Genome-Wide Association Study on the Early-Phase Insulin Response to a Liquid Mixed Meal: Results From the NEO Study

Ruifang Li-Gao,[1] Françoise Carlotti,[2] Renée de Mutsert,[1] Astrid van Hylckama Vlieg,[1] Eelco J.P. de Koning,[2,3,4] J. Wouter Jukema,[5,6] Frits R. Rosendaal,[1] Ko Willems van Dijk,[4,7] and Dennis O. Mook-Kanamori[1,8]

Diabetes 2019;68:2327–2336 | https://doi.org/10.2337/db19-0378

Early-phase insulin secretion is a determinant of postprandial glucose homeostasis. In this study, we aimed to identify novel genetic variants associated with the early-phase insulin response to a liquid mixed meal by a genome-wide association study using a discovery and replication design embedded in the Netherlands Epidemiology of Obesity (NEO) study. The early-phase insulin response was defined as the difference between the natural logarithm–transformed insulin concentrations of the postprandial state at 30 min after a meal challenge and the fasting state (Δinsulin). After Bonferroni correction, rs505922 (β: −6.5% [minor allele frequency (MAF) 0.32, $P = 3.3 \times 10^{-8}$]) located in the *ABO* gene reached genome-wide significant level ($P < 5 \times 10^{-8}$) and was also replicated successfully (β: −7.8% [MAF 0.32, $P = 7.2 \times 10^{-5}$]). The function of the *ABO* gene was assessed using in vitro shRNA-mediated knockdown of gene expression in the murine pancreatic β-cell line MIN6. Knocking down the *ABO* gene led to decreased insulin secretion in the murine pancreatic β-cell line. These data indicate that the previously identified elevated risk of type 2 diabetes for carriers of the *ABO* rs505922:C allele may be caused by decreased early-phase insulin secretion.

Impaired β-cell function is considered a key factor in the pathogenesis of type 2 diabetes (T2D) driven by insulin resistance (1). Insulin secretion in response to an intravenous glucose stimulus is a two-phase process: the first peak of insulin secretion occurs rapidly within 5–10 min after the glucose infusion, followed by a second peak depending on the degree and duration of glucose stimulus (1). Although the insulin response to ingested glucose (e.g., from a meal) does not exhibit a clear biphasic shape under physiological conditions, an early insulin response with rapid elevations of portal and peripheral insulin concentrations has been observed (2,3). A previous study found that the plasma insulin response at 30 min after an oral glucose load was inversely associated with the 2-h plasma glucose concentrations in patients with impaired glucose tolerance (4). This implies that the early-phase insulin secretion is a marker for postprandial glucose homeostasis and plays a role in the development of T2D.

In the past decade, the genetics of glycemic traits including fasting glucose, fasting insulin, 2-h glucose after an oral glucose tolerance test (OGTT), and HbA_{1c} have been extensively investigated by genome-wide association studies (GWAS) (5–7). Many of the genes associated with glycemic traits from these GWAS are thought to be related to β-cell function. Among >120 loci that were identified to associate with T2D, a few loci that resided in the *ABO* gene repeatedly popped up in the GWAS. *ABO* gene encodes proteins that determine the ABO blood group system, and

[1]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands
[2]Division of Nephrology, Department of Internal Medicine, Leiden University Medical Center, Leiden, the Netherlands
[3]University Medical Center Utrecht, Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences, Utrecht, the Netherlands
[4]Division of Endocrinology, Department of Internal Medicine, Leiden University Medical Center, Leiden, the Netherlands
[5]Einthoven Laboratory for Experimental Vascular and Regenerative Medicine, Leiden University Medical Center, Leiden, Netherlands
[6]Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands
[7]Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands
[8]Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, the Netherlands

Corresponding author: Ruifang Li-Gao, r.li@lumc.nl

Received 12 April 2019 and accepted 9 September 2019

This article contains Supplementary Data online at http://diabetes.diabetesjournals.org/lookup/suppl/doi:10.2337/db19-0378/-/DC1.

non-O blood group carriers showed an increased risk of T2D, myocardial infarction, and peripheral vascular disease as well as venous thromboembolism from the previous epidemiological studies (8,9). From all the GWAS on glycemic traits, few studies have focused on the dynamic measures of insulin secretion and action to identify genetic variation that is associated with the insulin response. To date, there has only been only one large-scale GWAS focused on the dynamic measures of insulin secretion, reflected by the insulin levels at 30 min after an OGTT (10). From this study, in addition to previously identified glycemic trait–associated loci, the imprinted *GRB10* gene was identified as an islet function regulator in men only. However, OGTT only assesses one component of metabolism, namely, glucose metabolism. It has been suggested that the response of gastrointestinal hormones to dietary proteins and fat has considerable effects on β-cell insulin secretion after ingestion of a mixed meal (11,12), which might be underestimated by studying OGTT alone. Thus far, the genetic basis of the early-phase insulin secretion response to a mixed meal has not been studied.

In the current study, we conducted a GWAS on early-phase insulin response to a liquid mixed meal. Discovery and replication were both performed in a large population-based cohort study, the Netherlands Epidemiology of Obesity (NEO) study. To further shed light on the function of the top genetic signal from the GWAS, we performed an in vitro shRNA knockdown gene expression experiment of the top genetic signal from the GWAS by measuring glucose-stimulated insulin secretion in the murine pancreatic β-cell line MIN6.

