Diab-20-0162

Type 1 diabetes in Africa: an immunogenetic study in the Amhara of NW Ethiopia

Running title: Type 1 diabetes in rural Africa

Shitaye A Balcha1\*, Abayneh G Demisse2\*, Rajashree Mishra3,4,5\*, Tanwi Vartak6\*, Diana L Cousminer3,5,7\*, Kenyaita M Hodge3,5, Benjamin F Voight 7,8, Kim Lorenz 7,8,Stanley Schwartz9, , Samuel T Jerram6, Arla Gamper10, Alice Holmes11, Hannah F Wilson12, Alistair JK Williams12,\*\*, Struan FA Grant3,5,7,13,14,\*\*, Richard David Leslie6,\*\*, David IW Phillips15\*\*, Elisabeth R Trimble16\*\* ϯ

1Department of Internal Medicine, Gondar University Hospital, Gondar, Ethiopia

2Department of Pediatrics and Child Health, School of Medicine, University of Gondar, Gondar, Ethiopia

3Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, PA, USA

4Graduate Group in Genomics and Computational Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

5Center for Spatial and Functional Genomics, The Children's Hospital of Philadelphia, Philadelphia, PA, USA

6Blizard Institute, Queen Mary University of London, London, UK

7Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

8 Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

9Main Line Health System, Wynnewood, PA, USA

10Severn Postgraduate School of Primary Care, HEE, Park House, Bristol, UK

11Avon and Wiltshire Mental Health Partnership NHS Trust, Clevedon, Somerset, UK

12Diabetes and Metabolism, Transitional Health Sciences, Southmead Hospital, University of Bristol, UK

13Institute for Diabetes, Obesity and Metabolism, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

14 Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

15Medical Research Council’s Lifecourse Epidemiology Unit, Southampton General Hospital, Southampton, UK

16Centre for Public Health, Institute of Clinical Science, Queen’s University Belfast, Belfast, UK

\*equal first authors; \*\*senior authors, ϯcorresponding author

Corresponding author: Professor Elisabeth R Trimble, Centre for Public Health, Queen’s University Belfast, Grosvenor Road, Belfast BT12 6BA, UK. Tel: +44 (0)2890660125 email: [e.trimble@qub.ac.uk](about:blank); ORCHID ID: 0000-0002-7385-0415

Word count: 3415 Tables /Figures: Main paper, 3 tables, 2Figs; Suppl, 3 tables,3Figs

**ABSTRACT**

**Aim**

To characterise the immunogenic background of insulin-dependent diabetes in a

resource-poor rural African community. The study was initiated because reports of low autoantibody prevalence and phenotypic differences from European-origin

cases with type 1 diabetes raise doubts as to the role of autoimmunity in this and similar populations.

**Methods**

A prospective study of consecutive, unselected cases of recently diagnosed, insulin-dependent diabetic subjects (n=236, ≤35 years) and controls (n=200) was carried out

in the ethnic Amhara of rural NW Ethiopia. We assessed their demographic and socio-economic characteristics, measured non-fasting C-peptide, diabetes-associated autoantibodies, and *HLA-DRB1* alleles. Leveraging genome-wide genotyping we performed both a principal component analysis (PCA) and, given the relatively modest sample size, a provisional genome-wide association study (GWAS). Type 1 diabetes genetic risk scores (GRS) were calculated to compare their genetic background with known European type 1 diabetes determinants.

**Results**

Patients presented with stunted growth, low BMI, and were insulin sensitive; only 15.3% had diabetes onset ≤15 years. C-peptide levels were low but not absent. With clinical diabetes onset at ≤15, 16-25, and 26-35 years, 86.1%, 59.7% and 50.0% were autoantibody positive, respectively. Most had GADA as a single antibody; the prevalence of IA-2A and ZnT8A positivity was low in all age groups. PCA showed that the Amhara genomes were distinct from modern European and African genomes. *HLA-DRB1\*03:01* (*P*=0.0014) and *HLA-DRB1\*04* (*P*=0.0001) were positively associated with this form of diabetes while *HLA-DRB1\*15* was protective (*P*<0.0001). The mean type 1 diabetes GRS (derived from European data) was higher in patients than controls (*P*=1.50x10-6). Interestingly, despite the modest sample size, autoantibody positive patients revealed evidence of association with SNPs in the well-characterised MHC region, already known to explain half of type 1 diabetes heritability in Europeans.

**Interpretation**

The majority of patients with insulin-dependent diabetes in rural NW Ethiopia have the immunogenetic characteristics of autoimmune type 1 diabetes. Phenotypic differences between type 1 diabetes in rural NW Ethiopia and the industrialised world remain unexplained.

**Key words:** Africa, autoantibodies, Ethiopia, genomes, HLA, rural, type 1 diabetes

**Research in Context**

**What is already known about the subject?**

In sub-Saharan Africa:

* Little is known about the immunogenetic basis of type 1 diabetes.

* There are suggestions that the disease is associated with low autoantibody prevalence, which raises doubts about the role of autoimmunity in its pathogenesis.
* Genetic diversity is enormous in Africa but few studies have defined the specific genetic risk factors for type 1 diabetes.

**What is the key question?**

Does insulin-dependent diabetes among the Amhara in NW Ethiopia, a typical resource-poor community in sub-Saharan Africa, have an autoimmune basis?

**What are the new findings?**

* Recently diagnosed type 1 diabetes patients in this community have a high percentage of autoantibody positivity.
* The genome of the Amhara is distinct from modern Europeans and Africans.
* The Amhara have the same immunogenetic features of autoimmune diabetes

as adolescent- and adult-onset type 1 diabetes in European-origin populations.

**How might this impact the clinical practice in the foreseeable future?**

This study sheds light on the characteristics of type 1 diabetes in rural populations of low and middle income countries in sub-Saharan Africa and underlines the need to develop sustainable infrastructure for its diagnosis and ongoing management.

Type 1 diabetes is poorly characterised in many low and middle- income countries of sub-Saharan Africa; specifically, there has been uncertainty about whether its pathogenesis is similar to the classic form of the disease found in industrialised countries. Striking phenotypic differences have been reported from several locations in sub-Saharan Africa; these include a low incidence in the pre-pubertal years with an age-specific peak in the third decade [1,2], strong associations with low socioeconomic status [3], skewing of gender ratios with male predominance in some settings [4], and reported low autoantibody prevalence, suggesting a diminished role for autoimmune mechanisms in its aetiology compared to classic type 1 diabetes [5-7].

Classic type 1 diabetes is an autoimmune disease resulting from the interaction between genetic susceptibility [8] and the environment. Many loci have been shown to confer risk, of which the HLA Class ll genes remain the most important. Although the vast genetic diversity of Africa is well known, [9,10] very little is known about how, or if, this alters the genetic risk for type 1 diabetes in this region. Non-genetic, environmental factors also contribute to the pathogenesis of type 1diabetes, their important role being inferred from the low concordance rate for the disease in identical twins, often less than 50%. [11] Environmental factors range from a putative role of viral infections in pancreatic β cell death, see review by Op de Beek and Eizirik, [12] to nutritional factors which may interact with genetic susceptibilities to determine disease risk. [13] There is a significant body of experimental evidence showing that moderate to severe, life-long under-nutrition starting *in utero* affects pancreatic development and function by, inter alia, silencing key growth and differentiation factors [14-16], and by having an effect on the immune system [17]. Ethiopia is an example of a country which has a long history of repeated famines and approximately 40% of children have evidence of nutritional stunting. [18] In these communities under-nutrition starts *in utero* and continues throughout life. The widespread experience of under-nutrition in some areas of sub-Saharan Africa, taken in conjunction with the low levels of islet-cell autoimmunity in many historic reports from Africa, [5-7] has raised the possibility that insulin-dependent diabetes in these populations may have a nutritional origin. Those with the lowest BMIs (15-16Kg/m2) were said to have ‘protein-deficient pancreatic diabetes’, [19] a type of malnutrition-related diabetes which has since been removed from WHO classifications. So, although there is agreement that environmental factors are important, it is not easy to prove their aetiological significance in type 1 diabetes, probably because genetic factors are stable and environmental factors usually change with time.

On account of uncertainties about the role of autoimmunity in the pathogenesis of type 1 diabetes in sub-Saharan Africa, we have carried out a prospective, immunogenetic investigation in a consecutive, unselected cohort of newly diagnosed insulin-dependent diabetic subjects from a community in Ethiopia. The genetic objectives were to investigate the genome-wide underpinnings of type 1 diabetes and the leading European type 1 diabetes risk loci in this population. The study was carried out in a rural community, the Amhara of NW Ethiopia, whose socioeconomic conditions are typical of many regions in sub-Saharan Africa.

**Patients and Methods**

**Setting.**

The study was based in Gondar, Ethiopia, 750km northwest of the capital, Addis Ababa, which has a central university hospital and nine stable satellite/peripheral health centres providing care for all patients with diabetes in a predominantly rural health zone comprising 2.6 million people. In order for the study to be representative of both the rural and urban population, the region chosen had to have a stable healthcare infrastructure that was inclusive of the 90% of the population who live in rural areas. The entire diabetes service in this region has been developed and overseen by the same consultant physician for more than 30 years; thus, diagnosis and treatment at both rural and urban clinics have been carried out under the oversight of the same clinical team with the same treatment and management plan. Epidemiological studies of the diabetes care and outcomes in this region have been extensively described [1, 4, 20].

