

# SCIENTIFIC REPORTS



OPEN

## A genome-wide association study links small-vessel ischemic stroke to autophagy

Tsong-Hai Lee<sup>1</sup>, Tai-Ming Ko<sup>2,3</sup>, Chien-Hsiun Chen<sup>2,4</sup>, Yeu-Jhy Chang<sup>1</sup>, Liang-Suei Lu<sup>2</sup>, Chien-Hung Chang<sup>1</sup>, Kuo-Lun Huang<sup>1</sup>, Ting-Yu Chang<sup>1</sup>, Jiann-Der Lee<sup>5</sup>, Ku-Chou Chang<sup>6</sup>, Jen-Tsung Yang<sup>5</sup>, Ming-Shien Wen<sup>1</sup>, Chao-Yung Wang<sup>1</sup>, Ying-Ting Chen<sup>2</sup>, Tsai-Chuan Chen<sup>2</sup>, Shu-Yu Chou<sup>1</sup>, Ming-Ta Michael Lee<sup>7</sup>, Yuan-Tsong Chen<sup>2,8</sup> & Jer-Yuarn Wu<sup>2,4</sup>

Genome-wide association studies (GWAS) can serve as strong evidence in correlating biological pathways with human diseases. Although ischemic stroke has been found to be associated with many biological pathways, the genetic mechanism of ischemic stroke is still unclear. Here, we performed GWAS for a major subtype of stroke—small-vessel occlusion (SVO)—to identify potential genetic factors contributing to ischemic stroke. GWAS were conducted on 342 individuals with SVO stroke and 1,731 controls from a Han Chinese population residing in Taiwan. The study was replicated in an independent Han Chinese population comprising an additional 188 SVO stroke cases and 1,265 controls. Three SNPs (*rs2594966*, *rs2594973*, *rs4684776*) clustered at 3p25.3 in *ATG7* (encoding Autophagy Related 7), with *P* values between  $2.52 \times 10^{-6}$  and  $3.59 \times 10^{-6}$ , were identified. Imputation analysis also supported the association between *ATG7* and SVO stroke. To our knowledge, this is the first GWAS to link stroke and autophagy. *ATG7*, which has been implicated in autophagy, could provide novel insights into the genetic basis of ischemic stroke.

Stroke is known to be the second leading cause of death and a major cause of disability worldwide<sup>1</sup>. Although traditional vascular risk factors such as hypertension, diabetes, atrial fibrillation, and cigarette smoking are common in stroke, stroke incidence<sup>2</sup> and subtype distribution<sup>3</sup> are different among ethnicities. It is possible that a non-traditional risk factor such as genetic predisposition might be important. Data from twin and family history studies have suggested a role for genetic factors in stroke risk<sup>4,5</sup>. A previous study of vascular disease reported that family history is an independent risk factor for SVO, especially in cases presenting before the age of 65, suggesting the involvement of underlying genetic components in the development of SVO<sup>6</sup>. A GWAS conducted using a Japanese cohort with ischemic stroke identified a genetic variant in *PRKCH*<sup>7</sup>; however, a meta-analysis of GWAS data from a Caucasian population found no association between ischemic stroke and *PRKCH* genetic variants<sup>8</sup>.

Compared with recent advances in high throughput genotyping for other subtypes of ischemic stroke<sup>9–11</sup>, gene discovery for SVO has progressed slowly because of etiologic heterogeneity and variations among different ethnic backgrounds. By exploiting these phenotypically more homogeneous classifications, a GWAS conducted using a single population may identify a genetic association specific to SVO. Thus, in the present study, we aimed to identify genetic correlations for SVO using a GWAS within a Han Chinese population.

<sup>1</sup>Chang Gung Memorial Hospital, Linkou Medical Center, and Chang Gung University College of Medicine, Taoyuan, Taiwan. <sup>2</sup>Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan. <sup>3</sup>Graduate Institute of Integrated Medicine, College of Chinese Medicine, China Medical University, Taichung, Taiwan. <sup>4</sup>School of Chinese Medicine, China Medical University, Taichung, Taiwan. <sup>5</sup>Chang Gung Memorial Hospital, Chiayi Branch, and Chang Gung University College of Medicine, Chiayi, Taiwan. <sup>6</sup>Chang Gung Memorial Hospital, Kaohsiung Medical Center, and Chang Gung University College of Medicine, Kaohsiung, Taiwan. <sup>7</sup>Genomic Medicine Institute, Geisinger Health System, Danville, Pennsylvania, USA. <sup>8</sup>Department of Pediatrics, Duke University Medical Center, Durham, North Carolina, USA. Tsong-Hai Lee and Tai-Ming Ko contributed equally to this work. Correspondence and requests for materials should be addressed to Yuan-Tsong Chen (email: [chen0010@ibms.sinica.edu.tw](mailto:chen0010@ibms.sinica.edu.tw)) or Jer-Yuarn Wu (email: [jywu@ibms.sinica.edu.tw](mailto:jywu@ibms.sinica.edu.tw))

## Methods

**Ethical statement.** All methods were performed in accordance with the relevant guidelines and regulations. The study was approved by the Institutional Review Board and the Ethics Committee of the Institutional Review Board of Chang Gung Healthcare System and Academia Sinica, Taiwan. Written informed consents were obtained from the subjects or their family members in accordance with institutional requirements and Declaration of Helsinki principles.

