MicroRNA-Related Genetic Variants Are Associated With Diabetic Retinopathy in Type 1 Diabetes Mellitus

Ebony Liu,1 Georgia Kaidonis,1 Bennet J. McComish,2 Mark C. Gillies,3 Sotoodeh Abhary,1 Rohan W. Essex,4 John H. Chang5,6 Bishwanath Pal,6 Mark Daniell,7 Stewart Lake,1 Nikolai Petrovsky,8 Alex W. Hewitt,9 Alicia Jenkins,10,11 Ecosse L. Lamoureux,9,12,13 Jonathan M. Gleadle,14 Jamie E. Craig,1 and Kathryn P. Burdon1,2

1Department of Ophthalmology, Flinders University, College of Medicine and Public Health, Flinders Medical Centre, Adelaide, South Australia, Australia
2Menzies Institute for Medical Research, University of Tasmania, Hobart, Tasmania, Australia
3Save Sight Institute, Clinical Ophthalmology and Eye Health, the University of Sydney, Sydney, New South Wales, Australia
4Academic Unit of Ophthalmology, Australian National University, Canberra, Australia
5School of Medical Sciences, University of NSW, Sydney, New South Wales, Australia
6Medical Retina Service, Moorfields Eye Hospital, London, United Kingdom
7Department of Ophthalmology, Royal Melbourne Hospital, Melbourne, Victoria, Australia
8Department of Endocrinology, Flinders University, Flinders Medical Centre, Adelaide, South Australia, Australia
9Centre for Eye Research Australia, University of Melbourne, Melbourne, Victoria, Australia
10NHMRC Clinical Trials Centre, University of Sydney, Camperdown, New South Wales, Australia
11St. Vincent’s Hospital, Fitzroy, Victoria, Australia
12Singapore Eye Research Institute, Singapore
13Duke-NUS Medical School, Singapore
14Renal Medicine, College of Medicine and Public Health, Flinders University, Flinders Medical Centre, Adelaide, South Australia

Correspondence: Ebony Liu, Flinders Medical Centre, Flinders Drive, Bedford Park, South Australia, Australia, 5042; ebony.liu@flinders.edu.au.
Submitted: August 22, 2018
Accepted: August 20, 2019


PURPOSE. Few studies have explored the association of genetic variants in microRNA genes and binding sites with diabetic retinopathy (DR) in type 1 diabetes. We conducted a genome-wide scan for single nucleotide polymorphisms (SNPs) in these genes by using data from a genome-wide association study (GWAS).

METHODS. All known SNPs were imputed from our GWAS data (n = 325) of DR cases and diabetic controls (no DR). Relevant SNPs were extracted using miRNASNP and PolymiRTS (version 2) databases. \( \chi^2 \) tests and logistic regression (adjusting for age, sex, duration of diabetes, HbA1c, and hypertension) were used to test the association between the imputed SNPs and DR phenotypes (any DR, nonproliferative DR [NPDR], proliferative DR [PDR], diabetic macular edema [DME], and sight-threatening DR defined as PDR, severe NPDR, or clinically significant macula edema (CSME)) compared with diabetic controls. Top-ranking SNPs were genotyped in a larger cohort (N = 560) to confirm their association with DR.

RESULTS. Three SNPs (rs10061133, rs1049835, rs9501255) were selected and genotyped in the final cohort. Rs10061133 in MIR449b was protective of sight-threatening DR (odds ratio [OR] = 0.32, \( P = 3.68 \times 10^{-4} \)) and PDR (OR = 0.30, \( P = 8.12 \times 10^{-4} \)), and the associations became more significant as the cohort increased in size.

CONCLUSIONS. Rs10061133 in MIR449b is significantly associated with a decreased risk of PDR and sight-threatening DR in Caucasian patients with type 1 diabetes. This can guide future studies on genetic risk profiling and on developing microRNA-related therapies for sight-threatening DR.

