MULTITRAIT ANALYSIS EXPANDS GENETIC RISK FACTORS IN CARDIOEMBOLIC STROKE

Running title: Multitrait analysis of cardioembolism

Jara Cárcel-Márquez^{1,2*} MSc, Elena Muiño^{1*} PhD, Cristina Gallego-Fabrega^{1,3} PhD, Natalia Cullell^{1,4} MSc, Miquel Lledós¹ MSc, Laia Llucà-Carol¹ MSc, Tomás Sobrino⁵ PhD, Francisco Campos⁵ PhD, José Castillo⁵ PhD, Marimar Freijo⁶ PhD, Juan Francisco Arenillas⁷ PhD, Victor Obach⁸ PhD, José Álvarez-Sabín⁹ PhD, Carlos A. Molina⁹ PhD, Marc Ribó⁹ PhD, Jordi Jiménez-Conde¹⁰ PhD, Jaume Roquer¹⁰ PhD, Lucia Muñoz-Narbona¹¹ PhD, Elena Lopez-Cancio¹² PhD, Mònica Millán¹¹ PhD, Rosa Diaz-Navarro¹³ PhD, Cristòfol Vives-Bauza¹³ PhD, Gemma Serrano-Heras¹⁴ PhD, Tomás Segura¹⁴ PhD, Laura Ibañez¹⁵ PhD, Laura Heitsch^{16,17} PhD, Pilar Delgado¹⁸ PhD, Rajat Dhar¹⁷ PhD, Jerzy Krupinski¹⁹ PhD, Raquel Delgado-Mederos³ PhD, Luis Prats-Sánchez³ PhD, Pol Camps-Renom³ PhD, Natalia Blay²⁰ PhD, Lauro Sumoy²¹ PhD, Rafael de Cid²⁰ PhD, Joan Montaner²² PhD, Carlos Cruchaga^{15,23} PhD, Jin-Moo Lee¹⁷ PhD, Joan Martí-Fàbregas³ PhD, Israel Fernández-Cadenas^{1†} PhD. On behalf GeneStroke Consortium and International Stroke Genetics Consortium.

1. Stroke Pharmacogenomics and Genetics Group, Biomedical Research Institute Sant Pau (IIB Sant Pau), Barcelona, Spain
2. Universitat Autònoma de Barcelona, Departament de Medicina
3. Stroke Unit, Department of Neurology, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain
4. Stroke Pharmacogenomics and Genetics Laboratory, Fundació Docència I Recerca Mútua Terrassa, Hospital Mútua Terrassa, Terrassa, Spain
5. Health Research Institute of Santiago de Compostela (IDIS), Clinical Neurosciences Research Laboratory, Hospital Clínico Universitario de Santiago de Compostela, Santiago de Compostela, Spain
6. Biocruces-Bizkaia Health Research Institute. Department of Neurology, Bilbao, Spain
7. Stroke Unit, Department of Neurology, University Hospital of Valladolid, Valladolid, Spain
8. Department of Neurology, Hospital Clínic de Barcelona, IDIBAPS, Barcelona, Spain

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

9. Stroke Unit, Department of Neurology, Hospital Universitari Vall d'Hebron, Barcelona, Spain
10. Department of Neurology, IMIM-Hospital del Mar; Neurovascular Research Group, IMIM (Institut Hospital del Mar d'Investigacions Mèdiques); Universitat Autònoma de Barcelona/DCEXS-Universitat Pompeu Fabra, Barcelona, Spain
11. Department of Neurosciences, Hospital Germans Trias i Pujol, Universitat Autònoma de Barcelona, Barcelona, Spain
12. Departament of Neurology, University Hospital Central de Asturias (HUCA). Oviedo, Spain
13. Department of Neurology, Son Espases University Hospital, Illes Balears Health Research Institute (IdISBa), Palma, Spain
14. Department of Neurology, University Hospital of Albacete, Albacete, Spain
15. Department of Psychiatry, Washington University School of Medicine, Saint Louis, MO, USA
16. Department of Emergency Medicine, Washington University School of Medicine, Saint Louis, MO, USA
17. Department of Neurology, Washington University School of Medicine, Saint Louis, MO, USA
18. Neurovascular Research Laboratory, Vall d'Hebron Institute of Research, Universitat Autònoma de Barcelona, Barcelona, Spain
19. Neurology Service, Hospital Universitari Mútua Terrassa, Terrassa, Spain
20. GenomesForLife-GCAT Lab Group, Germans Trias i Pujol Research Institute (IGTP), Badalona, Spain
21. High Content Genomics and Bioinformatics Unit, Germans Trias i Pujol Research Institute (IGTP), Badalona, Spain
22. Institute de Biomedicine of Seville, IBiS/Hospital Universitario Virgen del Rocío/CSIC/University of Seville & Department of Neurology, Hospital Universitario Virgen Macarena, Seville, Spain
23. Neurogenomics and Informatics Center at Washington University in St. Louis

* These authors contributed equally to this work.

†Corresponding author: Israel Fernández-Cadenas

Address: C/ Sant Antoni M^aClaret, 167, 08025 Barcelona, Spain

Phone: +34 932746000

e-mail: israelcadenas@yahoo.es

Twitter: @PharmaEpigenLab

Figure 1-4; Table 1

Word count: 5,986

ABSTRACT

Background and Purpose: The genetic architecture of cardioembolic stroke (CES) is still poorly understood. Atrial fibrillation (AF) is the main cause of CES, with which it shares heritability. We aimed to discover novel loci associated with CES by performing a Multitrait Analysis of the GWAS (MTAG) with atrial fibrillation genetic data.

Methods: For the MTAG analysis we used the MEGASTROKE cohort, which comprises European patients with CES and controls (n=362,661) and an AF cohort composed of 1,030,836 subjects. Regional genetic pleiotropy of the significant results was explored using an alternative Bayesian approach with GWAS-pairwise method. A replication was performed in an independent European cohort comprising 9,105 subjects using a Genome Wide Association Study (GWAS).