**Participants.**

Patients were entered sequentially into this prospective study without selection bias.All patients and controls were from the Amhara, the second largest ethnic group in Ethiopia (approximately 38% of the total Ethiopian population); their language is Amharic, a Semitic language in the Afroasian group of languages [21] and, phenotypically, the Amhara have many Caucasoid features. This report includes data from patients aged up to 35 years with insulin-requiring diabetes and attending the University of Gondar Diabetes and Paediatric Clinics and associated rural clinics. All patients presented with a high plasma glucose and degree of metabolic decompensation; all complained of weight loss (or parents worried about a very sick child), polydipsia, and polyuria; most of the children and a high percent of adults (50%) were ketoacidotic at presentation. All patients included in the study required insulin treatment continuously from first presentation. Controls were hospital attendees selected in the same manner as previously (ref 3); eligibility criteria for controls included no history of diabetes, normal random plasma glucose, aged up to 35 years, and belonging to the Amhara ethnic group. Control samples were used only for autoantibody, HLA, and genome investigations.

**Clinical characterisation.**

After metabolic stabilisation, a questionnaire was verbally administered in Amharic to record details of the duration and treatment of their diabetes, details of their education, occupation and socioeconomic circumstances, and height and weight measured as previously described [3]. Venous blood samples for study purposes were obtained at an average of 2.5 months (range 1-7) from diagnosis using vacutainers, and plasma separation was carried out on site; samples were stored at -70°C, initially in Gondar and subsequently in the UK.

**Laboratory analyses.**

Autoantibodies to GAD (GADA), IA-2 (IA-2A) and ZnT8 (ZnT8A) were measured by radiobinding assays as previously described [24,25]. All samples found GADA positive using full-length GAD65 were re-assayed using truncated GAD65(96-585) radiolabel [26], as 17 controls were positive for GADA in the full length GAD assay, the number was reduced to 4 in the truncated GAD assay; there was no difference in the absolute number of patients positive for GADA in the two assays. Only results for truncated GAD65(96-585) were used in subsequent analyses.C-peptide was measured by Roche ‘ECLIA’ C-peptide chemiluminescence assay on a Cobas 8000 E602 machine, with a minimal detection limit of 0.010 µg/L.

**Genotyping and Quality Control****.**

Patients and controls were genotyped on the Infinium OmniExpress Exome Beadchip platform (Illumina, San Diego, CA, USA) at the Children’s Hospital of Philadelphia Center for Applied Genomics (Philadelphia, PA, USA). Quality control was performed using PLINK [28], excluding individuals with discordant gender information, duplicate individuals, and individuals with missing genotype >5%. Single nucleotide polymorphisms (SNPs) with a call rate <95%, minor allele frequency <1%, and Hardy-Weinberg equilibrium *P*<10-5 were removed (708,143 SNPs remained).

*HLA* *DR* selected genotypes were measured by the method of Bunce et al [27]; due to the remaining limited volume of extracted DNA, 188 of 236 patients and 152 of 200 control patients were typed.

**Statistical analysis.**

BMI was calculated as weight (kg) divided by height squared (m2). Age and sex-adjusted height and BMI Z-scores were derived using WHO Anthro software (Version 3.2.2) [https://www.who.int/growthref/tools/en/](about:blank)) [22]. Whole body bioimpedence was measured using a Bodystat metre Bodystat Ltd, Douglas, Isle of Man) and fat mass was calculated by the method of Kotler et al. [23]. Differences in the anthropometric, metabolic, autoantibody or HLA status of patients were tested by t-tests or analysis of variance for continuously distributed variables (using log transformation where appropriate) or by the chi-squared tests for discrete variables. P values of <0.05 were considered to be statistically significant.

We carried out agenome-wide association study (GWAS) using a univariate linear mixed model within GEMMA [29], which accounts for population stratification and relatedness using the Wald test. Additionally, fifty-five established type 1 diabetes-implicated signals and their proxy SNPs were tested (11,748 SNPs) [30], and 403 established type 2 diabetes-implicated variants and their proxy SNPs were also tested (24,926 SNPs) [31]. Proxy SNPs were found using raggr ([http://raggr.usc.edu](about:blank)) with a linkage disequilibrium threshold of r2<0.8 in the European and African populations. The association tests were also performed in a restricted set of cases positive for at least one autoantibody.

A type 1 diabetes Genetic Risk Score (GRS) was calculated using PLINK by multiplying the number of risk-increasing alleles by the natural log of the odds ratio (OR) at each locus and summing the OR across risk loci for each individual. We included 19 SNPs (Supplementary Table 1) in the GRS, using weights as previously described [32]. The distribution of the type 1 diabetes GRS was compared for all diabetes cases versus controls and for autoantibody–positive cases using linear regression adjusting for sex and the first 4 principal components

Principal component (PC) analysis was performed using PLINK v.1.90Beta4.5 as follows. The 1000 Genomes [33] and Ethiopian genotype files were merged, removing 4277 SNPs with location conflicts. SNPs with a minor allele frequency <0.01 were removed, and LD pruning was performed at r2 < 0.2 between any two SNPs. We also removed one individual from each pair with an identity-by-descent value >0.3 (46 individuals removed). PLINK was then used to calculate PCs and R was used to plot the first three PCs.

**Ethical approval**

The study was approved by the institutional ethics review boards of Gondar College of Medicine and Health Sciences and the UK National Research Ethics Services Committee (REC reference: 14/WA/0132). Written informed consent was obtained from all participants or their parents as appropriate.

**Results**

The characteristics of the diabetes cases are shown in Table 1; results for antibody positivity in control subjects are shown separately below the patient results. Patients were investigated at an average of 2.5 months after clinical diagnosis. The overwhelming majority (89.4%) of the participants were born in rural areas around Gondar, Ethiopia. The median age at diagnosis was 21 yr.; most cases were in the 16-25 yr. age group, with only 36 (15.3%) diagnosed at 15 years of age or younger. There was a striking male preponderance in the older age groups: 72.6% in the post-pubertal, 16-25 yr. age group and 77.6% in the 26-35 yr. age group. The mean BMI SD Z scores were lower than WHO norms in all age groups; in the adult groups the BMIs were equivalent to 18.6(2.4) and 19.8(2.9) Kg/m2 in the 16-25 and 26-35 age groups, respectively. The percent of body fat was also low in all age groups. All cases had low, but detectable, non-fasting C-peptide levels. The insulin treatment doses were <1Unit/Kg for all age groups. Of the 236 cases, 112(47.5%) reported that their families’ source of income was subsistence farming or labouring while only 41 (17.4%) had paid employment or owned businesses. Educational levels were low with only 74(31.4%) reporting completed secondary education.

**Autoantibody status.**

Table 1 shows that 86.1% of cases in the 0-15 yr. age group, 59.7% of cases in the 16-25 yr. age group, and 50.0% of cases in the 26-35 yr. age group were autoantibody positive. GADA was the most common autoantibody and the most common when there was a single autoantibody; thus, 114 of 143 autoantibody positive cases had GADA as the only autoantibody. Of controls 2% were GADA positive. The prevalence of GADA declined from 80.6% in those aged ≤15 yr. to 55.6% in the 16-25 yr. age group and 43.4% in the 26-35 yr. age group. In contrast, the prevalence of ZnT8A and IA-2A were low, even in those of childhood-onset, without any age-specific trend. In comparison with the antibody (GADA) positive group, the negative group were older, had lower plasma glucose and insulin requirements, higher c-peptide levels and were more affluent with somewhat higher fat mass – supplementary Table 4.

**Genomic analysis.**

Principal component (PC) analysis (Fig1a, 1b): PC analysis is a technique which identifies the major axes of variation in genetic data. Leveraging the genome-wide genotyping data, the first PCs were plotted against the 1000 Genomes reference set [33] to visualise the relationship of the Ethiopian samples against worldwide super-populations. Comparing PC1 to PC2 (Fig. 1a), we observed a clear separation of the Ethiopian sample from the 1000 Genomes groups, and plotting PC2 vs PC3 (Fig. 1b) revealed clearly that the sample from the Amhara of Ethiopia was distinct from the European and other African populations.

HLA analysis (Table 2):*HLA-DRB1\*03:01* was positively associated with diabetes (81/188, 43.1%) compared to controls (40/152, 26.3%; OR 2.12, *P* = 0.0014), this association persisted when considering only GADA positive cases (54/102, 52.9%). *HLA-DRB1\*04* was also positively associated with diabetes (83/188, 44.1%) versus controls (37/152, 24.3%; OR 2.46; *P*=0.0001); this association persisted when GADA positive cases only were considered (51/102, 50.0%). *HLA-DRB1\*15* was strongly protective in the total diabetes group (7/188, 3.7%) compared to controls (25/152, 16.4% OR 0.20; *P*< 0.0001); this association persisted when GADA positive cases only were considered (2/102; 2.0%). In the GADA negative cases these three HLA disease-associations were less striking: *HLA-DRB1\*03:01* (31.4%, *P= 0.40*)*,* but, importantly, with significant risk with *HLA-DRB1\*04* (37.2%, *P* = 0.035) and significant protection with *HLA-DRB1\*15 (5.8%, P=0.018*).

