ESC European Society of Cardiology

Europace (2025) **27**, euaf042
https://doi.org/10.1093/europace/euaf042

CLINICAL RESEARCH

Identification of a new genetic locus associated with atrial fibrillation in the Taiwanese population by genome-wide and transcriptome-wide association studies

Guan-Wei Lee [1†], Jien-Jiun Chen [2†], Chih-Hsien Wang [3,4], Sheng-Nan Chang [2], Fu-Chun Chiu [2], Pang-Shuo Huang [2], Su-Kiat Chua [5], Eric Y. Chuang [1,6,7*], and Chia-Ti Tsai [4,8,9,10*]

[1]Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 106, Taiwan; [2]Division of Cardiology, Department of Internal Medicine, National Taiwan University College of Medicine and Hospital Yun-Ling Branch, Dou-Liu City, Taiwan; [3]Division of Cardiovascular Surgery, Department of Surgery, National Taiwan University College of Medicine and Hospital, Taipei, Taiwan; [4]Cardiovascular Center, National Taiwan University Hospital, Taipei, Taiwan; [5]Division of Cardiology, Department of Internal Medicine, Shin Kong Wu Ho-Su Memorial Hospital, Taipei, Taiwan; [6]Bioinformatics and Biostatistics Core, Centers of Genomic and Precision Medicine, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 106, Taiwan; [7]Research and Development Center for Medical Devices, National Taiwan University, No. 7, Chung-Shan S Rd, Taipei 100, Taiwan; [8]Division of Cardiology, Department of Internal Medicine, National Taiwan University College of Medicine and Hospital, Taipei, Taiwan; [9]Graduate Institute of Clinical Medicine, National Taiwan University College of Medicine, Taipei, Taiwan; and [10]Department of Geriatrics and Gerontology, National Taiwan University College of Medicine and Hospital, Taipei, Taiwan

Received 25 September 2024; accepted after revision 25 February 2025; online publish-ahead-of-print 28 February 2025

Aims

Genome-wide association studies (GWASs) identified common single-nucleotide polymorphisms (SNPs) in more than 100 genomic regions associated with atrial fibrillation (AF). We aimed to identify novel AF genes in Taiwanese population by multi-stage GWAS.

Methods and results

In exploratory stage, we did GWAS with whole-genome genotypes (4 512 191 SNPs) in 516 patients with AF from the National Taiwan University AF Registry and 5160 normal sinus rhythm controls from the Taiwan Biobank. Significant loci were replicated in 1002 independent patients and 2003 controls and in the UK Biobank. Expression quantitative trait locus (eQTL) mapping and transcriptome-wide association study (TWAS) were performed to implicate functional significance. Stage I GWAS revealed three loci associated with AF with a genome-wide significance level, including one close to *PITX2* gene (chromosome 4q25, rs2723329, minor allele frequency [MAF] 0.50 vs. 0.41, $P = 1.53 \times 10^{-10}$), another close to *RAP1A* gene (also to previous *KCND3*; chromosome 1p13.2, rs7525578, MAF 0.17 vs. 0.07, $P = 1.24 \times 10^{-26}$), and one novel locus close to *HNF4G* gene (chromosome 8q21.13, rs2980218, MAF 0.44 vs. 0.35, $P = 2.19 \times 10^{-9}$). They were validated in Stage II population. The eQTL analyses showed significant colocalization of 1p13.2 locus with *RAP1A* gene expression in fibroblasts and 8q21.13 locus with *HNF4G* expression in lymphocytes. There is a significant association of *RAP1A* gene expression in fibroblasts and *HNF4G* in lymphocytes and brain with AF in TWAS.

Conclusion

Genome-wide association study in Taiwan revealed *PITX2* and *RAP1A/KCND3* loci and novel AF locus (*HNF4G*) with the most significant locus in the *RAP1A* locus. *RAP1A* and *HNF4G* genes may implicate fibrosis, metabolic, and neurogenic pathways in pathogenesis of AF.

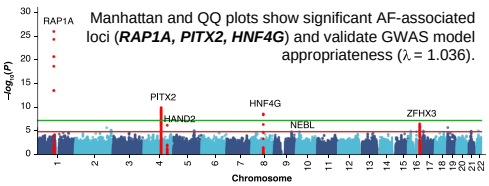
* Corresponding authors. Tel: +886-2-33663660. E-mail address: chuangey@ntu.edu.tw (E.Y.C); Tel: +886-2-23123456 ext 265002; fax: +886-2-82317099. *E-mail address*: cttsai1999@gmail.com; cttsai@ntuh.gov.tw (C.-T.T)

† The first two authors contributed equally to the study.

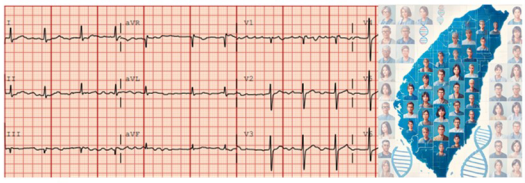
Graphical Abstract

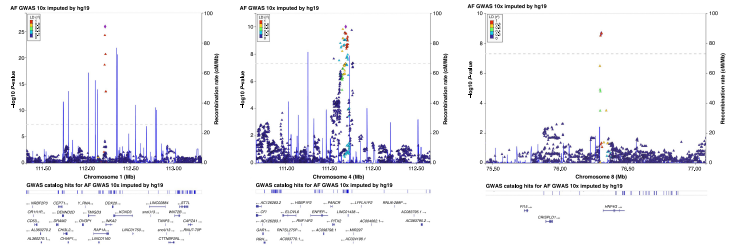
Identification of a new genetic locus associated with atrial fibrillation in the Taiwanese population by genome-wide and transcriptome-wide association studies

Methods: Multi-stage GWAS and TWAS analyzed 4 512 191 SNPs in 516 patients with AF and 5160 controls, validated in 1002 independent patients with AF and 2003 controls.



Manhattan and QQ plots show significant AF-associated loci (***RAP1A, PITX2, HNF4G***) and validate GWAS model appropriateness ($\lambda = 1.036$).





Regional plots highlight SNP associations with AF near ***RAP1A, PITX2,*** and ***HNF4G,*** displaying chromosomal positions and –log10(P) values.

Results: Identified three loci associated with AF: ***PITX2, RAP1A,*** and novel **HNF4G.** Functional analysis TWAS and eQTL linked ***RAP1A*** to fibrosis and ***HNF4G*** to inflammatory pathways.

Conclusion: GWAS and TWAS revealed genetic markers influencing AF, possibly via fibrosis, metabolism, and inflammation, providing insights into AF mechanisms and risk prediction.

Keywords Atrial fibrillation • Genome-wide association study • Transcriptome-wide association study • Expression quantitative trait locus • Whole genome • Taiwan Biobank • Taiwanese • Taiwan

What's new

- We identified three loci associated with AF, including the previously reported *PITX2* gene locus and *KCND3* gene locus (nearest gene is *RAP1A* in our study) and one novel *HNF4G* gene locus. The *RAP1A* and *HNF4G* genes are associated with fibrosis and inflammation, respectively—both of which are critical AF mechanisms.
- In previous AF GWAS, *PITX2* locus was usually the most significant locus associated with AF. In our Taiwanese population, we found the *RAP1A* locus to be even more significant than the *PITX2* locus.

