

SCIENTIFIC REPORTS



OPEN

Integrative Genome-Wide Association Studies of eQTL and GWAS Data for Gout Disease Susceptibility

Meng-tse Gabriel Lee¹, Tzu-Chun Hsu¹, Shyr-Chyr Chen¹, Ya-Chin Lee², Po-Hsiu Kuo^{2,3}, Jenn-Hwai Yang⁴, Hsiu-Hao Chang⁵ & Chien-Chang Lee¹

There is a paucity of genome-wide association study on Han Chinese gout patients. We performed a genome-wide association meta-analysis on two Taiwanese cohorts consisting of 758 gout cases and 14166 controls of Han Chinese ancestry. All the participants were recruited from the Taiwan Biobank. For pathway analysis, we applied ICSNPathway (Identify candidate Causal SNPs and Pathways) analysis, and to investigate whether expression-associated genetic variants contribute to gout susceptibility, we systematically integrated lymphoblastoid expression quantitative trait loci (eQTL) and genome-wide association data of gout using *Sherlock*, a Bayesian statistical framework. In the meta-analysis, we found 4 SNPs that reached genome-wide statistical significance ($P < 5.0 \times 10^{-8}$). These SNPs are in or close to *ABCG2*, *PKD2* and *NUDT9* gene on chromosome 4. ICSNPathway analysis identified rs2231142 as the candidate causal SNP, and *ABCG2* as the candidate gene. Sherlock analysis identified three genes, which were significantly associated with the risk of gout (*PKD2*, *NUDT9*, and *NAP1L5*). To conclude, we reported novel susceptible loci for gout that has not been previously addressed in the literature.

Gout is a common inflammatory arthritis that results from the deposition of monosodium urate crystals in joints. Epidemiological studies from a range of countries suggest a high prevalence of gout. Gout affects more than 3 million adults in the United States, and over 700,000 adults in the United Kingdom (UK)^{1–3}. In the UK, gout is a common inflammatory joint disease affecting 2.5% of the population in 2012, with prevalence as high as 14% in men aged over 75 years⁴.

A prerequisite for gout development is hyperuricaemia, which is caused by an imbalance in the rates of production and excretion of uric acid. Most gout patients have hyperuricaemia and a clear concentration-dependent relationship exists between serum urate concentrations and incident gout⁵. The familial nature of hyperuricaemia and gout has been recognized since 17th century⁶. Genome-wide data in Europeans estimate that the heritability of serum urate levels was 27–41% and heritability of gout was approximately 30%^{7,8}. The overall pattern of inheritance is best explained by a complex model incorporating interactions between more than one major gene, several modifying genes and environmental factors.

Genome-wide association studies (GWASs) have explored many genes associated with gout, for instance, *ABCG2*, *PKD2*, *SLC2A9*, *KCNQ1*, *SLC22A12* and *SLC17A1* for gout disease among individuals of European descent^{9–12}. However, the genetic studies conducted to date have largely been restricted to patients of European ancestry, and there were limited GWAS studies in the Han Chinese population^{13,14}. Therefore, conducting a GWAS study in the Han Chinese population may contribute to the understanding of the genetic causes of gout.

The standard data analysis of GWAS is based on a single SNP and may ignore the combined effect of modest SNPs/genes. To solve this problem, pathway-based analyses have been developed to extract more biological

¹Department of Emergency Medicine, National Taiwan University Hospital, Taipei, Taiwan. ²Department of Public Health and Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan. ³Research Center for Genes, Environment and Human Health, National Taiwan University, Taipei, Taiwan. ⁴National Center for Genome Medicine, Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan. ⁵Department of Pediatrics, National Taiwan University Hospital, National Taiwan University, College of Medicine, Taipei, Taiwan. Correspondence and requests for materials should be addressed to C.-C.L. (email: clee100@gmail.com)

information from existing GWAS datasets. The ICSNPPathway (identify candidate causal SNPs and pathways) analysis has been developed to identify candidate SNPs and their corresponding candidate pathways using GWAS data and by integrating linkage disequilibrium (LD) analysis, functional SNP annotation, and pathway-based analysis¹⁵. Thus, the integrative analysis using ICSNPPathway might provide new insights for the understanding on the genetic basis of gout.

In addition, recent studies have used integrative strategies to combine results from association studies and eQTL (expression quantitative trait loci) analyses to interrogate the potential regulatory effect of the susceptibility SNPs in GWAS. He *et al.* developed a tool called Sherlock to systematically explore the role of a gene in complex diseases by integrating not only eQTL cis- but also trans-effects of that gene in GWAS¹⁶. This tool has been found to uncover many new susceptible genes that cannot be identified using GWAS alone in different diseases such as Crohn's Disease, schizophrenia, and psoriasis^{16–18}. As far as we were aware of, an integrative analysis of lymphoblastoid eQTL and GWAS data for gout disease susceptibility has not been conducted in Han Chinese. Hence, one aim of this research is to explore susceptibility genes in lymphocytes with regulatory function in gout by using Sherlock.

