Association Testing of Previously Reported Variants in a Large Case-Control Meta-analysis of Diabetic Nephropathy


We formed the GENEtics of Nephropathy—an International Effort (GENIE) consortium to examine previously reported genetic associations with diabetic nephropathy (DN) in type 1 diabetes. GENIE consists of 6,366 similarly ascertained participants of European ancestry with type 1 diabetes, with and without DN, from the All Ireland-Warren 3-Genetics of Kidneys in Diabetes U.K. and Republic of Ireland (U.K.-R.O.I.) collection and the Finnish Diabetic Nephropathy Study (FinnDiane), combined with reanalyzed data from the Genetics of Kidneys in Diabetes U.S. Study (U.S. GoKinD). We found little evidence for the association of the EPO promoter polymorphism, rs161740, with the combined phenotype of proliferative retinopathy and end-stage renal disease in U.K.-R.O.I. (odds ratio [OR] 1.14, \( P = 0.19 \)) or FinnDiane (OR 1.06, \( P = 0.60 \)). However, a fixed-effects meta-analysis that included the previously reported cohorts retained a genome-wide significant association with that phenotype (OR 1.31, \( P = 2 \times 10^{-8} \)). An expanded investigation of the ELMO1 locus and genetic regions reported to be associated with DN in the U.S. GoKinD yielded only nominal statistical significance for these loci. Finally, top candidates identified in a recent meta-analysis failed to reach genome-wide significance. In conclusion, we were unable to replicate most of the previously reported genetic associations for DN, and significance for the EPO promoter association was attenuated.


Type 1 diabetes has continuously increased worldwide, and the highest incidence is found in Finland (1). Diabetic nephropathy (DN) is a complication that develops in approximately 25–40% of patients with type 1 and type 2 diabetes. DN is the leading cause of end-stage renal disease (ESRD) in the developed world. Currently, 44% of the new cases of ESRD in the U.S. annually are attributable to DN (2). A better understanding of the causal factors of DN and its pathogenesis may lead to new strategies to decrease its incidence, preemptively treat the disorder, attenuate morbidity and mortality, and would be a valuable contribution to global public health.

Several observations suggest that DN, one of the major complications of type 1 and type 2 diabetes, has an inherent genetic susceptibility. Familial clustering of DN is evident for both type 1 and type 2 diabetes (3–6), and genetic risk factors are being sought in multiple populations (7–9). Unfortunately, robust replication of many initial associations has not been forthcoming (10).

This study recruited a large collection of individuals with type 1 diabetes as part of the GENEtics of Nephropathy—an International Effort (GENIE) consortium and examined selected candidate loci associated with DN from genome-wide case-control studies or other association studies that reported high levels of statistical significance. The variants examined and the rationale for their inclusion are as follows:

1) A single nucleotide polymorphism (SNP) (rs1617640) within the promoter region of the EPO gene (encoding erythropoietin) was identified as having a genome-wide significant \((P < 5 \times 10^{-8})\) association with ESRD and
proliferative diabetic retinopathy (PDR) (11). Interestingly, erythropoietin levels were elevated sevenfold in the human vitreous fluid of nondiabetic individuals with the risk genotype TT compared with those with the wild-type GG genotype. In addition, EPO expression levels were significantly elevated above control in the tissues and vitreous fluid of animal models of DN (DN in db/db mice) and in proliferative retinopathy (murine oxygen-induced retinopathy model), respectively (11).

2) The engulfment and cell motility 1 gene (ELMO1) has been reported to be associated with DN in Japanese patients with type 2 diabetes (12). Recently, Pezzolesi et al. (13), using the Genetics of Kidneys in Diabetes U.S. Study (U.S. GoKinD) cohorts, also examined ELMO1 for association with DN and presented evidence of association of variants within this gene for the development of DN. However, the risk alleles for ELMO1 identified in their study differed from those reported in the original Japanese investigation. In the context of a genome-wide association study (GWAS), 118 SNPs were assessed in 1,705 individuals of European ancestry with type 1 diabetes (885 control subjects and 820 DN case subjects). The strongest associations in ELMO1 in the U.S. study occurred at rs11760038 (odds ratio [OR] 1.24; \( P = 1.7 \times 10^{-3} \)) and rs1882080 (OR 1.23; \( P = 3.2 \times 10^{-3} \)), located in intron 16. Two additional SNPs, located in introns 18 and 20, were also nominally associated with DN. In total, eight ELMO1 SNPs were reported to confer risk for DN, although none reached genome-wide significance (13). Supportive evidence was also found in African Americans with type 2 diabetes and ESRD (14).

