Table 1. Control Probes on the Illumina 450K methylation array.

Category	Description	No. of Types	No. of Probes
Bisulfite Conversion	Methylation at a site known to be methylated	3	10
Normalisation	Randomly permutated bisulphite-converted sequences containing no CpGs; Determines system background	4	186
Staining	Efficiency and sensitivity of staining step	2	2
Extension	Extension efficiency of A, T, C, and G nucleotides from a hairpin probe	4	4
Hybridisation	Hybridisation efficiency using synthetic targets instead of amplified DNA	3	3
Target Removal	Efficiency of stripping step after extension reaction	1	2
Specificity	Methylation at non-polymorphic T sites	3	9
Non-polymorphic	Methylation at a base in a non-polymorphic region of the genome	4	4

Table 2: Genomic Inflation Factor λ under the null hypothesis by permutation testing, using different levels of adjustment.

	Median	2.5 Percentile	97.5 Percentile
no adjustment	0.91	0.58	2.87
QN	0.98	0.80	1.44
+ Control Probe PCs	0.96	0.84	1.46
+ Gender, Age	0.96	0.84	1.47
+ WBC_{tot} + WBC_{est}	1.00	0.96	1.09
+ PC 1-5	1.00	0.97	1.05

Table 3. Simulations show significant performance improvements for each stage of the analysis pipeline (paired Wilcoxon rank test).

	% of spiked markers in top 100		
	Median	**2nd Decile**	**9th Decile**
no adjustment	46%	1%	97%
QN	66%	5%	98%
+ Control Probe PCs	71%	6%	98%
+ Gender, Age	71%	6%	98%
+ WBC_{tot} + WBC_{est}	75%	5%	99%
+ PC 1-5	76%	6%	99%

Table 4. Baseline characteristics of South Asian incident T2D cases and controls in the epigenome-wide association study. Results are presented as mean (SD) or as % (N).

	Incident T2D	**Controls**	**P**
N	1,074	1,590	
Follow-up (yrs)	8.6 (1.7)	8.4 (1.7)	0.94
Age (yrs)	52.5 (10.2)	49.9 (9.8)	<0.001
Sex (M)	67.3% (722)	68.2% (1,083)	0.61
Impaired fasting glucose	18.9% (207)	3.4% (55)	<0.001
Fasting glucose (mmol/L)	5.49 (0.59)	5.05 (0.47)	<0.001
HbA1c (%)	5.77 (0.49)	5.37 (0.48)	<0.001
Insulin (IU/L)	15.5 (10.9)	10.7 (9.2)	<0.001
Body mass index (kg/m^2)	28.9 (4.6)	26.7 (3.9)	<0.001
Waist circumference (cm)	101.0 (11.5)	94.9 (10.3)	<0.001
Waist-hip ratio	0.97 (0.07	0.94 (0.07)	<0.001

Table 5. Association of methylation markers with future T2D. Results are presented as relative risk for T2D [95% confidence interval] associated with a 1SD increase in respective methylation marker in the discovery phase (1,074 Indian Asians with incident T2D and 1,590 controls), in replication testing amongst 1,141 Europeans (377 with incident T2D), and in combined analysis. P_{het} is for heterogeneity of effect between discovery and replication.

Marker	Discovery	P	Replication	P	Combined	P	P_{het}
Marker 1	0.68 [0.62, 0.75]	1.0×10^{-13}	0.70 [0.60, 0.83]	2.5×10^{-5}	0.68 [0.63, 0.75]	1.5×10^{-18}	0.98
Marker 2	0.77 [0.70, 0.85]	9.3×10^{-8}	0.94 [0.82, 1.07]	0.32	0.83 [0.77, 0.89]	4.8×10^{-7}	0.04
Marker 3	0.68 [0.60, 0.78]	1.4×10^{-8}	0.98 [0.87, 1.10]	0.71	0.84 [0.78, 0.92]	4.8×10^{-5}	0.004
Marker 4	1.40 [1.26, 1.57]	8.4×10^{-9}	1.19 [1.05, 1.34]	5.4×10^{-3}	1.30 [1.20, 1.41]	3.0×10^{-10}	0.07
Marker 5	0.73 [0.66, 0.81]	2.1×10^{-9}	0.80 [0.70, 0.92]	1.2×10^{-3}	0.75 [0.70, 0.82]	4.1×10^{-12}	0.48
Marker 6	0.78 [0.71, 0.85]	2.1×10^{-7}	0.81 [0.72, 0.93]	1.6×10^{-3}	0.79 [0.73, 0.85]	4.7×10^{-10}	0.76
Marker 7	1.44 [1.31, 1.58]	2.2×10^{-13}	1.31 [1.14, 1.50]	1.2×10^{-4}	1.40 [1.29, 1.51]	1.1×10^{-17}	0.32

Table 6. . Replication testing for association with prevalent T2D in Indian Asians. Results are expressed as OR (95%CI) per 1SD increase in methylation. P is the P-value in combined analysis.

