npg

Variants in the inflammatory *IL6* and *MPO* genes modulate stroke susceptibility through main effects and gene–gene interactions

Helena Manso[1,2,3], Tiago Krug[2,4], João Sobral[1,2,3], Isabel Albergaria[1], Gisela Gaspar[1], José M Ferro[4,5], Sofia A Oliveira[2,4] and Astrid M Vicente[1,2,3]

[1]*Departamento Promoção da Saúde e Doenças Crónicas, Instituto Nacional de Saúde Dr Ricardo Jorge, Lisbon, Portugal;* [2]*Instituto Gulbenkian de Ciência, Oeiras, Portugal;* [3]*Center for Biodiversity, Functional and Integrative Genomics (BIOFIG), Lisbon, Portugal;* [4]*Instituto de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Lisbon, Portugal;* [5]*Serviço de Neurologia, Hospital de Santa Maria, Lisbon, Portugal*

There is substantial evidence that inflammation within the central nervous system contributes to stroke risk and recovery. Inflammatory conditions increase stroke risk, and the inflammatory response is of major importance in recovery and healing processes after stroke. We investigated the role of inflammatory genes *IL1B*, *IL6*, *MPO*, and *TNF* in stroke susceptibility and recovery in a population sample of 672 patients and 530 controls, adjusting for demographic, clinical and lifestyle risk factors, and stroke severity parameters. We also considered the likely complexity of inflammatory mechanisms in stroke, by assessing the combined effects of multiple genes. Two *interleukin 6* (*IL6*) and one *myeloperoxidase* (*MPO*) single-nucleotide polymorphisms were significantly associated with stroke risk ($0.022 <_{corrected}P < 0.042$), highlighting gene variants of low to moderate effect in stroke risk. An epistatic interaction between the *IL6* and *MPO* genes was also identified in association with stroke susceptibility ($P = 0.031$ after 1,000 permutations). In a subset of 546 patients, one *IL6* haplotype was associated with stroke outcome at 3 months ($_{corrected}P = 0.024$), an intriguing finding warranting further validation. Our findings support the association of the *IL6* gene and present novel evidence for the involvement of *MPO* in stroke susceptibility, suggesting a modulation of stroke risk by main gene effects, clinical and lifestyle factors, and gene–gene interactions.

Introduction

The brain was once regarded as an 'immune privileged' organ, neither susceptible to inflammation nor affected by systemic inflammatory responses. This view has, however, completely changed, and the brain is nowadays known to exhibit key features of inflammation, such as synthesis of cytokines and glial activation, and to intervene in the regulation of systemic inflammation and in acute phase response after brain injury (reviewed in Lucas *et al* (2006)). There is also substantial evidence that inflammation within the central nervous system has a role in many brain disorders including stroke, a major cause of death and significant disability in western countries.

Stroke pathophysiology is likely regulated by a combination of environmental/lifestyle and unclear genetic risk factors. Increasingly, research studies are suggesting that inflammation significantly contributes to stroke risk, progression, and outcome (Rodríguez-Yáñez and Castillo, 2008; Wang *et al*, 2007). For instance, known clinical risk factors for stroke, like atherosclerosis, diabetes, obesity, hypertension, and peripheral infection, are associated with an elevated systemic inflammatory profile (Bastard *et al*, 2006; Hansson and Libby, 2006; Moutsopoulos and Madianos, 2006). Atherosclerosis, in particular,

is an inflammatory disease and a major contributor to stroke, either through thromboembolism, which results from the rupture of atherosclerotic plaques, or indirectly through cardioembolism (Hansson and Libby, 2006).

Inflammation is equally of major importance in the acute phase of stroke and in the recovery process. It is known that the inflammatory response that follows ischemic or hemorrhagic stroke contributes to exacerbate the initial injury, but that neuroprotective and regenerative molecules are secreted at different stages after a stroke event (Correale and Villa, 2004; Lakhan *et al*, 2009; Lucas *et al*, 2006). Clearly, inflammation in central nervous system injury in general, and in stroke in particular, cannot be classified straightforwardly as harmful. Although there are many inflammatory mediators with detrimental effects, some can be beneficial and others may have dual roles, suggesting a complex orchestration in the acute and recovery phases after stroke (Lucas *et al*, 2006).

An inflammatory process is thus implicated in pathological conditions that increase stroke risk, in the injury mechanisms upon stroke and in the recovery pathways that mediate stroke outcome. It is likely that variants of genes encoding inflammatory molecules will influence not only individual stroke risk, but also the extension of the injury and the recovery process, and a number of studies have assessed this hypothesis. For instance, the *tumor necrosis factor* (*TNF*) gene has been associated with subarachnoid hemorrhage (Yamada *et al*, 2006), and polymorphisms in the *interleukin 1 beta* (*IL1B*) and *interleukin 6* (*IL6*) genes have been associated with ischemic stroke, and with ischemic stroke and intracerebral hemorrhage, respectively (Bis *et al*, 2008; Yamada *et al*, 2006). However, conflicting results have been obtained for the *IL6* gene (Tso *et al*, 2007). Other lines of evidence show that inflammatory molecules influence the extension of injury and the recovery process. Interleukin 1 beta and TNF-alpha are known to be released by neurons and endothelial cells in response to ischemia, initiating an inflammatory response and inducing IL-6 and IL-8, with deleterious consequences (Rodríguez-Yáñez and Castillo, 2008). The *myeloperoxidase* (*MPO*) gene is another intriguing candidate, as the encoded enzyme catalyzes the formation of MPO-derived reactive species that may contribute to atherosclerosis progression and destabilization of atherosclerotic plaques (reviewed in Schindhelm *et al* (2009)). Reinforcing the hypothesis of a role in stroke, *MPO* polymorphisms have been associated with the size of the brain infarct and functional outcome (Hoy *et al*, 2003).

