Human Molecular Genetics, 2010, Vol. 19, No. 9 ***1856–1862***
doi:10.1093/hmg/ddq057
Advance Access published on February 10, 2010

Genetic variants in *ABO* blood group region, plasma soluble E-selectin levels and risk of type 2 diabetes

Lu Qi[1,3,*], Marilyn C. Cornelis[1], Peter Kraft[3], Majken Jensen[1], Rob M. van Dam[1,2,3], Qi Sun[1], Cynthia J. Girman[4], Cathy C. Laurie[5], Daniel B. Mirel[6], David J. Hunter[2,3], Eric Rimm[1,2,3] and Frank B. Hu[1,2,3]

[1]Department of Nutrition and [2]Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA, [3]Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA, [4]Department of Epidemiology, Merck Research Labs, 305 S Sumneytown Pike, North Wales, PA 19454, USA, [5]Department of Biostatistics, University of Washington, Seattle, WA 98195, USA and [6]Center for Genotyping and Analysis, Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, MA 02142, USA

Received November 20, 2009; Revised January 11, 2010; Accepted February 5, 2010

Blood soluble E-selectin (sE-selectin) levels have been related to various conditions such as type 2 diabetes. We performed a genome-wide association study among women of European ancestry from the Nurses' Health Study, and identified genome-wide significant associations between a cluster of markers at the *ABO* locus (9q34) and plasma sE-selectin concentration. The strongest association was with rs651007, which explained ~9.71% of the variation in sE-selectin concentrations. SNP rs651007 was also nominally associated with soluble intracellular cell adhesion molecule-1 (sICAM-1) ($P = 0.026$) and TNF-R2 levels ($P = 0.018$), independent of sE-selectin. In addition, the genetic-inferred *ABO* blood group genotypes were associated with sE-selectin concentrations ($P = 3.55 \times 10^{-47}$). Moreover, we found that the genetic-inferred blood group B was associated with a decreased risk (OR = 0.44, 0.27–0.70) of type 2 diabetes compared with blood group O, adjusting for sE-selectin, sICAM-1, TNF-R2 and other covariates. Our findings indicate that the genetic variants at *ABO* locus affect plasma sE-selectin levels and diabetes risk. The genetic associations with diabetes risk were independent of sE-selectin levels.

INTRODUCTION

E-selectin is specifically synthesized by endothelial cells and plays an important role in mediating leukocyte-endothelial adhesion. A circulating form of E-selectin (soluble E-selectin or sE-selectin) may be released by enzymatic cleavage or result from shedding of damaged or activated endothelial cells. Plasma concentration of sE-selectin correlates with its expression on the surface of endothelial cells (1), and therefore, is a marker of endothelial dysfunction (2).

Elevated sE-selectin levels have been related to a plethora of metabolic disorders including type 2 diabetes (3–6), coronary heart disease (CHD) (7,8), hypertension (9) and Graves' disease (10). A recent genome-wide association study (GWAS) found that genetic variants at *ABO* locus significantly accounted for up to 19% of the variance in E-selectin levels in patients with type 1 diabetes and in non-diabetic subjects. The genetic-inferred ABO blood groups were also associated with sE-selectin levels (11).

The aim of the present study is to conduct a GWAS to identify common genetic variants associated with serum sE-selectin concentration. We particularly examined the associations between sE-selectin-associated genetic variants and the risk of type 2 diabetes.

*To whom correspondence should be addressed at: Department of Nutrition, Harvard School of Public Health, 665 Huntington Ave., Boston, MA 02115, USA. Tel: +1 6174324116; Fax: +1 6174322435; Email: nhlqi@channing.harvard.edu or luqi@hsph.harvard.edu