RESEARCH DESIGN AND METHODS

Study Design and Study Population

The study was performed in a population-based prospective cohort study, the NEO study (13). All participants gave written informed consent, and the Medical Ethical Committee of the Leiden University Medical Center (LUMC) approved the study design. Initiated from 2008, NEO was designed to study pathways that lead to obesity-related diseases. Detailed information about the study design and data collection has previously been published (13). Men and women aged between 45 and 65 years with a self-reported BMI of 27 kg/m^2 or higher living in the greater area of Leiden (in the west of the Netherlands) were eligible to participate in NEO. In addition, all inhabitants aged between 45 and 65 years from one municipality (Leiderdorp) were invited irrespective of BMI. Participants were invited for a baseline visit at the NEO study center in the LUMC after an overnight fast. Prior to the visit, participants completed a questionnaire at home with demographic, lifestyle, and clinical data. At the baseline visit, fasting blood samples were drawn. Within the next 5 min after the fasting blood draw, a liquid mixed meal (400 mL with 600 kcal, with 16% of energy [En%] derived from protein, 50 En% carbohydrates, and 34 En% fat) was consumed and subsequent blood samples were drawn 30 and 150 min after the liquid mixed meal. Individuals were excluded from the analyses (Supplementary Fig. 1) if they *1*) had self-reported type 1 or T2D or were taking any glucose-lowering medication, *2*) had fasting glucose measures >7 mmol/L at baseline, *3*) had HbA_{1c} >6.5% at baseline, *4*) violated overnight fasting, or *5*) violated liquid meal challenge protocol.

Genotyping and Imputation

DNA was extracted from venous blood samples obtained from the antecubital vein. Genotyping was performed in Centre National de Génotypage (Evry Cedex, France), using the Illumina HumanCoreExome-24 BeadChip (Illumina, San Diego, CA). The detailed quality-control process has previously been described (14). Genotypes were further imputed to the 1000 Genome Project reference panel (version 3, 2011) (15) using IMPUTE (version 2.2) software (16). No genetic variants with an imputation quality <0.4 or a minor allele frequency (MAF) <0.01 were considered for the analyses in the current study.

Laboratory Insulin Measurements and Early-Phase Insulin Response Definitions

Participants fasted for at least 10 h prior to the NEO baseline visits. Blood samples were drawn after this overnight fast as well as 30 and 150 min after the consumption of a liquid mixed meal. Blood was drawn in tubes that were immersed in ice after collection and frozen. The tube contained frozen serum with a volume of 350 μL. Insulin was measured by insulin kit 200T-L2KIN2, an immunometric sandwich assay method (Siemens DPC Immulite 2000 analyzer; Siemens Healthcare Diagnostics). Glucose concentrations were determined by using standard enzymatic methods (Modular Analytics P800; Roche Analytics, Almere, the Netherlands).

Early-phase insulin response measures were estimated by four different methods: *1*) insulin concentration at 30 min ($insulin_{30}$); *2*) insulin concentration ratio between 30 min after the liquid meal and fasting state ($insulin_{30}/insulin_0$); *3*) insulinogenic index (IGI), defined as a ratio of the area under the curve (AUC) of insulin to glucose, i.e., $AUC_{insulin(0–30–150\ min)}/AUC_{glucose(0–30–150\ min)}$, calculated by the trapezium rule; and *4*) insulin response to glucose during the first 30 min, adjusted for BMI ($IR_{BMIadj} = insulin_{30}/(glucose_{30} \times BMI)$) (10). The four derived measures were natural logarithm transformed to better approximate normally distributed variables, and the second measure by insulin concentration ratio between 30 min after the liquid meal and fasting state was therefore equivalent to the difference of the natural logarithm–transformed insulin concentrations between 30 min after the liquid meal and fasting state: $\Delta insulin = \ln(insulin_{30}) - (insulin_0)$. Extreme measures, i.e., when the natural logarithm–transformed values were >4 SD from the mean, were excluded, with one individual being removed in both the discovery and replication population for Δinsulin as well as IR_{BMIadj} separately and four individuals

removed in the replication population for Δinsulin (Supplementary Fig. 1).

GWAS

We conducted the GWAS on all the four predefined early-phase insulin response measures across 22 autosomal chromosomes. The previous GWAS on T2D did not show substantial heterogeneity of odds ratios estimated from obese and nonobese populations (17), so we divided the total NEO study population based on the geographical area of recruitment into a discovery cohort with oversampling on the overweight and obese population and a replication cohort with a reference population of no BMI restriction (13). Additive linear regression analyses were performed separately for the discovery and replication cohort by SNPTEST v2, with adjustment for age, sex, BMI, and the first four principal components to identify BMI-independent genetic signals. A P value $<5 \times 10^{-8}$ was considered to be genome-wide significant, and a P value $<1 \times 10^{-6}$ was considered a suggestive signal. Independent single nucleotide polymorphisms (SNPs) with a P value $<1 \times 10^{-6}$ in the discovery stage were validated in the replication cohort. A P value <0.05 for SNPs in the replication cohort were considered to indicate a successful replication. For further exploration of the BMI-dependent signals, association analyses were performed in the entire NEO cohort for the four predefined phenotypes, with adjustment for age, sex, and the first four principal components.