3) The U.S. GoKinD GWAS analyzed 359,193 SNPs in 820 case subjects (284 with proteinuria and 536 with ESRD) and 885 control subjects with type 1 diabetes but no evidence of DN. Although no risk variant achieved genome-wide significance, the primary association analysis identified 11 SNPs representing four distinct chromosomal regions (\( P < 1 \times 10^{-5} \)). The strongest association with DN reported in this study was on chromosome 9q with rs10868025 (OR 1.45; \( P = 5.0 \times 10^{-7} \)) (15).

4) Finally, in an effort to systematically explore and comprehensively capture common genetic variations that might be associated with DN, we reviewed the largest meta-analysis published to date studying genetic associations with the DN phenotype (7). In GENIE, we examined the top-reported SNP (or proxy) for each gene in that report for an association with DN.

In this study, we have assembled the largest reported case-control sample of DN in type 1 diabetes to evaluate the previously reported genetic associations in newly genotyped samples from the U.K., Republic of Ireland (R.O.I.), and Finland, plus pre-existing data from the U.S. GoKinD.

**TABLE 1**

<p>| Phenotypic characteristics of the GENIE cohorts (U.K.-R.O.I., FinnDiane, and U.S. GoKinD) |</p>
<table>
<thead>
<tr>
<th>U.K.-R.O.I.</th>
<th>FinnDiane</th>
<th>U.S. GoKinD†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>n</strong></td>
<td>903</td>
<td>1,001</td>
</tr>
<tr>
<td><strong>Sex (n)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>531</td>
<td>438</td>
</tr>
<tr>
<td>Female</td>
<td>372</td>
<td>563</td>
</tr>
<tr>
<td><strong>Type 1 diabetes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration (years)</td>
<td>32.9 ± 9.5</td>
<td>27.2 ± 8.7</td>
</tr>
<tr>
<td>Age at diagnosis (years)</td>
<td>14.6 ± 7.7</td>
<td>14.5 ± 7.8</td>
</tr>
<tr>
<td>HbA1c (%)</td>
<td>9.0 ± 1.9</td>
<td>8.7 ± 1.6</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26.2 ± 4.7</td>
<td>26.2 ± 4.2</td>
</tr>
<tr>
<td>ESRD (%)</td>
<td>28.0</td>
<td>0</td>
</tr>
</tbody>
</table>

Categoric data are shown as indicated; continuous data as mean ± SD. †Reanalysis of the U.S. GoKinD dataset using new quality control filters to account for published plate effects (see Research Design and Methods for complete details).
Center (JDC) using differing methods of ascertainment and recruitment (see Pezzolesi et al. [15] for details). Analysis of the U.S. GoKinD cohort was limited to individuals whose reported primary ethnicity was white.

**Phenotype definition: DN and EPO study outcomes.** DN was the primary outcome in EPO study, restricted to individuals with ESRD and PDR, including the previously reported cohorts (Fig. 1). As an additional experimental control, we examined the association of the EPO promoter polymorphism, rs1617640, with DN. For this analysis, case subjects had clinical evidence of PDR, whereas control subjects had none for sex and age. U.K.-R.O.I was adjusted for recruitment center, but the two U.S. GoKinD centers, GWU and JDC, were analyzed separately as reported by Pezzolesi et al. [15].

**SNP selection.** SNP markers with evidence for association with DN susceptibility in reference studies (11) were selected for genotyping in GENIE. For this analysis, case subjects had clinical evidence of PDR, whereas control subjects had none for sex and age. U.K.-R.O.I. was adjusted for recruitment center, but the two U.S. GoKinD centers, GWU and JDC, were analyzed separately as reported by Pezzolesi et al. [15]. Data from the GWAS genotyping was adjusted additionally for duration of type 1 diabetes and principal components from Eigenstrat analysis. The EPO locus was analyzed with the Pearson $\chi^2$ test in the FinnDiane dataset without adjusting for any covariates. Fixed-effects meta-analyses were conducted with the software package Comprehensive Meta-Analysis (Version 2.2.040, Englewood, NJ) and the software package METAL (http://www.sph.umich.edu/csg/abecasis/Metal/) (21) under the additive genetic model to determine the appropriate model for comparison for multiple testing, we calculated the total number of effective tests (because a large proportion of SNPs were in LD), using SNPSpD (http://gmp.imr.uu.se/general/data/NSNPspD/) (SNPSpD uses correlation between analyzed SNPs to calculate the total number of independent tests). The total number of tests tested is 2,199, with 113,7 effect-independent tests. Thus the experiment-wide cutoff for statistical significance was set at $4.4 \times 10^{-8}$ (0.05/113.7).