Results: MTAG-CES analysis revealed 40 novel and significant loci ($p\text{-value} < 5 \times 10^{-8}$) associated with CES, four of which had not previously been associated with AF. A significant replication was assessed for eight novel loci: *CAVI*, *IGF1R*, *KIAA1755*, *NEURL1*, *PRRX1*, *SYNE2*, *TEX41* and *WIPF1*, showing a $p\text{-value} < 0.05$ in the CES vs controls independent analysis. *KIAA1755*, a locus not previously described associated with AF. Interestingly, 51 AF risk loci were not associated with CES according to GWAS-pairwise analysis. Gene Ontology (GO) analysis revealed that these exclusive AF genes from the 51 loci participate in processes related mainly to cardiac development, whereas genes associated with AF and CES participate mainly in muscle contraction and the conduction of electrical impulses.

Conclusions: We found eight new loci associated with CES. In addition, this study provides novel insights into the pathogenesis of CES, highlighting multiple candidate genes for future functional experiments.

Non-standard Abbreviations and Acronyms.

AF: Atrial fibrillation

CES: Cardioembolic stroke

GWAS: Genome-wide association study

IS: Ischemic stroke

MTAG: Multitrait analysis of the GWAS

SNP: Single nucleotide polymorphism

QC: Quality control

INTRODUCTION

Cardioembolic stroke (CES) represents $\approx 25\%$ of all ischemic strokes (IS), is responsible for the most severe ones and its prevalence could triple by 2050¹. The main cause of CES is atrial fibrillation (AF), and the use of anticoagulants in AF patients could prevent $\approx 70\%$ of strokes¹, more than with antiplatelet treatment². Moreover, occult paroxysmal AF could be the cause of up to 16% of undetermined strokes³. Nevertheless, not everybody with AF will suffer a stroke, and the use of anticoagulation involves an increased risk of major bleeding.

Therefore, understanding the mechanisms that are common to both CES and AF patients in relation to the risk of having a stroke would make it possible to develop specific, more effective therapies with fewer side effects.

To date, there are only eight loci associated with CES: *PITX2*⁴, *ZHFX3*⁴, *RGS7*⁴, *NKX2-5*⁴, *ABO*^{4,5}, *GNAO1*⁶, *PHF20*⁶ and locus 5q22.3⁶. However, even the authors doubted the reliability of *RGS7* as a stroke-associated locus. Interestingly, the two most significant susceptibility loci (*PITX2* and *ZHFX3*) for CES are also the two most strongly associated with AF^{4,7}; in fact CES and AF have common heritability⁸.

Multitrait analysis of the GWAS (MTAG) integrates the information contained in correlated traits to boost genetic association signals for each individual trait, providing summary-level statistics⁹. This strategy has been successfully followed to identify novel loci associated with glaucoma¹⁰, bone mineral density¹¹ and small vessel stroke¹².

Our objective was to identify novel loci associated with CES and understand the biological mechanisms shared between CES and AF by performing a MTAG and additional GWAS analysis.

METHODS

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Cohorts' description

The summary statistics for the Cardioembolic Stroke analysis were obtained from the MEGASTROKE analysis (MEGASTROKE-CES) through the Cerebrovascular Disease Knowledge Portal (<http://cerebrovascularportal.org>). This cohort was composed of 7,193 CES patients and 355,468 controls of European ancestry. The summary statistics of the Atrial Fibrillation 2018 analysis (AF-2018) were obtained from the GWAS catalog portal (<https://www.ebi.ac.uk/gwas/>). The AF-2018 cohort was composed of 60,620 AF cases and 970,216 controls. The characteristics of the individuals in both studies are listed in the Online Supplement and publications^{4,7}.

Single Nucleotide Variation quality controls

A series of standard quality controls (QC) was applied to select the single nucleotide variants (SNVs) for the analysis. Variant exclusion criteria: 1) Not common to the summary statistics of the traits, 2) Minor allele frequency lower or equal to 0.01, 3) Missing values, 4) Negative standard error or not a number value, 5) p-value of 0, 6) Not SNVs, 7) Duplicated SNVs, 8) Strand ambiguity, and 8) Inconsistent allele pairs. After QCs, a total of 6,808,676 SNVs were selected (Online Figure I). Locus 15q21.3 prioritized genes *GCOM1* and *MYZAP* from AF-2018 was not evaluated due to absence of the significant SNVs of AF-2018 in the MEGASTROKE-CES analysis.

Multitrait analysis of genome-wide association studies

We applied multitrait analysis of the GWAS (MTAG)⁹, analyzing MEGASTROKE-CES and AF-2018 summary statistics. We considered associations with CES to be

significant if they attained a $p\text{-value} < 5 \times 10^{-8}$ in the MTAG results for CES (MTAG-CES) and had a $p\text{-value} < 0.05$ for association with CES in the MEGASTROKE study. Additionally, we tested whether the results from the MTAG-CES were consistent using an alternative Bayesian approach, to confirm pleiotropy of these loci using GWAS-pairwise¹³. This strategy reveals genomic regions that influence traits separately or pleiotropic regions that influence both traits. We used a posterior probability of pleiotropy (PPA pleiotropy) ≥ 0.6 to validate regions that influence both traits. Genomic inflation was estimated as lambda.

Loci associated exclusively with each trait were those that were considered significant in the MEGASTROKE-CES or AF-2018 and showed a PPA associated exclusively with one of the two traits ≥ 0.6 of the region mapped.

Identification of independent and novel loci associated with CES

Independent loci were defined as those > 1 megabase (Mb) apart in physical distance among SNVs with a genome-wide significance threshold of $p\text{-value} < 5 \times 10^{-8}$ ⁷. Loci were defined as novel when SNVs had an $r^2 < 0.1$ compared with the index SNVs of the loci *PITX2*⁴, *ZFHX3*⁴, *NKX2-5*⁴, *RGS7*⁴, *ABO*^{4,5}, *PHF20*⁶, *GNAOI*⁶ and 5q22.3⁶ that were GWAS significant in previous studies.

Replication stage in an independent European cohort

We performed genome-wide analyses in an independent cohort of 9,105 individuals (3,479 IS and 5,625 controls). IS patients were recruited if they had a measurable neurologic deficit on the NIHSS within 6 hours of the last known asymptomatic status, had been diagnosed with stroke by an experienced neurologist, which had been confirmed by neuroimaging, were over 18 years of age and were recruited via hospital-based studies between 2003 and 2020 in Spain. Controls were subjects without a history

of ischemic stroke, aged over 18 years, who declared they were free of neurovascular diseases before recruitment. The control cohort was collected in blood donation at primary care centers in Barcelona and in hospitals throughout Spain. An Institutional Review Board or Ethics Committee approved the study at each participating site. All patients or their relatives provided written informed consent. Further description of the cohorts is present in Online Supplement as well as array information, contribution of hospitals and clinical description (Online Tables I-III).