Introduction

Atrial fibrillation (AF) is the most prevalent sustained arrhythmia and a significant risk factor for stroke, heart failure, and cardiovascular mortality. Over the past decade, genome-wide association studies (GWASs) have identified common single-nucleotide polymorphisms (SNPs) in over 100 genomic regions associated with AF, including loci on chromosomes 4q25 (PITX2), 16q22 (ZFHX3), and 1q21 (KCNN3).[1–5] However, these loci do not fully account for the genetic risk of AF, suggesting the presence of additional genetic factors yet to be discovered. Moreover, several AF-related genes identified by GWAS in Caucasian populations have shown ethnic variations in Asian populations.[6–8]

While GWASs have been conducted in Japanese[6] and Korean[8] populations, Asian populations are highly diverse, comprising distinct ancestry groups such as East Asian, Southeast Asian, and South Asian, each with unique genetic structures and linkage disequilibrium (LD) patterns. Conducting GWAS across multiple Asian countries can help identify population-specific variants that might not be detected in a single country or region. For example, a variant strongly associated with a trait in Han Chinese individuals may not exhibit the same relevance in South Asians due to differences in allele frequencies and LD structures.[9] Our previous GWAS study on copy number variation (CNV) also identified a unique CNV in the *KCNIP1* gene associated with AF in the Taiwanese population, a finding not observed in other populations.[10] Consequently, the present study aims to identify additional novel AF loci or genes specific to patients with AF in the Taiwanese population.

Methods

Study populations

A multi-stage study design was employed to minimize false-positive findings while maximizing power and efficiency. The study design is illustrated in *Figure 1*. All the patients with AF in this study were recruited from the cardiovascular clinics within the National Taiwan University AF Registry (NTUAFR). Detailed patient selection criteria have been previously described.[10–13] Patients without documented AF in 12-lead ECG or Holter ECG, and with no reported history or diagnosis of AF, were selected as normal sinus rhythm (NSR) controls. The NSR controls were sourced from the general Taiwanese population [Taiwan Biobank (TWB)],[14,15] with whole-genome sequencing data for the Stage I exploratory population and from the NTUAFR for the Stage II validation population. The Stage I and II populations did not overlap. The TWB includes over 100 000 Taiwanese participants recruited from the community to investigate the effects of environmental and genetic factors on disease risks and to provide health information for the Taiwanese population.[14,15]

In the exploratory stage, a whole-genome GWAS was conducted on 516 patients with AF and 5160 NSR controls. In the replication stage, significant loci identified in the exploratory stage were validated in 1002 general patients with AF and 2003 NSR controls from our cardiovascular clinics (NTUAFR), as reported in our previous studies.[10–13] All patients were Taiwanese, with no population stratification ($\lambda = 1.036$). The study was approved by the Institutional Review Board (IRB) of the National Taiwan University Hospital (200911002R), and written informed consent was obtained from all participants in the National Taiwan University Hospital.

Validation of Taiwan AF loci in the Caucasian population and vice versa

In the third stage, to further validate the loci or genes identified in the Taiwanese population, we confirmed the association in the United

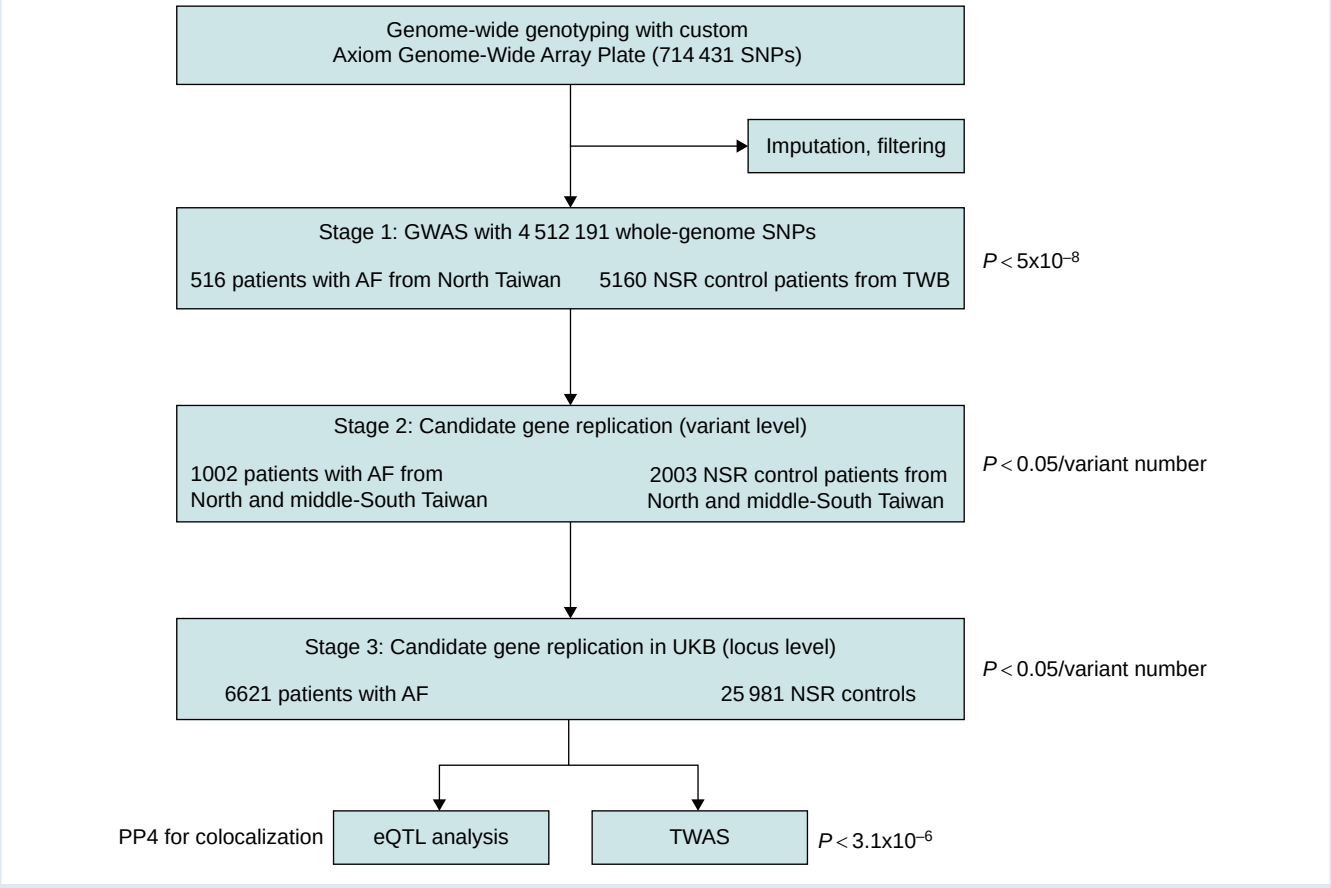


Figure 1 CONSORT flow diagram showing the study design of three-staged GWAS. AF, atrial fibrillation; GWAS, genome-wide association study; NSR, normal sinus rhythm; SNP, single-nucleotide polymorphism.

Kingdom (UK) population using genotype data from the UK Biobank. The UK Biobank is a prospective study of more than 500 000 individuals residing in the UK with long-term follow-up, aimed at exploring genetic and non-genetic determinants of diseases in middle and old age.[16] Participant DNA was genotyped using the UK BiLEVE and UKB Axiom arrays. Approximately 800 000 SNPs that passed quality control were imputed to the UK10 K reference panel, with genotyping called using Affymetrix Power Tools software. For this validation, 6621 patients with AF were used as the case population and 25 981 NSR patients as controls in the analysis.[17]

To validate the Taiwanese AF loci in the UK Biobank, we focused on locus-level association rather than variant-level association. In GWAS across different ethnic populations, the sentinel or lead variants at significant loci may differ due to factors such as variations in LD structure, allele frequencies, imputation quality, and distinct genetic architectures between populations.[18,19] Therefore, we prioritized SNPs located within 500 kb of each significant locus (sentinel SNP) and identified the most significant SNP associated with AF in the UK Biobank. For the validation of previously reported AF loci in our population, we used the same strategy, prioritizing SNPs within 500 kb of the reported loci and searching for the most significant SNPs associated with AF in our Taiwanese GWAS. A significance threshold of $P < 0.05$/tested SNP number was applied. The exploratory population was used for the validation of previously reported AF loci.