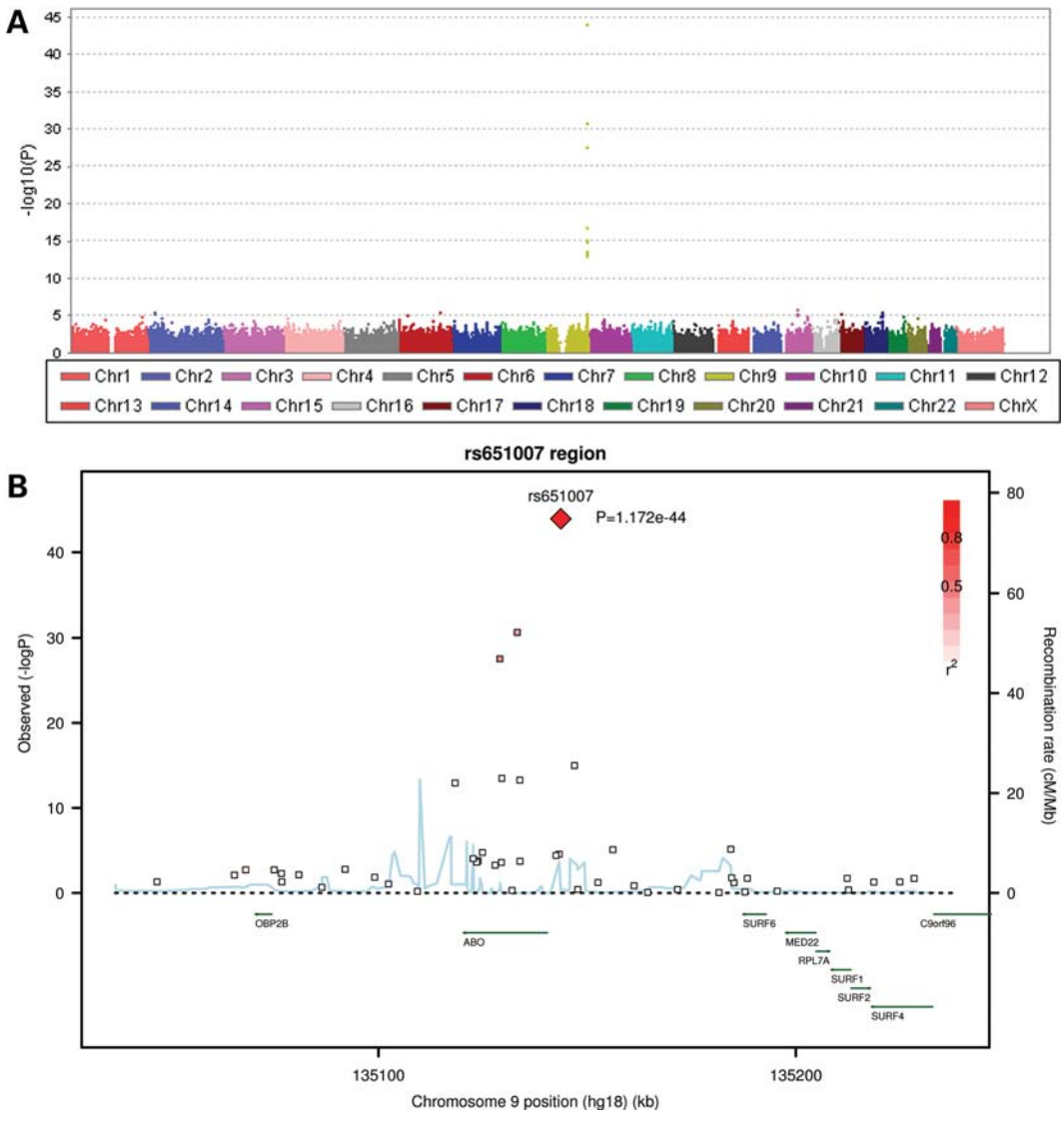

Figure 1. (**A**) Plot of the $-\log_{10}$ *P*-values for the analysis for sE-selectin level GWA trend tests. The analysis was adjusted for age, BMI and diabetes status. Each point represents a SNP from the 704 409 SNPs remaining after quality control filters. Different bands are used to differentiate SNPs on consecutive autosomal and X chromosomes; (**B**) association signals at chromosome 9, across a region centering on *ABO* gene. The vertical axis representing the $-\log_{10}$ *P*-values from the linear trend test. SNP rs651007 is shown in red, labeled with discovery stage *P*-values. Estimated recombination rates from HapMap are plotted to reflect the local LD structure. Genes were extracted from the HapMap Genome Browser.

RESULTS

We performed a GWA analysis in 1005 women (including 501 diabetic patients) from the Nurses' Health Study (NHS) in whom sE-selectin was measured. The characteristics of the participants are presented in Supplementary Material, Table S1. The participants were 43–69 years of age at the time of blood collection. The mean (SD) of sE-selectin concentration in the study sample was 58.2 (31.5) ng/ml. We fit a linear regression model for genotype trend effects (1 degree of freedom), adjusting for age, body mass index (BMI), fasting status and diabetes status. The quantile–quantile plot (Supplementary Material, Fig. S1) of the association test *P*-values suggest there was no systemic bias due to genotyping errors or population stratification (genomic inflation factor, $\lambda = 0.999$).

Figure 1A shows the plot of the $-\log_{10}$ *P*-values for the trend test in linear regression models. Seven SNPs passed the genome-wide significance threshold (5×10^{-8}). All of these SNPs cluster within a ~30 kb region on chromosome 9q34 (Fig. 1B and Table 1), which harbors the *ABO* gene encoding proteins related to the ABO blood group system. The correlations between these SNPs are presented in Supplementary Material, Table S2. The most strongly associated SNP at the *ABO* locus was rs651007 ($P = 1.17 \times 10^{-44}$), which explained ~9.71% of the residual variance in sE-selectin levels. To evaluate the stability of the associations, we performed sensitivity analyses in diabetic and non-diabetic women separately. The results were highly consistent in the sub-sample sets, although *P*-values were somewhat attenuated due to decreased statistical power (Supplementary Material, Table S3). We tested the interactions

Table 1. Genome-wide significant SNPs in ABO blood group region (9q34) for sE-selectin concentrations (ng/ml)

SNP	Position	Function	Alleles		MAF	HWE	Discovery ($n = 1005$)		Replication ($n = 518$)		Combined
			Minor	Major			Beta[a]	*P*-value	Beta[a]	*P*-value	*P*-value
rs651007	135143696	Upstream	T	C	0.22	0.37	−17.23	1.17E−44	−15.78	4.30E−41	2.37E−82
rs612169	135133263	Intronic	G	A	0.34	0.79	−12.58	2.28E−31	−11.76	2.81E−29	1.02E−57
rs657152	135129086	Intronic	A	C	0.38	0.76	−11.97	2.93E−28	−10.39	7.92E−23	7.83E−49
rs558240	135146954	Intergenic	A	G	0.39	0.16	−7.49	1.02E−15	−8.40	2.33E−17	2.04E−30
rs8176681	135129575	Intronic	C	T	0.42	0.18	7.65	3.24E−14	5.03	8.90E−08	1.65E−20
rs8176668	135133880	Intronic	T	A	0.4	0.16	7.89	5.31E−14	–	–	–
rs7857390	135118367	Downstream	A	G	0.4	0.39	7.68	1.14E−13	4.44	3.38E−06	2.33E−18

[a]Per-minor allele change in mean sE-selectin concentrations.