Keywords: diabetic retinopathy; microRNA; genetic variants; type 1 diabetes; Caucasian

Diabetic retinopathy (DR) is a leading cause of visual loss due to damage to the retina from diabetes. It is a common complication of type 1 diabetes, with a prevalence of 50% to 80%.1 The incidence of type 1 diabetes, and therefore diabetes-related vision loss, is increasing at a rate of 3% to 5% per year.2

Type 1 diabetes is caused by immune-mediated destruction of insulin-secreting beta cells of the pancreas. Genetic and environmental factors influence the development and progression of the disease and its complications. DR in type 1 diabetes has been demonstrated to have the highest sibling recurrence risk of all the microvascular complications in a large study of 8114 type 1 diabetics among 6707 families.3 The genes that contribute to DR risk are not well understood.4

Over the last 20 years, our understanding of the role of microRNAs in disease has developed. MicroRNAs are short, single-stranded RNA molecules that are involved in regulating gene expression and are believed to modulate the expression of one third of the genes in the genome.5 A single microRNA may
bind to hundreds of genes and induce mRNA degradation or transnational repression. Genetic variations in microRNA genes and their target binding sites can have significant consequences for these regulatory pathways. MicroRNAs have been implicated in a wide range of diseases, including cancer and cardiovascular and neurodegenerative diseases.6–8 Recently, differentially expressed microRNAs have been found to be associated with DR.9 as well as other microvascular diabetic complications.10 Few studies have explored whether genetic variations in microRNA-related genes are associated with DR. We previously studied microRNA 146a because of its identified functional polymorphism rs2910164 as being significantly associated with diabetic macular edema (DME) in type 2 diabetes but not type 1 diabetes.11

In this study, we investigated the association of all known microRNA genetic variations with DR in type 1 diabetes using a large Australian and British Caucasian population. We first performed an in silico study using imputed data from our previously described genome-wide association study (GWAS) for DR.12 We then confirmed the association of significant genetic variants with DR by genotyping a large population that included the original GWAS participants.

METHODS

Ethics Statement

This project was approved by the human research ethics committees (HRECs) in Australia (Southern Adelaide Clinical HREC, Royal Adelaide Hospital HREC, The Queen Elizabeth Hospital [Adelaide] HREC, Royal Melbourne Hospital HREC, Royal Victorian Eye and Ear Hospital HREC, St. Vincent’s Hospital [Melbourne] HREC, South Eastern Sydney Illawarra HREC, Tasmania Health and Medical HREC, Canberra Hospital HREC) and the NHS Health Research Authority in London. It adheres to the tenets of the Declaration of Helsinki. Written informed consent was obtained from each participant before study enrollment.

Recruitment of Patients and Data Collection

This study involved Caucasian participants (identifying as of European descent) recruited in the Australian Registry of Advanced Diabetic Retinopathy (RADAR) and the Genetic Study of Diabetic Retinopathy based at Flinders University, South Australia. Multiple recruitment centers were involved and included the following Australian hospitals: Flinders Medical Centre, The Repatriation General Hospital, The Royal Adelaide Hospital, The Queen Elizabeth Hospital, The Royal Melbourne Hospital, Royal Victorian Eye and Ear Hospital, St. Vincent’s Hospital, Sydney Eye Hospital, Canberra Hospital, Royal Hobart Hospital, and from the United Kingdom, The National Institute for Health Research Biomedical Research Centre at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology.

Eligible participants were actively recruited from ophthalmology, diabetes, and renal clinics, with the following inclusion criteria: previous diagnosis of type 1 diabetes and at least 18 years of age. All participants underwent a questionnaire and venous blood sample collection for DNA analysis. Clinical information was collected from case notes and electronic records, including hemoglobin A1c (HbA1c), renal, and lipid measures; medications; and the presence of hypertension and nonocular diabetic complications. HbA1c was recorded as the average of the three most recent available measurements or of three measurements immediately prior to a diagnosis of proliferative diabetic retinopathy (PDR). Hypertension was defined as systolic and diastolic blood pressure greater than 140 and 90 mm Hg, respectively, or pharmacological treatment for hypertension. DR grading (defined as the worst ever grading) and the presence of DME were determined from documented dilated fundus examinations performed by an ophthalmologist. DR grading was defined by the International Clinical DR Severity Scale.13 No DR was defined as no abnormalities found on dilated fundus examination. PDR was defined as neovascularization, vitreous, or preretinal hemorrhage. Other abnormalities between those two stages were defined as nonproliferative DR (NPDR). Severe NPDR was defined as any of the following: more than 20 intraretinal hemorrhages in all four quadrants of the retina, venous beading in at least two quadrants, or prominent microvascular abnormalities in at least one quadrant. Clinically significant macular edema (CSME) was defined according to the Early Treatment in Diabetic Retinopathy Study protocol.14 Sight-threatening DR was defined as severe NPDR, PDR, or CSME. Only individuals with complete clinical information were included in the analyses.

