Lack of association of genetic variants for diabetic retinopathy in Taiwanese patients with diabetic nephropathy

Ai-Ru Hsieh,[1] Yu-Chuen Huang,[2,3] Ya-Fei Yang,[4] Hui-Ju Lin,[2,5] Jane-Ming Lin,[5] Ya-Wen Chang,[3] Chia-Ming Wu,[3] Wen-Ling Liao ![ORCID],[6,7] Fuu-Jen Tsai[2,3,8]

For numbered affiliations see end of article.

Correspondence to
Dr Wen-Ling Liao;
wl0129@mail.cmu.edu.tw

ABSTRACT

Objective Diabetic nephropathy (DN) and diabetic retinopathy (DR) comprise major microvascular complications of diabetes that occur with a high concordance rate in patients and are considered to potentially share pathogeneses. In this case-control study, we sought to investigate whether DR-related single nucleotide polymorphisms (SNPs) exert pleiotropic effects on renal function outcomes among patients with diabetes.

Research design and methods A total of 33 DR-related SNPs were identified by replicating published SNPs and via a genome-wide association study. Furthermore, we assessed the cumulative effects by creating a weighted genetic risk score and evaluated the discriminatory and prediction ability of these genetic variants using DN cases according to estimated glomerular filtration rate (eGFR) status along with a cohort with early renal functional decline (ERFD).

Results Multivariate logistic regression models revealed that the DR-related SNPs afforded no individual or cumulative genetic effect on the nephropathy risk, eGFR status or ERFD outcome among patients with type two diabetes in Taiwan.

Conclusion Our findings indicate that larger studies would be necessary to clearly ascertain the effects of individual genetic variants and further investigation is also required to identify other genetic pathways underlying DN.

Significance of this study

What is already known about this subject?
► Diabetic nephropathy (DN) and diabetic retinopathy (DR) comprise major microvascular complications of diabetes that occur with a high concordance rate in patients and are considered to potentially share pathogeneses.
► Genetic loci with pleiotropic effects, whereby one genetic locus can affect more than one phenotype, may underlie the association between DN and DR phenotypes.

What are the new findings?
► We identified a total of 33 DR-related single nucleotide polymorphisms (SNPs) by replicating SNPs on candidate genes previously identified through meta-analyses and genome-wide association study (GWAS) (10 SNPs were replicated) and through GWAS performed on subjects with DR with estimated glomerular filtration rate (eGFR) >60 mL/min/1.73 m^2 and albumin-to-creatinine ratio <300 mg/g (23 SNPs were identified).
► These DR-related SNPs contributed no individual or cumulative genetic effect on the nephropathy risk, eGFR status or early renal functional decline outcome among patients with diabetes in Taiwan, as confirmed by multivariate logistic regression models.

How might these results change the focus of research or clinical practice?
► Further evidence for the genetic pleiotropy of diabetic microvascular complications should be sought using a cohort of a larger size in Taiwan.
► Further investigation is required to identify other genetic pathways underlying DN.

INTRODUCTION

Nephropathy constitutes a frequent and serious complication of diabetes mellitus. Approximately 20%–40% of patients with type two diabetes (T2D) will develop diabetic kidney disease (DKD)[1] and many will further progress to end-stage renal disease (ESRD) via a relentless decline in glomerular filtration rate (GFR).[2] According to Yang *et al*,[3] the increased prevalence of diabetic nephropathy (DN) in Taiwan is the main cause of the increases in prevalence and incidence of ESRD. Notably, analysis of the international statistics collected in the US Renal Data System indicates that Taiwan has the highest incidence and the second highest prevalence of ESRD worldwide.[4 5]

Several risk factors for developing DN have been identified, such as age, duration of diabetes, albuminuria grade, early GFR decline, increased or variability of hemoglobin A1c (HbA1c) and systolic blood pressure (SBP), serum uric acid, presence of concomitant microvascular complications and positive family history.[6] Additionally, hereditary components are considered to

constitute a significant risk factor for DN.[7] Numerous genes, such as *TCF7L2*, *ACE* and *SHROOM3* were identified to be associated with DN through a candidate gene approach, genome-wide association study (GWAS)[8 9] or meta-analysis.[10 11] However, the genetic mechanism of DN remains unclear. Furthermore, research suggests that the presence of diabetic retinopathy (DR) may be considered as a vital clinical biomarker for the diagnosis of DKD among patients with diabetes and microalbuminuria (MAU).[12 13] DN and DR comprise two major microvascular complications of diabetes and occur with a high concordance rate in patients with diabetes. Studies have suggested that these two diabetic complications might share common pathogeneses, such as glucose metabolism, angiogenesis, inflammation and oxidative stress.[14] Therefore, genetic loci with pleiotropic effects, whereby one genetic locus can affect more than one phenotype, may underlie the association between DR and DKD phenotypes.[15] However, to date only a few genetic polymorphisms have been reported to be associated with these two complications, including interleukin (IL)-10[16] and SLC2A1.[17]

The aim of this study is to investigate whether DR-related single nucleotide polymorphisms (SNPs) exert a pleiotropic effect on renal function outcomes among patients with diabetes. The DR-related SNPs were identified by replicating published SNPs and via GWAS among patients with T2D in Taiwan. Furthermore, we assessed their cumulative effects by creating a weighted genetic risk score (wGRS) and evaluated the discriminatory and predictive ability of these genetic variants using DN cases according to estimated GFR (eGFR) status along with a cohort with early renal functional decline (ERFD) in the Taiwanese population.

