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

Ai-Ru Hsieh, Yu-Chuen Huang, Ya-Fei Yang, Hui-Ju Lin, Jane-Ming Lin, Ya-Wen Chang, Chia-Ming Wu, Wen-Ling Liao, Fuu-Jen Tsa

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.

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) and many will further progress to end-stage renal disease (ESRD). According to Yang et al, 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.

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. Additionally, hereditary components are considered to
constitute a significant risk factor for DN. 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) or meta-analysis. However, to date only a few genetic polymorphisms have been reported to be associated with these two complications, including interleukin (IL)-10 and SLC2A1. Therefore, genetic loci with pleiotropic effects, whereby one genetic locus can affect more than one phenotype, may underlie the association between DR and DKD among patients with diabetes and microalbuminuria (MAU). 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. 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. However, to date only a few genetic polymorphisms have been reported to be associated with these two complications, including interleukin (IL)-10 and SLC2A1. 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 mL/min/1.73 m²) and 909 non-DN controls (eGFR >90 mL/min/1.73 m²) and stage III: verification of DR-related SNPs on diabetic cohort, verification of DR-related SNPs by using 417 subjects with ERFD (eGFR <60 mL/min/1.73 m²) and 909 subjects with eGFR >90 mL/min/1.73 m², were selected from our database. 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 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.

**Stage I**

DR-related genetic marker selection

- 33 SNPs → model building

**Stage II**

Verification in DN cases

- DN n=567
- Non-DN n=909

**Stage III**

Verification in ERFD cases

- ERFD n=417
- Non-ERFD n=909

**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.

**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 mL/min/1.73 m²) and 909 non-DN controls (eGFR >90 mL/min/1.73 m²) and stage III: verification of DR-related SNPs on diabetic cohort, verification of DR-related SNPs by using 417 subjects with ERFD (eGFR <60 mL/min/1.73 m²) and 909 subjects with eGFR >90 mL/min/1.73 m², were selected from our database. 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 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.
as >3.3 mL/min/1.73 m² decline in eGFR per year. The Modified Diet in Renal Disease equation 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 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/~cframan/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) and (6) excessive identity by descent (FDR > 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, using the PLINK module. The MDS method detects meaningful underlying dimensions that 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 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 the individuals were included as covariates in association test models to minimize spurious associations and maximize the detection power to identify genuine associations.
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 previously26 27; the formula is as follows: wGRS =33/77.189×((rs767763×2.025)+(rs16958803×1.992)+(rs1294232×2.036)+(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)+(rs7374657×1.665)+(rs34766496×2.453)+(rs2079628×1.809)+(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)+(rs18941515×5.049)+(rs6065957×1.494)). 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 χ² 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 χ² 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