In this study, we tested the genetic association of major inflammatory players *IL1B* (2q14), *IL6* (7p21), *TNF* (6p21.3), and *MPO* (17q23.1) with stroke susceptibility and stroke outcome at 3 months. The apparent complexity of the inflammatory mechanisms in stroke and the multiplicity of players involved suggest a concerted process, in which implicated molecules interact tightly to regulate each other. Still, nonadditive interactions or epistasis are generally overlooked in genetic studies. Epistasis is a plausible explanation for the lack of replication across different populations in candidate gene studies or in genome-wide association studies, where it is particularly difficult to assess because of the large dimension of the data (Lanktree *et al*, 2010; Moore, 2003). We, therefore, examined both independent gene effects and the occurrence of gene–gene interactions among the tested inflammatory genes in stroke risk and stroke recovery.

Materials and methods

Study Population

The study population included 672 first-ever stroke patients, recruited through Neurology and Internal Medicine Departments of several hospitals in Portugal. Stroke definition and the protocol for clinical assessment of patients were previously described (Krug *et al*, 2010; Manso *et al*, 2010). Five hundred thirty healthy controls with no clinical history of stroke were also enrolled. As stroke is a late-onset disease, we included older healthy individuals to reduce the probability of misclassification as controls. Information on clinical and lifestyle risk factors, matching the data available for patients, was obtained by direct interview of control subjects. A subset of 546 patients was included in the outcome analysis. These patients were classified into two groups based on the modified Rankin Scale (mRS) at 3 months: patients with mRS ≤ 1 were scored as 'good recovery' and with mRS > 1 were scored as 'poor recovery', as previously described (Manso *et al*, 2010).

The study was approved by the Ethics Committee of the Portuguese Dr Ricardo Jorge National Institute of Health and other hospitals involved, and participants gave their informed consent.

Single-Nucleotide Polymorphism Genotyping

To tag the genetic variation in the *IL1B*, *IL6*, *MPO*, and *TNF* gene regions, single-nucleotide polymorphisms (SNPs), located within and up to 5 kb upstream and downstream of those genes, were selected using the H-clust method (Rinaldo *et al*, 2005) (HapMap Release 21/phase II July 2006). Three SNPs in *IL1B*, six in *IL6*, two in *MPO*, and three in *TNF* were genotyped using Sequenom iPLEX assays with allele detection by mass spectroscopy, using Sequenom MassARRAY technology (Sequenom, San Diego, CA, USA) and following the manufacturer's protocol. Primer sequences were designed using Sequenom's MassARRAY Assay Design 3.0 software (Sequenom, San Diego, CA, USA). Quality control analyses were performed based on the genotyping of eight HapMap individuals, duplicated samples within and across genotyping plates, Mendelian segregation in three pedigrees and no-template samples. For each SNP, call rate < 90% and deviation from Hardy–Weinberg equilibrium ($P < 0.05$) were checked.

Statistical Analysis

To identify potential confounders, univariate analyses were performed comparing demographic and clinical and lifestyle risk factors between patients and controls, using the Pearson's χ^2 test and the Mann–Whitney test for discrete and continuous variables, respectively. Variables with a $P < 0.25$ in univariate analysis (Table 1) or of particular clinical relevance were included in a logistic regression model using forward selection (Hosmer and Lemeshow, 2000), and were maintained in the model if they were associated with stroke susceptibility at a $P \le 0.05$ level. The selected covariates were not correlated ($-0.5 <$ interaction $i < 0.5$). Logistic regression analyses were then used to determine the effect of each genetic variable on stroke susceptibility after adjustment for the significant covariates. Odds ratio (OR) and 95% confidence intervals (95% CIs) were computed for the log-additive model. A similar procedure was followed for the analysis of stroke outcome. Demographic and clinical data reflecting the severity of stroke were compared between patients with poor (mRS > 1) and good (mRS ≤ 1) outcome at 3 months to identify potential confounders (Table 2). Logistic regression analyses were then used to determine the effect of each genetic variable on patient's outcome after adjustment for the significant covariates. OR and 95% CI were also computed for the log-additive model.

Table 1 Demographic and clinical characteristics of the population sample

Characteristic	Controls	Patients	P[a]
Age, mean ± s.d. (years)	62.9 ± 6.8	52.2 ± 9.1	$< 10^{-4}$
Gender (male), *n/N* (%)	247/530 (46.6)	428/672 (63.7)	$< 10^{-4}$
Stroke type, n/N *(%)*			
Ischemic stroke	—	551/672 (82.0)	—
Hemorrhagic stroke	—	111/672 (16.5)	—
Unknown type of stroke	—	10/672 (1.5)	—
Stroke risk factors, n/N *(%)*			
Hypertension (> 85–140 mm Hg)	193/513 (37.6)	369/601 (61.4)	$< 10^{-4}$
Diabetes	59/501 (11.8)	102/628 (16.2)	0.033
Hypercholesterolemia (cholesterol > 200 mg/dL)	328/520 (63.1)	385/623 (61.8)	0.657
Smoking status	147/512 (28.7)	308/660 (46.7)	$< 10^{-4}$
Drinking status	218/505 (43.2)	388/662 (58.6)	$< 10^{-4}$

[a]Mann–Whitney test or Pearson's χ^2 test.