Quality control and imputation

DNA samples were genotyped on commercial arrays from Illumina (San Diego, CA) (Online Table II). Standard QCs were performed using PLINK v1.9 and KING v2.1.3 software. Imputation was performed in the Michigan Imputation Server Pipeline¹⁴ using Minimac4 And HRC r1.1 2016 panel. Further description of QCs and imputation are present in the Online Supplement.

Analysis

Analyses evaluated CES vs controls (CES = 1515, controls = 5626) and AF vs absence of AF (AF cases = 1110, AF controls = 7791). For the GWAS analyses we used an additive genetic model using fastGWA from gcta¹⁵. Age, sex and the first ten principal components were used as covariates. The results of the CES vs controls analysis were used to evaluate replicability (p-value<0.05) of the index SNVs and those significant SNVs (MTAG-CES p-value<5x10⁻⁸) in high LD ($r^2>0.8$), in the GWAS analysis of CES and to perform a Pearson correlation of the effect sizes (log(OR)) of the MTAG-CES results for the index SNVs of the loci validated by GWAS-pairwise, a strategy used in a previous study¹⁰. The results of the AF analysis were used to evaluate novel loci associated with AF in the MTAG.

Functional annotation and prioritization of genes

We used the HaploReg database to determine the functional annotation of the most strongly associated SNVs per locus. We also prioritized biological candidate genes generating a score based on: (1) nearest position to the index variant in a locus; (2) expression of the gene altered by the index variant in the eQTLGen¹⁶ or GTEx databases; (3) methylation of the gene altered by the index variant in Phenoscanner methylation data^{17,18}; (4) whether the gene harbored a protein-altering index variant itself or in high linkage disequilibrium (LD)($r^2 > 0.80$); (5) whether the gene interacted physically with the index variant in HiC data of cardiomyocytes¹⁹. For the missense SNVs, we determined the likelihood that amino acid substitution has a deleterious effect on protein function using SIFT.

Gene set analysis

We conducted a WebGestalt Overrepresentation Analysis of the selected prioritized genes associated in MTAG-CES. Gene Ontology (GO) of biological processes was performed, as well as a Benjamini Hochberg correction of the association p-value. We defined a biological process with a q-value < 0.05 as statistically significant.

Identification of potential drug repurposing targets

Prioritized genes were searched for in the DrugBank database²⁰ to find potential candidate drugs for repurposing.

RESULTS

MTAG analysis of CES

After QC (Online Figure I), there were 6,808,676 common qualified SNVs from the AF-2018 and MEGASTROKE-CES cohorts. MTAG software revealed mean χ^2 for AF-2018 and MEGASTROKE-CES in 1.39 and 1.12. The estimated equivalent GWAS sample size of the MTAG analysis for CES was of 861,823 individuals. A Manhattan plot of the MTAG-CES analysis and prioritized genes is shown in Figure 1; little evidence of genomic inflation was observed; $\lambda = 1.02$. The MTAG-CES results revealed a total of 44 associated loci ($p\text{-value} < 5 \times 10^{-8}$), 40 significant novel loci associated with CES and four previously found (Table 1). All loci previously associated in MEGASTROKE-CES (*ABO*, *NKX2-5*, *PITX2* and *ZFHX3*) are associated in this MTAG-CES, with the exception of the locus belonging to *RGS7* gene (top SNV rs146390073 MTAG $p\text{-value} = 0.001$, AF-2018 $p\text{-value} = 0.98$).

Other loci found significant in GWAS of CES different from MEGASTROKE, belonged to: *PHF20*, *GNAO1* and 5q22.3 region. For two of them, the association was more significant in our analysis. For the 5q22.3 region, top SNV rs2169955 MTAG-CES $p\text{-value} = 4.76 \times 10^{-7}$ vs MEGASTROKE-CES $p\text{-value} = 6.13 \times 10^{-3}$, and for mapped gene *PHF20*, top SNV rs11697087 MTAG-CES $p\text{-value} = 6.62 \times 10^{-5}$ vs MEGASTROKE-CES $p\text{-value} = 6.05 \times 10^{-4}$. *GNAO1* was not evaluated in our study due to absence of the index SNV in AF-2018.

New candidate loci associated with CES

After gene prioritization, 53 genes were selected from the 44 loci (Online Table IV). Novel loci showed a high degree of functionality of the SNVs as missense variants, eQTL, mQTLs, and HiC physical interaction (Online Tables V-VIII).

Replication analysis was performed in a new cohort of ischemic stroke patients and controls (n = 9,105). Evaluation of the index SNVs and SNVs in high LD revealed 56 SNVs with a p-value<0.05; these SNVs belong to 11 different loci, three already known *PITX2*, *ZFHX3*, *NKX2-5* and eight of them novel, prioritized genes: *CAVI*, *IGF1R*, *KIAA1755*, *NEURL1*, *PRRX1*, *SYNE2*, *TEX41* and *WIPF1* (Online Table IX). *NEURL1* locus was already found associated in the all types of stroke analysis of MEGASTROKE but not genome-wide significant in CES analysis ⁴.

Pearson correlation of the effect sizes (log (OR)) between index variants of the significant loci in the MTAG-CES and the CES analysis in an independent cohort was evaluated for 41 of the 44 index variants, due to the poor imputation value (INFO<0.6) of three of them in the Spanish cohort. A significant correlation was obtained; p-value=2.7x10⁻⁸, cor=0.74 (Figure 4). This association remained after removing the four SNVs from previously known loci (rs17042098, *PITX2*; rs2106261, *ZFHX3*; rs6891790, *NKX2-5* and rs635634, *ABO*) p-value=4.2x10⁻⁷, cor=0.72.

The top five most significant loci in the MTAG-CES analysis were *PITX2* and *ZFHX3* (previously known loci), and *KCNN3*, *NEURL1*, and *CAVI* (novel loci), the last two validated in the replication cohort. Interestingly, we found loci not previously described in AF or CES with four prioritized genes: *PHTF2*, *KIAA1755*, *NCOR2*, and *FILIP1*. *FILIP1* locus also reached genome-wide significance in the MTAG results for AF. Functional annotation of the index SNVs revealed rs3746471 as a missense variant of the *KIAA1755* gene coding for R1045W, and it was predicted to be deleterious with a SIFT score of 0.007.