MATERIALS AND METHODS

The present study was broadly divided into three stages: stage I, DR-related genetic marker selection, selection of genetic markers related to DR via the GWAS method and replication of reported SNPs among 206 DR cases and 206 non-DR controls; stage II, verification of DR-related SNPs by using DN cases, investigation of the effect of DR-related genes in diabetes by using 567 DN cases (eGFR $<60\,\text{mL/min/1.73\,m}^2$) and 909 non-DN controls (eGFR $>90\,\text{mL/min/1.73\,m}^2$) and stage III: verification of DR-related SNPs on diabetic cohort, verification of DR-related SNPs by using 417 subjects with ERFD (eGFR decline $>3.3\,\text{mL/year}$) and 417 subjects without ERFD. No overlap was allowed for the subjects among the three stages. Detailed information for study population selection for each stage is described in the 'Study population' section. An overall flow chart is shown in figure 1.

Study population

Subjects with T2D have been enrolled at the China Medical University Hospital since 2010, and were recruited from the Taiwan Biobank. Detailed information
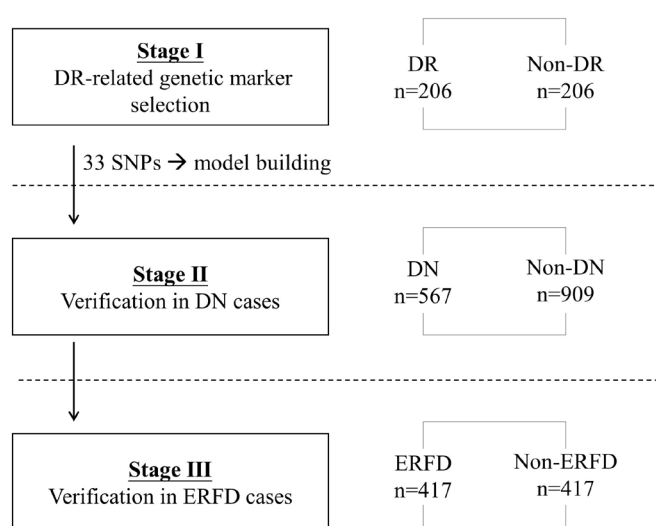

Figure 1 Flow chart for three stages of study design. DN, diabetic nephropathy; DR, diabetic retinopathy; ERFD, early renal functional decline; SNP, single nucleotide polymorphism.

about the number of subjects from different sources at each stage is provided in online supplementary table S1. Informed consent was obtained from all participants. The study was conducted in accordance with the Declaration of Helsinki.

DR cases for genetic marker selection

Among these subjects, 1947 subjects without retinopathy (non-DR group) and 245 subjects with retinopathy (DR group) exhibited eGFR $>60\,\text{mL/min/1.73\,m}^2$ and albumin-to-creatinine ratio (ACR) $<300\,\text{mg/g}$. The clinical information was collected via questionnaires, while the retinopathy status was self-reported. After pairwise matching according to HbA1c and disease duration, 206 DR and 206 non-DR cases were used for the selection of DR-related SNPs. The flow chart for patient selection is shown in online supplementary figure S1(a).

DN cases for replication

We used a cross-sectional study design to investigate the effect of DR-related genes on DN. After excluding 412 subjects with DR for marker selection, a total of 1476 subjects with T2D, including 567 subjects with eGFR $<60\,\text{mL/min/1.73\,m}^2$ and 909 subjects with eGFR $>90\,\text{mL/min/1.73\,m}^2$, were selected from our database. None of these subjects exhibited DR.

Cohort cases with rapid renal function decline for replication

We used a nested case-control study design to investigate the effect of DR-related genes on ERFD. A total of 950 subjects with T2D had available follow-up information from our database. After 1:1 matching according to gender, disease duration, HbA1c and eGFR, two groups of 417 subjects each were identified who exhibited or did not exhibit ERFD during the follow-up period (1.24±1.10 and 2.86±2.16 years of follow-up for ERFD and non-ERFD groups, respectively). ERFD was defined

as >3.3 mL/min/1.73 m^2 decline in eGFR per year.[18] The Modified Diet in Renal Disease equation[19] was used to estimate eGFR. The flow chart for patient selection is shown in online supplementary figure S1(b).

Data collection

Data regarding age, gender, age at T2D diagnosis and ocular history were collected from questionnaires. For each patient, SBP, diastolic blood pressure (DBP) and body mass index (BMI) were determined, and blood samples were collected by venipuncture for genomic DNA isolation and serological tests, including fasting glucose and HbA1c, at the time of enrollment in the study.

Genotyping, imputation, quality control and population substructure

Genomic DNA extracted from peripheral blood leukocytes with the Genomic DNA kit (Qiagen, California, USA) was genotyped using various platforms including Illumina HumanHap550-Duo BeadChips, Affymetrix Axiom genome-wide CHB array, and Affymetrix Axiom genome-wide TWB array, according to standard quality control procedures. Detailed information of the array used at each stage is provided in online supplementary table S1. Because the GWAS results were obtained using different genotyping platforms, genotype imputations were done separately in the three platforms and were performed according to a three-step genotype imputation approach. First, we used SHAPEIT[20] to prephase the study genotypes into full haplotypes. Second, we performed imputation using IMPUTE2 (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html) and the phase I 1000 Genomes Project reference panel (June 2011 interim release) consisting of 1094 phased individuals from multiple ancestry groups (The 1000 Genomes Project Consortium, 2010). Third, we used GTOOL software (http://www.well.ox.ac.uk/~cfreeman/software/gwas/gtool.html) to homogenize strand annotation by merging the imputed results obtained from each set of genotyped data.

Genotype and imputed genotype data were quality controlled (QC), and SNPs were excluded from further analysis if (1) only one allele appeared in cases and controls; (2) the total call rate was <95% for both cases and controls; (3) the minor allele frequency was <0.05 in the controls (the Han Chinese population); (4) they significantly departed from Hardy-Weinberg equilibrium proportions (p<0.05); (5) they had low imputation quality (info<0.4)[21] and (6) excessive identity by descent (ϖ>0.1875), which represented first-degree or second-degree relatives.