For identification of independent genetic loci that were associated with each early-phase insulin response measure, the linkage disequilibrium (LD)-based clumping procedure was applied using PLINK (18). As a start, all the SNPs with P values below a certain threshold (-clump-p1) were set as "index" SNPs, and the other SNPs were clustered into different clumps based on their LD and physical proximity to the index SNPs controlled by clump-r2 and clump-kb separately in the command (19). In the current analyses, the following parameters were adopted: –clump-p1, 1.0×10^{-6}; –clump-r2, 0.1; and –clump-kb, 1,000. The LD patterns were based on the 1000 Genome v3 20101123 reference set of Utah residents (CEPH) with Northern and Western European Ancestry population (15). For further verification of the independence of the selected loci from clumping, additional conditional analyses were performed using the genome-wide complex trait analysis (GCTA) tool with the command –cojo-cond (version 1.24.4) (20) in the entire chromosome region of the tag SNP, by conditioning on the tag SNP itself. The explained variance of the independent SNPs in each insulin response measure was estimated in the replication cohort. For each individual SNP, the explained variance was estimated as the partial R^2 from the linear regression model, with the SNP as the independent variable and insulin response (natural logarithm–transformed measures) as the dependent variable. In order to increase the statistical power, we pooled both discovery and replication cohorts to run the GWAS as a sensitivity analysis, followed by clumping and conditional analysis (similar to the split cohort analysis). As T2D is a heterogeneous disease with >150 genetic variants that have been identified from the previous GWAS (21), the associations between T2D-associated SNPs and insulin response were further highlighted to disentangle the potential insulin secretion–related T2D-associated loci.

Functional Annotation of Top SNPs

The independent top variants were searched in the human pancreatic islet *cis*-eQTL summary data from Varshney et al. (22) to identify the human pancreatic islet-specific expression patterns (expression quantitative trait loci [eQTLs]) of mRNAs associated with these variants. To further verify whether eQTL signals act as the putative effector transcripts of the GWAS findings, we performed colocalization analysis using (Approximate) Bayes Factor colocalization analyses (coloc.abf), in the R package coloc, by P values extracted from both our GWAS summary data and the eQTL summary data published by Varshney et al. (23,24). Bayesian colocalization analysis evaluated the posterior probabilities of five different configurations based on a prior probability on the SNP level, and the configuration with the highest posterior probability will be the most likely situation. Subsequently, Bayesian fine mapping analysis was performed by the method called "Bayesian fine mapping analysis" (finemap.abf) implemented in the R package coloc (24) to determine the probability of the top SNPs being causal for early-phase insulin response. All the SNPs with $P < 5 \times 10^{-8}$ in the same LD block as the top signal, which was tagged as lead SNP by plink clumping, were tested in coloc.abf and finemap.abf. The phenome-wide association study of independent top variants was performed by GeneAtlas based on 118 nonbinary and 599 binary traits of 408,455 related and unrelated UK Biobank participants (25). The regulatory function of the top variants were screened in HaploReg database (26), with the aim to explore their potential chromatin state and regulatory motif alternations.

ABO Blood Group Determined by Genotyping in the NEO Study

Since the GWAS results indicated that ABO blood type and insulin response to a mixed liquid meal were associated, four *ABO* SNPs (rs8176719:insC, rs7853989:G>C, rs8176749:G>A, and rs8176750:delC) were used to discriminate among the common *ABO* alleles: O^1, O^2, A^1, A^2, and B (27,28) (Supplementary Data). Six phenotypic ABO blood types (A_1, A_2, A_1B, A_2B, B, and O) were further derived from the 14 *ABO* allele combinations (29). Among them, A_1, A_2, A_1B, A_2B, and B were clustered as the non-O group in the analysis. The comparisons of insulin response difference (Δinsulin) between the O and non-O groups were performed by Student t test, and ANOVA was used to compare the mean differences among the six phenotypic ABO blood type groups on

Table 1—Characteristics of the discovery and replication cohort from the NEO study

	$Insulin_{30}$		ΔInsulin		IGI		IR_{BMIadj}	
Characteristics	Discovery	Replication	Discovery	Replication	Discovery	Replication	Discovery	Replication
N	3,526	1,239	3,518	1,233	3,441	1,202	3,503	1,228
Age (years)	56 (51, 61)	56 (51, 61)	56 (51, 61)	56 (51, 61)	56 (51, 61)	56 (51, 61)	56 (51, 61)	56 (51, 61)
Women, *N* (%)	1,757 (50.0)	678 (54.7)	1,753 (49.8)	674 (54.7)	1,701 (49.4)	654 (54.4)	1,743 (49.8)	669 (54.5)
BMI (kg/m^2)	30.1 (28.3, 32.6)	25.5 (23.2, 27.9)	30.1 (28.3, 32.6)	25.5 (23.3, 27.9)	30.1 (28.3, 32.6)	25.5 (23.2, 27.9)	30.1 (28.3, 32.6)	25.5 (23.2, 27.9)
Fasting serum measures								
Glucose (mmol/L)	5.5 (5.2, 5.9)	5.3 (5.0, 5.6)	5.5 (5.2, 5.9)	5.3 (5.0, 5.6)	5.5 (5.2, 5.9)	5.3 (5.0, 5.6)	5.5 (5.2, 5.9)	5.3 (5.0, 5.6)
Insulin (mU/L)	10.9 (7.5, 15.7)	7.8 (5.2, 11.7)	10.9 (7.5, 15.7)	7.8 (5.2, 11.7)	10.8 (7.4, 15.6)	7.7 (5.2, 11.7)	10.9 (7.5, 15.7)	7.8 (5.2, 11.7)
HbA_{1c} (%)	5.3 (5.2, 5.5)	5.3 (5.1, 5.4)	5.3 (5.2, 5.5)	5.3 (5.1, 5.4)	5.3 (5.2, 5.5)	5.3 (5.1, 5.4)	5.3 (5.2, 5.5)	5.3 (5.1, 5.4)
Serum measures at 30 min								
Glucose (mmol/L)	7.1 (6.2, 8.1)	6.2 (5.3, 7.2)	7.1 (6.2, 8.1)	6.2 (5.3, 7.2)	7.1 (6.2, 8.1)	6.2 (5.3, 7.2)	7.1 (6.2, 8.1)	6.2 (5.3, 7.2)
Insulin (mU/L)	61.4 (42.8, 89.6)	49.1 (34.9, 69.1)	61.3 (42.7, 89.6)	49.1 (34.9, 69.0)	61.3 (42.8, 89.4)	48.8 (34.6, 68.9)	61.4 (42.8, 89.6)	49.1 (34.8, 69.1)

Results are presented as median (interquartile range) unless otherwise indicated.