GRS for type 1 diabetes (Fig 2, Suppl. table 1): The average TID-GRS was significantly higher for diabetes cases than controls. The total diabetes group mean was 0.189 (0.064) and the autoantibody positive group mean was 0.199 (0.067), *P* = 6.71x10-8 and *P* = 1.54x10-9 compared to controls, respectively. The type 1 diabetes-associated SNPs used in the GRS analysis are shown in Suppl. Table 1. GRS for Ab neg people….

GWAS (Suppl. figs 1, 2, 3, Table 3; Suppl. Tables 2,3): No single SNP achieved genome-wide significance (*P*<5 x10-8) when all diabetic subjects were investigated as a single group irrespective of autoantibody positivity (Suppl. Fig. 1, Suppl. Fig. 2). However, despite the modest sample size, SNPs within the *HLA* region were border-line GWAS significant (*P*<1x10-6) for autoantibody positive patients (Table 3, Suppl. Fig. 2), with the strongest signal falling in the *HLA-DQB1* locus (rs9273363, *P* = 5.13 x 10-8, Table 3). We also observed suggestively associated signals on chromosomes 4, 16, and 3 (Table 3). The GWAS analysis did not show genomic inflation (λ = 1.01; Suppl. Fig.3). We then took a candidate SNP approach and extracted all type 1 and type 2 diabetes-associated SNPs; the strongest type 1 diabetes-associated signal was in the *HLA-DQB1* region (rs1063355, *P* = 6.28 x 10-6), (Supplementary Table 2); however, no type 2 diabetes-associated loci achieved significance (Suppl. Table 3).

**Discussion**

This detailed prospective study of newly diagnosed patients in an impoverished, mainly rural population in sub-Saharan Africa shows that the majority have low C-peptide levels and low BMI, as well as diabetes-associated autoantibodies and diabetes risk alleles for type 1 diabetes. As with previous studies in this and other locations in the region, the cases have a different disease phenotype; thus, the median age of onset is later than that observed in most industrialised countries and there is a male predominance in the post-pubertal age groups.

**Autoantibody prevalence.**

The prevalence of diabetes-related autoantibodies was high, falling from 86.1% in the youngest age group to 50.0% in the oldest group. This autoantibody frequency and age dependence is similar to some [34,35], if slightly lower [36] than, other reports from industrialised countries, with all demonstrating lower prevalence of autoantibody positivity with increasing age of onset. However, whereas the majority of type 1 diabetes patients in industrialised countries have multiple diabetes-related autoantibodies, in this present Ethiopian study most had GADA alone, which was evident even in the 0-15 age group. Of note, the prevalence of autoantibodies to full length GAD was relatively high in controls, at 8.5% [mentioned in Methods section], but in contrast to cases the majority of these autoantibodies were directed to the low risk N-terminal epitope [26].) The prevalence of other autoantibodies, IA-2A and ZnT8A, was low in all age groups including those with childhood-onset disease. Our findings contrast with two urban studies from Ethiopia which found much lower levels of GADA albeit with very low or absent IA-2A; however these subjects were studied six or more years after diagnosis [7, 37]. Unfortunately, there are very few studies from sub-Saharan Africa where blood sampling for autoantibody status took place close to the time of diagnosis and involved *both* urban and rural populations, these latter being numerically important. The enormous genetic diversity in this continent also complicates comparisons [9,10,38]. A recent study of type 1 diabetes from West Africa (Cameroon) has investigated an urban group shortly after clinical diagnosis and compared them with a Belgian population. As in Ethiopia, few (9%) had childhood-onset type 1 diabetes but, in contrast to Ethiopians, only 52% of children (≤ 15 years) and 32% of adults (median age 30 years) were autoantibody positive; the majority of these had GADA, many fewer had IA2A or ZnT8A [2]. The adult Belgian diabetes subjects had a much higher incidence of multiple autoantibodies than the Cameroonian subjects, 69% versus 23%, respectively. The West Africans of Cameroon are of Bantu background, however, and belong to a different genomic group to the Amhara [9,10,38]. These findings are echoed by studies in South Africa comparing the profiles of type 1 diabetes in Black (mixed ethnicity) and White subjects. Interestingly, the peak age of clinical onset was later in Black subjects, both populations were heavier than the Amhara (average BMI for Black and White subjects, 24.0 Kg/m2 and 22.4 Kg/m2, respectively). When studied several years after clinical diagnosis (mean duration more than 5 years), overall GADA positivity was 60% and 66%, for Black and White subjects respectively; while IA2A positivity was lower in the former than the latter (19% versus 41%, respectively), especially in those aged > 21 years at diagnosis (7.3% versus 33% respectively) [39]. Taken together with our data these results [2,39] from three different regions of sub-Saharan Africa suggest that there is considerable heterogeneity in the prevalence of autoantibody profiles both within Africa and between African and European populations. These differences appear to be independent of the age of onset of diabetes and, for the present, remain unexplained.

**Genomic analysis.**

Using genome-wide genotyping data, we demonstrated that our study population is distinct from the European and other African ancestral groups; based on what is known of their demographic and linguistic history this was not unexpected [10]. GWAS and *HLA* analysis confirmed the predominant *HLA* association with type 1 diabetes-associated risk and protective alleles. By contrast, no associations were found for type 2 diabetes-associated loci. However, our sample size was low and power limited, therefore the results should be seen as relatively preliminary. In view of the autoantibody and *HLA* results we applied a GRS, based on SNPs associated with type 1 diabetes on a European genetic background. This was done with the knowledge that the results may not be as robust when applied to a different ethnic group [40]. Despite the European background of the SNPs employed, the Ethiopian diabetes group showed a significantly increased GRS compared to controls. A mean T1D-GRS score of >0.280 is indicative of type 1 diabetes in Europeans using European background SNPs [41]; in the Amhara patients the highest TID-GRS score was 0.199 in the autoantibody positive group, which was significantly higher than that of the control group. In addition, these Amhara diabetes patients showed an association with *HLA* class II alleles *HLA-DRB1\*03:01* and *HLA-DRB1\*04*, in both the total diabetes cohort as well as those with GADA, while *HLA-DRB1\*15* was strongly protective for diabetes. In short, they have the same immunogenic features of autoimmune diabetes of adolescent and adult European-origin and Chinese populations [42], while the diabetes-associated autoantibody, GADA, was similarly the dominant autoantibody and associated with *HLA-DRB1\*03:01,* consistent with heterogeneity of type 1 diabetes endotypes [43, 44]. Our observations imply a widespread commonality in GADA-dominant, HLA-associated adult-onset autoimmune diabetes, despite global variation in the precise HLA-associated genotypes. Moreover, as with adult-onset autoimmune diabetes in Europe, China and Africa, both IA2A and ZnT8A were much less frequent and, in the present Amhara cohort, even in childhood-onset insulin-dependent diabetes.

**Limitations**

The sample size in this study is necessarily small because of the difficulty of obtaining a series of newly diagnosed patients in a rural African community. While the findings are specific to the Amhara group and may not necessarily apply elsewhere, the socioeconomic conditions of this community are typical of many sub-Saharan African countries.

**Implications**

Given that the majority of patients with diabetes in this rural Ethiopian population appear to have an autoimmune basis to their diabetes, the results do not support the hypothesis suggested by ourselves and others [1, 19] that the disease could have a direct relationship to undernutrition during prenatal and early post-natal life. Although type 1 diabetes is strongly associated with low socioeconomic status and skeletal disproportion in this community [3], it remains unclear whether nutrition or related aspects of poverty influence disease development and contribute to the relatively late peak age of onset of their diabetes.Of note, communities with diverse ethnic backgrounds in sub-Saharan Africa, but without reported malnutrition, report a later peak age-incidence [2, 39]. One line of evidence does, however, point to the involvement of a complex gene-environment interaction affecting the age of disease onset. Ethiopian Jews moved in large numbers from rural areas around Gondar to urban areas of Israel in the 1980s and 1990s. After immigration to Israel the age of onset of type 1 diabetes in the young offspring of these Ethiopian emigres slowly fell (in those with at least two high risk alleles for diabetes) in proportion to the time that their Ethiopian-born parents had been resident in Israel; at the same time the incidence of type 1 diabetes in these young people rose dramatically to be one of the highest within the Jewish communities of Israel [45].

It is possible that a proportion of cases of diabetes in the antibody negative group had diabetes with a nutritional origin, albeit based on small numbers. This negative group showed weaker HLA associations (Table 2), a GRS very similar to that of the controls (suppmentary Figure 2) and phenotypically had some of the features of type 2 diabetes, including lower insulin requirements, somewhat increased fat mass and more affluent circumstances –supplementary Table 4.

Another unexplained phenotypic feature in our study was the marked male preponderance in the post-pubertal but not pre-pubertal age groups; this is not a feature of type 2 diabetes in Ethiopia [4]. A less marked post-pubertal male preponderance in type 1 diabetes has been noted previously in European cohorts [46,47]. The cause of this gender bias is not clear but may include gender-related immune and differential epigenome effects [17, 48].

