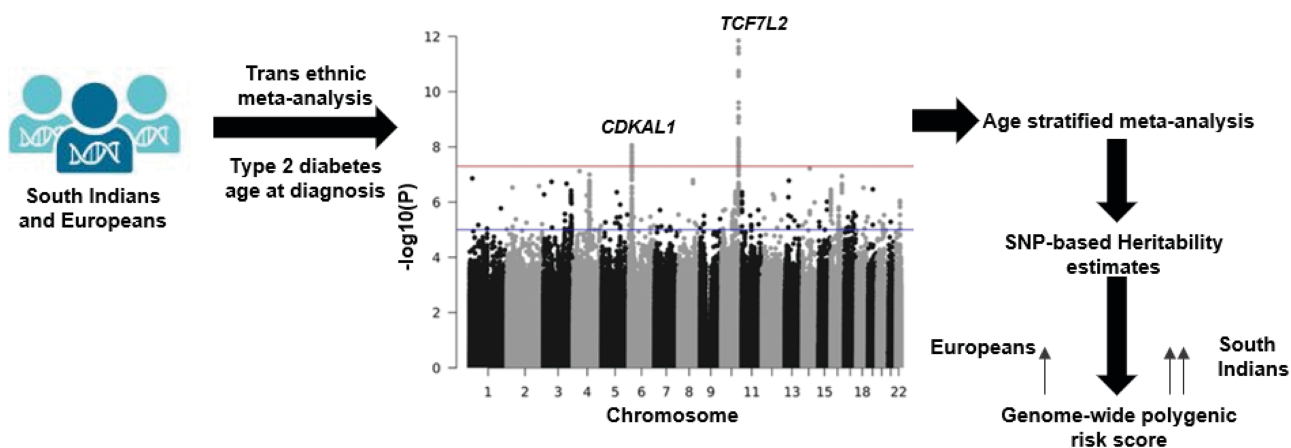


Common and Distinct Genetic Architecture of Age at Diagnosis of Diabetes in South Indian and European Populations

Sundararajan Srinivasan, Samuel Liju, Natarajan Sathish, Moneeza K. Siddiqui, Ranjit Mohan Anjana, Ewan R. Pearson, Alexander S.F. Doney, Viswanathan Mohan, Venkatesan Radha, and Colin N.A. Palmer

Diabetes Care 2023;46(8):1515–1523 | <https://doi.org/10.2337/dc23-0243>



Common and distinct genetic architecture of age at diagnosis of diabetes in South Indian and European populations. SNP, single nucleotide polymorphism.

ARTICLE HIGHLIGHTS

- Transethnic meta-analysis revealed an association of *TCF7L2* and *CDKAL1* with early age at diagnosis of type 2 diabetes (T2D) in South Asian Indians.
- The findings highlight the disparity in heritability estimates of age at diagnosis of T2D in South Indians and Europeans and explain the polygenic variation in South Indians.
- Genome-wide polygenic scores based on South Indian genome-wide association studies performed better in South Indians ($R^2 \sim 2\%$) than Europeans ($R^2 \sim 0.1\%$).



Common and Distinct Genetic Architecture of Age at Diagnosis of Diabetes in South Indian and European Populations

Diabetes Care 2023;46:1515–1523 | <https://doi.org/10.2337/dc23-0243>

Sundararajan Srinivasan,¹ Samuel Liju,² Natarajan Sathish,² Moneeza K. Siddiqui,¹ Ranjit Mohan Anjana,² Ewan R. Pearson,¹ Alexander S.F. Doney,¹ Viswanathan Mohan,² Venkatesan Radha,² and Colin N.A. Palmer¹

OBJECTIVE

South Asians are diagnosed with type 2 diabetes (T2D) more than a decade earlier in life than seen in European populations. We hypothesized that studying the genomics of age at diagnosis in these populations may give insight into the earlier age diagnosis of T2D among individuals of South Asian descent.

RESEARCH DESIGN AND METHODS

We conducted a meta-analysis of genome-wide association studies (GWAS) of age at diagnosis of T2D in 34,001 individuals from four independent cohorts of European and South Asian Indians.

RESULTS

We identified two signals near the *TCF7L2* and *CDKAL1* genes associated with age at the onset of T2D. The strongest genome-wide significant variants at chromosome 10q25.3 in *TCF7L2* (rs7903146; $P = 2.4 \times 10^{-12}$, $\beta = -0.436$; SE 0.02) and chromosome 6p22.3 in *CDKAL1* (rs9368219; $P = 2.29 \times 10^{-8}$; $\beta = -0.053$; SE 0.01) were directionally consistent across ethnic groups and present at similar frequencies; however, both loci harbored additional independent signals that were only present in the South Indian cohorts. A genome-wide signal was also obtained at chromosome 10q26.12 in *WDR11* (rs3011366; $P = 3.255 \times 10^{-8}$; $\beta = 1.44$; SE 0.25), specifically in the South Indian cohorts. Heritability estimates for the age at diagnosis were much stronger in South Indians than Europeans, and a polygenic risk score constructed based on South Indian GWAS explained $\sim 2\%$ trait variance.