Genome-wide genotyping

In the exploratory stage, both patients with AF and NSR controls were genotyped using a custom Axiom Genome-Wide Array Plate, known as the TWB chip 2, containing 714 431 SNPs based on Affymetrix technology. Single-nucleotide polymorphism quality control was performed by excluding SNPs with a missing call rate >0.05, Hardy–Weinberg equilibrium (HWE) P-value $< 1 \times 10^{-6}$ in controls, and a minor allele frequency (MAF) < 0.05.

Genotype imputation

Prior to imputation, we converted hg38 co-ordinates to hg19 co-ordinates using the HRC-1000G-check-bim tool (https://www.well.ox.ac.uk/~wrayner/tools/). Genotype imputation was performed using Minimac4 on the 22 chromosomes in the Michigan Imputation Server,[20] with the 1000 Genomes Project East Asian (EAS) Phase 3 Integrated Release Version 5 Haplotypes as the reference panel. For quality control after imputation, we selected SNPs with imputation quality $R^2 > 0.8$, HWE P-value $> 1 \times 10^{-6}$, MAF $\geq 5\%$, and complete genotypes in >99% of control subjects. A total of 4 512 191 SNPs passed quality control and were used for subsequent GWAS analysis.

Association analysis

Single-nucleotide polymorphism-based association analysis was conducted using PLINK (v1.9)[21] based on a binary outcome model of logistic regression. We adjusted for the top 10 principal components, age, and gender in the analysis. Manhattan plots and quantile–quantile (Q–Q) plots were generated using the qqman package. A genome-wide significance threshold of $P < 5 \times 10^{-8}$ was applied in the exploratory GWAS.

Table 1 Patient characteristics of the study population

Variables	Exploratory population		Replication population	
	AF ($n = 516$)	NSR ($n = 5160$)	AF ($n = 1002$)	NSR ($n = 2003$)
Age, mean ± SD	58.1 ± 8.7	57.8 ± 8.7	68.7 ± 12.1*	60.7 ± 14.2
BMI, mean ± SD	25.3 ± 3.8*	24.4 ± 3.6	24.4 ± 4.6	24.8 ± 4.0
Female, *n* (%)	245 (47.5)	2695 (52.3)	409 (40.8)	916 (45.7)
Male, *n* (%)	271 (52.5)	2460 (47.7)	593 (59.2)	1087 (54.3)
Diabetes, *n* (%)	18 (12.0)*	458 (8.9)	238 (23.8)*	327 (16.3)
Hyperlipidaemia, *n* (%)	25 (16.7)	668 (13.0)	134 (13.4)	278 (13.9)
Hypertension, *n* (%)	50 (33.3)*	1104 (21.4)	576 (57.5)*	819 (40.9)
CAD, *n* (%)	8 (5.3)*	150 (2.9)	246 (24.6)	402 (20.0)

AF, atrial fibrillation; BMI, body mass index; CAD, coronary artery disease; NSR, normal sinus rhythm.
*$P < 0.05$ compared with NSR.

Expression quantitative trait locus analyses

Expression quantitative trait locus (eQTL) mapping of novel AF genetic loci was performed using data from the publicly available Genotype-Tissue Expression (GTEx) database, version 8.[22,23] We treated the GWAS sentinel or lead SNP as a *cis*-eQTL and calculated the *P*-values, effect sizes, and standard errors using linear regression between the expression levels of the nearest gene and the genotypes of the sentinel SNP.[5]

Bayesian colocalization analysis was performed to evaluate whether the same genetic variant is likely responsible for the associations observed in both GWAS and eQTL datasets.[23] Posterior probabilities were calculated for the H4 hypothesis, indicating a shared association (colocalization) between GWAS and eQTL signals. Single-nucleotide polymorphisms available in only one database were also included, with LD adjustments applied. A posterior probability for H4 exceeding 0.75 was considered evidence of co-localization between GWAS and eQTL signals.

Additionally, MotifMap[24] and ChromHMM[25] were used to analyse whether these sentinel SNPs were located within transcription factor binding sites or chromatin-active regions, respectively, if they were significant *cis*-eQTLs.

Transcriptome-wide association study

Transcriptome-wide association study (TWAS) was conducted using MetaXcan.[26] In this approach, associations between predicted expression and AF were estimated based on gene prediction model weights, GWAS summary statistics, and an SNP-correlation LD matrix. The 1000 Genomes Project Phase 3 Version 5 EAS was used as the LD matrix. Genes with significant associations between predicted expression and AF were reported, highlighting potential candidate genes underlying the GWAS signal. In this exploratory TWAS phase, genes with a *P*-value of <0.05 were reported as suggestive associations, and a *P*-value of $<3.1 \times 10^{-6}$ (corrected by the total number of human genes) was considered a significant association. All tissues potentially related to AF mechanisms, including the right atrial appendage, left ventricle, aorta, coronary artery, brain, fibroblasts, and lymphocytes, were tested.

Results

Novel AF genetic loci in the Taiwanese population

The baseline characteristics of the study participants are summarized in *Table 1*. As expected, the mean age and the prevalence of hypertension, coronary artery disease, diabetes, and stroke were higher in patients with AF compared with NSR controls in both the exploratory and validation populations.

GWAS was performed on Stage I subjects using the Axiom Genome-Wide Array Chip (714 431 SNPs; *Figure 2*). After quality control filtering and imputation, a total of 4 512 191 SNPs were analysed. The Manhattan plot of the GWAS is shown in *Figure 2A*, alongside the Q–Q plot. Three loci were associated with AF at genome-wide significance ($P < 5 \times 10^{-8}$) in the exploratory stage (*Figure 2A*).

The most significant association was at 1p13.2/*RAP1A* [rs7525578; MAF: 0.167 in cases vs. 0.0682 in controls; odds ratio (OR): 2.74; 95% confidence interval (CI): 2.28–3.30; $P = 1.24 \times 10^{-26}$]. This was followed by 4q25/*PITX2* (rs2723329; MAF: 0.500 in cases vs. 0.411 in controls; OR: 1.54; 95% CI: 1.35–1.76; $P = 1.53 \times 10^{-10}$) and 8q21.13/*HNF4G* (rs2980218; MAF: 0.437 in cases vs. 0.345 in controls; OR: 1.49; 95% CI: 1.31–1.70; $P = 2.19 \times 10^{-9}$) (*Table 2*). *Figure 3* presents the regional plots for these significant loci.

While the *PITX2* locus has been consistently identified as the most significant AF locus in prior studies,[1–5] the *RAP1A* locus is more significant than the *PITX2* locus and is located <500 kb from the previously reported *KCND3* gene. Our study also identified a novel locus, *HNF4G*, significantly associated with AF in the Taiwanese population.

Validation of GWAS AF variants in a replication population

The significant variants identified in the exploratory GWAS were validated in a replication cohort comprising 1002 patients with AF and 2003 NSR controls (*Table 2*). The associations with AF were replicated for all three variants, with ORs and *P*-values of 1.72 (95% CI: 1.42–2.07; $P = 4.60 \times 10^{-9}$) for *RAP1A* (rs7525578), 1.41 (95% CI: 1.27–1.58; $P = 4.45 \times 10^{-10}$) for *PITX2* (rs2723329), and 1.25 (95% CI: 1.12–1.40; $P = 6.97 \times 10^{-5}$) for *HNF4G* (rs2980218). We also performed a meta-analysis combining the exploratory and validation populations. The association of these three variants with AF was even more significant in this meta-analysis (*Table 2*).