**Study subjects and phenotype definitions.** Individuals with SVO stroke ( $n = 530$ ) (comprising 342 SVO stroke cases from the GWAS and 188 SVO stroke cases from the replication study) were recruited from the three branch hospitals of Chang Gung Healthcare System, Linkou, Chiayi and Kaohsiung, in collaboration with the Translational Resource Center (TRC) for Genomic Medicine of Taiwan. These three branch hospitals cover a population of six million in Taiwan with a total of 3,300 annual ischemic stroke patients. SVO was defined by the presence of subcortical, hypodense lesions with a diameter of  $<15$  mm with accompanying clinical lacunar syndrome. The medical information and blood samples of all cases were centralized in Linkou CGMH, and the SVO stroke subtype was classified according to modified TOAST criteria<sup>12</sup> by single physician, TH Lee, to prevent from interobserver discrepancy. Besides the criteria of clinical presentations and lacunas in brain images, only cases with diameter stenosis  $<30\%$  in extracranial carotid artery confirmed by carotid ultrasound and/or in intracranial carotid artery by angiography (magnetic resonance, computed tomography or digital subtraction) were included for analysis. The control subjects (1,731 in the discovery study and 1,265 in the replication study) from the GWAS were randomly selected from the Taiwan Han Chinese Cell and Genome Bank in Taiwan. These controls were presumably disease-free as reported previously<sup>13</sup>.

**Genotyping and quality control.** Genomic DNA was extracted from blood using a Puregene DNA Isolation Kit (Gentra Systems). Each individual was genotyped using the Axiom Genome-Wide CHB (with a total of 642,832 SNPs) according to manufacturer's protocols by the National Center for Genome Medicine (NCGM) at Academia Sinica. All sample call rates were  $>98.69\%$ , and the mean individual sample call rate was  $99.5 \pm 0.26\%$ . First-degree relatives (parent-offspring and full sibling pairs) in SVO stroke cases and in control samples were identified by kinship analysis and were excluded from further analysis. Genotyping quality control for each SNP was further determined by the total call rate (successful call rate) and MAF in SVO stroke cases and controls. SNPs were excluded from further analysis if only one allele appeared in SVO stroke cases and controls, if the total call rate was  $<0.95$ , or if the total MAF was  $<0.05$  and the total call rate was  $<0.99$ .

**Statistical analysis.** The statistical method used for GWAS analysis has been well-established in our previous study<sup>9,14</sup>. Detection of possible population stratification that could influence association analysis was carried out using EIGENSTRAT 2.0. We estimated the variance inflation factor for genomic controls. Genome-wide association analysis and GC correction were carried out to compare allele and genotype frequencies between cases and controls using the Cochran-Armitage trend test. A quantile-quantile (Q-Q) plot was used to determine  $P$  value distribution (Fig. 1). The adjustment for principle components suggested that inflation was not due to population stratification.

**GWAS validation and replication.** The top three SNPs ( $P < 1 \times 10^{-4}$ ) from the genome-wide association analysis of the 342 SVO stroke cases and controls were further validated using MALDI-TOF mass spectrometry (MassARRAY, Sequenom) (Supplementary Table 1). In addition, we also validated other 30 SNPs ( $P < 1 \times 10^{-2}$ ) in the discovery stage (Supplementary Table 1). SNP genotypes with over 98% success rate and over 98% concordance between the two platforms were then genotyped. An additional 188 SVO stroke cases were used for the replication study.

