Electronic supplementary material

Methods

Studies and genotype quality control

Whitehall II Study (WHII) This study had 371 cases of type 2 diabetes and 5235 participants in total. Between 1985 and 1988, all civil servants aged between 35 and 55 years in 20 departments in London were invited to a medical examination at their workplace [1]. With 73% participation, the cohort included 10,308 participants at entry to the study. At phase 3 (1991–1993), all participants known to be alive and still in the UK were invited to the screening clinic for examinations including a 75 g OGTT; 6,058 men and 2,758 women (85.5% of the original sample) attended. For the analysis presented here phase 3 served as the baseline. Additional questionnaire-only phases assessed diabetes status at phase 4 (1995–1996) and phase 6 (2001). Of the baseline participants, 6,156 participated in phase 7 screening, at which DNA was collected. Diabetes status was determined on the basis of the OGTT (phases 3, 5 and 7), use of glucose-lowering drugs or self-report of a doctor's diagnosis (phases 3 to 7). Diabetes was defined according to WHO criteria by a 2 h glucose value of at least 11.1 mmol/l, fasting glucose of at least 7.0 mmol/l, self-report of doctor's diagnosis or use of glucose-lowering medication [2]. Only prevalent cases were considered. To minimise population stratification, we limited our analysis of gene association with traits to European white participants only (n=5,666).

The initial genotype calls were generated using the cluster file provided by Beadstudio software (Illumina). Call frequencies were below 98% for 1,798 of the 49,094 SNPs. These were reviewed by eye and re-clustered manually. SNPs that remained below 98% after re-clustering were discarded. A total of 113 duplicate pairs had a concordance rate >99.5%, one pair had a concordance rate of 99.1% and one pair had a concordance rate of 95%. Manual re-clustering of some SNPs resulted in duplicate concordance rates of 100%. Removal of 115 duplicates and one sample of ambiguous identity left 5,441 samples, of which 5,067 were whites. Eight outliers from the principal component analysis were removed on the basis of the genome-wide identity-by-state analysis implemented in PLINK [3], leaving 5,059 samples for further analysis.

British Women's Health and Heart Study (BWHHS) This study had 338 cases of type 2 diabetes and 3,410 participants overall. Between 1999 and 2001, 4,286 women aged 60 to 79 years were randomly selected from 23 British towns, and interviewed, examined and asked to complete medical questionnaires. Methods used at baseline assessment have been previously described [4]. Glucose, insulin and lipid levels were measured at baseline after a minimum 8 h fast. Women with a clinical diagnosis of diabetes and those with a fasting glucose concentration of $\geq$7 mmol/l were diagnosed as having type 2 diabetes. All cases of type 2 diabetes included in the analysis were prevalent.

Genotyping was successfully performed on 3,445 of 3,838 available samples using the Human CVD Beadchip (Illumina). Human CVD data were validated by 100% concordance by comparison to previous genotyping data for 59 SNPs to confirm sample identities. The following absolute cut-offs for Illumina BeadStudio variables were applied to the data: cluster separation <0.3; call frequency <0.95; AB R mean <0.3; AB T mean <0.2 or >0.8; and heterozygote excess <−0.3 or >0.1. SNPs at the borderline of these cut-offs were then manually checked and either re-clustered or discarded as necessary. Principal components analysis was used to confirm self-reported ancestry and 32 individuals were excluded as non-European, leaving 3,413 samples for analysis.

English Longitudinal Study of Aging (ELSA) There were 439 cases of type 2 diabetes and 5,530 participants in total in this study. The ELSA sample was drawn from participants in the Health Survey for England in 1998, 1999 or 2001. Participants were born before March 1952 and selected to be representative of people living in private households in England [5]. Fasting blood was collected in 2004 for genetic studies, and insulin and glucose measures. Diabetes was diagnosed from self-report or by a doctor through participation in the Health Survey for England, on the basis of a fasting glucose value of $\geq$7 mmol/l. For the analysis prevalent cases were considered.

