Diabetologia (2011) 54:1368–1374
DOI 10.1007/s00125-011-2063-2

ARTICLE

Replication of 13 genome-wide association (GWA)-validated risk variants for type 2 diabetes in Pakistani populations

S. D. Rees · M. Z. I. Hydrie · A. S. Shera · S. Kumar · J. P. O'Hare · A. H. Barnett · A. Basit · M. A. Kelly

Received: 22 October 2010 / Accepted: 5 January 2011 / Published online: 25 February 2011

Abstract
Aims/hypothesis Recent genome-wide association (GWA) studies and subsequent replication studies have greatly increased the number of validated type 2 diabetes susceptibility variants, but most of these have been conducted in European populations. Despite the high prevalence of the disease in South Asians, studies investigating GWA-validated type 2 diabetes risk variants in this ethnic group are limited. We investigated 30 single nucleotide polymorphisms (SNPs), predominantly derived from recent GWA studies, to determine if and to what extent these variants affect type 2 diabetes risk in two Punjabi populations, originating predominantly from the District of Mirpur, Pakistan.
Methods Thirty SNPs were genotyped in 1,678 participants with type 2 diabetes and 1,584 normoglycaemic control participants from two populations; one resident in the UK and one indigenous to the District of Mirpur.
Results SNPs in or near *PPARG*, *TCF7L2*, *FTO*, *CDKN2A/2B*, *HHEX/IDE*, *IGF2BP2*, *SLC30A8*, *KCNQ1*, *JAZF1*, *IRS1*, *KLF14*, *CHCHD9* and *DUSP9* displayed significant ($p<0.05$) associations with type 2 diabetes, with similar effect sizes to those seen in European populations. A constructed genetic risk score was associated with type 2 diabetes ($p=5.46\times10^{-12}$), BMI ($p=2.25\times10^{-4}$) and age at onset of diabetes ($p=0.002$).
Conclusions/interpretation We have demonstrated that 13 variants confer an increased risk of type 2 diabetes in our Pakistani populations; to our knowledge this is the first time that SNPs in or near *KCNQ1*, *JAZF1*, *IRS1*, *KLF14*, *CHCHD9* and *DUSP9* have been significantly associated with the disease in South Asians. Large-scale studies and meta-analyses of South Asian populations are needed to further confirm the effect of these variants in this ethnic group.

Keywords Genetic association study · Genetic risk score · Pakistani · Single nucleotide polymorphism · Type 2 diabetes

Abbreviations

AOD	Age at diagnosis/age at onset of type 2 diabetes
DGP	Diabetes Genetics in Pakistan
GRS	Genetic risk score
GWA	Genome-wide association

S. Kumar, J.P. O'Hare, A. H. Barnett on behalf of the UK Asian Diabetes Study.

Electronic supplementary material The online version of this article (doi:10.1007/s00125-011-2063-2) contains supplementary material, which is available to authorised users.

S. D. Rees (✉) · A. H. Barnett · M. A. Kelly
Diabetes Research Laboratory,
School of Clinical and Experimental Medicine,
College of Medical and Dental Sciences, The Medical School,
University of Birmingham,
Vincent Drive, Edgbaston,
Birmingham B15 2TT, UK
e-mail: s.d.rees@bham.ac.uk

M. Z. I. Hydrie · A. Basit
Baqai Institute of Diabetology and Endocrinology,
Karachi, Pakistan

A. S. Shera
Diabetic Association of Pakistan,
Karachi, Pakistan

S. Kumar · J. P. O'Hare
Warwick Medical School, University of Warwick,
Coventry, UK

A. H. Barnett
Heart of England NHS Foundation Trust,
Birmingham, UK

LD	Linkage disequilibrium
ROC	Receiver-operating characteristic
SNP	Single nucleotide polymorphism
UKADS	UK Asian Diabetes Study

Introduction

Type 2 diabetes is a major public health issue in the Indian subcontinent (India, Pakistan and Bangladesh), where it is predicted that the disease will affect approximately 76 million adults by 2025 [1]. A high prevalence of the disease is also observed in populations of South Asian ancestry living in other areas of the world. Although environmental and lifestyle factors undoubtedly contribute to the development of type 2 diabetes, they cannot fully explain the high prevalence of the disease in South Asians, suggesting that the disproportionate risk may be partly determined by the genetic makeup of this ethnic group.