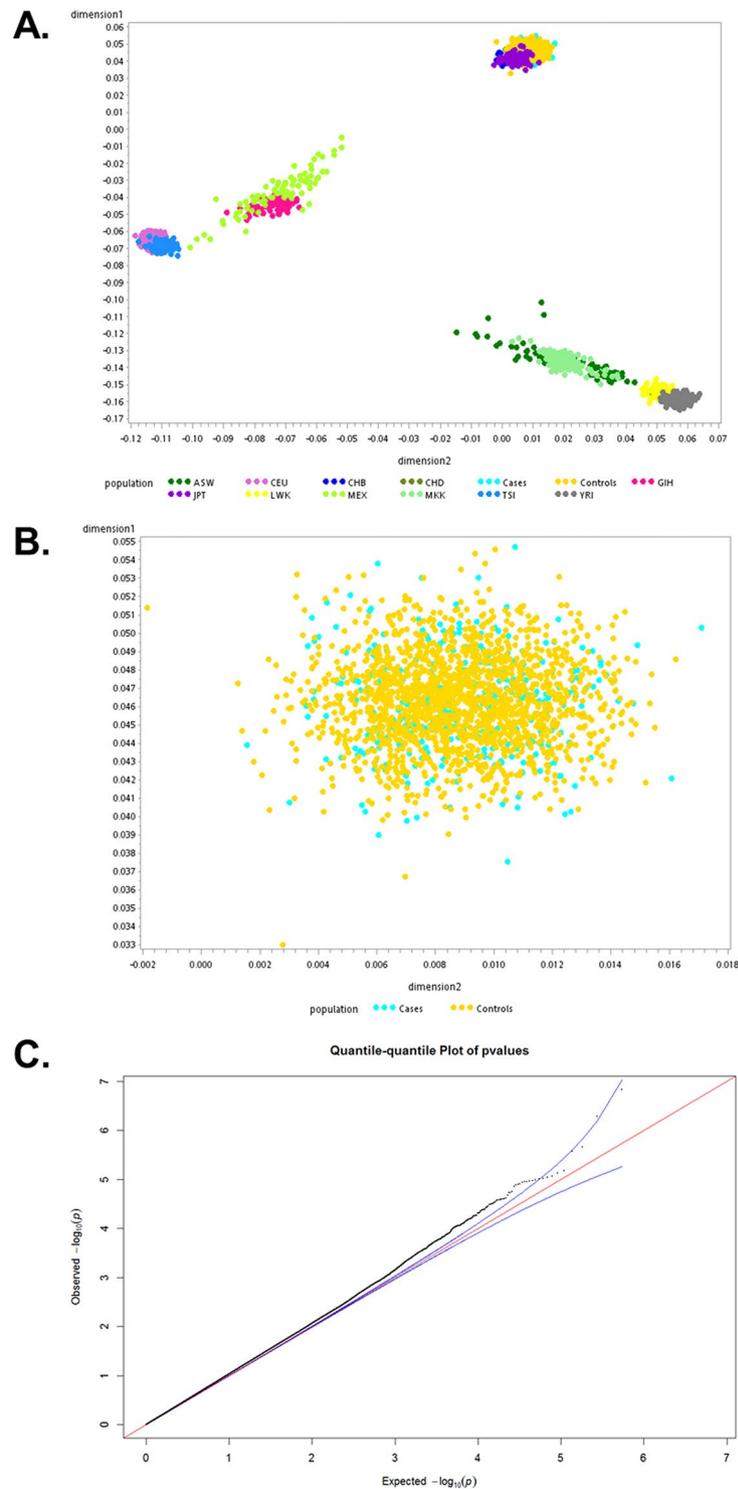
**Imputation.** For enhancement of the coverage, untyped SNPs were imputed by IMPUTE2 using 1000 Genomes reference panel<sup>15-17</sup>. In the pre-phasing step, we set up the haplotypes inferences via SHAPEIT method for optimizing the imputation procedure<sup>18</sup>. For elimination of edge effects, we expanded 500 kb buffer region on each side of imputation region. We determined the uncertainty of imputed genotypes based on likelihood scoring in SNPTEST v2 and frequentist association test of the additive model. We further validated the top imputed SNP by direct genotyping.

**Data availability.** The datasets generated during and/or analysed during the current study are available in the International Stroke Genetics Consortium repository, <http://www.cerebrovascularportal.org/>. The name of the dataset is "SVO-Han-population Taiwan-NCGM".

## Results

**Study populations.** Characteristics of the SVO stroke groups are shown in Table 1. The mean age of the SVO group in the GWAS (discovery) group and replication group was 57.9 years old and 56.0 years old, respectively. The ratio of males to females in the GWAS group and replication groups was 67.3% and 67.6%, respectively. For other stroke risk factors, there was no significant difference between the GWAS group and the replication group.

