Meta-analysis in more than 17,900 cases of ischemic stroke reveals a novel association at 12q24.12

`OPEN`

Laura L. Kilarski, PhD
Sefanja Achterberg, MD
William J. Devan, BS
Matthew Traylor, MSc
Rainer Malik, PhD
Arne Lindgren, MD
Guillame Pare, MD
Pankaj Sharma, FRCP
Agnieszcka Slowik, MD, PhD
Vincent Thijs, MD, PhD
Matthew Walters, MD
Bradford B. Worrall, MD, MSc
Michele M. Sale, PhD
Ale Algra, MD, PhD
L. Jaap Kappelle, MD
Cisca Wijmenga, PhD
Bo Norrving, MD
Johanna K. Sandling, PhD
Lars Rönnblom, MD
An Goris, PhD
Andre Franke, PhD
Cathie Sudlow, FRCPE
Peter M. Rothwell, FRCP
Christopher Levi, MBBS
Elizabeth G. Holliday, PhD
Myriam Fornage, PhD
Bruce Psaty, MD, PhD
Solveig Gretarsdottir, PhD
Unnar Thorsteinsdottir, PhD
Sudha Seshadri, MD
Braxton D. Mitchell, PhD
Steven Kittner, MD
Robert Clarke, MD
Jemma C. Hopewell, PhD
Joshua C. Bis, PhD
Giorgio B. Boncoraglio, MD
James Meschia, MD
M. Arfan Ikram, MD, PhD
Bjorn M. Hansen, MD
Joan Montaner, MD
Gudmar Thorleifsson, PhD
Kari Stefanson, PhD, Drmed
Jonathan Rosand, MD, MSc
Paul I.W. de Bakker, PhD
Martin Farrall, FRCPath

ABSTRACT

Objectives: To perform a genome-wide association study (GWAS) using the Immunochip array in 3,420 cases of ischemic stroke and 6,821 controls, followed by a meta-analysis with data from more than 14,000 additional ischemic stroke cases.

Methods: Using the Immunochip, we genotyped 3,420 ischemic stroke cases and 6,821 controls. After imputation we meta-analyzed the results with imputed GWAS data from 3,548 cases and 5,972 controls recruited from the ischemic stroke WTCCC2 study, and with summary statistics from a further 8,480 cases and 56,032 controls in the METASTROKE consortium. A final in silico "look-up" of 2 single nucleotide polymorphisms in 2,522 cases and 1,899 controls was performed. Associations were also examined in 1,088 cases with intracerebral hemorrhage and 1,102 controls.

Results: In an overall analysis of 17,970 cases of ischemic stroke and 70,764 controls, we identified a novel association on chromosome 12q24 (rs10744777, odds ratio [OR] 1.10 [1.07–1.13], $p = 7.12 \times 10^{-11}$) with ischemic stroke. The association was with all ischemic stroke rather than an individual stroke subtype, with similar effect sizes seen in different stroke subtypes. There was no association with intracerebral hemorrhage (OR 1.03 [0.90–1.17], $p = 0.695$).

Conclusion: Our results show, for the first time, a genetic risk locus associated with ischemic stroke as a whole, rather than in a subtype-specific manner. This finding was not associated with intracerebral hemorrhage. *Neurology® 2014;83:678–685*

GLOSSARY

GWAS = genome-wide association study; **ICH** = intracerebral hemorrhage; **LD** = linkage disequilibrium; **MAF** = minor allele frequency; **OR** = odds ratio; **QC** = quality control; **SNP** = single nucleotide polymorphism; **WTCCC2** = Wellcome Trust Case Control Consortium 2.

Genetic variation is now thought to play an important role in many diseases, including stroke. Genome-wide association studies (GWAS) have been applied to ischemic stroke directly, with *HDAC9* being identified as the first genetic risk factor specific to large artery ischemic stroke in the Wellcome Trust Case Control Consortium 2 (WTCCC2) study.[1] A 6p21.1 locus has also been associated with large artery stroke in a GWAS from Australia.[2] Subsequent replication of these associations in a large meta-analysis by the METASTROKE consortium confirmed that all were specific to individual ischemic stroke subtypes.[3]

GWAS arrays are designed to provide broad coverage of the entire human genome for non–hypothesis-driven studies. To bridge the gap between full GWAS arrays and targeted candidate gene studies, a smaller series of custom arrays has been developed. One such example is the Immunochip, which offers a targeted genome-wide array comprising ∼200,000 genetic variants spanning a range of immune-related genes.[4] Development of the Immunochip also included approximately 3,000 single nucleotide polymorphisms (SNPs) associated with ischemic stroke from an early-stage analysis of WTCCC2 data. However, inflammatory processes have been implicated in the pathogenesis of

---

Martin Dichgans, MD*
Hugh S. Markus, MD, FRCP*
Steve Bevan, PhD*
On behalf of the GARNET Collaborative Research Group, Wellcome Trust Case Control Consortium 2, Australian Stroke Genetic Collaborative, the METASTROKE Consortium, and the International Stroke Genetics Consortium

Correspondence to
Dr. Bevan:
snb31@medschl.cam.ac.uk

cardiovascular disease and stroke, suggesting the nonstroke content of the Immunochip may provide additional information when considering the stroke phenotype.[5,6] We report here the use of the Immunochip as the initial phase of a targeted GWAS, followed by meta-analysis with full GWAS data from WTCCC2 and an international collaboration of ischemic stroke GWAS data (METASTROKE). This is followed by in silico replication (i.e., ascertainment from previous data without the need for de novo genotyping) with data from the INTERSTROKE and VISP studies.