Over the last 4 years, large-scale association studies, in particular genome-wide association (GWA) studies, have increased the number of accepted type 2 diabetes susceptibility variants to over 30 [2–11]. Despite this progress, studies investigating the genetic basis of the disease in South Asian populations are still relatively limited and, to date, only four have attempted to replicate any of the findings of recent GWA studies [12–15]. The largest of these studies, investigating two Indian Asian populations [14], demonstrated that several variants displayed greater effect sizes than those previously seen in European studies, a fact that the authors attributed to higher genotype penetrance in Indian populations.

In this study, we investigated 30 single nucleotide polymorphisms (SNPs), all of which have been robustly confirmed as type 2 diabetes susceptibility variants in GWA/replication studies, in two Punjabi populations originating predominantly from the District of Mirpur, Pakistan. Our aims were to determine if these SNPs were associated with type 2 diabetes and whether or not they displayed a greater effect size than that seen in Europeans.

Methods

Study populations The study was performed using case–control collections from two populations of Pakistani origin. Type 2 diabetes was diagnosed according to WHO criteria [16]. Age at diagnosis of type 2 diabetes was ≥30 years in 97.7% of cases and the minimum age at onset was 20 years. We therefore felt that accidental inclusion of participants with type 1 diabetes was unlikely. UK-resident participants with type 2 diabetes (n=857) were recruited from Birmingham and Coventry as part of the UK Asian Diabetes Study (UKADS) [17], while Pakistan-based participants with the disease (n=821) were recruited from the Mirpur region of Azad Kashmir as part of the Diabetes Genetics in Pakistan (DGP) study. Ethnically matched control participants (UK-based, n=417; Pakistan-based, n=1,167) were recruited from the same geographical areas through community screening. Normoglycaemia was defined by a fasting blood glucose <5.6 mmol/l (DGP), fasting plasma glucose <6.1 mmol/l and 2 h plasma glucose <7.8 mmol/l on a 75 g OGTT or random blood glucose <7 mmol/l (UKADS). A range of clinical and anthropometric data was collected (see electronic supplementary material [ESM] Table 1). All study participants were of Punjabi ancestry, confirmed over three generations, and originated predominantly from the District of Mirpur. Informed consent was obtained from all participants and the study was approved by the Birmingham East, North and Solihull Research Ethics Committee and the Institutional Review Board of Baqai Institute of Diabetology and Endocrinology.

Genotyping All samples (3262) were genotyped for 30 SNPs (see ESM Table 2) using either the KASPar (KBioscience, Hoddesdon, UK) or TaqMan (AppliedBiosystems, Warrington, UK) method. Five SNPs (those in/near *TCF7L2*, *HHEX/IDE*, *CDKAL1*, *CDKN2A/2B* and *IGF2BP2*) were genotyped using TaqMan assays in the UKADS samples and the KASPar method in the DGP samples. All other SNPs were genotyped using the KASPar method for all participants. The *FTO* rs9939609 SNP was genotyped using both methods in the UKADS samples and the concordance rate was 100%. Genotyping success rates for all SNPs were above 97%. Error rates calculated from 384 duplicate samples were below 0.6% for all SNPs, with the exception of rs1111875, for which the error rate was 1.4%.

Statistical analyses Statistical analyses were performed using STATA IC version 10.1 (Stata Corporation, College Station, TX, USA) or PLINK v1.07 (http://pngu.mgh.harvard.edu/~purcell/plink/) [18]. SNPs were analysed for deviation from Hardy–Weinberg equilibrium (HWE) within the normoglycaemic group of each population, using an exact test. SNP×diabetes associations were analysed using logistic regression including all participants, adjusting for age, BMI, sex and country of residence. BMI data were available for 3,078 participants. Adjusting for BMI made little difference to the SNP×diabetes associations (ESM Table 3); therefore, in order to maintain power the individual SNP results shown (Table 1) are not adjusted for BMI. For the X-chromosome variant rs5945326

Table 1 Association of 28 SNPs with type 2 diabetes in the UKADS/DGP study populations