Table 2 Demographic and clinical characteristics of stroke patients analyzed for outcome at 3 months

Characteristic	Good recovery (mRS ≤ 1)	Poor recovery (mRS > 1)	P[a]
Age and gender			
Age, mean ± s.d. (years)	50.8 ± 9	52.5 ± 8.5	0.028
Gender (male), *n/N* (%)	174/276 (63.0)	174/270 (64.4)	0.734
History, n/N *(%)*			
Hypertension	159/241 (66.0)	143/240 (59.6)	0.147
Diabetes	36/259 (13.9)	47/246 (19.1)	0.115
Cardiac disease	37/264 (14.0)	43/257 (16.7)	0.390
Stroke type, n/N *(%)*			$< 10^{-4}$
Ischemic stroke	238/276 (86.2)	193/270 (71.5)	—
Hemorrhagic stroke	33/276 (12.0)	72/270 (26.7)	—
Unknown type of stroke	5/276 (1.8)	5/270 (1.9)	—
Stroke features, n/N *(%)*			
Aphasia	53/258 (20.5)	98/250 (39.2)	$< 10^{-4}$
Neglect	11/266 (4.1)	19/240 (7.9)	0.072
Dysphagia	15/270 (5.6)	25/251 (10.0)	0.059
Urinary incontinence	5/272 (1.8)	15/251 (6.0)	0.014
Paresis	203/273 (74.4)	244/269 (90.7)	$< 10^{-4}$
Altered consciousness	21/275 (7.6)	59/265 (22.3)	$< 10^{-4}$
Medical complications	18/265 (6.8)	82/254 (32.3)	$< 10^{-4}$
Neurologic complications	14/274 (5.1)	39/267 (14.6)	2.03×10^{-4}

mRS, modified Rankin Scale.
[a]Mann–Whitney test or Pearson's χ^2 test.

1754

Univariate and logistic regression analyses were performed using MASS and SNPassoc packages of the R software (R: A language and Environment for Statistical Computing, 2004) (v2.6.0). Haplotype blocks in the four genes were determined using the default method (Gabriel *et al*, 2002) of the Haploview software (Barrett *et al*, 2005) (v4.0) and haplotype-based association analyses were performed. This algorithm may select different haplotype blocks for the entire sample and the outcome study subset, as occurred specifically for the *IL6* gene, resulting in different haplotypes being tested for association with stroke susceptibility and outcome. Bonferroni correction for multiple testing was used to correct significant associations in individual SNP analysis, as well as haplotype-based association analysis.

Testing for genetic interactions in association with stroke susceptibility and outcome was performed using the multifactor-dimensionality reduction (MDR) method (Ritchie *et al*, 2001) (v2.0, beta 7.2), a nonparametric and genetic model-free approach. Briefly, by pooling multilocus genotypes into high- and low-risk groups, the MDR reduces the dimensionality of the data from *N* dimensions to one dimension. The new multilocus genotype attribute is then tested for its ability to classify and predict disease status, or good/poor outcome at 3 months. False-positive results due to multiple testing are reduced through combination of the cross-validation strategy and permutation testing (Moore, 2003; Ritchie *et al*, 2001). As the MDR method does not accept missing data, we imputed missing genotypes for each SNP using the PLINK software (http://pngu.mgh.harvard.edu/purcell/plink/) (Purcell *et al*, 2007) and CEU HapMap genotype data were used as reference. Best models for each group of genes were chosen from among the best 2, 3, and 4-factor models, based on the testing balanced accuracy (TBA) and the cross-validation consistency (CVC) of 10 cross-validation intervals. TBA measures how often individuals are correctly classified in relation to disease or outcome status, and CVC measures the number of times the MDR found the same set of loci across the cross-validation subsets. A model with a TBA > 0.6 is almost always statistically significant, while a TBA > 0.55 is considered interesting; TBA = 0.5 is random (http://compgen.blogspot.com/2006/12/mdr-101-part-4-results.html). The statistical significance of the best models was calculated after 1,000 permutations using the MDR Permutation Tool (v1.0, beta 2). Odds ratio-based MDR (OR-MDR) (v1.3-1) was used to determine the OR and 95% CI for each genotype combination as a quantitative measure of disease risk (Chung *et al*, 2007). To determine the gain in information about disease or outcome status by combining two variables together over that provided by the independent analysis of these variables, the MDR uses entropy measures (Jakulin and Bratko, 2003). Evidence for a synergistic interaction occurs when the combination of two or more SNPs gives a positive information gain. If the information gain is negative, then there is evidence for redundancy or correlation between SNPs; and if the information gain is 0, then the SNPs have independent effects. Entropy-based interaction dendrograms are used for interpreting epistasis models (Moore *et al*, 2006).

