Supplemental Material:

TXNIP DNA methylation is Associated with Glycemic Control Over 28 Years in Type 1 Diabetes: Findings from the Pittsburgh Epidemiology of Diabetes Complications (EDC) Study

Rachel G. Miller[1], Josyf C. Mychaleckyj[2], Suna Onengut-Gumuscu[2], Trevor J. Orchard[1], Tina Costacou[1]

[1]Department of Epidemiology, University of Pittsburgh, Pittsburgh, PA

[2]Center for Public Health Genomics, University of Virginia, Charlottesville, VA

Detailed DNA Methylation Quality Control and Data Processing Information

The EPIC output Intensity Data (IDAT) file contained raw probe intensity data, as well as the control probe information needed for quality control. The IDAT files were loaded and converted into an RGSet object and thence to a GMSet object based on the Illumina 1.0 B5 hg38 annotation release, using minfi. The minfiQC function was used to check for outlier samples of potentially poorer quality based on: 1) median intensity of methylated (M) versus unmethylated (U) probes; 2) median X, Y chromosome probe signals for karyotype prediction; 3) sample probe detection rate at p-value < 0.01; 4) repeated sample probe detection rate after dropping low quality probes (probe detection rate < 95%) and a curated list of probes recommended for exclusion by Zhou et al. (1). A second pipeline of quality control checks used functions in the SeSAMe package, after conversion of the RGset object to SigSet objects, to confirm and expand the minfi sample methylation data checks. The second pipeline checked: 1) mean sample probe intensity distribution by sample, mean(M+U); 2) sample bisulfite GCT score for completeness of bisulfite conversion; 3) inferred sex karyotype based on chromosome X and Y probes; 4) detection of duplicate samples using SNP genotype data extracted from a subset of the probes. Probe quality checks were also performed based on individual probe detection p-values. Any probe that was not detected in at least 95% of samples was dropped along with an exclusion set of probes previously identified to have a higher probability of generating artifactual methylation signals due to underlying genome sequence composition issues, cross-reactivity, or the likely presence of flanking SNP variation in predominantly European ancestry populations (1). Cryptic sample duplicates were inferred using KING (2) applied to the SNP genotype data extracted from the Illumina 'spiked-in' SNP probes (n=44) and additional probes that manifest a reproducible pattern of genotype-stratified methylation signals driven by a flanking SNP (n=865 SNPs in total).

References:

1. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. Nucleic Acids Res. 2017 Feb 28;45(4):e22.
2. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W. Robust relationship inference in genome-wide association studies. Bioinformatics. 2010;26(22):2867–73.