Population substructure was evaluated in the merged dataset using genotyped SNPs that passed QC with reference populations from the 1000 Genomes Project and based on our own samples by multidimensional scaling (MDS) analysis,[22–24] using the PLINK module. The MDS method detects meaningful underlying dimensions that
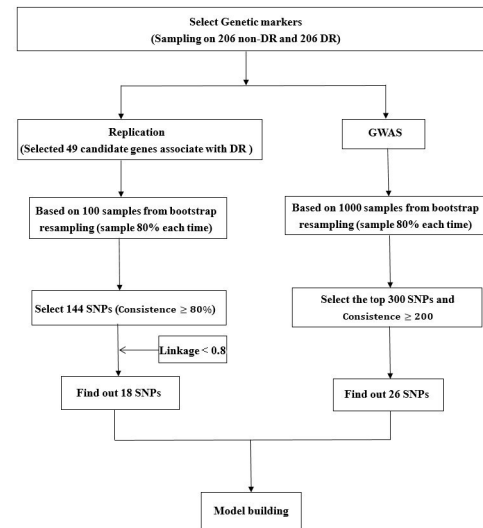

Figure 2 Flow chart for selection of genetic markers. DR, diabetic retinopathy; GWAS, genome-wide association study; SNP, single nucleotide polymorphism.

explain observed genetic distance. During the MDS of matrices, the graphs of mutual arrangement of studied populations in the two-dimensional space were obtained (see online supplementary figure S2 (a)(b)). Also, population outliers were evaluated by visual examination of MDS plots. And the components were evaluated based on the scree plot (see online supplementary figure S2 (c)). The first 10 MDS components which calculated based on our own samples were included as covariates in association test models to minimize spurious associations and maximize the detection power to identify genuine associations.

Genetic marker selection

The genetic markers related to DR status were selected using two methods: GWAS and a candidate gene approach to replicate reported SNPs. Detailed information is described in the following sections "Genome-wide association study" and "Replication", and the flow chart for genetic marker selection is presented in figure 2.

Genome-wide association study

To obtain a robust prioritization of SNPs for use in predictive models, we used the bootstrapping method[25] in GWAS. We resampled the data 1000 times, producing a SNP ranking for each bootstrap sampling based on p values from an additive model adjusting for the first 10 MDS components using PLINK 1.9.[22] Each bootstrap sample was generated by randomly selecting 80% of the individuals with replacement. For each bootstrap sample, the top 300 SNPs and those consistently present in 20% of bootstrapping results were identified. A total of 26 SNPs identified by GWAS were enrolled into model building. The Manhattan plot (see online supplementary figure S3) and quantile-quantile plot (see online supplementary figure S4) based on 206 DR cases and 206 non-DR controls are presented as online supplementary materials.

There was no indication of inflation of test statistics, with the lambda GC value being equal to 1.00713.

Replication

A total of 49 genetic loci (see online supplementary table S2) that were reported from published papers were replicated in the 206 DR cases and 206 non-DR controls mentioned in the 'DR cases for genetic marker selection' section. For the un-typed SNPs, imputation was performed, and the genotype data were quality controlled as described above. A 100 bootstrapping subsampling method was also applied. The significant SNPs (p values <0.05 under the additive inherited genetic model) were ranked in each bootstrapping subsampling and 144 SNPs located on 11 loci consistently present in >80% of bootstrapping results were selected. After removing highly correlated SNPs (D′>0.8), 18 SNPs were enrolled into model building.

DR-related GRS calculation and prediction model building

The wGRS in the prediction model was calculated among the DR cases. Among 44 SNPs (26 SNPs from GWAS and 18 from replication), 33 SNPs (23 SNPs from GWAS and 10 from replication) remained significant in the logistic regression model (with backward selection) after adjusting for low-density lipoprotein (LDL) and *ACR*. The genotype information for 33 SNP is shown in online supplementary table S3. These 33 SNPs were used to calculate the wGRS. Detailed description of the wGRS calculation has been reported previously[26 27]; the formula is as follows: wGRS =33/77.189×((rs767763×2.025)+(rs16958803×1.992)+(rs4129423×2.036)+(rs1559438×1.287)+(rs62324351×1.926)+(rs3791242×2.587)+(rs77625440×2.289)+(rs56170305×1.769)+(rs10055994×5.019)+(rs28610956×1.457)+(rs12076129×2.795)+(rs4665299×1.950)+(rs6433562×2.253)+(rs12991409×2.191)+(rs11889778×2.160)+(rs75759133×1.362)+(rs7374667×1.665)+(rs34766496 ×2.453)+(rs200796238×1.809)+(rs62328468×2.080)+(rs6841985×1.966)+(rs6554985×3.221)+(rs35019626×1.931)+(rs60421526×3.803)+(rs73357792×2.014)+(rs12680033×2.463)+(rs11318592×1.110)+(rs4618795×1.990)+(rs7940618×1.870)+(rs1263663×3.007)+(rs75631519×4.166)+(rs1894151×5.049)+(rs6065597×1.494)). Furthermore, the DR-related wGRS was calculated for each individual among the DN cases and the ERFD cohort. For investigating the cumulative effect of DR-related wGRS on DN complications, subjects were divided into three groups according to the distribution of wGRS in DN cases and ERFD cohort.