METHODS **Study design and participating studies.** The discovery sample consisted of 6 cohorts of patients of European ancestry with ischemic stroke. Participating centers were based in Belgium, Germany, the Netherlands (the PROMISe Study), Poland, Sweden, and the UK (2 cohorts, one from London [Imperial College; the BRAINS study] and one from Glasgow). All cohorts provided geographically and ancestry-matched controls. For the purposes of meta-analysis, the UK cohorts were treated as a single center in line with previous analyses undertaken in WTCCC2.[1]

Analysis plan. The analysis plan for this study was to perform a single meta-analysis of available data as follows: (1) association analysis of imputed Immunochip data; (2) meta-analysis with HAPMAP2-imputed WTCCC2 data and METASTROKE consortium data for which summary statistics were available; and (3): in silico look-up of significant SNPs from meta-analysis in the INTERSTROKE cohort[7] and the VISP cohort.[8] The populations used in both WTCCC2 and METASTROKE have been previously reported.[1,3] The WTCCC2 data have also been contributed to METASTROKE. Therefore, for this analysis the WTCCC2 data were removed from METASTROKE to prevent duplication of individuals, as was the BRAINS dataset, which overlapped with BRAINS cases contributing to the Immunochip discovery cohort. Table 1 includes full details of the discovery cohorts and outlines details of the WTCCC2, METASTROKE, INTERSTROKE, VISP, and intracerebral hemorrhage (ICH) cohorts. A GWAS standard a priori significance threshold of 5×10^{-8} was considered as a significant finding prior to analysis.

We additionally determined whether genome-wide associated SNPs from this analysis were also associated with primary ICH by in silico replication in GWAS data from a meta-analysis of 1,088 ICH cases and 1,102 controls (Genetics of Cerebral Hemorrhage with Anticoagulation [GOCHA] study).[9,10]

Full population details and demographics of all consortia are available in their original publications.[1,3,7,8,9,10] Full details of the Immunochip cohorts and the analysis methodology are available in the online supplementary material on the *Neurology*® Web site at Neurology.org.

Data genotyping, imputation, and statistical analysis. The Immunochip consortium developed as an immune-related targeted GWAS array comprising ~200,000 SNPs.[4] As part of

Table 1	Cohort sizes by study					
Study	Cohort	Controls	All	SVD	LVD	CE
Immunochip	Belgium (Leuven)	319	396	49	57	147
	Germany (Munich)	2,355	421	8	101	127
	Netherlands (Utrecht)	1,145	556	232	324	0
	Poland (Krakow)	255	384	28	33	119
	Sweden (Lund)	997	796	183	56	246
	UK (London & Glasgow)	1,790	867	257	152	130
	Total	6,861	3,420	757	723	771
WTCCC2		5,972	3,548	580	844	790
METASTROKE[a]		56,032[b]	8,480	1,177	1,203	1,586
INTERSTROKE		852	797	228	165	0
VISP		1,047	1,725	0	0	0
	Total	**70,764**	**17,970**	**2,742**	**2,935**	**3,147**
ICH (GOCHA)		390	389	—	—	—
ICH (GOCHA-Warfarin)		169	181	—	—	—
ICH—Europe		529	532	—	—	—

Abbreviations: CE = cardioembolic; ICH = intracerebral hemorrhage; LVD = large vessel disease; SVD = small vessel disease; WTCCC2 = Wellcome Trust Case Control Consortium 2.
The UK Immunochip cohort comprised samples from London and Glasgow with matched controls. In line with the WTCCC2 study, these were meta-analyzed as a single UK cohort. The WTCCC2 study comprised cohorts from Munich, Germany, and 3 UK sites (London, Oxford, and Edinburgh), as detailed previously. METASTROKE comprises 15 cohorts worldwide; full details have been published.
[a] METASTROKE in this study excludes those cohorts also included in WTCCC2 and BRAINS to prevent duplication of samples. For brevity both WTCCC2 and METASTROKE are referred to as single cohorts here.
[b] For METASTROKE, not all cohorts contained all subtypes. As such, the control figure in the table represents the maximal number of controls for the all ischemic stroke analysis.

this design of the Immunochip, WTCCC2 disease areas, including ischemic stroke, were able to suggest ~3,000 novel SNPs for incorporation into the array. The Immunochip therefore contains a subset of stroke-specific SNPs from an early analysis of WTCCC2 ischemic stroke data. However, for this study we used the entire Immunochip content.

The 6 discovery phase cohorts used the commercially available Immunochip array (Illumina, San Diego, CA). Genotyping for the PROMISe study (the Netherlands) was performed independently in Utrecht, the Netherlands. Genotyping for the remaining 5 case cohorts was performed at the Sanger Centre, Hinxton, Cambridge, UK. Swedish controls were provided and genotyped by the Swedish SLE network, Uppsala, Sweden. Belgian control samples were provided through the efforts of the International Multiple Sclerosis Genetics Consortium. Analysis and quality control (QC) of the PROMISe study (the Netherlands) was performed in Utrecht and for all other cohorts was performed at St George's, University of London, UK.