The *FILIP1* index variant was additionally evaluated in the AF GWAS in the independent cohort, revealing a suggestive p-value with consistent direction of the

effect of this novel association with AF (rs12211255-A, beta(se)=0.013(0.007), p-value=0.09).

Biological processes of loci associated with CES and AF and biological processes of loci associated exclusively to AF

GO of biological processes from the Genome-Wide loci of the MTAG-CES analysis revealed 85 enriched gene sets (Figure 2, Online Table X); the top biological processes were: cardiac muscle contraction, actin filament-based movement and cardiac conduction. A DrugBank search for the 53 prioritized genes associated with CES revealed a total of 87 drugs (Online Table XI) that target 10 of the prioritized genes.

The study of the 111 AF-2018 significant loci (Online Table XII) using GWAS-pairwise strategy suggests 51 loci that have an exclusive association with AF risk and a lack of association with CES. A biological processes analysis of the genes associated exclusively with AF (Figure 3, Online Table XIII) revealed 25 biological processes exclusive to AF risk and mainly associated with cardiac development processes (Figure 2, Online Table XIV).

DISCUSSION

Using a multitrait analysis (MTAG) with the two biggest cohorts of CES⁴ and AF⁷ to date, we expanded the number of loci associated with CES from eight to sixteen. In the MTAG analysis, 44 loci were found associated with CES. Four of them were previously known associations: *PITX2*, *ZFHX3*, *NKX2-5* and *ABO*. To explore replicability of the MTAG results, we conducted genome-wide analysis for CES and AF in an independent European cohort comprising 9,105 subjects. We found significant associations (p-value<0.05) of 56 SNVs from 11 loci (eight of them novel associations); prioritized known genes: *PITX2*, *ZFHX3*, *NKX2-5*, prioritized novel genes: *CAVI*, *IGF1R*,

KIAA1755, *NEURL1*, *PRRX1*, *SYNE2*, *TEX41* and *WIPF1*, in the CES analysis. Moreover, given the small sample size of the replication stage compared with the estimated sample size of the MTAG, we did not expect all the SNVs to be strongly associated, but rather that the odds ratios would be highly concordant, for that we performed Pearson correlation of the effect sizes for the index SNVs in MTAG-CES analysis. The Pearson correlation revealed a positive and significant correlation in this new cohort, showing most of the SNVs in the expected direction of the effect. Following the methodology used in other similar studies¹⁰, the presence of this significant correlation suggests potential validation of our MTAG-CES results.

Interestingly, *CAVI* has been associated with platelet count and systolic blood pressure, among other traits²¹. Besides, caveolin-1, the protein encoded by *CAVI* in endothelial cells, plays a key role in how gas6 exerts its prothrombotic role in the vasculature²². Regarding the other genes, and among other human traits in GWAS, *IGF1R* and *WIPF1* have been found associated with QRS duration^{23,24} and platelet crit^{25,26}. *NEURL1* is associated with white matter hyperintensity volume²⁷, red blood cell count²⁸ and to all types of stroke risk⁴. *PRRX1*, *SYNE2* and *TEX41* have been associated with PR interval²⁹⁻³¹ and *SYNE2* also associated with platelet count³⁰. It is worth noting, that of these replicated loci, locus 20q11.23, prioritized gene: *KIAA1755*, has not previously been found associated with AF. The index variant of this locus, rs3746471-A, codes for R1045W amino acid change, predicted to be deleterious according to SIFT. rs3746471-A, has been previously described associated with heart rate³²⁻³⁴ and PR interval³⁴, and remarkably suggestively associated with stroke infarct volume (p-value=6.80x10⁻⁷)^{35,36}. *KIAA1755* is predicted to encode an uncharacterized protein and is only characterized at the transcriptional level. The transcript is highly expressed in the brain and nerves and is also expressed in the heart.

These parallel associations of the replicated loci highlight the role that these genes could be playing in the risk conferring to CES, via altering AF, platelet count, coagulation cascade and different hematological traits.

Using MTAG strategy, we were also able to analyze the output of AF-2018 enhanced by MEGASTROKE-CES. In this direction, a novel locus, 6q14.1, was found significantly associated with AF. *FILIP1* is the gene prioritized; it encodes for a filamin A binding protein. Interestingly, *FILIP1* has been identified previously as a regulator of myogenesis differentiation, in human cells and in an *in vivo* mouse model³⁷. In the replication stage, this SNV was found suggestive (p-value=0.09), highly probable due to the small sample size in comparison to MTAG analysis.

The study of previous known loci associated with CES in the MTAG-CES analysis showed that the *RGS7*⁴ previously genome-wide significant in the MEGASTROKE was not associated with CES in our study. This could be explained by a causal mechanism other than AF in the risk of suffering CES or the possibility of a false positive result, as authors suspected in the MEGASTROKE study given the low allele frequency and poor imputation⁴. The evaluation of loci found in GWAS of CES different from MEGASTROKE (*PHF20*, *GNAO1* and 5q22.3 region), revealed 5q22.3 and *PHF20* loci with stronger significance in our study. This could be considered a reinforcement of these two signals and their relation to CES risk via AF. Sadly, *GNAO1* was not evaluated due to the lack of the index variant in AF-2018.

Globally, from the 44 loci genome-wide associated in the MTAG-CES, 53 genes were prioritized in our study: the top previously known genes were *PITX2* and *ZFHX3* and other top novel genes were *NEURL1* and *CAVI*. The prioritized genes were involved in biological processes, such as cardiac muscle contraction, actin filament-based

movement, cardiac conduction, and striated muscle contraction, which highlights their role in the risk of CES due to AF.

Remarkably, among the 111 loci previously associated with AF, a group of 51 loci had an exclusive genetic association with AF and a lack of association with CES using GWAS-pairwise. Apparently, no statistical differences were observed between the AF loci associated with CES and those with no association. GO of biological processes revealed 43 gene sets that were only enriched by the AF exclusive genes and not present in the biological processes of the 44 loci associated with CES; these biological processes were related mainly to cardiac development. These genes exclusively associated with AF could be important for understanding the risk of CES occurrence, and thereby help to develop more specific prevention drugs or to create better genetic scores for identifying AF patients with high risk of suffering a stroke.