Results

The demographic and clinical characteristics of our population sample are presented in Table 1. Univariate analysis showed that four stroke risk factors—hypertension, diabetes, smoking status, and alcohol consumption—were, as expected, significantly more frequent in patients than in control individuals. Gender and age were also significantly different between these two groups. During sample collection, the incidence of stroke was higher in males than in females, as expected in this age range; thus, the male/female ratio was higher in the patients' group. As stroke is a late-onset disease, we selected controls with a higher mean age than patients to reduce the probability of mis-classification as 'stroke free.' Multivariate analyses were performed after univariate analyses and nongenetic confounders were identified. Adjusting for these covariates was performed in the final logistic regression model, which also included genetic markers.

All SNPs tested were in Hardy–Weinberg equilibrium in controls and met quality control criteria, and were thus further analyzed. Two contiguous SNPs in the *IL6* gene were associated with stroke susceptibility under a log-additive model (rs2069837: $P = 0.005$, OR (95% CI) = 0.66 (0.50 to 0.89); rs2069861: $P = 0.007$, OR (95% CI) = 1.74 (1.15 to 2.63)) (Figure 1, Table 3, Supplementary Table 1), after adjusting for covariates significant in the multivariate analysis model—gender, hypertension, diabetes, and smoking status. These associations with stroke susceptibility remained significant after Bonferroni correction for multiple testing (rs2069837: $_{corrected}P = 0.032$; rs2069861: $_{corrected}P = 0.042$). A three-marker haplotype containing the two SNPs individually associated with stroke susceptibility and a third SNP (rs10242595) contiguous to rs2069861 conferred an increased risk of stroke (A(rs2069837)–T(rs2069861)–G(rs10242595), $P = 0.014$) (Supplementary Table 2). This association did not withstand a Bonferroni correction, although this method may be overconservative because these SNPs are not fully independent. Overall, these results highlight a region in the *IL6* gene as a likely susceptibility *locus*, with contiguous tagging SNPs associated with stroke susceptibility.

One SNP in the *MPO* gene was significantly associated with stroke susceptibility (rs8178406: $P = 0.011$, OR (95% CI) = 0.78 (0.65 to 0.95)), and this association survived Bonferroni correction ($_{corrected}P = 0.022$) (Table 3). Interestingly, restricting the analysis to ischemic patients showed a more significant association ($_{corrected}P = 0.006$). We did not attempt the independent analysis of hemorrhagic patients, as these were too few for adequate statistical power. None of the tested SNPs in *IL1B* and *TNF* was associated with stroke susceptibility in this sample (Supplementary Table 1).

Because gene interactions may have an important impact on complex phenotypes, including human
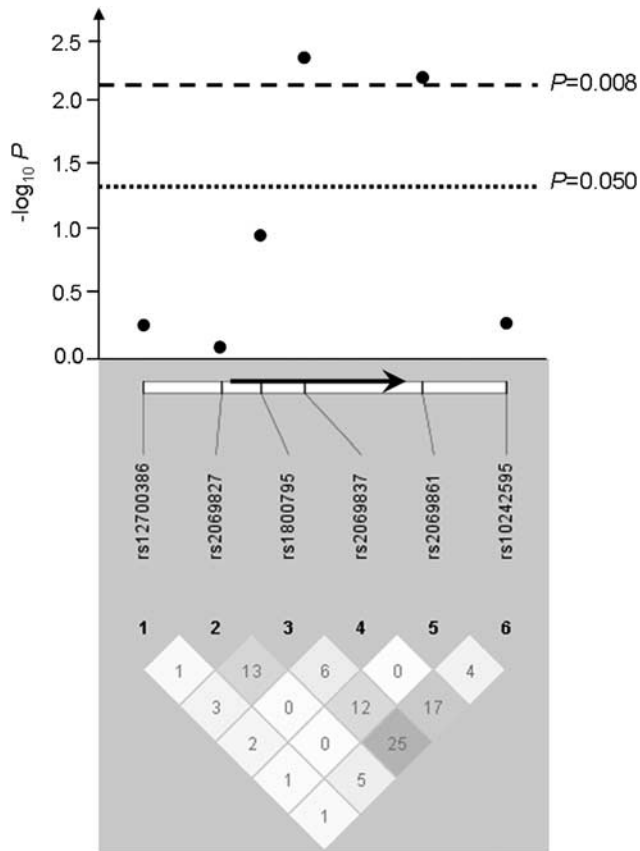
Figure 1 *IL6* association results ($-\log_{10} P$) with stroke susceptibility and pairwise r^2 among genotyped SNPs in our population sample. The positions of the six SNPs relative to the *IL6* gene (represented by an arrow) are indicated. The magnitude of linkage disequilibrium (r^2) is represented by the white–black gradient shading and the values within each diamond. Association results above the line $-\log_{10} P = 1.3$ are considered significant ($P < 0.050$); and those above $-\log_{10} P = 2.1$ survive Bonferroni correction ($P < 0.008$). Linkage disequilibrium blocks were generated using the Gabriel *et al* (2002) method.