Nearby gene	Chr	SNP	Risk/non-risk allele	RAF	Power	UKADS/DGP (n=3,262)		European studies	Indian Asian study[a]
						OR (95% CI)	p value	OR (95% CI)	OR (95% CI)
TCF7L2[b]	10	rs7903146	T/C	0.31	1.00	1.24 (1.11–1.38)	8.34×10^{-05}	1.37 (1.31–1.43)	1.89 (1.71–2.09)
IRS1[c]	2	rs2943641	C/T	0.73	0.96	1.24 (1.10–1.39)	3.92×10^{-04}	1.19 (1.13–1.25)	–
CDKN2A/2B[d]	9	rs10811661	T/C	0.88	0.29	1.31 (1.12–1.54)	9.89×10^{-04}	1.20 (1.14–1.25)	1.37 (1.18–1.59)
IGF2BP2[d]	3	rs4402960	T/G	0.38	0.78	1.18 (1.07–1.31)	1.43×10^{-03}	1.14 (1.11–1.18)	1.20 (1.09–1.33)
KCNQ1[e]	11	rs2237897	C/T	0.98	0.79	1.85 (1.23–2.79)	3.29×10^{-03}	1.36 (1.16–1.60)	–
JAZF1[f]	7	rs864745	T/C	0.67	0.57	1.16 (1.05–1.30)	6.01×10^{-03}	1.10 (1.07–1.13)	–
KLF14[g]	7	rs972283	G/A	0.57	0.35	1.14 (1.03–1.27)	9.66×10^{-03}	1.07 (1.05–1.10)	–
FTO[h]	16	rs9939609	A/T	0.30	0.87	1.15 (1.03–1.28)	0.012	1.15 (1.09–1.23)	–
HHEX/IDE[d]	10	rs1111875	C/T	0.43	0.69	1.13 (1.03–1.25)	0.013	1.13 (1.08–1.17)	1.27 (1.16–1.39)
DUSP9[g]	23	rs5945326	A/G	0.60	0.99	1.11 (1.02–1.21)	0.015	1.27 (1.18–1.37)	–
SLC30A8[d]	8	rs11558471[j]	A/G	0.74	0.66	1.13 (1.01–1.26)	0.041	1.12 (1.07–1.16)	1.34 (1.20–1.50)
PPARG[d]	3	rs1801282	C/G	0.87	0.40	1.17 (1.00–1.35)	0.044	1.14 (1.08–1.20)	1.37 (1.19–1.59)
CHCHD9[g]	9	rs13292136	C/T	0.89	0.21	1.18 (1.00–1.38)	0.046	1.11 (1.07–1.15)	–
WFS1[i]	4	rs10010131	G/A	0.68	0.74	1.08 (0.97–1.20)	0.176	1.12 (1.09–1.16)	–
TSPAN8/LGR5[f]	12	rs7961581	C/T	0.33	0.40	1.07 (0.96–1.19)	0.197	1.09 (1.06–1.12)	–
HNF1A[g]	12	rs7957197	T/A	0.94	0.20	1.14 (0.92–1.41)	0.226	1.07 (1.05–1.10)	–
CDKAL1[d]	6	rs10946398	C/A	0.26	0.75	1.07 (0.95–1.19)	0.265	1.12 (1.08–1.16)	1.18 (1.07–1.32)
KCNQ1[g]	11	rs231362	G/A	0.74	0.42	1.05 (0.94–1.18)	0.405	1.08 (1.06–1.10)	–
ZBED3[g]	5	rs4457053	G/A	0.24	0.34	1.05 (0.94–1.18)	0.405	1.08 (1.06–1.11)	–
ADAMTS9[f]	3	rs4607103	C/T	0.49	0.33	0.97 (0.88–1.08)	0.588	1.09 (1.06–1.12)	–
BCL11A[g]	2	rs243021	A/G	0.51	0.42	1.03 (0.93–1.14)	0.606	1.08 (1.06–1.10)	–
KCNJ11[d]	11	rs5219[k]	T/C	0.38	0.83	0.98 (0.88–1.08)	0.628	1.14 (1.10–1.19)	1.39 (1.26–1.54)
THADA[f]	2	rs7578597	T/C	0.86	0.50	0.97 (0.83–1.12)	0.636	1.15 (1.10–1.20)	–
CENTD2[g]	11	rs1552224	A/C	0.84	0.29	1.03 (0.90–1.18)	0.642	1.14 (1.11–1.17)	–
TP53INP1[g]	8	rs896854	T/C	0.41	0.24	0.98 (0.89–1.08)	0.675	1.06 (1.04–1.09)	–
NOTCH2[f]	1	rs10923931	T/G	0.17	0.68	1.02 (0.89–1.16)	0.777	1.13 (1.08–1.17)	–
PRC1[g]	15	rs8042680	A/C	0.61	0.28	0.99 (0.89–1.10)	0.881	1.07 (1.05–1.09)	–
CDC123/CAMK1D[f]	10	rs12779790	G/A	0.14	0.19	1.01 (0.88–1.17)	0.892	1.11 (1.07–1.14)	–