Statistical analysis

Continuous data are presented as means with SD, and categorical data are presented as proportions. We used t-test to compare mean values of continuous variables, and χ^2 test to compare the frequencies of categorical variables between two groups. We estimated ORs and 95% CIs of variables by using logistic regression to examine the independent association between GRS and the DN or ERFD end point. We used three steps to select independent

variables that result in a 'best' model. First, we conducted a univariate analysis of each variable. Second, we selected variables with p<0.05 as a candidate in the multivariate model. Third, we constructed a multivariate model with the candidate variables by using the backward selection method. Receiver operating characteristic (ROC) curves were generated to quantify the predictive accuracy of the models, and the area under the curve (AUC) was used to assess the discriminatory ability of the models. *The statistical significance of the difference* between the AUC values was determined using Z statistics.[28] All statistical analyses were performed using SPSS software V.21.0 for Windows (IBM, Armonk, New York, USA) or R V.3.4.4 (R Core Team, 2018). Additionally, to minimize the mean of paired distance on matching variables for each case and control pair, we performed 'pairmatch' in R, which retained most of the cases in our database and matched controls.[29]

RESULTS

DR-related genetic marker selection

We identified a total of 18 DR-related SNPs by replicating published SNPs and 26 through GWAS by using the matched 206 DR cases and 206 non-DR controls. The demographic information of subjects with DR is shown in online supplementary table S4. Then, we used a logistic regression model with backward selection to select the SNPs for calculating the wGRS after adjusting for LDL and ACR. A total of 33 SNPs presented a p value <0.1 and were included in GRS. The genotype information for each SNP is shown in online supplementary table S3. The mean number of risk alleles was 35.87±4.44 (range 24–47), and the mean wGRS was 39.05±4.30 (range 27.66–48.63). The distribution of risk alleles and wGRS is shown in online supplementary figure S5.

Verification of DR-related SNPs using DN cases

We enrolled a total of 1476 subjects with T2D including 567 with eGFR <60 mL/min/1.73 m² (as cases) and 909 with eGFR >90 mL/min/1.73 m² (as controls). The male-to-female ratios were 1.01 (457/452) for the control group and 1.26 (316/251) for the case group (p=0.041 by Pearson's χ^2 test). DN cases were significantly older, with long diabetes duration, having higher BMI, SBP and ACR values and lower HbA1c, DBP, high-density lipoprotein (HDL) and LDL compared with those of the controls (p=0.016, 0.002 and <0.001 for BMI, HbA1c and all other parameters, respectively; table 1). The individual effect of 33 DR-SNPs on eGFR status is shown in online supplementary table S5.

To determine the impact of the cumulative effect of DR-related SNPs on DN, a logistic regression model was used. wGRSs were divided into three groups based on the number of risk alleles. (The distribution of risk alleles and wGRS is shown in online supplementary figure S6(a)). All significant covariates in the univariate model and wGRS (forced in) were retained in the model. With

Table 1 Demographics of patients with diabetes by eGFR

	eGFR* >90 (n=909)	eGFR <60 (n=567)	P value†
Gender			
Male	457 (50.3%)	316 (55.7%)	
Female	452 (49.7%)	251 (44.3%)	0.041‡
Age (years)	54.62 (10.22)	69.31 (10.47)	<0.001§
DM duration (years)	5.90 (5.69)	10.02 (8.80)	<0.001§
BMI (kg/m^2)	25.48 (4.32)	26.06 (3.99)	0.016‡
HbA1c (%)	7.71 (1.54)	7.45 (1.52)	0.002‡
SBP (mm Hg)	127.61 (16.55)	135.68 (18.66)	<0.001§
DBP (mm Hg)	77.62 (11.05)	75.45 (12.34)	0.001‡
HDL (mg/dL)	49.58 (13.46)	46.64 (13.71)	<0.001§
LDL (mg/dL)	114.35 (34.48)	106.80 (37.10)	<0.001§
eGFR (mL/min/1.73 m^2)	115.17 (24.56)	40.26 (15.08)	0.001‡
Urine creatinine (mg/dL)	129.45 (220.83)	110.50 (64.27)	0.063
ACR (mg/g)	67.70 (307.72)	578.03 (1619.15)	<0.001§

Values are presented as N (%) or mean±SD.
*The Modified Diet in Renal Disease equation was used to estimate eGFR.
†P value for χ^2 test or two-sample independent t-test.
‡Represents a p value <0.05.
§Represents a p value <0.001.
ACR, albumin-to-creatinine ratio; BMI, body mass index; DBP, diastolic blood pressure; DM, diabetes mellitus; eGFR, estimated glomerular filtration rate; HbA1c, hemoglobin A1c; HDL, high-density lipoprotein; LDL, low-density lipoprotein; SBP, systolic blood pressure.

a backward selection, age, BMI, HDL, LDL, ACR and wGRS were selected into the final model. Compared with individuals in the lowest range of wGRS, the ORs with 95% CIs for those in the middle and high range were 1.65 (1.01 to 2.72) and 1.20 (0.72 to 2.01), respectively. These results suggest that the 33 DR-related SNPs exert no cumulative effect on the eGFR status among patients with T2D (table 2). To assess the discriminatory ability of the models, the AUC was calculated as 0.88 (95% CI 0.85 to 0.91). However, no significant increase in the AUC value was observed after adding genetic factors (wGRS) (Figure 3A).

Verification of DR-related SNPs using ERFD cases

A total of 950 subjects with T2D had follow-up information. After 1:1 matching based on gender,

Table 2 The cumulative effect of DR-related SNPs on patients with diabetes with different eGFR status in the multivariate logistic regression model

	eGFR* >90 N (%)	eGFR <60 N (%)	OR	95% CI	P value
Age (years)			1.18	1.15 to 1.21	<0.001†
BMI (kg/m^2)			1.10	1.04 to 1.16	0.001‡
HDL (mg/dL)			0.98	0.97 to 1.00	0.050
LDL (mg/dL)			1.00	0.99 to 1.00	0.135
ACR (mg/g)			1.001	1.001 to 1.002	<0.001†
Genetic risk score					
Q1 (<37.74)	251 (34.5)	147 (31.3)	Ref.	Ref.	Ref.
Q2 (37.74 to 40.42)	226 (31.1)	174 (37.1)	1.65	1.01 to 2.72	0.048‡
Q3 (>40.42)	250 (34.4)	148 (31.6)	1.20	0.72 to 2.01	0.483

