

## Genetic association analyses of non-synonymous single nucleotide polymorphisms in diabetic nephropathy

D. A. Savage · C. C. Patterson · P. Deloukas ·  
P. Whittaker · A. J. McKnight · J. Morrison ·  
A. J. Boulton · A. G. Demaine · S. M. Marshall ·  
B. A. Millward · S. M. Thomas · G. C. Viberti ·  
J. D. Walker · D. Sadlier · A. P. Maxwell · S. C. Bain

Received: 7 May 2008 / Accepted: 31 July 2008 / Published online: 5 September 2008  
© Springer-Verlag 2008

### Abstract

**Aims/hypothesis** Diabetic nephropathy, characterised by persistent proteinuria, hypertension and progressive kidney failure, affects a subset of susceptible individuals with diabetes. It is also a leading cause of end-stage renal disease (ESRD). Non-synonymous (ns) single nucleotide polymorphisms (SNPs) have been reported to contribute to genetic susceptibility in both monogenic disorders and common complex diseases. The objective of this study was to investigate whether nsSNPs are involved in susceptibility to diabetic nephropathy using a case-control design.

**Methods** White type 1 diabetic patients with (cases) and without (controls) nephropathy from eight centres in the UK and Ireland were genotyped for a selected subset of nsSNPs using Illumina's GoldenGate BeadArray assay. A  $\chi^2$  test for trend, stratified by centre, was used to assess differences in genotype distribution between cases and

controls. Genomic control was used to adjust for possible inflation of test statistics, and the False Discovery Rate method was used to account for multiple testing.

**Results** We assessed 1,111 nsSNPs for association with diabetic nephropathy in 1,711 individuals with type 1 diabetes (894 cases, 817 controls). A number of SNPs demonstrated a significant difference in genotype distribution between groups before but not after correction for multiple testing. Furthermore, neither subgroup analysis (diabetic nephropathy with ESRD or diabetic nephropathy without ESRD) nor stratification by duration of diabetes revealed any significant differences between groups.

**Conclusions/interpretation** The nsSNPs investigated in this study do not appear to contribute significantly to the development of diabetic nephropathy in patients with type 1 diabetes.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00125-008-1142-5) contains supplementary material, which is available to authorised users.

D. A. Savage (✉) · A. J. McKnight · A. P. Maxwell  
Nephrology Research Laboratory, Queen's University Belfast,  
c/o Regional Genetics Centre,  
Floor A, Tower Block, Belfast City Hospital, Lisburn Road,  
Belfast BT9 7AB, Northern Ireland, UK  
e-mail: d.savage@qub.ac.uk

C. C. Patterson  
Department of Epidemiology and Public Health,  
School of Medicine and Dentistry, Queen's University Belfast,  
Belfast, UK

P. Deloukas · P. Whittaker · J. Morrison  
Wellcome Trust Sanger Institute,  
Cambridge, UK

A. J. Boulton  
Manchester Royal Infirmary, University Department of Medicine,  
Manchester, UK

A. G. Demaine · B. A. Millward  
Molecular Medicine Research Group, Peninsula Medical School,  
Universities of Exeter and Plymouth,  
Plymouth, UK

S. M. Marshall  
Department of Diabetes and Metabolism,  
University of Newcastle upon Tyne,  
Newcastle upon Tyne, UK

**Keywords** Diabetic nephropathy · Genetic association study · Genetics of Kidneys in Diabetes · Genetic susceptibility · Non-synonymous single nucleotide polymorphisms

### Abbreviations

ESRD	end-stage renal disease
GoKinD	Genetics of Kidneys in Diabetes
MAF	minor allele frequency
ns	non-synonymous
SNP	single nucleotide polymorphism

### Introduction

Diabetic nephropathy is caused by hyperglycaemia-induced renal injury in genetically predisposed individuals. Patients affected with diabetic nephropathy have persistent proteinuria, hypertension and a declining glomerular filtration rate. Diabetic nephropathy is now the leading cause of end-stage renal disease (ESRD) and affected patients have a significantly increased cardiovascular mortality rate.