ORs are corrected for age, sex and population

Power calculated assuming a disease prevalence of 10%, ORs from European populations, risk allele frequencies from the Centre d'Etude du Polymorphisme (Utah residents with northern and western European ancestry) (CEU) HapMap and combined sample size of 3,262 individuals

Significant heterogeneity of ORs between the UKADS/DGP study and the Indian Asian study was observed for the *TCF7L2* and *KCNJ11* variants after correction for multiple testing (eight tests, $p_{corrected}=1.72\times10^{-7}$ and 8.88×10^{-6} respectively)

Indian Asian estimates taken from [a]Chauhan et al. [14]

European estimates taken from: [b]Scott et al. [10], [c]Rung et al. [6], [d]Zeggini et al. [4], [e]Unoki et al. [9], [f]Zeggini et al. [5], [g]Voight et al. [3], [h]Frayling et al. [2], [i]Franks et al. [11]

[j]In the current study the *SLC30A8* rs11558471 SNP was used as a proxy for the rs13266634 SNP used in Zeggini et al. [4] and Chauhan et al. [14] (r^2=0.949 in CEU HapMap data)

[k]In Zeggini et al. [4] the *KCNJ11* rs5215 SNP was used as a proxy for rs5219 (r^2 in the Wellcome Trust Case Control Consortium and UKT2D collections = 0.995)

RAF, risk allele frequency in normoglycaemic control group; Chr, chromosome

(*DUSP9*) the per allele OR was generated by coding all males as diploid homozygotes. Heterogeneity of OR values was analysed using Cochran's Q-statistics. To compare effect sizes between UKADS and DGP cohorts, logistic regression was first used to generate the UKADS- and DGP-specific ORs. A genetic risk score (GRS) was generated for each individual by taking the sum of the weighted number of observed risk alleles, each risk allele weighted by a European-derived SNP-specific per allele effect size ($\log_e$[OR]), and dividing by the mean European-derived per allele effect size for the successfully genotyped SNPs. Association between GRS and quantitative traits was

analysed using linear regression, correcting for age, BMI, sex, country of residence and type 2 diabetes, as appropriate. Age at diagnosis was used as the closest available proxy to age at onset of type 2 diabetes (AOD). To investigate the potential of the studied variants to discriminate between participants with and without diabetes, we constructed a receiver-operating characteristic (ROC) curve and calculated the AUC. Power was calculated using Genetic Power Calculator (http://ibgwww.colorado.edu/~pshaun/gpc/) [19].

Results

The clinical characteristics of our study populations are shown in ESM Table 1. The *HMGA2* rs1531343 variant deviated significantly from HWE after correcting for multiple testing (60 tests) whilst the *ZFAND6* rs11634397 SNP displayed heterogeneity of ORs between the UKADS and DGP groups (ESM Table 2). These SNPs were therefore excluded from further analyses.

In our Pakistani populations, 13 variants were significantly ($p<0.05$) associated with type 2 diabetes, although only the *TCF7L2*, *IGF2BP2*, *CDKN2A/2B* and *IRS1* SNPs retained significance after correction for multiple testing (Table 1). The GRS was strongly associated with type 2 diabetes, each effective risk allele contributing an OR of 1.08 ($p=5.46\times10^{-12}$, Fig. 1). Stratifying individuals into risk categories based upon numbers of effective risk alleles demonstrated that the 8.5% of individuals with 38 or more risk alleles had more than twice the odds of developing diabetes compared with the 9.5% of individuals with fewer than 28 risk alleles (OR 2.60 [95% CI 1.80–3.75] $p=4.04\times10^{-7}$). The GRS was also associated with AOD ($\beta=-0.13$ [95% CI −0.20, −0.05] $p=0.002$) and BMI ($\beta=-0.08$ [95% CI −0.13, −0.04] $p=2.25\times10^{-4}$) but not fasting glucose ($\beta=0.005$ 95% CI −0.002, 0.012] $p=0.195$; Fig. 1). The association with BMI was further enhanced by removing the obesity-related *FTO* variant from the GRS ($\beta=-0.10$ [95% CI −0.15, −0.06] $p=1.55\times10^{-5}$).