Locus	Migrant	Non-Migrant	Combined	P
Marker 1	0.25 (0.19 to 0.33)	0.35 (0.30 to 0.42)	0.32 (0.28 to 0.37)	9.7E-55
Marker 4	1.53 (1.25 to 1.87)	1.50 (1.30 to 1.74)	1.51 (1.34 to 1.70)	9.8E-12
Marker 5	0.83 (0.68 to 1.00)	0.71 (0.62 to 0.80)	0.74 (0.67 to 0.83)	5.2E-08
Marker 6	0.93 (0.76 to 1.13)	0.75 (0.66 to 0.85)	0.80 (0.72 to 0.89)	6.8E-05
Marker 7	1.15 (0.97 to 1.36)	1.71 (1.46 to 2.01)	1.42 (1.26 to 1.60)	4.4E-09

Table 7. Summary of current knowledge for candidate genes at the identified loci.

Marker 1	- Key component of pancreatic beta cell biology, nutrient sensing, energy metabolism and regulation of cellular redox. - Downregulates GLUT1, the major transmembrane glucose transporter, thereby acting as a negative feedback loop to regulate glucose entry and mitochondrial oxidative stress. - May contribute to regulation of adiposity and energy expenditure through hypothalamic pathways.
Marker 4	- Master transcriptional regulator of hepatic lipogenesis, capable of inducing the entire complement of genes necessary for the synthesis of monounsaturated fatty acids. - Decreased in insulin-deficient states, such as fasting, but increased in feeding, obesity and insulin resistance.
Marker 5	- Encodes a phosphatase involved in the generation of inorganic phosphate for bone mineralization; highly expressed at sites of mineralization in bone and cartilage. - Not previously associated with obesity, insulin action or diabetes.
Marker 6	- Major negative regulator of insulin signaling, and implicated in the pathogenesis of obesity and associated metabolic abnormalities. - In mouse model, knockout mice are protected against obesity induced hyperinsulinemia and insulin resistance.
Marker 7	- Involved in macrophage cholesterol and phospholipid transport, and promotes cholesterol efflux to HDL. - Knockout mice have impaired glucose tolerance and insulin secretion with normal insulin sensitivity. - Expression downregulated in humans with diabetes and upregulated by the insulin sensitizing agents such as thiazolidenediones.

Table 8. Positive predictive value and sensitivity for prediction of T2D based on models comprising known risk factors with or without addition of methylation score. In each model, positive predictive value and sensitivity are calculated at the 90th centile of risk score in people without T2D.

	Positive Predictive Value		**Sensitivity for T2D**	
	Without methylation	*With methylation*	*Without methylation*	*With methylation*
Risk factors				
Age, Sex, Family history T2D, Physical activity	16.7%	21.2%	18.0%	25.0%
+ Body mass index, Waist-hip ratio	20.8%	25.0%	23.6%	30.2%
+ Glucose, HbA1c	33.6%	36.0%	48.8%	51.3%

Table 9. Association of DNA methylation with gene expression in peripheral blood amongst Indian Asians and Europeans.

	Peripheral Blood					
	LOLIPOP (SA; n=907)		**EGM (EUR; n=591)**		**KORA (EUR; n=703)**	
Marker	**Beta**	**P**	**Beta**	**P**	**Beta**	**P**
Marker 1	2.20	3.00E-02	0.12	8.40E-01	-0.03	3.40E-01
Marker 4	-3.73	2.00E-04	-1.58	8.80E-05	-0.15	3.80E-03
Marker 5	1.69	9.00E-02	3.29	2.00E-02	0.08	5.50E-01
Marker 6	-1.55	1.20E-01	2.58	1.90E-03	0.09	3.00E-02
Marker 7	-9.73	3.80E-21	-2.51	1.40E-06	-5.20	1.50E-20

Table 10. Characteristics of South Asians in the family study

	Mean	**SD**
Age (years)	38.0	15.3
Sex (male)	52.8%	
Coronary heart disease (%)	15.0%	
Type 2 Diabetes (%)	9.8%	
Glucose (mmol/L)	5.5	1.8
HbA1c (%)	5.1	1.1
Insulin (IU/L)	10.3	7.6
Body mass index (kg/m2)	25.5	4
Waist (cm)	88.0	35.2
Waist-hip ratio	0.88	0.37
Systolic BP (mmHg)	123.1	20.4
Diastolic BP (mmHg)	71.0	11.7
Cholesterol (mmol/L)	5.1	1.0
Triglycerides (mmol/L)	1.5	0.9
HDL cholesterol (mmol/L)	1.3	0.3

Figure 1. Correlation between duplicate measurements of methylation in each sample, using different approaches to data-normalisation. Subset QN: Subset Quantile Normalization; Raw: No normalisation; BMIQ: Beta MIxture Quantile dilation; PBC: Peak-Based Correction; Illumina: Illumina Control Probe Normalisation; SWAN: Subset Within-Array Normalisation; QN-I6: Quantile Normalisation of Intensity-values (6 categories). Methods are sorted based on the median correlation coefficient.