CONCLUSIONS

Our findings provide a better understanding of ethnic differences in the age at diagnosis and indicate the potential importance of ethnic differences in the genetic architecture underpinning T2D.

Type 2 diabetes (T2D) is a multifactorial disease characterized by impaired insulin action and pancreatic islet dysfunction. The global prevalence of T2D is a pivotal driver of cardiovascular and renal disease (1–3), affecting hundreds of millions of people globally, and is responsible for long-term complications, decreased quality of life, and increased mortality (4–7). Improved understanding of the intrinsic genomic and phenotypic heterogeneity driving T2D has major potential for improvement of T2D clinical management and reducing morbidity and mortality. South

¹Division of Population Health and Genomics, Ninewells Hospital and School of Medicine, University of Dundee, Dundee, U.K.

²Madras Diabetes Research Foundation and Dr Mohan's Diabetes Specialities Centre, Chennai, India

Corresponding author: Colin N.A. Palmer, c.n.a.palmer@dundee.ac.uk

Received 10 February 2023 and accepted 15 May 2023

This article contains supplementary material online at <https://doi.org/10.2337/figshare.23063396>.

© 2023 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <https://www.diabetesjournals.org/journals/pages/license>.

Asian Indians have an earlier age of onset of diabetes compared with Europeans, and mounting evidence suggests that this is associated with earlier mortality, emphasizing the need to delay or prevent the onset of T2D in this ethnic group (8,9). South Asians with newly diagnosed diabetes may have a higher risk for microvascular complications than Europeans (10). Previous studies highlight the association between higher cardiovascular mortality and disease risk and early-onset T2D compared with delayed onset of the disease (11). South Asians (individuals originating from India, Pakistan, and Bangladesh) are genetically more diverse than White Europeans, and the prevalence of T2D is much higher in this ethnic group than other ethnic backgrounds (12–14).

Currently, nearly 250 genetic loci that influence T2D (>400 unique genetic variants) have been identified (2,15,16). Several of these genetic loci have only been identified in European study populations. A transethnic meta-analysis of European and East Asian populations reported several T2D risk variants with significant allelic frequency heterogeneity (12,13). Such frequency differences between ethnic populations affect the power to detect genomic signals within a specific ethnic subgroup. A recent study reported that migrant South Asians are more insulin resistant and have poorer β -cell function at a younger age than White Europeans. Previously identified genetic variants explained ~10% of the heritability of T2D (14).

Despite advancement in genetic research tools, South Asian Indian-specific studies are minimal compared with European ancestry studies. To our knowledge, no genome-wide association study (GWAS) has addressed the age at diagnosis of T2D in people of South Asian Indian ethnicity as compared with European populations. We aimed to identify novel genetic determinants that influence the risk of younger age at diagnosis in two distinct ethnic backgrounds, specifically in South Asian Indians and Europeans. We aimed to develop, evaluate, and understand a T2D age at diagnosis polygenic risk score (PRS) in cohorts from South India (Dr Mohan's Diabetes Specialties Centre [DMDSC]) and Europe (Genetics of Scottish Health Research Register [GoSHARE]). In this multicenter study, we focused on interancestry differ-

ences in the genetics of age at diagnosis of T2D that might influence ethnic-ancestry differences in health outcomes.

RESEARCH DESIGN AND METHODS

Study Participants

We included participants from four independent cohorts: Dr Mohan's Diabetes Specialties Centre (DMDSC) (17), Genetics of Diabetes Audit and Research in Tayside Scotland (GoDARTS) (18), Genetics of Scottish Health Research Register (GoSHARE) (19), and the UK Biobank (UKBB) (20). DMDSC is a chain of diabetes hospitals and clinics established in 1991 in Chennai, India, which currently includes 50 clinics in various locations across eight states (17). To date, a total of 560,000 patients with T2D have been provided a unique identification number at their first visit, and clinical, anthropometric, and biochemical data are updated at each subsequent visit. Each patient undergoes a comprehensive evaluation for screening and assessment of diabetes and presence of chronic complications at the time of their first registered visit, and these tests are repeated subsequently. All data are collected and stored in the common diabetes electronic medical record system. The Madras Diabetes Research Foundation is the research unit of DMDSC and is accredited by the College of American Pathologists and the National Accreditation Board for Testing and Calibration Laboratories for various biochemical tests. GoDARTS consists of 18,306 participants from the Tayside region of Scotland, of whom 10,149 were recruited based on their diagnosis of T2D (18). GoSHARE currently comprises a biobank of ~74,000 individuals across the National Health Service Fife and the National Health Service Tayside (19). Participants of both cohorts provided a sample of blood for genetic analysis and informed consent to link their genetic information to anonymized electronic health records. UKBB is a large, prospective, general population cohort. A total of 502,628 individuals aged 40–69 years were recruited between 2006 and 2010 from across the U.K. and provided electronically signed consent to use their self-reported answers on sociodemographic, lifestyle, ethnicity, a range of physical measures, and blood, urine, or saliva samples (20).

All research was conducted under the principles of the Declaration of Helsinki and approved by corresponding institutional

review boards. All study participants provided written informed consent, and institutional ethics committees approved the study. This study follows the Strengthening the Reporting of Genetic Association Studies (STREGA) reporting guideline (21) (Supplementary Table 2).