Therefore, the aim of this study is three-fold. The first aim was to identify genetic loci related to gout using GWAS in the Han Chinese population in Taiwan. The second aim is to conduct pathway analysis using the ICSNPPathway method, to identify SNP and pathways related to gout. Third, we aimed to explore susceptibility genes in lymphocytes with regulatory function in gout by using Sherlock, a tool that integrates not only eQTL cis but also trans-effects of that gene in GWAS.

Methods

Study population. This study incorporated 15,300 Taiwanese Han subjects randomly selected from the Taiwan Biobank. Taiwan Biobank is a population-based biomedical research database that has collected detailed health and lifestyle information on participants^{19,20}. Inclusion criteria were individuals who were aged between 30–70 years old and self-reported as being of Taiwanese Han Chinese ancestry. Patients diagnosed with cancer were excluded. In addition, aboriginal people and descents of foreigners were excluded to avoid population substructure. According to a recent study investigating the population admixture of Han Chinese residing in Taiwan, a high homogeneity was demonstrated among the Taiwanese subpopulations²⁰.

Study Variables. Participants of Taiwan Biobank were asked to fill a detailed questionnaire form. The detailed questionnaire form contained information on demographics, and previous medical history. Gout cases were identified from the self-reported questionnaire form, which has been evidenced to be the best test performance characteristics of existing definitions with sensitivity 80% and specificity 72%²¹. Controls were those without self-reported gout.

Genotyping and quality controls. Whole genome genotyping was performed using the customized Axiom-Taiwan Biobank Array Plate (TWB chip; Affymetrix Inc, CA, USA) for both the GWAS and replication samples. Containing 653,291 SNPs, TWB chip was designed to screen SNPs in genome-wide scale especially for Han-Chinese descent in Taiwan. The genotype information and linkage disequilibrium (LD) of healthy subjects have been released by the ethic and governance council of Taiwan Biobank (TaiwanView: <http://taiwanview.twbiobank.org.tw>).

Quality control procedures were done using plink with each individual, including gender concordance, sample quality, kinship, and population stratification (Supplementary Table 1). We did not observe any participant with sex mis-match for the discovery sample. No participants were removed at a call rate >0.97. However, when we searched for close relatives using identity-by-descent (IBD), 206 individuals with strong kinship (IBD > 0.8) were eliminated. To evaluate potential stratification in our study population, we also performed a principal component analysis (PCA). We identified no outliers from the scatter plot (Supplementary Figure 1). As a result, 7094 subjects, including 373 gout patients and 6721 healthy controls were retained. For the follow-up sample, following the same quality control procedures for individuals, 170 subjects were removed, resulting in a total of 7830 subjects, including 385 cases and 7445 controls. Quality control was also performed for SNPs. We removed markers if they failed Hardy-Weinberg tests with $P < 0.0001$, genotype missing rate >5%, and minor allele frequency (MAF) < 0.05. As a result, a total of 631,941 SNPs in the discovery samples, and 621,874 SNPs in the follow-up samples were retained (Supplementary Table 2).

Statistical Analyses. *GWAS analysis.* Using 607,675 SNPs after quality control, the association of SNPs with the phenotype was tested by multivariate logistic regression analysis with adjustment for age at recruitment, gender, and the first 10 principal components. Odds ratios were calculated by considering the non-risk allele as a reference. We determined the minimum P value under three genetic models (additive, recessive and dominant). Ten principal components were included as covariates in the logistic regression model to control for population stratification, although genomic inflation was acceptable (<1.006) even before this correction was applied. The genomic inflation factor was derived by applying P values from logistic regression in an additive model for all the tested SNPs. A quantile–quantile plot of GWAS was used to examine the P -value distribution (Supplementary Fig. 2).

We decided to use the significance threshold of $P = 5.0 \times 10^{-8}$ in the fixed effect meta-analysis combining both discovery and follow-up sample. Power analysis can be found in Supplementary Table 3. Heterogeneity among the studies was determined by Cochran's Q statistic. LocusZoom plots were created using the LocusZoom tool (found at <http://locuszoom.sph.umich.edu/locuszoom/>) and the “hg19/1000 Genomes Nov 2014 ASN” panel was selected²². For general statistical analysis, we used R statistical environment version 3.51 or PLINK version 1.9. This research project was approved by the ethics committee of National Taiwan University Hospital Institutional Review Board. The study was conducted in accordance with the principles of the Declaration of Helsinki and the Good Clinical Practice Guidelines, and all the participants were informed consent.