The Immunochip datasets were each imputed separately to the 1,000 Genomes Phase 1 integrated variant set (March 2012) using IMPUTE v2.2.2.[11] Standard parameters were used with the exception of the number of haplotypes (k), which was increased to 100 to maximize accuracy. In total, between 123,920 and 135,006 SNPs were directly genotyped for the Belgian, German, Polish, Swedish, and British cohorts. After imputation and QC filtering on IMPUTE-info scores with filtering threshold <0.3 and minor allele frequency (MAF) <0.01, there were between 3,601,403 and 4,170,444 autosomal SNPs for final analysis. For the PROMISe study (the Netherlands), stricter QC parameters on the directly genotyped SNPs resulted in 88,511 directly genotyped SNPs and 3,524,203 SNPs for final analysis after imputation. Full details of SNPs at all stages for the discovery cohorts are listed in table e-1. The lambda value (a measure of genomic control to account for overinflation of false-positive results) for the imputed Immunochip cohorts (filtered for heterogeneity or missingness as in the meta-analysis) was $\lambda = 1.165$, equating to $\lambda_{1000} = 1.036$ using the method of de Bakker et al.[12] Selecting the stroke-specific subset of SNPs produced $\lambda = 1.252$ ($\lambda_{1000} = 1.055$) and as a common null set the "reading and writing SNP subset" from WTCCC2 $\lambda = 1.300$ ($\lambda_{1000} = 1.066$), showing little evidence for inflation in the stroke subset or the Immunochip overall. QQ plots are shown in figure e-1.

Association analysis of clean imputed datasets was performed for each Immunochip cohort individually using the frequentist test under an additive model as implemented in SNPTEST v2,[13] including sex and 10 principal components as covariates. Imputed genotype probabilities were taken into account using a missing data likelihood score test or an expectation-maximization method for SNPs with low MAF or high uncertainty. Association analyses were performed on all ischemic stroke cases and for the defined subtypes of large artery stroke, small vessel disease, and cardioembolic stroke.

Meta-analyses were performed using an inverse-variance weighted fixed-effects model as implemented in METAL.[14] SNPs were taken into consideration only if they were present in at least 50% of datasets, were genotyped or imputed in all phases, and if the p value for Cochrane Q test for heterogeneity exceeded 1×10^{-3}. For targeted replication with data from INTERSTROKE and VISP, summary statistics were provided for 2 SNPs (rs17696736 and rs10744777) and meta-analyzed as above.

Conditional analysis. Conditional analysis of the Chr12 locus was performed by including genotype dosage of each of the 10 genome-wide significant SNPs as a covariate in the logistic regression model independently, as well as by inclusion of all 10 SNPs together. This was performed on Immunochip and WTCCC2 data only since individual level genotypes were not available for METASTROKE.

Risk factor defined analysis. A case-only risk factor defined analysis was performed by classifying presence (case) or absence (control) of defined risk factors in the Immunochip and WTCCC2 cohorts (individual genotype level data were not available for the METASTROKE, INTERSTROKE, or VISP cohorts). The 5 cardiovascular risk factors hypertension, diabetes, hypercholesterolemia, coronary artery disease/ischemic heart disease, and smoking status were then assessed independently across the Chr12 locus.

RESULTS **Meta-analysis of Immunochip data.** Analysis of the 6 Immunochip cohorts comprising 3,420 cases and 6,821 controls resulted in identification of 3 SNPs spanning 2 independent loci on chromosomes 10q26 and 19q13 exceeding a genome-wide significance threshold of 5×10^{-8} for all ischemic stroke, and a further 9 SNPs spanning 5 loci in large artery ischemic stroke (top SNPs in each of the 7 loci shown in table 2, block 1, full findings in table e-2). No SNP exceeded a genome-wide threshold in cardioembolic stroke or small vessel disease when examining Immunochip cohorts alone.

Meta-analysis with WTCCC2 and METASTROKE data confirmed previously published associations between the *HDAC9* locus on chromosome 7p21 and large artery stroke, and between *PITX2* and *ZFHX3* loci at 4q25 and 16q22 and cardioembolic stroke. Meta-analysis also revealed a novel locus at 12q24.12 (top SNP rs17696736, $p = 6.06 \times 10^{-10}$). All of these loci exceeded a genome-wide threshold of 5×10^{-8}. The 3 SNPs identified in the Immunochip data alone showed no replication (table 2, block 2).

Targeted in silico replication of the 12q24.12 region in INTERSTROKE and VISP strengthened the association, revealing a new top SNP at this locus (rs10744777, $p = 7.12 \times 10^{-11}$, odds ratio [OR] 1.10 [1.07–1.13]). A forest plot displaying results for all cohorts at rs10744777 is shown in figure 1A. The association was similar across all ischemic stroke subtypes as measured by effect size, with no evidence of subtype specificity (table 3).