<table>
<thead>
<tr>
<th></th>
<th>eGFR* &gt;90 (n=909)</th>
<th>eGFR &lt;60 (n=567)</th>
<th>P value†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>457 (50.3%)</td>
<td>316 (55.7%)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>452 (49.7%)</td>
<td>251 (44.3%)</td>
<td>0.041‡</td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td>54.62 (10.22)</td>
<td>69.31 (10.47)</td>
<td>&lt;0.001§</td>
</tr>
<tr>
<td><strong>DM duration (years)</strong></td>
<td>5.90 (5.69)</td>
<td>10.02 (8.80)</td>
<td>&lt;0.001§</td>
</tr>
<tr>
<td><strong>BMI (kg/m²)</strong></td>
<td>25.48 (4.32)</td>
<td>26.06 (3.99)</td>
<td>0.016‡</td>
</tr>
<tr>
<td><strong>HbA1c (%)</strong></td>
<td>7.71 (1.54)</td>
<td>7.45 (1.52)</td>
<td>0.002‡</td>
</tr>
<tr>
<td><strong>SBP (mm Hg)</strong></td>
<td>127.61 (16.55)</td>
<td>135.68 (18.66)</td>
<td>&lt;0.001§</td>
</tr>
<tr>
<td><strong>DBP (mm Hg)</strong></td>
<td>77.62 (11.05)</td>
<td>75.45 (12.34)</td>
<td>0.001‡</td>
</tr>
<tr>
<td><strong>HDL (mg/dL)</strong></td>
<td>49.58 (13.46)</td>
<td>46.64 (13.71)</td>
<td>&lt;0.001§</td>
</tr>
<tr>
<td><strong>LDL (mg/dL)</strong></td>
<td>114.35 (34.48)</td>
<td>106.80 (37.10)</td>
<td>&lt;0.001§</td>
</tr>
<tr>
<td><strong>eGFR (mL/min/1.73m²)</strong></td>
<td>115.17 (24.56)</td>
<td>40.26 (15.08)</td>
<td>0.001‡</td>
</tr>
<tr>
<td><strong>Urine creatinine (mg/dL)</strong></td>
<td>129.45 (220.83)</td>
<td>110.50 (64.27)</td>
<td>0.063</td>
</tr>
<tr>
<td><strong>ACR (mg/g)</strong></td>
<td>67.70 (307.72)</td>
<td>578.03 (1619.15)</td>
<td>&lt;0.001§</td>
</tr>
</tbody>
</table>

Values are presented as N (%) or mean±SD.
*The Modified Diet in Renal Disease equation was used to estimate eGFR.
†P value for χ² 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.

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

<table>
<thead>
<tr>
<th></th>
<th>eGFR* &gt;90 N (%)</th>
<th>eGFR &lt;60 N (%)</th>
<th>OR</th>
<th>95% CI</th>
<th>P value†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td>1.18 (1.15 to 1.21)</td>
<td>0.98 (0.97 to 1.00)</td>
<td>1.00 (0.99 to 1.00)</td>
<td>1.001 (1.001 to 1.002)</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td><strong>BMI (kg/m²)</strong></td>
<td>1.10 (1.04 to 1.16)</td>
<td>0.98 (0.97 to 1.00)</td>
<td>1.00 (0.99 to 1.00)</td>
<td>1.001 (1.001 to 1.002)</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td><strong>HDL (mg/dL)</strong></td>
<td>1.00 (0.99 to 1.00)</td>
<td>0.98 (0.97 to 1.00)</td>
<td>1.00 (0.99 to 1.00)</td>
<td>1.001 (1.001 to 1.002)</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td><strong>LDL (mg/dL)</strong></td>
<td>1.00 (0.99 to 1.00)</td>
<td>0.98 (0.97 to 1.00)</td>
<td>1.00 (0.99 to 1.00)</td>
<td>1.001 (1.001 to 1.002)</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td><strong>ACR (mg/g)</strong></td>
<td>1.00 (0.99 to 1.00)</td>
<td>0.98 (0.97 to 1.00)</td>
<td>1.00 (0.99 to 1.00)</td>
<td>1.001 (1.001 to 1.002)</td>
<td>&lt;0.001†</td>
</tr>
</tbody>
</table>

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.

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,
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

