SUPPLEMENTARY DATA

ABO blood group determined by genotyping in the NEO study

In the NEO genotype data, only rs8176749:G>A was directly genotyped on the array. Rs7853989:G>C was derived by imputation, with an imputation quality of 0.99. Rs8176719:insC and rs8176750:delC were not available and replaced by rs8176645:A>T (R2=0.98 with rs8176719:insC, imputation quality of 0.80) and rs8176704:G>A (R2=0.99 with rs8176750:delC, imputation quality of 0.99) separately. The best guess genotypes for the imputed SNPs were determined by the highest probability of the genotype with a minimum threshold of 0.7. If the highest probability of the genotype was below the minimum threshold, the genotype of that SNP was set to missing. Theoretically, 15 combinations of the four ABO alleles were possible, and 14 out of the 15 combinations were observed in the current NEO study population.

Modelling the effects of ABO gene expression changes on insulin secretion in vitro

The knock-down efficiency was assessed by Q-PCR (Light Cycler 480-II Real-time PCR system - Roche) on total RNA extracted using RNeasy kit (Qiagen), and reversed transcribed using M-MLV reverse transcriptase (Invitrogen). Fold induction was calculated using the DeltaCT method with GAPDH as housekeeping gene. The primers used were the following: ABO-Fw: ACAGACACTGAACCATCCTGGGTT; ABO-Rv: AGACAAACACTGCGAAGGGAAGGA; GAPDH-Fw: ATCATCCCTGCATCCACTG; GAPDH-Rv: ATCATACTTGGCAGGTTTCTCC.

Transduced MIN6 cells were subjected to a GSIS assay. Briefly, cells were pre-incubated in a modified Krebs-Ringer bicarbonate buffer with HEPES (KRBH) containing 2 mmol/l glucose for 1.5 h at 37°C. They were then incubated in 2 mmol/l glucose KRBH buffer for 1 h at 37°C, then switched to 25 mmol/l glucose KRBH buffer for 1 h at 37°C, and again switched to 2 mmol/l glucose KRBH buffer for 1 h at 37°C. Insulin concentrations were determined in supernatant fractions by ELISA (Mercodia, Uppsala, Sweden). At the end of the assay, MIN6 cells were lysed by sonication in distilled water and total DNA content was determined by Quant-iT PicoGreen dsDNA kit (Invitrogen) for normalization.

SUPPLEMENTARY DATA

Supplementary Table S1. Independent suggestive SNPs by the four different measures of early phase insulin response.

Chr	SNP	Phenotype	Position	Location	Gene	Effect/ Non-effect allele	Effect allele frequency[a]	Imputation quality	Discovery cohort			Replication cohort		
									Effect size per allele[b] (%)	SE	P-value	Effect size per allele[b] (%)	SE	P-value
Suggestive SNPs														
19	rs74889068	$Insulin_{30}$	46199363	Intron	*QPCTL*	A/G	0.15	0.94	-9.7	0.019	6.1E-8	-3.0	0.029	0.28
9	rs657152	$Insulin_{30}$	136139265	Intron	*ABO*	A/C	0.34	1	7.2	0.014	3.2E-7	1.0	0.021	0.64
3	rs62254949	$Insulin_{30}$	67933644	Intron	*SUCLG2-AS1*	G/A	0.12	0.95	10.5	0.020	7.7E-7	0.9	0.032	0.78
3	rs115404340	IGI	67839179	intron	*SUCLG2-AS1*	G/A	0.010	0.44	49.5	0.081	6.0E-7	-5.0	0.10	0.62
12	rs117236472	IGI	28121063	intron	*PTHLH*	T/C	0.014	0.65	32.5	0.057	7.6E-7	-7.3	0.093	0.41
17	rs34958340	IGI	48709115	nc transcript	*LOC101927253*	T/C	0.013	0.53	37.2	0.064	9.6E-7	-5.1	0.075	0.49
5	rs147351017	IR_{bmiadj}	132373452	intergenic	*ZCCHC10/HSPA4*	G/A	0.015	0.56	-28.2	0.066	5.1E-7	4.3	0.12	0.68
9	rs657152	IR_{bmiadj}	136139265	intron	*ABO*	A/C	0.34	1	6.5	0.013	6.1E-7	1.3	0.020	0.45
18	rs112752032	IR_{bmiadj}	24428770	intron	*AQP4-AS1*	G/A	0.011	0.41	54.4	0.088	7.8E-7	14.3	0.15	0.41
17	rs141652399	IR_{bmiadj}	78397039	intron	*ENDOV*	T/C	0.019	0.55	34.8	0.061	8.0E-7	-4.0	0.088	0.69

Threshold for the suggestive signals is 1×10^{-6}.
*A successful replication with p-value< 0.05.

[a] In the discovery cohort.
[b] Beta coefficient expressed as the percentage difference in the outcome by one copy of the effect allele.
Chr, chromosome; SNP, single nucleotide polymorphism; SE, standard error; nc transcript, transcript variant of a non-coding RNA gene.

SUPPLEMENTARY DATA

Supplementary Table S2. Phewas on rs505922 in the UK Biobank, with all the genome-wide significant phenotypes reported.