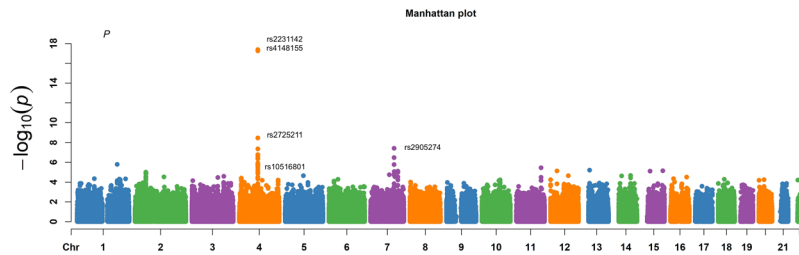


Figure 1. Manhattan plots for genome-wide SNPs associated with gout. Results of genome-wide association analysis ($-\log_{10} P$) shown in chromosomal order for 631, 941 SNPs tested for association in initial sample of 373 cases and 6721 controls. The x axis represents each of the SNPs used in the primary scan. The y axis represents the $-\log_{10} P$ -value obtained by logistic regression analysis (additive model) with adjustment for age, gender, and 10 principal components.

ICSNPathway using GWAS data. We applied ICSNPathway analysis to the full list of gout GWAS SNPs p value¹⁵. ICSNPathway analysis involves two stages: (1) SNP clumping, which pruned SNPs by LD while prioritizing by p-value; (2) annotation of the biological mechanisms to pre-selected candidate SNPs using a pathway-based algorithm named i-GSEA (improved-gene set enrichment analysis). To avoid stochastic bias and the testing to general biological processes, we discarded pathways that contained <5 or >20 genes.

Sherlock. Using the web-based tool Sherlock, we implemented the integrated analysis of GWAS data and public lymphoblastoid eQTL data¹⁶. Lymphoblastoid B cells are selected as these cells are involved in the acute stage of gout. The underlying assumption is that the expression level of a specific gene(s) may influence the risk of a disease (eg, gout). Therefore, genetic variation (both in *cis*, and in *trans*) that perturbs gene expression may affect the risk of this disease. *Sherlock* first searches for all eSNPs of each gene using the whole genome eQTL data from lymphoblastoid B cells. For each eSNP, *Sherlock* will then evaluate its association with gout using genome-wide association (GWA) data of gout. There can be three scenarios: (1) If the eSNP of a specific gene is also associated with gout in GWAS, a positive score would be given; (2) If the eSNP of this gene is not associated with gout, a negative score would be assigned; and (3) association only in GWAS (ie, non-eSNPs) does not alter the score. The total score of a gene increases along with the increase in the number of SNPs with combined evidence. For each gene, *Sherlock* performs a Bayesian inference to test whether the expression change of this gene has any impact on the risk of gout by using the collective information of the putative eSNPs of the gene. Based on the combined evidence from GWAS and lymphoblastoid eQTL, *Sherlock* infers gout-associated genes by calculating the logarithm of the Bayes factor of each gene. Compared with traditional analysis, which usually ignores SNPs with a moderate association (e.g., SNPs with P-values ranging from 1×10^{-6}), *Sherlock* utilizes both strong and moderate SNPs in the eQTL and GWAS data through using a powerful statistical model. *Sherlock* makes the statistical inference by aggregating the information from both strong SNPs and moderate SNPs (strong SNPs have a larger contribution to the final score).

Results

We performed a genome-wide association meta-analysis on two Taiwanese cohorts consisting of 758 gout cases and 14166 controls of Han Chinese ancestry. Characteristics of the study subjects are shown in Supplementary Table 4. After performing a standard quality control procedure, we analyzed 373 individuals with gout (cases) and 6721 controls without gout from Taiwan Biobank in the discovery stage. In the discovery stage, we identified 4 SNPs that showed significant association with gout at the genome-wide level ($P = 5.0 \times 10^{-8}$). (Fig. 1 and Table 1) All of these SNPs are located in previously identified regions on chromosome 4. The only exception that we found was rs2905274 ($P = 3.91 \times 10^{-8}$; OR, 1.87), which was located on chromosome 7. The top-associated SNP in chromosome 4 were rs2231142 ($P = 4.25 \times 10^{-18}$; OR, 2.00) and rs4148155 ($P = 5.49 \times 10^{-18}$; OR, 2.00), which have been mapped to the *ABCG2* gene. We also found that rs2725211 on chromosome 4 was also associated with increased risk of gout ($P = 3.42 \times 10^{-9}$; OR, 1.64), and was located within a genomic region that encodes both the *ABCG2* and *PKD2* gene. The regional association plot showed that all the strongly associated SNPs were confined to regions around *ABCG2* and *PKD2* gene (Fig. 2).