Validation of Taiwanese AF loci in the UK population

We further validated the associations of 4q25/*PITX2*, 1p13.2/*RAP1A*, and 8q21.13/*HNF4G* with AF in the UK Biobank. All SNPs within 500 kb of these loci were analysed, with the most significant SNPs

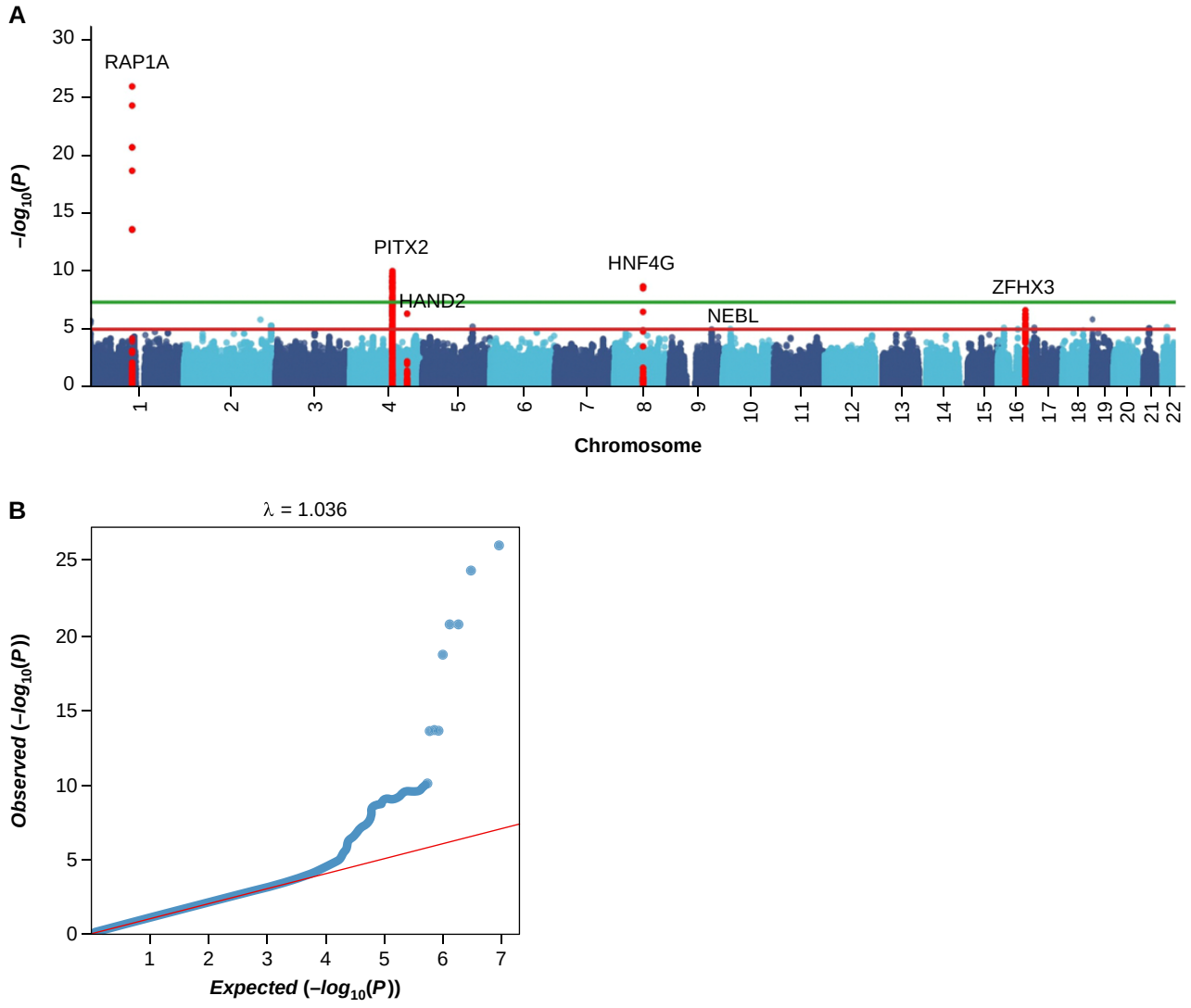


Figure 2 Manhattan and Q–Q plots of GWAS results. (*A*) Three loci/genes (*RAP1A*, *PITX2*, and *HNF4G*) show associations with AF with $P < 5 \times 10^{-8}$ (above the green line), and two loci/genes (*HAND2* and *ZFHX3*) show associations with AF with $P < 10^{-5}$ (above the red line) in the exploratory stage. (*B*) Q–Q plot of the GWAS scan. Most *P*-values were similar to the expected diagonal in the Q–Q plot, which indicates the appropriateness of the GWAS model ($\lambda = 1.036$).

reported. Only the 4q25/*PITX2* locus was significantly associated with AF in the UK population, but not the other two loci. The *P*-values were 2.93×10^{-43} for 4q25/*PITX2* (rs200657990, 126 kb from rs2723329, $r^2 < 0.1$), 0.0015 for 1p13.2/*RAP1A* (rs141883790, 33 kb from rs7525578, $r^2 < 0.1$), and 0.0022 for 8q21.13/*HNF4G* (rs114415293, 1.2 kb from rs2980218, $r^2 < 0.1$).

Validation of previous AF loci in the Taiwanese population

More than 100 genetic loci have been associated with AF in multi-ethnic GWAS meta-analyses.[5,6] We validated these loci in the Taiwanese population, and the results are given in *Table 3*. Loci with *P*-values <0.05 were listed, with significance determined using Bonferroni's correction ($P < 0.05$/tested SNP number for each locus). Among the 138 reported AF loci (tested SNP numbers from 297 to 3542),[5] three genes—*PITX2* (rs2723329), *ZFHX3* (rs2106261), and *HAND2* (rs78164752)—were significantly associated with AF in the Taiwanese population. The most significant association was at *PITX2* as expected, followed by *ZFHX3* (rs2106261; OR: 1.42; 95% CI: 1.24–1.62; $P = 2.42 \times 10^{-7}$) and *HAND2* (rs78164752; OR: 1.43; 95% CI: 1.20–1.70; $P = 4.92 \times 10^{-6}$).

eQTL analyses

The effects of novel susceptibility SNPs on the expression of nearby genes across various human tissues were assessed by investigating eQTLs and colocalization between GWAS and eQTL loci.[23,27] Initial analyses focused on whether these sentinel SNPs were associated with gene expression in the GTEx atrial appendage or left ventricular (LV) samples, but no significant associations were observed. Subsequent analyses in other tissues revealed significant colocalization between the *RAP1A* locus and eQTL in fibroblasts (see Supplementary

Table 2 Summary of genetic study results

Gene	Exploratory population (516 AF vs. 5160 controls)		Validation population (1002 AF vs. 2003 controls)		Combined population (1518 AF vs. 7163 controls)	
Lead SNP	**OR (95% CI)**	***P***	**OR (95% CI)**	***P***	**OR (95% CI)**	***P***
RAP1A rs7525578 (C/T)[a]	2.74 (2.28–3.30)	1.24×10^{-26}	1.72 (1.42–2.07)	4.60×10^{-9}	2.07 (1.83–2.35)	1.13×10^{-28}
PITX2 rs2723329 (T/C)[a]	1.54 (1.35–1.76)	1.53×10^{-10}	1.41 (1.27–1.58)	4.45×10^{-10}	1.51 (1.40–1.63)	2.15×10^{-21}
HNF4G rs2980218 (G/A)[a]	1.49 (1.31–1.70)	2.19×10^{-9}	1.25 (1.12–1.40)	6.97×10^{-5}	1.33 (1.23–1.44)	7.89×10^{-11}

AF, patients with atrial fibrillation; OR, odds ratio per minor allele (additive model); CI, confidence interval; SNP, single-nucleotide polymorphism.
[a]Major allele/minor allele

material online, *Figure S1A*) and between the *HNF4G* locus and eQTL in lymphocytes (see Supplementary material online, *Figure S1B*).