## RESULTS

### EPO promoter polymorphism.

The association of the EPO promoter polymorphism, rs1617640, with DN was evaluated by de novo genotyping of the SNP in GENIE. For this analysis, case subjects were defined as having both ESRD and PDR because the initial report showed the polymorphism was robustly associated with DN when both of these “extreme” phenotypes were coexpressed. Significant association was not observed in the U.K.-R.O.I. (both $P = 0.19$) or FinnDiane (where $P = 0.60$), although the directions of effect were consistent with the original report. Fixed-effects meta-analysis of the association of rs1617640 with ESRD/PDR, including the previously reported cohorts (a total of 3,162 case and 3,845 control subjects across five separate cohorts of European and European-American ancestry) retained genome-wide statistical significance (OR 1.31 [95% CI 1.20–1.44], $P = 2 \times 10^{-8}$, Fig. 1).

As an additional experimental control, we examined the potential association of the EPO promoter polymorphism with the development of PDR in case subjects, irrespective of ESRD status. No association was observed between EPO and PDR for the individual cohorts or in the meta-analysis of the combined results for FinnDiane (OR 0.95 [95% CI 0.85–1.04], $P = 0.25$) or and U.K.-R.O.I. (0.96 [0.88–1.04], $P = 0.29$). Furthermore, no association was observed after inclusion of the restricted phenotype, PDR, in U.S. GoKinD case and control subjects, separately, or in the meta-analysis of all cohorts combined (results not shown).

**ELMO1.** Neither single-center nor meta-analysis of de novo genotyping in U.K.-R.O.I, nor GWAS data for FinnDiane, revealed a significant association in subjects with type 1 diabetes between rs741301, the previously reported risk variant within ELMO1, and DN (OR 1.04 [95% CI 0.95–1.13], $P = 0.46$; Supplementary Fig. 1). Pezzolesi et al. [15] also tested rs741301 but did not replicate the reported association. They went on to test other SNPs in the region and reported nominal associations ($P = 0.002–0.05$) with eight other SNPs. We examined LD between rs741301 and the
SNPs reported to be associated with DN by these investigators. The $r^2$ statistic between rs741301 and the other SNPs revealed only low to moderate LD, ranging from 0.38 to 0.65 (Supplementary Table 1).

As an additional and more extensive test of variants in this region, we performed an expanded analysis capturing all available SNPs 20 kb upstream and downstream of the ELMO1 locus to account for LD differences presumably due to ancestry. Results of the expanded analysis did not reveal any significant SNPs for either cohort individually or for the two cohorts meta-analyzed after correcting for multiple testing ($P < 4.3 \times 10^{-4}$). Furthermore, no SNP achieved significance with inclusion of the U.S. GoKinD results in the meta-analysis (Supplementary Table 2).

**Risk variants reported from U.S. GoKinD for DN in type 1 diabetes.** Eleven DN susceptibility SNPs that were first reported by the U.S. GoKinD investigators as highly associated with the risk of developing DN in type 1 diabetes were parsed into eight candidate loci. We selected one representative SNP for each region of strong LD in which multiple SNPs represented the same association signal. After performing additional QC checks (see RESEARCH DESIGN AND METHODS), we first tested the eight U.S. GoKinD potential DN susceptibility SNPs by reanalyzing the U.S. GoKinD dataset downloaded from dbGAP (19). As in the original report, none of the eight SNPs were associated with DN at genome-wide statistical significance (Table 2). The two SNPs in the FERM (F [Band 4.1], E [Ezrin], R [Radixin], M [Moesin]) domain 3 (FRMD3) region showed similar $P$ values as in the original report ($2.1 \times 10^{-7}$ and $1.6 \times 10^{-6}$, respectively; Table 2). In our analysis, the statistical significance of the $P$ values for SNPs in the CPVL/CHN2 and CARS regions was reduced from $6.5 \times 10^{-7}$ to $2 \times 10^{-5}$ and $6.4 \times 10^{-6}$ to $2.2 \times 10^{-3}$, respectively. $P$ values for the 6 SNPs in the 13q region were also changed from $1.8\text{–}7.0 \times 10^{-6}$ to $1.4\text{–}9.5 \times 10^{-3}$.