Phenotyping: Age at Diagnosis of T2D

We included 8,295 patients with T2D from the DMDSC cohort of South Indians whose first clinical visit was within 1 year of diagnosis; age at diagnosis was recorded at first visit. For individuals who were diagnosed at DMDSC, the age at diagnosis was recorded at that time. Diabetes is diagnosed by general practitioners using World Health Organization criteria (22) either the oral glucose tolerance test or fasting and/or random glucose test of HbA_{1c}. All study participants underwent a structured assessment, including detailed family history, at the DMDSC. We excluded patients with type 1 diabetes (T1D) or if positive for GAD65 and ZnT8 antibodies during treatment and follow-up. Although this study is cross sectional, diabetes classifications were applied retrospectively to ensure that the population under study only included individuals with T2D. We selected the study participants in GoDARTS and GoSHARE based on the following inclusion criteria: aged 20–80 years and T2D status monitored continuously and updated by a primary or secondary physician or community nurses. Algorithms track the use of insulin and other oral hypoglycemic agents. Individuals who were originally classified as having T2D can be reclassified as having T1D if they were aged <30 years when diagnosed while also having routine insulin use as per World Health Organization guidelines and recorded in the Scottish Care Information–Diabetes system (23). The age at diagnosis is available as part of the Scottish Care Information–Diabetes Collaboration data recording system, which is centrally updated. This estimate of when the disease was diagnosed is the most precise (24). We identified 14,552 participants with T2D within the UKBB cohort with physician-diagnosed diabetes (data field code 2443), having started insulin 1 year after diagnosis (2986), and self-reported ethnicity (21000) and excluded participants with outlying principal components.

Genotyping, Quality Control, and Imputation

Blood samples were collected from DMDSC participants. A total of 5,801 patients with T2D were genotyped using Illumina global screening arrays version 1.0 (GSA v1.0), and 2,494 patients with T2D were genotyped using GSA v2.0. All genotyped samples were converted to PLINK format files using Illumina Genome Studio version 2.04. We excluded samples with a call rate <95% and genetically inferred sex discordance with phenotype data, batch effects, heterozygosity >3 SDs, and sample duplicates (identity by descent [IBD] score >0.8). We excluded single nucleotide polymorphisms (SNPs) with <97% call rate and Hardy-Weinberg equilibrium $P < 1 \times 10^{-6}$ (autosomal variants only). Quality control assessment was performed independently for DMDSC cohorts before and after phasing and imputation against the Haplotype Reference Consortium (HRC) version r1.1 panel.

Genotyping of the GoDARTS and GoSHARE cohorts was derived from various platforms: Affymetrix 6.0 (Santa Clara, CA), Illumina Omni Express-12V1 platform, and GSA v2.0. A total of 11,154 (6,999 GoDARTS and 4,155 GoSHARE) participants were considered after excluding individuals who did not meet quality control criteria. The overall individual genotype call rate (<95%), heterozygosity >3 SDs from the mean, and the highly related sample's identity by descent. We then performed SNP-level quality control by excluding markers with a <97% call rate and Hardy-Weinberg $P < 1 \times 10^{-6}$. PLINK versions 1.7 and 1.9 were used for quality control assessment and data preprocessing for imputation. Ancestry outliers were identified by principal component analysis in each cohort. The genotype data from all three cohorts were imputed against the HRC r1.1 reference panel. Monomorphic markers or imputation quality score <0.4 were excluded in the postimputation data.

Ethnic-Specific Meta-analysis of GWAS

GWAS were performed independently for each cohort using an additive model while adjusting for sex (Supplementary Figs. 5–8). Of note, previous studies have highlighted that South Asians in general have the weakest age-adjusted association between BMI and T2D or no diabetes (25). In our own data sets, we also found that BMI is the weakest predictor of age of diagnosis of T2D in a

South Indian cohort but is a predictor in White Europeans (Supplementary Fig. 10 and Supplementary Table 4). We estimated allelic effects using a linear mixed model as implemented in BOLT-LMM version 2.3.2, which accounts for relatedness and any population stratification, and SNPTEST version 2.5 in each cohort accordingly. We performed a meta-analysis based on ancestry: South Asian Indian-specific analyses include the DMDSC cohort, a unique South Indian population, and migrated South Asians in the UKBB, and the European-specific analyses include the GoDARTS, GoSHARE, and White Europeans in the UKBB. We performed the meta-analyses using a fixed-effects method in METAL software (26), which assumes that the effect allele is the same for each study within an ancestry. We then conducted transancestry meta-analyses using the HRC-imputed data of up to 26.2 million SNPs directly genotyped or successfully imputed at high quality across all the study cohorts. Heterogeneity across these studies was assessed using I^2 (low to high) and Cochran Q statistics as reported by METAL. Forest plots were generated using the metafor package. We annotated the genetic variants using the University of California, Santa Cruz, genome resource based on Genome Reference Consortium Human Build 37.

Conditional Analysis

We performed conditional analyses to identify additional secondary signals across the lead SNPs within the South Indian population.