In summary, the majority of insulin-dependent diabetes in the Amhara of North-West Ethiopia is autoimmune in nature and the genetic risk and protective factors for type 1 diabetes are largely common to those found in Europeans with type 1 diabetes. The results of this study and reports from other areas of sub-Saharan Africa highlight the need for a wider understanding of the gene-environment interactions giving rise to differences in timing of peak incidence, male preponderance (post-pubertal onset), and autoantibody profile in type 1 diabetes.

**Acknowledgements**

The authors thank Cornelius Self, research student, for assistance with computing and Margaret McDonnell and Alastair Magill, Department of Clinical Biochemistry, Royal Victoria Hospital, Belfast for laboratory support. We are grateful to Dr Vito Lampasona, San Raffaele Scientific Institute, Milan and Professor Ake Lernmark, Lund University, Sweden for the plasmids used for islet autoantibody measurement.

Funding for the study:

DIWP and ERT received a grant from the Association of Physicians of Great Britain and Ireland; D.L.C.is supported by a grant from the NIH, NIH1K99HD099330-01; S.F.A.G.is supported by the NIH (RO1 DK085212) and the Daniel B. Burke endowed Chair for Diabetes Research; S.T.J. and RDL are funded by Type 1 Diabetes UK (DUK 19/0005951) RDL is funded by the EFSD(MMBP1C3R), EU (EU-FP7:282510), St Bartholomew’s Hospital, London, Charity (OGA011582), T.V. funded by the British Diabetic Twin Trust.

Duality of Interest. None relevant to this project

Author Contributions S.A.B., A.G.D., A.G., A.H., identified patients and controls, were responsible for measuring and recording demographic data and oversaw the procurement and initial storage of patient samples. R.M., D.L.C. K.M.H., S.S., and S.F.A.G. responsible for designing, measuring, and analysis of ~~immuno~~genetic assays. R.M. and D.L.C, and KL. were also involved in manuscript preparation. T.V., data analysis; S.T.J., genetic and data analysis; R.D.L. genetic design and analysis, manuscript preparation; A.J.K.W. and H.F.W., autoantibody assay design and analysis; D.I.W.P., E.R.T., conception, design, data collection, data analysis, manuscript preparation.

References

1. Balcha S A, Phillips DIW, Trimble ER (2018) Type 1 diabetes in a resource-poor setting: malnutrition related, malnutrition modified, or just diabetes? Curr Diab Rep 18: 47 doi: 10.007/s11892-018-1003-7

2. Asanghanwa M, Gorus FK, Weets I, et al (2014) Clinical and biological characteristics of diabetic patients under age 40 in Cameroon: relation to autoantibody status and comparison with Belgian patients. Diabetes Res Clin Pract 103: 97-105

3. Fekadu S, Yigzaw M, Alemu S, et al. (2010) Insulin-requiring diabetes in Ethiopia: associations with poverty, early undernutrition and anthropometric disproportion. Europ J Clin Nutr 64: 1192-1198

4. Alemu S, Dessie A, Seid E, et al (2009) Insulin-requiring diabetes in rural Ethiopia: should we reopen the case for malnutrition-related diabetes? Diabetologia 52: 1842-1845

5. Panz VR, Kalk WJ, Zouvanis M, Joffe BI (2000) Distribution of autoantibodies to glutamic acid decarboxylase across the spectrum of diabetes mellitus seen in South Africa. Diabet Med 17: 524-527

6. Lutale JJK, Thordarson H, Holm PI, Eide GE, Vetvik K (2007).Islet cell autoantibodies in African patients with Type 1 and Type 2 diabetes in Dar es Salaam Tanzania: a cross sectional study. J Autoimmune Dis 4:4 doi: 10.1186/1740-2557-4.4

7. Siraj ES, Gupta MK, Yifter H, et al (2016) Islet-cell associated autoantibodies in Ethiopians with diabetes mellitus. J Diabetes and its Complications 30: 1039-42

8. Barrett JC, Clayton D, Concannon P, et al (2009) Genome-wide association study and meta-analysis finds over 40 loci affect the risk of type 1 diabetes. Nat Genet.41:703-707

9. Campbell MC and Tishkoff SA (2008) African Genetic Diversity: Implications for human demographic history, modern human origins, and complex disease mapping. Annu Rev Genomics Hum Genet 9: 403-433

10. Gurdasani D, Carstensen T, Tekola-Ayele F et al. (2015) The African Genome Variation Project shapes medical genetics in Africa. Nature 517:327-332

11. Redondo MJ, Yu L, Hawa M, et al (2001) Heterogeneity of type 1 diabetes: analysis of monozygotic twins in Great Britain and the United States. Diabetologia 44:354-362

12**.** Op de Beeck A, Eizirik DL (2016) Viral infections in type 1 diabetes mellitus - why the β cells?

13 Carlsson A, Kockum I, Lindblad B et al (2012) Low risk HLA-DQ and increased body mass index in newly diagnosed type 1 diabetes children in the Better Diabetes Diagnosis study in Sweden. Int J Obes (Lond) 36: 718-724

14. Waterland RA, Lin JR, Smith CA, Jirtle RL (2006) Post-weaning diet affects genomic imprinting at the insulin-like growth factor 2 (Igf2) locus. Hum Mol Genet 15: 705-716

15. Park JH, Stoffers DA, Nicholls RD, Simmons RA (2008) Development of type 2 diabetes following intrauterine growth retardation in rats is associated with progressive epigenetic silencing of Pdx1. J Clin Invest `118: 2316-2324

16. Sandovici I, Smith NH, Nitert MD, et al (2011) Maternal diet and aging alter the epigenetic control of a promoter-enhancer interaction at the HNF4α gene in rat pancreatic islets. Proc Natl Acad Sci USA 108: 5449-5454

17. Moore SE, Cole TJ, Collinson AC, Poskitt EM, McGregor IA, Prentice AM (1999). Prenatal or early postnatal events predict infectious deaths in young adulthood in rural Africa. Int J Epidemiol 28:1088-95

18. Amare ZY, Endris M, Mehari AB (2019) Determinants of nutrirional status among children under age 5 in Ethiopia : further analysis of the 2016 Ethiopia demographic and health survey. Globalization and Health 15:62/ doi.org/10.1186/s12992-019-0505-7

19. Abdulkadir J, Mengesha B, Welde Gebriel Z, et al(1990) The clinical and hormonal (C-peptide and glucagon) profile and liability to ketoacidosis during nutritional rehabilitation in Ethiopian patients with malnutrition-related diabetes mellitus. Diabetologia 33: 223-227

20. Alemu S, Dessie A, Tsegaw A, et al.(2015) Retinopathy in Type 1 diabetes: major differences between rural and urban dwellers in northwest Ethiopia. Diabetes Res Clin Pract 109: 191-198

21. Fort M, de Stefano G-F, Cambon-Thomsen A, et al (1998) HLA class ll allele and haplotype frequencies in Ethiopian Amhara and Oromo populations. Tissue Antigens 51: 327-336

22. WHO Geneva 2009.WHO AnthroPlus software for personal computers

([https://www.who.int/growthref/tools/en/](about:blank))

23. Kotler DP, Burastero S, Wang J, Pierson Jr RN (1996) Prediction of body cell mass, fat-free mass, and total body water with bioelectrical impedance analysis: effects of race, sex and disease. Am J Clin Nutr 64: 489s-497s

24. Bonifacio E, Yu L, Williams AK, et al (2010). Harmonization of glutamic acid decarboxylase and Islet Antigen-2 autoantibody assays for National Institute of Diabetes and Digestive and Kidney Diseases consortia. J Clin Endocrinol Metab 95: 3360-7

25. Long AE, Gillespie KM, Rokni S, Bingley PJ, Williams AJK (2012) Rising incidence of Type 1 diabetes is associated with altered immunophenotype at diagnosis. Diabetes 61: 683-686

26. Williams AJK, Lampasona V, Wyatt R, et al (2015) Reactivity to N-terminally truncated GAD65(96-585) identifies GAD autoantibodies that are more closely associated with diabetes progression in relatives of patients with type 1 diabetes. Diabetes 64: 3247-3252

27. Bunce M, O’Neill CM, Barnardo MC, et al (1995) Phototyping : comprehensive DNA typing for HLA-A, B, C, DRB1, DRB3, DRB4, DRB5 & DQB1 by PCR with 144 primer mixes utilizing sequence-specific primers (PCR-SSP).Tissue Antigens 46: 355-367

28. 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

29. Zhou X, Stephens M (2012) Genome-wide efficient mixed-model analysis for association studies. Nat Genet. 44: 821-824

30. Mishra R, Chesi A, Cousimer DL et al (2017) Relative contribution of type 1 and type 2 diabetes loci to the genetic etiology of adult-onset, non-insulin-requiring autoimmune diabetes. BMC Med https://doi.org/10.1186/s12916-017-0846-0

31. Mahajan A, Taliun D, Thurner M et al (2018) Fine-mapping of an expanded set of type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nat Genet.50: 1505-1513