**Assessment of population stratification.** We performed a case-control GWAS to identify loci associated with increased risk of small-vessel ischemic stroke in the Han Chinese population using an Affymetrix Axiom CHB array containing 642,832 SNP probes. We initially enrolled 342 SVO stroke and 1,731 controls from a Han Chinese population residing in Taiwan. After kinship analysis and strict quality control filtering, we analyzed 552,090 SNPs (representing 87.9% of array SNPs) for the samples from the GWAS group. Multidimensional scaling analysis (Fig. 1) and results of permutation tests for identity-by-state revealed no differences between the

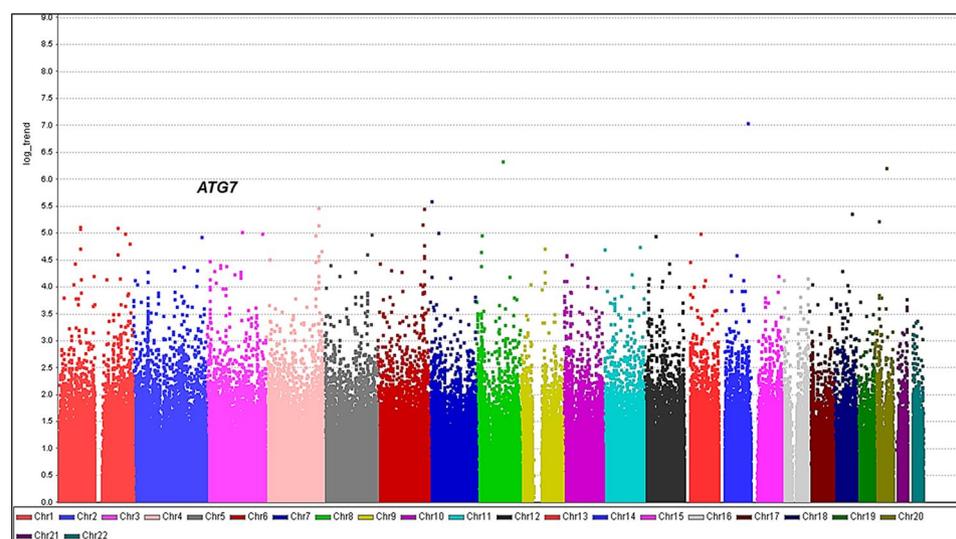


**Figure 1.** Multidimensional scaling analysis. **(A)** Results of the multidimensional scaling analysis of the GWAS samples with HapMap populations. **(B)** Results of the multidimensional scaling analysis of the GWAS samples with the GWAS samples only. **(C)** Q-Q plot of the  $P$  values in a Cochran-Armitage trend test. Lambda value is 1.09.

SVO and control groups, providing no evidence for strong population stratification. Quantile-quantile (Q-Q) plots were used to examine  $P$  value distributions (Fig. 1), and the lambda value was 1.09. In total, we found three top validated SNPs associated with SVO ( $P < 1 \times 10^{-4}$ ) (Supplementary Table 1). The IBS sharing method implemented in PLINK showed no cryptic family relationships among SVO stroke cases and controls.

	n		Discovery (n = 342)	n		Replication (n = 188)
Median Age (IQR)	342	57.9	(48.3 – 67.0)	183	56	(44.0 – 67.0)
Sex-Male (%)	342	67.30%		188	67.60%	
Hypertension (%)	340	77.60%		187	78.10%	
Diabetes mellitus (%)	340	29.10%		187	32.60%	
Alcohol (%)	329	22.20%		173	15.00%	
Family history of stroke (%)	327	41.90%		166	42.80%	
Median HDL-C (IQR)	331	44	(36.0 – 53.0)	175	42	(35.0 – 49.0)
Median LDL-C (IQR)	328	110	(93.0 – 135.0)	174	115	(91.0 – 137.0)
Median VLDL (IQR)	133	26	(19.0 – 42.0)	91	30	(20.0 – 37.3)
Median Triglyceride (IQR)	341	131	(99.0 – 188.0)	188	145	(105.0 – 185.5)
Median Cholesterol (IQR)	341	185	(164.0 – 212.0)	188	187.5	(162.8 – 216.3)
Median Uric acid (IQR)	329	5.9	(4.9 – 7.1)	177	6.6	(5.1 – 7.6)

**Table 1.** Baseline Demographic Summary of Patients.



**Figure 2.** Results of genome-wide association analysis ( $-\log_{10} P$ ) shown in chromosomal order for 552,090 SNPs tested for association in initial samples from 342 patients with SVO stroke and 1,731 controls. X-axis represents each of the SNPs used in the primary scan. Y-axis represents the  $-\log_{10} P$  value of the trend test. Signals in the *ATG7* loci are indicated.

**GWAS and cross-platform validation.** The analysis was first performed with samples from 342 individuals with SVO stroke and 1,731 controls (Fig. 2 and Table 2). Qualified Affymetrix calling (>99%) of clustering in both SVO stroke cases and controls, located in (or within) 50 kb of known genes, was selected for cross-platform validation using a Sequenom MassARRAY or direct sequencing.