Phenotype	Beta	P-values
venous thromboembolic disease	-0.007	6.95E-86
deep venous thrombosis (dvt)	-0.006	6.27E-72
Haematocrit percentage	0.108	3.59E-68
Haemoglobin concentration	0.036	2.88E-66
I80 Phlebitis and thrombophlebitis	-0.004	4.54E-55
Red blood cell (erythrocyte) count	0.011	2.25E-48
Monocyte count	0.006	2.62E-40
pulmonary embolism +/- dvt	-0.003	9.56E-39
I26 Pulmonary embolism	-0.003	1.33E-37
I26-I28 Pulmonary heart disease and diseases of pulmonary circulation	-0.003	9.96E-33
White blood cell (leukocyte) count	0.043	2.75E-21
Impedance of arm (left)	-0.774	6.80E-21
Neutrophill count	0.028	5.79E-20
Impedance of arm (right)	-0.726	2.59E-19
Monocyte percentage	0.050	5.28E-18
Mean reticulocyte volume	-0.130	5.60E-16
high cholesterol	-0.006	2.58E-15
I84 Haemorrhoids	0.004	4.72E-14
Arm fat percentage (left)	-0.118	1.25E-13
Platelet distribution width	0.008	2.17E-13
Arm fat percentage (right)	-0.116	2.77E-13
Mean sphered cell volume	-0.076	8.28E-13
K57 Diverticular disease of intestine	0.004	1.76E-11
Trunk fat percentage	-0.102	2.66E-11
Trunk fat-free mass	0.039	2.80E-09
Trunk predicted mass	0.037	5.47E-09
Arm fat-free mass (left)	0.006	1.32E-08
Arm predicted mass (left)	0.005	1.76E-08

SUPPLEMENTARY DATA

Eosinophill count	0.002	1.94E-08
K55-K64 Other diseases of intestines	0.005	3.39E-08
M79 Other soft tissue disorders, not elsewhere classified	-0.002	3.60E-08
Arm predicted mass (right)	0.005	3.71E-08
Arm fat-free mass (right)	0.005	4.05E-08
Body fat percentage	-0.073	4.07E-08
Lymphocyte percentage	-0.087	4.38E-08

SUPPLEMENTARY DATA

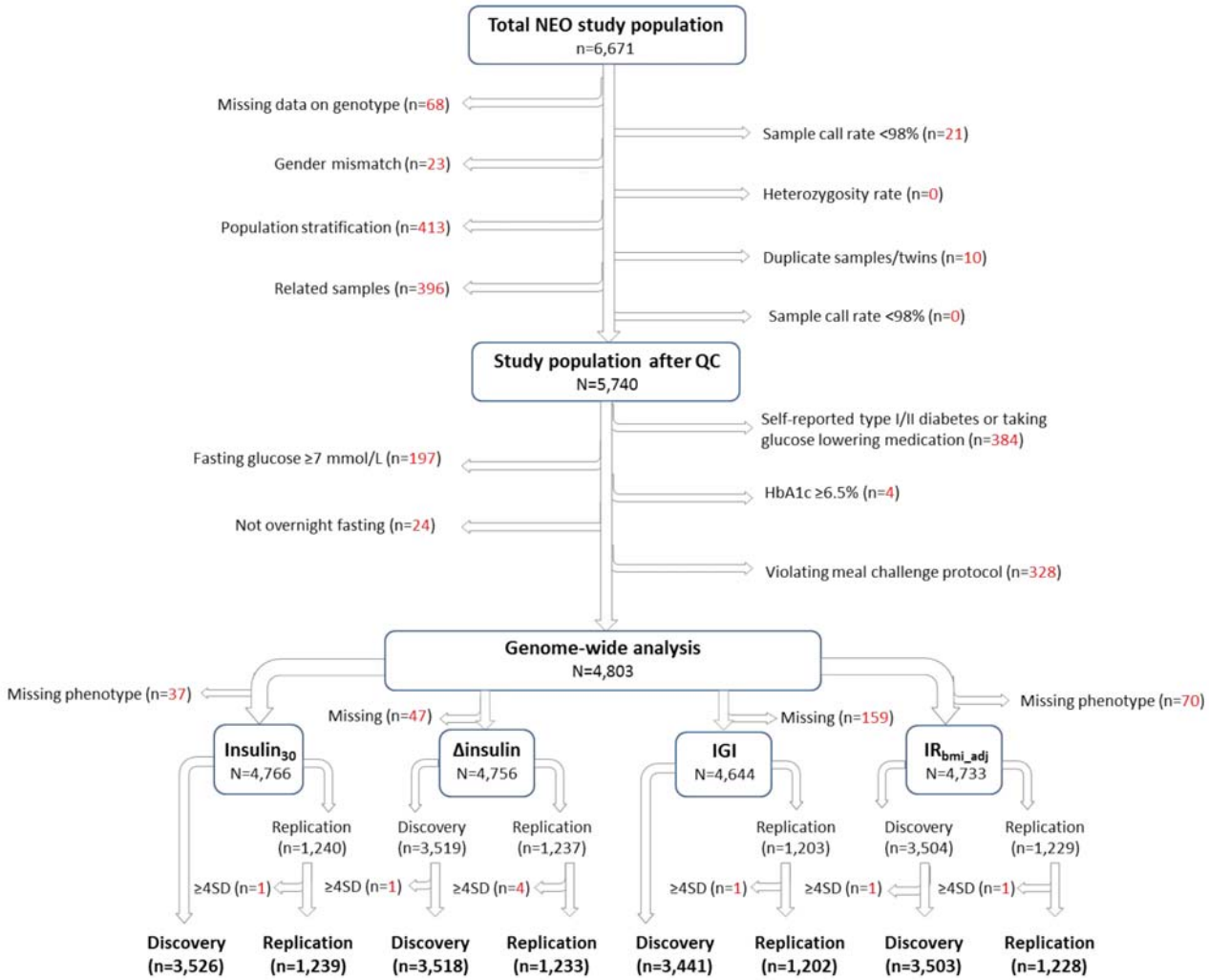
Supplementary Table S3. Genome-wide significant SNPs on Δinsulin in the entire NEO cohort (N=4,751).

Chr	SNP	Position	Location	Gene	Effect/ Non-effect allele	Effect allele frequency[a]	Imputation quality	Discovery cohort (N=3,518)			Entire NEO cohort (N=4,751)		
								Effect size per allele[b] (%)	SE	P-value	Effect size per allele[b] (%)	SE	P-value
9	rs676996*	136146077	intron	*ABO*	G/T	0.32	0.98	-6.4	0.012	6.2E-8	-6.9	0.011	1.3E-11
19	rs2287019	46202172	intron	*QPCTL*	T/C	0.19	1	-5.6	0.015	1.4E-4	-6.9	0.013	3.3E-8

Threshold for genome-wide significance is 5×10^{-8}.
* The LD between rs676996 and rs505922 is 0.97 in the entire NEO population.
[a] Coding allele frequency in the entire NEO cohort.
[b] Beta coefficient expressed as the percentage difference in the outcome by one copy of the effect allele.
Chr, chromosome; SNP, single nucleotide polymorphism; SE, standard error.