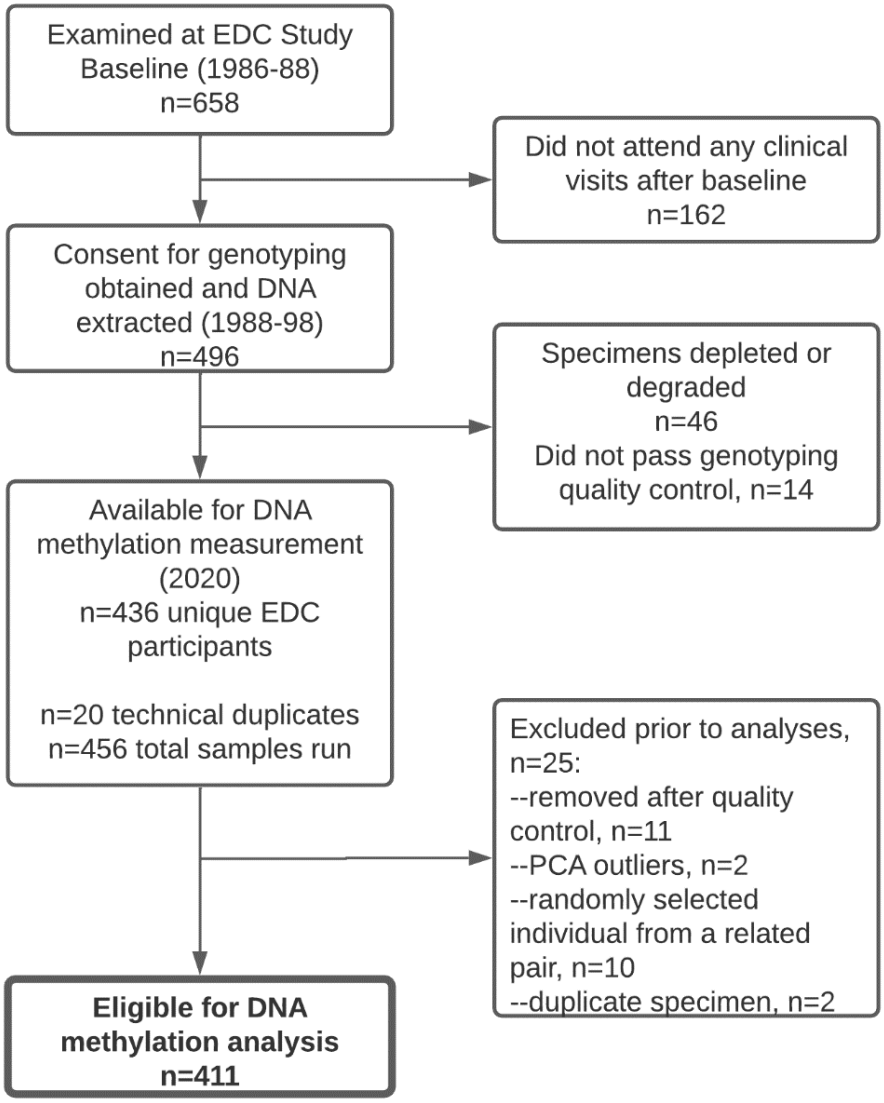
Ascertainment of Clinical Variables

Data on clinical variables were taken from the study visit of DNA collection. Fasting blood samples were obtained to measure HbA1, lipids, and serum creatinine. HbA1 values were converted to DCCT-aligned values using a regression equation derived from duplicate assays (DCCT HbA1c=0.14 + 0.83[EDC HbA1]) (1). Total cholesterol and triglycerides were determined enzymatically (2,3). HDL cholesterol was determined using a modified precipitation technique (4). Non-HDL cholesterol was calculated by subtracting HDL cholesterol from total cholesterol. Blood pressure was measured according to the Hypertension Detection and Follow-Up protocol with a random-zero sphygmomanometer (5). Pulse rate (beats/minute) was determined by palpitating the radial pulse for 30 seconds and multiplying by two. Urinary albumin was measured by immunonephelometry (6). Albumin excretion rate (AER) was calculated for each of three timed urine samples (24-hour, overnight, and 4-hour collections obtained over a two-week period); the median of the three AERs was used in analyses. Serum creatinine was measured using an Ectachem 400 Analyzer (Eastman Kodak Co.) and glomerular filtration rate was estimated by the CKD-EPI creatinine equation (7). Height and weight were measured using standard methods to calculate BMI. Smoking history, insulin regimen, lipid-lowering and blood pressure-lowering medication use, history of hypoglycemia requiring assistance, and history of diabetic ketoacidosis were self-reported via questionnaire. Insulin dose was calculated as insulin units per day divided by body weight (kg).

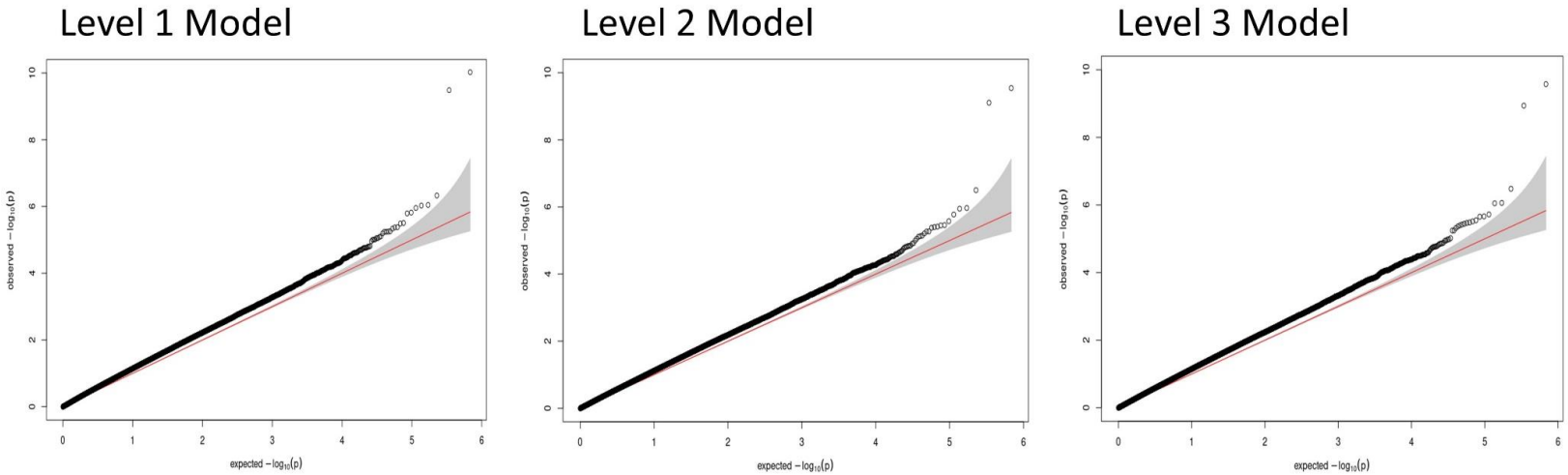
References:

