerotic lesions. The precise molecular mechanism by which the M1(V213) variant reduces the risk of LAS, however, remains to be determined.

Primate sequences, including sequences of archaic hominids, indicate that the M1(V213) change, which does not occur in other mammals, arose in Africa before the divergence of modern humans (40, 41). In contemporary populations, M1(V213) is least frequent in Africa, more frequent in Europe, and almost fixed in East Asians and Native Americans. Whether this difference in allele frequencies relates to differences in LAS risk across different populations remains to be determined.

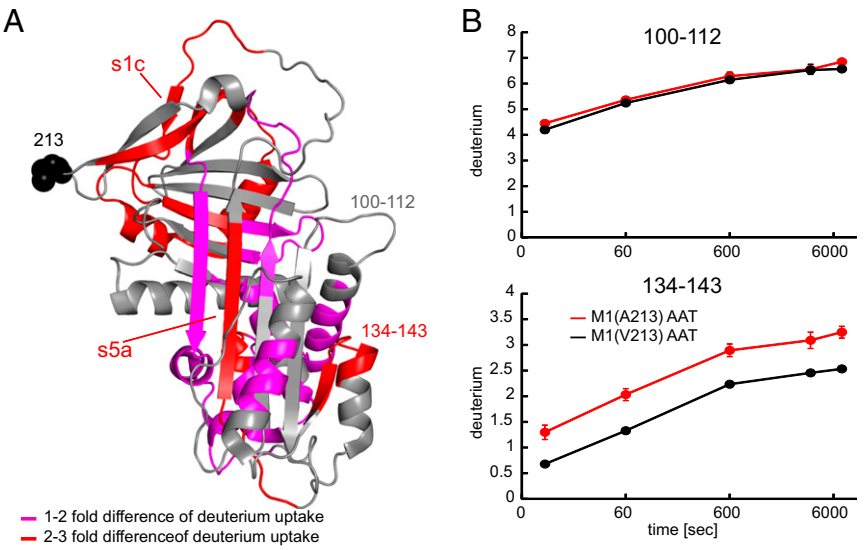

Fig. 4. Differences in hydrogen/deuterium exchange kinetics between M1(A213) and M1(V213) AAT. (*A*) Shown is the cumulative difference across five time points (10 s to 120 min) in deuterium uptake between M1(A213) and M1(V213) mapped onto the crystal structure of AAT (Protein Data Bank ID code 1QLP). Peptides displaying a significant increase in deuterium uptake in M1(A213) compared with M1(V213) are colored red (two- to threefold increase) and pink (one- to twofold increase). Otherwise, there was no significant difference in deuterium uptake between the two variants, such as the stretch between positions 100 and 112 (gray arrow). Substitution of Ala for Val at position 213 results in subtle but widespread destabilization of AAT. (*B*) Example kinetic curves for the deuterium uptake with two peptides. Peptide 134–143 (*Lower*) displays a significantly higher deuterium uptake in the M1(A213) AAT variant compared with M1(V213), whereas no significant difference is seen with peptide 100–112 (*Upper*). Values represent means from three different measurements.

Materials and Methods

Study Participants, Genotyping, and QC. We studied 3,202 participants with LAS involved in the following cross-sectional stroke studies: (*i*) German samples from Munich and the Westphalian Stroke Registry; (*ii*) UK samples from London, Edinburgh, and Oxford, including samples from the Wellcome Trust Case Control Consortium 2; (*iii*) Australian samples from the Australian Stroke Genetics Consortium study; and (*iv*) South Asian samples from the Risk Assessment of Cerebrovascular Events study (Pakistan). LAS was defined according to the Trial of Org 10172 in Acute Stroke Treatment classification, with radiological confirmation of stroke subtype. A total of 9,862 stroke-free control samples were selected from the Kooperative Gesundheitsforschung in der Region Augsburg S4/F4 study to match the German sample, from the 1958 Birth Cohort/Wellcome Trust Case Control Consortium 2 to match the UK and Australian sample, and from the Pakistan Risk of Myocardial Infarction study to match the South Asian sample (Table S1 and *SI Materials and Methods*). The relevant institutional review boards and ethics committees of participating institutions (medical faculty of Ludwig Maximilians University, John Hunter Hospital, Queen Elizabeth Hospital, Royal Perth Hospital, St. George's Hospital, University of Oxford, University of Edinburgh, and University of Münster) approved these studies, and written or oral consent was obtained from all participants. Genotyping of cases and controls was done using the Illumina HumanExome-12v1 or HumanExome-12v2 Beadchip, holding information on more than 240,000 functional exonic variants. Further information on the genotyping, QC, and statistical analysis can be found in *SI Materials and Methods*. In total, 75 cases and 84 controls were removed after QC, resulting in 3,127 cases and 9,778 controls being included in the final analysis.

