Diabetologia (2015) 58:543–548
DOI 10.1007/s00125-014-3459-6

SHORT COMMUNICATION

SORBS1 gene, a new candidate for diabetic nephropathy: results from a multi-stage genome-wide association study in patients with type 1 diabetes

Marine Germain · Marcus G. Pezzolesi · Niina Sandholm · Amy J. McKnight · Katalin Susztak · Maria Lajer · Carol Forsblom · Michel Marre · Hans-Henrik Parving · Peter Rossing · Iiro Toppila · Jan Skupien · Ronan Roussel · Yi-An Ko · Nora Ledo · Lasse Folkersen · Mete Civelek · Alexander P. Maxwell · David-Alexandre Tregouet · Per-Henrik Groop · Lise Tarnow · Samy Hadjadj

Received: 6 August 2014 / Accepted: 30 October 2014 / Published online: 6 December 2014

Abstract
Aims/hypothesis The genetic determinants of diabetic nephropathy remain poorly understood. We aimed to identify novel susceptibility genes for diabetic nephropathy.
Methods We performed a genome-wide association study using 1000 Genomes-based imputation to compare type 1 diabetic nephropathy cases with proteinuria and with or without renal failure with control patients who have had diabetes for more than 15 years and no evidence of renal disease.
Results None of the single nucleotide polymorphisms (SNPs) tested in a discovery cohort composed of 683 cases and 779 controls reached genome-wide statistical significance. The 46 top hits ($p<10^{-5}$) were then sought for first-stage analysis in the Genetics of Kidneys in Diabetes US (US-GoKinD) study, an independent population of 820 cases and 885 controls. Two SNPs in strong linkage disequilibrium with each other and located in the *SORBS1* gene were consistently and significantly ($p<10^{-4}$) associated with diabetic nephropathy. The minor

Electronic supplementary material The online version of this article (doi:10.1007/s00125-014-3459-6) contains peer-reviewed but unedited supplementary material, which is available to authorised users.

M. Germain · D.-A. Tregouet
Sorbonne Université, UPMC Université Paris 06, UMR_S 1166, Team Genomics & Pathophysiology of Cardiovascular Diseases, Paris, France

M. Germain · D.-A. Tregouet
Inserm, UMR_S 1166, Team Genomics & Pathophysiology of Cardiovascular Diseases, Paris, France

M. Germain · D.-A. Tregouet
ICAN Institute for Cardiometabolism and Nutrition, 75013 Paris, France

M. G. Pezzolesi · J. Skupien
Research Division, Joslin Diabetes Center & Department of Medicine, Harvard Medical School, Boston, MA, USA

N. Sandholm · C. Forsblom · I. Toppila · P.-H. Groop
Folkhälsan Institute of Genetics, Folkhälsan Research Center, Biomedicum Helsinki, Helsinki, Finland

N. Sandholm · C. Forsblom · I. Toppila · P.-H. Groop
Division of Nephrology, Department of Medicine, Helsinki University Central Hospital, Helsinki, Finland

N. Sandholm
Department of Biomedical Engineering and Computational Science, Aalto University, Espoo, Finland

A. J. McKnight · A. P. Maxwell
Nephrology Research, Centre for Public Health, Queen's University of Belfast, Belfast, UK

K. Susztak · Y.-A. Ko · N. Ledo
Renal Electrolyte and Hypertension Division, University of Pennsylvania, Philadelphia, PA, USA

M. Lajer · P. Rossing · L. Tarnow
Steno Diabetes Center, Gentofte, Denmark

M. Marre · R. Roussel
Université Paris-Diderot, Sorbonne Paris-Cité, Paris, France

M. Marre · R. Roussel
Diabetologie, AP-HP, Hôpital Bichat, Paris, France

M. Marre · R. Roussel
Departement Hospitalo-Universitaire FIRE, Paris, France

rs1326934-C allele was less frequent in cases than in controls (0.34 vs 0.43) and was associated with a decreased risk for diabetic nephropathy (OR 0.70; 95% CI 0.60, 0.82). However, this association was not observed in a second stage with two additional diabetic nephropathy cohorts, the All Ireland-Warren 3-Genetics of Kidneys in Diabetes UK and Republic of Ireland (UK-ROI; p=0.15) and the Finnish Diabetic Nephropathy (FinnDiane; p=0.44) studies, totalling 2,142 cases and 2,494 controls. Altogether, the random-effect meta-analysed rs1326934-C allele OR for diabetic nephropathy was 0.83 (95% CI 0.72, 0.96; p=0.009).