For each participant, approximately 8 ml blood was collected in EDTA blood collection tubes and underwent DNA extraction using a kit (Qiagen Blood DNA Maxi Kit; Qiagen, Venlo, The Netherlands). Details regarding the data collection method have been described previously.15

In Silico Analysis of MicroRNA Single Nucleotide Polymorphisms (SNPs) in the Type 1 Diabetes Discovery Cohort

A genome-wide association study has been previously performed on 354 type 1 diabetic participants. Genotyping was conducted with a kit (OmniExpress Array; Illumina, San Diego, CA, USA) as described previously.15 The overall study design is illustrated in Supplementary Fig. S1. Following data cleaning, principal components were calculated using EIGENSTRAT software, and individuals were removed if they were further than six standard deviations from the mean of any principal component, limiting the analysis to participants of predominately European descent. Additional outliers were removed based on visualizing the plot of PC1 versus PC2 (Supplementary Fig. S2). Relatedness was calculated using PLINK, a whole genome data analysis toolset, and one of each pair with pi HAT > 0.2 were removed. After quality control, a total of 325 samples remained. The autosomal genotype data were phased using Eagle software (version 2.3.5).14 and genotypes were then imputed on the basis of the 1000 Genomes Project reference panel (EUR reference, Phase III, version 5)15 using Minimac3 genotype imputation software (version 2.0.1).16 Indels, SNPs within 5 bp of an indel, rare variants (minor allele frequency [MAF] < 0.01), and variants with poor imputation quality (R² < 0.8) were removed. Known microRNA genes and their binding sites were downloaded from the miRNASNP and PolymiRTS (version 2) databases (accessed August 31, 2017) and used to filter the imputed SNPs.

The most likely genotype was used for the association analyses with the imputed SNPs. Association tests with DR phenotypes were performed in PLINK (version 1.07). Controls were defined as those with diabetes but no DR. Analysis for association with DR phenotypes was repeated for different definitions of cases: (1) any DR, (2) PDR, (3) DME, and (4) sight-threatening DR (severe NPDR, PDR, or CSME). Logistic regression for multivariate analysis incorporated age, sex, duration of diabetes, HbA1c, and hypertension. A quantile-quantile plot of the P values in the association study for any DR...
compared with no DR in the GWAS discovery cohort is shown in Supplementary Figure S3.

Stage 2 Genotyping and Combined Analysis

Top-ranking SNPs were selected for genotyping in stage 2 to confirm that their association with DR as a two-stage design with joint analysis is both cost effective and more powerful than discovery and replication. For variations in microRNA genes, SNPs were selected if $P < 0.01$ in the imputed data, and for variations in microRNA binding sites, SNPs were selected if $P < 0.0001$. All SNPs chosen had minor allele frequencies greater than 1% in the imputed data.

Selected SNPs were genotyped in the total sample ($N = 560$) by the Australian Genome Research Facility, using the MassARRAY platform (Agena Bioscience, San Diego, CA, USA). The same statistical analyses as described above were performed: firstly, in the original GWAS group ($n = 325$) to validate the analysis with the imputation results, secondly with the additional samples ($n = 235$) for replication, and thirdly in all samples combined ($N = 560$) to maximize the study’s power in the two-stage design.

RESULTS

Demographic details of the total study population stratified by DR phenotype are shown in Table 1 for no DR and any DR, and Supplementary Table S1 for PDR, DME, and sight-threatening DR. As expected, diabetes duration and HbA1c increased as the severity of DR increased, and this was significantly different between cases and controls.