Conditional analysis of 12q24.12. The significant SNPs in 12q24.12 spanned 2 Mb of DNA encompassing 16 genes. Across this region, 10 SNPs reached significance levels of $p < 5 \times 10^{-8}$ (figure e-2A). To investigate whether there was a single signal or multiple signals across this locus, we performed a conditional analysis on rs10744777 in those cohorts for which we had genotypic level data (Immunochip and WTCCC2). None of the 9 other genome-wide significant SNPs remained significant after controlling for rs10744777 (table e-3 and figure e-2B). The same effect was

Table 2 Top SNP for each locus exceeding a genome-wide threshold of $5e^{-08}$

rs# ID	Locus	Gene	Subtype	IC	IC + WT + MS	All data
rs8113518	19q13	*KRTDAP*	All IS	$4.20e^{-09a}$	$7.03e^{-02}$	—
rs2281673	10q26	*PLEKHA1*	All IS	$6.30e^{-09a}$	$5.54e^{-06}$	—
rs12323577	14q31	—	LVD	$7.55e^{-19a}$	$4.45e^{-03}$	—
rs17007400	4q31	*ILI5*	LVD	$8.13e^{-16a}$	$1.10e^{-04}$	—
rs4597201	13q31	*NDF1P2*	LVD	$3.33e^{-12a}$	$1.07e^{-03}$	—
rs2930144	3p25	*SLC6A11*	LVD	$6.10e^{-09a}$	$2.71e^{-01}$	—
rs12240184	1p34	*SLC6A9*	LVD	$2.09e^{-08a}$	$7.14e^{-01}$	—
rs17696736	12q24.12	*NAA25*	All IS	$1.17e^{-03}$	$6.06e^{-10a}$	—
rs2023938	7p21	*HDAC9*	All IS	$2.29e^{-03}$	$2.37e^{-08a}$	—
rs12646447	4q25	*PITX2*	All IS	$2.39e^{-01}$	$2.07e^{-08a}$	—
rs2107595	7p21	*HDAC9*	LVD	$1.93e^{-05}$	$2.18e^{-14a}$	—
rs6843082	4q25	*PITX2*	CE	$1.96e^{-03}$	$3.31e^{-18a}$	—
rs879324	16q22	*ZFHX3*	CE	$7.03e^{-03}$	$3.10e^{-10a}$	—
rs10744777	12q24.12	*ALDH2*	All IS	$5.48e^{-06}$	$1.24e^{-09a}$	$7.12e^{-11a}$

Abbreviations: CE = cardioembolic; IC = Immunochip analysis; IS = ischemic stroke; LVD = large vessel disease; MS = METASTROKE; SNP = single nucleotide polymorphism; SVD = small vessel disease; WT = Wellcome Trust Case Control Consortium 2 (WTCCC2) analysis.

All data include INTERSTROKE and VISP replication where applicable. Block 1 represents the lead SNP from 7 loci exceeding $5e^{-8}$ in Immunochip genome-wide association studies. Block 2 represents the lead SNP from 6 loci after meta-analysis. Block 3 shows the lead SNP after in silico replication of rs10744777. All SNPs exceeding $1e^{-06}$ in any analysis are shown in table e-2.

[a] Denotes SNPs exceeding genome-wide significance.

identified when conditioning on each of the 9 other SNPs independently (data not shown).

Risk factor defined analysis. A lack of risk factor data in all controls prevented risk factors from being used in a conventional stratified analysis. However, to examine the effect of underlying risk factors on the 12q24.12 region driving the observed association with all ischemic stroke, we performed a case-only risk factor defined analysis in which cases were subdivided on the basis of presence (case) or absence (control) of 5 available cardiovascular risk factors—hypertension, diabetes mellitus, hypercholesterolemia, smoking, and past history of symptomatic coronary artery disease/ischemic heart disease. These are defined in the online supplementary material. Regional analysis of the Chr12 locus was performed on Immunochip and WTCCC2 data (a lack of individual level genotypes prevented this analysis from being conducted in METASTROKE data). Table 4 shows the lack of association, when classifying cases and controls by the presence of risk factors, for rs10744777. Therefore we found no evidence that the association with stroke was mediated via a single conventional cardiovascular risk factor.

12q24.12 and intracerebral hemorrhage. We also performed in silico replication of the 12q24.12 locus in ICH. Examining 1,088 cases of ICH and 1,102 controls revealed no association with rs10744777 (OR = 1.03, 95% confidence interval 0.90–1.17, $p = 0.695$) (figure 1B).

DISCUSSION Adopting a GWAS approach with direct genotyping and imputation with the Immunochip array, followed by meta-analysis with existing GWAS data in ischemic stroke, we identified a novel risk locus for ischemic stroke on chromosome 12q24.12. Unlike all previous GWAS-identified ischemic stroke loci, this locus does not appear to be associated with a single subtype but rather is associated with ischemic stroke as a whole.