1. Prince CT, Becker DJ, Costacou T, Miller RG, Orchard TJ. Changes in glycaemic control and risk of coronary artery disease in type 1 diabetes mellitus: findings from the Pittsburgh Epidemiology of Diabetes Complications Study (EDC). Diabetologia. 2007 Nov;50(11):2280–8.
2. Bucolo G, David H. Quantitative determination of serum triglycerides by the use of enzymes. Clin Chem. 1973;19(5):476–82.
3. Allain CC, Poon LS, Chan CSG, Richmond W, Fu PC. Enzymatic determination of total serum cholesterol. Lipids. 1974;20(4):470–5.
4. Warnick GR, Albers JJ. Heparin--Mn2+ quantitation of high-density-lipoprotein cholesterol: an ultrafiltration procedure for lipemic samples. Clin Chem. 1978;24(6):900–4.
5. The hypertension detection and follow-up program. Hypertension detetection and follow-up program cooperative group. Prev Med. 1976;5:207–15.

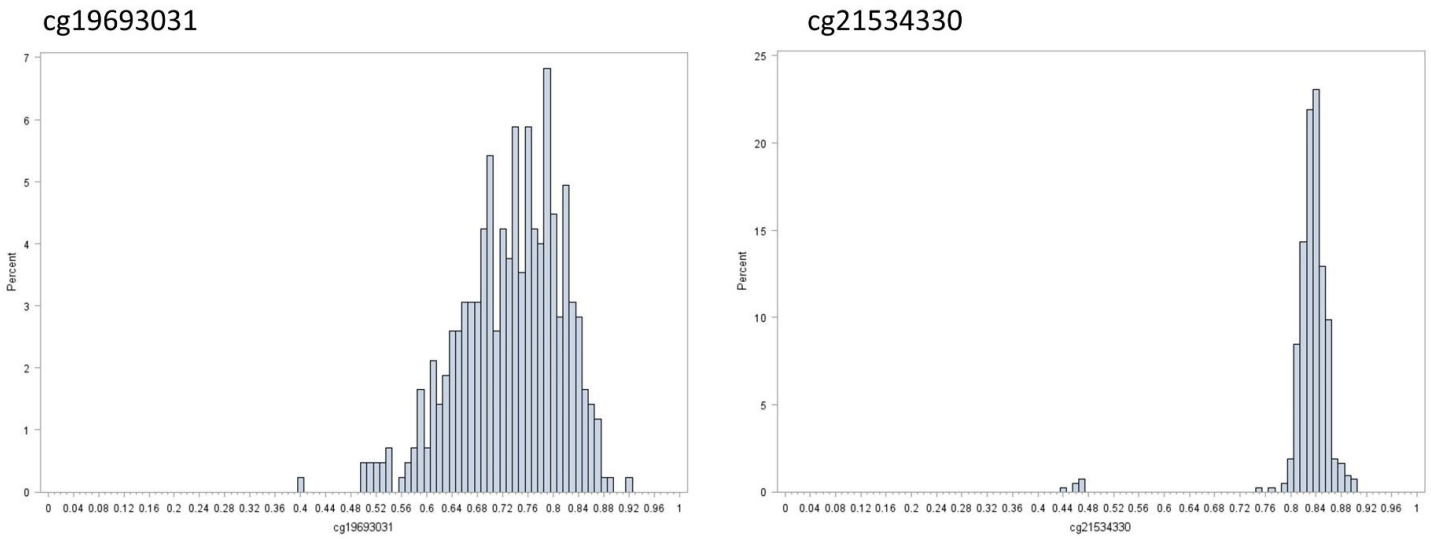
6. Ellis D, Coonrod BA, Dorman JS, Kelsey SF, Becker DJ, Avner ED, et al. Choice of urine sample predictive of microalbuminuria in patients with insulin-dependent diabetes mellitus. Am J kidney Dis Off J Natl Kidney Found. 1989;13(4):321–8.
7. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150(9):604–12.