The genetic risk score was calculated based on 33 SNPs listed in online supplementary table S2.
*The Modified Diet in Renal Disease equation was used to estimate eGFR.
†Represents a p value <0.001.
‡Represents a p value <0.05.
ACR, albumin-to-creatinine ratio; BMI, body mass index; DR, diabetic retinopathy; eGFR, estimated glomerular filtration rate; HDL, high-density lipoprotein; LDL, low-density lipoprotein; Q, quantile; Ref, reference; SNP, single nucleotide polymorphisms.
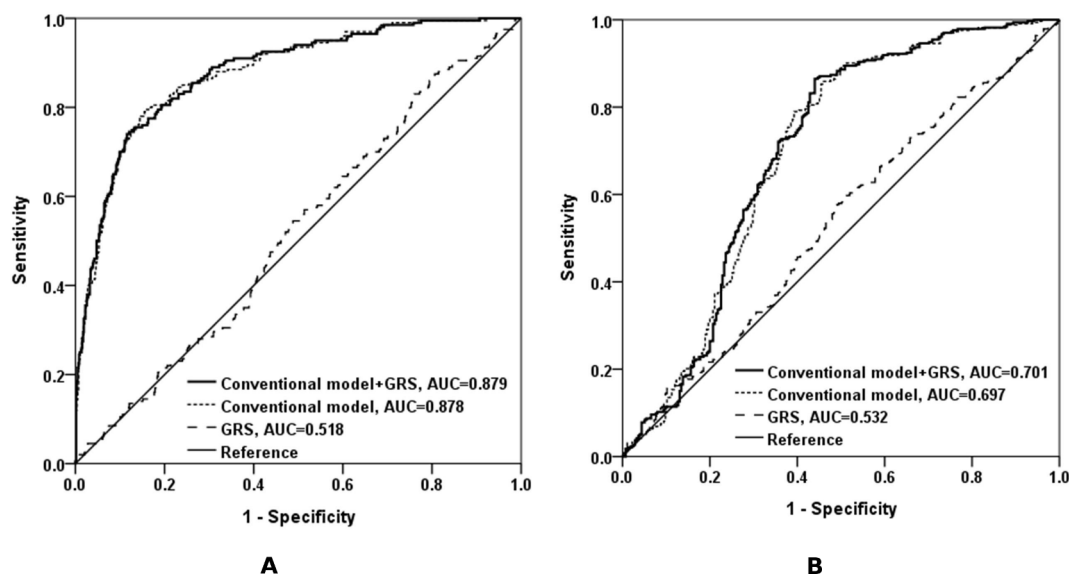
Figure 3 Receiver operating characteristic (ROC) curve and area under the curve (AUC) in the patients with diabetes. ROC curves and AUC of the models built for estimated glomerular filtration rate status (A) and early renal functional decline outcome (B) (AUC=0.879 and 0.701, respectively). The diagonal line indicates zero predictive value of the model. GRS, genetic risk score.

disease duration, HbA1c and eGFR at baseline, 417 exhibited and 417 did not exhibit ERFD during the follow-up period (1.24±1.10 and 2.86±2.16 years of follow-up for ERFD and non-ERFD groups, respectively). However, on matching according to disease duration, the subjects in the ERFD group were older than those in the non-ERFD group (mean (SD) age

of 57.99 (9.99) vs 56.15 (10.58), respectively). Moreover, SBP and DBP were significantly higher, and LDL was lower in the ERFD group when compared with those in the non-ERFD group (p<0.001 for all parameters; table 3). The individual effect of 33 DR-SNPs on ERFD status is shown in online supplementary table S6.

Table 3 Demographics of patients with diabetes by ERFD* status

	Non-ERFD (n=417)	ERFD (n=417)	P value†
Gender			
Male	258 (61.9%)	257 (61.6%)	
Female	159 (38.1%)	160 (38.4%)	0.943
Age (years)	56.15 (10.58)	57.99 (9.99)	0.010‡
DM duration (years)	8.32 (7.23)	8.97 (6.39)	0.225
Follow-up duration (years)	2.86 (2.16)	1.24 (1.10)	<0.001§
BMI (kg/m^2)	26.04 (3.67)	26.22 (3.86)	0.501
HbA1c (%)	7.32 (1.53)	7.48 (1.31)	0.110
SBP (mm Hg)	129.58 (19.56)	135.39 (18.33)	<0.001§
DBP (mm Hg)	79.08 (13.00)	82.91 (11.60)	<0.001§
HDL (mg/dL)	49.88 (12.86)	48.09 (13.93)	0.054
LDL (mg/dL)	115.63 (34.70)	105.16 (32.72)	<0.001§
eGFR (mL/min/1.73 m^2)	88.70 (28.09)	90.52 (28.20)	0.352
Urine creatinine (mg/dL)	141.68 (91.63)	120.52 (72.73)	0.001‡
ACR (mg/g)	96.12 (353.09)	157.47 (472.41)	0.046*

Values are presented as N (%) or mean±SD.
*Early renal function decline was defined as >3.3 mL/min/1.73 m^2 decline in the eGFR per year.
†P value for χ^2 test or two independent t-tests.
‡Represents a p value <0.05.
§Represents a p value <0.001.
ACR, albumin-to-creatinine ratio; BMI, body mass index; DBP, diastolic blood pressure; DM, diabetes mellitus; eGFR, estimated glomerular filtration rate; ERFD, early renal functional decline; HbA1c, hemoglobin A1c; HDL, high-density lipoprotein; LDL, low-density lipoprotein; SBP, systolic blood pressure.