Given that these three sentinel SNPs were identified as potential *cis*-eQTLs, further investigation was conducted using MotifMap and ChromHMM to determine whether they were located within transcription factor binding sites or chromatin-active regions. SNP rs2723329 was localized in a region containing transcription factor binding sites for *IRF8, HNF4, NFAT2*, and *MTF1* (MotifMap). Chromatin state analyses (ChromHMM) suggested an inactive promoter or weak enhancer region (see Supplementary material online, *Figure S2A*). The region 60 kb upstream of rs2980218 contained predicted binding sites for *TEF-1, AP-3*, and *IRF8*, while the region 50 kb downstream showed binding sites for *TEF-1* and *STAT6* (MotifMap). SNP rs2980218 was situated in a chromatin region indicative of an active promoter or weak enhancer (ChromHMM) (see Supplementary material online, *Figure S2B*).

SNP rs7525578 was localized within an intron of the *RAP1A* gene, with chromatin state analyses indicating an actively transcribed region (ChromHMM) (see Supplementary material online, *Figure S2C*). Additionally, rs7525578 was found to be near (100 kb) a transcription factor binding site region associated with *HSF2, HNF4, LEF1, TEF-1*, and *SOX10* (MotifMap).

TWAS analysis

The TWAS results are summarized in *Table 4*. The most significant gene identified in the atrial appendage was *GPHN* ($P = 7.23 \times 10^{-9}$). Although *GPHN* did not demonstrate significant expression in LV tissue, it also exhibited high significance in brain tissue ($P = 3.30 \times 10^{-9}$). Of note, while GPHN was the top hit in TWAS, it was not a genome-wide significant locus. The *CASQ2* gene showed a suggestive association in both atrial appendage ($P = 0.010$) and LV ($P = 0.015$) tissues, consistent with findings from two large multi-ethnic AF GWAS and TWAS studies.[5,6]

Among the three GWAS-significant loci, the 8q21.13 and 1p13.2 loci also demonstrated significant associations in the TWAS. Expression of the *HNF4G* gene in brain tissue and lymphocytes was significantly associated with AF ($P = 3.30 \times 10^{-7}$ in brain tissue and 4.22×10^{-8} in lymphocytes). Similarly, expression of the *RAP1A* gene in fibroblasts was significantly associated with AF ($P = 1.46 \times 10^{-9}$). In contrast, TWAS analysis of the 4q25 locus did not identify any significant gene expression associated with AF.

Discussion

In this study, we identified a novel 8q21.13/*HNF4G* locus associated with AF specifically in the Taiwanese population, which had not been previously reported. Alongside Japanese and Korean AF GWAS, our study represents the third AF GWAS conducted in the Asia–Pacific region. Our sample size was similar to that of the Korean AF GWAS. In all three AF GWAS studies in the Asia–Pacific region, the 4q25/*PITX2* locus was consistently and highly significantly associated with AF. In our Taiwanese AF GWAS, the 1p13.2/*RAP1A* locus showed even greater significance than the 4q25/*PITX2* locus. The association between the novel 8q21.13/*HNF4G* locus and AF was also significant in TWAS.

Thus, we contribute a new AF locus (8q21.13/*HNF4G*) to the AF GWAS catalog. In Japanese AF GWAS studies, six new loci (with nearest genes *KCND3, PPFIA4, SLC1A4-CEP68, HAND2, NEBL*, and *SH3PXD2A*) were identified as novel AF loci.[7] Among these, *PPFIA4* and *HAND2* were validated in the Korean AF GWAS.[8] Although *HAND2* gene was not significantly associated with AF in our AF GWAS ($P > 5 \times 10^{-8}$), the locus within 1 Mb of the *HAND2* gene was identified as a significant region in the replication analysis of previously reported AF loci in our study (*Table 3*). *HAND2* was first found as an AF gene/locus in the Japanese population[7] and then validated in the Korean population,[8] both of which are Asian populations. *HAND2* locus was also significant in our Taiwanese population, also an Asian population.

There has been significant progress in the genetic study of cardiac arrhythmias, including both hereditary channelopathies and common arrhythmias such as AF.[28] Genetic research has greatly advanced the diagnosis and treatment of cardiac arrhythmias, particularly hereditary channelopathies. However, due to AF as a complex trait, applying genetic study results to AF treatment remains challenging. At the current stage, these findings are more effectively utilized in guiding clinical genetic testing for diagnosing or predicting the risk of AF. Our discovery of *HNF4G* and *RAP1A* establishes an understandable connection between genetic studies and mechanistic insights, potentially offering a more straightforward path for developing novel AF treatments based on these findings.

A plausible explanation for the association of the *HNF4G* gene with AF lies in its previously reported role as a risk gene for obesity in obesity-related GWAS.[29,30] Obesity is a well-established risk factor for AF,[31,32] and weight reduction by itself has been shown to decrease the burden of AF.[33] In our AF GWAS cohort, patients with AF had a significantly higher mean body mass index (BMI) compared with NSR controls. This suggests the possibility that the association of *HNF4G* with AF may be body-weight dependent. However, in our replication cohort, the mean BMI of patients with AF did not differ significantly from that of NSR controls, yet a significant association with the *HNF4G* gene was still observed. This suggests that the association of *HNF4G* with AF may extend beyond BMI or obesity[34,35] and may involve metabolic or inflammatory pathways.[36,37] This is supported by