Statistical Analysis. Single-variant association analysis was conducted using PLINK (42) by logistic regression adjusted for the first two principal components using genomic control to correct for population genetic substructure. Common variants (MAF > 5%; $n = 26{,}543$), low-frequency variants ($1\% < \text{MAF} < 5\%$; $n = 7{,}042$) and rare variants (MAF < 1%; $n = 59{,}187$) were studied separately, setting Bonferroni-corrected *P*-value cutoffs at $P < 1.88\text{E-}6$, $P < 7.10\text{E-}6$, and $P < 8.45\text{E-}7$, respectively. Metaanalysis was performed using fixed-effects models based on inverse variance-weighted effect size for the German and UK/Australian cohorts. Han and Eskin's (43) random effects model optimized to detect associations under heterogeneity was used for the transethnic metaanalysis. Effects of nonsynonymous variants were checked by PolyPhen2 (44), SIFT (45), and PROVEAN (46). The significance level of gene-based association tests was set at $P < 2.91\text{E-}6$ (accounts for Bonferroni correction for the number of genes).

Determination of the Dissociation Constants. Information about expression and purification of recombinant proteins, labeling of NE, and determination of the concentration of recombinant AAT is provided in *SI Materials and Methods*.

Before each experiment, all protein stocks and plasma samples were centrifuged for 5 min at 20,000 × *g* at 4 °C. Plasma and fluorescently labeled catalytically inactive elastase were mixed in buffer 1 [20 mM Na_2HPO_4, 300 mM NaCl (pH 7.4)] and 0.02% Tween-20. Antiphotobleaching enzyme and substrate components were added according to the manufacturer's protocol (Monolith Anti Photobleach Kit; NanoTemper Technologies). A dilution series of both AAT variants in buffer 1 was prepared. The AAT dilutions were then mixed with buffer 1 at a volume ratio of 1:1, yielding a final concentration of 7.5% (vol/vol) AAT-deficient plasma and 4.5 nM elastase. The relatively high background of blood plasma necessitates a background subtraction. The background mixture includes everything but labeled elastase and AAT. The samples were incubated at 22 °C for 2 h. Samples were measured in NT.115 MST Premium-coated capillaries (NanoTemper Technologies) on a Monolith NT.115 Pico instrument (NanoTemper Technologies) at 22 °C using 10% light-emitting diode and 40% infrared (IR) laser powers with IR laser on/off times of 20/5 s. Each dilution point was measured in triplicate. For each plasma sample, the whole procedure was performed three times to yield independent triplicates. After subtracting the background, the relative fluorescence depletion values were normalized to the saturation value, at which value all molecules were in the bound state. Normalized depletion values corresponding to the fraction of bound elastase molecules of three technical replicates were plotted on a linear *y* axis against the concentration of the serially diluted AAT on the $\log_{10}$ *x* axis, resulting in binding curves. In Igor Pro-5.03, a global fit of at least three replicates to the quadratic solution of the mass action law was performed to yield the dissociation constant of each curve with a weighted error fit (further details are provided in *SI Materials and Methods*).

Hydrogen/Deuterium Exchange Measurements Using Mass Spectrometry. Deuterium/hydrogen exchange and quantification were performed as previously described (47). In brief, exchange reactions were initiated by diluting 5 μL of 25 μM AAT with 45 μL of D_2O containing buffer, and the reaction was stopped at different time points (10 s, 1 min, 10 min, 60 min, and 120 min) by further addition of 50 μL of quench buffer (100 mM phosphate buffer, pH 2.5). The quench reactions were immediately injected in a Waters HDX system with in-line pepsin digestion. The peptides were separated on a C18 column before analysis on a Waters Synapt G2 mass spectrometer. The back-exchange reaction of hydrogens was corrected using a fully deuterated sample that was obtained by incubating AAT in D_2O containing high-concentration deuterated guanidine DCl before quenching. At all time points, undeuterated and fully deuterated controls were analyzed in triplicate. Peptide identification was performed with the ProteinLynx Global Server. Coverage maps and deuterium uptake for each peptide was obtained with DynamX 3.0.