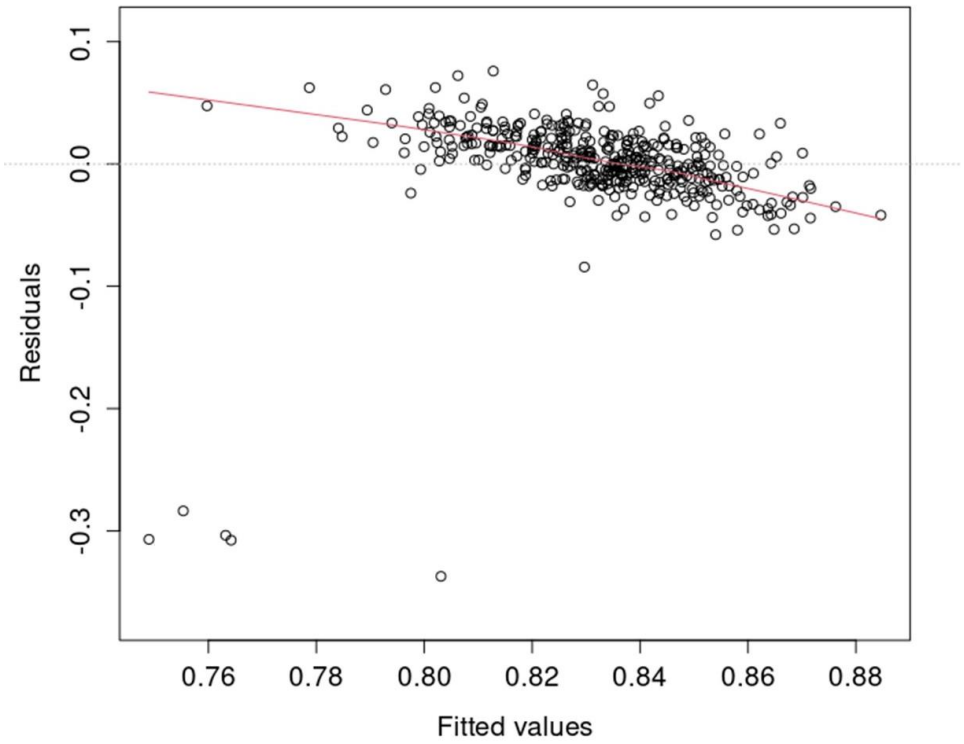
Online Supplemental Figure 1. Flow diagram of the analytic sample for epigenome-wide associations between DNA methylation and baseline HbA1c in the EDC study



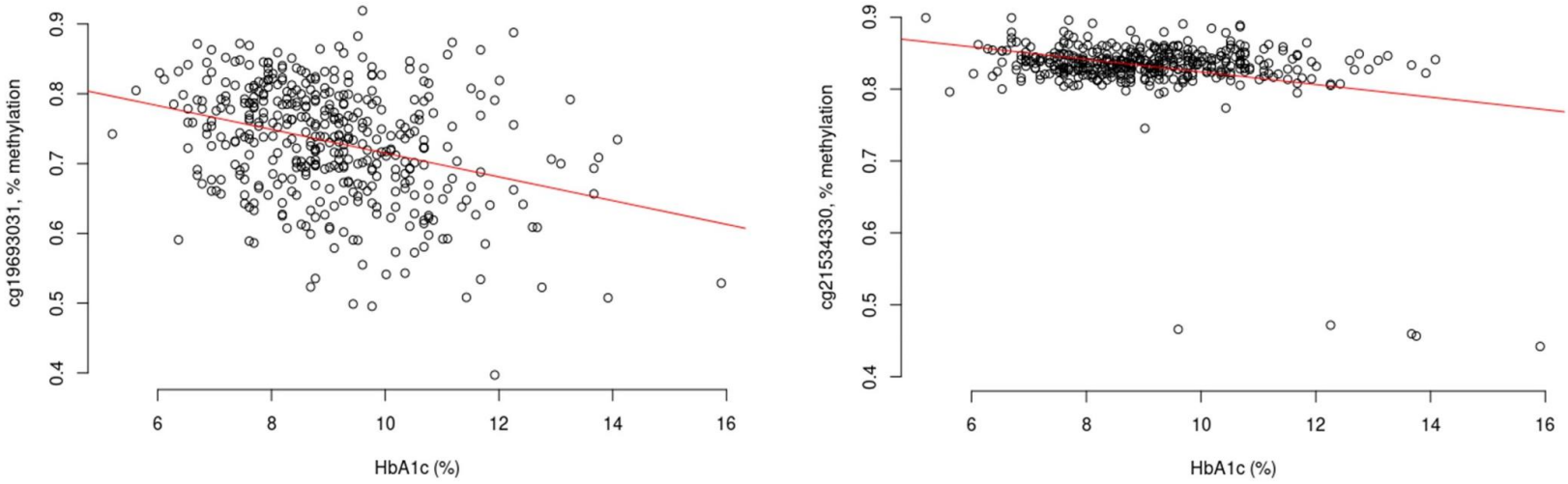
Online Supplemental Figure 2. QQ plots for epigenome-wide association of DNA methylation with baseline HbA1c in the EDC cohort by the three models