The 12q24.12 locus had been included in the Immunochip due to its association with type 1 diabetes mellitus and was not one of the 3,000 stroke SNPs included from initial analysis of the WTCCC2 stroke study.[15] Type 1 diabetes, although associated with an increased risk of stroke and other premature cardiovascular disease, is rare and therefore accounts for only a very small proportion of total stroke risk on a population basis. In addition to type 1 diabetes, the 12q24.12 locus has been associated with a number of cardiovascular risk factors, including blood pressure and cholesterol levels. To investigate whether the association might be mediated via these risk factors, we performed a case-only risk factor defined analysis, which showed no evidence that the association with stroke was mediated via conventional cardiovascular

Figure 1   Forest plot of rs10744777
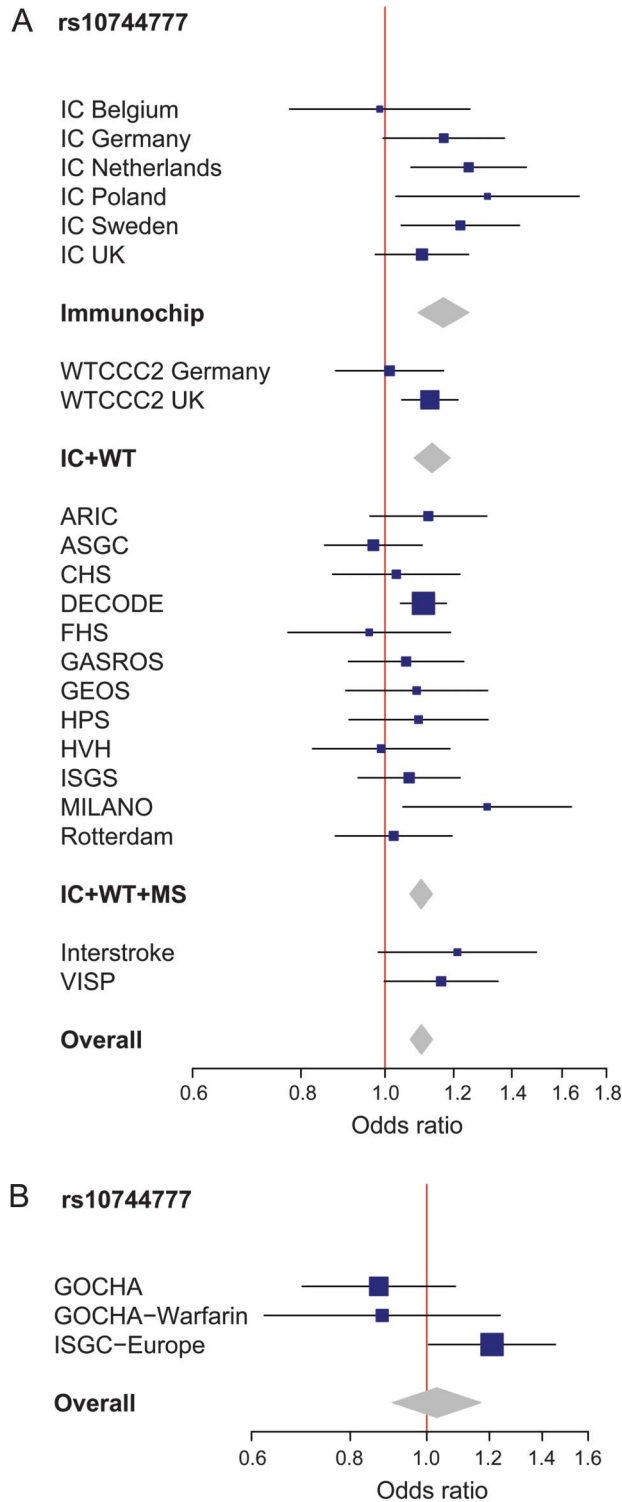
Forest plot for rs10744777 in (A) all ischemic stroke cohorts and (B) intracerebral hemorrhage cohorts. IC = Immunochip; MS = METASTROKE; WT = Wellcome Trust Case Control Consortium 2 (WTCCC2).

The mechanism by which this variant might increase risk of all ischemic stroke without increasing ICH risk is uncertain. However, as the mechanisms of arterial disease differ between stroke subtypes, this finding would be consistent with a systemic risk factor such as altered coagulation rather than a risk factor associated with a single ischemic stroke subtype. Chromosome 12q24.12 has been identified as a region likely to have undergone positive selection in Europeans about 3,000–4,000 years ago, and as such features a complex long-ranging linkage disequilibrium (LD) pattern that does not facilitate identification of a causal variant among the 10 genome-wide significant SNPs identified in this study.[16,17] However, the lead SNP, rs10744777, has been identified as an expression quantitative trait locus for *ALDH2* in monocytes.[18] *ALDH2* codes for mitochondrial aldehyde dehydrogenase 2, which plays a key role in ethanol metabolism but has also emerged as a potentially protective agent in myocardial ischemia.[19] The only nonsynonymous SNP in LD with any of the genome-wide significant variants is rs3184504 in *SH2B3*. rs3184504 has previously been associated with blood pressure and coronary artery disease/myocardial infarction.[20,21] The missense mutation is classified as benign (PolyPhen-2) and tolerated (SIFT); however, it is thought to change 2 transcription factor motifs. *SH2B3* (Src homology 2-B3, also: Lnk) has been implicated in inflammation and innate immunity and has been shown to influence endothelial cell migration and adhesion in vitro.[22] Detailed functional work will be required to elucidate the possibility of these 2 or any of the other 14 genes as potential candidates in ischemic stroke. We did perform a conditional analysis on rs3184504 in the Immunochip and WTCCC2 data, whereby rs10744777 remained significant (OR = 1.12, 95% confidence interval 1.06–1.18, $p = 3.52 \times 10^{-5}$). We therefore cannot exclude the possibility of independent signals in this region underlying the range of reported significant phenotypes.