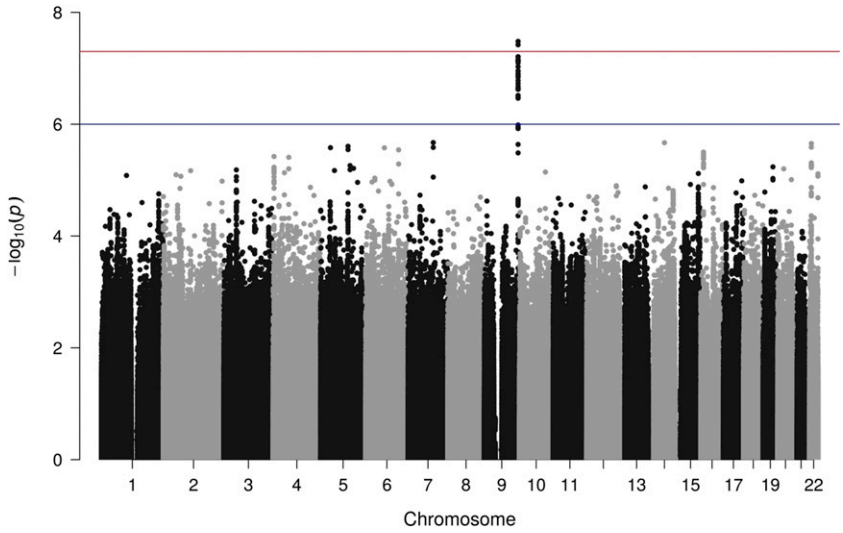

Figure 1—Manhattan plot for the GWAS on Δinsulin, ln(insulin30) − ln(insulin0), in the discovery cohort (n = 3,518).

insulin response (Δinsulin). For estimation of the effect sizes, and with 95% CI, of ABO blood type groups on Δinsulin, linear regression models were used with age and sex adjusted for in the model. For the classification of O and non-O groups, the non-O group was considered the reference group; for the six phenotypic ABO blood group types, blood group A_1 was used as the reference.

Modeling the Effects of *ABO* Gene Expression Changes on Insulin Secretion In Vitro

We assessed the effects of lentivirus-mediated shRNA knockdown of *ABO* gene expression on glucose-stimulated insulin secretion in the murine pancreatic β-cell line MIN6. Lentivirus vectors were obtained from the Sigma-Aldrich MISSION library (murine ABO [shRNA ABO] clone TRCN0000110442 and nontarget control [shRNA control] clone SHC-002) and produced in the facility of R.C. Hoeben (Department of Cell and Chemical Biology, LUMC) as previously described (30). Cells were used 4 days post-transduction, and all experiments were repeated three times (Supplementary Data).

Data and Resource Availability

The data sets generated (all the GWAS summary statistics results) during and analyzed during the current study are available from the corresponding author on request.

RESULTS

BMI was higher in the discovery cohort than in the replication (30.1 kg/m^2 vs. 25.5 kg/m^2, respectively) in all the four derived insulin response measures (Table 1). In addition, fasting glucose and insulin were also slightly higher in the discovery cohort compared with the replication (glucose 5.5 mmol/L vs. 5.3 mmol/L and insulin 10.8–10.9 mU/L vs. 7.7–7.8 mU/L). However, there was no difference in HbA_{1c} between the two subpopulation (5.3% vs. 5.3%).

Three out of the four early-phase insulin response measures were highly correlated (Supplementary Fig. 2), whereas Δinsulin, $\ln(\text{insulin}_{30}) - \ln(\text{insulin}_0)$ showed the most dissimilarity to the other three measures (Pearson correlations to insulin_{30}, IGI, and IR_{BMIadj}: 0.38, and 0.21, 0.46 separately). The Manhattan plot for the GWAS in the discovery population on Δinsulin is shown in Fig. 1, and the other three Manhattan plots, for insulin_{30}, IGI, and IR_{BMIadj}, are shown in Supplementary Figs. 4–6. After clumping and conditional analysis on the lead SNPs, six independent genome-wide significant SNPs ($P < 5 \times 10^{-8}$) by four different insulin response measures were identified in the discovery phase (Table 2). However, only rs505922:T>C from the GWAS on Δinsulin (β: −6.5% [MAF 0.32, $P = 3.3 \times 10^{-8}$]) was successfully replicated (β: −7.8% [MAF 0.32, $P = 7.2 \times 10^{-5}$]). Another ten SNPs across four different measures reached suggestive significance ($P < 1 \times 10^{-6}$) levels, which are shown in Supplementary Table 1. However, none of these SNPs were replicated.