In the follow-up GWAS study using 385 independent gout cases and 7,445 controls, we still observed significant associations for the three SNPs on chromosome 4. However, rs2905274 on chromosome 7 failed to replicate. In the combined analysis of the discovery and follow-up cohorts, we identified significant associations for rs2231142 ($P = 5.06 \times 10^{-35}$; OR, 2.00), rs4148155 ($P = 3.74 \times 10^{-35}$; OR, 2.00), and rs2725211 ($P = 6.88 \times 10^{-17}$; OR, 1.63) in the additive model, without any heterogeneity between the two stages.

Candidate causal SNPs and pathways from the meta-analysis data of GWASs. Utilizing the SNPs p-values from the genome-wide association meta-analysis analysis as input, ICSNPathway analysis identified one candidate causal SNP (rs2231142), one gene (*ABCG2*), and three candidate causal pathways. rs2231142 is not in LD with any SNP, and the candidate causal pathways provide three related hypothetical biological mechanisms of gout: ABC transporters; ATPASE activity coupled; and ATPASE activity coupled to movement of substances (Supplementary Table 5).

SNP	Chr	Allele 1/2 ^a	Stage	Cases				Controls				Additive ^b		Dominant ^b		Recessive ^b		P _{het}
				11	12	22	RAF	11	12	22	RAF	P-value	Risk allele OR (95% CI)	P-value	Risk allele OR (95% CI)	P-value	Risk allele OR (95% CI)	
rs2231142	4	T/G	Discovery	82	181	110	0.46	630	2885	3200	0.31	4.25e-18	2.00	5.27e-11	2.18	2.66e-15	3.00	0.9644
			Follow-up	78	197	109	0.46	696	3202	3538	0.31	1.498e-18	1.99	2.86e-13	2.366	2.96e-12	2.65	
			Combined ^c	160	378	219	0.46	1326	6087	6738	0.31	5.06e-35	2.00	1.10e-22	2.27	6.50e-26	2.82	
rs4148155	4	G/A	Discovery	82	181	110	0.46	633	2889	3195	0.31	5.49e-18	2.00	6.27e-11	2.17	3.16e-15	2.99	0.9893
			Follow-up	79	197	109	0.46	702	3204	3536	0.31	8.593e-19	2.00	2.65e-13	2.37	1.29e-12	2.68	
			Combined ^c	161	378	219	0.46	1335	6093	6731	0.31	3.74e-35	2.00	1.21e-22	2.27	3.29e-26	2.83	
rs2725211	4	T/C	Discovery	48	153	172	0.33	386	2418	3900	0.24	3.42e-09	1.64	8.52e-06	1.63	9.38e-09	2.69	0.9417
			Follow-up	45	168	172	0.34	426	2764	4244	0.24	3.78e-09	1.62	1.22e-06	1.69	9.96e-07	2.35	
			Combined ^c	93	321	344	0.33	812	5182	8144	0.24	6.88e-17	1.63	4.75e-11	1.66	5.43e-14	2.52	
rs2905274	7	A/G	Discovery	12	80	279	0.14	44	1027	5640	0.083	3.91e-08	1.87	2.46e-06	1.84	4.19e-06	5.21	0.0041
			Follow-up	5	71	306	0.11	58	1231	6138	0.091	0.26	1.15	0.34	1.14	0.2829	1.69	
			Combined ^c	17	151	585	0.12	102	2258	11778	0.087	1.46e-06	1.50	4.81e-05	1.46	1.42e-05	3.50	

Table 1. Results of association analyses of gout. We analyzed 758 gout cases (in the GWAS and in replication) and 14,166 controls (6,721 in the GWAS and 7,445 in replication). Chr., chromosome; RAF, risk allele frequency. ^aAllele 1, risk allele; allele 2, non-risk allele. ^bP values and ORs were calculated by logistic regression analysis, with age, gender, and 10 principal components as covariates. Non-risk alleles were considered as references in the three genetic models: additive, 1 versus 2; recessive, 11 versus 12 + 22; dominant, 11 + 12 versus 22. Heterogeneity across the two stages was examined by Cochran Q test under a genetic model which provided the minimum P value in the screening stage. ^cORs and P values were calculated using the Mantel-Haenszel fixed-effects model.