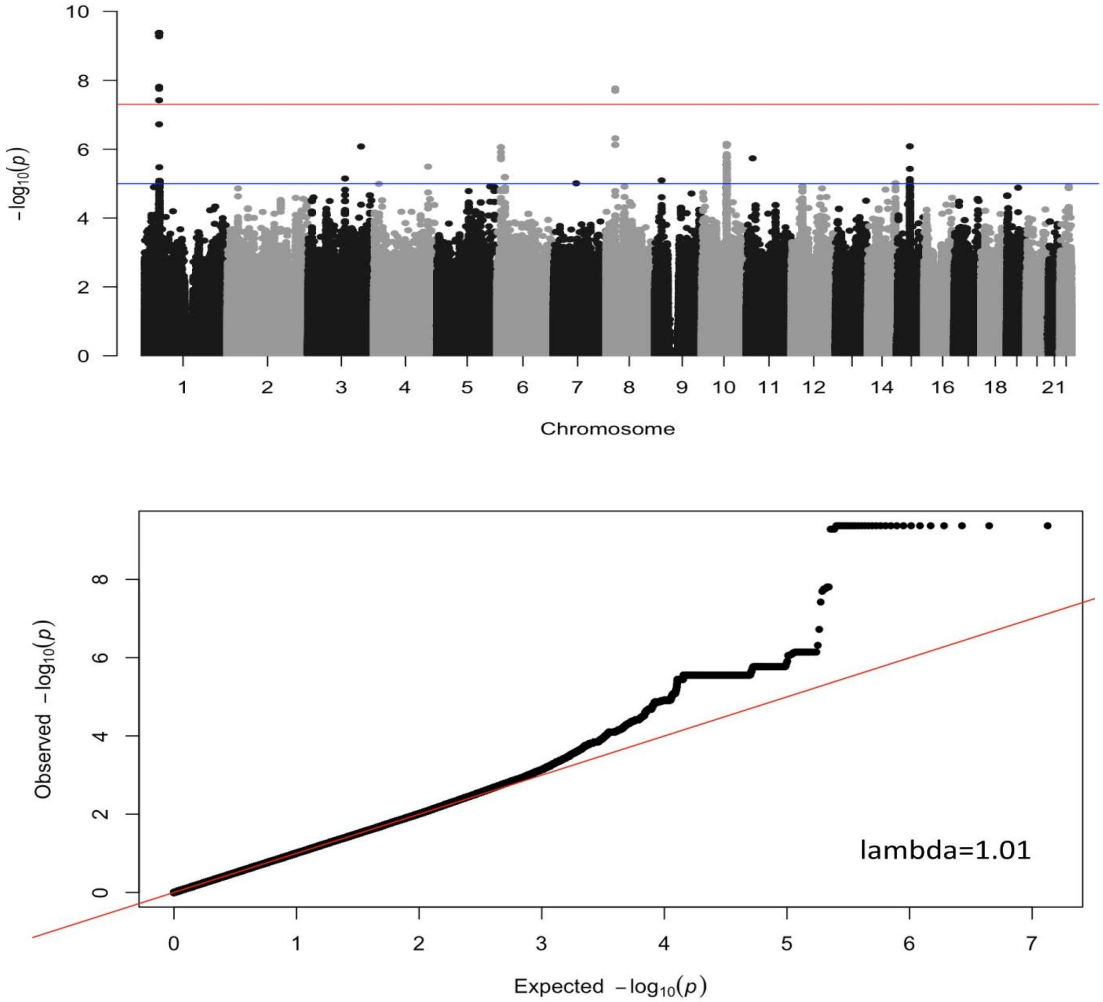
Online Supplemental Figure 3. Distribution of probe β values for CpGs significantly associated with HbA1c in the EDC study



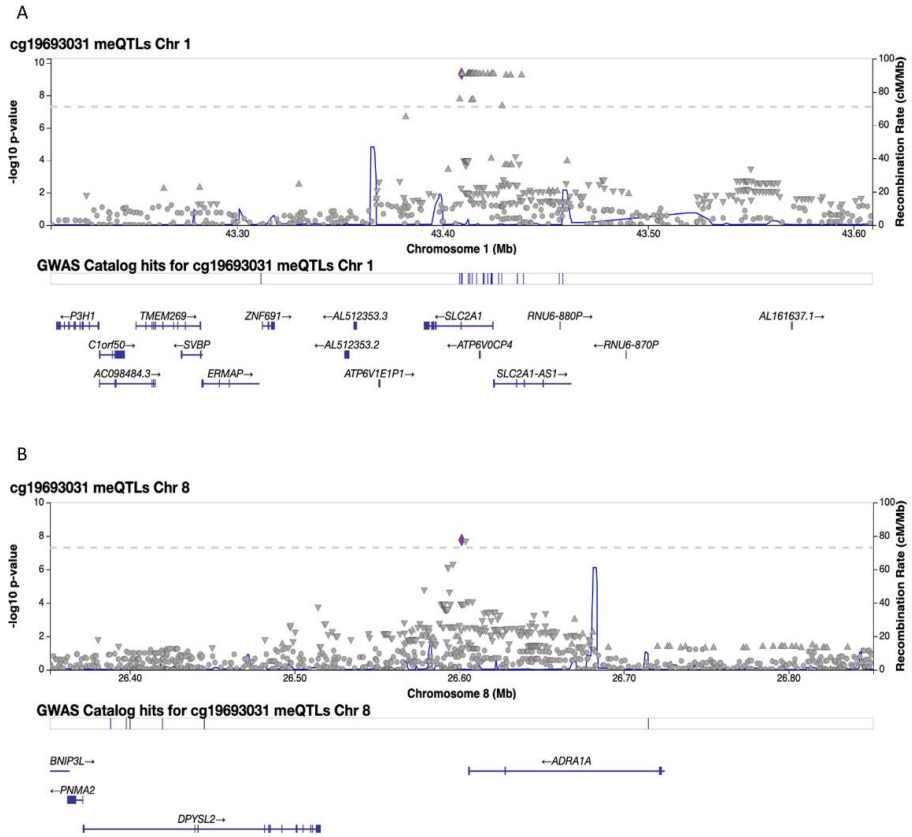
Online Supplemental Figure 4. Plot of residuals versus fitted values from the linear model for cg21534330 and HbA1c