<table>
<thead>
<tr>
<th>Gender</th>
<th>Non-ERFD (n=417)</th>
<th>ERFD (n=417)</th>
<th>P value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>258 (61.9%)</td>
<td>257 (61.6%)</td>
<td>0.943</td>
</tr>
<tr>
<td>Female</td>
<td>159 (38.1%)</td>
<td>160 (38.4%)</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>56.15 (10.58)</td>
<td>57.99 (9.99)</td>
<td>0.010‡</td>
</tr>
<tr>
<td>DM duration (years)</td>
<td>8.32 (7.23)</td>
<td>8.97 (6.39)</td>
<td>0.225</td>
</tr>
<tr>
<td>Follow-up duration (years)</td>
<td>2.86 (2.16)</td>
<td>1.24 (1.10)</td>
<td>&lt;0.001§</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26.04 (3.67)</td>
<td>26.22 (3.86)</td>
<td>0.501</td>
</tr>
<tr>
<td>HbA1c (%)</td>
<td>7.32 (1.53)</td>
<td>7.48 (1.31)</td>
<td>0.110</td>
</tr>
<tr>
<td>SBP (mm Hg)</td>
<td>129.58 (19.56)</td>
<td>135.39 (18.33)</td>
<td>&lt;0.001§</td>
</tr>
<tr>
<td>DBP (mm Hg)</td>
<td>79.08 (13.00)</td>
<td>82.91 (11.60)</td>
<td>&lt;0.001§</td>
</tr>
<tr>
<td>HDL (mg/dL)</td>
<td>49.88 (12.86)</td>
<td>48.09 (13.93)</td>
<td>0.054</td>
</tr>
<tr>
<td>LDL (mg/dL)</td>
<td>115.63 (34.70)</td>
<td>105.16 (32.72)</td>
<td>&lt;0.001§</td>
</tr>
<tr>
<td>eGFR (mL/min/1.73m²)</td>
<td>88.70 (28.09)</td>
<td>90.52 (28.20)</td>
<td>0.352</td>
</tr>
<tr>
<td>Urine creatinine (mg/dL)</td>
<td>141.68 (91.63)</td>
<td>120.52 (72.73)</td>
<td>0.001‡</td>
</tr>
<tr>
<td>ACR (mg/g)</td>
<td>96.12 (353.09)</td>
<td>157.47 (472.41)</td>
<td>0.046*</td>
</tr>
</tbody>
</table>

Values are presented as N (%) or mean±SD.
*Early renal function decline was defined as >3.3 mL/min/1.73 m² decline in the eGFR per year.
†P value for χ² 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.
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 m² 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. 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). El-Asrar et al reported that patients with T1D and DR were 13.39 times more likely to develop DKD than those without DR. Moreover, Yang et al reported that the urine proteome specific for eye damage could predict chronic renal insufficiency (eGFR, 60 mL/min/1.73 m²) in a 5.3 years prospective cohort for T2D.

However, only a few genetic reports support possible pleiotropy between DN and DR. Hosseini et al 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, cell proliferation and differentiation and vascular endothelial relaxation. These pathophysiological associations may be related to nephropathy or diabetic complications, however, none of the genetic variants significantly correlated with DN risk among patients with T2D in previous studies 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

### Table 4

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

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

Values are presented as N (%) or mean±SD.

*Early renal function decline was defined as >3.3 mL/min/1.73 m² 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.

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

Acknowledgements The authors would like to thank all the subjects who were involved in the study and the people who helped in collecting the blood samples. The authors would like to thank Taiwan Biobank for providing samples or related data Genome Medicine of the National Core Facility Program for Biotechnology, Ministry (all anonymous) for the research. The authors would also like to thank the National Center for of Science and Technology for the technical/bioinformatics support.

Contributors W-LL contributed to the conception and design of the study, the acquisition and interpretation of data, Y-CH, Y-FY, H-JL and J-ML contributed to the acquisition of data, A-RR, C-MW and Y-WC analyzed the data, F-JT contributed to the reagents, materials and analysis tools, W-LL wrote and W-LL and A-RR revised the manuscript. All the coauthors have read and approved the final version of the manuscript.

Funding This work was supported in part by research grants from the Biosignature project (grant number BM10701010022) and Biomarker project (grant number AS-BD-108-9), Academia Sinica, Taiwan; from Ministry of Science and Technology of Taiwan (grant number MOST 107-2314-B-039-052-MY2) and from China Medical University (grant number CMU108-MF-83).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The study was approved by the China Medical University Hospital Institutional Review Board (CMUH103-REC2-071).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. All data relevant to the study are included in the article or uploaded as online supplementary information.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

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

REFERENCES