disease susceptibility, we investigated the existence of genetic interactions that could contribute to stroke risk, using the MDR method. The most significant model for interaction was a two-marker combination between rs10242595 in the *IL6* gene and rs8178406 in the *MPO* gene (Table 4). This model shows a moderately increased TBA of 0.556, thus correctly classifying 55.6% of the individuals tested ($P = 0.031$, based on 1,000-fold permutations), but a high CVC of 9/10, i.e., the model was selected 9 times out of 10 cross-validation subsets. The global OR for this model was 1.69 (95% CI = 1.31 to 2.19). Two genotype combinations of these SNPs contributed to stroke: AA(rs10242595)-CC(rs8178406) and GA(rs10242595)-TT(rs8178406) (OR (95% CI) = 2.80 (1.17 to 7.53) and 1.53 (1.22 to 1.99), respectively) (Table 5). The *IL6* rs10242595 SNP is part of the three-marker haplotype associated with stroke susceptibility (Supplementary Table 2), while the *MPO* rs8178406 SNP was independently associated with stroke (Table 3). Interpretation of the genetic effects identified by the MDR is possible using the interaction dendrogram shown in Figure 2. Interaction between rs10242595 and rs8178406 shows a positive information gain, indicating a nonlinear, synergistic relationship between the *IL6* (rs10242595) and the *MPO* (rs8178406) genes (i.e., epistasis). Our finding thus suggests that an interaction between two genetic variants in the *IL6* and *MPO* genes contributes to stroke susceptibility, warranting confirmation in independent populations. Additional models were not significant, although there are trends possibly suggesting that interactions between *IL6* (rs10242595) and *TNF* (rs909253), as well as *MPO* (rs2071590) and *TNF* (rs8178406) (permuted $P = 0.054$ and $P = 0.060$, respectively) may contribute to stroke susceptibility (Table 4).

The impact of *IL1B*, *IL6*, *MPO*, and *TNF* genetic variants in patient's outcome at 3 months was

Table 3 Genotype and allele frequency distribution, and association with stroke susceptibility for the *IL6* and *MPO* SNPs

Gene	SNP	Genotype	Genotype frequency		OR (95% CI)	P[a]
			Controls, n (%)	Cases, n (%)		
IL6	**rs2069837**					
		A/A	365 (76.4)	461 (81.2)	0.66 (0.50–0.89)	0.005[b]
		G/A	105 (22.0)	102 (18.0)		
		G/G	8 (1.7)	5 (0.9)		
IL6	**rs2069861**					
		C/C	442 (91.9)	497 (86.9)	1.74 (1.15–2.63)	0.007[b]
		T/C	39 (8.1)	70 (12.2)		
		T/T	0 (0.0)	5 (0.9)		
MPO	**rs8178406**					
		T/T	151 (31.4)	221 (38.9)	0.78 (0.65–0.95)	0.011[b]
		T/C	254 (52.8)	262 (46.1)		
		C/C	76 (15.8)	85 (15.0)		

CI, confidence interval; OR, odds ratio; SNPs, single-nucleotide polymorphisms.
Only associated SNPs are shown.
OR > 1 indicates increased probability of having a stroke for the carriers of the minor allele.
[a]OR (95% CI) and P for the log-additive genetic model after adjustment for significant covariates (gender, history of hypertension, diabetes, and smoking status).
[b]Significant result after Bonferroni correction.

Table 4 Gene × gene interaction models obtained using the multifactor-dimensionality reduction (MDR) method in stroke susceptibility

Genes	Best model				CVC	TBA	P*
	SNP1	SNP2	SNP3	SNP4			
IL1B_TNF	rs1143643(IL1B)	rs16944(IL1B)	rs2071590(TNF)	—	8/10	0.517	0.487
IL6_TNF	rs10242595(IL6)	rs909253(TNF)	—	—	10/10	0.549	0.054
IL6_IL1B	rs2069837(IL6)	rs10242595(IL6)	rs1143643(IL1B)	—	5/10	0.541	0.145
MPO_TNF	rs2071590(TNF)	rs8178406(MPO)	—	—	9/10	0.547	0.060
IL6_MPO	rs10242595(IL6)	rs8178406(MPO)	—	—	9/10	0.556	**0.031**
MPO_IL1B	rs1143643(IL1B)	rs16944(IL1B)	rs8178406(MPO)	rs4401102(MPO)	9/10	0.538	0.160
IL1B_IL6_TNF_MPO	rs10242595(IL6)	rs1143643(IL1B)	rs8178406(MPO)	rs4401102(MPO)	8/10	0.527	0.323

CVC, cross-validation consistency; SNP, single-nucleotide polymorphism; TBA, testing balanced accuracy.
*1,000 permutations P.
Bold values are significant results.

Table 5 Odds ratio (OR) of each genotype combination of IL6 rs10242595 and MPO rs8178406 obtained using the OR-based MDR

rs10242595	rs8178406	Frequency (case:control)	OR (95% CI)
AA	CC	17:5	**2.80 (1.17–7.53)**
GA	CC	42:44	0.79 (0.59–1.18)
GG	CC	28:32	0.72 (0.51–1.18)
AA	TC	48:54	0.73 (0.56–1.06)
GA	TC	136:124	0.90 (0.76–1.12)
GG	TC	108:94	0.95 (0.78–1.21)
AA	TT	32:33	0.80 (0.57–1.28)
GA	TT	132:71	**1.53 (1.22–1.99)**
GG	TT	77:53	1.20 (0.92–1.66)