Table 4 The cumulative effect of DR-related SNPs on diabetic patients with ERFD* outcomes in the multivariate logistic regression model

	Non-ERFD N (%)	ERFD N (%)	OR	95% CI	P value
Follow-up duration (years)			0.59	0.51 to 0.67	<0.001[†]
Urine creatinine (mg/dL)			0.998	0.996 to 1.000	0.074
Genetic risk score					
Q1 (<37.58)	111 (32.6)	117 (34.0)	Ref.	Ref.	Ref.
Q2 (37.58 to 40.23)	105 (30.8)	124 (36.0)	1.16	0.76 to 1.77	0.502
Q3 (>40.23)	125 (36.7)	103 (29.9)	0.89	0.58 to 1.37	0.597

Values are presented as N (%) or mean±SD.
The genetic risk score was calculated based on 33 SNPs listed in online supplementary table S2.
*Early renal function decline was defined as >3.3 mL/min/1.73 m^2 decline in the eGFR per year.
†Represents a p value <0.001.
DR, diabetic retinopathy; eGFR, estimated glomerular filtration rate; ERFD, early renal functional decline; Q, quantile; Ref, reference; SNP, single nucleotide polymorphisms.

To determine the impact of the cumulative effect of DR-related SNPs on ERFD, a logistic regression model was used. wGRSs were divided into three groups based on the number of risk alleles among subjects. (The distribution of risk alleles and wGRS is shown in online supplementary figure S6(b).) All significant covariates in the univariate model and GRS (force in) were included in the model. With backward selection, follow-up time, urine creatinine and wGRS were selected into the final model. Compared with individuals in the lowest range of wGRS, the ORs with 95% CIs for those in the middle and high range were 1.16 (0.76 to 1.77) and 0.89 (0.58 to 1.37), respectively. These results suggested that the 33 DR-related SNPs afforded no cumulative effect on the ERFD risk (table 4). The AUC value for the discriminatory ability of the models was 0.70 (95% CI 0.6 to 0.75). Additionally, no significant increase in the AUC value was observed after adding genetic factors (wGRS) (figure 3B).

DISCUSSION

Here, we investigated whether DR-related SNPs affect renal function among patients with diabetes. We identified a total of 33 DR-related SNPs by replicating SNPs on candidate genes previously identified through meta-analyses and GWAS (10 SNPs were replicated) and through GWAS performed on our subjects with DR with eGFR >60 mL/min/1.73 m^2 and ACR <300 mg/g (23 SNPs were identified). The DR-related SNPs contributed no individual or cumulative genetic effect on the nephropathy risk, eGFR status or ERFD outcome among patients with diabetes, as confirmed by multivariate logistic regression models.

Current literature suggests that the presence of one pre-existing microvascular complication (retinopathy or nephropathy) may contribute to the development of another.[30 31] Retinal vascular geometry can independently predict the incidence of renal dysfunction and may be a useful tool to identify individuals at high risk for renal disease early in the course of type one diabetes (T1D).[32] El-Asrar et al[33] reported that patients with T1D and DR were 13.39 times more likely to develop DKD than those without DR. Moreover, Yang et al[12] reported that the urine proteome specific for eye damage could predict chronic renal insufficiency (eGFR, 60 mL/min/1.73 m^2) in a 5.3 years prospective cohort for T2D.

However, only a few genetic reports support possible pleiotropy between DN and DR.[16 17 34 35] Hosseini et al[36] tested previous suggestive signals for DN for association with severe DR, but none of the loci showed a significant association after multiple testing. In the present study, we identified DR-related SNPs from GWAS and performed replication by using patients with DR and MAU among the Han population in Taiwan. In addition, some of the identified genes, such as guanylate cyclase 1 soluble subunit beta 1, signal transducer and activator of transcription-4 and G protein-coupled receptor 35 were reportedly related to renal perfusion and renin release,[37] cell proliferation and differentiation[38–40] and vascular endothelial relaxation.[41] These pathophysiological associations may be related to nephropathy or diabetic complications[42–46]; however, none of the genetic variants significantly correlated with DN risk among patients with T2D in previous studies[47] or the present study. This may be owing to limited statistical power or might suggest the lack of genetic pleiotropy effect among these two tested diabetic complications.

We recognize several limitations that may contribute to the lack of significant findings in the present study. First, a limited number of DR-related SNPs were identified, which may be due to the small sample size for the process of genetic marker selection, although a matched case-control study design was used in the present study to control for potential confounding factors. However, the effect sizes obtained from such matching may differ from the effect sizes in the whole sample. Therefore, we performed an analysis in which we replicated 33 identified DR-SNPs in a larger population (234 DR and 2368

non-DR), from which 3 DR-related SNPs (rs6841985 from GWAS; rs1559438 and rs56170305 from replication) remained significantly related to DR status. The wGRS was calculated based on the effect sizes for the three DR-related SNPs obtained from the larger population. However, these three DR-related SNPs still contributed to no individual (see online supplementary table S7) or cumulative genetic effect on the nephropathy risk, eGFR status or ERFD outcome in the multivariate logistic regression models (see online supplementary table S8–9). Second, the validity of association (with DR) at many of these loci is unconfirmed or lacks robust replication. This may be due to differences among studies regarding the phenotype definition, the method used to identify retinopathy, ethnic population or other confounding factors. Although the self-reported DR status was used in the present study, we confirmed high consistency in our limited database between the self-reported DR status and that reported by an ophthalmologist (n=749; accuracy: 612/749=81.7%). Furthermore, a high validity of self-report of laser treatments of DR, as compared with fundus photography grading, was observed in the EDIC study T1D cohort.[48] Third, inadequate statistical power may have contributed to the lack of significant associations between single loci and eGFR status and ERFD outcome. Dichotomized outcomes (for eGFR and ERFD) were used in the present analyses, which may reduce the power as compared with the analysis of the outcomes as quantitative measures. However, when the continuous quantitative eGFR was considered as the outcome, and the continuous quantitative wGRS was used as predictor in the linear regression models (with the sample size increased to 3100), no significant effect of wGRS was observed on eGFR outcome even after controlling for other potential confounding factors (see online supplementary table S10). Moreover, the logistic regression models based on conventional variables have high discriminatory power with AUC of 0.88 (on eGFR status) and 0.70 (on ERFD risk) in this study, and inclusion of genetic predictor variants (wGRS) did not significantly change the AUCs. Based on these results, there could be either no pleiotropic effect between these two diabetic complications or the effect sizes of DR-related SNPs on eGFR status and ERFD outcomes may be small. Therefore, larger sample sizes are warranted in these two evaluations. The use of modified equations for calculating eGFR based on Asian populations[49] may also be worth considering in future studies. Lastly, the information of medical use, such as antihypertension or statin therapy, was not collected for the subjects during enrollment. This is potentially relevant to blood pressure and lipids levels and might have confounded the results.