In addition to the identification of chromosome 12q24.12 as a novel stroke locus, we show for the first time genome-wide significant association of SNPs in *PITX2* and *HDAC9* with all ischemic stroke; previously, associations had only been detected with cardioembolic and large artery stroke subtypes, respectively. This is likely to be an effect of increased sample size, however, as there is no evidence for these loci being risk alleles in subtypes other than cardioembolic stroke and large artery stroke.

There are a number of limitations to this analysis that would benefit from further exploration. The Immunochip is a chip focused preferentially on immune-related genes. As such, this study cannot be considered to be a "full" GWAS in a true sense,

risk factors. Furthermore, given the strong role hypertension plays as a risk factor for hemorrhage, the lack of association with ICH is additional evidence that hypertension does not mediate the association.

Table 3 Association between rs10744777 and ischemic stroke and its subtypes by analysis stage

Subtype	Analysis stage	No. cases	No. controls	p Value	OR (95% CI)	I2	Het p value
All IS	IC	3,420	6,861	$5.48e^{-06}$	1.17 (1.09-1.25)	0.0%	0.46
	IC + WT + MS	15,448	68,865	$1.24e^{-09}$	1.10 (1.06-1.13)	0.0%	0.38
LVD	IC	723	6,861	0.02	1.15 (1.02-1.29)	0.0%	0.50
	IC + WT + MS	2,770	56,020	$1.36e^{-03}$	1.11 (1.04-1.18)	0.0%	0.81
CE	IC	771	5,716	0.02	1.15 (1.03-1.30)	0.0%	0.38
	IC + WT + MS	3,147	61,856	0.03	1.07 (1.01-1.13)	0.0%	0.62
SVD	IC	757	6,861	$4.24e^{-06}$	1.31 (1.17-1.47)	0.0%	0.31
	IC + WT + MS	2,514	58,837	$1.47e^{-05}$	1.15 (1.08-1.23)	39.7%	0.02

Abbreviations: CE = cardioembolic; CI = confidence interval; IC = Immunochip; IS = ischemic stroke; LVD = large vessel disease; MS = METASTROKE; OR = odds ratio; SVD = small vessel disease; WT = Wellcome Trust Case Control Consortium 2 (WTCCC2).

since regions of the genome lacking immune-related genes will not be covered. This effect is negated somewhat by imputation, but it is possible there are other risk alleles for ischemic stroke that remain to be identified in a cohort of this size. Although we included almost 18,000 stroke cases in the largest stroke GWAS meta-analysis to date, the numbers in individual subtypes were smaller; therefore, we cannot completely exclude the possibility that the association with chromosome 12q24.12 is predominantly mediated by a single subtype. Although the effect sizes in the subtype analyses are similar, we are unable to test the possibility of no significant difference between them directly due to the use of shared controls between subtypes. This could be overcome by subtype-specific analyses as a primary endpoint in a future study. We are also unable to definitively exclude the possibility of this association being driven by underlying risk factors we are not powered to detect. The lack of risk factor data in controls prevented a more conventional stratified analysis.

However, we performed a case-only risk factor defined analysis and this showed no evidence that the association with stroke was mediated via conventional cardiovascular risk factors. There is also the limitation of all such case-control studies that controls may become cases in the future, and in the case of stroke it is possible controls may have had a clinically undiagnosed "silent" stroke. Larger sample sizes are one mechanism that can be used to negate this possibility. In line with other genetic studies, the ORs associated with this finding are highly significant but small. Further functional genetic studies will be required to elucidate the mechanism of action associated with this finding, leading to patient benefit.

We have identified, through the largest meta-analysis of ischemic stroke GWAS data to date, a novel locus on Chr12 increasing risk in all subtypes of ischemic stroke but not ICH. Previous GWAS associations with stroke have been subtype specific, and this represents the first genome-wide association increasing risk of all ischemic stroke subtypes.

Table 4 Risk factor defined analysis in Immunochip and WTCCC2 cases

rs10744777	No. cases	No. controls	OR	95% CI	Significance	I2	Het p value
Nonstratified (IC + WT + MS)	15,448	68,865	1.10	1.06-1.13	$1.24e^{-09}$	0.0%	0.38
Nonstratified (IC + WT)	6,968	12,833	1.13	1.08-1.19	$1.75e^{-07}$	0.0%	0.34
	No. cases with risk factor	No. cases without risk factor	OR	95% CI	Significance	I2	Het p value
HT stratified (IC + WT)	4,554	2,364	1.00	0.92-1.08	0.91	0.0%	0.82
Diabetes stratified (IC + WT)	1,256	5,660	1.02	0.93-1.12	0.70	8.0%	0.20
HC stratified (IC + WT)	3,002	3,260	1.01	0.93-1.10	0.81	0.0%	0.85
CAD/IHD stratified (IC + WT)	1,786	4,400	1.07	0.98-1.18	0.12	28.0%	0.09
Smoking stratified (IC + WT)	3,115	3,457	1.02	0.94-1.11	0.57	0.0%	0.91

Abbreviations: CAD/IHD = coronary artery disease/ischemic heart disease; CI = confidence interval; HC = hypercholesterolemia; HT = hypertension; IC = Immunochip; MS = METASTROKE; OR = odds ratio; WT = Wellcome Trust Case Control Consortium 2 (WTCCC2).
Cases are defined as presence of the risk factor; controls are defined as absence of the risk factor. A lack of individual level genotypes and risk factor data for METASTROKE precluded a risk factor analysis in this data.