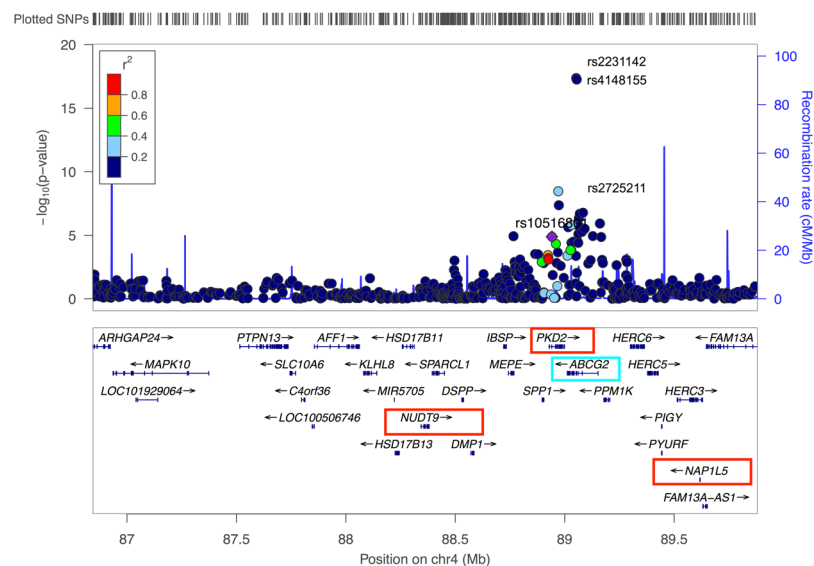


Figure 2. Regional association plot and linkage disequilibrium (LD) on chromosome 4.

Integrative analysis of eQTL and GWAS results using Sherlock. Through systematic integration of lymphoblastoid eQTL and SNP associations from our discovery GWAS analysis, *PKD2* expression showed the most significant association with gout (LBF = 6.89, $P_{\text{sher}} = 1.08 \times 10^{-5}$) followed by *NUTD9*, *NAP1L5*, and *BRE* (Table 2). In the follow-up analysis, we still observed significant associations for all the genes identified by Sherlock, with the only exception for *BRE*. Interestingly, *PKD2*, *NUTD9* and *NAP1L5* are all in the 4q22.1 locus.

One eSNP (rs4148155) showed significant association with both *PKD2* (LBF = 7.02, $P_{\text{eQTL}} = 2.00 \times 10^{-5}$) and *NAP1L5* expression (LBF = 5.62, $P_{\text{eQTL}} = 2.00 \times 10^{-4}$) and strong evidence for association with gout ($P_{\text{GWAS}} = 5.49 \times 10^{-18}$) (Table 2). We also identified a novel cis eSNP (rs4148155), which showed significant association with of *NUTD9* (LBF = 5.62, $P_{\text{eQTL}} = 8.00 \times 10^{-5}$) but moderate evidence for association with gout ($P_{\text{GWAS}} = 1.31 \times 10^{-5}$; $P_{\text{Replication}} = 2.33 \times 10^{-6}$; $P_{\text{combined}} = 1.34 \times 10^{-10}$).

Gene	SNP	Proximity	Location	Result in Discovery		Result in Follow-up	
				LBF	P*	LBF	P
PKD2			4q22.1	6.89	1.08e-05	6.93	8.98e-06
	rs4148155	cis	chr4:89273691	7.02	2.00e-05	7.02	2.00e-05
NUDT9			4q22.1	6.41	2.87e-05	5.79	6.29e-05
	rs10516801	cis		5.62	8.00e-05	6.02	8.00e-05
NAP1L5			4q22.1	5.54	9.52e-05	6.10	4.67e-05
	rs4148155	cis	chr4:89273691	5.62	2.00e-04	5.62	2.00e-04
BRE			2p23.2	5.87	6.11e-05	2.57	5.76e-03
	rs4148155	trans	chr4:89273691	2.66	9.90e-06	2.66	9.90e-06
	rs4129943	trans	chr2:241402488	2.25	2.00e-06	-0.102	2.00e-06

Table 2. Predicted regulatory genes and SNPs for the risk of gout in lymphocytes. *The gene p-value refers to the Sherlock p value, but the SNP p-value refers to the eQTL P value.

Discussion

In this study, we sought to identify novel genetic variations that predisposed individuals to gout among 15,300 Han Chinese residing in Taiwan. From 2 independent cohorts, we found 3 SNPs (rs2231142, rs4148155, and rs2725211) that reached genome-wide statistical significance, and these SNPs are in cis of the *ABCG2* and *PKD2* gene. ICSNPathway analysis identified rs2231142 as the candidate causal SNP, and *ABCG2* located in 4q22.1 as the candidate gene. In order to identify other susceptibility genes exhibiting regulatory function underlying gout, we correlated the signatures of expression data of lymphoblastoid B cell with that of GWASs in gout. We identified three genes, which were significantly associated with the risk of gout (*PKD2*, *NUDT9*, and *NAP1L5*), with *NUDT9*, and *NAP1L5* reported at the first time.