Conclusions/interpretation These data suggest that *SORBS1* might be a gene involved in diabetic nephropathy.

Keywords Diabetic nephropathy · GWAS · Kidney · Sorbin · Type 1 diabetes

Abbreviations

1000G	1000 Genomes
CA	Cochran–Armitage
eGFR	Estimated glomerular filtration rate
FinnDiane	Finnish Diabetic Nephropathy study
GENIE	Genetics of Nephropathy an International Effort consortium
GWAS	Genome-wide association study
SNP	Single nucleotide polymorphism
SORBS1	Sorbin and SH3 domain-containing protein 1
UK-ROI	All Ireland-Warren 3-Genetics of Kidneys in Diabetes UK and Republic of Ireland study
US-GoKinD	Genetics of Kidneys in Diabetes US study

Introduction

Diabetic nephropathy is a frequent condition affecting a substantial proportion of diabetic patients, and is a leading cause of end-stage renal disease. The determinants of diabetic nephropathy are complex and include genetic factors, as supported by a strong familial aggregation of diabetic nephropathy [1].

The precise nature of the genetic burden involved in diabetic nephropathy remains poorly understood. An alternative approach to candidate gene studies is to perform unbiased genome-wide association studies (GWAS). To date, only a small number of GWAS for diabetic nephropathy have been performed. Recently, the Genetics of Nephropathy an International Effort (GENIE) consortium identified two loci associated with a population attributable risk of 0.5–10.5% [2]. Therefore, additional studies are needed to discover novel loci associated with diabetic nephropathy.

The aim of our research was to conduct a GWAS in a large-scale case–control study including French and Danish patients, based on the latest 1000 Genomes (1000G) imputation techniques, and to replicate our findings in available cohorts of patients with type 1 diabetes from European ancestry.

Methods

Studied participants were of white European origin [2–4]. Briefly, patients were classified as cases or controls at the time of recruitment in all but the Finnish Diabetic Nephropathy (FinnDiane) study, where they were selected among a population-based consecutive collection of people with type

M. Marre · R. Roussel
Inserm U1138, Centre de Recherche des Cordeliers, Paris, France

H.-H. Parving · P. Rossing · L. Tarnow
Faculty of Health Sciences, University of Aarhus, Aarhus, Denmark

H.-H. Parving
Department of Medical Endocrinology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

P. Rossing
Novo Nordisk Foundation, Center for Basic Metabolic Research, Faculty of Health, University of Copenhagen, Copenhagen, Denmark

J. Skupien
Department of Metabolic Diseases, Jagiellonian University Medical College, Krakow, Poland

L. Folkersen
Department of PharmacoGenetics, Novo Nordisk Park 9.1.21, Copenhagen, Denmark

L. Folkersen
Atherosclerosis Research Unit, Karolinska Hospital, Stockholm, Sweden

M. Civelek
Department of Medicine, University of California Los Angeles, Los Angeles, CA, USA

A. P. Maxwell
Regional Nephrology Unit, Belfast City Hospital, Belfast, UK

P.-H. Groop
Baker IDI, Heart and Diabetes Institute, Melbourne, VIC, Australia

L. Tarnow
Nordsjaellands Hospital, Hilleroed, Denmark

S. Hadjadj
Université de Poitiers, UFR Médecine Pharmacie, Centre d'Investigation Clinique, Poitiers, France

S. Hadjadj (✉)
Service d'Endocrinologie-Diabetologie and Centre d'Investigation Clinique, CHU de Poitiers, BP 577, 86021 Poitiers Cedex, France
e-mail: samy.hadjadj@gmail.com

S. Hadjadj
Inserm, CIC1402 and U1082, Poitiers, France

1 diabetes, and the nephropathy status of the patients was defined based on the latest available data. All cohorts were multi-centric except the single-centre LEwis blood type and Angiotensin Converting Enzyme (LEACE) study. Local ethics committees approved the study protocols. All participants gave written informed consent.