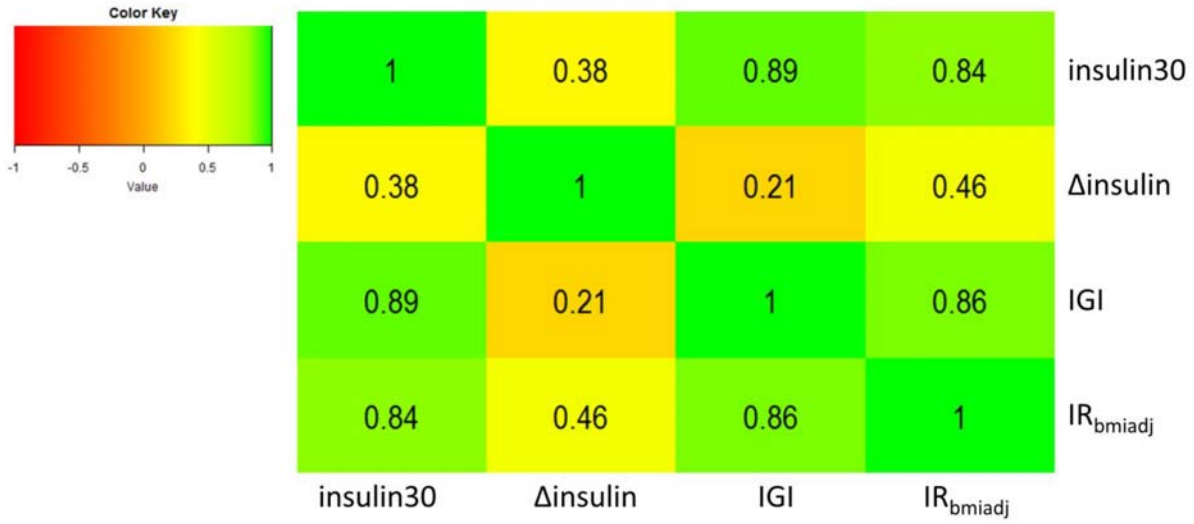
SUPPLEMENTARY DATA

Supplementary Figure S1. Quality control steps for the genome-wide association analyses with four different measures of insulin response and exclusion criteria for the discovery and replication cohort.



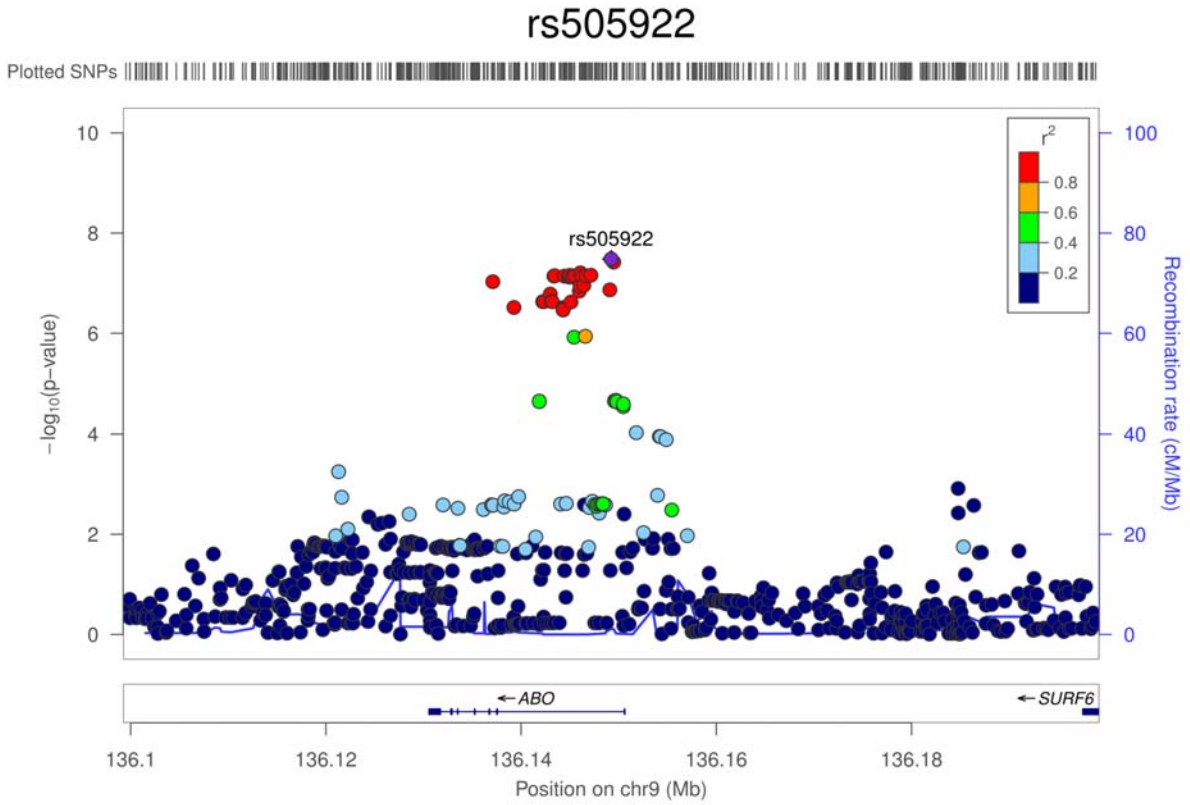
SUPPLEMENTARY DATA

Supplementary Figure S2. Phenotypic correlation matrix of four different measures of the early phase insulin response.