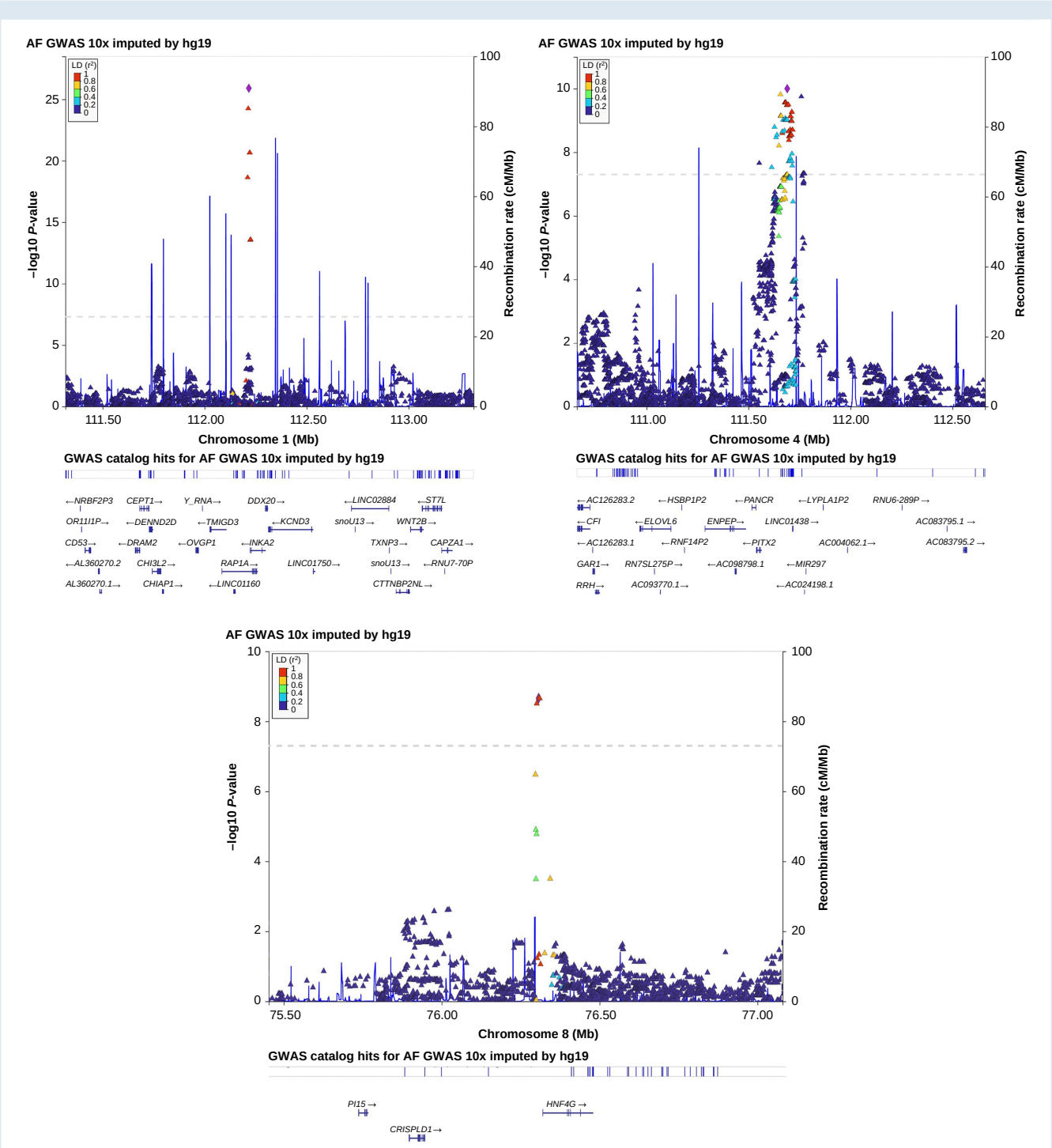


Figure 3 Regional plots of variants that show associations with AF with $P < 5 \times 10^{-8}$. Regional plots of SNPs centred on the lead SNP around the *RAP1A* gene (left upper panel), *PITX2* gene (right upper panel), and *HNF4G* gene (lower panel) are shown. For each SNP, the chromosomal location is shown on the *x*-axis and the significance level for association with AF is indicated by a −log10 *P*-value on the *y*-axis. *P*-values are expressed as −log10 (*P*) (*y*-axis) for every tested SNP ordered by chromosomal location (*x*-axis).

eQTL colocalization and TWAS results in our study that revealed a significant association between *HNF4G* gene expression in lymphocytes and AF.

In our study, the 1p13.2/*RAP1A* locus was more significantly associated with AF than the 4q25/*PITX2* locus when considering the *P*-value, a finding not previously reported. While *P*-values are important, the effect size (OR) of an association is also critical and can sometimes hold greater significance. In our study, the effect size was 1.5, which aligns with findings from the combined ancestral GWAS of Caucasian populations.[4] In contrast, the OR reported in a Japanese

Table 3 Replication of known AF loci in the Taiwanese population

#Chr	Position	SNP	Nearest gene	Allele[a]	OR	95% CI	*P*	#SNP
1	22282619	rs7529220	*HSPG2*	C/T	0.772	0.655–0.910	0.0020505	2243
1	112392360	rs12044963	*KCND3*	G/T	1.164	1.020–1.328	0.0242902	2223
1	116297758	rs4073778	*CASQ2*	C/A	1.203	1.051–1.377	0.0072332	2281
1	170591310	rs651386	*PRRX1*	A/T	0.860	0.751–0.984	0.0282370	1723
1	203026591	rs17461925	*PPFIA4*	A/G	0.803	0.680–0.949	0.0100012	2295
3	179170494	rs4855075	*GNB4*	C/T	1.220	1.023–1.454	0.0266717	1190
4	111658394	rs2723329	*PITX2*[b]	T/C	1.542	1.350–1.763	1.53E-10	1736
4	174448143	rs78164752	*HAND2*[b]	T/G	1.434	1.248–1.702	4.92E-06	1115
5	113736416	rs716845	*KCNN2*	G/A	1.223	1.009–1.482	0.0401149	1384
8	141746324	rs4355822	*PTK2*	A/G	0.835	0.727–0.960	0.0110472	1245
10	21157621	rs2296610	*NEBL*	G/T	1.349	1.144–1.591	0.0003795	1881
14	32981484	rs2145587	*AKAP6*	G/A	1.277	1.121–1.455	0.0002357	1488
16	73051620	rs2106261	*ZFHX3*[b]	C/T	1.421	1.244–1.624	2.42E-07	1649
17	44019712	rs242557	*MAPT*	A/G	1.141	1.001–1.299	0.0478401	368

AF, atrial fibrillation; #Chr, chromosome number; CI, confidence interval; OR, odds ratio per minor allele (additive model); SNP, single-nucleotide polymorphism; #SNP, SNP number tested for association.
[a]Major allele/minor allele.
[b]Significant replicated gene.

Table 4 Genes significantly associated with AF in transcriptome-wide association study

Gene name	RAP *P*	LV *P*	Brain *P*	Chr	Sentinel SNP	SNP position	GWAS *P*	Nearest gene	Risk locus
GPHN	7.23E−09	3.40E−01	3.30E−09	14	rs992474	66699865	0.00251793	*ENSG00000287833*	—
TRPC1	9.96E−08	1.05E−01	3.92E−02	3	3:142312975:C:CTT	142312975	4.37E−05	*PLS1*	—
IKBKAP	5.06E−06	2.22E−05	1.76E−05	9	9:111661422:C:CA	111661422	1.06E−05	*ELP1*	—
FMO2	1.34E−05	1.62E−05	1.49E−05	1	rs72714189	171266804	0.00587934	*FMO1*	—
RP11-188D8.1	2.38E−04	2.55E−05	6.84E−10	1	1:118026324:C:A	118026324	4.47E−05	*ENSG00000279513*	—
SNX24	4.21E−03	3.17E−01	1.15E−08	5	rs56071867	121831707	0.00300478	*MGC32805*	—
TMEM59L	9.63E−03		2.76E−06	19	rs12610537	18687225	2.29E−05	*UBA52*	—
CASQ2	1.04E−02	1.47E−02	1.59E−07	1	rs9428221	116300137	0.0029283	*CASQ2*	—
PGRMC2	9.71E−02	5.69E−02	5.87E−08	4	rs79555688	129709201	0.00038062	*JADRR*	—
RNF126	2.70E−01		4.56E−08	19	rs8113305	1088017	8.11E−05	*ARHGAP45*	—
OCIAD1	4.02E−01	1.03E−01	8.57E−07	4	rs6814786	48313005	0.0366267	*SLAIN2*	—
ALPK1	5.29E−01	5.41E−01	2.54E−15	4	rs13108156	113115904	0.00073831	*FAM241A*	—
UBE4A	5.92E−01	1.93E−10	2.54E−09	11	rs572460646	118262515	0.00158631	*ENSG00000254873*	—
GTF2E2	6.54E−01	5.30E−02	2.85E−06	8	8:30386410:A:AT	30386410	2.13E−05	*ENSG00000279041*	—
USP8	6.64E−01	4.98E−01	1.89E−06	15	rs1294724317	51122054	0.0339529	*ENSG00000273674*	—
CYTH3	7.55E−01		2.81E−06	7	rs145450974	6093078	0.00156747	*RNU6-218P*	—
HNF4G			3.30E−07	8	rs2980218	76306804	2.19E−09	*HNF4G*	8q21.13
ZAR1			9.38E−09	4	rs10938521	48113219	0.00440531	*RNU6-868P*	—
SPSB3			8.12E−07	16	rs7199735	1765171	0.00130256	*MAPK8IP3-AS1*	—

Chr, chromosome; GWAS, genome-wide association study; LV, left ventricular tissue; RAP, right atrial appendage tissue; SNP, single-nucleotide polymorphism.

population AF GWAS was 2.0,[7] as observed in a Korean AF GWAS study.[8] Therefore, compared with other Asian populations, the effect size in our study was smaller and more closely resembled that of Caucasian populations. This difference may partly explain why the 4q25 locus was not the most significant in our study. However, racial differences in genetic susceptibility to AF may also play a role.