Despite convincing support for the existence of susceptibility and protective genes in diabetic nephropathy, genetic association studies have not yet identified consistently replicated associations in well-powered studies with large sample sizes [1]. Non-synonymous (ns) single nucleotide polymorphisms (SNPs) resulting in amino acid changes may directly affect protein structure and/or function. Evidence of association between nsSNPs and common multifactorial diseases has recently been reported [2–4]. It is also noteworthy that missense and nonsense

mutations are the predominant type of mutation (~60%) reported in monogenic disorders [5]. We therefore reasoned that nsSNPs could directly contribute to genetic susceptibility to diabetic nephropathy. In the present study we tested our hypothesis by using Illumina's GoldenGate BeadArray assay (<http://www.illumina.com/pages.ilmn?ID=11>; [6]) to assess the largest panel of nsSNPs to date for association with diabetic nephropathy in over 1,700 individuals with type 1 diabetes.

### Methods

**Participants** The type 1 diabetic participants with (cases) and without (controls) diabetic nephropathy who were recruited in this study were white and had parents and grandparents born in the UK or Ireland. All participants were diagnosed with type 1 diabetes before the age of 35 years and required insulin from time of diagnosis. Cases and controls were derived from the Warren 3 ([http://www.diabetes.org.uk/en/About\\_us/News\\_Landing\\_Page/2865/](http://www.diabetes.org.uk/en/About_us/News_Landing_Page/2865/)) and Genetics of Kidneys in Diabetes (GoKinD; <http://www.gokind.org>) collections recruited from seven centres in the UK (Belfast, Birmingham, Edinburgh, London, Manchester, Newcastle and Plymouth), as well as additional cases and controls from both the Belfast centre and a single centre (Dublin) in the Republic of Ireland.

Cases with nephropathy were defined by development of persistent proteinuria (>0.5 g protein/24 h) at least 10 years after diagnosis of diabetes, hypertension (BP >135/85 mmHg and/or treatment with antihypertensive agents) and presence of diabetic retinopathy. In contrast, controls were defined as patients with type 1 diabetes of at least 15 years' duration, but with urinary albumin in the normal range and not receiving antihypertensive treatment. Patients with microalbuminuria were excluded from both groups. Ethical approval was obtained from the appropriate research ethics committees, and written informed consent was obtained from all participants prior to conducting this study.

**Non-synonymous SNP selection and genotyping** The nsSNPs surveyed were derived from candidate genes (1) identified following extensive literature reviews, (2) from reports on transcription profiling, (3) inferred by function or (4) known to be expressed in the kidney.

Bioinformatic interrogation was applied to the gene list to identify nsSNPs with a minor allele frequency (MAF)  $\geq 0.05$  in the HapMap CEPH (Utah residents with ancestry in northern and western Europe) collection (<http://www.hapmap.org>) to be considered for inclusion in a custom Illumina GoldenGate array comprising nsSNPs with an assay design score >0.4. Genotyping of SNPs was

S. M. Thomas  
Department of Diabetes and Endocrinology,  
Guy's and St Thomas' Hospitals,  
London, UK

G. C. Viberti  
Unit for Metabolic Medicine, KCL School of Medicine,  
University of London,  
London, UK

J. D. Walker  
Medical Unit, St John's Hospital,  
Livingston, UK

D. Sadlier  
Conway Institute, University College Dublin,  
Dublin, Ireland

S. C. Bain  
Department of Medicine, University of Wales,  
Swansea, UK

performed on DNA samples (250 ng) in 96-well plates using the Illumina GoldenGate assay as described elsewhere [6]. Each plate contained a combination of case and control DNA samples, one non-template control and duplicates of three different samples in positions unique to a given plate. Genotyping data were then imported into Illumina's Gencall data analysis software, which employs algorithms for clustering, calling and scoring of genotypes [6]. Assignment of GenCall scores permits ranking of the data and where appropriate assists in the identification and subsequent exclusion of failed DNAs, genotypes and/or loci; only nsSNPs with GenCall scores >0.25 were considered acceptable in this study.