AUTHOR AFFILIATIONS

From the Stroke and Dementia Research Centre (L.L.K., M.T.), St George's University of London, UK; Department of Neurology and Neurosurgery (S.A., A.A., L.J.K.), Utrecht Stroke Centre, Rudolf Magnus Institute of Neuroscience, Julius Centre for Health Sciences and Primary Care (A.A., P.I.W.d.B.), and Department of Medical Genetics (G.T., P.I.W.d.B.), University Medical Centre, Utrecht, the Netherlands; Center for Human Genetic Research (W.J.D., J.R.), Massachusetts General Hospital, Boston; Program in Medical and Population Genetics (W.J.D., J.R., P.I.W.d.B., M.D.), Broad Institute of Harvard and MIT, Cambridge, MA; Institute for Stroke and Dementia Research (R.M.), Klinikum der Universität München, Ludwig-Maximilians-Universität, Munich, Germany; Department of Clinical Sciences Lund, Neurology (A.L., B.N., B.M.H.), Lund University, and Department of Neurology (A.L., B.N., B.M.H.), Skåne University Hospital, Lund, Sweden; Department of Pathology & Molecular Medicine (G.P.), Population Health Research Institute, McMaster University, Hamilton, Ontario, Canada; Imperial College Cerebrovascular Research Unit (ICCRU) (P.S.), Imperial College London, UK; Department of Neurology (A.S.), Jagiellonian University, Krakow, Poland; Laboratory of Neurobiology (V.T.), Vesalius Research Centre, VIB, Leuven, Belgium; Experimental Neurology and Leuven Research Institute for Neurodegenerative Diseases (LIND) (V.T.) and Laboratory for Neuroimmunology, Section of Experimental Neurology, Department of Neurosciences (A.G.), University of Leuven (KU Leuven), Belgium; Department of Neurology (V.T.), University Hospitals Leuven, Belgium; Institute of Cardiovascular and Medical Sciences (M.W.), University of Glasgow, UK; Departments of Neurology (B.B.W.), Public Health Science (B.B.W.), Medicine (M.M.S.), and Biochemistry and Molecular Genetics (M.M.S.) and Center for Public Health Genomics (M.M.S.), University of Virginia, Charlottesville; Department of Genetics (C.W.), University Medical Centre Groningen, University of Groningen, the Netherlands; Department of Medical Sciences, Molecular Medicine and Science for Life Laboratory (J.K.S.) and Rheumatology (L.R.), Uppsala University, Sweden; Institute of Clinical Molecular Biology (A.F.), Christian-Albrechts University, Kiel, Germany; Division of Clinical Neurosciences and Institute of Genetics and Molecular Medicine (C.S.), University of Edinburgh, UK; Stroke Prevention Research Unit, Nuffield Department of Clinical Neuroscience (P.M.R.), Clinical Trial Service Unit and Epidemiological Studies Unit (R.C., J.C.H.), Wellcome Trust Centre for Human Genetics (M. Farrall), and Department of Cardiovascular Medicine (M. Farrall), University of Oxford, UK; Center for Translational Neuroscience and Mental Health Research (C.L.), University of Newcastle, New South Wales, Australia; School of Medicine and Public Health (C.L., E.G.H.), University of Newcastle and Hunter Medical Research Institute, Newcastle, Australia; University of Texas Health Science Center at Houston (M. Fornage), TX; Departments of Epidemiology, Medicine, and Health Services (B.P.) and Cardiovascular Health Research Unit (J.C.B.), University of Washington, Seattle; Group Health Research Institute (B.P.), Group Health, Seattle, WA; deCODE Genetics (S.G., U.T., K.S.), Reykjavik, Iceland; Faculty of Medicine (U.T., K.S.), University of Iceland, Reykjavik; Department of Neurology (S.S.), Boston University School of Medicine, MA; Departments of Medicine (B.D.M.) and Neurology (S.K.), University of Maryland School of Medicine; Department of Neurology (S.K.), Veterans Affairs Medical Center, Baltimore, MD; Department of Cerebrovascular Disease (G.B.B.), IRCCS Istituto Neurologico Carlo Besta, Milan, Italy; Department of Neurology (J.Meschia), Mayo Clinic, Jacksonville, FL; Departments of Epidemiology, Neurology, and Radiology (M.A.I.), Erasmus University Medical Center, Rotterdam, the Netherlands; Netherlands Consortium for Healthy Aging (M.A.I.), Leiden, the Netherlands; Neurovascular Research Laboratory (J. Montaner), Neurology Department-Universitat Autònoma de Barcelona, Vall d'Hebrón Institute of Research (VHIR) and Hospital, Barcelona, Spain; Division of Genetics (P.I.W.d.B.), Brigham and Women's Hospital, Harvard Medical School, Boston, MA; Munich Cluster for Systems Neurology (SyNergy) (M.D.), Munich, Germany; and Clinical Neurosciences (H.S.M., S.B.), University of Cambridge, UK.