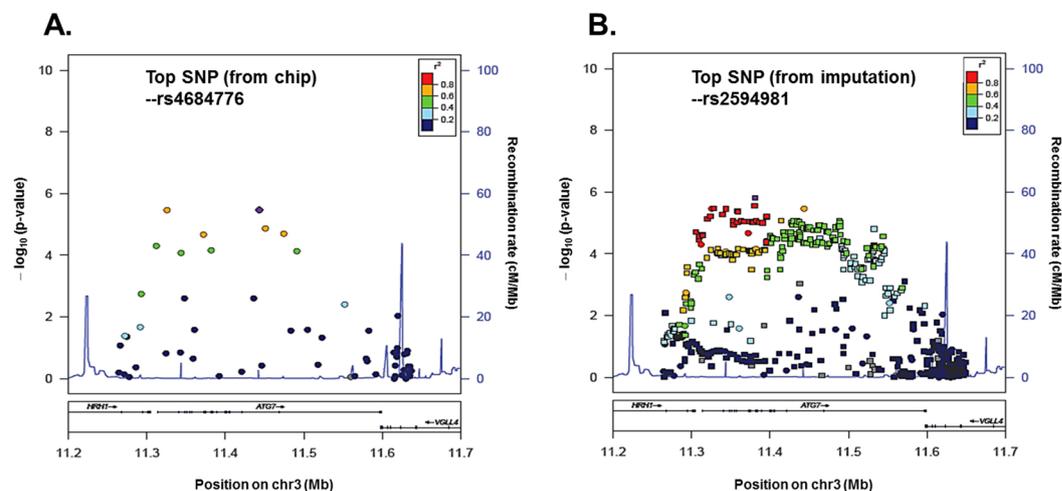
**Replication of top variants for SVO.** In the replication stage, 33 SNPs (Supplementary Table 1) were replicated in an independent cohort of 188 patients with SVO stroke and 1,265 controls (Supplementary Table 2). In a combined analysis of the GWAS and replication cohorts,  $P$  values for 8 of the identified SNPs were lower than  $10^{-5}$  (Table 2).

We found that the SNPs rs2594966 ( $P = 2.52 \times 10^{-6}$ ), rs2594973 ( $P = 2.52 \times 10^{-6}$ ), and rs4684776 ( $P = 2.52 \times 10^{-6}$ ) located at 3p25.3 in *ATG7* (encoding Autophagy Related 7). Other 5 SNPs were located at 3q27.3 (rs4686799,  $P = 8.9 \times 10^{-6}$ ) in *KNG1* (encoding Kininogen-1), at 4q28.1 (rs78868369,  $P = 7.6 \times 10^{-6}$ ) in *ANKRD50* (encoding Ankyrin repeat domain-containing protein 50), at 6q27 (rs536348,  $P = 1.1 \times 10^{-6}$ ) in *LOC441178*, at 8q22.3 (rs17201317,  $P = 6.7 \times 10^{-6}$ ) in *ATP6VIC1* (encoding ATPase H<sup>+</sup> transporting V1 subunit C1), and at 14q31.1 (rs11846182,  $P = 9.5 \times 10^{-6}$ ) in *DIO2-AS1* (encoding DIO2 antisense RNA 1), respectively. These were all replicated in the independent Han Chinese population (Table 1).

In addition, to enhance the SNP coverage, whole gene region of *ATG7* was identified using discovery GWAS dataset (Fig. 3). The imputation demonstrated a strong association within one LD with identified the top SNPs including rs2594966, rs2594973, and rs4684776. The top imputed SNP, rs2594981, was further validated by direct genotyping in discovery GWAS. These data also supported the association between *ATG7* and SVO stroke.