In the discovery cohort, 2420 SNPs in microRNA genes and 401,000 SNPs in microRNA binding sites of target genes were retrieved from the imputed SNP data ($n = 325$ type 1 diabetes patients). After quality control, 213 microRNA variants and 401,000 SNPs in microRNA binding sites of target genes were selected for genotyping, however only three (one in a microRNA gene and two in binding sites) were successfully genotyped in the final study population: rs10061133 (MIR449b), rs9501255 (HLA-DPB1), and rs1049835 (GPM6A). Table 2 shows the number of individuals in each stage of analysis and with each phenotype.

Two of the three SNPs were associated with a DR phenotype after multivariable analysis (adjusting for age, sex, duration of diabetes, HbA1c, and hypertension) of the genotyped data in the original discovery cohort (Table 3). The lack of validation of rs9501255 in HLA-DPB1 suggests the imputation of this SNP was of poor quality, as is often seen with HLA SNPs. Of the two remaining SNPs associated with sight-threatening DR in the discovery cohort, neither were associated with sight-threatening DR ($P > 0.05$) in the additional samples genotyped ($n = 235$). When all samples were combined, rs10061133 (MIR449b) and rs1049835 (GPM6A) were associated with sight-threatening DR (Table 3). Rs1006113 (MIR449b) in particular became more significantly associated with sight-threatening DR as the sample size increased ($P = 3.68 \times 10^{-5}$) and showed consistent OR between the discovery and the additional samples (Table 3). This SNP was protective against sight-threatening DR with an OR of 0.51 to 0.37 across the three stages.

The two associated SNPs (rs10061133 and rs1049835) were then analyzed for association with DME and subtypes of sight-threatening DR (PDR and CSME) in the final combined sample. Rs10061133 in MIR449b was strongly associated with a decreased risk of PDR (OR = 0.50, 95% CI 0.15–0.61, $P = 8.12 \times 10^{-4}$) but not the other phenotypes.

DISCUSSION

In our discovery cohort, no SNPs reached significance after multiple hypothesis testing ($P < 3.00 \times 10^{-7}$). Therefore, we undertook a two-stage design, selecting the top-ranking SNP for evaluation in the second cohort and a combined analysis between the two. The two-stage design with joint analysis is both cost effective and more powerful than discovery and replication. We found that rs10061133 in MIR449b was most consistently associated with sight-threatening DR and PDR. The minor allele G was protective against sight-threatening DR and PDR after adjustment for covariates and was confirmed with greater significance in the larger sample. A large proportion of the sight-threatening DR group consisted of PDR samples (181 out of 223), and thus this association is likely to be more important in PDR risk, rather than CSME. Rs10061133 in MIR449b has been reported to be associated with other diseases such as esophageal cancer, thyroid

| Table 1. Demographics of Total Study Population Stratified Between No DR and Any DR |
|---------------------------------|-------------|-------------|
| Demographic                | No DR      | Any DR      | $P$ Value |
| $n$                          | 228        | 332         |           |
| Female, $n$ (%)             | 99 (43.8)  | 178 (47.3)  | 0.44      |
| Age, y, median (range)      | 34 (17–83) | 47 (18–88)  | 3.33 $\times 10^{-22}$ |
| Diabetes duration, y, median (range) | 15 (5–60) | 28 (5–70)   | 7.57 $\times 10^{-49}$ |
| HbA1c, %, median (range)    | 8.06 (5–15) | 8.6 (5–15) | 4.00 $\times 10^{-6}$ |
| Hypertension, $n$ (%)       | 53 (23.5)  | 225 (59.8)  | 2.26 $\times 10^{-18}$ |

| Table 2. Number of Individuals in Each Stage of Analysis by Phenotype Included in Logistic Regression Analysis |
|-------------------------------------------------|---------|---------|---------|
| Discovery Cohort | Additional Cohort | Total |
| No DR            | 123     | 105     | 228     |
| Any DR           | 202     | 150     | 352     |
| DME              | 48      | 41      | 89      |
| PDR              | 125     | 56      | 181     |
| Sight-threatening DR | 155    | 68      | 223     |
TABLE 3. Validation and Replication of Association Results