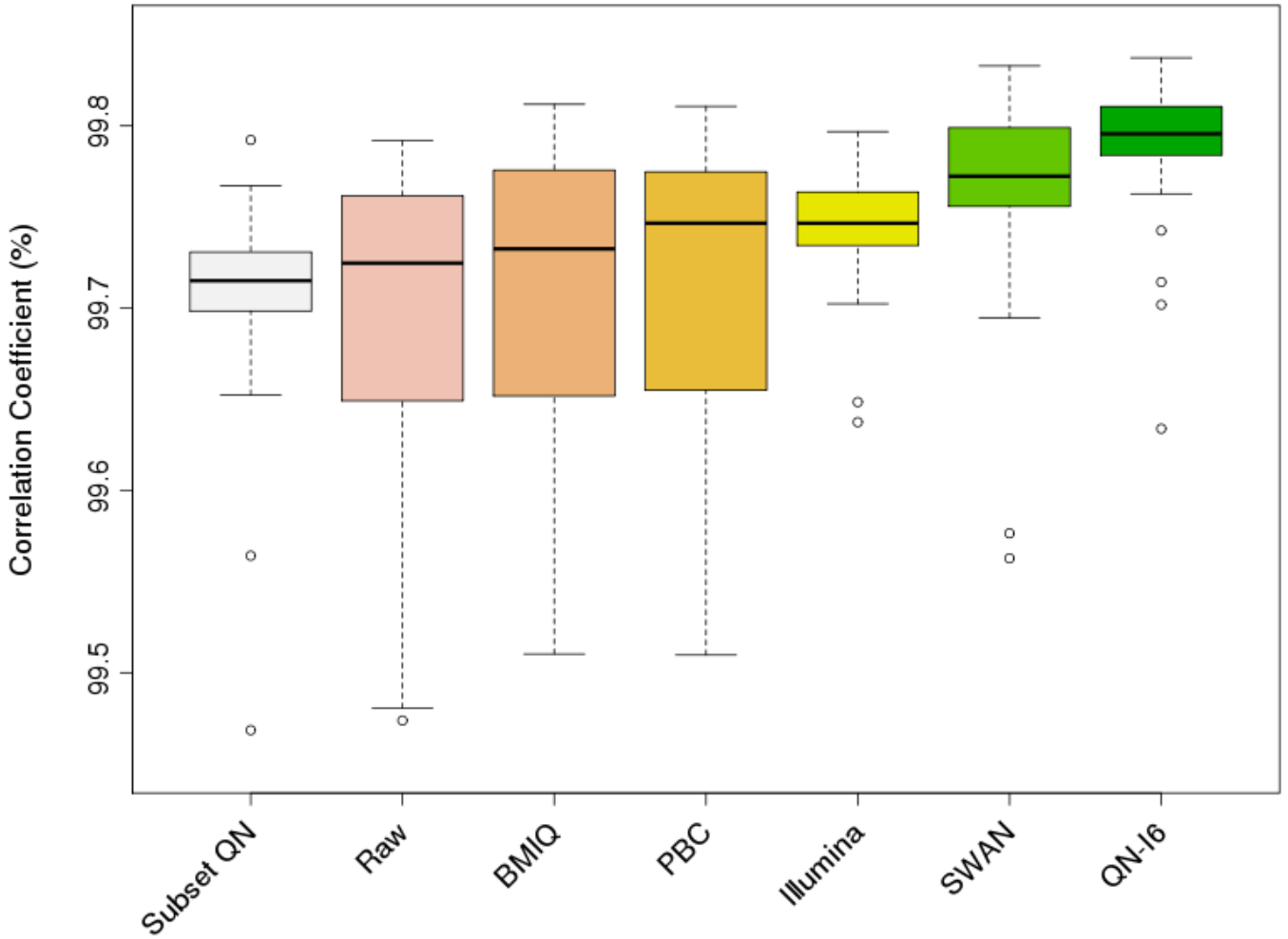


Figure 2. Simulation analysis comparing different approaches to data normalisation. For each normalisation method we increased ("spiked") beta-values of 100 randomly selected markers and determined the proportion of spiked markers that were ranked amongst the top 100. SWAN: Subset Within-Array Normalisation; Illumina: Illumina Control Probe Normalisation; Raw: No normalisation; BMIQ: Beta MIxture Quantile dilation; PBC: Peak-Based Correction; Subset QN: Subset Quantile Normalization; QN-I6: Quantile Normalisation of Intensity-values (6 categories). Methods in the figure legend are sorted based on the average percentage of top 100 spike-markers.