SNP-Based Heritability

We used summary statistics data from the South Indian- and European-specific meta-analyses to estimate the SNP-based heritability in a liability scale using linkage disequilibrium score regression software (27).

Genome-Wide PRSs for Age at Diagnosis of T2D

For PRSs, we considered summary statistics of DMDSC samples. The PRSice tool generates the scores by the weighted sum of the risk allele carried by individuals based on effect estimate. We removed DNA polymorphisms with ambiguous strands (A/T or C/G) from the score derivation. SNPs were clumped to a more significant SNP in linkage disequilibrium ($r^2 \geq 0.10$) within a 500-kb window. The

PRS calculation considered several P value thresholds (0.001, 0.05, and 0.1).

Data and Resource Availability

Summary data might be made available upon reasonable request via e-mail to the lead and corresponding authors.

RESULTS

A total of 34,001 participants with T2D were included for this study after quality control filtering: 8,295 with T2D of South Indian ancestry from DMDSC, a unique and homogenous population as shown in Supplementary Fig. 1; 6,999 of European ancestry from GoDARTS (18); 4,155 of European ancestry from GoSHARE (19); and 14,552 from UKBB (20). We identified participants of European ($n = 13,744$) and South Asian Indian ($n = 808$) descent in the UKBB using principal component analysis of genome-wide data and found that this was consistent with self-reported ancestry information (Supplementary Fig. 2). The population characteristics of the cohorts are described in Supplementary Table 1. Notably, we observed that the mean age of diagnosis of T2D in South Asian Indians was 40 years, whereas in White Europeans, it was 59 years in GoDARTS, 58.2 years in GoSHARE, and 54.6 years in UKBB.

SNP-Based Heritability

Using linkage disequilibrium score regression tools, we estimated that the SNP-based heritability for age of diagnosis of T2D in South Indians was 17% (SE 6%) but was only 5% (SE 2%) for Europeans.

Transethnic Meta-analysis of GWAS for Age at T2D Diagnosis

Our meta-analysis revealed two previously known T2D loci at chromosome 10q25.2 near transcription factor 7-like 2 (*TCF7L2*) (rs79603146, $P < 2.4 \times 10^{-12}$, $\beta = -0.436$; SE 0.02; P -heterogeneity = 0.01) and at chromosome 6p22.3 cyclin-dependent kinase 5 (*CDK5*) regulatory subunit-associated protein 1-like 1 (*CDKAL1*) (rs9368219, $P < 2.29 \times 10^{-8}$, $\beta = -0.053$; SE 0.01; P -heterogeneity = 0.007) associated with age at T2D diagnosis (Fig. 1 and Supplementary Fig. 9). The allelic frequency of *CDKAL1* was more common in the South Indian cohort (minor allele frequency [MAF] 0.26) compared with the White European cohort (MAF 0.18). The lead SNPs at the *TCF7L2* and *CDKAL1* (Fig. 1 and Supplementary Fig. 9) loci demonstrated