Our study has limitations, as there is a known overlap of individuals between the MEGASTROKE-CES and AF-2018 studies. However, MTAG permits sample overlap between the GWAS summary data of the traits by conducting bivariate LD score regression of the traits⁹. Additionally, the difference in the sample size between the two original studies is a concern, on the other hand MTAG estimation of χ^2 revealed an scenario expected to be strong against false positives, as tested in the original publication⁹. To control this, we performed a replication analysis in an independent cohort to validate the results obtained. Another limitation is that the MTAG approach relies on the relatively strong assumption that associated variants act on both traits, which may not always be the case for CES and AF. To control this, we only considered SNVs reaching the GWAS significant p-value, showing an association with both traits and having a PPA pleiotropy of the genomic region $\Rightarrow >0.6$ in the GWAS-pairwise. In addition, we performed a replication step in an independent cohort to further evaluate

these loci. Although this replication is objectively weak compared to the MTAG analysis, we hypothesize that more loci found in our study will be validated when an equivalent cohort in the future is analyzed. It is also important to mention that through this strategy we have only found loci associated to CES risk due to AF, and therefore further multitrait analysis should be performed with different traits to uncover the different high-risk sources of CES.

To summarize, in our study, we identified 40 novel loci associated with CES through a multitrait analysis enhanced by AF, and we were able to replicate eight of them in an independent cohort. This study provides novel insights into the pathogenesis of CES, its relation to AF and highlights multiple candidates for use in future functional experiments to identify specific mechanisms conferring a risk of CES that could be targeted therapeutically.

ACKNOWLEDGMENTS

The Genotype-Tissue Expression (GTEx) Project was funded by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from the GTEx Portal on 03/30/20. This study uses data generated by the GCAT | Genomes for Life. Cohort study of the Genomes of Catalonia, IGTP. A full list of the investigators who contributed to the generation of the data is available from <http://www.genomesforlife.com/>. IGTP is part of the CERCA Program/Generalitat de Catalunya.

FUNDING

J. Cárcel-Márquez has received funding through an AGAUR Contract (Agència de Gestió d'Ajuts Universitaris i de Recerca; FI_DGR 2019, grant number 2020FI_B1 00157) co-financed with Fons Social Europeu (FSE). From Instituto de Salud Carlos III: E. Muiño is funded by a Río Hortega Contract (CM18/00198), M. Lledós is funded by a PFIS Contract (Contratos Predoctorales de Formación en Investigación en Salud, FI19/00309), C. Gallego-Fabrega is supported by a Sara Borrell Contract (CD20/00043) and Fondo Europeo de Desarrollo Regional (ISCIII-FEDER), T. Sobrino (CPH17/00027) and F. Campos (CPH19/00020) are recipients of research contracts from the Miguel Servet Program. This study has been funded by Carlos III Institute PI15/01978, PI17/02089, PI18-01338, and RETICS INVICTUS PLUS RD16/0019), by Marató TV3 support of the Epigenesis study by the Fundació Docència i Recerca FMT grant for the Epigenesis project, by Eranet-Neuron of the Ibiostroke project (AC19/00106) and by Boehringer Ingelheim of the SEDMAN Study. GCAT Cession Research Project PI-2018-01.

DECLARATION OF INTERESTS

Nothing to report.

REFERENCES

1. Kamel H, Healey JS. Cardioembolic Stroke. *Circ. Res.* 2017;120:514–526.
2. Evans A, Perez I, Yu G, Kalra L. Should stroke subtype influence anticoagulation decisions to prevent recurrence in stroke patients with atrial fibrillation? *Stroke.* 2001;32:2828–2832.
3. Gladstone DJ, Spring M, Dorian P, Panzov V, Thorpe KE, Hall J, et al. Atrial fibrillation in patients with cryptogenic stroke. *N. Engl. J. Med.* 2014;370:2467–2477.
4. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat. Genet.* 2018;50:524–537.
5. Williams FMK, Carter AM, Hysi PG, Surdulescu G, Hodgkiss D, Soranzo N, et al. Ischemic stroke is associated with the ABO locus: the EuroCLOT study. *Ann. Neurol.* 2013;73:16–31.
6. von Berg J, van der Laan SW, McArdle PF, Malik R, Kittner SJ, Mitchell BD, et al. Alternate approach to stroke phenotyping identifies a genetic risk locus for small vessel stroke. *Eur. J. Hum. Genet.* [Internet]. 2020;28:963–972. Available from: <http://dx.doi.org/10.1038/s41431-020-0580-5>
7. Nielsen JB, Thorolfsson RB, Fritsche LG, Zhou W, Skov MW, Graham SE, et al. Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. *Nat. Genet.* [Internet]. 2018 [cited 2020 Mar 9];50:1234–1239. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/30061737>
8. Pulit SL, Weng L-C, McArdle PF, Trinquart L, Choi SH, Mitchell BD, et al. Atrial fibrillation genetic risk differentiates cardioembolic stroke from other stroke subtypes. *Neurol. Genet.* [Internet]. 2018 [cited 2020 Apr 2];4:e293. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/30584597>
9. Turley P, Walters RK, Maghzian O, Okbay A, Lee JJ, Fontana MA, et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat. Genet.* [Internet]. 2018 [cited 2020 Mar 23];50:229–237. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29292387>
10. Craig JE, Han X, Qassim A, Hassall M, Cooke Bailey JN, Kinzy TG, et al. Multitrait analysis of glaucoma identifies new risk loci and enables polygenic prediction of disease susceptibility and progression. *Nat. Genet.* 2020;52:160–166.
11. Pei YF, Liu L, Liu T Le, Yang XL, Zhang H, Wei XT, et al. Joint Association