Chr.	SNP	Position	Gene	Allele format	Risk allele	Stage	control/case	RAF controls	RAF cases	Trend P	OR	95%CI	
3	rs2594966	11325276	ATG7	GA		GWAS	1731/342	0.6155	0.6959	7.00E-05	1.43	1.198	1.706
				GA	A	Replication	1265/188	0.6105	0.6765	1.31E-02	1.334	1.059	1.681
				GA	Combined	2996/530	0.6133	0.689	2.52E-06	1.397	1.214	1.607	
3	rs2594973	11395821	ATG7	CG		GWAS	1731/342	0.5948	0.6795	3.24E-05	1.445	1.212	1.722
				CG	G	Replication	1265/188	0.5918	0.6489	3.38E-02	1.275	1.017	1.598
				CG	Combined	2996/530	0.5935	0.6686	3.81E-06	1.381	1.203	1.587	
3	rs4684776	11443223	ATG7	CT		GWAS	1731/342	0.6096	0.6877	1.08E-04	1.41	1.183	1.681
				CT	T	Replication	1265/188	0.6067	0.6729	1.19E-02	1.334	1.06	1.678
				CT	Combined	2996/530	0.6084	0.6824	3.59E-06	1.383	1.203	1.59	
3	rs4686799	186451236	KNG1	TC		GWAS	1731/342	0.6598	0.7108	9.95E-03	1.268	1.057	1.521
				TC	C	Replication	1265/188	0.6454	0.7486	1.06E-04	1.637	1.276	2.099
				TC	Combined	2996/530	0.6537	0.7244	8.91E-06	1.392	1.202	1.612	
4	rs78868369	126077143	ANKRD50, FAT4	TC		GWAS	1731/342	0.08449	0.1268	4.81E-04	1.574	1.219	2.033
				TC	T	Replication	1265/188	0.08796	0.133	4.55E-03	1.59	1.146	2.208
				TC	Combined	2996/530	0.08595	0.129	7.61E-06	1.575	1.288	1.927	
6	rs536348	168098215	LOC441178, C6orf123	TC		GWAS	1731/342	0.6407	0.7132	2.52E-04	1.395	1.165	1.67
				TC	C	Replication	1265/188	0.6395	0.7234	1.26E-03	1.474	1.16	1.874
				TC	Combined	2996/530	0.6402	0.7169	1.12E-06	1.423	1.232	1.643	
8	rs17201317	104096256	ATP6V1C1, BAALCOS	CT		GWAS	1731/342	0.1096	0.1485	3.45E-03	1.417	1.118	1.796
				CT	C	Replication	1265/188	0.1093	0.1738	2.99E-04	1.715	1.276	2.304
				CT	Combined	2996/530	0.1095	0.1575	6.67E-06	1.521	1.265	1.829	
14	rs11846182	80872021	DIO2-AS1	CT		GWAS	1731/342	0.602	0.6754	2.58E-04	1.376	1.156	1.638
				CT	T	Replication	1265/188	0.6135	0.6818	9.98E-03	1.35	1.07	1.702
				CT	Combined	2996/530	0.6069	0.6777	9.48E-06	1.362	1.185	1.565	

**Table 2.** SNPs with  $P$  values  $< 1 \times 10^{-5}$  in the Joint Analysis. Chr., chromosome; gene, genes containing the SNP or the closest gene up to 50 kb upstream or downstream of the SNP; RAF controls, risk allele frequency in controls; RAF cases, risk allele frequency in SVO stroke; OR, odds ratio; 95% CI, 95% confidence interval.



**Figure 3.** Association plots for the *ATG7* locus. Regional association plot for the *ATG7* locus on chromosome 3 (A) with gene annotations superimposed. Each SNP is plotted with respect to its chromosomal location (x-axis) and its  $-\log_{10} P$  values (left y-axis) for the trend test from the primary GWAS scan at that region of the chromosome. After imputation (B), squares represent imputed SNPs, and circles represent genotyped SNPs. Colors denote the strength of the linkage disequilibrium of the SNPs to *ATG7*.

## Discussion

In this study, we identified novel genetic variants associated with SVO stroke susceptibility. This represents the first report of a GWAS for SVO stroke conducted on a Han Chinese population. Based on two independent Han Chinese groups and without significant difference in other risk factors, several novel loci for SVO stroke were identified and replicated. These findings suggest that SVO is a heritable trait and provide new insights into the genetic basis of SVO stroke.

Although there have been studies showing that autophagy may be involved in stroke<sup>19,20</sup>, the genetic association between autophagy-related genes and SVO stroke was never reported. In the present study, we identified five SVO stroke-associated SNPs—rs2594966, rs2594973, rs4684776, rs34843621, and rs12637318—in the same linkage disequilibrium (LD) of chromosome region 3p25.3. The SNPs are located in the *ATG7* gene (Supplementary Fig. 1), which encodes an ubiquitin-activating E1-like enzyme critical for autophagy<sup>21</sup>. This newly identified genetic link may reveal novel molecular insights into the pathogenesis of SVO stroke. A recent study revealed that deletion of *ATG7* was strongly protective against neuronal damage in the brain<sup>22</sup>. A megakaryocyte- and platelet-specific deletion of *ATG7* caused modest defects in platelet aggregation and granule cargo packaging in a mouse model<sup>23</sup>, suggesting that *ATG7* may be a contributing factor for thrombosis. Moreover, *ATG7*-dependent autophagy has been related to hepatic lipid metabolism<sup>24</sup>, which has in turn been found to be associated with stroke<sup>25,26</sup>. It is therefore possible that enhancement of SVO stroke by alteration of the lipid profile may be mediated by *ATG7*-dependent autophagy. The potential role of *ATG7* in occlusion and SVO stroke pathogenesis will require further investigation. Another SVO-associated SNP, rs4686799, was identified within an intron of the *KNG1* gene. *KNG1* was identified from a GWAS for plasma factor XI levels as a genetic determinant of activated partial thromboplastin time<sup>27</sup>. Elevated plasma levels of FXI have been correlated with venous thrombosis and ischemic stroke; therefore, *KNG1* could be involved in the pathogenesis of SVO stroke via regulation of plasma FXI levels.