After correction for multiple testing (30 tests) no significant heterogeneity of ORs was observed between the UKADS/DGP cohorts and published data from European populations [2–6, 9–11] (Table 1). Conversely, in our Pakistani populations the *TCF7L2* and *KCNJ11* ORs were significantly lower than those previously reported in Indian Asian populations (Table 1; $p_{\text{corrected}}=1.72\times10^{-7}$ and 8.88×10^{-6} respectively) [14].

Calculating the AUC demonstrated that the discriminatory power of the 28 variants included in the analysis was low (ESM Fig. 1). When including only BMI, age, sex and country of residence in the model the AUC was 0.71. This increased slightly, but significantly ($p=2.36\times10^{-7}$), to 0.74, when the 28 SNPs were included in the model.

Discussion

In this study we have investigated the association of type 2 diabetes with 30 SNPs, derived predominantly from European GWA studies, in two independently ascertained Pakistani populations. Of the variants studied, 13 displayed nominally significant associations with the disease (Table 1). Seven of these variants are in regions that have previously demonstrated association with type 2 diabetes in South Asian populations; *PPARG*, *TCF7L2*, *FTO*, *CDKN2A/2B*, *HHEX/IDE*, *IGF2BP2* and *SLC30A8* [12, 14, 15]. To our knowledge, however, this study is the first to report significant associations between the disease and variants in or near *KCNQ1*, *JAZF1*, *IRS1*, *KLF14*, *CHCHD9* and *DUSP9* in a sizeable sample of South Asians.

Only rs7903146 (*TCF7L2*), rs4402960 (*IGF2BP2*), rs10811661 (*CDKN2A/2B*) and rs2943641 (*IRS1*) remained significantly associated with type 2 diabetes after a classical Bonferroni correction for the number of tests performed; however, we believe that this type of correction is overly conservative. All the variants studied have been robustly associated with type 2 diabetes in European populations. Assuming similar linkage disequilibrium (LD) patterns between our study populations and European populations, this provides an a priori hypothesis that these SNPs would also be associated with the disease in South Asians.

In our study, despite 83% statistical power, we found a complete lack of association between the *KCNJ11* rs5219 (E23K) SNP and type 2 diabetes. It is possible that our finding is due to differing LD between rs5219 and another SNP, one that may be the true aetiological variant; for example E23K has been shown to be in tight LD with the S1368A variant in the *ABCC8* gene [20]. Previous studies provide evidence of association between rs5219 and type 2 diabetes in both European and South Asian populations, however, suggesting that a difference in LD patterns may not be responsible for our finding. It may be that the effect size of the rs5219 SNP is slightly lower in our Pakistani populations compared with previously studied populations; power is reduced dramatically to 56% if the OR is actually 1.10 rather than the 1.14 assumed in this study. We suggest that a number of the variants that display no statistically significant association with type 2 diabetes in the current study may truly confer disease risk, as 9 out of the 15 had effect sizes in the same direction as those seen in European studies. Larger sample sizes will be needed in order to elucidate the effects of these variants on type 2 diabetes risk in South Asian populations.

A recent study demonstrated that a number of common type 2 diabetes risk variants have larger effect sizes in Indian Asians than in European populations [14]. In contrast, the ORs observed in the current study do not