CONCLUSIONS

In conclusion, we identified no individual or cumulative genetic effect of the DR-related SNPs on nephropathy risk, eGFR status or ERFD outcome among patients with diabetes in Taiwan using multivariate logistic regression models. Further evidence for the pleiotropy of microvascular complications should be sought using a cohort of a larger size in the near future.

Author affiliations
[1]Department of Statistics, Tamkang University, Taipei, Taiwan
[2]School of Chinese Medicine, China Medical University, Taichung, Taiwan
[3]Human Genetic Center, Department of Medical Research, China Medical University Hospital, Taichung, Taiwan
[4]Kidney Institute and Division of Nephrology, Department of Internal Medicine, China Medical University Hospital, Taichung, Taiwan
[5]Department of Ophthalmology, China Medical University Hospital, Taichung, Taiwan
[6]Graduate Institute of Integrated Medicine, China Medical University, Taichung, Taiwan
[7]Center for Personalized Medicine, China Medical University Hospital, Taichung, Taiwan
[8]Department of Health and Nutrition Biotechnology, Asia University, Taichung, Taiwan

ORCID iD
Wen-Ling Liao http://orcid.org/0000-0003-1223-8844

REFERENCES

1 American Diabetes Association. 9. microvascular complications and foot care. *Diabetes Care* 2015;38:S58–66.
2 Reutens AT. Epidemiology of diabetic kidney disease. *Med Clin North Am* 2013;97:1–18.
3 Yang W-C, Hwang S-J, Taiwan Society of Nephrology. Incidence, prevalence and mortality trends of dialysis end-stage renal disease in Taiwan from 1990 to 2001: the impact of national health insurance. *Nephrol Dial Transplant* 2008;23:3977–82.
4 System URD. *Annual data report*. In: National Institutes of Diabetes DaKD, editor. Bethesda, MD, 2002.

5   System URD. *Annual data report*. In: National Institutes of Diabetes DaKD, editor. Bethesda, MD, 2005.

6   Radcliffe NJ, Seah J-mine, Clarke M, *et al*. Clinical predictive factors in diabetic kidney disease progression. *J Diabetes Investig* 2017;8:6–18.

7   Rizvi S, Raza ST, Mahdi F. Association of genetic variants with diabetic nephropathy. *World J Diabetes* 2014;5:809–16.

8   Regele F, Jelencsics K, Shiffman D, *et al*. Genome-wide studies to identify risk factors for kidney disease with a focus on patients with diabetes. *Nephrol Dial Transplant* 2015;30:iv26–34.

9   Liao L-N, Chen C-C, Wu F-Y, *et al*. Identified single-nucleotide polymorphisms and haplotypes at 16q22.1 increase diabetic nephropathy risk in Han Chinese population. *BMC Genet* 2014;15:113.

10 Tziastoudi M, Stefanidis I, Hadjigeorgiou GM, *et al*. A systematic review and meta-analysis of genetic association studies for the role of inflammation and the immune system in diabetic nephropathy. *Clin Kidney J* 2017;10:293–300.

11 Ahmad N, Jamal R, Shah SA, *et al*. Renin–Angiotensin–Aldosterone system gene polymorphisms and type 2 diabetic nephropathy in Asian populations: an updated meta-analysis. *Curr Diabetes Rev* 2019;15:263–76.

12 Yang J-K, Wang Y-Y, Liu C, *et al*. Urine proteome specific for eye damage can predict kidney damage in patients with type 2 diabetes: a case-control and a 5.3-Year prospective cohort study. *Diabetes Care* 2017;40:253–60.

13 Penno G, Solini A, Zoppini G, *et al*. Rate and determinants of association between advanced retinopathy and chronic kidney disease in patients with type 2 diabetes: the renal insufficiency and cardiovascular events (RIACE) Italian multicenter study. *Diabetes Care* 2012;35:2317–23.

14 Wei L, Xiao Y, Li L, *et al*. The susceptibility genes in diabetic nephropathy. *Kidney Diseases* 2018;4:226–37.

15 Hu C, Zhang R, Yu W, *et al*. *CPVL/CHN2* genetic variant is associated with diabetic retinopathy in Chinese type 2 diabetic patients. *Diabetes* 2011;60:3085–9.

16 Shu Y, Chen Y, Luo H, *et al*. The roles of IL-10 gene polymorphisms in diabetes mellitus and their associated complications: a meta-analysis. *Horm Metab Res* 2018;50:811–5.

17 Siokas V, Fotiadou A, Dardiotis E, *et al*. SLC2A1 tag SNPs in Greek patients with diabetic retinopathy and nephropathy. *Ophthalmic Res* 2019;61:26–35.

18 Lindeman RD, Tobin J, Shock NW. Longitudinal studies on the rate of decline in renal function with age. *J Am Geriatr Soc* 1985;33:278–85.

19 Levey AS, Bosch JP, Lewis JB, *et al*. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. modification of diet in renal disease Study Group. *Ann Intern Med* 1999;130:461–70.