approval, statistical analysis. Joshua C. Bis: analysis or interpretation of data, accepts responsibility for conduct of research and will give final approval, statistical analysis. Giorgio B. Boncoraglio: drafting/revising the manuscript, accepts responsibility for conduct of research and will give final approval, acquisition of data. James Meschia: drafting/revising the manuscript, accepts responsibility for conduct of research and will give final approval, acquisition of data. M. Arfan Ikram: drafting/revising the manuscript, accepts responsibility for conduct of research and will give final approval, acquisition of data, study supervision, obtaining funding. Bjorn M. Hansen: drafting/revising the manuscript, accepts responsibility for conduct of research and will give final approval, acquisition of data. Joan Montaner: study concept or design, analysis or interpretation of data, accepts responsibility for conduct of research and will give final approval, acquisition of data. Gudmar Thorleifsson: analysis or interpretation of data, accepts responsibility for conduct of research and will give final approval, statistical analysis. Kari Stefanson: analysis or interpretation of data, accepts responsibility for conduct of research and will give final approval, acquisition of data, study supervision. Jonathan Rosand: drafting/revising the manuscript, study concept or design, accepts responsibility for conduct of research and will give final approval, acquisition of data, study supervision, obtaining funding. Paul I. W. de Bakker: drafting/revising the manuscript, study concept or design, accepts responsibility for conduct of research and will give final approval, acquisition of data, statistical analysis. Martin Farrall: study concept or design, analysis or interpretation of data, accepts responsibility for conduct of research and will give final approval, acquisition of data, obtaining funding. Martin Dichgans: drafting/revising the manuscript, analysis or interpretation of data, accepts responsibility for conduct of research and will give final approval, acquisition of data. Hugh S. Markus: drafting/revising the manuscript, study concept or design, analysis or interpretation of data, accepts responsibility for conduct of research and will give final approval, acquisition of data, study supervision, obtaining funding. Steve Bevan: drafting/revising the manuscript, study concept or design, analysis or interpretation of data, accepts responsibility for conduct of research and will give final approval, acquisition of data, statistical analysis, study supervision.

STUDY FUNDING

DISCLOSURE

REFERENCES

1. Bellenguez C, Bevan S, Gschwendtnew A, et al. Genetic analysis identifies a new susceptibility locus in HDAC9 for large vessel ischemic stroke, and supports genetic heterogeneity across stroke subtypes. Nat Genet 2012;44: 328–333.
2. Holliday EG, Maguire J, Evans TJ, et al. Common variants at 6p21.1 are associated with large artery atherosclerotic stroke. Nat Genet 2012;44:1147–1151.
3. Traylor M, Farrall M, Holliday E, et al. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. Lancet Neurol 2012;11:951–962.
4. Cortes A, Brown MA. Promise and pitfalls of the immunochip. Arthritis Res Ther 2011;13:101–103.
5. Tousoulis D, Kampoli AM, Papageorgiou N, et al. Pathophysiology of atherosclerosis: the role of inflammation. Curr Pharm Des 2011;17:4089–4110.
6. Grau AJ, Urbanek C, Palm F. Common infections and the risk of stroke. Nat Rev Neurol 2010;6:681–694.
7. O'Donnell M, Xavier D, Diener C, et al. Rationale and design of INTERESTROKE: a global case-control study of risk factors for stroke. Neuroepidemiology 2010;35:36–44.
8. Spence JD, Howard VJ, Chambless LE, et al. Vitamin intervention for stroke prevention (VISP) trial: rationale and design. Neuroepidemiology 2001;20:16–25.
9. Falcone GJ, Biffi A, Devan WJ, et al. Burden of blood pressure-related alleles is associated with larger hematoma volume and worse outcome in intracerebral hemorrhage. Stroke 2013;44:321–326.
10. Falcone GJ, Biffi A, Devan WJ, et al. Burden of risk alleles for hypertension increases risk of intracerebral hemorrhage. Stroke 2012;43:2877–2883.
11. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 2009;5: e1000529.
12. de Bakker PIW, Ferreira MA, Jia X, Neale B, Raychaudhuri S, Voight BF. Practical aspects of imputation-driven meta-analysis of genome-wide association studies. Hum Mol Genet 2008;17:R122–R128.
13. Marchini J, Howie BN, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 2007;39: 906–913.
14. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190–2191.
15. Todd JA, Walker NM, Cooper JD, et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat Genet 2007;39:857–864.
16. Soranzo N, Spector TD, Mangino M, et al. A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in hte HaemGen consortium. Nat Genet 2009;41:1182–1190.
17. Zhernakova A, Elbers C, Ferwerda B, et al. Evolutionary and functional analysis of celiac risk loci reveals SH2B3 as a protective factor against bacterial infection. Am J Hum Genet 2010;86:970–977.
18. Seller T, Wild P, Szymczak S, et al. Genetics and beyond - the transcriptome of human monocytes and disease susceptibility. PLoS One 2010;5:e10693.
19. Budas G, Disatnik M, Mochly-Rosen D. Aldehyde dehydrogenase 2 in cardiac protection: a new therapeutic target? Trends Cardiovasc Med 2009;19:158–164.
20. Gudbjartsson D, Bjornsdottir U, Halapi E, et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. Nat Genet 2009;41:342–347.
21. Deloukas P, Kanoni S, Willenborg C, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet 2012;45:25–33.
22. Devalliere J, Chatelais M, Fitau J, et al. LNK (SH2B3) is a key regulator of integrin signalling in endothelial cells and targets alpha-parvin to control cell adhesion and migration. FASEB J 2012;26:2592–2606.