The *RAP1A* locus lies within 500 kb of the previously reported *KCND3* gene, though it remains uncertain which gene is causal in this locus. Our eQTL and TWAS results suggest *RAP1A* as a potential causal gene. *RAP1A* encodes for Rap1 GTPase, which plays a role in reducing fibrotic gene expression and myofibroblast activation to counteract cardiac fibrosis.[38] Our eQTL and TWAS results also indicated a significant association of *RAP1A* expression in fibroblasts with AF, suggesting its involvement in atrial fibrosis and AF.[39,40] Furthermore, *RAP1A* has been implicated in cardiac development and function,[38] with studies showing that knockdown of *RAP1A* in zebrafish heart results in reduced connexin 43 and conduction block, key mechanisms in AF.[38]

In our TWAS analysis, the 4q25/*PITX2* locus did not show any significant gene associations. Similarly, recent large multi-ethnic AF GWAS studies have failed to link the 4q25 locus to *PITX2* expression in any tissue, including cardiac tissue samples. Negative TWAS results may arise for various reasons and do not necessarily rule out a functional role for the nearest gene. Therefore, functional studies continue to explore the relationship between *PITX2* and AF,[41] even though no definitive AF risk link to *PITX2* expression levels has been confirmed by TWAS.

Our novel 8q21.13 locus was also linked to a significant TWAS result for *HNF4G* in brain tissue. As mentioned, *HNF4G* is associated with obesity, and interestingly, TWAS on obesity has revealed significant genes in the brain,[42] suggesting a neurogenic contribution to obesity. The association of *HNF4G* in the brain with AF could indicate a neurogenic mechanism underlying AF. In our TWAS analysis, the most significant gene was *GPHN*, which encodes gephyrin, a neuronal assembly protein that anchors inhibitory neurotransmitter receptors.[43] Furthermore, we previously reported the association of *KCNIP1*, a potassium channel gene highly expressed in the brain, with AF.[10]

Despite a limited number of cases in the exploratory phase of our GWAS, we identified a novel locus that was later confirmed in a larger cohort of over 1000 patients with AF. This finding highlights the stochastic nature of GWAS, where certain causal genes may be identified in one study but not in others, irrespective of sample size. Additionally, our cohort consisted entirely of patients with symptomatic AF treated in tertiary medical centres, in contrast to larger GWAS studies that often include patients from biobanks, some of whom may have low AF burden or asymptomatic AF. Furthermore, we excluded patients with atrial flutter, which many large AF GWAS studies include.

This study has several limitations. First, we did not directly prove that *RAP1A* and *HNF4G* are involved in the mechanism of AF. Future studies, including knockout mice models, may be necessary to establish the causal relationship between *RAP1A* and atrial fibrosis or *HNF4G* and obesity-related inflammatory pathways in AF. Second, in our eQTL analyses, although *PITX2* and *RAP1A* are known to be expressed in the left atrium, the significant SNPs for these genes were not associated with *PITX2* and *RAP1A* in atrial or ventricular tissues. Direct analyses in human surgical left atrium tissue are warranted.

In conclusion, in this first Taiwanese AF GWAS, we identified a novel *HNF4G* locus associated with AF, and its significance was further confirmed by TWAS. Additionally, we found that the 1p13.2/*RAP1A* locus was more significant than the well-established 4q25/*PITX2* locus. These genes may involve fibrosis, inflammatory, and metabolic pathways in the mechanism of AF, thereby contributing to a deeper understanding of AF pathology.

Supplementary material

Supplementary material is available at *Europace* online.

Authors' contribution

G.-W.L. designed the study, analysed the genetic data, and wrote the manuscript. J.-J.C. conceived the study, collected patients' data, and revised the manuscript. C.-H.W. collected surgical samples for functional studies. S.-N.C. and F.-C.C. collected patients' data. P.-S.H. collected patients' data and analysed the data. S.-K.C. revised the manuscript. E.Y.C., designed the genetic panel and analysed the genetic data. C.-T.T. designed the study and genetic panel, collected patients' data, revised the manuscript, and was in charge of the whole study.

Acknowledgements

This research was conducted using data from the UK Biobank (www.ukbiobank.ac.uk), a major biomedical database, approved under project no. 54423.

Funding

This work was partially supported by grants from the National Science and Technology Council (110-2314-B-002-217-MY2, 110-2314-B-002-198-MY3, 111-2218-E-002-043 and 112-2314-B-002-281).

Conflict of interest: The authors declare that they have no competing interests.

Data availability

The raw data supporting the conclusions of this article will be made available by the authors upon specific request.

Ethic approval

The study was approved by the IRB of the National Taiwan University Hospital (200911002R).

Consent to participate

Written informed consent was obtained from all the participating individuals in the National Taiwan University Hospital.

References

1. Benjamin EJ, Rice KM, Arking DE, Pfeufer A, van Noord C, Smith AV *et al.* Variants in ZFHX3 are associated with atrial fibrillation in individuals of European ancestry. *Nat Genet* 2009;**41**:879–81.
2. Ellinor PT, Lunetta KL, Glazer NL, Pfeufer A, Alonso A, Chung MK *et al.* Common variants in KCNN3 are associated with lone atrial fibrillation. *Nat Genet* 2010;**42**:240–4.
3. Gudbjartsson DF, Arnar DO, Helgadottir A, Gretarsdottir S, Holm H, Sigurdsson A *et al.* Variants conferring risk of atrial fibrillation on chromosome 4q25. *Nature* 2007;**448**:353–7.
4. Roselli C, Chaffin MD, Weng LC, Aeschbacher S, Ahlberg G, Albert CM *et al.* Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet* 2018;**50**:1225–33.
5. Roselli C, Rienstra M, Ellinor PT. Genetics of atrial fibrillation in 2020: GWAS, genome sequencing, polygenic risk, and beyond. *Circ Res* 2020;**127**:21–33.
6. Miyazawa K, Ito K, Ito M, Zou Z, Kubota M, Nomura S *et al.* Cross-ancestry genome-wide analysis of atrial fibrillation unveils disease biology and enables cardioembolic risk prediction. *Nat Genet* 2023;**55**:187–97.
7. Low SK, Takahashi A, Ebana Y, Ozaki K, Christophersen IE, Ellinor PT *et al.* Identification of six new genetic loci associated with atrial fibrillation in the Japanese population. *Nat Genet* 2017;**49**:953–8.
8. Lee JY, Kim TH, Yang PS, Lim HE, Choi EK, Shim J *et al.* Korean atrial fibrillation network genome-wide association study for early-onset atrial fibrillation identifies novel susceptibility loci. *Eur Heart J* 2017;**38**:2586–94.
9. Gorlova OY, Xiao X, Tsavachidis S, Amos CI, Gorlov IP. SNP characteristics and validation success in genome wide association studies. *Hum Genet* 2022;**141**:229–38.
10. Tsai CT, Hsieh CS, Chang SN, Chuang EY, Ueng KC, Tsai CF *et al.* Genome-wide screening identifies a KCNIP1 copy number variant as a genetic predictor for atrial fibrillation. *Nat Commun* 2016;**7**:10190.
11. Tsai CT, Lai LP, Lin JL, Chiang FT, Hwang JJ, Ritchie MD *et al.* Renin-angiotensin system gene polymorphisms and atrial fibrillation. *Circulation* 2004;**109**:1640–6.
12. Chang SN, Lai LP, Chiang FT, Lin JL, Hwang JJ, Tsai CT. C-reactive protein gene polymorphism predicts the risk of thromboembolic stroke in patients with atrial fibrillation: a more than 10-year prospective follow-up study. *J Thromb Haemost* 2017;**15**:1541–6.
13. Tsai CT, Hsieh CS, Chang SN, Chuang EY, Juang JM, Lin LY *et al.* Next-generation sequencing of nine atrial fibrillation candidate genes identified novel de novo mutations in patients with extreme trait of atrial fibrillation. *J Med Genet* 2015;**52**:28–36.