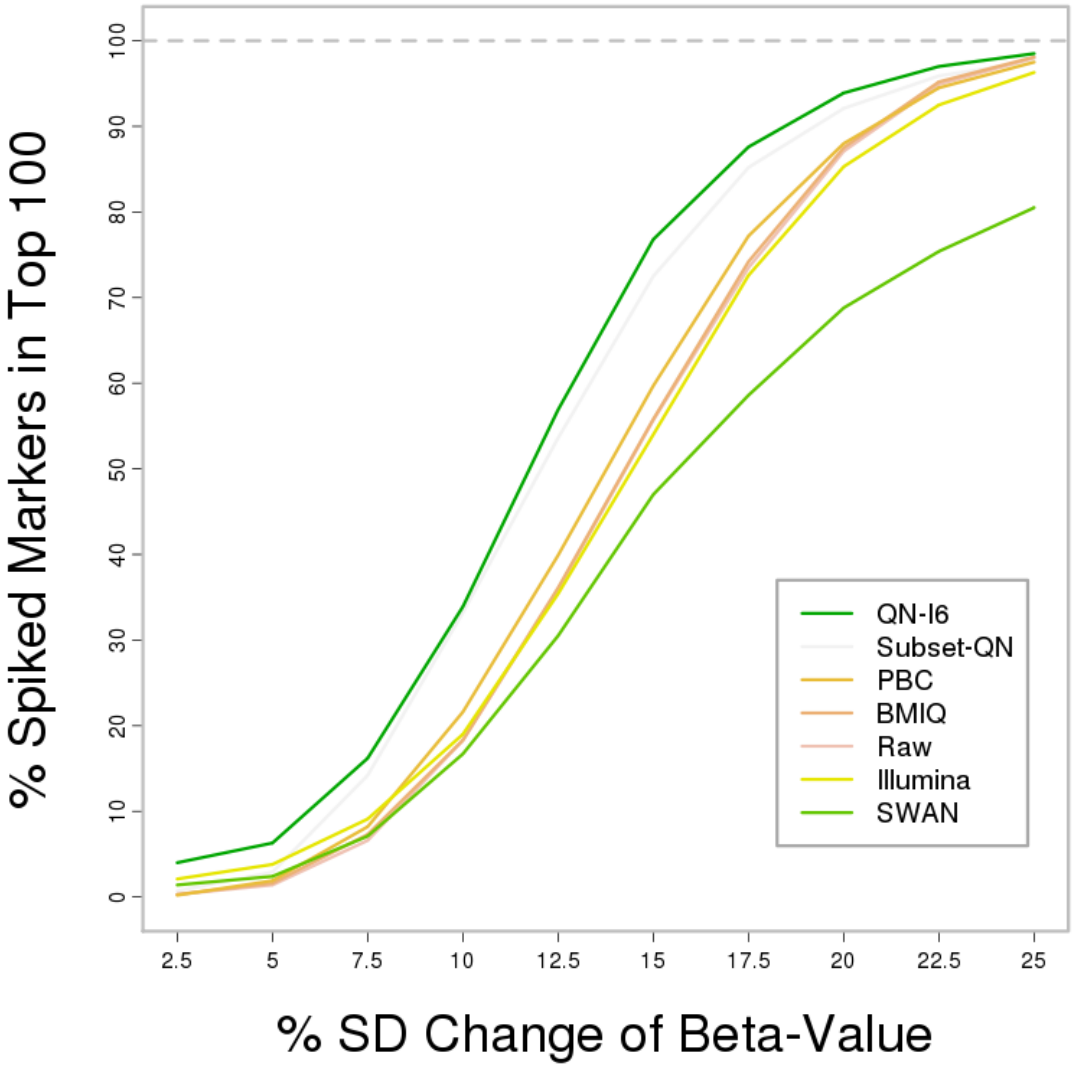


Figure 3. Correcting for statistical inflation due to technical biases. Quantile-Quantile (QQ) plot for the comparison of 36 samples measured in duplicate reveals high statistical inflation after quantile normalisation (λ_{QN}=2.11; green points) due to technical biases. Batch-correction based on control probes removes technical biases and statistical inflation (λ_{CP}=1.01; green points).