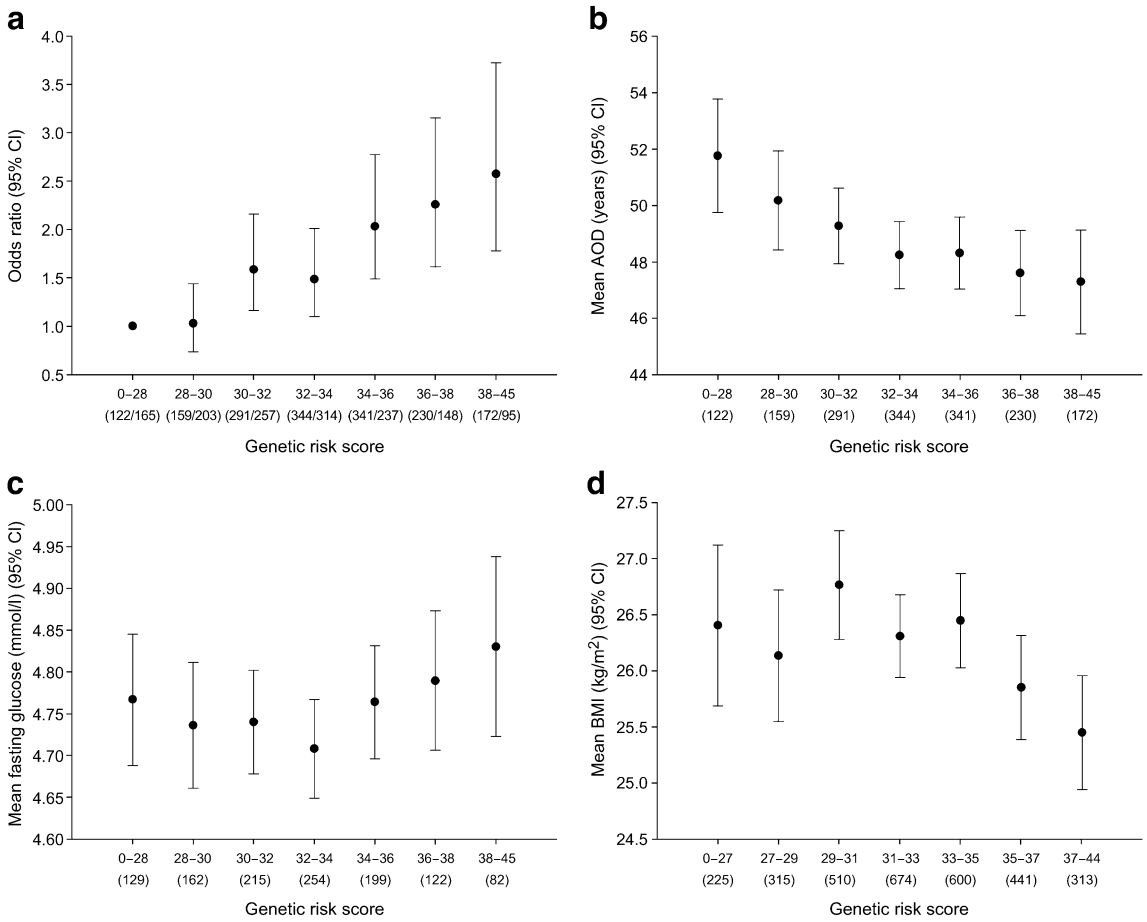

Fig. 1 Association of GRS formed from 28 SNPs for the analysis of (**a**) type 2 diabetes, (**b**) age at onset of type 2 diabetes and (**c**) fasting glucose, but formed from 27 SNPs (removing the obesity-associated *FTO* rs9939609 variant) for the analysis of BMI (**d**). Per effective risk allele effect sizes were: (**a**) OR 1.08 (95% CI 1.06–1.10) $p=5.46\times10^{-12}$ (**b**) $\beta=-0.13$ (95% CI −0.20, −0.05) $p=0.002$, (**c**) $\beta=0.005$ (95% CI −0.002, 0.012) $p=0.195$ and (**d**) $\beta=-0.10$ (95% CI −0.15, −0.06) $p=1.55\times10^{-5}$. Values in parentheses indicate the sample size within each GRS category; either cases/controls (**a**), cases only (**b**), controls only (**c**) or cases and controls (**d**). For the analysis of type 2 diabetes (adjusted for age, sex, country of residence and BMI) and BMI (adjusted for age, sex, country of residence and type 2 diabetes) 3,078 participants with available BMI data were used. For the age at onset analysis (adjusted for age, sex, country of residence and BMI) 1,659 type 2 diabetes participants with available BMI data were used. For the fasting glucose analysis (adjusted for age, sex and BMI) 1,163 normoglycaemic controls with available BMI data from the DGP collection were used

differ significantly from those seen in Europeans [2–6, 9–11]. Our results suggest that the large ORs demonstrated by Chauhan et al. [14] are not necessarily applicable to all South Asian populations. The observed difference in effect sizes may be an artefact of participant ascertainment; variation in BMI may mediate the effect of genetic variants on type 2 diabetes risk, lower age at disease onset may reflect a greater genetic load, and differing criteria for the selection of control participants (i.e. age thresholds, blood glucose tests) will influence the chances of including individuals who have impaired glucose tolerance or are likely to develop type 2 diabetes later on in life. It is difficult to assess the potential cumulative effect that this multitude of factors may have on results reported by different studies. Another possibility is that the large effect sizes reported by Chauhan et al. [14] are a phenomenon specific to Indian Asians.