Rs505922:T>C, located in the first intron of *ABO* gene, explained 1.2% of the total variation in the natural logarithm–transformed insulin response (Δinsulin) and had a per-allele decrease in Δinsulin of 6.5%. This SNP is in high LD with the ABO blood type determining SNP rs657152:C>A (r^2 = 0.90 in NEO). However, the LDs to the previously reported glycemic traits–associated SNPs located in *ABO* gene are relatively moderate (LDs to rs651007:T>C 0.40, rs579459:C>T 0.40, rs635634:T>C 0.48, and rs507666:G>A 0.49). Rs505922 has been reported as a human pancreatic islet eQTL for *ABO* gene expression in two independent studies with significant P values (31,32). In total, 34 SNPs located in the same LD block as rs505922:T>C were tested for the colocalization to the pancreatic islet eQTL reported by Varshney et al. (22) and their probability of being causal for the

Table 2—Independent lead SNPs that reached genome-wide significance under four different measures of early-phase insulin response

									Discovery cohort			Replication cohort		
Chr	SNP	Phenotype	Position	Location	Gene	Effect/ noneffect allele	Effect allele frequency[a]	Imputation quality[a]	Effect size per allele (%)[b]	SE	*P*	Effect size per allele (%)[b]	SE	*P*
3	rs115404340	$insulin_{30}$	67839179	Intron	*SUCLG2-AS1*	G/A	0.01	0.44	78.5	0.098	3.8E-9	−15.7	0.13	0.19
9	rs505922	ΔInsulin	136149229	Intron	*ABO*	C/T	0.32	1	−6.5	0.012	3.3E-8	−7.8	0.021	7.2E-5*
9	rs657152	IGI	136139265	Intron	*ABO*	A/C	0.34	0.98	8.1	0.011	3.0E-12	2.7	0.017	0.11
19	rs74889068	IGI	46199363	Intron	*QPCTL*	A/G	0.15	0.93	−8.8	0.016	3.8E-9	−2.1	0.024	0.36
3	rs115404340	IR_{BMIadj}	67839179	Intron	*SUCLG2-AS1*	G/A	0.01	0.42	72.5	0.092	3.4E-9	−9.8	0.12	0.41
19	rs74889068	IR_{BMIadj}	46199363	Intron	*QPCTL*	A/G	0.15	0.93	−10.9	0.018	4.4E-11	−4.0	0.028	0.14

Threshold for genome-wide significance is 5×10^{-8}. Chr, chromosome. *A successful replication, with *P* value < 0.05. [a]In the discovery cohort. [b]β-Coefficient expressed as the percentage difference in the outcome by one copy of the effect allele.

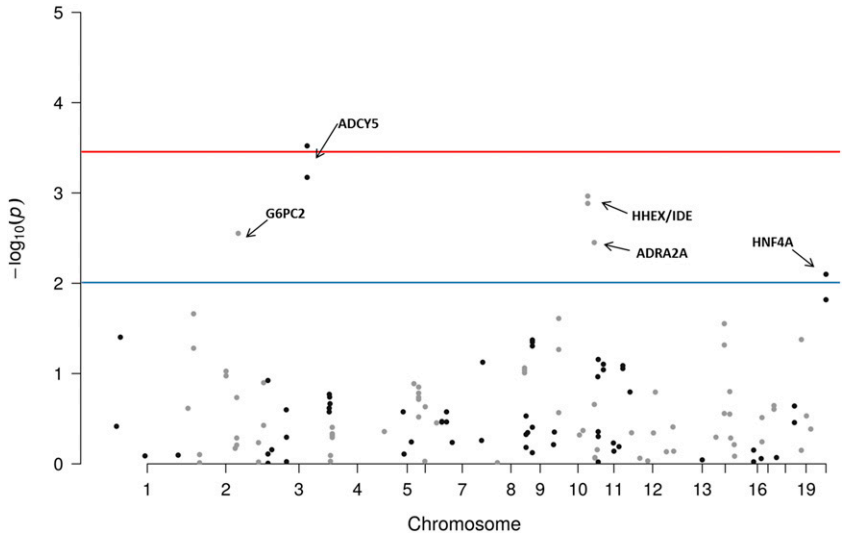

Figure 2—The associations between T2D-associated SNPs (N = 143) and Δinsulin, ln(insulin30) − ln(insulin0), in the entire NEO cohort (N = 4,751).

early-phase insulin response. When we performed the Bayesian colocalization analysis on the 34 SNPs located in the same LD block as rs505922:T>C, the hypothesis H4, indicating the posterior probability of both traits being associated and sharing a single causal variant, was 96.7%, indicating a shared causal variant associated with both early-phase insulin response to a meal and pancreatic islet ABO gene expression. The median probability of these tested 34 SNPs was 3%, with the highest probability, 8%, assigned to rs505922:T>C. From the regulatory annotations in the HaploReg database, rs505922:T>C resides in a DNase I–hypersensitive site in the pancreas and has enhancer marks (e.g., H3K4me1 and H3K4me3) recorded in the ENCODE (Encyclopedia of DNA Elements) consortium data. By the phenome-wide association studies of rs505922:T>C in the UK Biobank, 35 traits reached the genome-wide significant threshold ($P < 5 \times 10^{-8}$), and the top associated phenotype is the venous thromboembolic disease ($P = 7.0 \times 10^{-86}$) (Supplementary Table 2).

Among 151 previously identified T2D-associated SNPs, 143 were found in the current NEO imputed genotype data (5 SNPs were removed due to MAF <1% and 3 SNPs were not present in the imputed data). Figure 2 shows the associations of these T2D-associated SNPs with Δinsulin in the entire NEO study population. None of the signals reached genome-wide significance ($P < 5 \times 10^{-8}$). However, when the significance level was set to the candidate SNP-wide significance ($P < 5.2 \times 10^{-4}$, 0.05/96, where 96 is the number of independent loci), rs2877716:T>C, located in the intron of *ADCY5*, was significant. Another six T2D-assocaited SNPs reached candidate SNP-wide suggestive significance ($P < 1 \times 10^{-2}$).