<table>
<thead>
<tr>
<th>SNP</th>
<th>Allele</th>
<th>Phenotype</th>
<th>Controls OR 95% CI</th>
<th>Discovery, Imputed</th>
<th>Genotyped Additional</th>
<th>All Samples Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10061133</td>
<td>MIR449B</td>
<td>G</td>
<td>Sight-threatening DR</td>
<td>0.08 vs. 0.14</td>
<td>0.36 (0.17–0.74)</td>
<td>5.32</td>
</tr>
<tr>
<td>rs1049835</td>
<td>GPM6A</td>
<td>G</td>
<td>Sight-threatening DR</td>
<td>0.42 vs. 0.27</td>
<td>2.17 (1.38–3.44)</td>
<td>2.40</td>
</tr>
<tr>
<td>rs9501255</td>
<td>HLA-DPB1</td>
<td>T</td>
<td>Any DR</td>
<td>0.02 vs. 0.08</td>
<td>0.26 (0.22–0.88)</td>
<td>0.017</td>
</tr>
</tbody>
</table>

P-value and OR with 95% CI are given for all successfully genotyped SNPs at each stage of the analysis including both the imputation and genotyping of the discovery cohort along with the additional samples genotyped and the combined cohort. All analyses are adjusted for age, sex, duration of diabetes, HbA1c, and hypertension. CI, confidence interval; OR, odds ratio; MAF, minor allele frequency.

Polymorphisms of microRNA genes can affect either the seed region (binding site) or the processing of precursor microRNA to mature microRNA and therefore the overall production and function of the molecule. Evidence supporting this has largely come from bioinformatics and statistical analyses. Rs10061135 is a single nucleotide polymorphism, where the replacement of G over A in the MIR449b gene is predicted to alter a Dicer cleavage site according to the miRvar database. Dicer is the enzyme that cleaves the precursor microRNA to the mature form. Therefore, rs10061135 could have functional consequences on microRNA-449b through this mechanism. Further functional studies are required to explore this hypothesis.

Despite our intention to comprehensively study all known microRNA variants, some microRNAs could not be studied. We attempted to study 2420 known microRNA gene variants and 401,000 microRNA binding site variants, but only 213 and cancer, premature ovarian insufficiency, and recurrent pregnancy loss. To our knowledge, this is the first study to report an association between rs10061135, microRNA-449b, and DR in type 1 diabetes.

Previous studies have explored microRNA involvement in DR. Kovacs et al. reported 86 differentially expressed microRNAs in retinal endothelial cells of streptozotocin-induced diabetic rats. The majority of these were NF-κB, VEGF, and p53 responsive and included microRNA-449a but not microRNA-449b. Other studies have compared expression levels of microRNA in serum between cases and controls using microRNA arrays. These studies were small, and only one study has been performed in type 1 diabetes. Barutta et al. found 25 differentially expressed microRNAs between 312 diabetic complication cases and 143 diabetic controls in type 1 diabetes. Specifically, microRNA-126 levels were negatively associated with proliferative retinopathy (0.77 vs. 1.29, P = 0.02). Studies profiling microRNA using microarrays have identified microRNA-21, 200b, 15a, 320a, 320b, 93, 29a, and 423-5p to be significantly associated with proliferative disease in both types of diabetes.

MicroRNA-449b is part of a family of microRNAs with similar sequences and secondary structures including microRNA-449a, 449c, and 38. They are all involved in cell proliferation, tumorigenesis, and angiogenesis and are importantly hypoxia-regulated microRNAs. MicroRNA-449a and 449b have been found in retinal tissues of mice. In an oxygen-induced DR mouse model, microRNA-449a was significantly downregulated, which is in keeping with a protective role against DR. MicroRNA-449b has not been reported in human vitreous; however, the studies investigating vitreous microRNAs are few, and a small number of specimens (fewer than five PDR samples) have been analyzed. The angiogenesis pathway that microRNA-449b regulates is the E2F pathway. E2F is a transcription factor essential for cell proliferation. E2F increases the levels of microRNA-449b that then inhibit E2F in a negative feedback loop. Activation of E2F regulates the expression of other genes involved in angiogenesis, such as VEGF and PIGF and breakdown of the blood-retinal barrier such as E-cadherin, and hence there is good biological plausibility for the involvement of microRNA-449b in DR susceptibility.