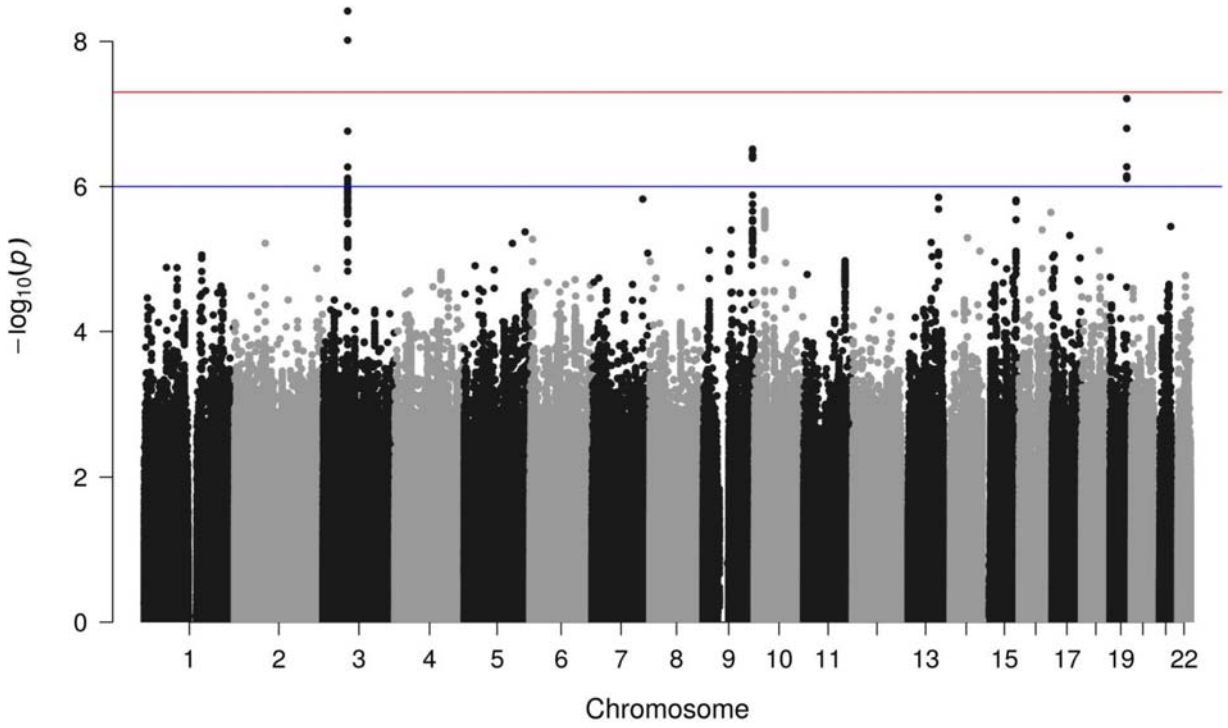
SUPPLEMENTARY DATA

Supplementary Figure S3. Regional association plot for the independent genome-wide significant SNP identified from Δinsulin, i.e. rs505922. The purple diamond indicates the lead SNP for the locus.



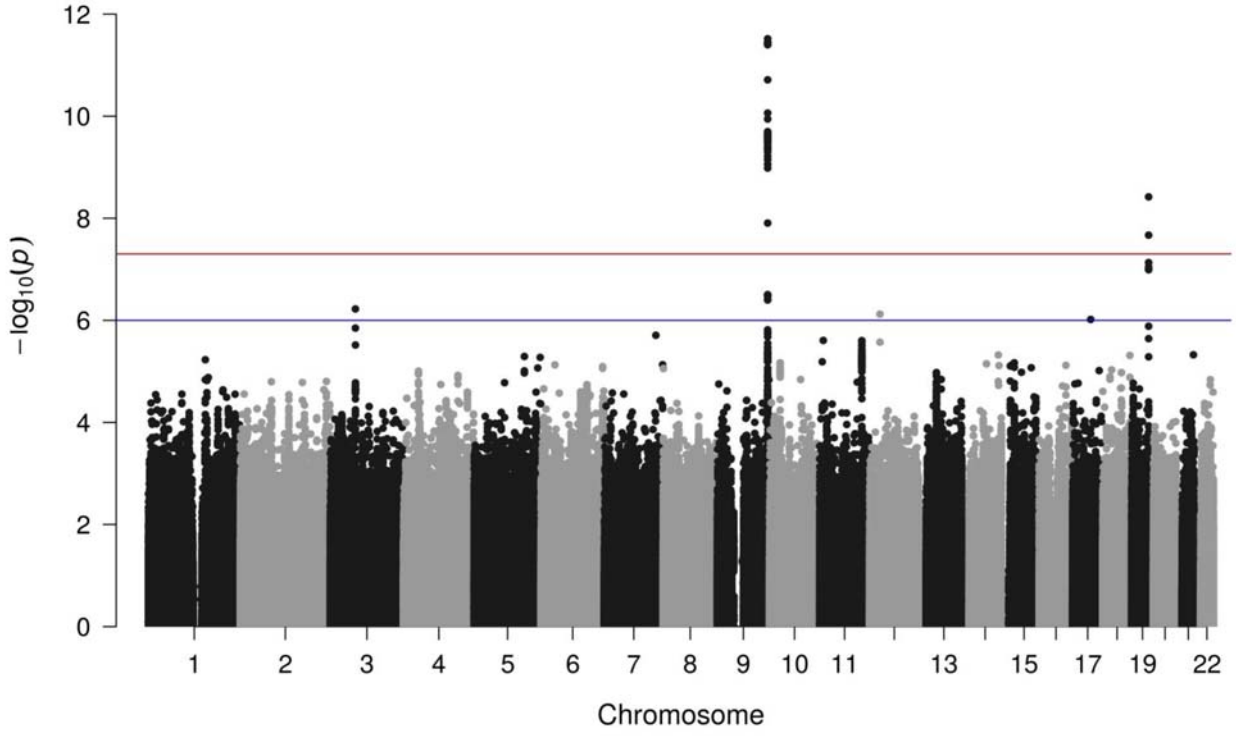
SUPPLEMENTARY DATA

Supplementary Figure S4. Manhattan plot for the genome-wide association study on insulin measures at 30 minutes ($insulin_{30}$) in the discovery cohort (n=3,526).