- Analysis Identified 18 New Loci for Bone Mineral Density. *J. Bone Miner. Res.* [Internet]. 2019 [cited 2020 Apr 13];34:1086–1094. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/30690781>
12. Chung J, Marini S, Pera J, Norrving B, Jimenez-Conde J, Roquer J, et al. Genome-wide association study of cerebral small vessel disease reveals established and novel loci. *Brain*. 2019;142:3176–3189.
13. Pickrell JK, Berisa T, Liu JZ, Séguérel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. *Nat. Genet.* 2016;48:709–717.
14. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. *Nat. Genet.* 2016;48:1284–1287.
15. Jiang L, Zheng Z, Qi T, Kemper KE, Wray NR, Visscher PM, et al. A resource-efficient tool for mixed model association analysis of large-scale data. *Nat. Genet.* [Internet]. 2019 [cited 2021 Jun 1];51:1749–1755. Available from: <https://doi.org/10.1038/s41588-019-0530-8>
16. Võsa U, Claringbould A, Westra H-J, Bonder MJ, Deelen P, Zeng B, et al. Unraveling the polygenic architecture of complex traits using blood eQTL metaanalysis. *bioRxiv* [Internet]. 2018 [cited 2020 Mar 30];447367. Available from: <https://www.biorxiv.org/content/10.1101/447367v1>
17. Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, et al. PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics* [Internet]. 2016 [cited 2020 Mar 30];32:3207–3209. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27318201>
18. Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics* [Internet]. 2019 [cited 2020 Mar 30];35:4851–4853. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/31233103>
19. Choy MK, Javierre BM, Williams SG, Baross SL, Liu Y, Wingett SW, et al. Promoter interactome of human embryonic stem cell-derived cardiomyocytes connects GWAS regions to cardiac gene networks. *Nat. Commun.* 2018;9:1–10.
20. Wishart D, Y. F, A. G, Lo E, Marcu A, Grant J, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* 2017 [Internet]. DrugBank. 2017 [cited 2020 Jun 30]; Available from: <https://academic.oup.com/nar/article/46/D1/D1074/4602867>
21. Common Metabolic Diseases Knowledge Portal. CAV1. 2021 Jun 17. <https://hugeamp.org/gene.html?gene=CAV1>.
22. Laurance S, Aghourian MN, Jiva Lila Z, Lemarié CA, Blostein MD. Gas6-induced tissue factor expression in endothelial cells is mediated through caveolin-1-enriched microdomains. *J. Thromb. Haemost.* [Internet]. 2014 [cited 2021 Jun 17];12:395–408. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1111/jth.12481>
23. Prins BP, Mead TJ, Brody JA, Sveinbjornsson G, Ntalla I, Bihlmeyer NA, et al.

- Exome-chip meta-analysis identifies novel loci associated with cardiac conduction, including ADAMTS6. *Genome Biol.* [Internet]. 2018 [cited 2021 Jun 17];19. Available from: <https://pubmed.ncbi.nlm.nih.gov/30012220/>
24. van der Harst P, van Setten J, Verweij N, Vogler G, Franke L, Maurano MT, et al. 52 Genetic Loci Influencing Myocardial Mass. *J. Am. Coll. Cardiol.* [Internet]. 2016 [cited 2021 Jun 17];68:1435–1448. Available from: <https://pubmed.ncbi.nlm.nih.gov/27659466/>
25. Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, et al. The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell* [Internet]. 2016 [cited 2021 Jun 17];167:1415–1429.e19. Available from: <https://pubmed.ncbi.nlm.nih.gov/27863252/>
26. Common Metabolic Diseases Knowledge Portal. WIPF1. 2021 Jun 17. <https://hugeamp.org/gene.html?gene=WIPF1>.
27. Verhaaren BFJ, DeBette S, Bis JC, Smith JA, Ikram MK, Adams HH, et al. Multiethnic Genome-Wide Association Study of Cerebral White Matter Hyperintensities on MRI. *Circ. Cardiovasc. Genet.* [Internet]. 2015 [cited 2021 Jun 17];8:398–409. Available from: <https://pubmed.ncbi.nlm.nih.gov/264427240/>
28. Common Metabolic Diseases Knowledge Portal. NEURL1. 2021 Jun 17. <https://hugeamp.org/gene.html?gene=NEURL1>.
29. Common Metabolic Diseases Knowledge Portal. PRRX1. 2021 Jun 17. <https://hugeamp.org/gene.html?gene=PRRX1>.
30. Common Metabolic Diseases Knowledge Portal. SYNE2. 2021 Jun 17. <https://hugeamp.org/gene.html?gene=SYNE2>.
31. Ntalla I, Weng LC, Cartwright JH, Hall AW, Sveinbjornsson G, Tucker NR, et al. Multi-ancestry GWAS of the electrocardiographic PR interval identifies 202 loci underlying cardiac conduction. *Nat. Commun.* [Internet]. 2020 [cited 2021 Jun 17];11. Available from: <https://pubmed.ncbi.nlm.nih.gov/32439900/>
32. Nolte IM, Munoz ML, Tragante V, Amare AT, Jansen R, Vaez A, et al. Genetic loci associated with heart rate variability and their effects on cardiac disease risk. *Nat. Commun.* 2017;8.
33. van den Berg ME, Warren HR, Cabrera CP, Verweij N, Mifsud B, Haessler J, et al. Discovery of novel heart rate-associated loci using the Exome Chip. *Hum. Mol. Genet.* [Internet]. 2017 [cited 2020 Apr 13];26:2346–2363. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28379579>
34. Common Metabolic Diseases Knowledge Portal. rs3746471. 2021 Jun 17. <https://hugeamp.org/variant.html?variant=rs3746471>.
35. Common Metabolic Diseases Knowledge Portal. KIAA1755. 2021 Jun 17. <https://hugeamp.org/gene.html?gene=KIAA1755>.
36. Pirruccello JP, Bick A, Wang M, Chaffin M, Friedman S, Yao J, et al. Analysis of cardiac magnetic resonance imaging in 36,000 individuals yields genetic insights into dilated cardiomyopathy. *Nat. Commun.* [Internet]. 2020 [cited 2021 Jun 18];11. Available from: <https://pubmed.ncbi.nlm.nih.gov/32382064/>

37. Militello G, Hosen MR, Ponomareva Y, Gellert P, Weirick T, John D, et al. A novel long non-coding RNA Myolinc regulates myogenesis through TDP-43 and Filip1. *J. Mol. Cell Biol.* [Internet]. 2018 [cited 2020 Apr 13];10:102–117. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29618024>

Figures and Tables.

Figure 1. Manhattan plot of MTAG-CES. The X axis represents chromosome location, and the Y axis represents the minus logarithm on base 10 of p-value. The red line represents the GWAS-significance threshold. The novel loci are shown in blue and the established loci in yellow.

Figure 2. GO Biological processes enriched in CES prioritized gene set.

Figure 3. GO Biological processes enriched exclusively in analysis of AF associated genes independently of CES risk.

Figure 4. Comparison of the effect sizes (log ORs) for 41 genome-wide significant index SNVs identified from the MTAG-CES analysis versus those in CES analysis of the Spanish independent cohort.