Another downstream target of microRNA-449a/b is SERPIN1E1, also known as plasminogen activator inhibitor-1 (PAI-1). Under hypoxic conditions, downregulation of microRNA-449a/b increases expression levels of PAI-1. The main role of PAI-1 is to inhibit fibrinolysis, but it has also been shown to play an important role in facilitating retinal angiogenesis. PAI-1 levels are increased in the serum of DR patients compared with diabetic controls, and polymorphisms of the PAI-1 gene have been previously linked with DR.

Polymorphisms of microRNA genes can affect the processing of precursor microRNA to mature microRNA and therefore the overall production and function of the molecule. Evidence supporting this has largely come from bioinformatics and statistical analyses. Rs10061135 is a single nucleotide polymorphism, where the replacement of G over A in the MIR449b gene is predicted to alter a Dicer cleavage site according to the miRvar database. Dicer is the enzyme that cleaves the precursor microRNA to the mature form. Therefore, rs10061135 could have functional consequences on microRNA-449b through this mechanism. Further functional studies are required to explore this hypothesis.

Despite our intention to comprehensively study all known microRNA variants, some microRNAs could not be studied. We attempted to study 2420 known microRNA gene variants and 401,000 microRNA binding site variants, but only 213 and...
41,578 SNPs, respectively, were imputed with good quality. While we had many type 1 diabetic participants from our registry, after stratifying the group into different phenotypes, the numbers for analyses were much smaller. This clearly limits the scope of the current study. Future work should focus on directly genotyping microRNA SNPs in a suitably powered cohort, with replication in independent cohorts. Some of the top-ranked SNPs in microRNA genes also failed genotyping in the replication dataset. It is possible that as microRNAs form secondary structures, the DNA template could have also formed secondary structures during the PCR stage and potentially inhibited primer binding. This study could be strengthened by analysis of miRNAs in biological tissues (such as plasma/serum) of patients with type 1 diabetes; however, as the genetic association results do not reach significance after multiple hypotheses testing ($P < 3.00 \times 10^{-5}$), the results should be replicated by independent studies before resources are directed to functional assays. Interpretation of genetic associations is also complicated by the clinical picture of DR. Age and duration of diabetes are significant risk factors for worsening DR, resulting in the selection of a younger and healthier control group. We have corrected for hypertension, given its strong association with DR, but have not adjusted for other comorbidities, such as hyperlipidemia or macrovascular complications (e.g., heart disease). Strengths of this study include sample population from multiple sites, rigor of DR status characterization and categorization, analyses with adjustments for known clinical risk factors, and meta-analyses after a two-stage design.

**Conclusions**

Through a genome-wide approach, we have identified novel microRNA and microRNA binding site genetic variants associated with DR phenotypes in a large Caucasian population of type 1 diabetic participants. Rs10061133 in MIR449b was found to be strongly and consistently associated with sight-threatening DR and PDR, when compared with diabetic controls (no DR), although independent replication is required. Bioinformatics predict rs10061133 has functional consequences on microRNA449b and therefore can affect regulation of inflammatory pathways in DR. These results have helped elucidate the complex genetic mechanisms involved in DR susceptibility in type 1 diabetes and can guide future studies on genetic risk profiling and developing microRNA-related therapies for sight-threatening DR.

**Acknowledgments**

Supported by the National Health and Medical Research Council (NHMRC) of Australia (Project Grant no. 595918), the Ophthalmic Research Institute of Australia, the National Institute for Health (NIHR) Biomedical Research Centre at Moorfields Eye Hospital, NHS Foundation Trust and UCL Institute of Ophthalmology. The funding organization had no role in the design or conduct of this study.

Disclosure: E. Liu, None; G. Kaidonis, None; B.J. McComish, None; M.C. Gillies, None; S. Abhary, None; R.W. Essex, None; J.H. Chang, None; B. Pal, None; M. Daniele, None; S. Lake, None; N. Petrovsky, None; A.W. Hewitt, None; A. Jenkins, None; E.L. Lamoureux, None; J.M. Gleadle, None; J.E. Craig, None; K.P. Burdon, None

**References**