Online Supplemental Figure 5. Scatter plots of methylation by HbA1c for significant CpGs



Online Supplemental Figure 6. Manhattan and QQ plots for SNP genome-wide associations with methylation at cg19693031 in the EDC cohort



Online Supplemental Figure 7. Locus zoom plots for meQTLs associated with cg19693031 methylation in the EDC cohort.

Online Supplemental Table 1. Baseline characteristics of the EDC study DNA methylation sub-cohort (n=411)

	Mean (SD)
Age, years	29.7 (7.9)
T1D Duration, years	21.4 (7.7)
Age at T1D Onset, years	8.3 (4.1)
Female Sex, % (n)	47.0% (193)
≥Bachelor's Degree, % (n)	34.3% (141)
HbA1c, %	9.1 (1.6)
HbA1c, mmol/mol	75.6 (17.0)
Number of Longitudinal HbA1c Measures*	5 (1-8)
Smoking Status, % (n)	
Never	64.2% (263)
Past	14.8% (61)
Current	21.3% (87)
Smoking, pack-years†	0 (0, 3.3)
Body Mass Index, kg/m^2	24.1 (3.3)
Insulin Dose, insulin units/kg body weight	0.78 (0.25)
MDI‡ or Insulin Pump Use, % (n)	14.2% (58)
Self-monitoring of blood glucose, % (n)	72.3% (295)
History of Hypoglycemia Requiring Assistance, % (n)	44.0% (180)
Estimated Glucose Disposal Rate, mg/kg/min	7.4 (2.1)
Total Cholesterol (mg/dl)	188.8 (41.9)
HDLc (mg/dl)	53.2 (12.7)
Non-HDLc (mg/dl)	135.6 (40.8)
Triglycerides (mg/dl)‡	82.0 (57.0, 123.5)
Lipid Medication, % (n)	0.49% (2)
Systolic Blood Pressure (mmHg)	115.3 (18.0)
Diastolic Blood Pressure (mmHg)	72.9 (11.0)
Hypertension, % (n)	21.4% (88)
Blood Pressure Medication, % (n)	13.6% (56)
Pulse Rate, bpm	75.9 (10.9)
Albumin Excretion Rate, µg/min†	12.2 (6.4, 105.7)
Estimated Glomerular Filtration Rate, ml/min/1.73m^2	108.9 (35.1)
White Blood Cell Count, x 10^9 cells/l	7.0 (2.2)

*median (range), †median (p25, p75), ‡Multiple daily injection therapy (≥3 insulin injections per day)

Online Supplemental Table 2. Variants associated with cg19693031 methylation (meQTLs) at genome-wide significance in the EDC cohort

CHR	Accession Num	Gene Symbol	Variant Type	SNP	BP	A1	BETA	SE	L95	U95	P	meQTL in GoDMC[a]	eQTL in GTEx[b]
1	rs5773795	SLC2A1	Indel	1:43412360	43412360	CCAAT	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	no
1	rs200895692	SLC2A1	Indel	1:43421956	43421956	C	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs3216016	SLC2A1	Indel	1:43418056	43418056	A	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs3841834	SLC2A1	Indel	1:43413384	43413384	TA	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs893766801	SLC2A1	Intron SNV	1:43414738	43414738	C	0.03911	0.006799	0.02578	0.05243	1.76E-08	yes	no
1	rs1108902	SLC2A1	Intron SNV	1:43413212	43413212	A	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs12718444	SLC2A1	Intron SNV	1:43409179	43409179	T	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs17387733	SLC2A1	Intron SNV	1:43414285	43414285	C	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs17387775	SLC2A1	Intron SNV	1:43416133	43416133	C	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs17387886	SLC2A1	Intron SNV	1:43419705	43419705	C	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs2297976	SLC2A1	Intron SNV	1:43415465	43415465	T	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs34964576	SLC2A1	Intron SNV	1:43409364	43409364	C	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs35022307	SLC2A1	Intron SNV	1:43409420	43409420	T	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs3768036	SLC2A1	Intron SNV	1:43412655	43412655	A	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs3768037	SLC2A1	Intron SNV	1:43412662	43412662	A	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs3768042	SLC2A1	Intron SNV	1:43413735	43413735	G	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs57247989	SLC2A1	Intron SNV	1:43414370	43414370	G	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs61296119	SLC2A1	Intron SNV	1:43414046	43414046	C	0.03882	0.006762	0.02556	0.05207	1.87E-08	yes	yes
1	rs6657798	SLC2A1	Intron SNV	1:43408279	43408279	G	0.03935	0.006829	0.02597	0.05274	1.66E-08	yes	yes
1	rs71654266	SLC2A1	Intron SNV	1:43414656	43414656	G	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs74742820	SLC2A1	Intron SNV	1:43414369	43414369	C	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs7512557	SLC2A1	Intron SNV	1:43413324	43413324	T	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs7512565	SLC2A1	Intron SNV	1:43413361	43413361	T	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs7522674	SLC2A1	Intron SNV	1:43416781	43416781	T	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs7534555	SLC2A1	Intron SNV	1:43413319	43413319	C	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs75366795	SLC2A1	Intron SNV	1:43414447	43414447	G	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs900835	SLC2A1	Intron SNV	1:43412455	43412455	T	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs114530232	SLC2A1; SLC2A1-AS1	Intron SNV; 2KB Upstream SNV	1:43424051	43424051	G	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs3754226	SLC2A1; SLC2A1-	Non-coding transcript	1:43424812	43424812	G	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes

		AS1	SNV; 2KB Upstream SNV										
1	rs80200046	SLC2A1; SLC2A1-AS1	Intron SNV; 2KB Upstream SNV	1:43423932	43423932	A	0.04494	0.00703	0.03116	0.05872	4.58E-10	yes	yes
1	rs12750584	SLC2A1-AS1	Intron SNV	1:43430883	43430883	A	0.04477	0.007042	0.03097	0.05858	5.58E-10	yes	yes
1	rs58062906	SLC2A1-AS1	Intron SNV	1:43438399	43438399	T	0.04477	0.007042	0.03097	0.05858	5.58E-10	yes	yes
1	rs71654267	SLC2A1-AS1	Intron SNV	1:43433066	43433066	A	0.04477	0.007042	0.03097	0.05858	5.58E-10	yes	yes
1	rs77841501	SLC2A1-AS1	Intron SNV	1:43428964	43428964	G	0.04981	0.008922	0.03232	0.0673	4.39E-08	yes	yes
8	rs7843809	ADRA1A	3' UTR SNV	8:26601545	26601545	A	-0.03643	0.006331	-0.04883	-0.02402	1.74E-08	no	no
8	rs9314327	ADRA1A	3' UTR SNV	8:26603955	26603955	C	-0.03626	0.006333	-0.04868	-0.02385	2.01E-08	no	no

[a] GoDMC Database http://mqtldb.godmc.org.uk/search?query=cg19693031 (Accessed 25 February 2022)

[b] GTEx Portal https://gtexportal.org/ (Accessed 25 February 2022)

Online Supplemental Table 3. Variants associated with cg21534330 methylation (meQTLs) at genome-wide significance in the EDC cohort

CHR	Accession	Gene Symbol	Variant	SNP	BP	A1	BETA	SE	L95	U95	STAT	P	meQTL in	eQTL in

	Num		Type										GoDMC[a]	GTEx[b]
2	rs72819137	AFF3	Intron SNV	2:100500167	100500167	T	-0.04412	0.006786	-0.05742	-0.03082	-6.502	2.37E-10	no	no
7	rs9655138	ANKMY2	Intron SNV	7:16649835	16649835	T	-0.03627	0.005292	-0.04664	-0.0259	-6.854	2.74E-11	no	no
7	rs71540757	ANKMY2	Intron SNV	7:16653094	16653094	A	-0.03652	0.005337	-0.04698	-0.02606	-6.843	2.93E-11	no	no
7	rs11772957	ANKMY2	Intron SNV	7:16651317	16651317	C	-0.03043	0.004843	-0.03992	-0.02094	-6.284	8.64E-10	no	no
7	rs10282627	ANKMY2	Intron SNV	7:16652216	16652216	C	-0.02986	0.004795	-0.03926	-0.02046	-6.227	1.21E-09	no	no
7	rs2286230	ANKMY2	Intron SNV	7:16648885	16648885	C	-0.02998	0.004826	-0.03944	-0.02053	-6.214	1.31E-09	no	no
7	rs370078	ATP6V0A4	Intron SNV	7:138432523	138432523	T	-0.03357	0.005696	-0.04473	-0.0224	-5.893	8.06E-09	no	no
17	rs16943912	LOC105371536	Intron SNV	17:10818236	10818236	C	-0.03255	0.005776	-0.04387	-0.02123	-5.636	3.30E-08	no	no
20	rs147213238	n/a	Indel	20:50664946	50664946	GA	-0.04086	0.006068	-0.05275	-0.02897	-6.734	5.76E-11	no	no
20	rs73281721	n/a	SNV	20:50665214	50665214	T	-0.04086	0.006068	-0.05275	-0.02897	-6.734	5.76E-11	no	no
20	rs76950404	n/a	SNV	20:50670387	50670387	C	-0.04026	0.006026	-0.05207	-0.02845	-6.682	7.97E-11	no	no
20	rs80294592	n/a	SNV	20:50681096	50681096	C	-0.03877	0.005938	-0.0504	-0.02713	-6.528	2.03E-10	no	no
8	rs113203945	n/a	SNV	8:111571815	111571815	A	-0.02562	0.004351	-0.03415	-0.01709	-5.888	8.31E-09	no	no
1	rs72716780	n/a	SNV	1:88503824	88503824	G	-0.03534	0.006094	-0.04728	-0.0234	-5.799	1.36E-08	no	no
8	rs17677264	TNKS	2KB Upstream SNV	8:9412050	9412050	C	-0.03582	0.006328	-0.04822	-0.02342	-5.661	2.88E-08	no	no