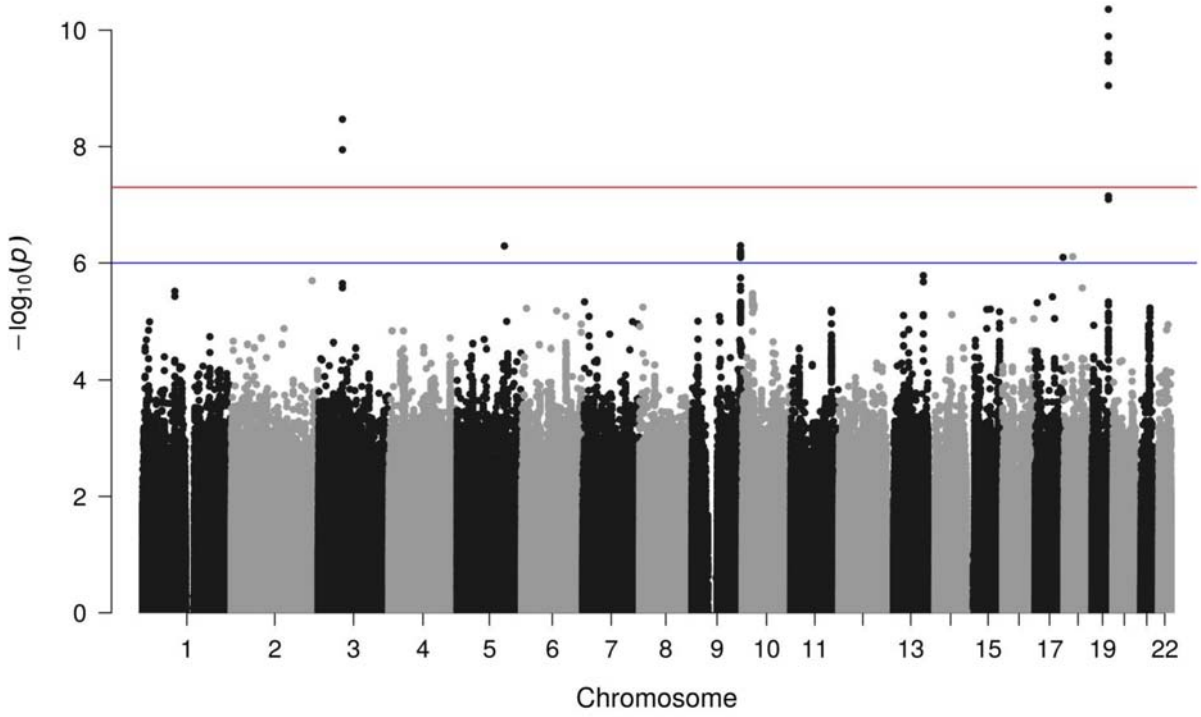
SUPPLEMENTARY DATA

Supplementary Figure S5. Manhattan plot for the genome-wide association study on insulinogenic index (IGI) in the discovery cohort (n=3,441).



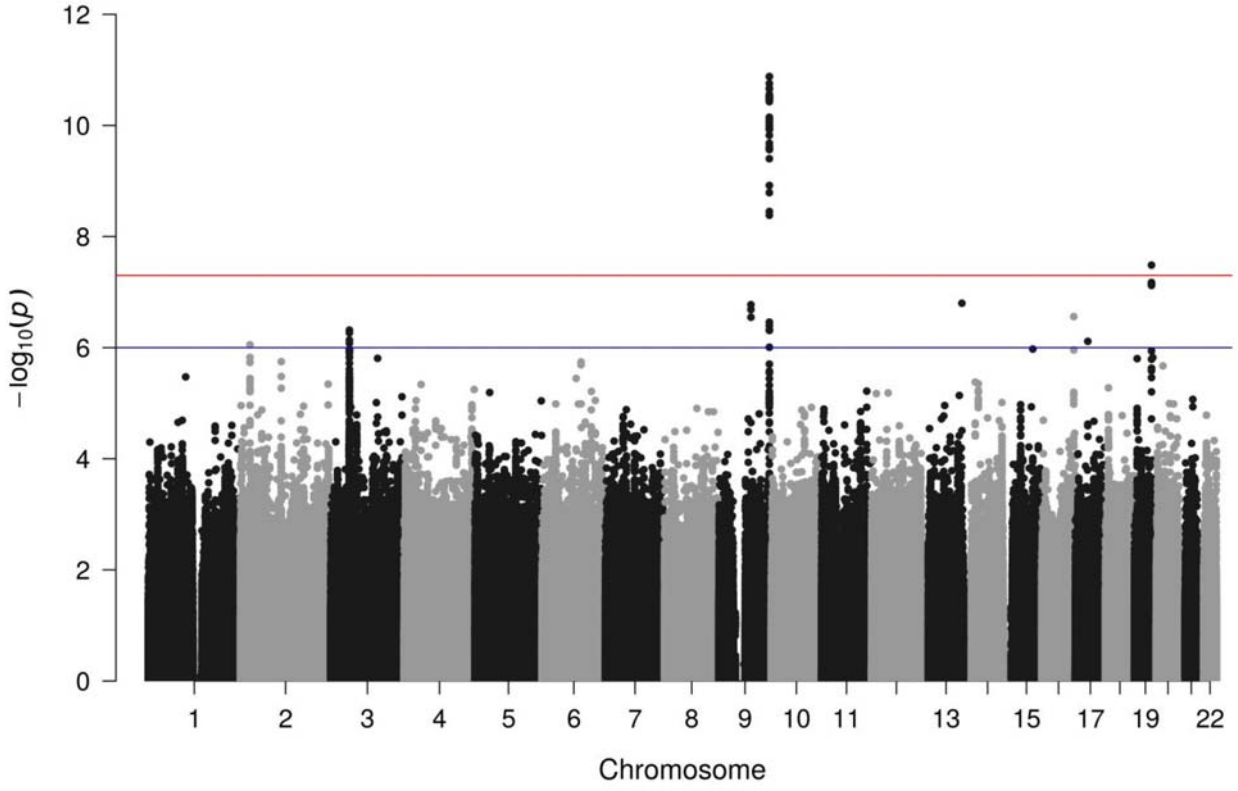
SUPPLEMENTARY DATA

Supplementary Figure S6. Manhattan plot for the genome-wide association study on insulin response to glucose at 30 minutes, adjusted on BMI (IR_{bmiadj}) in the discovery cohort (n=3,503).