Table 2. sE-selectin concentrations (ng/ml) by ABO blood group genotypes

ABO genotypes	Frequency (%)	Means	SD
A1/A1	3.98	40.2*	8.5
A1/A2	2.09	42.7*	11.7
A1/B	1.69	54*	13.0
A1/O	27.09	44.7*	3.3
A2/A2	0.3	48.3	30.9
A2/B	1.39	57.6	14.3
A2/O	8.37	65.2	5.8
B/B	0.5	45.8*	24.0
B/O	10.56	59.6*	5.2
O/O	44.02	67.4	2.5

$P = 3.55\text{E}-47$ for testing the difference across ABO genotypes.
*$P < 0.05$ when compared with O/O genotype.

Table 3. Association of genetic-inferred and self-reported blood ABO groups with levels of sE-selectin, sICAM-1 and TNF-R2

	Blood groups	Genetic-inferred			Self-reported		
		Means	SD	*P*-value	Means	SD	*P*-value
sE-selectin (ng/ml)	O	66.2	2.2	–	65.9	2.4	–
	A	50.1	2.3	5.27E−37	51	2.6	2.43E−21
	B	57.6	4.5	0.003	59	4.5	0.18
	AB	53.3	8.4	0.0002	51.1	6.0	1.00E−06
sICAM-1 (ng/ml)	O	261.8	7.9	–	256.2	8.0	–
	A	282.6	8.2	0.006	281	8.8	0.03
	B	274.3	15.3	0.18	272.9	14.4	0.42
	AB	289.5	28.6	0.04	290.6	19.3	0.02
TNF-R2 (pg/ml)	O	2560	78	–	2569	82	–
	A	2732	80	0.04	2613	90	0.26
	B	2577	153	0.64	2617	149	0.71
	AB	2657	284	0.85	2754	200	0.88

Analyses are adjusted for age, BMI, fasting status, diabetes and biomarkers (sICAM-1 and TNF-R2 for sE-selectin; and sE-selectin for sICAM-1 and TNF-R2).
sE-selectin, soluble E-selectin; sICAM-1, soluble intracellular cell adhesion molecule-1; TNF-R2, tumor necrosis factor-alpha receptor 2; SD, standard deviation.

with diabetes status and did not find significant interactions ($P > 0.05$).

We conducted an *in silico* replication for the SNPs reaching the genome-wide significance in an independent GWA scan of a second set of 518 women from the NHS (Supplementary Material, Table S1 and Table 1). All the tested SNPs showed directionally consistent associations with sE-selectin levels, adjusting for age, BMI, fasting status and CHD status. In the combined analysis, the strongest association (rs651007) with sE-selectin concentration had a *P*-value of 2.37×10^{-82}.

The ABO histo-blood group antigen is the most important blood group system in transfusion medicine. In our genotyping set, we identified perfect surrogate SNPs tagging *ABO* alleles: rs8176704 for the A2 allele, rs612169 for the O allele and rs8176672 for the B allele. There was significant difference in sE-selectin level among the genetic-inferred ABO genotypes (Table 2; $P = 3.55 \times 10^{-47}$), similar to a previous report (11). Individuals with one or more A1 alleles had significantly lower sE-selectin levels than individuals with O/O genotype, which was associated with the highest sE-selectin concentration. B/O and B/B genotypes were also associated with lower sE-selectin than O/O genotype. Genetic-inferred ABO blood groups were consistent with self-reported ABO blood groups in the NHS (concordance = 0.8). Table 3 presents the associations between genetic-inferred and self-reported ABO blood groups and sE-selectin levels. The genetic-inferred blood groups A ($P=5.27 \times 10^{-37}$), B ($P = 0.003$), and AB ($P = 0.0002$) were associated with significantly lower sE-selectin concentrations than blood group O. Similar associations were observed in the self-reported ABO blood groups in the same samples.

ABO histo-blood group antigens are known to vary in frequency among Caucasian sub-populations (12). However, further adjustment for the top three eigenvectors determined by the principal component analysis did not alter the associations. SNP rs651007 is strongly correlated with the *ABO* A1 allele ($r^2 = 0.76$). The statistical significance for the association between sE-selectin levels and rs651007 was dramatically reduced after adjusting for *ABO* genotypes ($P = 0.005$).

SNPs at the *ABO* locus were previously associated with soluble intracellular cell adhesion molecule-1 (sICAM-1) (12) and tumor necrosis factor alpha (TNF-alpha) concentrations (13). In the GWA samples, we measured sICAM-1 and tumor necrosis factor-alpha receptor 2 (TNF-R2), which reflects TNF system activation (14). The correlations (r^2) of sICAM-1 and TNF-R2 with sE-selectin were 0.6 and 0.3, respectively. Adjustment for sICAM-1 and TNF-R2 did not appreciably change the associations of rs651007 (Fig. 2) and ABO blood groups with sE-selectin level (Table 3). After adjustment of sE-selectin and other covariates, SNP