[a] GoDMC Database http://mqtldb.godmc.org.uk/search?query=cg19693031 (Accessed 25 February 2022)

[b] GTEx Portal https://gtexportal.org/ (Accessed 25 February 2022)

Online Supplemental Table 4. Association between cg19693031 methylation and baseline clinical characteristics, independent of HbA1c

	Individual Models[*]	Final Model[†]

	β (standard error)	p-value	β (standard error)	p-value
Total Cholesterol	-0.00019 (0.0001)	0.052	*NR*	-
HDLc	-0.00004 (0.0003)	0.897	*NO*	-
Non-HDLc	-0.00019 (0.0001)	0.054	*NR*	-
ln(Triglycerides)	-0.0209 (0.0072)	0.004	*NR*	-
Systolic BP	-0.00004 (0.0002)	0.869	*NO*	-
Diastolic BP	-0.00040 (0.0004)	0.272	*NO*	-
Pulse rate	-0.00093 (0.0004)	0.009	-0.00093 (0.0004)	0.009
BMI	0.00033 (0.0012)	0.776	*NO*	-
Insulin Dose	-0.01128 (0.0164)	0.491	*NO*	-
Multiple daily injections or Insulin Pump	0.01438 (0.0110)	0.190	*NR*	-
Hypoglycemia Requiring Assistance	0.00994 (0.0077)	0.198	*NR*	-
Any History of Diabetic Ketoacidosis	-0.02409 (0.0156)	0.124	*NR*	-
ln(ACR)	-0.00362 (0.0019)	0.059	*NR*	-
eGFR	-0.00017 (0.0001)	0.187	*NR*	-

*Data are the change in methylation associated with a 1-unit increment in each clinical variable modeled separately, adjusting for HbA1c, type 1 diabetes duration, sex, pack years of smoking, and estimated cell type composition.

†All clinical variables with p<0.2 in the individual models were offered for final model selection using backward selection. NR=offered but not retained in the final model (p>0.1), NO=not offered, individual model p>0.20.

Online Supplemental Table 5. Associations between 28-Year longitudinal clinical risk factors and cg19693031 methylation, independent of HbA1c

	Minimally adjusted β per 5% cg19693031 methylation* (standard error)	**Minimally adjusted p-value***	**Fully adjusted β per 5% cg19693031 methylation* (standard error)**	**Fully adjusted p-value†**
Total Cholesterol	**-2.37 (0.92)**	**0.010**	-0.68 (0.93)	0.468
HDLc	-0.06 (0.34)	0.858	n/a	n/a
Non-HDLc	**-2.65 (0.91)**	**0.004**	-1.01 (0.92)	0.272
ln(Triglycerides)	**-0.06 (0.01)**	**<0.0001**	**-0.05 (0.01)**	**<0.0001**
Systolic BP	-0.55 (0.74)	0.452	n/a	n/a
Diastolic BP	-0.29 (0.23)	0.213	-0.27 (0.70)	0.698
Pulse rate	**-0.98 (0.22)**	**<0.0001**	**-0.84 (0.22)**	**<0.0001**
BMI	-0.04 (0.10)	0.663	n/a	n/a
Insulin Dose	-0.01 (0.01)	0.073	n/a	n/a
ln(ACR)	**-0.24 (0.05)**	**<0.0001**	**-0.20 (0.05)**	**<0.0001**
eGFR	0.89 (0.73)	0.221	n/a	n/a

Bold text indicates p<0.05

*Minimally adjusted for HbA1c, type 1 diabetes duration, sex, pack years of smoking, and cell type composition

†For factors significant at p<0.05 after minimal adjustment, fully adjusted for HbA1c, type 1 diabetes duration, sex, pack years of smoking, and cell type composition and all other factors with p-value<0.05 (total cholesterol model not adjusted for non-HDLc and vice versa)