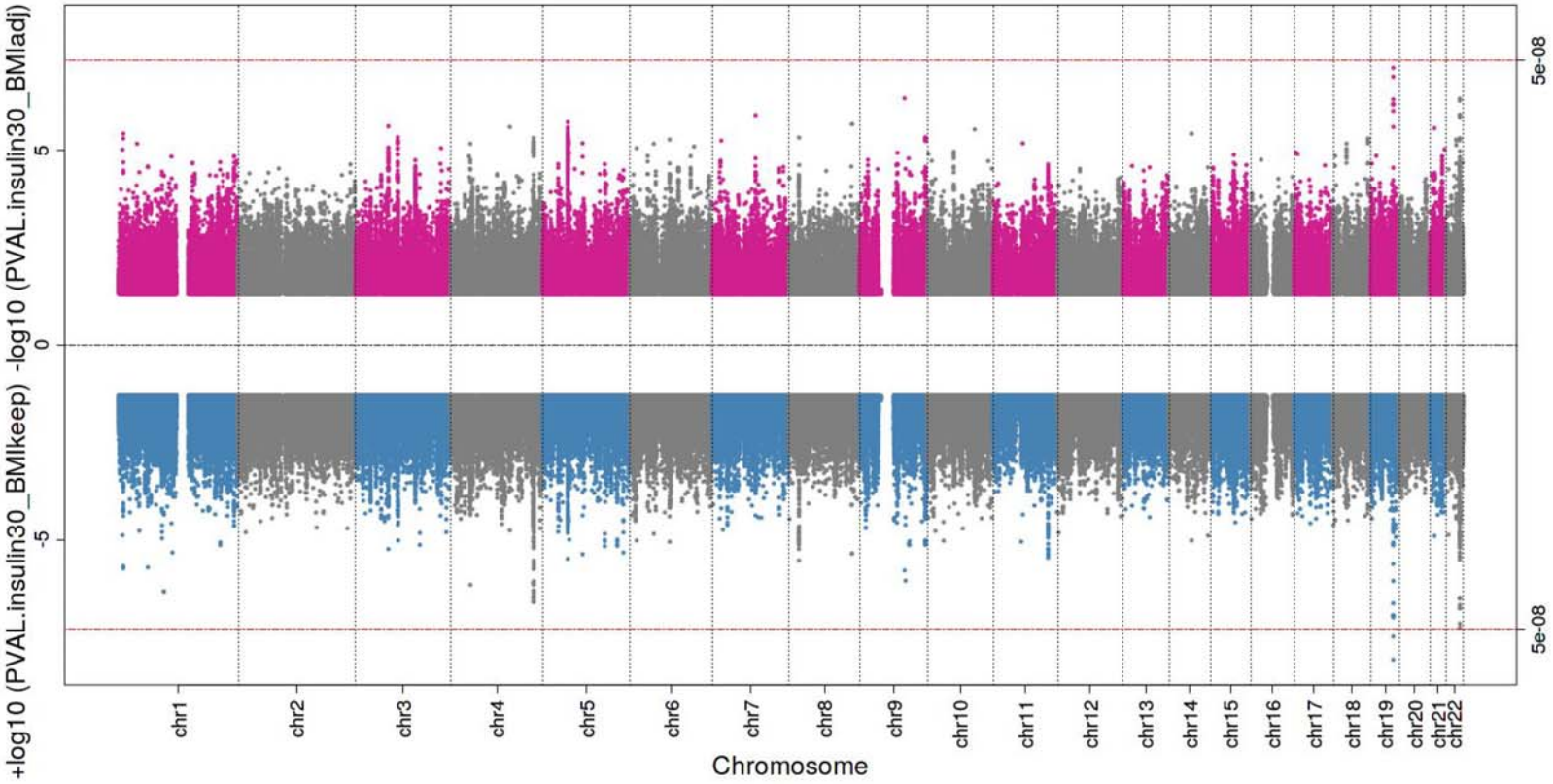
SUPPLEMENTARY DATA

Supplementary Figure S7. Manhattan plot for the genome-wide association study on Δinsulin (Δinsulin=ln(insulin$_{30}$)- ln(insulin$_0$)) in the entire NEO cohort (n=4,751).