CI, confidence interval; MDR, multifactor-dimensionality reduction.
Bold values are significant results.
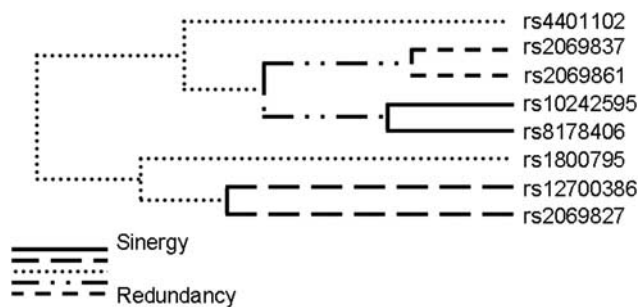
Figure 2 Interaction dendrogram for the IL6 and MPO polymorphisms in stroke susceptibility. The length of the dendrogram branch that connects two polymorphisms indicates the strength of interaction (the shorter the branch, the stronger is the interaction).

investigated in the subset of 546 patients for whom clinical information during hospitalization at 3 months was available. The demographic and clinical characteristics of our population sample are presented in Table 2. Occurrence of aphasia, urinary incontinence, paresis, altered consciousness and medical and neurologic complications during hospitalization, which reflect stroke severity, was identified in univariate analysis as significant predictors of poor outcome. One SNP in the IL6 gene was associated with stroke outcome at 3 months (rs1800795: $P = 0.011$, OR (95% CI) = 1.52 (1.10 to 2.12)) (Supplementary Table 3), after adjusting for covariates significant in the multivariate model—type of stroke, history of hypertension, and occurrence of aphasia, paresis, altered consciousness, and complications during hospitalization, but did not remain significant after Bonferroni correction ($_{corrected}P = 0.066$). One two-marker haplotype containing this SNP was associated with an increased probability of good recovery at 3 months (G(rs1800795)-A(rs2069837), $P = 0.008$), and this haplotypic association survived Bonferroni correction ($_{corrected}P = 0.048$) (Supplementary Table 4). We found no evidence for an association of IL1B, MPO, or TNF with stroke outcome at 3 months (Supplementary Tables 3 and 4). No significant interaction model was found for stroke outcome at 3 months (Supplementary Table 5).

Discussion

The objective of the present study was to investigate the role of selected inflammatory genes in stroke susceptibility and recovery. We found evidence for a main effect of the IL6 and MPO genes in stroke risk, with specific polymorphisms significantly associated with stroke susceptibility, after adjustment for confounding demographic, clinical or lifestyle risk factors. We also report an epistatic gene interaction effect between IL6 and MPO in stroke susceptibility. Our genetic findings thus support previous evidence from other research areas for a role of inflammatory molecules in stroke.

Association analysis of the IL6 gene showed that two SNPs survived Bonferroni correction, highlighting a region in the IL6 gene that is likely to harbor risk variants of moderate to low effect size. The associated SNPs are contiguous to the IL6 SNP (rs1800795) that has been widely tested in multiple

population sets, but are not in linkage disequilibrium with this functional polymorphism in our sample. We did not replicate the association with this SNP, suggesting that the present results are signaling a different causative variant in the *IL6* gene, but still reinforce a role of the *IL6* gene in stroke susceptibility. Accordingly, several previous studies have failed to confirm the association of rs1800795 with stroke, while others showed heterogeneity regarding the associated allele or genotype (Tso *et al*, 2007). These conflicting results may be due to allelic or genetic heterogeneity and/or limitations in study designs, or reflect true differences in stroke etiology between populations. Our study also provided novel evidence for the association of the *MPO* gene with stroke. Of the two SNPs tested, covering genetic variability in this region, one was associated with stroke risk. This effect seemed to be largely driven by the ischemic stroke subset where the strength of association was improved, perhaps reflecting somewhat distinct pathological mechanisms for the hemorrhagic and ischemic subtypes. Validation in independent populations is now warranted.

The identification of a synergistic interaction between *IL6* and *MPO* contributing to stroke susceptibility highlights the importance of testing for epistasis and illustrates the complexity of the inflammatory processes in stroke. It indicates that susceptibility may be modulated not only by a variety of genetic factors but also by nonlinear gene–gene interactions, as had been previously shown by others (Flex *et al*, 2004; Liu *et al*, 2009; Palmer *et al*, 2010). Liu *et al* (2009), in particular, investigated the existence of gene–gene interactions between five candidate genes and stroke and found that individuals with a combination of polymorphisms in three of these genes had an increased risk of thrombotic stroke. Two additional studies also report that the risk of stroke increases with the number of high risk genotypes in proinflammatory gene polymorphisms carried by an individual, suggesting that such polymorphisms act synergistically (Flex *et al*, 2004; Palmer *et al*, 2010). Finally, our finding is in agreement with a previous *in vitro* functional study, showing that enzymatically inactive MPO induced IL-6 secretion in a dose and time-dependent manner by endothelial cells (Lefkowitz *et al*, 2000).