Phenotypic determination

Type 1 diabetes Type 1 diabetes was diagnosed using American Diabetes Association criteria: rapid definitive insulin requirement (within 1 year of diagnosis) and age at diabetes onset below 31 or 36 years, differing according to each cohort (Electronic supplementary material [ESM] Table 1).

Diabetic nephropathy status Phenotype determination was based on the clinical criteria proposed by the Genetics of Kidneys in Diabetes US (US-GoKinD) study: cases had proteinuria with or without renal failure. Proteinuria was uniformly defined on at least two of three sterile urine collections (ESM Table 1). Controls had long-term duration of diabetes (over 15 years), normo-albuminuria without renal failure and were not prescribed drugs that block the renin-angiotensin system. Microalbuminuric patients were not considered for the current analyses.

Study organisation

The general research strategy adopted in this project is summarised in ESM Fig. 1. Brief descriptions of the studied cohorts and patients can be found in ESM Tables 1 and 2.

We ultimately assessed expression of the identified gene in a post-hoc analysis of micro-dissected glomerular and tubule samples [5].

Genotype, quality control and imputation

Genotyping platforms, quality control and imputation are detailed in ESM Methods.

Statistical analysis

A logistic regression analysis was conducted to evaluate the association of each imputed single nucleotide polymorphism (SNP) with diabetic nephropathy under an additive genetic model (allele dosage was used as the covariate for characterising the tested SNP). Analyses were adjusted for age, sex and the first four principal components derived from the genome-wide genotyped SNPs computed by the Eigenstrat programme (http://www.hsph.harvard.edu/alkes-price/software/) [6]. Association analyses were performed using the mach2dat (version 1.0.19) software (http://www.unc.edu/~yunmli/software.html).

All SNPs with suggestive evidence of an association with diabetic nephropathy ($p<10^{-5}$) were moved forward in the first-stage study. The centre- and sex-adjusted association of these SNPs with diabetic nephropathy was investigated using the SNPTEST programme (https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html#introduction).

In the second-stage samples, the association of candidate SNP with diabetic nephropathy was tested using the Cochran–Armitage (CA) trend test. ORs derived from application of the CA test in the four case–control samples were finally combined into a random-effect model based meta-analysis using the GWAMA programme (http://www.well.ox.ac.uk/gwama/download.shtml) [7].

Results

Discovery analysis

None of the 11,133,962 tested SNPs achieved the statistical threshold of 5×10^{-8} for declaring genome-wide significance (ESM Figs 2 and 3). All SNPs with suggestive evidence for association at $p<10^{-5}$ ($n=46$) were selected for further association testing in a first independent sample of 820 cases and 885 controls in the US-GoKinD study (ESM Fig. 1).

First-stage study

After Bonferroni correction, three SNPs demonstrated significant association with diabetic nephropathy: rs11188343 ($p=9.06\times10^{-5}$), rs1326934 ($p=9.85\times10^{-5}$) and rs4917695 ($p=1.27\times10^{-4}$) (ESM Table 3). These SNPs mapped to the *SORBS1* gene and were in near complete linkage disequilibrium ($r^2\sim1$).

In the discovery GWAS, the minor rs1326934-C allele was less frequent in cases than in controls (0.34 vs 0.43) and was associated with a decreased risk for diabetic nephropathy (OR 0.70 [95% CI 0.60, 0.82]; $p=7.87\times10^{-6}$). In the first-stage population, a consistent association was observed with the rs1326934-C allele at a lower frequency in cases (combined statistical evidence for association of $p=3.52\times10^{-9}$) compared with controls (Table 1).