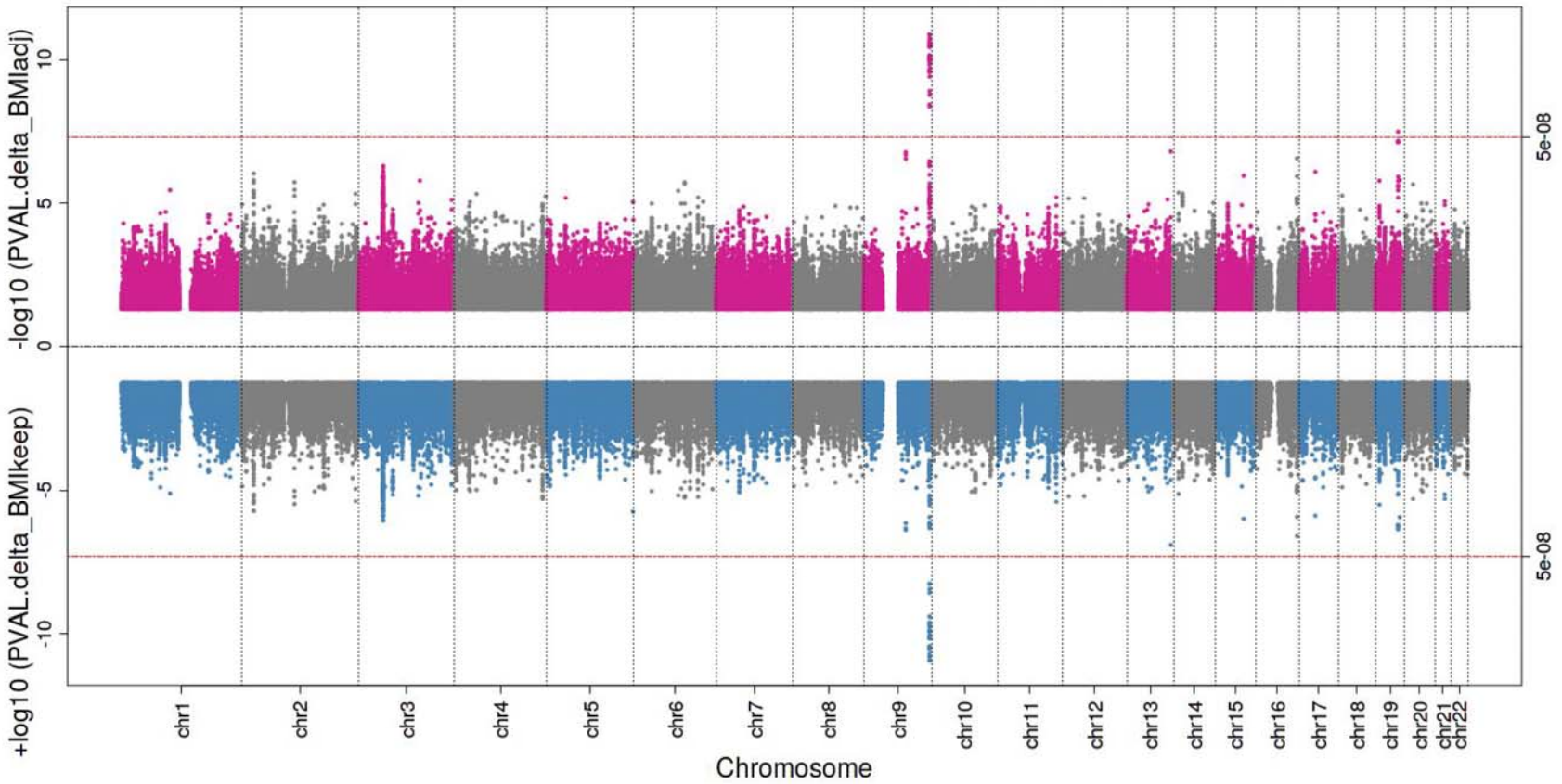
SUPPLEMENTARY DATA

Supplementary Figure S8. Miami plot for the genome-wide association study on insulin measures at 30 minutes ($insulin_{30}$) in the entire NEO cohort (n=4,765), with (top panel) /without (bottom panel) adjustment on BMI.