ACKNOWLEDGMENTS. Samples were processed in the Genetics Core Laboratory of the WT Clinical Research Facility, Edinburgh. Neuroimaging was performed at the Scottish Funding Council Brain Imaging Research Centre Division of Clinical Neurosciences, University of Edinburgh, a core area of the WT Clinical Research Facility and part of the Scottish Imaging Network, a Platform for Scientific Excellence collaboration. This study makes use of data generated by the Wellcome Trust Case Control Consortium. A full list of investigators who contributed to the generation of the data is available (https://www.wtccc.org.uk/). This work was supported by grants from the German Federal Ministry of Education and Research (e:Med program e:AtheroSysMed), the FP7 European Union (EU) project CVgenes{at}target(261123),

the Deutsche Forschungsgemeinschaft (DFG) (CRC 1123, B3), the Corona Foundation, and the Fondation Leducq (Transatlantic Network of Excellence on the Pathogenesis of Small Vessel Disease of the Brain) (all to M.D.), as well as by grants from the Australian National and Medical Health Research Council (Grant 569257), Australian National Heart Foundation (NHF; Grant G 04S 1623), Gladys M. Brawn Fellowship scheme, and Vincent Fairfax Family Foundation in Australia. E.G.H. was supported by a Fellowship from the NHF and National Stroke Foundation of Australia (Grant 100071). The Wellcome Trust (WT) Case Control Consortium 2 (WTCCC2) was funded by the WT (WTCCC2 Projects 085475/B/08/Z, 085475/Z/08/Z, and WT084724MA). The Stroke Association provided support for collection of some of the St. George's Hospital cases. H.S.M. is supported by a National Institute for Health Research (NIHR) Senior Investigator Award and by the Cambridge University Hospitals National Health Service Trust NIHR Comprehensive Biomedical Research Centre. The work of H.S.M. and S.B. is supported by the Evelyn Trust. The Oxford cases were collected as part of the Oxford Vascular Study (funded by the Medical Research Council, Stroke Association, Dunhill Medical Trust, NIHR, and Biomedical Research Centre, Oxford). The Edinburgh Stroke Study was supported by the WT (clinician scientist award to C.L.M.S.) and the Binks Trust. Additional funding for this study was provided by the WT (Awards 076113 and 090355), a joint DFG grant (JE194/4-1 and BR 2152/2-1), the EU Horizon 2020 research and innovation program [Grant Agreement 668036, relapses prevention (RELENT)], and the University of Maryland Baltimore, School of Pharmacy Mass Spectrometry Center (Grant SOP1841-IQB2014).