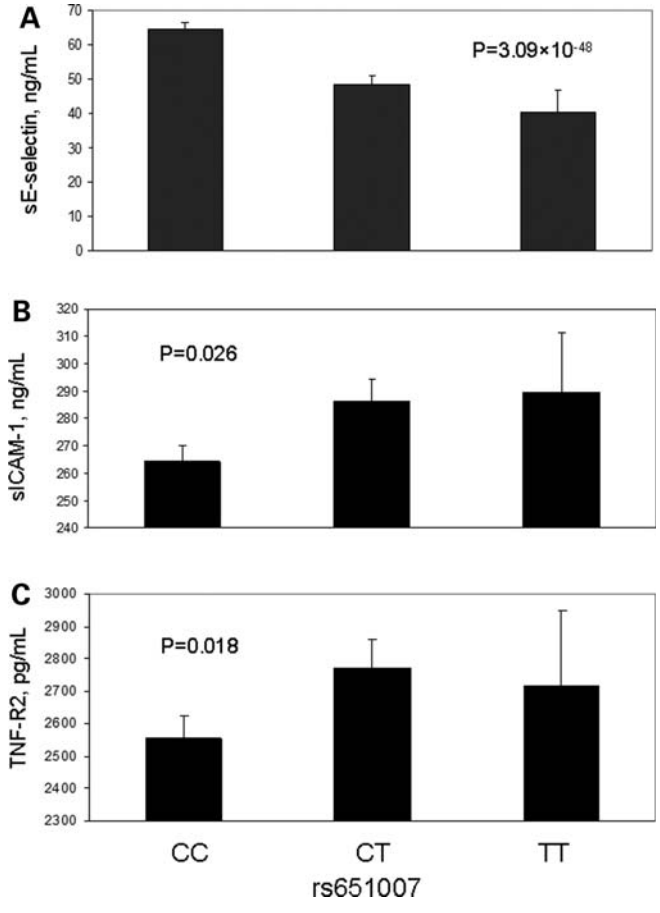

Figure 2. The levels of (**A**) sE-selectin, (**B**) sICAM-1 and (**C**) TNF-R2 by the genotypes of SNP rs651007. The *P*-values are from the trend test in linear regression model. The analyses were adjusted for age, BMI, fasting status, diabetes status and biomarkers (sICAM-1 and TNF-R2 for sE-selectin analysis; and sE-selectin for sICAM-1 and TNF-R2 analyses). Error bars denote standard deviation.

rs651007 was nominally associated with higher sICAM-1 ($P = 0.026$) and TNF-R2 levels ($P = 0.018$) (Fig. 2); and the genetic-inferred blood group A was also significantly associated with higher sICAM-1 ($P = 0.006$) and TNF-R2 levels ($P = 0.04$) than blood group O (Table 3).

Because sE-selectin levels have been associated with the risk of type 2 diabetes (3,4), we therefore further examined the associations of ABO variants and blood groups with diabetes risk. None of the ABO SNPs was significantly associated with diabetes risk. When compared with the genetic-inferred blood group O, group B was associated with a decreased risk (OR = 0.44; 0.27–0.70; $P = 0.0006$) of type 2 diabetes; while blood groups A and AB were not significantly associated with diabetes risk (Fig. 3). The association for blood group B remained significant after correcting for the number to comparisons ($n = 3$), with $P = 0.0018$. Further adjustment for the top three eigenvectors determined by the principal component analysis did not appreciably alter these associations. Similarly, the self-reported blood group B was associated with a decreased diabetes risk compared with blood group O (OR = 0.57; 0.36–0.90; $P = 0.015$); while blood groups A and AB were not significantly associated with diabetes risk.

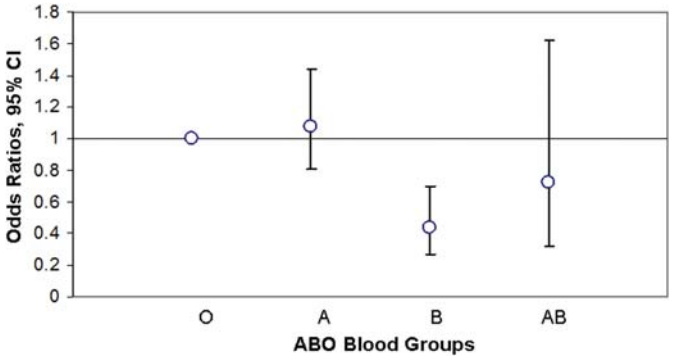

Figure 3. Odds ratios of diabetes associated with the genetic-inferred ABO blood groups A, B and AB compared with blood group O. The analyses were adjusted for age, BMI, smoking, alcohol consumption, physical activity, family history of diabetes and menopausal status. Error bars denote 95 percent confidence intervals.

We also examined the associations of the sub-types of blood group A (A1/O and A2/O) and did not find significant difference in their relations with diabetes risk. The odds ratio of diabetes comparing blood genotype A2/O with A1/O was 0.88 (0.52–1.48). In addition, adjustment for ABO variants and blood groups did not appreciably change the associations between sE-selectin and diabetes risk (data not shown).

DISCUSSION

In the genome-wide analyses, we confirmed that common variations near/within *ABO* gene and genetic-inferred ABO blood groups were associated with plasma concentrations of sE-selectin at a genome-wide significance level.