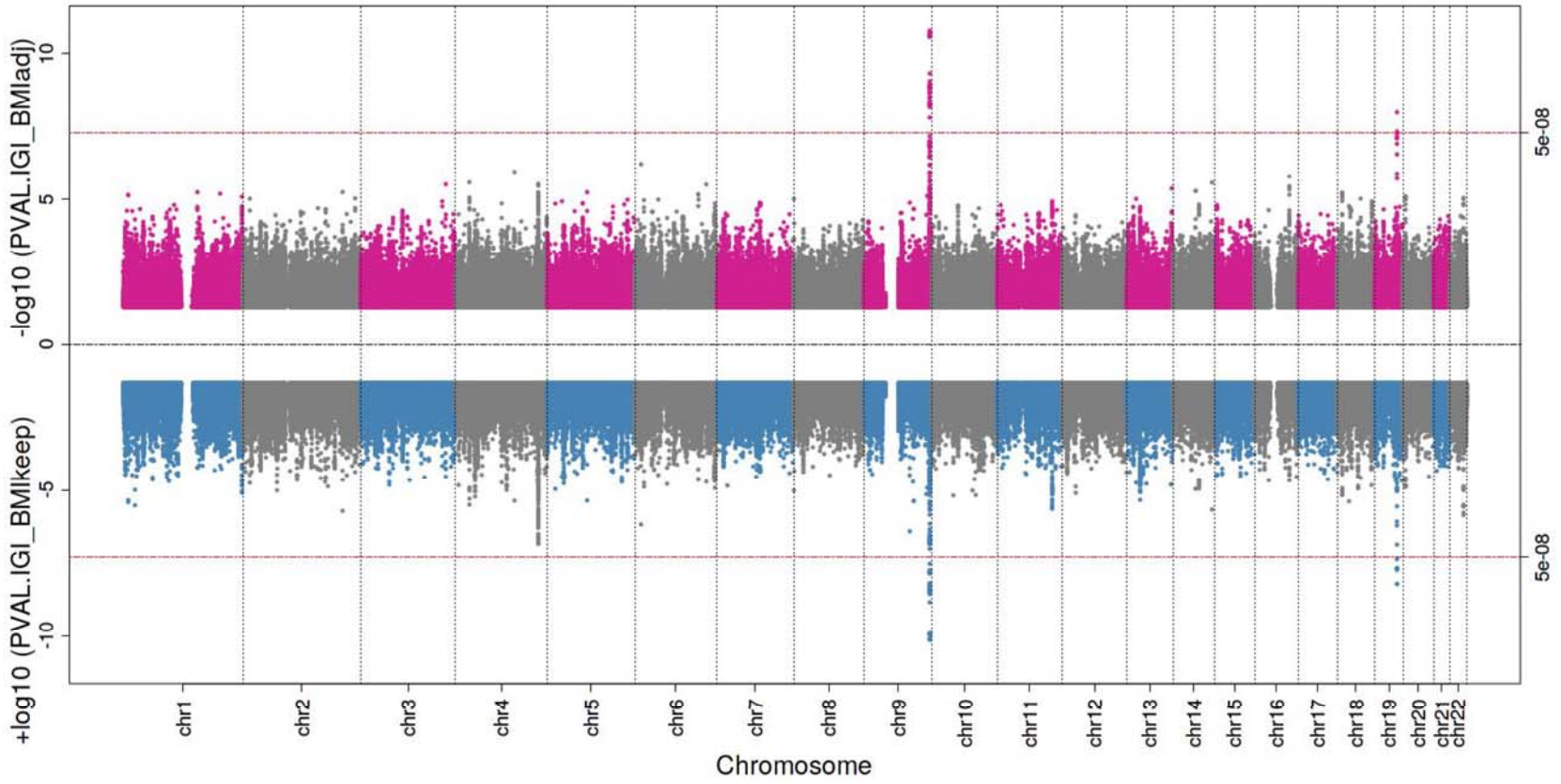
SUPPLEMENTARY DATA

Supplementary Figure S9. Miami plot for the genome-wide association study on Δinsulin (Δinsulin=ln(insulin$_{30}$)- ln(insulin$_0$)) in the entire NEO cohort (n=4,751), with (top panel) /without (bottom panel) adjustment on BMI.



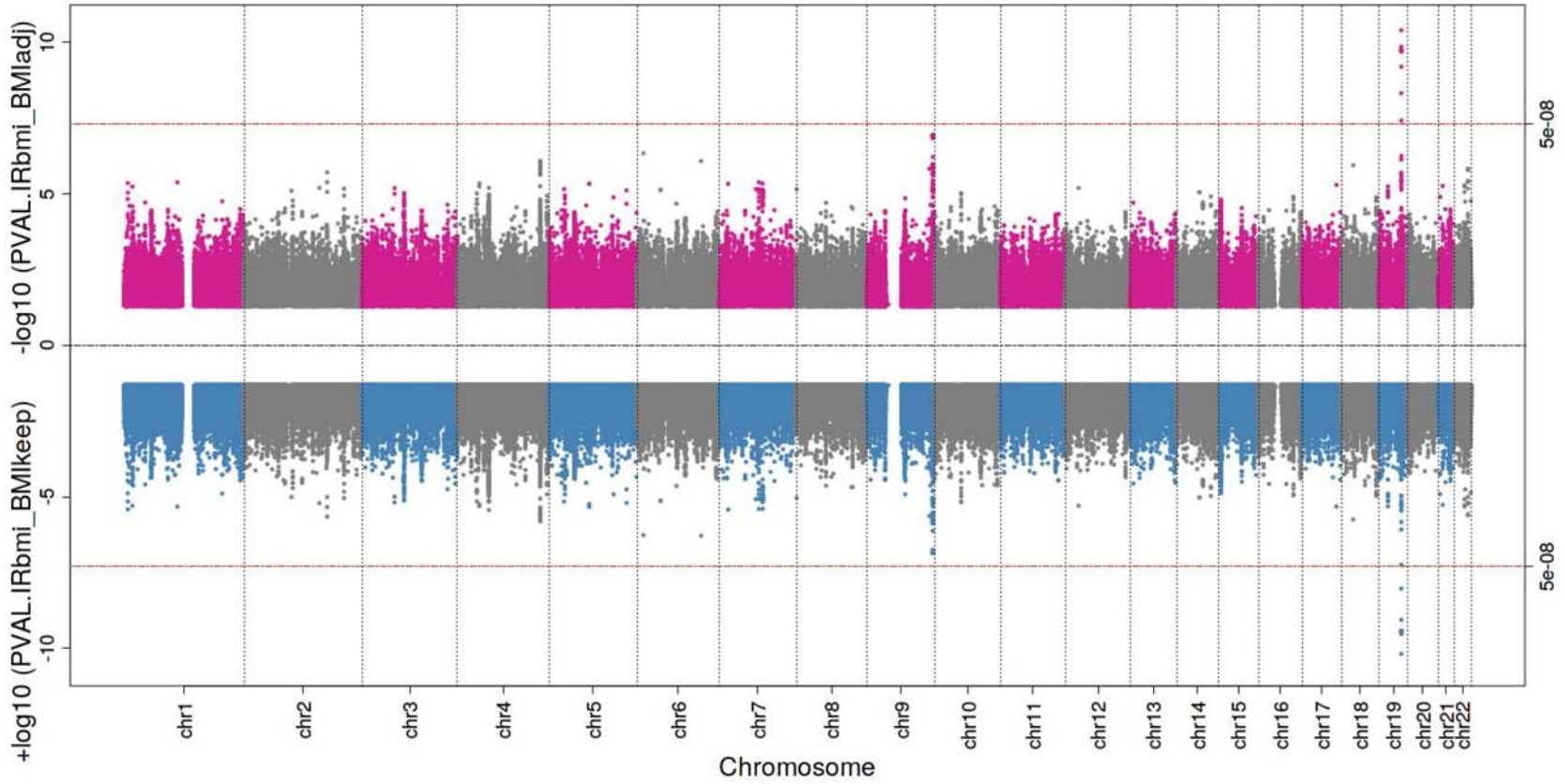
SUPPLEMENTARY DATA

Supplementary Figure S10. Miami plot for the genome-wide association study on insulinogenic index (IGI) in the entire NEO cohort (n=4,642), with (top panel) /without (bottom panel) adjustment on BMI.



SUPPLEMENTARY DATA

Supplementary Figure S11. Miami plot for the genome-wide association study on insulin response to glucose at 30 minutes, adjusted on BMI (IR_{bmiadj}) in the entire NEO cohort (n=4,731), with (top panel) /without (bottom panel) adjustment on BMI.



SUPPLEMENTARY DATA

Supplementary Figure S12. Modelling the effects of ABO gene expression changes on insulin secretion in vitro: (a) *ABO* gene expression is reduced by 60% in MIN6 cells transduced with the shRNA ABO lentivirus as compared with non-target shRNA ctrl. (b) Glucose-stimulated insulin secretion (GSIS) tests performed on MIN6 cells transduced with shRNA ABO or shRNA ctrl lentivirus.