1. Mozaffarian D, et al.; American Heart Association Statistics Committee and Stroke Statistics Subcommittee (2015) Heart disease and stroke statistics–2015 update: A report from the American Heart Association. *Circulation* 131(4):e29–e322.
2. WHO (2014) The top 10 causes of death. Available at www.who.int/mediacentre/factsheets/fs310/en/. Accessed December 7, 2015.
3. Schneider AT, et al. (2004) Ischemic stroke subtypes: A population-based study of incidence rates among blacks and whites. *Stroke* 35(7):1552–1556.
4. Traylor M, et al.; Australian Stroke Genetics Collaborative; Wellcome Trust Case Control Consortium 2 (WTCCC2); International Stroke Genetics Consortium (2012) Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): A meta-analysis of genome-wide association studies. *Lancet Neurol* 11(11):951–962.
5. Weber C, Noels H (2011) Atherosclerosis: Current pathogenesis and therapeutic options. *Nat Med* 17(11):1410–1422.
6. Bevan S, et al. (2012) Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genomewide associations. *Stroke* 43(12):3161–3167.
7. Holliday EG, et al.; Australian Stroke Genetics Collaborative; International Stroke Genetics Consortium; Wellcome Trust Case Control Consortium 2 (2012) Common variants at 6p21.1 are associated with large artery atherosclerotic stroke. *Nat Genet* 44(10):1147–1151.
8. Falcone GJ, Malik R, Dichgans M, Rosand J (2014) Current concepts and clinical applications of stroke genetics. *Lancet Neurol* 13(4):405–418.
9. Malik R, et al.; ISGC Analysis Group; METASTROKE collaboration; Wellcome Trust Case Control Consortium 2 (WTCCC2); NINDS Stroke Genetics Network (SiGN) (2016) Low-frequency and common genetic variation in ischemic stroke: The METASTROKE collaboration. *Neurology* 86(13):1217–1226.
10. NINDS Stroke Genetics Network; International Stroke Genetics Consortium (2015) Loci associated with ischaemic stroke and its subtypes (SiGN): A genome-wide association study. *Lancet Neurol* 15(2):174–184.
11. Bellenguez C, et al.; International Stroke Genetics Consortium (ISGC); Wellcome Trust Case Control Consortium 2 (WTCCC2) (2012) Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. *Nat Genet* 44(3):328–333.
12. Gschwendtner A, et al.; International Stroke Genetics Consortium (2009) Sequence variants on chromosome 9p21.3 confer risk for atherosclerotic stroke. *Ann Neurol* 65(5):531–539.
13. Auer PL, et al. (2012) Imputation of exome sequence variants into population-based samples and blood-cell-trait-associated loci in African Americans: NHLBI GO Exome Sequencing Project. *Am J Hum Genet* 91(5):794–808.
14. 1000 Genomes Project Consortium; et al. (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature* 491(7422):56–65.
15. Henriksen PA, Sallenave JM (2008) Human neutrophil elastase: Mediator and therapeutic target in atherosclerosis. *Int J Biochem Cell Biol* 40(6-7):1095–1100.
16. Talmud PJ, et al.; Diabetes Atherosclerosis Intervention Study Investigators (2003) Progression of atherosclerosis is associated with variation in the alpha1-antitrypsin gene. *Arterioscler Thromb Vasc Biol* 23(4):644–649.
17. Dollery CM, et al. (2003) Neutrophil elastase in human atherosclerotic plaques: Production by macrophages. *Circulation* 107(22):2829–2836.
18. Soehnlein O (2012) Multiple roles for neutrophils in atherosclerosis. *Circ Res* 110(6):875–888.
19. Ferrarotti I, et al. (2012) Serum levels and genotype distribution of α1-antitrypsin in the general population. *Thorax* 67(8):669–674.
20. Nukiwa T, et al. (1987) Characterization of the M1(Ala213) type of alpha 1-antitrypsin, a newly recognized, common "normal" alpha 1-antitrypsin haplotype. *Biochemistry* 26(17):5259–5267.
21. Traylor M, et al.; METASTROKE, International Stroke Genetics Consortium; Wellcome Trust Case Consortium 2 (WTCCC2) (2014) A novel MMP12 locus is associated with large artery atherosclerotic stroke using a genome-wide age-at-onset informed approach. *PLoS Genet* 10(7):e1004469.
22. Schunkert H, et al.; Cardiogenics; CARDIoGRAM Consortium (2011) Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet* 43(4):333–338.
23. Simino J, et al.; LifeLines Cohort Study (2014) Gene-age interactions in blood pressure regulation: A large-scale investigation with the CHARGE, Global BPgen, and ICBP Consortia. *Am J Hum Genet* 95(1):24–38.
24. Newton-Cheh C, et al.; Wellcome Trust Case Control Consortium (2009) Genome-wide association study identifies eight loci associated with blood pressure. *Nat Genet* 41(6):666–676.
25. Lippok S, et al. (2012) Direct detection of antibody concentration and affinity in human serum using microscale thermophoresis. *Anal Chem* 84(8):3523–3530.
26. Tsutsui Y, Liu L, Gershenson A, Wintrode PL (2006) The conformational dynamics of a metastable serpin studied by hydrogen exchange and mass spectrometry. *Biochemistry* 45(21):6561–6569.