Given the similarity of our results to those reported in investigations of European populations, the studied variants are unlikely to contribute greatly to the increased risk of type 2 diabetes experienced by Pakistani individuals compared with Europeans. It may be that this risk is conferred by genetic variation not investigated in this study (such as rare SNPs, undiscovered common SNPs or copy number variants), the interaction of genetics with environment, or epigenetic modification.

The additive effect of multiple type 2 diabetes susceptibility variants has been widely reported [14, 21, 22]. In our study each effective risk allele conferred a small but significant increase in disease risk, and individuals with >38 effective risk alleles had a 2.6-fold increased risk compared with individuals with <28 effective risk alleles. We also demonstrated that the GRS was significantly

associated with a lower AOD and BMI. The power of these variants to discriminate between individuals with and without diabetes, however, is still low, adding just 0.03 to the AUC of 0.71 produced by a model including only age, sex, BMI and country of residence.

The observed decrease in BMI with increasing GRS requires some explanation, as increased BMI represents a substantial environmental risk factor for type 2 diabetes development. Such an association within the case group may be explained relatively simply, as those participants with lower BMI may have developed the disease due to an increased genetic load. In our study, however, the association was present in both case and control groups (β=−0.09 [95% CI −0.15, −0.03] p=0.002 and β=−0.11 [95% CI −0.18, 0.04] p=0.002 respectively, using the GRS with the *FTO* SNP excluded). This may reflect the fact that those individuals with a high GRS and increased BMI are more likely to have developed type 2 diabetes and therefore would not be present in the control group. Speculatively, it is also possible that the observed association in the control group is linked to gene function. Most of the recently discovered type 2 diabetes risk variants are thought to contribute to disease development through beta cell dysfunction rather than insulin resistance. Those participants with a high GRS may therefore have slightly impaired insulin production/secretion which, although not severe enough to result in detectable impaired glucose homeostasis, may affect BMI through the role of insulin in fat metabolism.

Our study has limitations, mainly due to moderate sample size. Although false positive results are unlikely, due to the strong a priori hypothesis that the studied variants would be associated with type 2 diabetes, statistical power is low and it is likely that our results include false negatives. This is exacerbated by the fact that we did not use OGTT to classify most of our normoglycaemic participants, and it is possible that our control groups include individuals who have impaired glucose tolerance and may develop diabetes in the future. Other possible explanations for the observed lack of association for some SNPs may include differences in LD patterns between our studied populations and Europeans, as most of the SNPs identified through European GWA studies are not aetiological variants, or the possibility that the risk allele of the aetiological variant is absent or very rare in South Asians. Another limitation is that we did not have enough genetic data to test for population stratification or cryptic relatedness, although all participants were of Punjabi ancestry confirmed over three generations, and the study protocols included the caveat that related individuals should not be recruited. Because of these recruitment criteria, we feel that population stratification is unlikely. In addition, if related individuals have unintentionally been recruited it is likely that the level of relatedness is similar between case and control groups, which may negate the effect of any possible cryptic relatedness. Our strict ethnic selection criteria may have introduced another limitation; as our participants all originate from the Punjab our dataset may not be representative of all Pakistani populations.

In conclusion, we have performed the most comprehensive analysis to date of type 2 diabetes susceptibility variants in South Asians. Our results demonstrate that variants in or near 13 genes are significantly associated with the disease in two Pakistani populations, with similar effect sizes to those reported previously in Europeans. A GRS constructed from 28 SNPs was significantly associated with type 2 diabetes, AOD and BMI, although even this relatively large number of SNPs demonstrated poor power to discriminate between participants with and without diabetes. To our knowledge, this is the first time that variants in or near *KCNQ1*, *JAZF1*, *IRS1*, *KLF14*, *CHCHD9* and *DUSP9* have been significantly associated with type 2 diabetes in a sizeable South Asian cohort, although further studies and meta-analyses will be needed to confirm these SNPs as true disease susceptibility variants in South Asians.