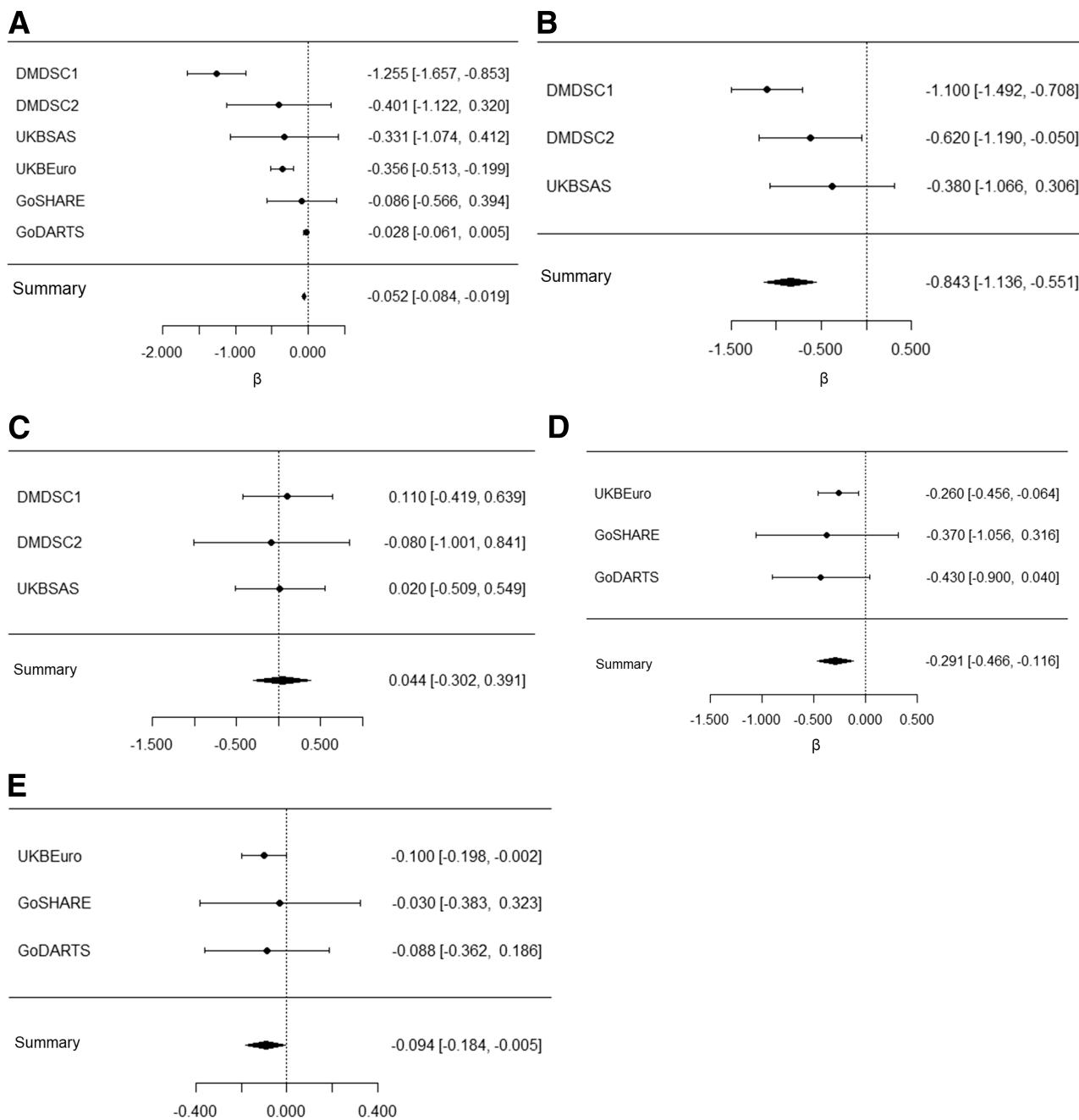


Figure 1—Forest plot for the top significant SNP rs7903146 near *TCF7L2* (effect allele T) in individuals with T2D. *A*: Overall meta-analysis of GWAS of age at diagnosis of T2D. *B*: South Asian Indians with early-onset T2D (aged 20–55 years at diagnosis). *C*: South Asian Indians with late-onset T2D (aged >55 years at diagnosis). *D*: Europeans with early-onset T2D (aged 20–55 years at diagnosis). *E*: Europeans with late-onset T2D (aged at >55 years at diagnosis). UKBEuro, UKBB Europeans; UKBSAS, UKBB South Asians.

consistent allelic direction across all cohorts, with the risk alleles associated with younger age at diagnosis; however, a large difference was observed in the size of the estimate of the effects between the South Indian and European cohorts, explaining that variation in allelic effect estimates is presumably due to their genetic ancestry. Interestingly, the effect size of the variants was much lower in the cohorts of European descent. Ethnic-specific meta-analysis

results are presented in Supplementary Tables 6–8. There was no evidence of population stratification in the meta-analysis (genomic inflation factor $\lambda = 1.007$).

Stratification of T2D by Age at Diagnosis

As the two ethnic groups were very different in the mean age at diagnosis, we explored the extent to which the observed differences in allelic effect size

may be determined by the heterogeneity in age at onset between the populations. DMDSC participants with T2D stratified by various age-groups are shown in Supplementary Table 5. Only a small proportion of individuals were diagnosed after 55 years of age.

Based on the European mean age at diagnosis, study participants of both ethnicities were stratified into an early-onset T2D group (aged 20–55 years) and a late-onset

T2D group (aged ≥ 55 years) (Fig. 1). We found that the effect size of both the *TCF7L2* and *CDKAL1* variants was more pronounced in the early-onset group, regardless of ethnicity (Fig. 1 and Supplementary Fig. 3). These variants had very little effect on age at diagnosis (Table 1) in those with diabetes diagnosed after 55 years of age in either ethnicity (Fig. 1). Although this variant is also nominally associated with age at diagnosis in the late-onset group in Europeans, we observed high heterogeneity ($I^2 = 79\%$) between younger and older age at onset (Supplementary Fig. 11).

Role of Other T2D Variants in Age at T2D Diagnosis

We identified several previously reported T2D variants as suggestive signals ($P < 1 \times 10^{-5}$) in these transethnic meta-analyses of age at diagnosis of T2D (Supplementary Table 6). In particular, the risk variant nearby *SEC24B* at

chromosome location 4q25 (rs76170449, $P < 1.79 \times 10^{-7}$) is also associated with cardiovascular traits, and 3p24.3 *ZNF385D* (rs17011243, $P < 1.13 \times 10^{-5}$) has been associated with T2D in prior GWAS. In addition to the other suggestive signals, we detected potential common variants at chromosome location 16p13.3 (*TPSD1*, rs1977100, $P < 3.40 \times 10^{-6}$) and 17q21.2 (*MLX*, rs684214, $P < 2.40 \times 10^{-6}$), with no difference in their effect estimates between two distinct ethnic groups (Supplementary Table 6). We replicated previously reported South Asian T2D genome-wide signals (15,28–31) with suggestive evidence or a nominal association for age at diagnosis of T2D in the transethnic meta-analyses and ancestry-specific groups. Most of the formerly associated T2D loci from earlier GWAS showed consistent effect estimates in South Indian and European participants. These include *LPL*, *SLC30A8*, *GCKR*, *THADA*, *HNF1A*, *TPCN2*,

GRB14, *SIX3*, *WDR11*, *SPC25*, *CENTD2*, *MLX*, *APS32*, *WFS1*, *ST6GAL1*, *KNCQ1*, and *IGF2BP2*.