27. Azghandi S, et al. (2015) Deficiency of the stroke relevant HDAC9 gene attenuates atherosclerosis in accord with allele-specific effects at 7p21.1. *Stroke* 46(1):197–202.
28. Cao Q, et al. (2014) Histone deacetylase 9 represses cholesterol efflux and alternatively activated macrophages in atherosclerosis development. *Arterioscler Thromb Vasc Biol* 34(9):1871–1879.
29. Ioannidis JP, Thomas G, Daly MJ (2009) Validating, augmenting and refining genome-wide association signals. *Nat Rev Genet* 10(5):318–329.
30. Inouye M, et al. (2012) Novel loci for metabolic networks and multi-tissue expression studies reveal genes for atherosclerosis. *PLoS Genet* 8(8):e1002907.
31. Larionov S, Dedeck O, Birkenmeier G, Thal DR (2007) Expression of alpha2-macroglobulin, neutrophil elastase, and interleukin-1alpha differs in early-stage and late-stage atherosclerotic lesions in the arteries of the circle of Willis. *Acta Neuropathol* 113(1):33–43.
32. Dahl M, et al. (2003) Blood pressure, risk of ischemic cerebrovascular and ischemic heart disease, and longevity in alpha(1)-antitrypsin deficiency: The Copenhagen City Heart Study. *Circulation* 107(5):747–752.
33. Dahl M, Tybjaerg-Hansen A, Nordestgaard BG (2003) Risk of ischemic heart and ischemic cerebrovascular disease is not increased in S, Z, and 11478A alpha1-antitrypsin carriers of the Copenhagen City Heart Study. *Arterioscler Thromb Vasc Biol* 23(11):e55; author reply e55.
34. Elzouki AN, Ryden Ahlgren A, Lanne T, Sonesson B, Eriksson S (1999) Is there a relationship between abdominal aortic aneurysms and alpha1-antitrypsin deficiency (PiZ)? *Eur J Vasc Endovasc Surg* 17(2):149–154.
35. Thun GA, et al. (2013) Causal and synthetic associations of variants in the SERPINA gene cluster with alpha1-antitrypsin serum levels. *PLoS Genet* 9(8):e1003585.
36. Dau T, et al. (2016) Quantitative analysis of protease recognition by inhibitors in plasma using microscale thermophoresis. *Sci Rep* 6:35413.
37. Mashiba S, et al. (2001) In vivo complex formation of oxidized alpha(1)-antitrypsin and LDL. *Arterioscler Thromb Vasc Biol* 21(11):1801–1808.
38. Moreno JA, et al. (2014) High-density lipoproteins potentiate α1-antitrypsin therapy in elastase-induced pulmonary emphysema. *Am J Respir Cell Mol Biol* 51(4):536–549.
39. Gordon SM, et al. (2015) Rosuvastatin alters the proteome of high density lipoproteins: Generation of alpha-1-antitrypsin enriched particles with anti-inflammatory properties. *Mol Cell Proteomics* 14(12):3247–57.
40. Reich D, et al. (2010) Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* 468(7327):1053–1060.
41. Meyer M, et al. (2012) A high-coverage genome sequence from an archaic Denisovan individual. *Science* 338(6104):222–226.
42. Purcell S, et al. (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81(3):559–575.
43. Han B, Eskin E (2011) Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *Am J Hum Genet* 88(5):586–598.
44. Adzhubei IA, et al. (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7(4):248–249.
45. Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 4(7):1073–1081.
46. Choi Y, Chan AP (2015) PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* 31(16):2745–2747.
47. Hughes VA, Meklemburg R, Bottomley SP, Wintrode PL (2014) The Z mutation alters the global structural dynamics of α1-antitrypsin. *PLoS One* 9(9):e102617.
48. Grove ML, et al. (2013) Best practices and joint calling of the HumanExome BeadChip: The CHARGE Consortium. *PLoS One* 8(7):e68095.
49. Ionita-Laza I, Lee S, Makarov V, Buxbaum JD, Lin X (2013) Sequence kernel association tests for the combined effect of rare and common variants. *Am J Hum Genet* 92(6):841–853.
50. Lee S, et al.; NHLBI GO Exome Sequencing Project—ESP Lung Project Team (2012) Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. *Am J Hum Genet* 91(2):224–237.
51. Lee S, Teslovich TM, Boehnke M, Lin X (2013) General framework for meta-analysis of rare variants in sequencing association studies. *Am J Hum Genet* 93(1):42–53.
52. Verhoeven BA, et al. (2004) Athero-express: differential atherosclerotic plaque expression of mRNA and protein in relation to cardiovascular events and patient characteristics. Rationale and design. *Eur J Epidemiol* 19(12):1127–1133.
53. van den Borne P, et al. (2014) Leukotriene B4 levels in human atherosclerotic plaques and abdominal aortic aneurysms. *PLoS One* 9(1):e86522.
54. Perera NC, et al. (2013) NSP4 is stored in azurophil granules and released by activated neutrophils as active endoprotease with restricted specificity. *J Immunol* 191(5):2700–2707.
55. Dau T, Sarker RS, Yildirim AO, Eickelberg O, Jenne DE (2015) Autoprocessing of neutrophil elastase near its active site reduces the efficiency of natural and synthetic elastase inhibitors. *Nat Commun* 6:6722.