After pooling the discovery and replication cohort, a GWAS was performed on the Δinsulin measure in all the NEO samples (Supplementary Fig. 7). After clumping and conditional analysis, two independent signals were identified (Supplementary Table 3). The top signal, rs676996:T>G, is in the same locus as rs505922:T>C (LD 0.97, β: −6.9% [MAF 0.32, $P = .3 \times 10^{-11}$]), located within the first intron of *ABO* gene. A new signal was identified in the second intron of *QPCTL* (rs2287019:C>T) (β:−6.9% [MAF 0.19, $P = 3.3 \times 10^{-8}$]). However, rs2287019 had a much weaker signal in the discovery cohort alone (β: −5.5% [MAF 0.19, $P = 1.4 \times 10^{-4}$). Compared with the top signals after BMI adjustment, BMI-dependent GWAS (without BMI adjustment in the model) for all the four predefined early-phase insulin response measures identified nearly the same loci (Supplementary Figs. 8–11).

The Δinsulin was further compared across six ABO blood types as well as the combined O and non-O blood type groups in the entire NEO population (Fig. 3*A*-*C*). On average, insulin concentrations were lower among the O blood type group than non-O both at fasting baseline and 30 min after the mixed-meal challenge (Fig. 3*A*), with a slope difference test P value of 5.3×10^{-4}. Similarly, a stronger insulin response (Δinsulin) was observed in the O blood group than in the non-O group (Fig. 3*C*) (t test $P = 1.8 \times 10^{-4}$), indicating a higher potential of postprandial insulin concentrations to increase among O blood group versus non-O blood group carriers. Compared with non-O group, individuals in the O blood type group had a larger Δinsulin: 0.090 (95% CI 0.060–0.12).

Next, in vitro experiments were performed in the murine pancreatic β-cell line MIN6. *ABO* gene expression was reduced by 60% in MIN6 cells transduced with the shRNA ABO lentivirus compared with the nontarget shRNA controls (Supplementary Fig. 12*A*). After glucose stimulation, the increase in insulin secretion was lower after transduction with the shRNA ABO lentivirus than in

the controls, indicating that downregulation of the *ABO* gene led to a reduced glucose-stimulated insulin secretion (Supplementary Fig. 12*B*).

DISCUSSION

Rs505922 located in the first intron of the *ABO* gene was associated with Δinsulin and was also successfully replicated in another subsample of the NEO population. Further analyses revealed that ABO blood type and Δinsulin were associated. Blood group O carriers had a lower fasting insulin level but showed a larger increase upon the mixed meal than the non-O blood group carriers. By in vitro *ABO* gene downregulation, a decreased insulin secretion was observed.

A recent study reported the association between rs505922:C in the *ABO* gene and an increased risk of T2D (odds ratio 1.06 [95% CI 1.04–1.09]) (33). Rs505922 is in strong LD ($r^2 = 0.90$) with the ABO blood type determining SNP rs8176719, and in a previous observational study (8), it was shown that the non-O blood group carriers have an increased risk of developing T2D. It can therefore be hypothesized that β-cell function and insulin secretion capacity are different between the different ABO blood groups, which subsequently leads to different risks