The genetic factors influencing outcome after a stroke event are far less studied than genetic risk factors. The importance of inflammation after stroke onset and the correlation between inflammatory marker levels and infarct volume or patient's outcome (Smith *et al*, 2004; Sotgiu *et al*, 2006), led us to investigate the role of several inflammatory genes in stroke outcome at 3 months. We found an *IL6* two-marker haplotype associated with patient's outcome at 3 months. These results are intriguing, as IL-6 is one of the cytokines induced after stroke, having a fundamental role in the inflammatory injury that follows a stroke event, but equally known to have neuroprotective effects in later stages after stroke

(Herrmann *et al*, 2003). Interleukin-6 levels have been correlated with stroke severity, 12 months mortality, clinical outcome, and brain infarct volume (Smith *et al*, 2004). However, our results require validation in independent, larger population samples.

Interleukin-6 is a pleiotropic cytokine, with both pro and antiinflammatory functions and a low level of expression in the brain under normal physiological conditions (Luheshi and Rothwell, 1996). However, increased levels of IL-6 have been detected after a stroke event (Clark *et al*, 1999). Increased expression of IL-6 has also been found in atherosclerotic plaques (Schieffer *et al*, 2000), suggesting that the identified *IL6* association with stroke susceptibility may be mediated by atherosclerosis progression. Likewise, abundant MPO-positive cells are present in sites of atherosclerotic plaque rupture, and this molecule may contribute to stroke through destabilization of the atherosclerotic plaques (Sugiyama *et al*, 2001). It would be very interesting to correlate *IL6* and *MPO* genetic variants with the carotid intimal-media wall thickness, which is a marker for atherosclerosis (Mattace Raso *et al*, 1999). However, this data are only available for a small percentage of our patients, precluding this analysis for now.

The complex interplay between genetic background, clinical and lifestyle factors and the environment may ultimately regulate the onset, acute phase, and outcome of stroke. In the present study, we present supporting evidence for a role of the *IL6* and *MPO* inflammatory genes in stroke susceptibility, and show that stroke risk is modulated by main gene effects together with clinical and lifestyle factors as well as by gene–gene interactions. Our findings are compatible and strengthen previous genetic and biological observations, highlighting the need of further functional studies, particularly in view of the possible utility of IL-6 as a diagnostic and prognostic biomarker for stroke.

Acknowledgements

1758

Disclosure/conflict of interest

The authors declared no conflict of interest.

References

Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–5

Bastard JP, Maachi M, Lagathu C, Kim MJ, Caron M, Vidal H, Capeau J, Feve B (2006) Recent advances in the relationship between obesity, inflammation, and insulin resistance. *Eur Cytokine Netw* 17:4–12

Bis JC, Heckbert SR, Smith NL, Reiner AP, Rice K, Lumley T, Hindorff LA, Marciante KD, Enquobahrie DA, Monks SA, Psaty BM (2008) Variation in inflammation-related genes and risk of incident nonfatal myocardial infarction or ischemic stroke. *Atherosclerosis* 198:166–73

Chung Y, Lee SY, Elston RC, Park T (2007) Odds ratio based multifactor-dimensionality reduction method for detecting gene-gene interactions. *Bioinformatics* 23:71–6

Clark WM, Rinker LG, Lessov NS, Hazel K, Eckenstein F (1999) Time course of IL-6 expression in experimental CNS ischemia. *Neurol Res* 21:287–92

Correale J, Villa A (2004) The neuroprotective role of inflammation in nervous system injuries. *J Neurol* 251:1304–16

Flex A, Gaetani E, Papaleo P, Straface G, Proia AS, Pecorini G, Tondi P, Pola P, Pola R (2004) Proinflammatory genetic profiles in subjects with history of ischemic stroke. *Stroke* 35:2270–5

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D (2002) The structure of haplotype blocks in the human genome. *Science* 296:2225–9

Hansson GK, Libby P (2006) The immune response in atherosclerosis: a double-edged sword. *Nat Rev Immunol* 6:508–19

Herrmann O, Tarabin V, Suzuki S, Attigah N, Coserea I, Schneider A, Vogel J, Prinz S, Schwab S, Monyer H, Brombacher F, Schwaninger M (2003) Regulation of body temperature and neuroprotection by endogenous interleukin-6 in cerebral ischemia. *J Cereb Blood Flow Metab* 23:406–15

Hosmer DW, Lemeshow S (2000) *Applied Logistic Regression*. New York: John Wiley and Sons, Inc

Hoy A, Leininger-Muller B, Poirier O, Siest G, Gautier M, Elbaz A, Amarenco P, Visvikis S (2003) Myeloperoxidase polymorphisms in brain infarction. Association with infarct size and functional outcome. *Atherosclerosis* 167:223–30

Jakulin A, Bratko I (2003) Analyzing attribute interactions. *Lect Notes Artif Intell* 2838:229–40

Krug T, Manso H, Gouveia L, Sobral J, Xavier JM, Albergaria I, Gaspar G, Correia M, Viana-Baptista M, Simões RM, Pinto AN, Taipa R, Ferreira C, Fontes JR, Silva MR, Gabriel JP, Matos I, Lopes G, Ferro JM, Vicente AM, Oliveira SA (2010) Kalirin: a novel genetic risk factor for ischemic stroke. *Hum Genet* 127:513–23

Lakhan SE, Kirchgessner A, Hofer M (2009) Inflammatory mechanisms in ischemic stroke: therapeutic approaches. *J Transl Med* 7:97

Lanktree MB, Dichgans M, Hegele RA (2010) Advances in genomic analysis of stroke: what have we learned and where are we headed? *Stroke* 41:825–32