**Statistical analyses** The clinical characteristics of the patient groups were compared by the independent samples *t* test. The Cochran–Armitage  $\chi^2$  test for trend, stratified by centre, was used to identify differences in genotype distribution between cases and controls [7]. Possible inflation of the  $\chi^2$  test statistics was assessed by a quantile–quantile plot of observed  $\chi^2$  test statistics against expected order statistics [8]. The lowest 90% of test statistics were averaged and the result was divided by the average of the lowest 90% of expected order statistics from a  $\chi^2$  distribution to derive an inflation factor,  $\lambda$ . The False Discovery Rate method was used to account for multiple testing [9].

## Results

The age at diagnosis of diabetes was not significantly different between the cases and controls, with a mean duration of diabetes of at least 27 years for each group (Table 1). As expected the average blood pressure was higher in cases than in controls, consistent with the renal disease phenotype and despite the concurrent use of antihypertensive drug treatment by patients with diabetic nephropathy (Table 1). Within the diabetic nephropathy group, a subset of 30% ( $n=265$ ) had ESRD that was being managed by dialysis or renal transplantation. As reported for most national registries of renal disease, men were more likely to have diabetic nephropathy than women [10].

The nsSNPs with individual GenCall scores >0.25 were filtered on the basis of our criteria for call rates of at least 95% for cases and controls (of those meeting these criteria, 88% had call rates of at least 99%) and a differential in data completeness no greater than 2.0% between cases and controls. Further criteria were then applied that excluded nsSNPs with more than one duplicate discrepancy ( $n=16$ ), a MAF <1% ( $n=16$ ; nine of which were monomorphic) and departures from Hardy–Weinberg equilibrium ( $p<0.001$ ) in

**Table 1** Clinical characteristics of cases and controls

Clinical characteristic	Cases ( $n=894$ )	Controls ( $n=817$ )
Male sex, <sup>a</sup> $n$ (%)	510 (57)	359 (44)
Age at diagnosis of diabetes (years)	14.6 (7.6)	14.7 (7.8)
Duration of diabetes (years) <sup>b</sup>	33.5 (9.4)	27.7 (8.7)
HbA <sub>1c</sub> (%) <sup>c</sup>	8.84 (1.79)	7.86 (1.74)
SBP (mmHg) <sup>c</sup>	144.7 (21.0)	124.1 (14.1)
DBP (mmHg) <sup>c</sup>	81.6 (11.6)	75.1 (7.6)
ESRD, $n$ (%)	265 (30)	Not applicable

Unless otherwise stated, values are mean (SD)

<sup>a</sup> Designation of sex missing for eight participants

<sup>b</sup> Duration of diabetes calculated from date of diagnosis to date of recruitment

<sup>c</sup> Values are average of three most recent measurements prior to recruitment

DBP, diastolic blood pressure; SBP, systolic blood pressure

$p<0.001$  for all comparisons, except age at diagnosis of diabetes (non-significant)

either cases or controls ( $n=17$ ). This resulted in high-quality genotyping data for 1,111 nsSNPs in 894 cases and 817 controls (Electronic Supplementary Material [ESM] Table 1).