Meta-analysis of South Indian Cohorts

In the meta-analysis of only the South Indian cohorts, we also found an additional novel genome-wide signal at chromosome 10q26.12 near the *WDR11* region (rs3011366, $P < 3.255 \times 10^{-8}$, $\beta = 1.44$, SE 0.26). However, this variant was not associated with age at diagnosis in the European cohorts (Table 2). *WDR11* encodes the WD repeat domain family, which involves signal transduction and cell cycle progression. Previous GWAS in the European populations and UKBB participants with T2D have reported that *WDR11* (rs3011366) is associated primarily with fasting glucose (32). In silico lookups in the Common Metabolic Diseases Knowledge Portal indicate that this SNP near *WDR11* is also associated

Table 1—Summary statistics of the lead SNPs from meta-analysis using the stratified age at diagnosis cohort

Cohort with diabetes	SNP	EA/NEA	Study cohort	Sample size, n	β	SE	EAF	P	P-heterogeneity	Gene		
Younger onset of diabetes (aged 20–55 years)	rs7903146	T/C	DMDSC data freeze 1	5,191	−1.10	0.20	0.35	8.6×10^{-8}	0.13	<i>TCF7L2</i>		
			DMDSC data freeze 2	2,102	−0.62	0.29	0.34				0.03	
			UKBB (SAS)	542	−0.38	0.35	0.35				0.4	0.78
			Meta-analysis (SAS)	7,835	−0.84	0.15	0.34				<0.0001	
			GoDARTS	768	−0.43	0.26	0.34				0.11	
			GoSHARE	543	−0.37	0.35	0.36				0.0008	
			UKBB (Europeans)	4,991	−0.26	0.01					0.015	
			Meta-analysis (Europeans)	6,302	−0.29	0.08					0.001	
Older onset of diabetes (aged >55 years)	rs7903146	T/C	DMDSC data freeze 1	610	0.11	0.27	0.35	0.64	0.93	<i>TCF7L2</i>		
			DMDSC data freeze 2	392	−0.08	0.47	0.33				0.54	
			UKBB (SAS)	266	0.02	0.27	0.36				0.70	
			Meta-analysis (SAS)	1,268	0.05	0.17	0.34				0.80	
			GoDARTS	6,231	−0.08	0.14	0.33				0.55	
			GoSHARE	3,612	−0.03	0.18	0.36				0.83	
			UKBB (Europeans)	8,753	−0.10	0.05					0.06	
			Meta-analysis (Europeans)	18,596	−0.09	0.04					0.03	
Younger onset of diabetes (aged 20–55 years)	rs9368219	T/C	DMDSC data freeze 1	5,191	−1.06	0.19	0.25	4.6×10^{-8}	0.08	<i>CDKAL1</i>		
			DMDSC data freeze 2	2,102	−0.11	0.30	0.25				0.06	
			UKBB (SAS)	542	−0.53	0.38	0.27				0.04	0.48
			Meta-analysis (SAS)	7,835	−0.74	0.29	0.18				<0.001	
			GoDARTS	768	−0.05	0.30	0.19				0.08	
			GoSHARE	543	−0.50	0.39	0.17				0.07	
			UKBB (Europeans)	4,991	−0.008	0.12					0.9	
			Meta-analysis (Europeans)	6,302	−0.05	0.01					0.63	
Older onset of diabetes (aged >55 years)	rs9368219	T/C	DMDSC data freeze 1	610	0.11	0.27	0.25	0.68	0.95	<i>CDKAL1</i>		
			DMDSC data freeze 2	392	0.02	0.47	0.25				0.54	
			UKBB (SAS)	266	0.16	0.25	0.27				0.5	0.21
			Meta-analysis (SAS)	1,268	0.12	0.29	0.18				0.39	
			GoDARTS	6,231	−0.04	0.15	0.19				0.08	
			GoSHARE	3,612	−0.38	0.21	0.17				0.07	
			UKBB (Europeans)	8,753	0.002	0.06					0.9	
			Meta-analysis (Europeans)	18,596	−0.02	0.12					0.59	

EA, effect allele; EAF, effect allele frequency; NEA, noneffect allele; SAS, South Asians.

with youth-onset T2D, with a nominal significance level in transancestry cohorts (33). A recent case-control meta-analysis highlighted the association of the *WDR11* gene with T2D in East Asians (13), but the association was shown for a different allele in Europeans and East and South Asians. The conditional analyses conducted for South Indian ancestry (Supplementary Table 3) indicated two independent secondary signals at *TCF7L2* (rs570193324, $q_{25.2}$, $P < 3.2E-05$, $\beta = 9.8$, MAF 0.002, $R^2 = 0.0006$) and *CDKAL1* (rs143316471, $P < 0.0054$, $\beta = -5.3$, MAF 0.003, $R^2 = 0.005$). Allelic frequency for both independent signals were rare in European cohorts compared with South Indian cohorts. The regional plot for an independent association of *TCF7L2* and *CDKAL1* is shown in Supplementary Fig. 4.