Lefkowitz DL, Roberts E, Grattendick K, Schwab C, Stuart R, Lincoln J, Allen RC, Moguilevsky N, Bollen A, Lefkowitz SS (2000) The endothelium and cytokine secretion: the role of peroxidases as immunoregulators. *Cell Immunol* 202:23–30

Liu J, Sun K, Bai Y, Zhang W, Wang X, Wang Y, Wang H, Chen J, Song X, Xin Y, Liu Z, Hui R (2009) Association of three-gene interaction among MTHFR, ALOX5AP and NOTCH3 with thrombotic stroke: a multicenter case-control study. *Hum Genet* 125:649–56

Lucas SM, Rothwell NJ, Gibson RM (2006) The role of inflammation in CNS injury and disease. *Br J Pharmacol* 147(Suppl 1):S232–40

Luheshi G, Rothwell N (1996) Cytokines and fever. *Int Arch Allergy Immunol* 109:301–7

Manso H, Krug T, Sobral J, Albergaria I, Gaspar G, Ferro JM, Oliveira SA, Vicente AM (2010) Variants of the Matrix Metalloproteinase-2 but not the Matrix Metalloproteinase-9 genes significantly influence functional outcome after stroke. *BMC Med Genet* 11:40

Mattace Raso F, Rosato M, Talerico A, Cotronei P, Mattace R (1999) Intimal-medial thickness of the common carotid arteries and lower limbs atherosclerosis in the elderly. *Minerva Cardioangiol* 47:321–7

Moore JH (2003) The ubiquitous nature of epistasis in determining susceptibility to common human diseases. *Hum Hered* 56:73–82

Moore JH, Gilbert JC, Tsai CT, Chiang FT, Holden T, Barney N, White BC (2006) A flexible computational framework for detecting, characterizing, and interpreting statistical patterns of epistasis in genetic studies of human disease susceptibility. *J Theor Biol* 241:252–61

Moutsopoulos NM, Madianos PN (2006) Low-grade inflammation in chronic infectious diseases: paradigm of periodontal infections. *Ann NY Acad Sci* 1088:251–64

Palmer CN, Kimber CH, Doney AS, Proia AS, Morris AD, Gaetani E, Quarta M, Smith RC, Pola R (2010) Combined effect of inflammatory gene polymorphisms and the risk of ischemic stroke in a prospective cohort of subjects with type 2 diabetes: a Go-DARTS study. *Diabetes* 59:2945–8

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559–75

R: A language and Environment for Statistical Computing (2004) Vienna

Rinaldo A, Bacanu SA, Devlin B, Sonpar V, Wasserman L, Roeder K (2005) Characterization of multilocus linkage disequilibrium. *Genet Epidemiol* 28:193–206

Ritchie MD, Hahn LW, Roodi N, Bailey LR, Dupont WD, Parl FF, Moore JH (2001) Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *Am J Hum Genet* 69:138–47

Rodríguez-Yáñez M, Castillo J (2008) Role of inflammatory markers in brain ischemia. *Curr Opin Neurol* 21:353–7

Schieffer B, Schieffer E, Hilfiker-Kleiner D, Hilfiker A, Kovanen PT, Kaartinen M, Nussberger J, Harringer W, Drexler H (2000) Expression of angiotensin II and interleukin 6 in human coronary atherosclerotic plaques: potential implications for inflammation and plaque instability. *Circulation* 101:1372–8

Schindhelm RK, van der Zwan LP, Teerlink T, Scheffer PG (2009) Myeloperoxidase: a useful biomarker for cardio-vascular disease risk stratification? *Clin Chem* 55: 1462–70

Smith CJ, Emsley HC, Gavin CM, Georgiou RF, Vail A, Barberan EM, del Zoppo GJ, Hallenbeck JM, Rothwell NJ, Hopkins SJ, Tyrrell PJ (2004) Peak plasma inter-leukin-6 and other peripheral markers of inflammation in the first week of ischaemic stroke correlate with brain infarct volume, stroke severity and long-term outcome. *BMC Neurol* 4:2

Sotgiu S, Zanda B, Marchetti B, Fois ML, Arru G, Pes GM, Salaris FS, Arru A, Pirisi A, Rosati G (2006) Inflamma-tory biomarkers in blood of patients with acute brain ischemia. *Eur J Neurol* 13:505–13

Sugiyama S, Okada Y, Sukhova GK, Virmani R, Heinecke JW, Libby P (2001) Macrophage myeloperoxidase regulation by granulocyte macrophage colony-stimulating factor in human atherosclerosis and implications in acute coronary syndromes. *Am J Pathol* 158:879–91

Tso AR, Merino JG, Warach S (2007) Interleukin-6 174G/C polymorphism and ischemic stroke: a systematic review. *Stroke* 38:3070–5

Wang Q, Tang XN, Yenari MA (2007) The inflammatory response in stroke. *J Neuroimmunol* 184:53–68

Yamada Y, Metoki N, Yoshida H, Satoh K, Ichihara S, Kato K, Kameyama T, Yokoi K, Matsuo H, Segawa T, Watanabe S, Nozawa Y (2006) Genetic risk for ischemic and hemorrhagic stroke. *Arterioscler Thromb Vasc Biol* 26:1920–5