Acknowledgements The authors would like to thank all the diabetic patients and control participants for agreeing to participate in this study. We are grateful to S. Mughal, K. Johal, A. Dixon, S. Bellary and T. Mehrali within the UK and A. Fawwad, W. Iqbal and A. Razzaq in Pakistan for recruiting the study participants and coordinating data collection. The UK Asian Diabetes Study was conceived and managed by A. Barnett, S. Kumar and J. P. O'Hare and was supported by Pfizer, sanofi-aventis, Servier Laboratories UK, Merck Sharp & Dohme/Schering-Plough, Takeda UK, Roche, Merck Pharma, Daiichi-Sankyo UK, Boehringer Ingelheim, Eli Lilly, Novo Nordisk, Bristol-Myers Squibb, Solvay Health Care and Assurance Medical Society UK. Funding for the collection of DNA and clinical data from the Pakistan-resident population was provided by Diabetes UK (project number 07/0003512). Funding for genetic analyses was provided by Diabetes UK (project number 09/0003926).

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

References

1. King H, Aubert RE, Herman WH (1998) Global burden of diabetes, 1995–2025: prevalence, numerical estimates, and projections. Diab Care 21:1414–1431
2. Frayling TM, Timpson NJ, Weedon MN et al (2007) A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science 316:889–894
3. Voight BF, Scott LJ, Steinthorsdottir V et al (2010) Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 42:579–589
4. Zeggini E, Weedon MN, Lindgren CM et al (2007) Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science 316:1336–1341
5. Zeggini E, Scott LJ, Saxena R et al (2008) Meta-analysis of genome-wide association data and large-scale replication identifies

additional susceptibility loci for type 2 diabetes. Nat Genet 40:638–645
6. Rung J, Cauchi S, Albrechtsen A et al (2009) Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. Nat Genet 41:1110–1115
7. Sladek R, Rocheleau G, Rung J et al (2007) A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 445:881–885
8. Grant SF, Thorleifsson G, Reynisdottir I et al (2006) Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. Nat Genet 38:320–323
9. Unoki H, Takahashi A, Kawaguchi T et al (2008) SNPs in KCNQ1 are associated with susceptibility to type 2 diabetes in East Asian and European populations. Nat Genet 40:1098–1102
10. Scott LJ, Mohlke KL, Bonnycastle LL et al (2007) A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 316:1341–1345
11. Franks PW, Rolandsson O, Debenham SL et al (2008) Replication of the association between variants in WFS1 and risk of type 2 diabetes in European populations. Diabetologia 51:458–463
12. Sanghera DK, Ortega L, Han S et al (2008) Impact of nine common type 2 diabetes risk polymorphisms in Asian Indian Sikhs: PPARG2 (Pro12Ala), IGF2BP2, TCF7L2 and FTO variants confer a significant risk. BMC Med Genet 9:59
13. Sanghera DK, Been L, Ortega L et al (2009) Testing the association of novel meta-analysis-derived diabetes risk genes with type II diabetes and related metabolic traits in Asian Indian Sikhs. J Hum Genet 54:162–168
14. Chauhan G, Spurgeon CJ, Tabassum R et al (2010) Impact of common variants of PPARG, KCNJ11, TCF7L2, SLC30A8, HHEX, CDKN2A, IGF2BP2, and CDKAL1 on the risk of type 2 diabetes in 5,164 Indians. Diabetes 59:2068–2074
15. Chidambaram M, Radha V, Mohan V (2010) Replication of recently described type 2 diabetes gene variants in a South Indian population. Metabolism 59:1760–1766
16. Alberti KG, Zimmet PZ (1998) Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. Diabet Med 15:539–553
17. Bellary S, O'Hare JP, Raymond NT et al (2008) Enhanced diabetes care to patients of south Asian ethnic origin (the United Kingdom Asian Diabetes Study): a cluster randomised controlled trial. Lancet 371:1769–1776
18. Purcell S, Neale B, Todd-Brown K et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575
19. Purcell S, Cherny SS, Sham PC (2003) Genetic power calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 19:149–150
20. van de Bunt M, Gloyn AL (2010) From genetic association to molecular mechanism. Curr Diab Rep 10:452–466
21. Lango H, Palmer CN, Morris AD et al (2008) Assessing the combined impact of 18 common genetic variants of modest effect sizes on type 2 diabetes risk. Diabetes 57:3129–3135
22. Weedon MN, McCarthy MI, Hitman G et al (2006) Combining information from common type 2 diabetes risk polymorphisms improves disease prediction. PLoS Med 3:e374