14. Wong HS, Tsai SY, Chu HW, Lin MR, Lin GH, Tai YT *et al.* Genome-wide association study identifies genetic risk loci for adiposity in a Taiwanese population. *PLoS Genet* 2022;**18**:e1009952.
15. Huang PS, Hsieh CS, Chang SN, Chen JJ, Chiu FC, Wu CK *et al.* Prevalence of sudden arrhythmic death syndrome-related genetic mutations in an Asian cohort of whole genome sequence. *Europace* 2020;**22**:1287–97.
16. Allen NE, Sudlow C, Peakman T, Collins R; UK Biobank. UK biobank data: come and get it. *Sci Transl Med* 2014;**6**:224ed4.
17. Lane JM, Vlasac I, Anderson SG, Kyle SD, Dixon WG, Bechtold DA *et al.* Genome-wide association analysis identifies novel loci for chronotype in 100,420 individuals from the UK Biobank. *Nat Commun* 2016;**7**:10889.
18. Wojcik GL, Graff M, Nishimura KK, Tao R, Haessler J, Gignoux CR *et al.* Genetic analyses of diverse populations improves discovery for complex traits. *Nature* 2019;**570**:514–8.
19. Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S *et al.* Human demographic history impacts genetic risk prediction across diverse populations. *Am J Hum Genet* 2017;**100**:635–49.
20. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A *et al.* Next-generation genotype imputation service and methods. *Nat Genet* 2016;**48**:1284–7.
21. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;**81**:559–75.
22. GTEx Consortium; Laboratory, Data Analysis & Coordinating Center (LDACC)—Analysis Working Group; Statistical Methods groups—Analysis Working Group; Enhancing GTEx (eGTEx) groups; NIH Common Fund; NIH/NCI *et al.* Genetic effects on gene expression across human tissues. *Nature* 2017;**550**:204–13.
23. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C *et al.* Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet* 2014;**10**:e1004383.
24. Xie X, Rigor P, Baldi P. MotifMap: a human genome-wide map of candidate regulatory motif sites. *Bioinformatics* 2009;**25**:167–74.
25. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods* 2012;**9**:215–6.
26. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet* 2016;**48**:245–52.
27. Hormozdiari F, van de Bunt M, Segre AV, Li X, Joo JWJ, Bilow M *et al.* Colocalization of GWAS and eQTL signals detects target genes. *Am J Hum Genet* 2016;**99**:1245–60.
28. Crotti L, Brugada P, Calkins H, Chevalier P, Conte G, Finocchiaro G *et al.* From gene-discovery to gene-tailored clinical management: 25 years of research in channelopathies and cardiomyopathies. *Europace* 2023;**25**:euad180.
29. Ayari S, Gil-lturbe E, le Gleau L, Osinski C, Kapel N, Soula HA *et al.* Hnf4g invalidation prevents diet-induced obesity via intestinal lipid malabsorption. *J Endocrinol* 2021;**252**:31–44.
30. Berndt SI, Gustafsson S, Magi R, Ganna A, Wheeler E, Feitosa MF *et al.* Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nat Genet* 2013;**45**:501–12.
31. Wang TJ, Parise H, Levy D, D'Agostino RB Sr, Wolf PA, Vasan RS *et al.* Obesity and the risk of new-onset atrial fibrillation. *JAMA* 2004;**292**:2471–7.
32. Wanahita N, Messerli FH, Bangalore S, Gami AS, Somers VK, Steinberg JS. Atrial fibrillation and obesity–results of a meta-analysis. *Am Heart J* 2008;**155**:310–5.
33. Pathak RK, Middeldorp ME, Meredith M, Mehta AB, Mahajan R, Wong CX *et al.* Long-term effect of goal-directed weight management in an atrial fibrillation cohort: a long-term follow-up study (LEGACY). *J Am Coll Cardiol* 2015;**65**:2159–69.
34. Kolwicz SC J, Purohit S, Tian R. Cardiac metabolism and its interactions with contraction, growth, and survival of cardiomyocytes. *Circ Res* 2013;**113**:603–16.
35. McCauley MD, Hong L, Sridhar A, Menon A, Perike S, Zhang M *et al.* Ion channel and structural remodeling in obesity-mediated atrial fibrillation. *Circ Arrhythm Electrophysiol* 2020;**13**:e008296.
36. Packer M, Lam CSP, Lund LH, Maurer MS, Borlaug BA. Characterization of the inflammatory-metabolic phenotype of heart failure with a preserved ejection fraction: a hypothesis to explain influence of sex on the evolution and potential treatment of the disease. *Eur J Heart Fail* 2020;**22**:1551–67.
37. Esposito K, Giugliano D. Diet and inflammation: a link to metabolic and cardiovascular diseases. *Eur Heart J* 2006;**27**:15–20.
38. Dong W, Yang Z, Yang F, Wang J, Zhuang Y, Xu C *et al.* Suppression of Rap1 impairs cardiac myofibrils and conduction system in zebrafish. *PLoS One* 2012;**7**:e50960.
39. Burr SD, Stewart JA Jr. Rap1a overlaps the AGE/RAGE signaling cascade to Alter expression of alpha-SMA, p-NF-kappaB, and p-PKC-zeta in cardiac fibroblasts isolated from type 2 diabetic mice. *Cells* 2021;**10**:557.
40. Burr SD, Stewart JA Jr. Rap1a regulates cardiac fibroblast contraction of 3D diabetic collagen matrices by increased activation of the AGE/RAGE cascade. *Cells* 2021;**10**:1286.
41. Bai J, Zhu Y, Lo A, Lu Y, Zhao J. In silico assessment of genetic variation in PITX2 reveals the molecular mechanisms of calcium-mediated cellular triggered activity in atrial fibrillation. *Annu Int Conf IEEE Eng Med Biol Soc* 2020;**2020**:2353–6.
42. Ang MY, Takeuchi F, Kato N. Deciphering the genetic landscape of obesity: a data-driven approach to identifying plausible causal genes and therapeutic targets. *J Hum Genet* 2023;**68**:823–33.
43. Choii G, Ko J. Gephyrin: a central GABAergic synapse organizer. *Exp Mol Med* 2015;**47**:e158.