These 1,111 SNPs, covering 1,021 unique genes (some SNPs overlap more than one gene, see ESM Table 2), were first compared in cases and controls using the unstratified  $\chi^2$  test for trend, and a factor  $\lambda=1.121$  (12.1% inflation, 95% CI 4.9–19.3) was obtained. Stratifying the  $\chi^2$  statistics by centre reduced this inflation factor to  $\lambda=1.079$  (7.9% inflation, 95% CI 1.1–14.7) (ESM Fig. 1). This indicated that the observed test statistics were 7.9% greater than would have been anticipated from the expected distribution of test statistics; the test statistics were adjusted accordingly. Before making allowance for the large number of tests ( $n=1,111$ ) performed, 49 nsSNPs generated results that were significant at the  $p<0.05$  level (55.5 expected by chance alone) and 13 nsSNPs gave results that were significant at the  $p<0.01$  level (11.1 expected by chance alone) (Table 2). None of the  $p$  values for these nsSNPs fell below the False Discovery Rate cut-off for significance when taking into account the number of tests (Table 2). We then analysed the data in subgroups defined by the presence of ESRD, i.e. diabetic nephropathy cases with ESRD ( $n=265$ ) vs controls ( $n=817$ ) and diabetic nephropathy cases without ESRD ( $n=629$ ) vs controls ( $n=817$ ). These additional subgroup analyses also revealed no significant differences in allele frequencies between groups.

Given the relevance of diabetes duration in identifying susceptibility factors of moderate effect [11], the data were

**Table 2** Distribution of allele frequencies (%) for 13 top-ranked nsSNPs (uncorrected  $p < 0.01$ ) in 894 cases and 817 controls

Gene symbol	Gene name	SNP	Chromosome	Base change	MAF (%)		$\chi^2$ statistic (1 df) stratified by centre	$p$ uncorrected adjusted for inflation	False Discovery Rate cut-off for $p$ uncorrected
					Cases	Controls			
<i>GML</i>	GPI anchored molecule-like protein	rs3764795	8q24.3	C>T	3.2	5.6	14.52	0.00024	0.00005
<i>IMPG1</i>	Interphotoreceptor matrix proteoglycan 1	rs3734311	6q14.2–q15	C>G	40.4	36.5	10.81	0.00155	0.00009
<i>RPP21</i>	Ribonuclease P protein subunit p21	rs6986	6p21.3	C>G	16.7	20.2	10.40	0.00190	0.00014
<i>TEX15</i>	Testis expressed 15	rs323346	8p12	C>T	15.3	20.0	10.40	0.00190	0.00018
<i>TEX15</i>	Testis expressed 15	rs323347	8p12	A>G	15.4	19.9	9.72	0.00269	0.00023
<i>BPI</i>	Bactericidal/permeability-increasing protein	rs5743509	20q11.23–q12	C>T	7.3	5.3	8.95	0.00398	0.00027
<i>DNAH5</i>	Dynein, axonemal, heavy chain 5	rs4701997	5p15.2	A>T	51.6	46.8	8.20	0.00582	0.00032
<i>TSPAN8</i>	Tetraspanin 8	rs3763978	12q14.1–q21.1	C>G	38.5	42.2	7.92	0.00673	0.00036
<i>ITIH2</i>	Inter-alpha (globulin) inhibitor H2	rs7084817	10p15	C>G	6.3	4.4	7.90	0.00681	0.00041
<i>ERCC6</i>	Excision repair cross-complementing rodent repair deficiency, complementation group 6	rs2228527	10q11.23	A>G	20.9	16.5	7.71	0.00752	0.00045
<i>TBC1D1</i>	TBC1 (tre-2/USP6, BUB2, cdc16) domain family, member 1	rs10501	4p14	G>T	39.8	44.5	7.53	0.00825	0.00050
<i>FARP2</i>	FERM, RhoGEF and pleckstrin domain protein 2	rs757978	2q37.3	A>G	9.0	11.8	7.51	0.00833	0.00054
<i>FECH</i>	Ferrochelatase (protoporphyrin)	rs1041951	18q21.3	C>T	14.4	18.2	7.32	0.00918	0.00059

also analysed following stratification according to duration of diabetes (<25, 25–34 and  $\geq 35$  years). This stratification also failed to identify an association between any of the nsSNPs and diabetic nephropathy irrespective of diabetes duration.