The human *ABO* gene encodes a glycosyltransferase that catalyzes the transfer of carbohydrates to H antigen and form the antigenic structure of the ABO blood groups (15). A and B antigens are formed by the action of glycosyltransferases encoded by functional alleles of the *ABO* gene. The A allele encodes A transferase, which synthesizes the A antigen. Similarly, the B allele encodes B transferase, which synthesizes the B antigen. The O allele does not produce an active enzyme. Blood type A has variations in subgroups, of which A1 and A2 are the most important. Blood type A with a normal quantity of antigen is named A1, which comprises ~80% of blood type A in Europeans. Blood type A2 has a single base deletion near the carboxyl terminal, resulting in a loss of A2 transferase activity (16). The A1 allele has 30–50-fold higher A transferase activity than A2 allele (17).

Our results are highly consistent with the findings from a recent GWA analysis in type 1 diabetic patients and non-diabetic controls (11), and also in line with previous report that human ABO blood groups were related to sE-selectin levels (18). Although the mechanism underlying the associations between genetic variants at the *ABO* locus and sE-selectin concentration is unknown, the high consistency in data from different studies provides solid evidence for *ABO* locus as a major genetic determinant for plasma sE-selectin concentration.

In addition, we found the variations at *ABO* locus and ABO blood groups were associated with sICAM-1 and/or TNF-R2

levels, independent of sE-selectin levels. These results are consistent with previous GWA studies in which SNPs in the same region were related to levels of sICAM-1 (12) and TNF-alpha (13). E-selectin is transcriptionally regulated by TNF-alpha (19), and both sICAM-1 and/or TNF-R2 are positively correlated with sE-selectin levels. However, the associations of ABO variants with sICAM-1 and TNF-R2 were to the opposite direction to their associations with sE-selectin. The data suggest that *ABO* variants may affect these markers through different mechanisms. Further investigations are warranted to understand the functional alterations contributing to the changes in these correlated markers.

Several previous studies have examined the relations between human ABO blood groups with the risk of type 2 diabetes. Although some studies suggested a link between ABO blood groups and diabetes (20–22), the associations were not observed in others (23,24). The controversial results from these studies were partly due to their retrospective design and small case numbers, and the variation in the genetic structure among different ethnic groups. We found that blood group B was associated with a decreased risk compared with blood group O. Although associated with even lower sE-selectin, the association between blood group A and diabetes risk was not significant, maybe partly explained by its associations with higher levels of sICAM-1 and TNF-R2, both are risk factors for type 2 diabetes (3,14). In addition, data from some studies indicate that *ABO* locus might also affect other biomarkers such as factor VIII and throumbomodulin (18,25). Therefore, the associations between ABO blood groups and diabetes risk may reflect the combined effects of multiple risk factors. We did not find significant associations between individual SNPs at *ABO* locus with diabetes risk. The results were in line with the failure in identifying genetic variants associated with diabetes in GWAS studies (26–30). These data suggest that the combination of multiple alleles (haplotypes) at this locus, rather than individual genetic variant, may affect diabetes risk.

The major strengths of our study include high quality genotype data, careful quality control and minimal population stratification. We acknowledge several study limitations, including errors in biomarker measurements and genotyping. Nevertheless, we employed strict quality control criteria in genotyping, and these errors more likely bias the association toward null because the measurement errors for biomarker assays and genotyping are uncorrelated and, thus, random. The case–control sample may not represent a random sample from the general population. We have controlled the diabetes status in the analyses to avoid the potential sampling bias. In addition, we performed sensitivity analyses in the cases and controls separately. The associations with e-selectin levels from all the analyses were highly consistent. The study sample size is relatively small to identify the associations between blood groups and diabetes risk at genome-wide significance level. For example, given the effect size observed and the frequency for the genetic-inferred blood group B, we estimated a sample size of $\sim$1700 (850 diabetes cases) is needed to reach genome-wide significance. Our study populations exclusively consisted of Caucasian women with European ancestry. Therefore, the findings may not be generalizable to men and other ethnicities.

In conclusion, by examining a GWA scan, we confirmed that the *ABO* locus was a major determinant for plasma sE-selectin levels. We found that the variants at *ABO* locus and the genetic-inferred ABO blood groups were associated with the risk of type 2 diabetes, independent of sE-selectin levels. We found that blood group B was associated with a decreased risk compared with blood group O. The mechanism underlying the observed association remains unknown and our findings warrant the need for further replications in other ethnic groups and functional investigations.

MATERIALS AND METHODS

Study population

Details of the NHS have been described previously (31). Briefly, The NHS was established in 1976 when 121 700 female registered nurses aged 30–55 years and residing in 11 large US states completed a mailed questionnaire on their medical history and lifestyle. The lifestyle factors, including smoking, menopausal status and postmenopausal hormone therapy, and body weight, have been updated by validated questionnaires every 2 years. On the 1996 questionnaire, participants were asked their blood types (A, B, AB, O or unknown) (32). A total of 32 826 women provided blood samples between 1989 and 1990. The present study was approved by the institutional review board at Brigham and Women's Hospital and returning the questionnaires was supposed to imply informed consent. Participants for the current study were a subset of women ($n = 1005$) included in a nested case–control study of sE-selectin in relation to type 2 diabetes in the NHS. The diagnosis of type 2 diabetes has been described in detail else where (3,33).

For replication, we used a case–control study on risk of CHD, also nested within the NHS. The details in selection of CHD cases and controls were previously described (34). We restricted the sample for the present study to the participants with sE-selectin available ($n = 518$).