Previous studies have reported polymorphism in *ABCG2* to be associated with gout in several populations, such as, European Americans, African Americans, Mexican Americans, Americans Indians, German, Japanese and Han Chinese^{9,10,23–25}. The rs2231142 (Arg141Lys) genetic variant at *ABCG2* is a common missense genetic variants, and meta-analysis of existing study found the rs2231142 Arg141Lys carriers was associated with 1.73 fold increased susceptibility of gout. It has been reported that the Arg141Lys variant of *ABCG2* causes instability in the nucleotide-binding domain of *ABCG2*, and lead to decreased surface expression and function of *ABCG2*²⁶. As a result, rs2231142 Arg141Lys carriers have decreased uric acid excretion through both the kidney and the gut with the potential for hyperuricemia. Besides leading to hyperuricemia, *ABCG2* dysfunction was also found to be involved in subsequent steps in gout formation. Knock down of *ABCG2* by siRNA led to gouty inflammation involving the release of IL-8 upon MSU crystals-stimulation²³. In addition, in a Taiwanese study of Han Chinese with gout, rs2231142 Arg141Lys carriers were associated with 1.51 fold increased risk of tophi²⁷.

Besides the rs2231142 variant, this study also found that the rs4148155 variant of *ABCG2* was associated with gout. This is likely due to the fact that rs2231142 and rs4148155 are completely in LD in the Han Chinese. The rs4148155 genetic variant was reported to be an intron variant of *ABCG2*, and was also found to be associated with uric acid formation in both Han Chinese and Japanese population^{28,29}. Interestingly, in our Sherlock analysis using lymphocytes, the rs4148155 variant was associated with the eQTL of *PKD2*, and *NAP1L5*. *ABCG2*, *PKD2*, *NAP1L5* are all located in the 4q22.1 region, and we hypothesize a cis acting epistatic interactions between these genes. In fact, a previous study in Han-Chinese found also found a positive correlation between *ABCG2* mRNA expression and *PKD2* mRNA expression³⁰. Currently, the biological mechanism on how *ABCG2* interacts with *PKD2*/*NAP1L5* in the pathogen of gout is unclear. But there are strong clinical and genetic reports linking gout and *PKD2*. *PKD2* encodes Polycystin-2, which is the protein mutated in autosomal dominant polycystic kidney disease (ADPKD)³¹. It is well recognized that patients with ADPKD develop renal failure and progress to hyperuricemia and increased risk of gout^{32,33}. In addition, our Sherlock analysis validated Genecards' report that *PKD2* is expressed in lymphocytes. The role of lymphocytes related to gout development has been well recognized, but it unclear how *ABCG2* interact with *PKD2* in the inflammation stage of gout^{34,35}. Future research into the functional role of *PKD2* in lymphocytes may also help explain why not all ADPKD patients develop gout. As for *NAP1L5*, it has been implicated in IL-8 release, and IL-8 release has been found to be an important activator for monosodium urate crystal-induced arthritis^{36–38}. Interestingly, siRNA knock down of *ABCG2* also increases IL-8 release²³.

Results of this study have to be interpreted in light of several limitations. First, this was a case-control study conducted in a Han Chinese population residing in Taiwan. Future investigations using other populations will be critical to clarify whether these newly identified susceptible genes are shared in other populations. Second, this study focused on only common SNPs and did not consider the contributions of rare variants. Future studies on rare variants should also be conducted to fully understand the role of rare variants in the pathogenesis of gout. Third, this study did not conduct a functional study to identify the causal variant for gout disease, and functional study should be conducted by follow-up studies.

In summary, we have identified several genetic loci related to gout using GWAS in the Han Chinese population in Taiwan. We performed single-marker as well as pathway analyses to identify genetic associations with gout. The rs2231142 (Arg141Lys) genetic variant at *ABCG2* was identified to be the causal SNP, but this SNP was also found to be in complete LD with rs4148155. In addition, we conducted Sherlock analysis to identify susceptibility genes in with regulatory function in gout, and identify three genes, which were significantly associated with the risk of gout (*PKD2*, *NUDT9*, and *NAP1L5*). To conclude, the results of our study may contribute to the

understanding of the genetic causes of gout, and future studies are needed to confirm and explore the role of NUTD9, and NAP1L5 in the pathogenesis of gout.

Data Availability

The data used in this study are available for purchase from Taiwan Biobank. To gain access, interested individuals should contact “biobank@gate.sinica.edu.tw”. The GWAS data will be deposited in the GWAS catalog upon manuscript acceptance by a peer-reviewed journal.