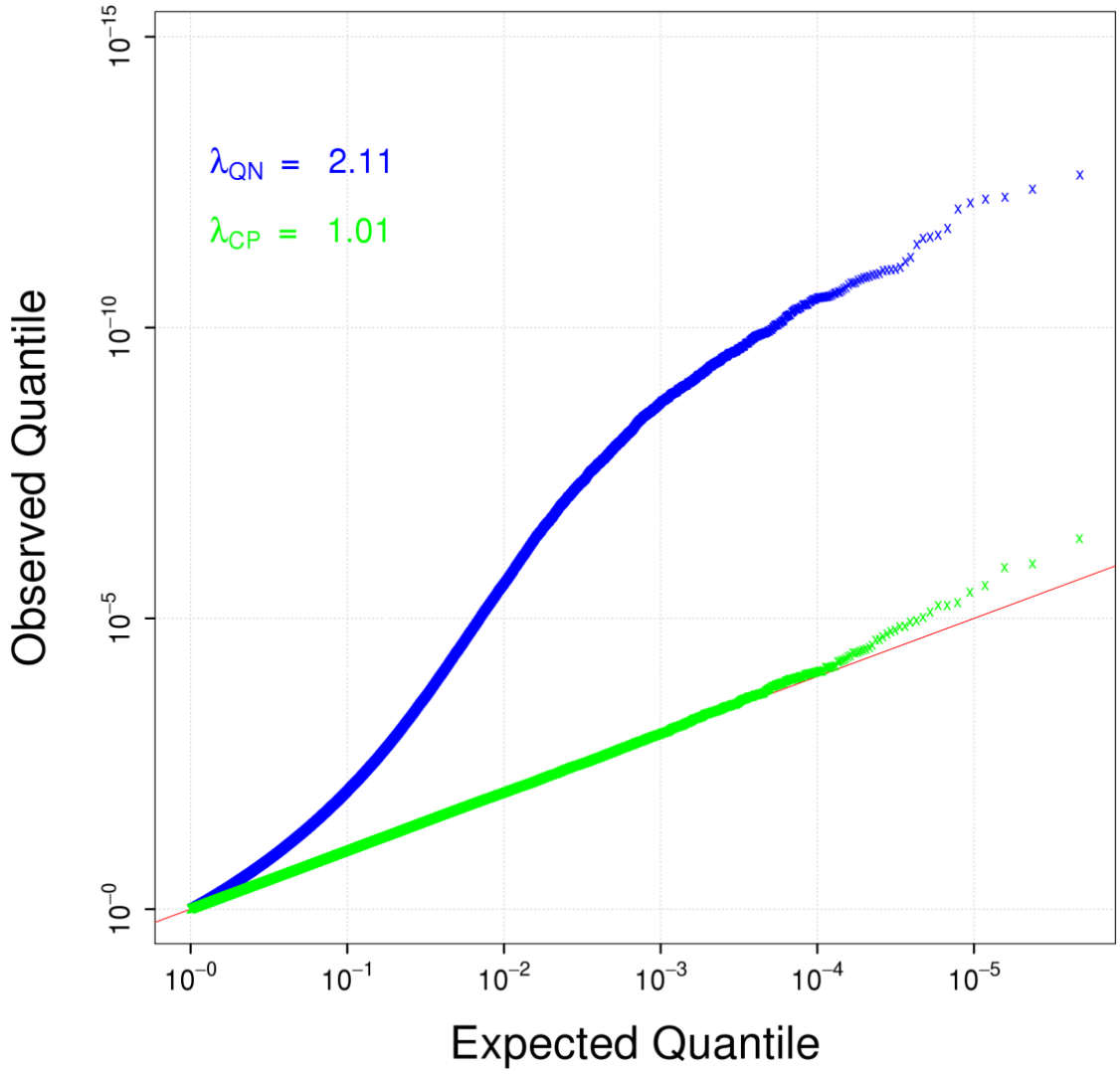


Figure 4. PCA of control probe intensities reveals strong correlations between Principal Components and technical factors in the population study.

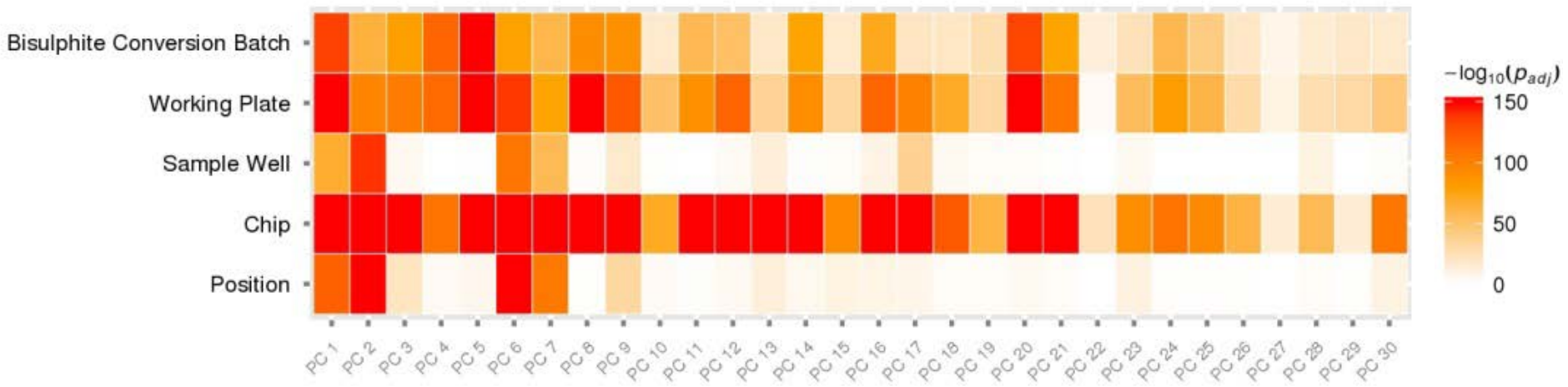


Figure 5. QQ plot showing results for the Epigenome-wide association study. Blue dots represent results for individual CpG sites for association with T2D. The red line represents the null hypothesis of no association; the grey shading provides the 95% confidence interval for the null hypothesis.