32. Oram RA, Patel K, Hill A, et al.(2016) A Type 1 diabetes Genetic Risk score can aid discrimination between Type 1 and Type 2 diabetes in young adults. Diabetes Care 39: 337-344

33. Clarke L, Fairley S, Zheng-Bradley X et al (2017) The international Genome sample resource (IGSR): A worldwide collection of genome variation incorporating the 1000 Genomes project data. Nucleic Acids Res 45: D854-D859

34. Wang J, Miao D, Babu S et al (2007) Prevalence of autoantibody negative diabetes is not rare at all ages and increases with older age and obesity. J Clin Endocrinol Metab 92: 88-92

35. Kim CS, Song MK, Park JS et al (2007) The clinical and immunogenetic characteristics of adult-onset type 1 diabetes mellitus in Korea. Acta Diabetol 44: 45-54

36. Bravis V, Kaur A, Walkey HC et al. (2018) Relationship between islet autoantibody status and the clinical characteristics of children and adults with incident type 1diabetes in a UK cohort. BMJ Open 8: e020904. doi.10.1136/bmjopen-2017-020904

37. Gill GV, Tekle A, Reja a et al (2011) Immunological and C-peptide studies of patients with diabetes in northern Ethiopia: existence of an unusual subgroup possibly related to malnutrition. Diabetologia 54: 51-57

38. Fan S, Kelly DE, Beltram MH, et al (2019) African evolutionary history inferred from whole genome sequence data of 44 indigenous African populations. Genome Biol 20:82 doi.org/10.1186/s13059-019-1679-2

39. Padoa CJ, Rheeder P, Pirie FJ, Motala AA, van Dyk JC, Crowther NJ (2019)Identification of a subgroup of black South Africanswith type 1 diabetes who are older at diagnosis but have lower levels of glutamic acid decarboxylase and islet antigen 2 autoantibodies. Diab Med doi: 10.1111/dme.14204ePub ahead of print.

40. Perry DJ, Wasserfall CH, Oram RA, et al (2018) Application of a genetic risk score to racially diverse Type 1 diabetes populations demonstrates the need for diversity in risk-modelling. Sci Rep. 8:4529.doi: 10.1038/s41598-018-22574-5

41. Patel KA, Oram RA, Flanagan SE et al. (2016) Type 1 diabetes genetic risk score: a novel tool to discriminate monogenic and type1 diabetes. Diabetes 65: 2094-2099

42. Zhu M, Xu K, Chen Y et al (2019) Identification of novel T1D risk loci and their association with age and islet function at diagnosis in autoantibody-positive T1D individuals: Based on a two-stage genome-wide association study. Diabetes Care 42:1414-1421.

43. Leslie RD, Grant SFA (2018) The dynamic origins of Type 1 diabetes. Diabetes Care 41: 2441-2

44. Battaglia M, Ahmed S, Anderson MA et al. (2020) Introducing the endotype concept to address the challenge of disease heterogeneity in type 1 diabetes. Diabetes Care 43: 5-12

45. Zung A, Elizur M, Weintrob N et al. (2004) Type 1 diabetes in Jewish Ethiopian immigrants to Israel: HLA Class II immunogenetics and contribution of new environment.

46. Pundziute-Lycka A, Dahlquist G, Nystrom L, et al (2002) The incidence of Type 1 diabetes has not increased but shifted to a younger age at diagnosis in the 0-34 years group in Sweden 1983-1998. Diabetologia 45: 784-791

47. Kyvik KO, Nystrom L, Gorus F et al (2004) The epidemiology of Type 1 diabetes mellitus is not the same in young adults as in children. Diabetologia 47: 377-384

48. Khulan B, Cooper WN, Skinner BM et al (2012) Periconceptual maternal micronutrient supplementation is associated with widespread gender related changes in the epigenome: a study of a unique resource in the Gambia. Hum Mol Genet 21:2086-2101

Table 1

Metabolic characteristics and autoantibody status at presentation of Ethiopian patients with type 1 diabetes

ƚ Based on WHO standards, see Methods. §GADA – antibodies against truncated GAD ie. GAD65(96-585)

ǂ Note: Non-diabetic controls n=200: positive for autoantibodies, GADA n=4, IA-2A n=2, ZnT8A n=5; any antibody n=11

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Age at onset (yr.) | | | | *P*-value |
|  | 0-15 | 16-25 | 26-35 | All |  |
| Number tested | 36 | 124 | 76 | 236 |  |
| Male (%) | 16(44.4) | 90(72.6) | 59(77.6) | 165(69.9) | 0.001 |
| Blood glucose at diagnosis, mmol/l, median (IQR) | 29.3(27.1-33.3) | 29.7(24.0-33.3) | 26.1(20.4-32.2) | 28.3(22.2-33.3) | 0.03 |
| Insulin dose after stabilisation, units/kg, mean (SD) | 0.92(0.37) | 0.79(0.23) | 0.66(0.18) | 0.77(0.26) | <0.001 |
| Diabetes duration, months, median (IQR) | 3(1-7) | 2(1-6) | 2(1-7) | 2.5(1-7) | ns |
| C-peptide, μg/L, median (IQR) | 0.46(0.32-1.09) | 0.77(0.33-1.35) | 0.98(0.46-1.87) | 0.80(0.34-1.42) | 0.03 |
| Height, SD Z score, mean(SD) ƚ | -1.49(1.09) | -1.18(0.91) | -1.09(0.83) | -1.20(0.92) | ns |
| BMI, SD Z score, mean (SD) ƚ | -1.20(1.14) | -1.44(1.11) | -0.97(1.16) | -1.25(1.15) | 0.02 |
| % body fat, mean (SD) | 6.2(8.7) | 11.9(7.7) | 13.7(6.5) |  | <0.001 |
| Rural birth, n (%) | 30(83.3) | 115(92.7) | 66(86.8) | 211(89.4) | ns |
|  |  |  |  |  |  |
| Autoantibody prevalence, n (%)ǂ |  |  |  |  |  |
| GADA§ | 29(80.6) | 69(55.6) | 33(43.4) | 131(55.5) | 0.001 |
| IA2A | 1(2.8) | 5(4.0) | 2(2.6) | 8(3.4) | ns |
| ZnT8A  Any antibody present | 6(16.7)  31(86.1) | 11(8.9)  74(59.7) | 7(9.2)  38(50.0) | 24(10.2)  143(60.6) | ns  0.001 |



Table 3.

GWAS: Border-line genome-wide significant signals associated with autoantibody positive diabetes cases (*P*<5x10-6).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **SNP** | **Chromosome** | **Position\*** | **Minor/Major Allele** | **MAF ƚ in cases** | **MAF in controls** | **OR** | **CI** | **P** | **Locus** |
| rs9273363 | 6 | 32626272 | A/C | 0.445 | 0.217 | 1.281 | 1.226-1.339 | 5.13E-08 | *HLA-DQB1* |
| rs2760985 | 6 | 32566398 | A/G | 0.261 | 0.088 | 1.363 | 1.288-1.443 | 1.16E-07 | *HLA-DRB1* |
| rs9268528 | 6 | 32383108 | G/A | 0.407 | 0.358 | 1.261 | 1.208-1.316 | 1.57E-07 | *BTNL2* |
| rs9268542 | 6 | 32384721 | G/A | 0.407 | 0.365 | 1.256 | 1.203-1.312 | 3.23E-07 | *BTNL2* |
| rs11947273 | 4 | 92544404 | T/C | 0.336 | 0.445 | 0.811 | 0.778-0.846 | 1.25E-06 | *CCSER1* |
| rs2187818 | 6 | 32395568 | G/T | 0.360 | 0.427 | 0.805 | 0.770-0.841 | 1.48E-06 | *HLA-DRA* |
| rs9268585 | 6 | 32397403 | T/G | 0.360 | 0.427 | 0.805 | 0.770-0.841 | 1.48E-06 | *HLA-DRA* |
| rs4784939 | 16 | 58468211 | T/C | 0.309 | 0.146 | 1.280 | 1.216-1.347 | 2.43E-06 | *GINS3* |
| rs996482 | 3 | 20583674 | T/C | 0.318 | 0.172 | 0.817 | 0.782-0.853 | 4.06E-06 | *SGOL1* |

\*Base pair position reported for genome build 37.

MAF = minor allele frequency. Note that minor allele is the effect allele.