Meta-analysis of European Cohorts

In the analyses unique to White Europeans, we did not observe any genome-wide signal in the European-specific meta-analyses. However, we observed a suggestive association of a missense variant rs2232328 near *SPC25*, an established variant for fasting blood glucose and T2D (34). Several other

SNPs reached suggestive significance for age at diagnosis of T2D, and the direction of the effect was consistent across all cohorts of European descent (Supplementary Table 8). Notably, rs17843614 near *HLA-DQB1* reached suggestive significance. While *HLA-DQB1* is a well-established T1D locus, there has been strong and consistent evidence about the association of rs17843614 with T2D (35).

PRS Analysis Reveals Polygenic Effects for Age at the Onset of T2D

PRS is emerging as a more informative clinical screening and prediction tool, with an increasing number of robust genomic variants identified through more extensive genetic association studies (36). To investigate whether different genetic variants shared between ethnicities were conferring the risk of the onset of T2D, the PRS was derived as the weighted sum of risk alleles based on β -values from the DMDSC 1 GWAS summary statistics. Results are presented based on 8,232,187 SNPs after quality control and clumped based on linkage disequilibrium ($r^2 \geq 0.1$) within a 500-kb window. We then validated this using the DMDSC 2 of South

Indian data, which contained no overlapping participants. Next, we assessed the performance of this South Indian genetic risk score in the European GoSHARE cohort (Fig. 2). The PRS replicated strongly between the South Indian cohorts. On the other hand, the South Indian-derived PRS explained $<0.1\%$ of the variance in age at diagnosis of T2D in the GoSHARE cohort (European ancestry).

CONCLUSIONS

In this study, we undertook a transethnic meta-analysis of age at diagnosis of T2D in 34,001 individuals from two diverse ancestral backgrounds, European and South Asian Indian, revealing a differential role for established T2D susceptibility loci in determining the age at onset of diabetes. Interestingly, the well-established T2D signal at *TCF7L2* was much more strongly associated with age at onset in the South Indian population compared with the European population, despite the allele frequency not differing between these ancestral groups. We show that this difference is due to the distribution of age of onset of diabetes within the two ancestral groups, with the *TCF7L2* effect

Table 2—Summary statistics of the most significant SNPs from the meta-analysis

SNP	CHR	POS	EA/NEA	Study cohort	β	SE	EAF	P	P -heterogeneity	Gene
rs7903146	10	114758349	T/C	DMDSC data freeze 1	-1.26	0.20	0.35	1.0×10^{-10}		<i>TCF7L2</i>
				DMDSC data freeze 2	-0.40	0.36	0.34	2.7×10^{-03}	0.02	
				UKBB (SAS)	-0.33	0.37	0.35	0.3	2.7×10^{-06}	
				Meta-analysis (SAS)	-0.92	0.16	0.34	1.1×10^{-08}	0.01	
				GoDARTS	-0.02	0.17	0.34	0.11		
				GoSHARE	-0.90	0.26	0.32	0.0008		
				UKBB (Europeans)	-0.35	0.08	0.35	1.3×10^{-05}		
				Meta-analysis (Europeans)	-0.02	0.02	0.34	0.004		
				Transethnic meta-analysis	-0.05	0.08	0.35	2.4×10^{-12}		
				rs9368219	6	20674691		DMDSC data freeze 1	-1.20	
DMDSC data freeze 2	-0.38	0.38	0.25					6.1×10^{-02}	0.002	
UKBB (SAS)	-0.59	0.40	0.27					0.14	0.02	
Meta-analysis (SAS)	-0.92	0.17	0.26					6.6×10^{-08}	0.007	
GoDARTS	-0.03	0.02	0.18					0.004		
GoSHARE	-0.80	0.30	0.19					0.007		
UKBB (Europeans)	-0.17	0.09	0.17					0.07		
Meta-analysis (Europeans)	-0.05	0.02	0.19					0.03		
Transethnic meta-analysis	-0.05	0.23	0.21					2.3×10^{-08}		
rs3011366	10	122554701	G/A					DMDSC data freeze 1	1.35	0.32
				DMDSC data freeze 2	1.02	0.57	0.10	0.07	0.25	
				UKBB (SAS)	2.44	0.68	0.08	3.4×10^{-04}	0.85	
				Meta-analysis (SAS)	1.44	0.26	0.09	3.3×10^{-08}	0.0001	
				GoDARTS	0.21	0.11	0.01	0.36		
				GoSHARE	-0.31	1.09	0.01	0.77		
				UKBB (Europeans)	-0.21	0.41	0.01	0.60		
				Meta-analysis (Europeans)	-0.09	0.08	0.01	0.20		
				Transethnic meta-analysis	0.21	0.07	0.01	0.01		

CHR, chromosome; EA, effect allele; EAF, effect allele frequency; NEA, noneffect allele; POS, position; SAS, South Asians.

being largely observed in those diagnosed before 50 years of age in both ancestral groups (Table 1). This finding is consistent with the concept that early-onset disease would have a stronger genetic component; indeed, when we looked at the overall heritability estimates for age at diagnosis of T2D, the heritability was much stronger in the younger South Indian population with diabetes compared with the more elderly European population with diabetes. We also found evidence for ethnic-specific signals that were associated with an early age at diagnosis of T2D in South Asian Indians that were very rare in the European cohort. Our findings emphasize and support our recently reported finding that South Indians have greater genetic β -cell dysfunction compared with Europeans (37).