Second-stage study

To validate further the association of rs13626934 with diabetic nephropathy, we studied this SNP in two additional case–control studies composed of 823 cases and 903 controls (All Ireland-Warren 3-Genetics of Kidneys in Diabetes UK and Republic of Ireland [UK-ROI]) and 1,335 cases and 1,633 controls (FinnDiane; Table 1). In UK-ROI, the same trend of association, though not significant ($p=0.15$), was observed: the

Table 1 Association of *SORBS1* rs1326934 with diabetic nephropathy in the four studied cohorts

	Cases				Controls				
	TT	TC	CC	MAF	TT	TC	CC	MAF	p^a
Discovery cohort	291	314	78	0.344	247	398	134	0.427	$2.94\ 10^{-6}$
First-stage analysis	289	407	124	0.399	260	428	197	0.464	$1.32\ 10^{-4}$
Second-stage analysis									
UK-ROI[b]	263	426	134	0.422	269	462	170	0.445	0.155
FinnDiane	681	539	99	0.279	819	640	132	0.284	0.756

[a] CA trend test for association

[b] Two patients with low call-rate for rs1326934 in controls from the UK-ROI cohort

MAF, minor allele frequency

rs13626934-C allele was less frequent in cases than in controls (0.42 vs 0.45). However, no association was observed in the FinnDiane population (p=0.44) where the allele frequency of the C allele was similar in cases and controls (0.27 vs 0.28).

Combining the results of the four studies into a fixed-effect meta-analysis led to an overall OR for diabetic nephropathy of 0.84 (95% CI 0.79, 0.90; $p=5.69\times10^{-7}$). However, this association was statistically heterogeneous across the four samples ($p=3.94\times10^{-3}$) and the random-effect meta-analysed OR for diabetic nephropathy was, therefore, 0.83 (95% CI 0.72, 0.96; p=0.009). When FinnDiane patients were excluded the meta-analysed OR was 0.795 (95% CI 0.733, 0.861; $p=2.40\times10^{-8}$).

Gene expression analysis

SORBS1 gene overexpression was observed in tubules of type 2 diabetes patients compared with controls (Fig. 1a, ESM Tables 4 and 5). We also observed a significant inverse correlation between *SORBS1* expression and estimated glomerular filtration rate (eGFR) ($R=-0.493$; $p=1.44\times10^{-3}$) (Fig. 1b). Among the other genes mapped within 250 kb of rs1326934, *SORBS1* demonstrated the strongest correlation with eGFR in both control and diabetic nephropathy tubule samples (Fig. 1c). Furthermore, at a protein level, sorbin and SH3 domain-containing protein 1(SORBS1) was also highly expressed in renal tubules and medium expressed in glomeruli (www.proteinatlas.org).

Discussion

In this study, a multi-stage based GWAS looked for novel susceptibility genes associated with diabetic nephropathy in patients with type 1 diabetes. Although we did not detect any new loci passing genome-wide significance, we observed promising evidence for an association between *SORBS1* rs1326934 (or any SNP in complete linkage disequilibrium with it) and diabetic nephropathy in three of four of the studied populations. Given the frequency of the risk allele observed in our control populations (~0.55) and the associated OR for diabetic nephropathy (~1.25), the population attributable risk of the identified polymorphisms would be 12%.

No association was observed in the Finish population, which exhibited a marked difference in allele frequency at the candidate SNP compared with the three other studied populations. We further interrogated databases with genome-wide genotype and gene expression but did not observe any association in monocytes, macrophages, hepatocytes, adipocytes [8] or endothelial cells [9] with *SORBS1* gene expression. Nevertheless, the *SORBS1* rs1188343 (in complete linkage disequilibrium with the rs1326934) was predicted in the RegulomeDB database (http://regulome.stanford.edu) to be located at a potential binding site for either *RFX3* or histone interaction.

The sorbin protein, coded by *SORBS1*, was found to be differentially upregulated in glomeruli of rats with diabetic nephropathy compared with rats without diabetic nephropathy [10]. High tubular and moderate glomerular expression of sorbin protein was observed in human kidney samples. These findings from type 2 diabetes patients provide additional support for the association of *SORBS1* with diabetic nephropathy. Gene expression changes of *SORBS1* were easier to detect in tubules, as *SORBS1* has a higher tubular expression. Although *SORBS1* expression was significantly upregulated only in tubules, we cannot exclude the importance of glomerular *SORBS1*. Diabetic nephropathy not only involves glomeruli but also tubules [11], and genes identified through GWAS are likely to impact both renal structures.