Table 1. MTAG-CES results of the independent and significant loci

				MEGASTROKE CES		AF-2018		MTAG-CES		MTAG-AF		PPA pleio
Index SNV and assessed allele	Locus	Gene	Novelty of Locus	Z	P	Z	P	Z	P	Z	P	
rs17042098-A	4q25	PITX2	Known ⁴	12.72	3.72x10 ⁻³⁷	38.01	8.97x10 ⁻³¹⁸	32.86	7.73x10 ⁻²³⁷	37.93	0	1.00
rs2106261-T	16q22.3	ZFHX3	Known ⁴	6.75	1.63x10 ⁻¹¹	20.23	4.97x10 ⁻⁹¹	17.48	2.02x10 ⁻⁶⁸	20.19	1.20x10 ⁻⁹⁰	1.00
rs11264280-T	1q21.3	KCNN3	Novel	2.49	1.28x10 ⁻⁰²	18.97	3.07x10 ⁻⁷⁹	14.22	7.17x10 ⁻⁴⁶	18.44	5.88x10 ⁻⁷⁶	0.65
rs11598047-A	10q24.33	NEURL1	Novel	-3.27	1.06x10 ⁻⁰³	-17.08	8.95x10 ⁻⁶⁶	-13.38	7.50x10 ⁻⁴¹	-16.73	7.42x10 ⁻⁶³	0.93
rs3807989-A	7q31.2	CAVI	Novel	-4.50	6.75x10 ⁻⁰⁶	-15.63	1.24x10 ⁻⁵⁴	-13.10	3.43x10 ⁻³⁹	-15.50	3.30x10 ⁻⁵⁴	1.00
rs680084-A	1q24.2	PRRX1	Novel	-4.10	4.02x10 ⁻⁰⁵	-13.54	3.31x10 ⁻⁴²	-11.46	2.10x10 ⁻³⁰	-13.46	2.84x10 ⁻⁴¹	0.99
rs883079-T	12q24.21	TBX5	Novel	3.55	3.95x10 ⁻⁰⁴	13.26	2.84x10 ⁻⁴⁰	10.96	5.97x10 ⁻²⁸	13.12	2.60x10 ⁻³⁹	0.97
		<i>FAM13B,</i> <i>KLHL3,</i> <i>KDM3B,</i> <i>GFRA3,</i>										
rs17171711-T	5q31.2	KIF20A	Novel	3.79	1.56x10 ⁻⁰⁴	12.48	1.95x10 ⁻³⁵	10.57	4.04x10 ⁻²⁶	12.41	2.35x10 ⁻³⁵	0.99
rs4385527-A	9q22.32	AOPEP	Novel	3.92	8.56x10 ⁻⁰⁵	12.03	6.16x10 ⁻³³	10.34	4.69x10 ⁻²⁵	11.99	3.89x10 ⁻³³	0.95
rs78249997-T	10q22.2	SYNPO2L	Novel	-3.82	1.37x10 ⁻⁰⁴	-12.08	8.75x10 ⁻³⁴	-10.32	5.80x10 ⁻²⁵	-12.03	2.45x10 ⁻³³	0.98
rs7172038-T	15q24.1	HCN4	Novel	-2.73	6.43x10 ⁻⁰³	-12.58	4.78x10 ⁻³⁶	-10.04	1.02x10 ⁻²³	-12.37	3.76x10 ⁻³⁵	0.74
rs2738413-A	14q23.2	SYNE2	Novel	2.97	3.03x10 ⁻⁰³	11.61	2.55x10 ⁻³¹	9.52	1.67x10 ⁻²¹	11.47	1.80x10 ⁻³⁰	0.85
rs6891790-T	5q35.1	NKX2-5	Known ⁴	-4.97	6.67x10 ⁻⁰⁷	-9.59	4.53x10 ⁻²²	-9.29	1.57x10 ⁻²⁰	-9.80	1.16x10 ⁻²²	1.00
rs2857265-A	2q31.2	TTN	Novel	2.76	5.85x10 ⁻⁰³	10.16	4.58x10 ⁻²⁴	8.42	3.65x10 ⁻¹⁷	10.06	8.29x10 ⁻²⁴	0.70
rs10753933-T	1q32.1	PPFIA4	Novel	3.72	2.09x10 ⁻⁰⁴	9.09	9.84x10 ⁻²⁰	8.24	1.70x10 ⁻¹⁶	9.16	5.30x10 ⁻²⁰	0.99
rs74399915-T	11q24.3	KCNJ5	Novel	3.38	7.20x10 ⁻⁰⁴	9.05	1.23x10 ⁻¹⁹	8.03	1.01x10 ⁻¹⁵	9.08	1.11x10 ⁻¹⁹	0.92
rs13191450-A	6q22.31	HSF2	Novel	2.93	3.51x10 ⁻⁰³	8.88	9.97x10 ⁻¹⁹	7.65	1.95x10 ⁻¹⁴	8.86	7.96x10 ⁻¹⁹	0.81
rs2834618-T	21q22.12	RUNX1	Novel	3.31	9.48x10 ⁻⁰⁴	8.43	3.41x10 ⁻¹⁷	7.56	3.96x10 ⁻¹⁴	8.47	2.37x10 ⁻¹⁷	0.96