Northwick Park Heart Study II (NPHSII) There were 159 cases of type 2 diabetes and a total of 2,695 participants in this study. Between 1989 and 1994, 3,012 healthy European white men aged 50 to 64 years and registered with nine primary care practices in the UK were recruited for prospective surveillance [6]. Eligible participants were free of CHD. Blood samples were collected for plasma and DNA analysis. Self-report by questionnaire identified

69 diabetic participants at baseline (participants requiring insulin or oral hypoglycaemic drugs), who were excluded in the current analyses. New cases were identified up to the end of 2005 by searching general practice notes for physician-diagnosed and -treated type 2 diabetes, according to current national guidelines. For this study, both prevalent and incident cases were analysed.

Details of SNPs used in the analysis

A total of 19 *ADRA2A* SNPs were included in the chip, namely rs553668, rs638019, rs11195417, rs17128356, rs35468157, rs7096359, rs491589, rs35844011, rs11195418, rs521674, rs36022820, rs33931318, rs11195419, rs13306146, rs36069680, rs602618, rs35607325, rs35941526 and rs11815669. Full details of call rates, concordance, locus position and minor allele frequencies in WHII are given in ESM Table 2. rs10885122 had been genotyped previously in WHII and BWHHS as part of the MAGIC study [7]. In ELSA and NPHSII, rs553688 and rs10885122 were genotyped by KASPar (Kbioscience, Hoddesdon, UK) and TaqMan (Applied Biosystems, Carlsbad, CA, USA) assays, respectively (with call rates >96% and >99%, respectively).

Adjustment for lipid-lowering medication

We adjusted for lipid-lowering medications on the basis of the effects observed for longitudinal data in WHII (F. Drenos, unpublished data). For triacylglycerol this adjustment involved multiplying by 1.21 before log transformation.

Haplotype analysis

Haplotype analysis was performed using a maximum likelihood model based on the stochastic-EM algorithm implemented in the THESIAS program [8]. THESIAS allows the simultaneous estimation of haplotype frequencies and their associated effects on the phenotype of interest. A global *p* value was calculated using differences in log-likelihood assuming an additive model of haplotype effects. In addition each haplotype was compared with the common haplotype. Due to low frequency for some of the inferred haplotypes, we ran a permutation analysis to check the reliability of the results. The haplotypes were inferred using PHASE [9]. A totel of 20 haplotype reconstruction datasets were randomly drawn from

the posterior distributions of the PHASE results. For each reconstructed dataset, we tested the association between the haplotypes and the phenotype of interest through 1,000 permutations of the data in order to deal with the problem of low frequency for some of the inferred haplotypes. The results for all 20 alternative haplotypic reconstructions were finally combined using Rubin's rule [10].

References

1. Marmot MG, Smith GD, Stansfeld S, et al (1991) Health inequalities among British civil servants: the Whitehall II study. Lancet 337:1387–1393

2. 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

3. Purcell S, Neale B, Todd-Brown K et al (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. Am J Hum Genet 81:559–575

4. Lawlor DA, Bedford C, Taylor M, Ebrahim S (2003) Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. J Epidemiol Community Health 57:134–140

5. Netuveli G, Wiggins RD, Hildon Z, Montgomery SM, Blane D (2006) Quality of life at older ages: evidence from the English longitudinal study of aging (wave 1). J Epidemiol Community Health 60:357–363

6. Miller GJ, Bauer KA, Barzegar S, et al (1995) The effects of quality and timing of venepuncture on markers of blood coagulation in healthy middle-aged men. Thromb Haemost 73:82–86

7. Dupuis J, Langenberg C, Prokopenko I, et al (2010) New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet 42:105–116

8. Tregouet DA, Tiret L (2004) Cox proportional hazards survival regression in haplotype-based association analysis using the Stochastic-EM algorithm. Eur J Hum Genet 12:971–974

9. Scheet P, Stephens M (2006) A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. Am J Hum Genet 78:629–-644

10. Rubin DB (1987) Multiple imputation for nonresponse in surveys. Wiley, New York