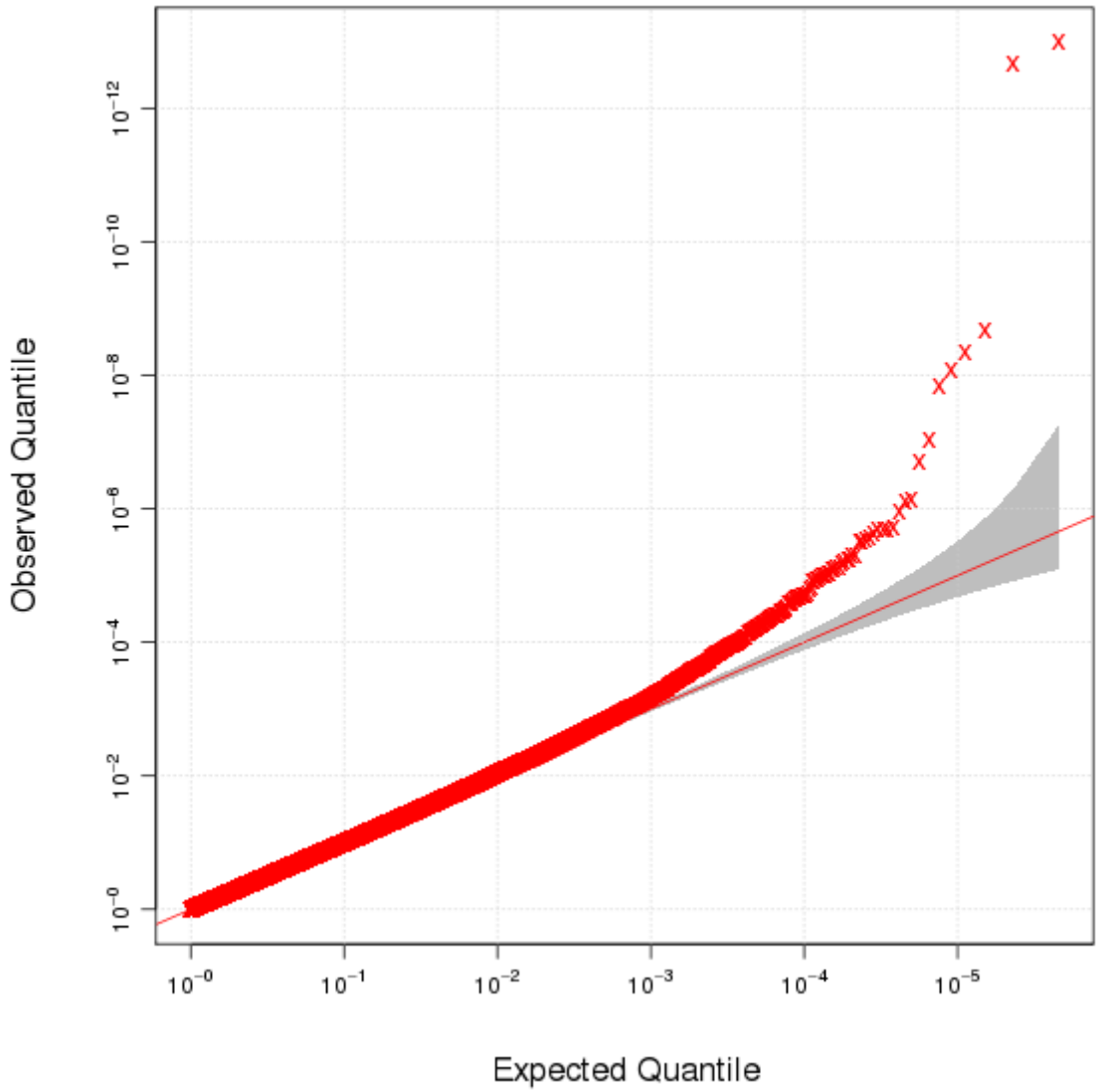


Figure 6. Manhattan plot for the Epigenome-wide association study

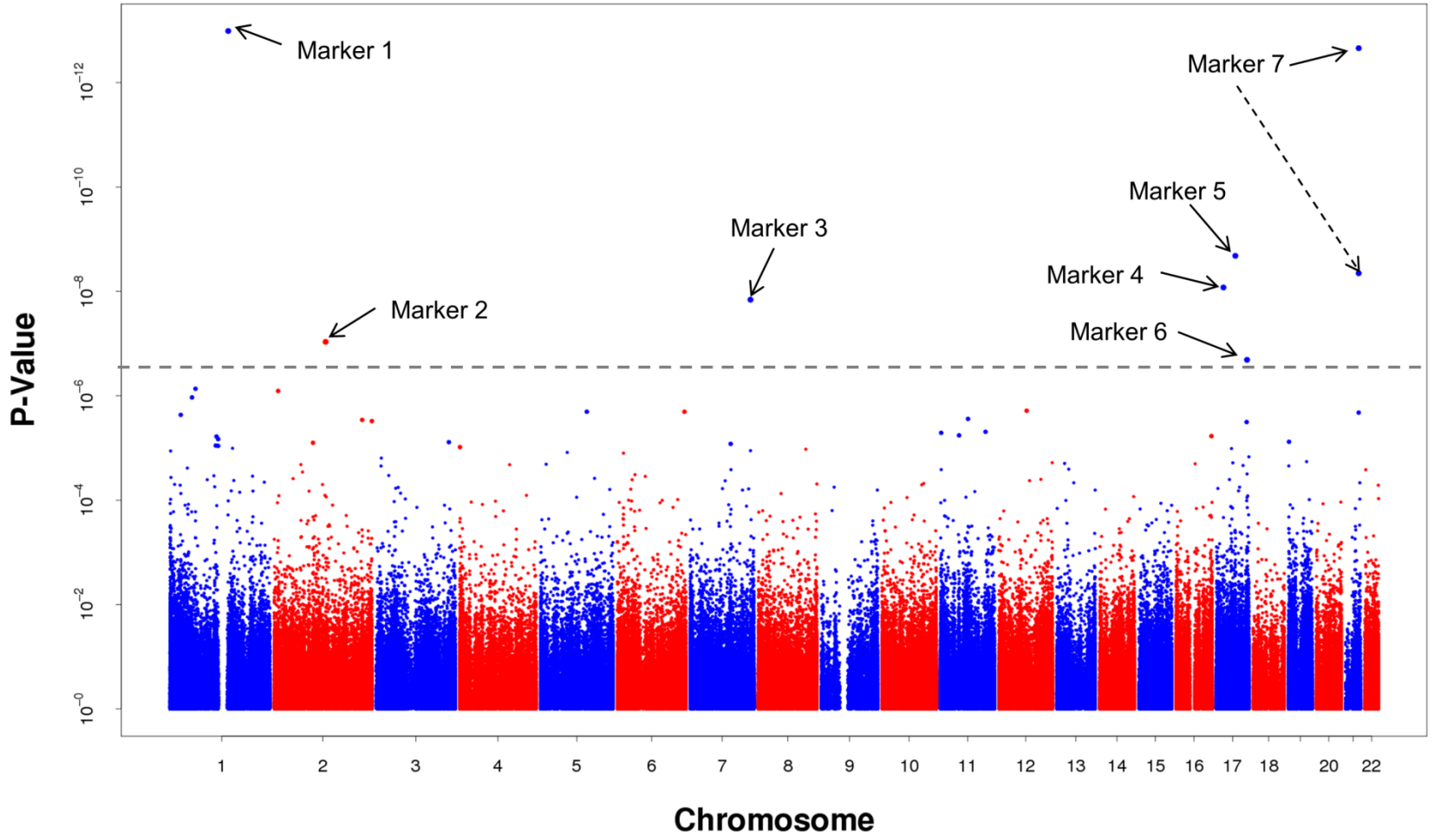


Figure 7. Relative risk of T2D (with 95% Confidence Interval) between top and bottom quartiles of methylation for replicated markers.