A meta-analysis showed that SNP 1425 G/A in *PRKCH* was associated with ischemic stroke, particularly lacunar infarction, in Chinese and Japanese populations<sup>28</sup> and *ALDH2*<sup>29,30</sup> and *FOXP2*<sup>29,30</sup> with small vessel disease and white matter hyperintensity in Caucasians, respectively. The disparity in risk genes between Asian and Caucasian populations could be due to the inconsistency in diagnostic and stroke subtyping criteria among centers and nations. The current study was conducted in a single healthcare system using similar diagnostic tools, and the SVO subtype was classified by single doctor.

A major limitation of the current study could be the sample size. However, because SVO was reported to be common in patients with intracranial artery stenosis<sup>31</sup>, the current study used strict recruitment criteria to exclude cases with extracranial and/or intracranial artery stenosis; hence, the sample size was reduced. An additional independent larger group, such as the International Stroke Genetics Consortium, CHARGE, or METASTROKE, would strengthen our findings with detailed stroke subtyping. Further genome-wide association tests assessing whether the potential susceptibility loci have genome-wide significance in different populations will elucidate the genetic contribution in SVO stroke pathogenesis.

In this study, we provide the first genome-wide evidence showing in two independent cohorts, thirty-three SNPs located in novel genetic loci that were found to be associated with SVO stroke in a Han Chinese population. The novel risk loci for SVO stroke contained genes, especially for *ATG7*, that have been implicated in autophagy and thrombosis, which may provide insights into future studies to identify the therapeutic targets for SVO stroke.

## References

- Feigin, V. L. *et al.* Global and regional burden of stroke during 1990–2010: findings from the Global Burden of Disease Study 2010. *Lancet* **383**, 245–54 (2014).
- Feigin, V. L., Lawes, C. M., Bennett, D. A., Barker-Collo, S. L. & Parag, V. Worldwide stroke incidence and early case fatality reported in 56 population-based studies: a systematic review. *Lancet Neurol* **8**, 355–69 (2009).
- White, H. *et al.* Ischemic stroke subtype incidence among whites, blacks, and Hispanics: the Northern Manhattan Study. *Circulation* **111**, 1327–31 (2005).
- Flossmann, E., Schulz, U. G. & Rothwell, P. M. Systematic review of methods and results of studies of the genetic epidemiology of ischemic stroke. *Stroke* **35**, 212–27 (2004).
- Markus, H. S. Unravelling the genetics of ischaemic stroke. *PLoS Med* **7**, e1000225 (2010).
- Jerrard-Dunne, P., Cloud, G., Hassan, A. & Markus, H. S. Evaluating the genetic component of ischemic stroke subtypes: a family history study. *Stroke* **34**, 1364–9 (2003).
- Kubo, M. *et al.* A nonsynonymous SNP in *PRKCH* (protein kinase C eta) increases the risk of cerebral infarction. *Nat Genet* **39**, 212–7 (2007).
- Traylor, M. *et al.* Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol* **11**, 951–62 (2012).
- Lee, T. H. *et al.* Identification of *PTCSC3* as a Novel Locus for Large-Vessel Ischemic Stroke: A Genome-Wide Association Study. *J Am Heart Assoc* **4**, e003003 (2016).
- Holliday, E. G. *et al.* Common variants at 6p21.1 are associated with large artery atherosclerotic stroke. *Nat Genet* **44**, 1147–51 (2012).
- International Stroke Genetics, C. *et al.* Genome-wide association study identifies a variant in *HDAC9* associated with large vessel ischemic stroke. *Nat Genet* **44**, 328–33 (2012).
- Adams, H. P. Jr. *et al.* Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke* **24**, 35–41 (1993).
- Pan, W. H. *et al.* Han Chinese cell and genome bank in Taiwan: purpose, design and ethical considerations. *Hum Hered* **61**, 27–30 (2006).
- Lee, Y. C. *et al.* Two new susceptibility loci for Kawasaki disease identified through genome-wide association analysis. *Nat Genet* **44**, 522–5 (2012).
- Howie, B., Marchini, J. & Stephens, M. Genotype imputation with thousands of genomes. *G3 (Bethesda)* **1**, 457–70 (2011).
- Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* **5**, e1000529 (2009).
- Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* **39**, 906–13 (2007).
- Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G. R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* **44**, 955–9 (2012).
- Tian, F. *et al.* *In vivo* imaging of autophagy in a mouse stroke model. *Autophagy* **6**, 1107–14 (2010).
- Puyal, J. & Clarke, P. G. Targeting autophagy to prevent neonatal stroke damage. *Autophagy* **5**, 1060–1 (2009).
- Hong, S. B. *et al.* Insights into noncanonical E1 enzyme activation from the structure of autophagic E1 Atg7 with Atg8. *Nat Struct Mol Biol* **18**, 1323–30 (2011).
- Xie, C. *et al.* Neuroprotection by selective neuronal deletion of Atg7 in neonatal brain injury. *Autophagy* **12**, 410–23 (2016).