rs56181519-T	2q31.1	WIPF1	Novel	-2.73	6.31x10 ⁻⁰³	-8.60	6.46x10 ⁻¹⁸	-7.35	1.98x10 ⁻¹³	-8.56	1.11x10 ⁻¹⁷	0.82
rs12908004-A	15q25.1	ARNT2	Novel	-3.28	1.03x10 ⁻⁰³	-8.13	4.12x10 ⁻¹⁶	-7.35	2.03x10 ⁻¹³	-8.19	2.65x10 ⁻¹⁶	0.96
rs3176326-A	6p21.2	CDKN1A	Novel	-3.98	6.66x10 ⁻⁰⁵	-7.36	1.42x10 ⁻¹³	-7.23	5.01x10 ⁻¹³	-7.54	4.59x10 ⁻¹⁴	1.00
		<i>GMCL1,</i>										
		<i>ANXA4,</i>										
		<i>SNRNP27,</i>										
rs6747542-T	2p13.3	ASPRV1	Novel	2.61	8.86x10 ⁻⁰³	8.27	1.10x10 ⁻¹⁶	7.06	1.63x10 ⁻¹²	8.23	1.82x10 ⁻¹⁶	0.75
rs337705-T	5q22.3	KCNN2	Novel	-2.56	1.04x10 ⁻⁰²	-8.29	1.63x10 ⁻¹⁶	-7.05	1.77x10 ⁻¹²	-8.25	1.57x10 ⁻¹⁶	0.72
rs41292535-A	1p32.2	EPSI5	Novel	3.81	1.39x10 ⁻⁰⁴	7.16	7.42x10 ⁻¹³	6.99	2.73x10 ⁻¹²	7.33	2.34x10 ⁻¹³	0.96
rs140185678-A	16p13.3	RPL3L	Novel	2.92	3.54x10 ⁻⁰³	7.61	2.43x10 ⁻¹⁴	6.79	1.13x10 ⁻¹¹	7.64	2.13x10 ⁻¹⁴	0.89
rs76774446-A	17q21.32	GOSR2	Novel	4.38	1.20x10 ⁻⁰⁵	6.15	8.72x10 ⁻¹⁰	6.63	3.32x10 ⁻¹¹	6.43	1.25x10 ⁻¹⁰	0.99
rs55754224-T	4q26	CAMK2D	Novel	2.80	5.05x10 ⁻⁰³	7.39	2.15x10 ⁻¹³	6.57	4.92x10 ⁻¹¹	7.41	1.22x10 ⁻¹³	0.80
rs79187193-A	1q21.2	GJA5	Novel	-2.37	1.78x10 ⁻⁰²	-7.59	3.15x10 ⁻¹⁴	-6.47	9.79x10 ⁻¹¹	-7.56	4.08x10 ⁻¹⁴	0.71
rs12260801-T	10q25.2	RBM20	Novel	4.56	5.22x10 ⁻⁰⁶	5.53	2.94x10 ⁻⁰⁸	6.31	2.79x10 ⁻¹⁰	5.86	4.62x10 ⁻⁰⁹	1.00
rs2269001-A	7q36.1	KCNH2	Novel	-2.89	3.94x10 ⁻⁰³	-6.63	4.01x10 ⁻¹¹	-6.11	1.00x10 ⁻⁰⁹	-6.70	2.06x10 ⁻¹¹	0.73
rs6598541-A	15q26.3	IGF1R	Novel	2.69	7.17x10 ⁻⁰³	6.35	2.22x10 ⁻¹⁰	5.81	6.21x10 ⁻⁰⁹	6.41	1.48x10 ⁻¹⁰	0.72
rs11125871-T	2p15	USP34	Novel	-3.24	1.17x10 ⁻⁰³	-5.79	6.42x10 ⁻⁰⁹	-5.75	9.19x10 ⁻⁰⁹	-5.95	2.71x10 ⁻⁰⁹	0.87
rs635634-T	9q34.2	ABO	Known ^{4,5}	5.10	3.31x10 ⁻⁰⁷	4.20	2.74x10 ⁻⁰⁵	5.72	1.04x10 ⁻⁰⁸	4.66	3.13x10 ⁻⁰⁶	1.00
rs10272350-A	7q11.23	PHTF2	Novel	4.53	6.13x10 ⁻⁰⁶	4.62	3.41x10 ⁻⁰⁶	5.68	1.32x10⁻⁰⁸	4.99	5.93x10 ⁻⁰⁷	0.99
rs116600817-A	17p13.1	TNFSF12	Novel	2.88	3.92x10 ⁻⁰³	6.00	2.45x10 ⁻⁰⁹	5.68	1.33x10 ⁻⁰⁸	6.10	1.07x10 ⁻⁰⁹	0.85
rs13010313-T	2q22.3	TEX41	Novel	3.52	4.21x10 ⁻⁰⁴	5.45	5.72x10 ⁻⁰⁸	5.67	1.41x10 ⁻⁰⁸	5.65	1.57x10 ⁻⁰⁸	0.90
rs2885697-T	1p34.2	SCMH1	Novel	-2.54	1.09x10 ⁻⁰²	-6.27	2.88x10 ⁻¹⁰	-5.67	1.41x10 ⁻⁰⁸	-6.32	2.70x10 ⁻¹⁰	0.72
rs11057583-A	12q24.31	NCOR2	Novel	3.82	1.35x10 ⁻⁰⁴	5.19	2.10x10 ⁻⁰⁷	5.67	1.46x10⁻⁰⁸	5.45	5.13x10 ⁻⁰⁸	0.74
rs11782313-T	8q24.3	PTK2	Novel	-3.05	2.34x10 ⁻⁰³	-5.81	5.71x10 ⁻⁰⁹	-5.64	1.65x10 ⁻⁰⁸	-5.94	2.94x10 ⁻⁰⁹	0.78
rs12211255-A	6q14.1	FILIP1	Novel	3.42	6.24x10 ⁻⁰⁴	5.44	5.06x10 ⁻⁰⁸	5.61	2.06x10⁻⁰⁸	5.63	1.78x10 ⁻⁰⁸	0.96
rs1898096-A	10q22.3	LRMDA	Novel	-2.64	8.30x10 ⁻⁰³	-6.08	1.39x10 ⁻⁰⁹	-5.60	2.13x10 ⁻⁰⁸	-6.15	7.96x10 ⁻¹⁰	1.00
		<i>FGF5,</i>										
rs11099098-T	4q21.21	PRDM8	Novel	2.77	5.67x10 ⁻⁰³	5.94	2.96x10 ⁻⁰⁹	5.58	2.38x10 ⁻⁰⁸	6.03	1.62x10 ⁻⁰⁹	0.78
rs55985730-T	7q32.1	CALU,	Novel	-2.84	4.47x10 ⁻⁰³	-5.82	5.24x10 ⁻⁰⁹	-5.54	3.06x10 ⁻⁰⁸	-5.92	3.19x10 ⁻⁰⁹	0.87

<i>FLNC</i>												
rs3746471-A	20q11.23	KIAA1755	Novel	3.56	3.58x10 ⁻⁰⁴	5.18	2.30x10 ⁻⁰⁷	5.51	3.50x10⁻⁰⁸	5.40	6.59x10 ⁻⁰⁸	0.96

SNV: Single Nucleotide Variant, Z: Z-score; P: P-value. Loci highlighted in bold are the ones not previously associated with AF.