Assessment of biomarkers

Women providing blood samples were sent a phlebotomy kit, returning the sample by overnight mail in a frozen water bottle. On arrival, samples were processed and frozen in liquid nitrogen until analysis; 97% arrived within 26 h of phlebotomy. Quality control samples were routinely frozen with study samples. Study samples were analyzed in randomly ordered case–control pairs to further reduce systematic bias and interassay variation. Levels of sE-selectin, sICAM-1 and TNF-R2 were measured by commercial enzyme-linked immunosorbent assay (R & D Systems, Minneapolis, MN, USA). The coefficients of variation for analyses were 4.5–6.2% for sE-selectin; 3.3–4.8% for sICAM-1 and 2.6–4.8% for TNF-R2 (3,35).

Genome-wide scan and quality control

For the GWA samples, DNA was extracted from white blood cells using the QIAmpTM (QIAGEN Inc., Chatsworth, CA, USA) blood protocol. Genotyping was done using the

Affymetrix Genome-Wide Human 6.0 array and the Birdseed calling algorithm (36). Genotypic data for a total of 3,429 NHS samples passed laboratory technical quality control criteria which included SNP fingerprints for sample tracking and early detection of sample misidentification, missing call rate (MCR), the use of HapMap controls to check genotype quality independent of study samples and the tracking of reagent and instrumental performance.

Relatedness was evaluated using pairwise identity-by descent using 80K SNPs in a method of moments approach implemented in PLINK software (37). Five pairs of duplicate samples were identified and removed. One pair of full siblings and eight sets (six pairs and two triplets) of possible first cousins were also identified. Gender was confirmed by examining the mean of the intensities of SNP probes on the X and Y chromosomes. One male sample was mis-identified as a female sample and was excluded. Twenty-seven subjects with gross chromosomal anomalies, determined by analyzing relative intensity ('LogRRatio') and allelic imbalance ('BAlleleFreq', BAF) (38) and 22 samples having a MCR $\geq$2% were also removed.

More than 96% (879 071) of the 909 622 SNP probes on the array passed the quality control standards of the genotyping center (Broad Institute of MIT and Harvard) for NHS samples. We further excluded SNPs which were monomorphic, had a MCR $\geq$2%, more than one discordance, a HWE P-value $<1 \times 10^{-4}$ or a MAF $<$0.02. Duplicate SNPs (assayed with different probes) were also removed. A total of 704 409 SNPs passed quality control and were included for the analysis.

The replication samples from NHS were genotyped at Rosetta/Merck using the Affymetrix Genome-Wide Human 6.0 array and the Birdseed calling algorithm (36). A total of 907 370 SNPs were successfully genotyped. The same QC standards used in the discovery GWA scan were applied in the replication set, and 518 participants (160 CHD cases and 358 controls) with sE-selectin measurement passed the QC. The QC on genotyping left a total of 721 316 SNPs for final analysis.

Determination of population structure

Population structure was investigated by principal component analysis (39). We used a set of 12 021 SNPs selected to have very low levels of linkage disequilibrium (LD) and to have MAF $>$0.05 in Caucasians (40). Unrelated genetically inferred European ancestral women passing QC were included in the current study. To control for potential confounding by population stratification, we adjusted for the top three eigenvectors in the analyses.

Statistical analyses

GWA analysis was performed with the linear regression model, using PLINK software (37). Analyses were adjusted for age, BMI, fasting status and diabetes status. Plasma concentrations of sE-selectin, sICAM-1 and TNF-R2 were logarithmically transformed to achieve normal distributions. A quantile–quantile (Q-Q) plot analysis was carried out to determine whether the distribution of the inflation corrected P-values deviated from the expected distribution under the null hypothesis of no genetic association. LD analysis was carried out by use of the Haploview software (41). A conservative P-value cut-off of 5×10^{-8} was used to correct for the roughly 1 000 000 independent statistical tests (42). We combined study-specific β-estimates from GWA and replication analyses using the inverse of the variance of the study-specific β-estimates to weight the contribution of each study, with the fixed effect model. The associations with diabetes risk were analyzed using logistic regression model, adjusting for the covariates. These analyses were performed in SAS 9.1 (SAS Institute, Inc., Cary, NC, USA).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

Conflict of Interest statement. None declared.

FUNDING

The NHS type 2 diabetes GWA study (U01HG004399) is a component of a collaborative project that includes 13 other GWA studies (U01HG004738, U01HG004422, U01HG004 402, U01HG004729, U01HG004726, U01HG004735, U01H G004415, U01HG004436, U01HG004423, U01HG004728, RFAHG006033; National Institute of Dental & Craniofacial Research: U01DE018993, U01DE018903) funded as part of the Gene Environment-Association Studies (GENEVA) under the NIH Genes, Environment and Health Initiative (GEI). Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center (U01HG004446). Assistance with data cleaning was provided by the National Center for Biotechnology Information. Genotyping was performed at the Broad Institute of MIT and Harvard, with funding support from the NIH GEI (U01HG04424), and Johns Hopkins University Center for Inherited Disease Research, with support from the NIH GEI (U01HG004438) and the NIH contract 'High throughput genotyping for studying the genetic contributions to human disease' (HHSN268200782096C). Additional funding for the current research was provided by the National Cancer Institute (P01CA087969, P01CA055075), the National Institute of Diabetes and Digestive and Kidney Diseases (R01DK058845) and Merck & Co., Inc. L.Q. is supported by National Institutes of Health grants RO1 HL71981, American Heart Association Scientist Development Award and the Boston Obesity Nutrition Research Center (DK46200). M.C. Cornelis is a recipient of a Canadian Institutes of Health Research Fellowship.