The function of sorbin is not fully established but we speculate it plays a key role in several processes involved in diabetic nephropathy, including insulin resistance and cytoskeleton architecture. Sorbin acts in the genesis of stress fibres and might, therefore, be involved in podocyte alterations of the slit diaphragm barrier.

The lack of homogenous replicated associations in our work is unlikely to be due to clinical heterogeneity of the studied

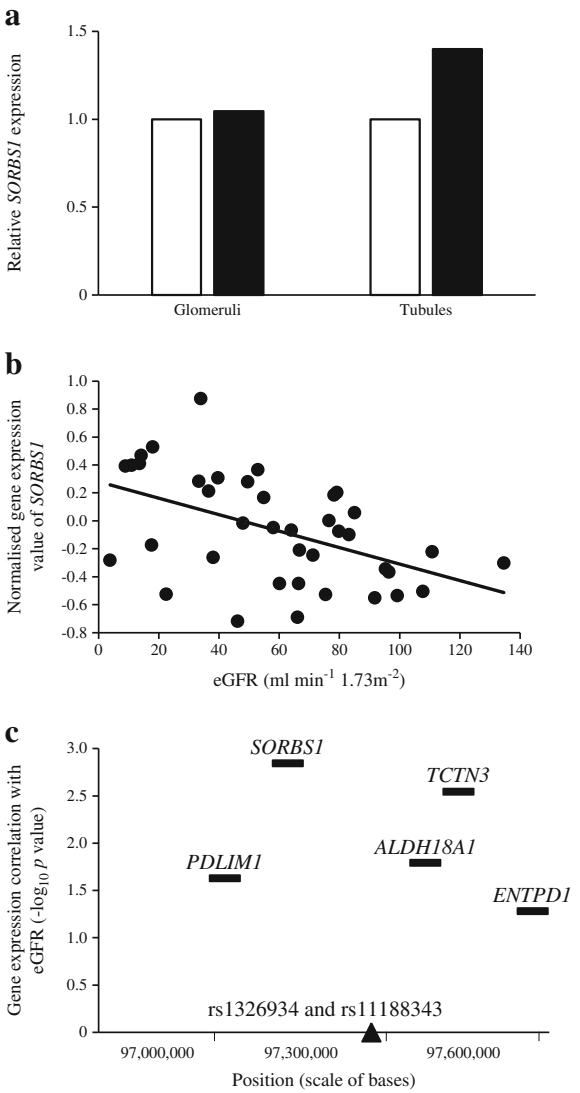

Fig. 1 Expression profile of *SORBS1* in diabetic nephropathy. (**a**) Tubular expression of *SORBS1* is significantly upregulated in diabetic kidney disease (black bars) compared with control samples (white bars; $p=0.0006$). (**b**) Each dot represents transcript levels and eGFR values from a single kidney sample. The line represents the fitted linear correlation values. (**c**) The genomic position of each gene on chromosome 10 (q24.1) in the 500 kb vicinity of rs1326934 and rs11188343 loci (triangle) is shown. Not only the *SORBS1* transcript, but also other transcripts, correlate with renal function in the vicinity of rs1326934 and rs11188343 loci

populations as the same definitions for type 1 diabetes and diabetic nephropathy were used. Conversely, the different patterns of allelic heterogeneity and linkage disequilibrium across European populations may explain the lack of replication of the *SORBS1* signal in the Finnish population. Characterising the exact variability at the *SORBS1* locus is needed to validate *SORBS1* as a new susceptibility gene for diabetic nephropathy and to identify the disease-associated functional variant(s).