References

- Mikuls, T. R. *et al.* Gout epidemiology: results from the UK General Practice Research Database, 1990–1999. *Ann Rheum Dis* **64**, 267–272, <https://doi.org/10.1136/ard.2004.024091> (2005).
- Lawrence, R. C. *et al.* Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. *Arthritis Rheum* **58**, 26–35, <https://doi.org/10.1002/art.23176> (2008).
- Helmick, C. G. *et al.* Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part I. *Arthritis Rheum* **58**, 15–25, <https://doi.org/10.1002/art.23177> (2008).
- Kuo, C. F., Grainge, M. J., Mallen, C., Zhang, W. & Doherty, M. Rising burden of gout in the UK but continuing suboptimal management: a nationwide population study. *Ann Rheum Dis* **74**, 661–667, <https://doi.org/10.1136/annrheumdis-2013-204463> (2015).
- Campion, E. W., Glynn, R. J. & Delabry, L. O. Asymptomatic hyperuricemia. Risks and consequences in the Normative Aging Study. *The American journal of medicine* **82**, 421–426 (1987).
- Syndenham, T. The Works of Thomas Syndenham, MD on Acute and Chronic Diseases. Vol. II. *GJ & J. Robinson* (1853).
- Köttgen, A. *et al.* Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. *Nature genetics* **45**, 145 (2013).
- Cadzow, M., Merriman, T. R. & Dalbeth, N. Performance of gout definitions for genetic epidemiological studies: analysis of UK Biobank. *Arthritis Res Ther* **19**, 181, <https://doi.org/10.1186/s13075-017-1390-1> (2017).
- Ly, X. *et al.* The association between the polymorphism rs2231142 in the ABCG2 gene and gout risk: a meta-analysis. *Clinical rheumatology* **33**, 1801–1805 (2014).
- Dehghan, A. *et al.* Association of three genetic loci with uric acid concentration and risk of gout: a genome-wide association study. *Lancet* **372**, 1953–1961, [https://doi.org/10.1016/S0140-6736\(08\)61343-4](https://doi.org/10.1016/S0140-6736(08)61343-4) (2008).
- Sulem, P. *et al.* Identification of low-frequency variants associated with gout and serum uric acid levels. *Nature genetics* **43**, 1127 (2011).
- Justice, A. E. *et al.* In *BMC proceedings*. 56 (BioMed Central).
- Li, C. *et al.* Genome-wide association analysis identifies three new risk loci for gout arthritis in Han Chinese. *Nat Commun* **6**, 7041, <https://doi.org/10.1038/ncomms8041> (2015).
- Dong, Z. *et al.* Effects of multiple genetic loci on the pathogenesis from serum urate to gout. *Sci Rep* **7**, 43614, <https://doi.org/10.1038/srep43614> (2017).
- Zhang, K. *et al.* ICSNPathway: identify candidate causal SNPs and pathways from genome-wide association study by one analytical framework. *Nucleic Acids Res* **39**, W437–443, <https://doi.org/10.1093/nar/gkr391> (2011).
- He, X. *et al.* Sherlock: detecting gene-disease associations by matching patterns of expression QTL and GWAS. *Am J Hum Genet* **92**, 667–680, <https://doi.org/10.1016/j.ajhg.2013.03.022> (2013).
- Yin, X. *et al.* Five regulatory genes detected by matching signatures of eQTL and GWAS in psoriasis. *J Dermatol Sci* **76**, 139–142, <https://doi.org/10.1016/j.jdermsci.2014.07.007> (2014).
- Yang, C.-P. *et al.* Comprehensive integrative analyses identify GLT8D1 and CSNK2B as schizophrenia risk genes. *Nature communications* **9**, 838 (2018).
- Fan, C.-T., Lin, J.-C. & Lee, C.-H. Taiwan Biobank: a project aiming to aid Taiwan’s transition into a biomedical island. *Pharmacogenomics* **9**, 235–246, (2008).
- Chen, C. H. *et al.* Population structure of Han Chinese in the modern Taiwanese population based on 10,000 participants in the Taiwan Biobank project. *Hum Mol Genet* **25**, 5321–5331, <https://doi.org/10.1093/hmg/ddw346> (2016).
- Dalbeth, N. *et al.* Survey definitions of gout for epidemiologic studies: comparison with crystal identification as the gold standard. *Arthritis care & research* **68**, 1894–1898 (2016).
- Pruim, R. J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337, <https://doi.org/10.1093/bioinformatics/btq419> (2010).
- Chen, C. J. *et al.* ABCG2 contributes to the development of gout and hyperuricemia in a genome-wide association study. *Sci Rep* **8**, 3137, <https://doi.org/10.1038/s41598-018-21425-7> (2018).
- Woodward, O. M. *et al.* Identification of a urate transporter, ABCG2, with a common functional polymorphism causing gout. *Proc Natl Acad Sci USA* **106**, 10338–10342, <https://doi.org/10.1073/pnas.0901249106> (2009).
- Zhang, L. *et al.* Association of functional polymorphism rs2231142 (Q141K) in the ABCG2 gene with serum uric acid and gout in 4 US populations: the PAGE Study. *Am J Epidemiol* **177**, 923–932, <https://doi.org/10.1093/aje/kws330> (2013).
- Wong, K., Briddon, S. J., Holliday, N. D. & Kerr, I. D. Plasma membrane dynamics and tetrameric organisation of ABCG2 transporters in mammalian cells revealed by single particle imaging techniques. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1863**, 19–29 (2016).
- Tu, H. P. *et al.* Joint effects of alcohol consumption and ABCG2 Q141K on chronic tophaceous gout risk. *J Rheumatol* **41**, 749–758, <https://doi.org/10.3899/jrheum.130870> (2014).
- Kamatani, Y. *et al.* Genome-wide association study of hematological and biochemical traits in a Japanese population. *Nat Genet* **42**, 210–215, <https://doi.org/10.1038/ng.531> (2010).
- Yu, K.-H. *et al.* A comprehensive analysis of the association of common variants of ABCG2 with gout. *Scientific Reports* **7**, 9988 (2017).
- Dong, Z. *et al.* PKD2 influence uric acid levels and gout risk by interacting with ABCG2. *bioRxiv*, 278713 (2018).
- González-Perrett, S. *et al.* Polycystin-2, the protein mutated in autosomal dominant polycystic kidney disease (ADPKD), is a Ca²⁺-permeable nonselective cation channel. *Proceedings of the National Academy of Sciences* **98**, 1182–1187 (2001).
- Mejias, E., Navas, J., Llubes, R. & Martínez-Maldonado, M. Hyperuricemia, gout, and autosomal dominant polycystic kidney disease. *Am J Med Sci* **297**, 145–148 (1989).
- Puig, J. G. *et al.* Hereditary nephropathy associated with hyperuricemia and gout. *Arch Intern Med* **153**, 357–365 (1993).
- Lai, S. & Zhou, X. Inflammatory cells in tissues of gout patients and their correlations with comorbidities. *The open rheumatology journal* **7**, 26 (2013).
- Kanevets, U., Sharma, K., Dresser, K. & Shi, Y. A role of IgM antibodies in monosodium urate crystal formation and associated adjuvanticity. *J Immunol* **182**, 1912–1918, <https://doi.org/10.4049/jimmunol.0803777> (2009).
- Warner, N., Burberry, A., Pliakas, M., McDonald, C. & Nunez, G. A genome-wide small interfering RNA (siRNA) screen reveals nuclear factor-kappaB (NF-kappaB)-independent regulators of NOD2-induced interleukin-8 (IL-8) secretion. *J Biol Chem* **289**, 28213–28224, <https://doi.org/10.1074/jbc.M114.574756> (2014).