REFERENCES

1. Leeuwenberg, J.F., Smeets, E.F., Neefjes, J.J., Shaffer, M.A., Cinek, T., Jeunhomme, T.M., Ahern, T.J. and Buurman, W.A. (1992) E-selectin and intercellular adhesion molecule-1 are released by activated human endothelial cells in vitro. *Immunology*, **77**, 543–549.
2. Blann, A. and Seigneur, M. (1997) Soluble markers of endothelial cell function. *Clin. Hemorheol. Microcirc.*, **17**, 3–11.
3. Meigs, J.B., Hu, F.B., Rifai, N. and Manson, J.E. (2004) Biomarkers of endothelial dysfunction and risk of type 2 diabetes mellitus. *JAMA*, **291**, 1978–1986.

4. Song, Y., Manson, J.E., Tinker, L., Rifai, N., Cook, N.R., Hu, F.B., Hotamisligil, G.S., Ridker, P.M., Rodriguez, B.L., Margolis, K.L. *et al.* (2007) Circulating levels of endothelial adhesion molecules and risk of diabetes in an ethnically diverse cohort of women. *Diabetes*, **56**, 1898–1904.
5. Stranges, S., Rafalson, L.B., Dmochowski, J., Rejman, K., Tracy, R.P., Trevisan, M. and Donahue, R.P. (2008) Additional contribution of emerging risk factors to the prediction of the risk of type 2 diabetes: evidence from the Western New York Study. *Obesity (Silver Spring)*, **16**, 1370–1376.
6. Thorand, B., Baumert, J., Chambless, L., Meisinger, C., Kolb, H., Doring, A., Lowel, H. and Koenig, W. (2006) Elevated markers of endothelial dysfunction predict type 2 diabetes mellitus in middle-aged men and women from the general population. *Arterioscler. Thromb. Vasc. Biol.*, **26**, 398–405.
7. Malik, I., Danesh, J., Whincup, P., Bhatia, V., Papacosta, O., Walker, M., Lennon, L., Thomson, A. and Haskard, D. (2001) Soluble adhesion molecules and prediction of coronary heart disease: a prospective study and meta-analysis. *Lancet*, **358**, 971–976.
8. Hwang, S.J., Ballantyne, C.M., Sharrett, A.R., Smith, L.C., Davis, C.E., Gotto, A.M. Jr and Boerwinkle, E. (1997) Circulating adhesion molecules VCAM-1, ICAM-1, and E-selectin in carotid atherosclerosis and incident coronary heart disease cases: the Atherosclerosis Risk In Communities (ARIC) study. *Circulation*, **96**, 4219–4225.
9. Nadar, S., Blann, A.D. and Lip, G.Y. (2004) Endothelial dysfunction: methods of assessment and application to hypertension. *Curr. Pharm. Des.*, **10**, 3591–3605.
10. Hara, H., Sugita, E., Sato, R. and Ban, Y. (1996) Plasma selectin levels in patients with Graves' disease. *Endocr. J.*, **43**, 709–713.
11. Paterson, A.D., Lopes-Virella, M.F., Waggott, D., Boright, A.P., Hosseini, S.M., Carter, R.E., Shen, E., Mirea, L., Bharaj, B., Sun, L. *et al.* (2009) Genome-wide association identifies the ABO blood group as a major locus associated with serum levels of soluble E-selectin. *Arterioscler. Thromb. Vasc. Biol.*, **29**, 1958–1967.
12. Pare, G., Chasman, D.I., Kellogg, M., Zee, R.Y., Rifai, N., Badola, S., Miletich, J.P. and Ridker, P.M. (2008) Novel association of ABO histo-blood group antigen with soluble ICAM-1: results of a genome-wide association study of 6,578 women. *PLoS Genet.*, **4**, e1000118.
13. Melzer, D., Perry, J.R., Hernandez, D., Corsi, A.M., Stevens, K., Rafferty, I., Lauretani, F., Murray, A., Gibbs, J.R., Paolisso, G. *et al.* (2008) A genome-wide association study identifies protein quantitative trait loci (pQTLs). *PLoS Genet.*, **4**, e1000072.
14. Hu, F.B., Meigs, J.B., Li, T.Y., Rifai, N. and Manson, J.E. (2004) Inflammatory markers and risk of developing type 2 diabetes in women. *Diabetes*, **53**, 693–700.
15. Greenwell, P. (1997) Blood group antigens: molecules seeking a function? *Glycoconj. J.*, **14**, 159–173.
16. Hosoi, E. (2008) Biological and clinical aspects of ABO blood group system. *J. Med. Invest.*, **55**, 174–182.
17. Yamamoto, F., McNeill, P.D. and Hakomori, S. (1992) Human histo-blood group A2 transferase coded by A2 allele, one of the A subtypes, is characterized by a single base deletion in the coding sequence, which results in an additional domain at the carboxyl terminal. *Biochem. Biophys. Res. Commun.*, **187**, 366–374.
18. Blann, A.D., Daly, R.J. and Amiral, J. (1996) The influence of age, gender and ABO blood group on soluble endothelial cell markers and adhesion molecules. *Br. J. Haematol.*, **92**, 498–500.
19. Kneuer, C., Ehrhardt, C., Radomski, M.W. and Bakowsky, U. (2006) Selectins—potential pharmacological targets? *Drug Discov. Today*, **11**, 1034–1040.
20. Okon, U.A., Antai, A.B., Osim, E.E. and Ita, S.O. (2008) The relative incidence of diabetes mellitus in ABO/Rhesus blood groups in south-eastern Nigeria. *Niger. J. Physiol. Sci.*, **23**, 1–3.