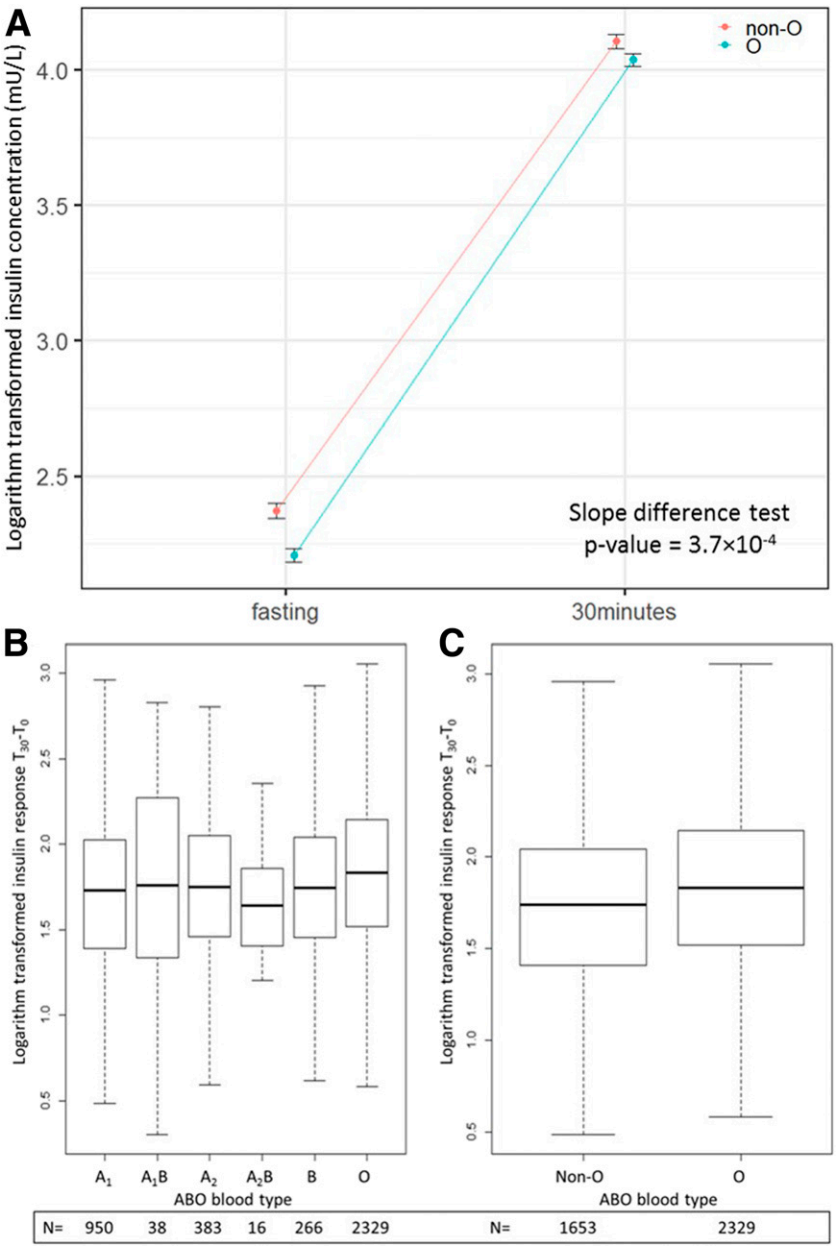

Figure 3—The distributions of natural logarithm–transformed insulin response across different ABO blood groups in the NEO study.

of T2D. The ABO gene encodes a glycosyltransferase that determines the glycosylation status of the glycoprotein H antigen that is expressed on all normal blood red cells. The A and B alleles encode enzymes with different activity, whereas the O allele is a null variant. The ABO gene is expressed in a variety of tissues and thus in theory could affect the glycosylation status of numerous proteins. Our results indicate the activity of one or more of the proteins that are directly or indirectly involved in insulin secretion or clearance after a meal is affected by ABO-mediated glycosylation status. Since the substrate specificity of the ABO protein beyond the glycoprotein H antigen is poorly defined, the identity of these proteins requires additional work. However, given the pleiotropic nature of the disease associations observed with the ABO gene locus, it is likely that ABO affects a variety of substrates. Although there are extensive differences between in vitro and in vivo conditions, the in vitro experiment described in the current study showed a clear effect on glucose-stimulated insulin secretion after *ABO* gene downregulation in a mouse cell line. Taking all these factors together, the previously reported high risk of T2D in those with non-O blood type may be mediated through an increased baseline insulin level and decreased capacity to increase insulin levels in the early-phase insulin response.

In addition to rs505922, alternative loci near the *ABO* gene have been associated with either fasting glucose levels (rs651007, rs579459, and rs507666) or the risk of T2D (rs635634) (7,34). All of these four SNPs are located in the intergenic region near the *ABO* gene. However, these SNPs show modest LDs to rs505922, which resides in the first intron of the *ABO* gene. This observation may very well be explained by the omnigenic model for complex traits in the GWAS (35). In this hypothetical model, the disease risk is affected by a few strong effect "core" functional genetic variants as well as numerous small effect "peripheral" genes. Together they form a highly connected network. Some perturbation on the peripheral genes will transmit to the "core" gene and affect its function. It has also been shown that disease-associated genetic signals are significantly enriched in regions that are transcribed actively, which partially explained the observation of a large amount of genetic variants identified from GWAS located in the gene regulatory regions. Therefore, rs505922 may be another "peripheral" genetic signal that is involved in T2D pathophysiological pathways, similar to other genetic markers identified previously in the intron and intergenic regions of *ABO* gene and other T2D-associated genes.

Several methodological aspects should be considered. The main strength of this study is the liquid meal that was provided to all the NEO participants. Moreover, epidemiological data were further investigated with functional analyses to verify the role of the *ABO* gene in the early-phase insulin response. Nonetheless, the sample size of the GWAS was too small to identify genetic variants with low-frequency and rare variants. Finally, another strength is that we followed up the top signal of the GWAS with an in vitro experiment.

Conclusion

A genetic variant, rs505922, in an intron of the *ABO* gene, showed an association with the early-phase insulin response to a liquid meal measured by Δinsulin, log ($insulin_{30}/insulin_0$), and a subsequent in vitro analysis showed that knocking down the *ABO* gene affected glucose-stimulated insulin secretion. A phenotypic difference in Δinsulin between O and non-O blood type groups was found, which may explain the role of ABO blood type in the risk of developing T2D.

Acknowledgments. The authors thank all individuals who participated in the NEO study and all participating general practitioners for inviting eligible participants. The authors also thank the NEO study group, Pat van Beelen (Department of Clinical Epidemiology, Leiden University Medical Center) and all research nurses for data collection, Petra Noordijk (Department of Clinical Epidemiology, Leiden University Medical Center) and her team for laboratory management, and Ingeborg de Jonge (Department of Clinical Epidemiology, Leiden University Medical Center) for data management of the NEO study. The authors thank Nutricia Research, Utrecht, the Netherlands, for providing the liquid mixed meal.

Funding. The NEO study is supported by the Netherlands Cardiovascular Research Initiative, an initiative with support of the Dutch Heart Foundation (CVON2014-02 ENERGISE). D.O.M.-K. is supported by the Dutch Science Organization (ZonMW-VENI grant 916.14.023).

Duality of Interest. D.O.M.-K. is a part-time clinical research consultant for Metabolon, Inc. No other potential conflicts of interest relevant to this article were reported.

Author Contributions. R.L.-G. performed the analysis and wrote and edited the manuscript. F.C. conducted the in vitro experiment and drafted and reviewed the manuscript. R.d.M. and F.R.R. contributed to study design and manuscript review. A.v.D.V., E.J.P.d.K., and J.W.J. contributed to results interpretation and manuscript review. K.W.v.D. and D.O.M.-K. contributed to study design, conceived the idea of the current study, interpreted the results, and reviewed the manuscript. R.L.-G. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Parts of this study were presented in abstract form at the Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium meeting, Rotterdam, the Netherlands, 18–19 April 2018.