23. Ouseph, M. M. *et al.* Autophagy is induced upon platelet activation and is essential for hemostasis and thrombosis. *Blood* **126**, 1224–33 (2015).
24. Xiao, Y. *et al.* MAPK1/3 regulate hepatic lipid metabolism via ATG7-dependent autophagy. *Autophagy*, 1–2 (2016).
25. Adibhatla, R. M., Dempsy, R. & Hatcher, J. F. Integration of cytokine biology and lipid metabolism in stroke. *Front Biosci* **13**, 1250–70 (2008).
26. Zhang, L. *et al.* Proprotein convertase subtilisin/kexin type 9 (PCSK9) in lipid metabolism, atherosclerosis and ischemic stroke. *Int J Neurosci* **126**, 675–80 (2016).
27. Sabater-Lleal, M. *et al.* A genome-wide association study identifies KNG1 as a genetic determinant of plasma factor XI Level and activated partial thromboplastin time. *Arterioscler Thromb Vasc Biol* **32**, 2008–16 (2012).
28. Li, J., Luo, M., Xu, X. & Sheng, W. Association between 1425G/A SNP in PRKCH and ischemic stroke among Chinese and Japanese populations: a meta-analysis including 3686 cases and 4589 controls. *Neurosci Lett* **506**, 55–8 (2012).
29. NINDS Stroke Genetics Network (SiGN) & International Stroke Genetics Consortium (ISGC). Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. *Lancet Neurol* **15**, 174–184 (2016).
30. Neurology Working Group of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium, Stroke Genetics Network (SiGN) & International Stroke Genetics Consortium (ISGC). Identification of additional risk loci for stroke and small vessel disease: a meta-analysis of genome-wide association studies. *Lancet Neurol* **15**, 695–707 (2016).
31. Kwon, H. M. *et al.* Frequency, Risk Factors, and Outcome of Coexistent Small Vessel Disease and Intracranial Arterial Stenosis: Results From the Stenting and Aggressive Medical Management for Preventing Recurrent Stroke in Intracranial Stenosis (SAMMPRIS) Trial. *JAMA Neurol* **73**, 36–42 (2016).

## Acknowledgements

We thank all affected individuals and their families who devoted their time and effort to participate in this study. We thank doctors in the three branch hospitals of Chang Gung Healthcare System, Taiwan, for their contributions in recruiting stroke patients. We gratefully acknowledge the members of the Translational Resource Center (TRC) (NSC102-2325-B-001-023) of National Research Program for Biopharmaceuticals (NRPB) and the National Center for Genome Medicine (NCGM) (NSC102-2319-B-001-001) of National Core Facility Program for Biotechnology (NCFPB), National Science Council, at Academia Sinica for their support in subject recruitment, genotyping, and statistical analysis. This study was supported by the Academia Sinica Genomic Medicine Multicenter Study, Taiwan (40-05-GMM) and Chang Gung Memorial Hospital (CMRPG35072, 35073, 39082, 39083, CMRPG3B0611, CMRPG3A0352, CMRPG3B0111 and BMRP 274). The funders had no role in study design, data collection or analysis, the decision to publish or preparation of the manuscript.

## Author Contributions

T.-H.L., Y.-T.C., and J.-Y.W. are the principal investigators who conceived and obtained funding for this project. T.-M.K., C.-H.C. and J.-Y.W. organised and supervised the GWAS and replication genotyping pipeline and devised the overall analysis plan. T.-M.K. wrote the first draft of the manuscript. T.-H.L., T.-M.K., Y.-T.C., and J.-Y.W. contributed to the writing of the manuscript. Replication and validation data were provided by Y.-T.C and L.-S.L., T.-H.L., Y.-J.C., C.-H.C., K.-L.H., T.-Y.C., J.-D.L., K.-C.C., J.-T.Y., M.-S.W., C.-Y.W., T.-C.C., S.-Y.C., M.-T.L., Y.-T.C., and J.-Y.W. coordinated and contributed subject and database.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-017-14355-3>.

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017