21. Williams, D.R. and Cartwrigth, R.A. (1979) Genetic polymorphisms in diabetics and non-diabetics. *J. Med. Genet.*, **16**, 351–357.
22. Sidhu, L.S., Malhotra, P. and Singh, S.P. (1988) ABO and Rh blood groups in diabetes mellitus. *Anthropologischer Anzeiger; Bericht. uber. die. biologisch-anthropologische. Literatur.*, **46**, 269–275.
23. Nemesure, B., Wu, S.Y., Hennis, A. and Leske, M.C. (2006) Hypertension, type 2 diabetes, and blood groups in a population of African ancestry. *Ethn. Dis.*, **16**, 822–829.
24. Iyengar, S., Hamman, R.F., Marshall, J.A., Baxter, J., Majumder, P.P. and Ferrell, R.E. (1989) Genetic studies of type 2 (non-insulin-dependent) diabetes mellitus: lack of association with seven genetic markers. *Diabetologia*, **32**, 690–693.
25. O'Donnell, J. and Laffan, M.A. (2001) The relationship between ABO histo-blood group, factor VIII and von Willebrand factor. *Transfus. Med.*, **11**, 343–351.
26. Zeggini, E., Scott, L.J., Saxena, R., Voight, B.F., Marchini, J.L., Hu, T., de Bakker, P.I., Abecasis, G.R., Almgren, P., Andersen, G. *et al.* (2008) Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.*, **40**, 638–645.
27. Zeggini, E., Weedon, M.N., Lindgren, C.M., Frayling, T.M., Elliott, K.S., Lango, H., Timpson, N.J., Perry, J.R., Rayner, N.W., Freathy, R.M. *et al.* (2007) Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science (New York, NY)*, **316**, 1336–1341.
28. Steinthorsdottir, V., Thorleifsson, G., Reynisdottir, I., Benediktsson, R., Jonsdottir, T., Walters, G.B., Styrkarsdottir, U., Gretarsdottir, S., Emilsson, V., Ghosh, S. *et al.* (2007) A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat. Genet.*, **39**, 770–775.
29. Sladek, R., Rocheleau, G., Rung, J., Dina, C., Shen, L., Serre, D., Boutin, P., Vincent, D., Belisle, A., Hadjadj, S. *et al.* (2007) A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*, **445**, 881–885.
30. Scott, L.J., Mohlke, K.L., Bonnycastle, L.L., Willer, C.J., Li, Y., Duren, W.L., Erdos, M.R., Stringham, H.M., Chines, P.S., Jackson, A.U. *et al.* (2007) A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science (New York, NY)*, **316**, 1341–1345.
31. Colditz, G.A., Manson, J.E. and Hankinson, S.E. (1997) The Nurses' Health Study: 20-year contribution to the understanding of health among women. *J. Womens Health*, **6**, 49–62.
32. Wolpin, B.M., Chan, A.T., Hartge, P., Chanock, S.J., Kraft, P., Hunter, D.J., Giovannucci, E.L. and Fuchs, C.S. (2009) ABO blood group and the risk of pancreatic cancer. *J. Nat. Cancer Inst.*, **101**, 424–431.
33. Qi, L., Rifai, N. and Hu, F.B. (2007) Interleukin-6 receptor gene variations, plasma interleukin-6 levels, and type 2 diabetes in U.S. Women. *Diabetes*, **56**, 3075–3081.
34. Pai, J.K., Pischon, T., Ma, J., Manson, J.E., Hankinson, S.E., Joshipura, K., Curhan, G.C., Rifai, N., Cannuscio, C.C., Stampfer, M.J. *et al.* (2004) Inflammatory markers and the risk of coronary heart disease in men and women. *New Engl. J. Med.*, **351**, 2599–2610.
35. Qi, L., van Dam, R.M., Liu, S., Franz, M., Mantzoros, C. and Hu, F.B. (2006) Whole-grain, bran, and cereal fiber intakes and markers of systemic inflammation in diabetic women. *Diabetes Care*, **29**, 207–211.
36. Korn, J.M., Kuruvilla, F.G., McCarroll, S.A., Wysoker, A., Nemesh, J., Cawley, S., Hubbell, E., Veitch, J., Collins, P.J., Darvishi, K. *et al.* (2008) Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat. Genet.*, **40**, 1253–1260.
37. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.
38. Peiffer, D.A., Le, J.M., Steemers, F.J., Chang, W., Jenniges, T., Garcia, F., Haden, K., Li, J., Shaw, C.A., Belmont, J. *et al.* (2006) High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. *Genome Res.*, **16**, 1136–1148.
39. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A. and Reich, D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.*, **38**, 904–909.
40. Petersen, K.F., Oral, E.A., Dufour, S., Befroy, D., Ariyan, C., Yu, C., Cline, G.W., DePaoli, A.M., Taylor, S.I., Gorden, P. *et al.* (2002) Leptin reverses insulin resistance and hepatic steatosis in patients with severe lipodystrophy. *J. Clin. Invest.*, **109**, 1345–1350.
41. Barrett, J.C., Fry, B., Maller, J. and Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics (Oxford, England)*, **21**, 263–265.
42. Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Gibbs, R.A., Belmont, J.W., Boudreau, A., Hardenbol, P., Leal, S.M. *et al.* (2007) A second generation human haplotype map of over 3.1 million SNPs. *Nature*, **449**, 851–861.