## Discussion

Statistical analysis of the 1,111 nsSNPs in 894 cases and 817 controls revealed no significant associations with diabetic nephropathy after correction for multiple genotype comparisons. Further subgroup analyses examining the influence of diabetes duration or renal status (diabetic nephropathy with ESRD or diabetic nephropathy without ESRD) also showed no significant associations between genotypes and diabetic nephropathy. However, individually a number of SNPs had statistically significant associations before correction for multiple testing (Table 2) and may warrant further follow-up, even though a large number of false positive findings are likely.

Several explanations might account for the lack of statistically significant associations between the nsSNPs analysed in the present study and diabetic nephropathy. One possibility is that we failed to identify an association due to a false negative outcome(s) (i.e. type 2 errors). The sample sizes employed in our analysis (894 cases, 817 controls) gave at least 80% power to detect, at  $p=5.0 \times 10^{-5}$  ( $=0.05/1,000$ ), a disease effect at a typed locus with odds ratio equal to 1.5 for a MAF of no less than 20%; however, for the same MAF we only had 3% and 18% power to detect associations for odds ratios of 1.2 and 1.3 respectively.

Another possibility for the lack of association is technical errors in genotyping. However, this is extremely unlikely, as we employed the highly accurate (>99%) and reproducible (>99%) GoldenGate assay from Illumina [6], in conjunction with strict quality control criteria.

In conclusion, it is likely that the nsSNPs analysed in this case–control study do not contribute significantly to a predisposition to diabetic nephropathy. Nevertheless 1,111 nsSNPs represent only a small proportion of nsSNPs [3], and the hypothesis, that nsSNPs may be involved in susceptibility to diabetic nephropathy remains plausible.

To comprehensively test this hypothesis, a much larger nsSNP panel would have to be fully evaluated for association with diabetic nephropathy.

**Acknowledgements** We thank Diabetes UK for supporting this project, and the Juvenile Diabetes Research Foundation and Diabetes UK for access to DNA samples from the GoKinD UK and Warren 3 collections respectively. We also thank T. Dibling and C. Hind at the Wellcome Trust Sanger Institute for genotyping. The work of P. Whittaker, J. Morrison and P. Deloukas was supported by the Wellcome Trust.

**Duality of interest** The authors declare that there is no duality of interest associated with this manuscript.

## References

1. Savage DA, Bain SC, McKnight AJ, Maxwell AP (2007) Gene discovery in diabetic nephropathy. *Curr Diab Rep* 7:139–145
2. Smyth DJ, Cooper JD, Bailey R et al (2006) A genome-wide association study of nonsynonymous SNPs identifies a type 1 diabetes locus in the interferon-induced helicase (IFIH1) region. *Nat Genet* 38:617–619
3. Wellcome Trust Case Control Consortium & The Australo–Anglo–American Spondylitis Consortium (2007) Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 39:1329–1337
4. Hampe J, Franke A, Rosenstiel P et al (2007) A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet* 39:207–211
5. Botstein D, Risch N (2003) Discovering genotypes underlying human phenotypes: past successes for Mendelian disease, future approaches for complex disease. *Nat Genet* 33:228–237
6. Fan JB, Oliphant A, Shen R et al (2003) Highly parallel SNP genotyping. *Cold Spring Harb Symp Quant Biol* 68:69–78
7. Schlesselman JJ (1982) Basic methods of analysis. In: Schlesselman JJ (ed) Case-control studies. Oxford University Press, New York, pp 171–226
8. Clayton DG, Walker NM, Smyth DJ et al (2005) Population structure, differential bias and genomic control in a large-scale, case-control association study. *Nat Genet* 37:1243–1246
9. Sabatti C, Service S, Freimer N (2003) False discovery rate in linkage and association genome screens for complex disorders. *Genetics* 164:829–833
10. Ansell D, Feehally J, Feest TG, Tomson C, Williams AJ, Warwick G (2007) UK Renal Registry Report 2007. UK Renal Registry, Bristol, UK
11. Rogus JJ, Warram JH, Krolewski AS (2002) Genetic studies of late diabetic complications: the overlooked importance of diabetes duration before complication onset. *Diabetes* 51:1655–1662