20 O'Connell J, Gurdasani D, Delaneau O, *et al*. A general approach for haplotype phasing across the full spectrum of relatedness. *PLoS Genet* 2014;10:e1004234.

21 Chung J, Wang X, Maruyama T, *et al*. Genome-wide association study of Alzheimer's disease endophenotypes at prediagnosis stages. *Alzheimer's & Dementia* 2018;14:623–33.

22 Purcell S, Neale B, Todd-Brown K, *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559–75.

23 Li Q, Yu K. Improved correction for population stratification in genome-wide association studies by identifying hidden population structures. *Genet Epidemiol* 2008;32:215–26.

24 Marees AT, de Kluiver H, Stringer S, *et al*. A tutorial on conducting genome-wide association studies: quality control and statistical analysis. *Int J Methods Psychiatr Res* 2018;27:e1608.

25 Manor O, Segal E. Predicting disease risk using bootstrap ranking and classification algorithms. *PLoS Comput Biol* 2013;9:e1003200.

26 Ding K, Bailey KR, Kullo IJ. Genotype-informed estimation of risk of coronary heart disease based on genome-wide association data linked to the electronic medical record. *BMC Cardiovasc Disord* 2011;11:66.

27 Liao W-L, Lin J-M, Chen W-L, *et al*. Multilocus genetic risk score for diabetic retinopathy in the Han Chinese population of Taiwan. *Sci Rep* 2018;8:14535.

28 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–45.

29 Hansen BB, Klopfer SO. Optimal full matching and related designs via network flows. *J Comput Graph Stat* 2006;15:609–27.

30 He F, Xia X, Wu XF, *et al*. Diabetic retinopathy in predicting diabetic nephropathy in patients with type 2 diabetes and renal disease: a meta-analysis. *Diabetologia* 2013;56:457–66.

31 Chen S-Y, Hsu Y-M, Lin Y-J, *et al*. Current concepts regarding developmental mechanisms in diabetic retinopathy in Taiwan. *BioMedicine* 2016;6:8.

32 Benitez-Aguirre PZ, Sasongko MB, Craig ME, *et al*. Retinal vascular geometry predicts incident renal dysfunction in young people with type 1 diabetes. *Diabetes Care* 2012;35:599–604.

33 El-Asrar AMA, Al-Rubeaan KA, Al-Amro SA, *et al*. Retinopathy as a predictor of other diabetic complications. *Int Ophthalmol* 2001;24:1–11.

34 Abhary S, Hewitt AW, Burdon KP, *et al*. A systematic meta-analysis of genetic association studies for diabetic retinopathy. *Diabetes* 2009;58:2137–47.

35 Mooyaart AL, Valk EJJ, van Es LA, *et al*. Genetic associations in diabetic nephropathy: a meta-analysis. *Diabetologia* 2011;54:544–53.

36 Hosseini SM, Boright AP, Sun L, *et al*. The association of previously reported polymorphisms for microvascular complications in a meta-analysis of diabetic retinopathy. *Hum Genet* 2015;134:247–57.

37 Theilig F, Bostanjoglo M, Pavenstädt H, *et al*. Cellular distribution and function of soluble guanylyl cyclase in rat kidney and liver. *J Am Soc Nephrol* 2001;12:2209–20.

38 Marrero MB, Banes-Berceli AK, Stern DM, *et al*. Role of the JAK/STAT signaling pathway in diabetic nephropathy. *Am J Physiol Renal Physiol* 2006;290:F762–8.

39 Zheng Z, Chen H, Zhao H, *et al*. Inhibition of JAK2/STAT3-mediated VEGF upregulation under high glucose conditions by PEDF through a mitochondrial ROS pathway in vitro. *Invest Ophthalmol Vis Sci* 2010;51:64–71.

40 Lee H-P, Chen P-C, Wang S-W, *et al*. Plumbagin suppresses endothelial progenitor cell-related angiogenesis in vitro and in vivo. *J Funct Foods* 2019;52:537–44.

41 Wang Y, Liu H, McKenzie G, *et al*. Kynurenine is an endothelium-derived relaxing factor produced during inflammation. *Nat Med* 2010;16:279–85.

42 Yi J, Fang X, Wan Y, *et al*. Stat4 polymorphisms and diabetes risk: a meta-analysis with 18931 patients and 23833 controls. *Int J Clin Exp Med* 2015;8:3566–72.

43 Cheng S, Rhee EP, Larson MG, *et al*. Metabolite profiling identifies pathways associated with metabolic risk in humans. *Circulation* 2012;125:2222–31.

44 Sallée M, Dou L, Cerini C, *et al*. The aryl hydrocarbon receptor-activating effect of uremic toxins from tryptophan metabolism: a new concept to understand cardiovascular complications of chronic kidney disease. *Toxins* 2014;6:934–49.

45 Rhee EP, Clish CB, Ghorbani A, *et al*. A combined epidemiologic and metabolomic approach improves CKD prediction. *J Am Soc Nephrol* 2013;24:1330–8.

46 Kalim S, Rhee EP. An overview of renal metabolomics. *Kidney Int* 2017;91:61–9.

47 Huda N, Hosen MI, Yasmin T, *et al*. Genetic variation of the transcription factor GATA3, not STAT4, is associated with the risk of type 2 diabetes in the Bangladeshi population. *PLoS One* 2018;13:e0198507.

48 Grassi MA, Sun W, Gangaputra S, *et al*. Validity of self-report in type 1 diabetic subjects for laser treatment of retinopathy. *Ophthalmology* 2013;120:2580–6.

49 Ji H, Zhang H, Xiong J, *et al*. eGFRs from Asian-modified CKD-EPI and Chinese-modified CKD-EPI equations were associated better with hypertensive target organ damage in the community-dwelling elderly Chinese: the Northern Shanghai study. *Clin Interv Aging* 2017;12:1297–308.