Statistics: Linear mixed model with Wald test in GEMMA

Supplementary Table 1. Genetic Risk Score calculation : SNPs with established type 1 diabetes association (Oram et al, ref 25)

used in genetic risk score calculation

|  |  |  |  |
| --- | --- | --- | --- |
| Chromosome | SNP | Risk Allele | Weight |
| 1 | rs3024505 | G | 0.17 |
| 2 | rs3087243 | G | 0.2 |
| 2 | rs1990760 | T | 0.15 |
| 6 | rs2395029 | T | 0.92 |
| 6 | rs9388489 | G | 0.16 |
| 10 | rs10509540 | T | 0.29 |
| 10 | rs11594656 | T | 0.17 |
| 12 | rs2292239 | T | 0.3 |
| 14 | rs1465788 | C | 0.15 |
| 15 | rs3825932 | C | 0.15 |
| 15 | rs17574546 | C | 0.13 |
| 16 | rs4788084 | G | 0.15 |
| 18 | rs1893217 | G | 0.18 |
| 18 | rs763361 | T | 0.15 |
| 19 | rs425105 | T | 0.15 |
| 20 | rs2281808 | C | 0.1 |
| 22 | rs5753037 | T | 0.1 |
| 6 | rs2187668,rs7454108 | *DR3/DR4-DQ8* | 3.87 |
| 6 | rs2187668,rs7454108 | *DR3/DR3* | 3.05 |
| 6 | rs2187668,rs7454108 | *DR4-DQ8/DR4-DQ8* | 3.09 |
| 6 | rs2187668,rs7454108 | *DR4-DQ8/X* | 1.95 |
| 6 | rs2187668,rs7454108 | *DR3/X* | 1.51 |

Supplementary Table 2

Type 1 diabetes-related signals in all cases with diabetes at an uncorrected P<0.05.

No signals remained significant after multiple test correction (*P*<4.26 x10-6).

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **SNP** | **Locus** | | **Chr** | **Position** | **Minor/Major Allele** | **MAF**\* **in Cases** | **MAF**\* **in Controls** | **OR** | **CI** | **P-value** |
| rs3024493 | *IL10* |  | 1 | 206943968 | A/C | 0.083 | 0.135 | 0.88 | 0.83-0.93 | 3.21 x 10-2 |
| rs1063355 | *HLA-DQB1* |  | 6 | 32627714 | T/G | 0.275 | 0.420 | 0.85 | 0.81-0.88 | 6.04 x 10-5 |
| rs4235991 | *DLL1* |  | 6 | 170382923 | G/A | 0.233 | 0.321 | 0.90 | 0.86-0.95 | 1.76 x 10-2 |
| rs667899 | *PRKCQ* |  | 10 | 6464156 | A/G | 0.294 | 0.201 | 1.13 | 1.07-1.18 | 9.96 x 10-3 |
| rs229541 | *C1QTNF6* |  | 22 | 37591318 | G/A | 0.401 | 0.329 | 1.10 | 1.05-1.14 | 3.06 x 10-2 |

**T**ype 1 diabetes-related signals in cases positive for at least one autoantibody at an uncorrected P<0.05.

No signal significant after multiple test correction (*P*<4.26 x10-6).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **SNP** | **Locus** | **Chr** | **Position** | **Minor/Major Allele** | **MAF**\* **in Cases** | **MAF**\* **in Controls** | **OR** | **CI** | **P-value** |
| rs1063355 | *HLA-DQB1* | 6 | 32627714 | T/G | 0.229 | 0.420 | 0.81 | 0.77-0.85 | 6.28 x 10-6 |
| rs667899 | *PRKCQ* | 10 | 6464156 | A/G | 0.297 | 0.201 | 1.13 | 1.07-1.19 | 1.8 x 10-2 |
| rs11170466 | *ITGB7* | 12 | 53585859 | T/C | 0.131 | 0.069 | 1.17 | 1.09-1.26 | 2.7 x 10-2 |
| rs12161793 | *CSAD* | 12 | 53552475 | G/A | 0.161 | 0.095 | 1.15 | 1.08-1.23 | 3.2 x 10-2 |
| rs4235991 | *WDR27* | 6 | 170382923 | G/A | 0.241 | 0.321 | 0.91 | 0.86-0.95 | 4.7 x 10-2 |
|  |  |  |  |  |  |  |  |  |  |

\*MAF= minor allele frequency Statistics: Linear mixed model with Wald test in GEMMA

Supplementary Table 3 Type 2 diabetes-related signals in total diabetes group at an uncorrected P<0.05.

No signals were significant after multiple test correction (*P*<2.00 x10-6)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **SNP** | **Locus** | **Chr** | **Position** | **Major/Minor Allele** | **MAF**\* **in Cases** | **MAF**\* **in Controls** | **OR** | **CI** | **P-value** |
| rs1493694 | *NOTCH2* | 1 | 120526982 | T/C | 39.6% | 30.7% | 1.12 | 1.07-1.16 | 7.71 x 10-3 |
| rs4402960 | *IGF2BP2* | 3 | 185511687 | T/G | 51.9% | 44.5% | 2.19 | 2.10-2.28 | 4.33 x 10-2 |
| rs6813195 | *TMEM154* | 4 | 153520475 | T/C | 39.6% | 49.6% | 0.90 | 0.87-0.94 | 8.64 x 10-3 |
| rs3129948 | *BTNL2* | 6 | 32354644 | A/C | 27.5% | 18.3% | 1.13 | 1.08-1.19 | 7.69 x 10-3 |
| rs2968553 | *ADCK2* | 7 | 140382156 | T/C | 43.9% | 52.9% | 0.45 | 0.43-0.47 | 4.27 x 10-2 |
| rs6476842 | *GLIS3* | 9 | 4291268 | C/T | 16.6% | 23.4% | 2.49 | 2.37-2.62 | 4.82 x 10-2 |
| rs4146894 | *PLEKHA1* | 10 | 124155381 | C/T | 46.5% | 38.3% | 2.27 | 2.18-2.36 | 4.15 x 10-2 |
| rs10400343 | *HSD17B12* | 11 | 43832505 | G/A | 27.8% | 20.6% | 2.56 | 2.46-2.67 | 3.84 x 10-2 |
| rs4275659 | *ABCB9* | 12 | 123447928 | T/C | 33.7% | 47.5% | 0.87 | 0.84-0.91 | 6.27 x 10-4 |
| rs1051434 | *MPHOSPH9* | 12 | 123641200 | C/T | 33.9% | 48.9% | 0.88 | 0.84-0.91 | 1.66 x 10-4 |
| rs1879379 | *C12orf65* | 12 | 123727443 | G/A | 35.3% | 51.4% | 0.85 | 0.82-0.89 | 5.75 x 10-5 |
| rs12970134 | *MC4R* | 18 | 57884750 | G/A | 18.7% | 11.7% | 1.12 | 1.07-1.18 | 2.52x 10-2 |

Type 2 diabetes-related signals in autoantibody positive cases at an uncorrected P<0.05.

No signals were significant after multiple test correction (*P*<2.00x10-6).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **SNP** | **Locus** | **Chr** | **Position** | **Minor/Major Allele** | **MAF in Cases** | **MAF in Controls** | **OR** | **CI** | **P** |
| rs35265698 | *HLA-DRB1* | 6 | 32561334 | G/C | 28.0% | 12.5% | 1.26 | 1.19-1.33 | 2.50x10-5 |
| rs984748 | *RAB1A* | 2 | 65369964 | T/C | 24.6% | 40.5% | 0.86 | 0.83-0.90 | 4.10x10-4 |
| rs4451914 | *SIN3A* | 15 | 75739824 | C/T | 49.6% | 37.2% | 1.17 | 1.12-1.22 | 6.82x10-4 |
| rs1051431 | *MPHOSPH9* | 12 | 123645803 | A/G | 41.1% | 44.2% | 0.86 | 0.83-0.90 | 9.64x10-4 |
| rs1493694 | *NOTCH2* | 1 | 120526982 | T/C | 43.2% | 30.7% | 1.15 | 1.10-1.20 | 1.78x10-3 |
| rs2937121 | *WWP2* | 16 | 69870409 | G/T | 30.5% | 19.7% | 1.17 | 1.11-1.23 | 1.81x10-3 |
| rs1035061 | *PITPNC1* | 17 | 65647063 | G/A | 27.1% | 39.1% | 0.87 | 0.83-0.91 | 3.47x10-3 |
| rs7559672 | *ACVR2A* | 2 | 147855299 | C/T | 10.6% | 4.0% | 1.27 | 1.17-1.38 | 3.58x10-3 |
| Suppl Table 3 continued |  |  |  |  |  |  |  |  |  |
| SNP | Locus | Chr | Position | Minor/major  Allele | MAF in cases | MAF in Controls | OR | CI | P |
| rs13212734 | *RREB1* | 6 | 7037637 | A/G | 21.2% | 33.6% | 0.87 | 0.83-0.91 | 3.95x10-3 |
| rs12778642 | *HHEX* | 10 | 94464307 | T/G | 21.2% | 32.5% | 0.88 | 0.84-0.92 | 6.87x10-3 |
| rs11068780 | *WSB2* | 12 | 118476079 | T/C | 4.2% | 9.5% | 0.80 | 0.73-0.87 | 1.18x10-2 |
| rs10276674 | *DGKB* | 7 | 14922007 | C/T | 36.9% | 28.5% | 1.13 | 1.07-1.18 | 1.49x10-2 |
| rs1296328 | *PCDH18* | 4 | 137083193 | C/A | 46.6% | 36.0% | 1.12 | 1.07-1.17 | 1.50x10-2 |
| rs1552224 | *ARAP1* | 11 | 72433098 | C/A | 4.6% | 1.1% | 1.40 | 1.21-1.60 | 1.76x10-2 |
| rs1358980 | *VEGFA* | 6 | 43764551 | T/C | 27.4% | 36.5% | 0.89 | 0.85-0.94 | 1.79x10-2 |
| rs4946812 | *BEND3* | 6 | 107431688 | A/G | 14.4% | 21.9% | 0.87 | 0.82-0.92 | 1.89x10-2 |
| rs10788575 | *PTEN* | 10 | 89768584 | A/G | 12.3% | 6.6% | 1.19 | 1.10-1.28 | 1.95x10-2 |
| rs12454712 | *BCL2* | 18 | 60845884 | C/T | 19.5% | 28.8% | 0.88 | 0.84-0.93 | 1.99x10-2 |
| rs11635472 | *HERC1* | 15 | 63975690 | C/T | 38.6% | 47.1% | 0.90 | 0.86-0.94 | 2.16x10-2 |
| rs11786992 | *ESRP1* | 8 | 95685147 | C/A | 33.9% | 44.9% | 0.91 | 0.87-0.95 | 3.04x10-2 |
| rs2303108 | *ZC3H4* | 19 | 47589895 | T/C | 42.4% | 33.2% | 1.10 | 1.05-1.15 | 3.33x10-2 |
| rs6575984 | *MARK3* | 14 | 103876454 | A/G | 47.5% | 39.1% | 1.10 | 1.05-1.15 | 3.60x10-2 |
| rs943005 | *TFAP2B* | 6 | 50865820 | T/C | 12.3% | 18.6% | 0.88 | 0.83-0.94 | 3.74x10-2 |
| rs7593685 | *THADA* | 2 | 43215855 | C/T | 30.3% | 22.4% | 1.11 | 1.06-1.17 | 3.78x10-2 |
| rs10803762 | *ITGB6* | 2 | 161105876 | G/A | 33.1% | 42.7% | 0.91 | 0.87-0.95 | 3.91x10-2 |
| rs10011174 | *FBXW7* | 4 | 153495515 | G/A | 35.6% | 43.8% | 0.91 | 0.87-0.95 | 4.01x10-2 |
| rs7124681 | *CELF1* | 11 | 47529947 | A/C | 36.0% | 46.4% | 0.92 | 0.88-0.96 | 4.22x10-2 |
| rs2183825 | *LINGO2* | 9 | 28412375 | C/T | 18.6% | 25.9% | 0.90 | 0.85-0.95 | 4.41x10-2 |
| rs13332406 | *RBL2* | 16 | 53489705 | G/A | 47.0% | 36.9% | 1.09 | 1.05-1.14 | 4.69x10 |