The main limitation of the present study is its design. We adopted a multi-stage strategy using all available GWAS resources imputed for 1000G reference dataset at the time this work was launched. A more powerful approach would have been to conduct a comprehensive meta-analysis of the four populations. An international initiative has been set up to overcome this limitation. Another caveat is that the diabetic nephropathy phenotype was defined by the presence of proteinuria, regardless of renal function, and genetic susceptibility to proteinuria might differ from genetic predisposition to renal failure. However, the diabetic nephropathy cohorts we used all had this combined diabetic nephropathy phenotype permitting international collaboration.

In conclusion, our study provides preliminary support for *SORBS1* as a new susceptibility gene for diabetic nephropathy.

Acknowledgements We would like to thank all the type 1 diabetes patients who took part in the different cohorts considered for the current work. The respective acknowledgements for individual cohorts are detailed in ESM Acknowledgements. H. Nickerson (JDRF, New York, NY, USA), M. Harden (Broad Institute, Cambridge, MA, USA) and Jeremy Bonassies (Inserm Aquitaine, Bordeaux, France) are acknowledged here for their administrative support in organising the work. The CHU Poitiers biobanking facility (CRB Poitiers, BB0033-00068) is acknowledged for handling biological samples used in the present work. The following staff are specially acknowledged: S. Brishoual and E. Rogeon for organising the transfer of biological material to the Broad Institute. Statistical analyses of the discovery dataset were performed using the C2BIG computing cluster, funded by the Région Ile de France, Pierre and Marie Curie University, and the ICAN Institute for Cardiometabolism and Nutrition (ANR-10-IAHU-05).

Funding The genotyping of the French/Danish participants was supported by a grant from the JDRF (Spring Research 2010 Innovative Complications grant).

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

Contribution statement MG, MGP, NS, AJMK, JS, YK, NL performed analysis and edited the manuscript.

KS performed analysis, and partly wrote and edited the manuscript. ML, CF, MM, HHP, PR, IT, RR, LF, MC, APM, PHG, LT researched data and edited the manuscript. DAT performed analysis and wrote the manuscript. SH researched data and wrote the manuscript. All authors approved the final version to be published. SH is the guarantor of the work.

References

1. Harjutsalo V, Katoh S, Sarti C, Tajima N, Tuomilehto J (2004) Population-based assessment of familial clustering of diabetic nephropathy in type 1 diabetes. Diabetes 53:2449–2454
2. Sandholm N, Salem RM, McKnight AJ et al (2012) New susceptibility loci associated with kidney disease in type 1 diabetes. PLoS Genet 8:e1002921
3. Pezzolesi MG, Poznik GD, Mychaleckyj JC et al (2009) Genome-wide association scan for diabetic nephropathy susceptibility genes in type 1 diabetes. Diabetes 58:1403–1410
4. Tarnow L, Groop PH, Hadjadj S et al (2008) European rational approach for the genetics of diabetic complications–EURAGEDIC:

patient populations and strategy. Nephrol Dial Transplant 23:161–168
5. Woroniecka KI, Park AS, Mohtat D, Thomas DB, Pullman JM, Susztak K (2011) Transcriptome analysis of human diabetic kidney disease. Diabetes 60:2354–2369
6. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38:904–909
7. Magi R, Morris AP (2010) GWAMA: software for genome-wide association meta-analysis. BMC Bioinforma 11:288
8. Folkersen L, van't Hooft F, Chernogubova E et al (2010) Association of genetic risk variants with expression of proximal genes identifies novel susceptibility genes for cardiovascular disease. Circ Cardiovasc Genet 3:365–373
9. Erbilgin A, Civelek M, Romanoski CE et al (2013) Identification of CAD candidate genes in GWAS loci and their expression in vascular cells. J Lipid Res 54:1894–1905
10. Nakatani S, Kakehashi A, Ishimura E et al (2011) Targeted proteomics of isolated glomeruli from the kidneys of diabetic rats: sorbin and SH3 domain containing 2 is a novel protein associated with diabetic nephropathy. Exp Diabetes Res 2011:979354
11. Gilbert RE, Cooper ME (1999) The tubulointerstitium in progressive diabetic kidney disease: more than an aftermath of glomerular injury? Kidney Int 56:1627–1637