References

1. Gerich JE. Is reduced first-phase insulin release the earliest detectable abnormality in individuals destined to develop type 2 diabetes? Diabetes 2002; 51(Suppl. 1):S117–S121

2. Caumo A, Luzi L. First-phase insulin secretion: does it exist in real life? Considerations on shape and function. Am J Physiol Endocrinol Metab 2004;287: E371–E385

3. Vollmer K, Holst JJ, Baller B, et al. Predictors of incretin concentrations in subjects with normal, impaired, and diabetic glucose tolerance. Diabetes 2008;57: 678–687

4. Mitrakou A, Kelley D, Mokan M, et al. Role of reduced suppression of glucose production and diminished early insulin release in impaired glucose tolerance. N Engl J Med 1992;326:22–29

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

6. Saxena R, Hivert MF, Langenberg C, et al.; GIANT consortium; MAGIC investigators. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. Nat Genet 2010;42:142–148
7. Scott RA, Lagou V, Welch RP, et al.; DIAbetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. Nat Genet 2012;44:991–1005
8. Fagherazzi G, Gusto G, Clavel-Chapelon F, Balkau B, Bonnet F. ABO and Rhesus blood groups and risk of type 2 diabetes: evidence from the large E3N cohort study. Diabetologia 2015;58:519–522
9. Larson NB, Bell EJ, Decker PA, et al. ABO blood group associations with markers of endothelial dysfunction in the Multi-Ethnic Study of Atherosclerosis. Atherosclerosis 2016;251:422–429
10. Prokopenko I, Poon W, Mägi R, et al. A central role for GRB10 in regulation of islet function in man. PLoS Genet 2014;10:e1004235
11. Brown JC, Otte SC. Gastrointestinal hormones and the control of insulin secretion. Diabetes 1978;27:782–787
12. Ebert R, Creutzfeldt W. Gastrointestinal peptides and insulin secretion. Diabetes Metab Rev 1987;3:1–26
13. de Mutsert R, den Heijer M, Rabelink TJ, et al. The Netherlands Epidemiology of Obesity (NEO) study: study design and data collection. Eur J Epidemiol 2013;28:513–523
14. Blauw LL, Li-Gao R, Noordam R, et al. CETP (cholesteryl ester transfer protein) concentration: a genome-wide association study followed by Mendelian randomization on coronary artery disease. Circ Genom Precis Med 2018;11:e002034
15. 1000 Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, et al. An integrated map of genetic variation from 1,092 human genomes. Nature 2012;491:56–65
16. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 2007;39:906–913
17. Voight BF, Scott LJ, Steinthorsdottir V, et al.; MAGIC investigators; GIANT Consortium. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 2010;42:579–589
18. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–575
19. Draisma HHM, Pool R, Kobl M, et al. Genome-wide association study identifies novel genetic variants contributing to variation in blood metabolite levels. Nat Commun 2015;6:7208
20. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011;88:76–82
21. Prasad RB, Groop L. Genetics of type 2 diabetes-pitfalls and possibilities. Genes (Basel) 2015;6:87–123
22. Varshney A, Scott LJ, Welch RP, et al.; NISC Comparative Sequencing Program. Genetic regulatory signatures underlying islet gene expression and type 2 diabetes. Proc Natl Acad Sci U S A 2017;114:2301–2306
23. Giambartolomei C, Vukcevic D, Schadt EE, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet 2014;10:e1004383
24. Coloc: a package for colocalisation analyses [Internet], 2019. Available from https://cran.r-project.org/web/packages/coloc/vignettes/vignette.html
25. Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. Nat Genet 2018;50:1593–1599
26. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res 2012;40:D930–D934
27. Olsson ML, Chester MA. Polymorphism and recombination events at the ABO locus: a major challenge for genomic ABO blood grouping strategies. Transfus Med 2001;11:295–313
28. Ocak G, Vossen CY, Verduijn M, et al. Risk of venous thrombosis in patients with major illnesses: results from the MEGA study. J Thromb Haemost 2013;11:116–123
29. El-Zawahri MM, Luqmani YA. Molecular genotyping and frequencies of A1, A2, B, O1 and O2 alleles of the ABO blood group system in a Kuwaiti population. Int J Hematol 2008;87:303–309
30. Carlotti F, Bazuine M, Kekarainen T, et al. Lentiviral vectors efficiently transduce quiescent mature 3T3-L1 adipocytes. Mol Ther 2004;9:209–217
31. Fadista J, Vikman P, Laakso EO, et al. Global genomic and transcriptomic analysis of human pancreatic islets reveals novel genes influencing glucose metabolism. Proc Natl Acad Sci U S A 2014;111:13924–13929
32. van de Bunt M, Manning Fox JE, Dai X, et al. Transcript expression data from human islets links regulatory signals from genome-wide association studies for type 2 diabetes and glycemic traits to their downstream effectors. PLoS Genet 2015;11:e1005694
33. Bonàs-Guarch S, Guindo-Martínez M, Miguel-Escalada I, et al. Re-analysis of public genetic data reveals a rare X-chromosomal variant associated with type 2 diabetes. Nat Commun 2018;9:321
34. Wessel J, Chu AY, Willems SM, et al.; EPIC-InterAct Consortium. Low-frequency and rare exome chip variants associate with fasting glucose and type 2 diabetes susceptibility. Nat Commun 2015;6:5897
35. Boyle EA, Li YI, Pritchard JK. An expanded view of complex traits: from polygenic to omnigenic. Cell 2017;169:1177–1186