\*MAF= minor allele frequency

Statistics: Linear mixed model with Wald test in GEMMA

|  |  |  |  |
| --- | --- | --- | --- |
|  | Patients with diabetes | | P-value |
|  | GADA positive | GADA negative |  |
| No of patients | 131 | 105 |  |
| Age, yr, mean(SD) | 20.4(6.1) | 23.6(6.7) | <0.001 |
| Male gender, n(%) | 89(67.9) | 76(72.4) | ns |
| Plasma glucose at diagnosis, mmol/l, median (IQR) | 29.8(24.6-33.3) | 26.7(20.9-33.3) | 0.01 |
| Insulin dose after stabilisation, units/kg, mean (SD) | 0.81(0.27) | 0.72(0.22) | 0.012 |
| Diabetes duration, months, median (IQR) | 3(1-7) | 2(1-6) | ns |
| C-peptide, µg/L, median (IQR) | 0.71(0.31-1.26) | 0.96(0.39-1.76) | 0.015 |
| Height, SD zscore, mean(SD) | -1.17(0.98) | -1.23(0.83) | ns |
| BMI, SD zscore, mean(SD) | -1.25(1.07) | -1.25(1.23) | ns |
| BMI(kg/m2)$ | 19.3(2.3) | 19.3(2.9) | ns |
| % body fat, mean(SD) | 10.8(7.7) | 12.7(7.9) | 0.065 |
| Rural Birth, n (%) | 121(92.4) | 90(85.7) | ns |
| Subsistence farming/labouring, n (%) | 68(52.0) | 43(41.0) | 0.03 |
| Employed/business owners, n(%) | 15(11.5) | 26(24.8) | 0.007 |

Supplementary Table 4: Comparison of the metabolic, anthropometric and phenotypic characteristics of the patients with diabetes according to the presence of absence of GADA antibody.

$Adults >= 18 yrs

Principal Component Analysis

Figure 1a.

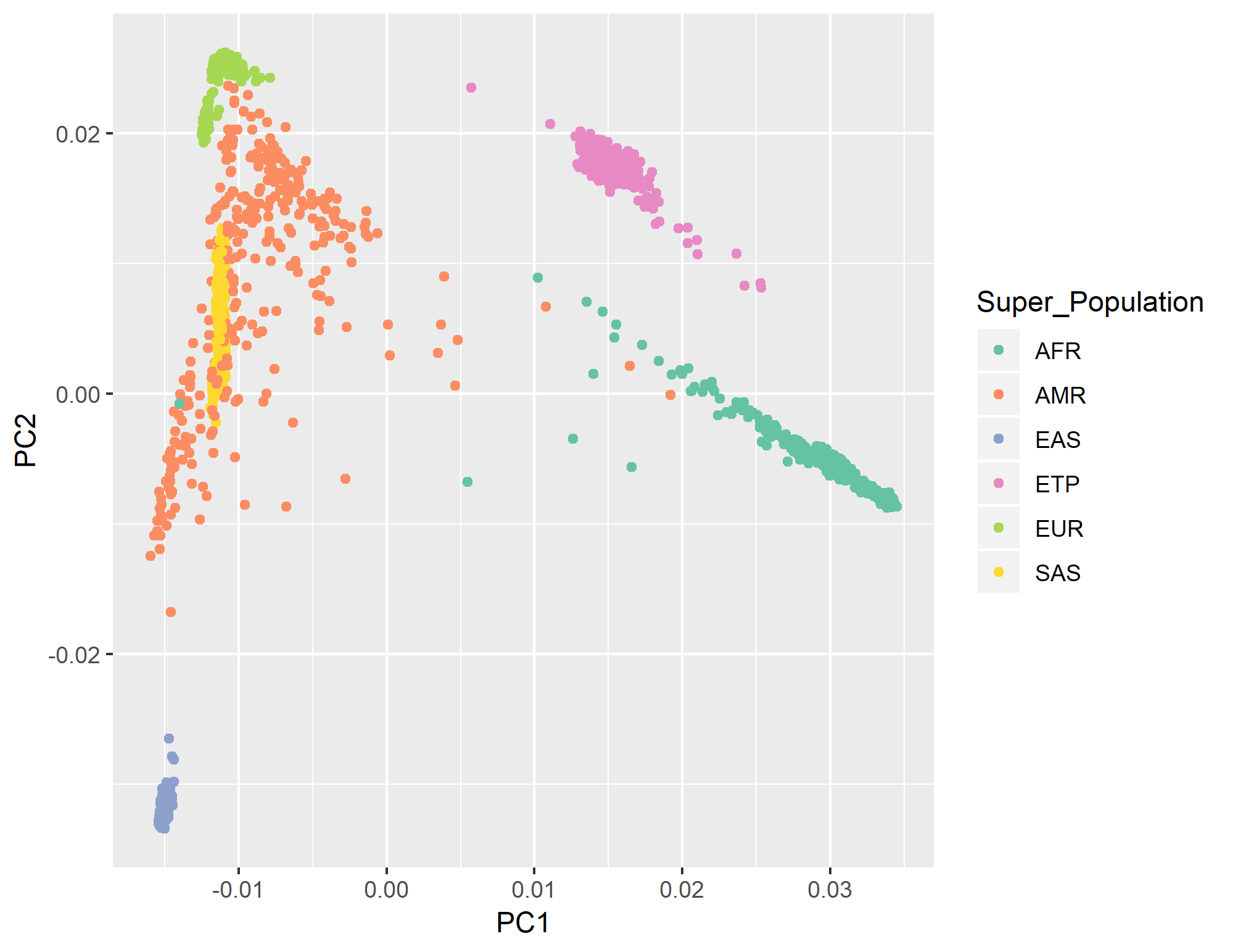
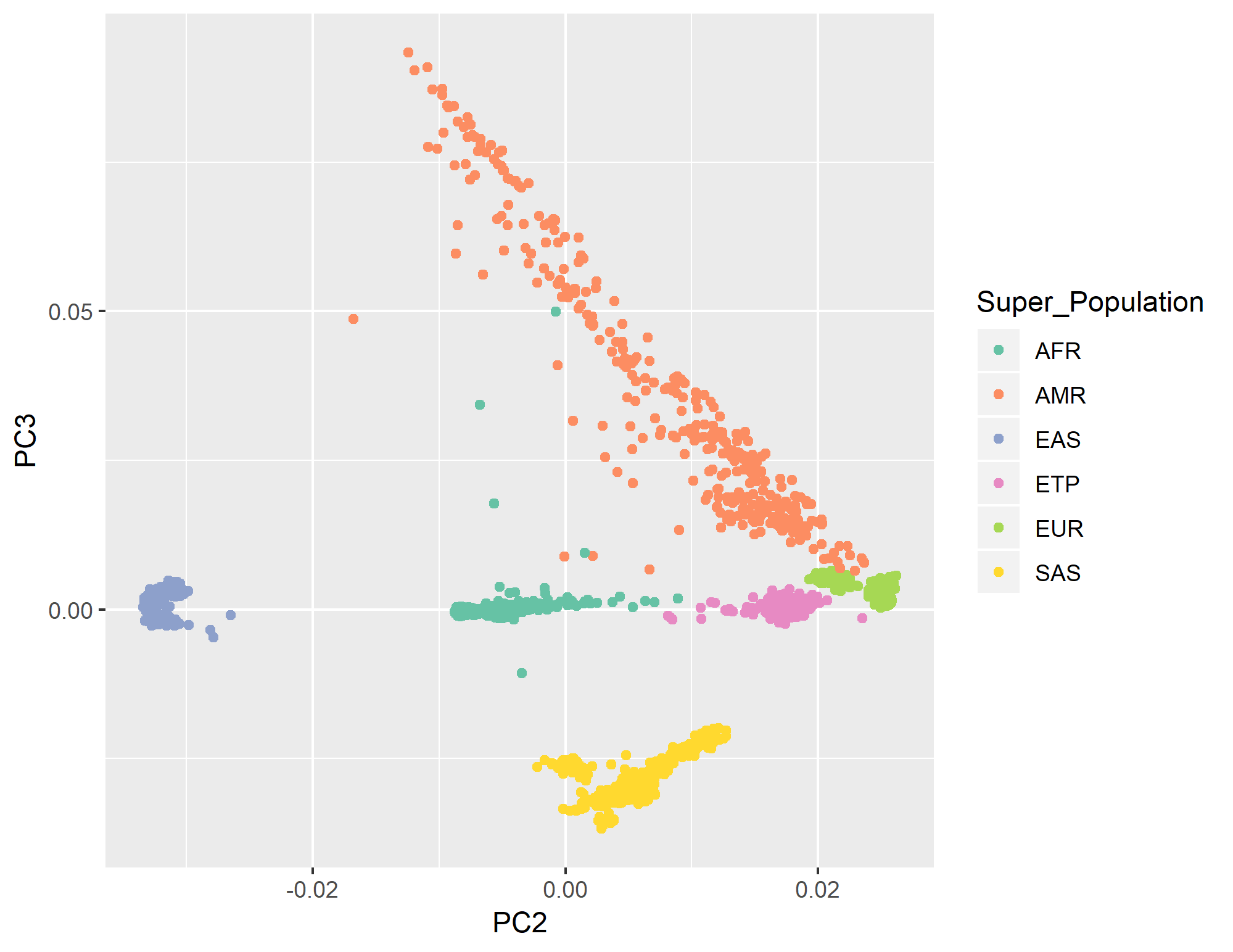