We next examined these SNPs in our newly genotyped samples. Case-control association analysis for these loci in the U.K.-R.O.I. and FinnDiane samples revealed no significant associations in either cohort. The strongest signal was observed at rs39075 near CPVL/CHN2 in U.K.-R.O.I., (OR 1.12, $P = 0.08$) and in meta-analysis of the two replication cohorts (U.K.-R.O.I. and FinnDiane; OR 1.06, $P = 0.06$; Table 2). In expanded locus-region analyses (plus or minus 20 kb of the locus of interest), no SNP reached significance after adjustment for multiple testing (one-tailed $P = 4.3 \times 10^{-4}$) for the two cohorts separately or via meta-analysis. The combined meta-analysis including U.S. GoKinD revealed two SNPs downstream of FRMD3, rs1888747 ($P = 1.5 \times 10^{-4}$) and rs13288659 ($P = 9.7 \times 10^{-4}$), which showed significance after adjusting for experiment-wide multiple testing ($P < 4.3 \times 10^{-4}$). However, neither SNP achieved genome-wide significance ($P < 5 \times 10^{-8}$, Supplementary Table 3).

**Pooled meta-analyses examining variants associated with DN in type 1 and type 2 diabetes.** In the most comprehensive literature search for DN associated genetic variants to date, Mooyaart et al. (7) identified 24 loci. In GENIE, we examined all the available top-reported SNPs (or their proxies) for each gene in that report. Three SNPs were nominally associated ($P < 0.05$) with DN: rs13293564 at UNC13B ($P = 0.01$) and rs179975 at the ACE polymorphism remained nominally significant ($P = 0.04$). Including the U.S. GoKinD results, the FRMD3 signal at rs1888747 emerged as noted above; no other signals were significant after adjusting for multiple comparisons (see Supplementary Table 4 for full details).

**DISCUSSION**

Using a large, homogeneous population sample of European ancestry subjects with type 1 diabetes in the GENIE consortium, we were unable to replicate most of the previously reported genetic associations with DN that we examined.

Our findings do not support previously reported genetic associations with DN in type 1 diabetes in the largest GWAS published to date (15). None of those signals reached genome-wide statistical significance with the addition of larger, similarly ascertained datasets. Using ORs at the lower limit of the 95% CI from the original publication, our


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>rs741301</td>
<td>EPO</td>
<td>C</td>
<td>2.87E-08</td>
<td>1.13E-07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs390752</td>
<td>CPVL/CHN2</td>
<td>A</td>
<td>1.25E-03</td>
<td>1.01E-02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs451041</td>
<td>CARS</td>
<td>C</td>
<td>1.22E-03</td>
<td>1.25E-02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1888747*</td>
<td>APOL1</td>
<td>T</td>
<td>1.20E-04</td>
<td>1.12E-03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1411765</td>
<td>FRMD3</td>
<td>G</td>
<td>1.23E-03</td>
<td>1.12E-02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1411766</td>
<td>ELMO1</td>
<td>A</td>
<td>1.41E-04</td>
<td>1.23E-03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1741288</td>
<td>ELMO1</td>
<td>A</td>
<td>1.43E-04</td>
<td>1.23E-03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs6492208</td>
<td>No gene (13q)</td>
<td>T</td>
<td>1.37E-03</td>
<td>1.20E-02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs7989848</td>
<td>No gene (13q)</td>
<td>A</td>
<td>1.37E-03</td>
<td>1.19E-02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Results in the reanalyzed U.S. GoKinD and additional GENIE cohorts for eight SNPs with strongest reported associations to DN in U.S. GoKinD.
of a false-negative finding (type II error), even accounting for the likely overestimation of effect sizes due to the winner’s curse phenomenon (25). We also harmonized the ascertainment criteria for case-control definitions across all the study populations (including U.S. GoKinD), making it unlikely that phenotypic heterogeneity across study populations explains the lack of replication.

A crucial issue that bears on the interpretation of case-control studies of the genetics of DN concerns the adequacy of phenotype definition. In this and most studies cited to date, there is the presumption that long-duration diabetes exposure and the presence of frank protein in the urine—macroalbuminuria—defines DN and that phenotypic heterogeneity has been well controlled through this classification. These definitions are derived in large measure from the classic studies of Parving et al. (26), Viberti et al. (27), and Mogensen and Christensen (28), who documented 30 years ago a virtually inexorable progression to ESRD in patients who developed microalbuminuria after approximately 2 decades of exposure to the diabetic metabolic milieu. However, these longitudinal findings were based on small numbers of patients. The plasticity of DN phenotypes is reflected in more recent and much larger longitudinal studies showing that most patients with type 1 diabetes, categorized initially as having microalbuminuria, undergo regression to normoalbuminuria with preservation of renal function (29). It is not entirely clear that microalbuminuria versus macroalbuminuria, stage of chronic kidney disease and attendant renal function, the rate of renal decline, or the occurrence of extreme phenotypes, such as ESRD/PDR, represent one disease process along a continuum or many distinct disease states, each of which may be under distinct genetic control. As pointed out recently (24), genetic variants, such as those in MYH9 and APOL1 that are common in certain ethnic groups, may mask the effects at other loci unless methods such as multilocus modeling and interaction analyses are used to control for these effects.