37. Nishimura, A. *et al.* Attenuation of monosodium urate crystal-induced arthritis in rabbits by a neutralizing antibody against interleukin-8. *J Leukoc Biol* **62**, 444–449 (1997).
38. Matsukawa, A. *et al.* Analysis of the cytokine network among tumor necrosis factor alpha, interleukin-1beta, interleukin-8, and interleukin-1 receptor antagonist in monosodium urate crystal-induced rabbit arthritis. *Laboratory investigation; a journal of technical methods and pathology* **78**, 559–569 (1998).

Acknowledgements

This study was funded by the National Taiwan University and National Taiwan University Hospital Collaboration Grant NTUH 106-050; National Taiwan University Hospital Grant NTUH.106-P04; Taiwan National Ministry of Science and Technology Grants MOST 104-2314-B-002 -039 -MY3, and MOST 106-2811-B-002-048. No funding bodies had any role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

All investigators contributed to the design or execution of the study, and approved the final version. C.-C.L. designed the study, obtained funding, drafted the analytical plan, guided the statistical analysis, interpreted the data, and critically revised the manuscript. Shyr-Chyr Chen and Po-Hsiu Kuo was involved in the grant proposal writing. Ya-Chin Lee performed most of the quality control analysis. M.-T.L. was involved in the design of study and analytical plan and interpretation of the results. M.-T.L. and T.-C.H. performed the statistical analysis. Shyr-Chyr Chen, Po-Hsiu Kuo, Jenn-Hwai Yang, and Hsiu-Hao Chang reviewed the manuscript and provided insights in the contents.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-41434-4>.

Competing Interests: The authors declare 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) 2019