The role of β -cell function as a driver for the early age at onset of T2D in South Indians is well supported from our ethnic-specific *TCF7L2* and *CDKAL1* signals. The *TCF7L2* variation has led to an upsurge risk of early-onset T2D among African Americans (38). It is worth highlighting that the current study reinforces earlier studies on South Asian T2D genetics (39). In addition, *WDR11* has previously been associated with fasting glucose (32), T2D susceptibility (32), and youth-onset T2D (33).

Strengths and Limitations

One of the strengths of this study is that we address the genetic basis for the large differences in age at diagnosis of T2D between two ethnic groups. To date, this study is the first that demonstrates the genome-wide PRS of age at diagnosis of T2D in South Indians. We believe that our PRSs derived from Asian I

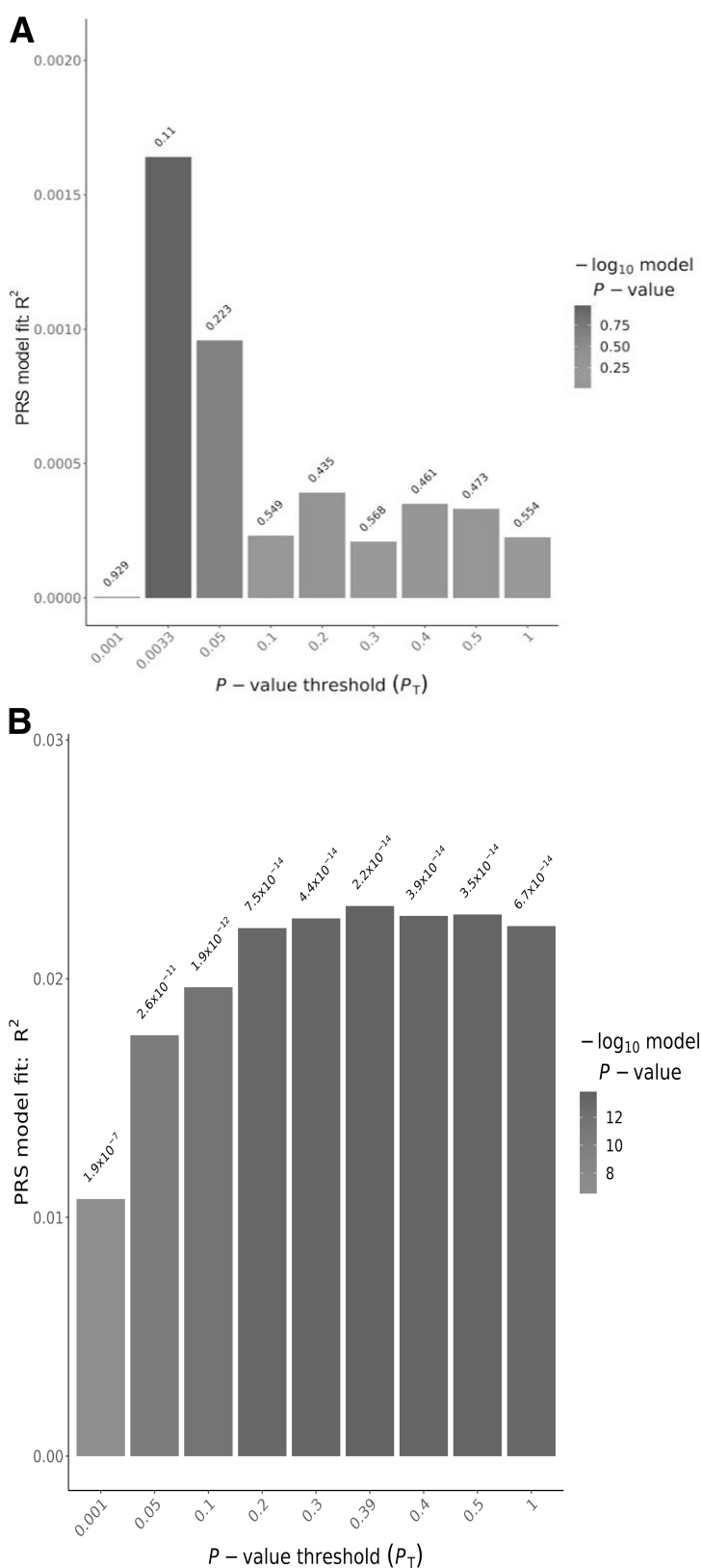


Figure 2—Performance of South Indian PRS in the European population. PRSs were generated using South Indian summary data from DMDSC 1 and tested in European samples GoSHARE ($N = 4,155$) and a South Indian independent cohort DMDSC 2 ($N = 2,494$) for polygenic risk prediction of age at onset of T2D.