In addition, phenotypic variation may be a function of ethnicity and disease-specific gene expression. For example, Pima Indians with type 2 diabetes have very early-onset DN, characterized by an accelerated loss of renal function and progression to ESRD despite lower blood pressures and lipid levels, factors thought to be protective (30). This has been postulated to be due to structural differences in the nephron–podocyte number and density per glomerulus ("podocyte insufficiency"), a decrease in net nephron mass (glomerulopenia) resulting in glomerulomegaly, increased intraglomerular capillary pressure, and ultimately, hyperfiltration injury (30). Whether these structural and intrarenal hydraulic changes could be genetically regulated is ultimately a testable hypothesis; they warrant further investigation to continue the inquiry why certain populations have an apparent disproportional susceptibility to ESRD and, particularly, DN.

In summary, we have presented evidence that several previously reported genetic associations with DN in type 1 diabetes could not be replicated in a large, homogeneous sample of subjects with type 1 diabetes. Our failure to replicate these associations underscores the need to apply stringent statistical thresholds of significance, maximize power through meta-analysis of all available data, and seek replication in independent samples, as has been proposed by a number of different authors (31,32). Finally, the applicability and generalizability of DN risk loci from type 1 diabetes to type 2 diabetes, and the related question of shared genetic susceptibility for nephropathy between type 1 and type 2 diabetes, remain unresolved.

ACKNOWLEDGMENTS

Funding for this study was provided by National Institutes of Health National Institute of Diabetes and Digestive and Kidney Diseases Grant R01-DK-081923 to J.N.H., P.-H.G., and J.C.F., a Science Foundation Ireland U.S.-Ireland R&D partnership, and The Health Research Board Ireland. R.M.S. was supported by a Juvenile Diabetes Research Foundation postdoctoral fellowship (#3-2011-70). The FinnDiane Study was supported by grants from the Folkhälsoan Research Foundation, the Wilhelm and Else Stockmann Foundation, Liv och Hälsa Foundation, Helsinki University Central Hospital Research Funds (EVO), the Sigrid Juselius Foundation, the Signe and Arne Gyllenberg Foundation, Finska Läkaressällskapet, the European Commission, and the European Union’s Seventh Framework Program (FP7/2007-2013) for the Innovative Medicine Initiative under Grant Agreement No. IMI/115006 (the Surrogate Markers for Micro- and Macrovascular Hard Endpoints for Innovative Diabetes Tools [SUMMIT] consortium).

J.C.F. has received consulting honoraria from Novartis, Eli Lilly, and Pfizer. No other potential conflicts of interest relevant to this article were reported.

W.W.W. contributed to study design and management, data acquisition, statistical and data analysis, and manuscript preparation and review. R.M.S. contributed to data acquisition, genotyping, statistical and data analysis, and manuscript review. A.J.M. contributed to data acquisition, statistical and data analysis, and manuscript preparation. N.S. contributed to data acquisition, statistical and data analysis, and manuscript preparation and review. C.F. contributed to data acquisition, statistical and data analysis, and manuscript review. A.T., C.Gu., and M.P. contributed to genotyping and manuscript review. J.B.M., E.F., V.H., R.L., D.G., K.H., J.K., M.R.-B., N.T., M.S., J.W., and B.H. contributed to data acquisition and manuscript review. G.J.M., T.I., E.P.B., D.M.S., C.P., S.B., F.M., C.Go., and A.P.M. contributed to manuscript review. J.S. contributed to genotyping and statistical and data analysis, and manuscript review. L.T., J.P., C.S., J.T., and A.M.O. contributed to data acquisition. A.S. contributed to data acquisition, genotyping, and manuscript review. K.T. contributed to statistical and data analysis, and manuscript review. J.N.H. contributed to study design and manuscript review. P.-H.G. contributed to study design, statistical and data analysis, and manuscript review. J.C.F. contributed to study design and management, statistical and data analysis, and manuscript preparation and review. J.C.F. is the guarantor of this work and, as such, had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.


PRIOR GENETIC ASSOCIATIONS WITH DN