Figure 1b.

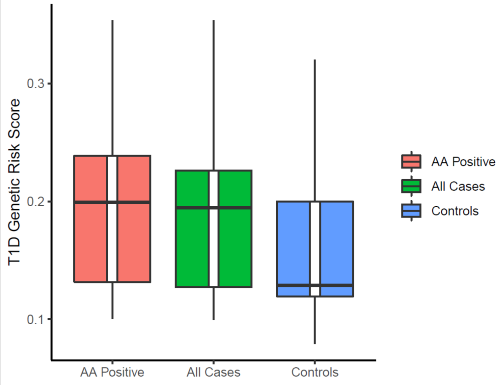


**Figure 1 Legend:** Plots (a) the first two principal components (PCs) and (b) the second and third PCs based on genome-wide

genotypes of the Ethiopian (ETP, Amhara) subjects compared to 1000 Genomes ancestral groups. AFR, African.

AMR, Americas. EAS, East Asian. ETP, Ethiopian. EUR, European. SAS, Southeast Asian.

Figure 2 Genetic Risk Score (GRS) for Type 1 diabetes

****

Legend Fig 2. Diabetes patients Autoantibody (AA) positive (n=121), All diabetes cases (n=187), and Controls(n=137)

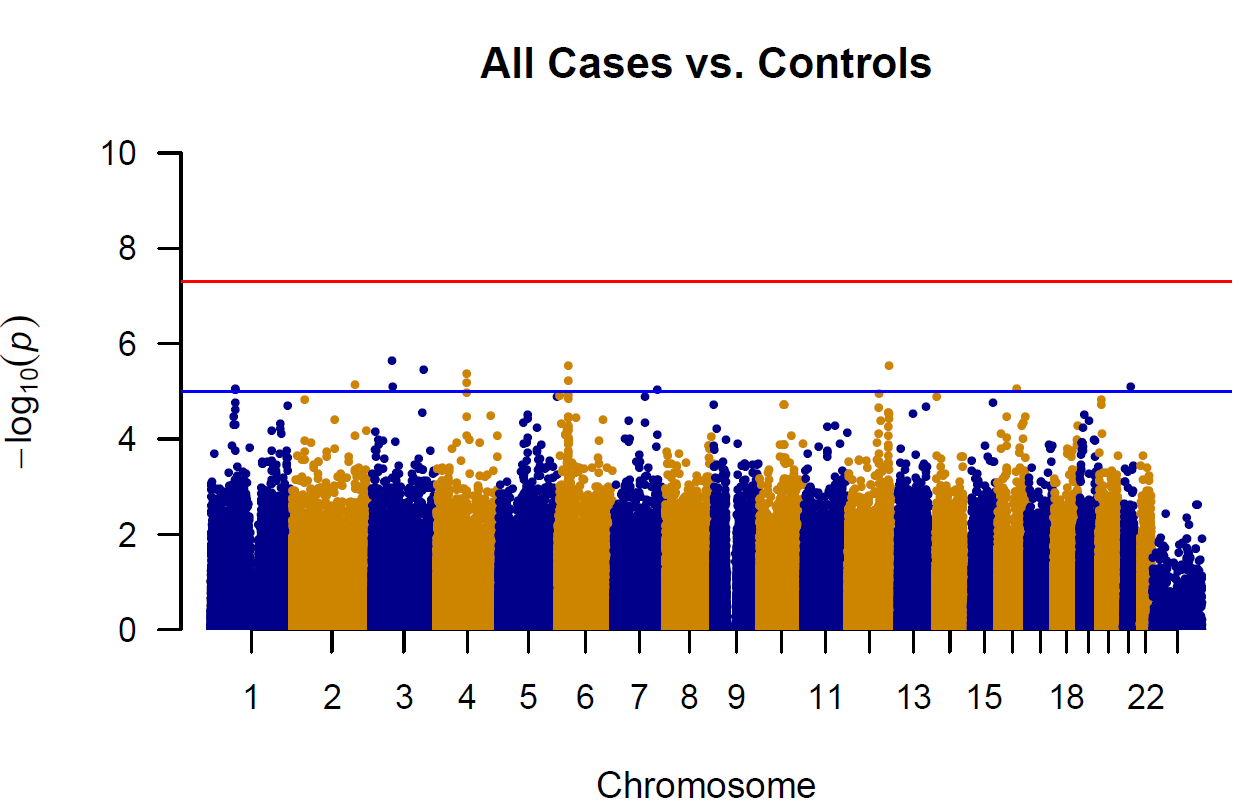
Nineteen established Type 1 diabetes - associated SNPs (Oram et al, ref 25) were used to calculate GRS;

variants, risk allele, and weights are listed in Suppl Table 1.

AA vs Controls, *P*= 1.54x10-9, All diabetic cases vs Controls, *P*=1.6x10-7

(linear regression of GRS vs case/control status, adjusting for sex and first 4 principal components)

Supplementary Figure 1. Manhattan Plot of GWAS results of all cases with diabetes vs controls

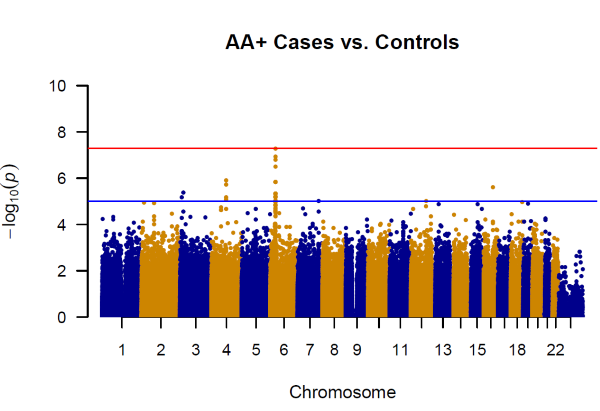


Legend: Supplementary Figure 1. Red line is genome-wide significance *P*-value threshold (*P* < 5x10-8) and

Blue line is line for suggestive significance (P<1x 10- 5)

**Supplementary Figure 2.**

Manhattan Plot of GWAS results for autoantibody positive (AA+) cases with diabetes vs controls



Legend : Supplementary Figure 2 Red line is genome-wide significance *P*-value threshold (*P* < 5x10-8) and

Blue line is line for suggestive significance (P<1x 10- 5)

Supplementary Figure 3. QQ Plot of GWAS for Autoantibody positive cases (AA+) with diabetes vs controls



Legend: Supplementary Figure 3. The QQ plot shows there is no genomic inflation, Lambda =1.01