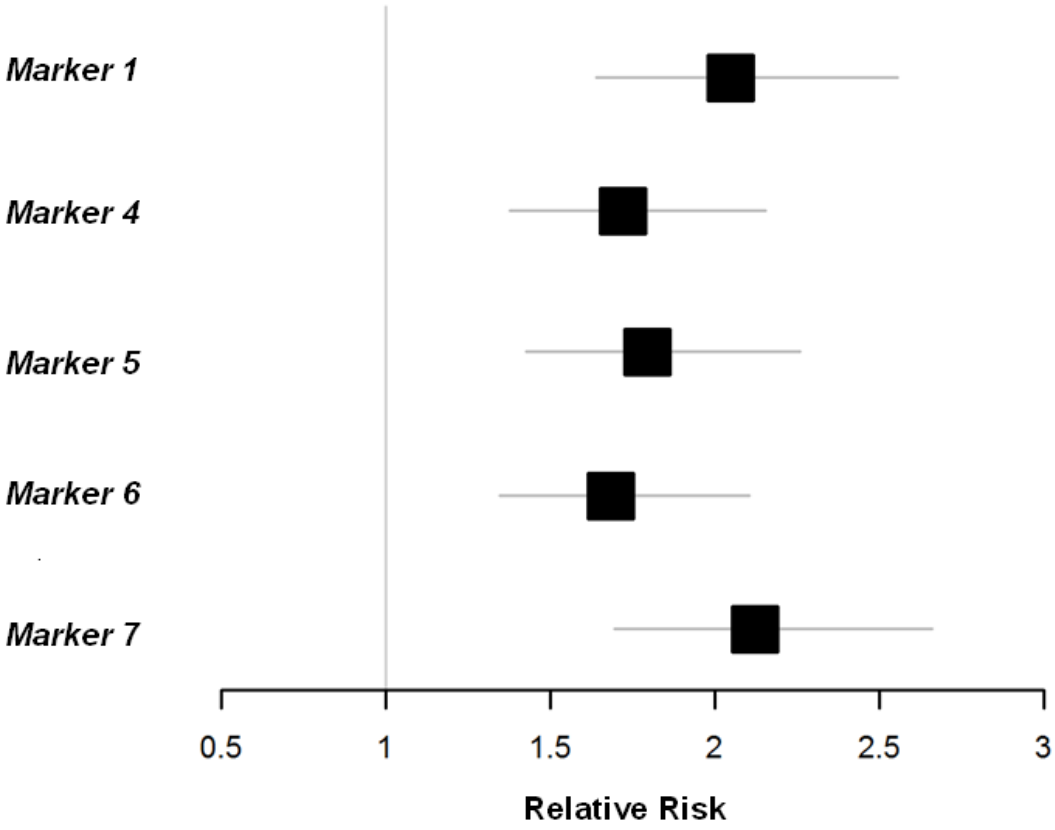


Figure 8. Incidence of T2D (8 year follow-up) by quartile of methylation score amongst normal weight (Body mass index [BMI] 18.5-24.9kg/m^2), overweight (BMI 25.0-29.9kg/m^2) and obese (BMI ≥30.0kg/m^2) Indian Asians with normoglycaemia (HbA1c<6% and fasting glucose<6mmol/l).

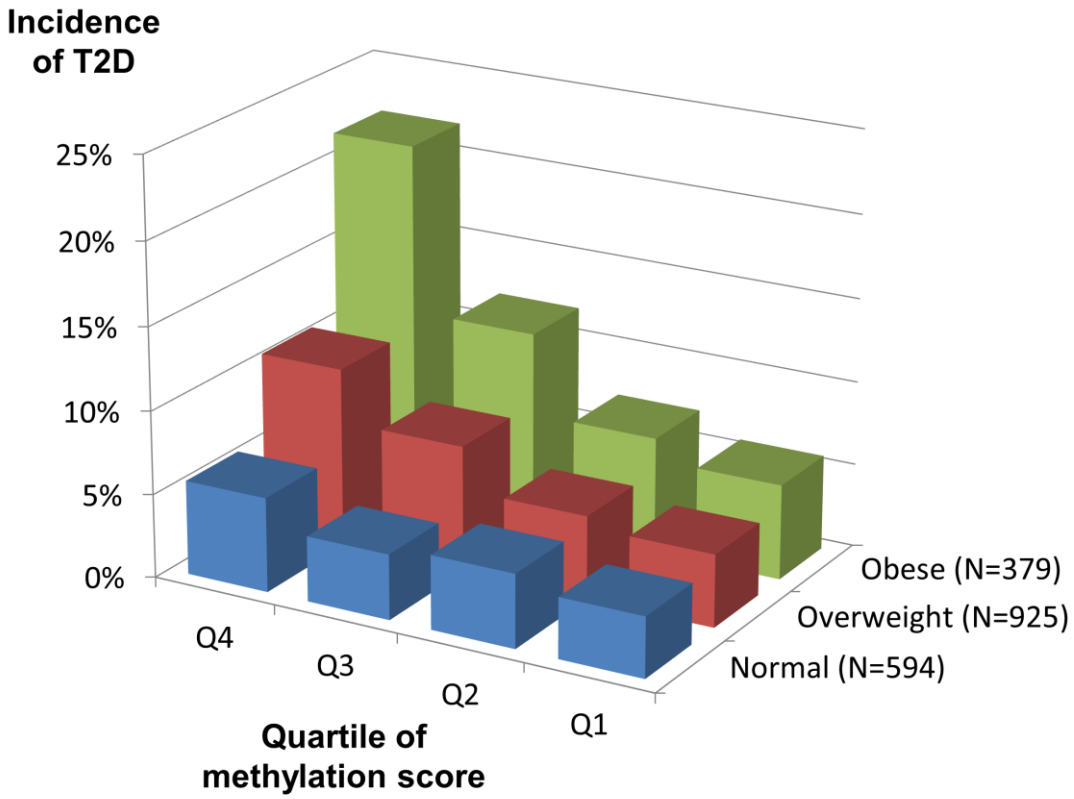


Figure 9. Targeted resequencing of the top locus by next generation sequencing. Bars present mean methylation at the CpG sites evaluated; the sentinel marker identified by EWAS is shown as a purple bar. The correlation track provides the correlation between methylation at each CpG site with the sentinel marker.

Inset: Relative risk for T2D for the methylation markers at the top locus identified by targeted resequencing. Results shown for: i. the 8 individual CpG sites assayed by pyrosequencing (green; light green for sentinel marker); ii. Sentinel marker by microarray (blue); and iii. Regional methylation scores (red). Results are as relative risk for T2D associated with 1SD reduction in methylation or methylation score. Score: sum of all 8 methylation markers.

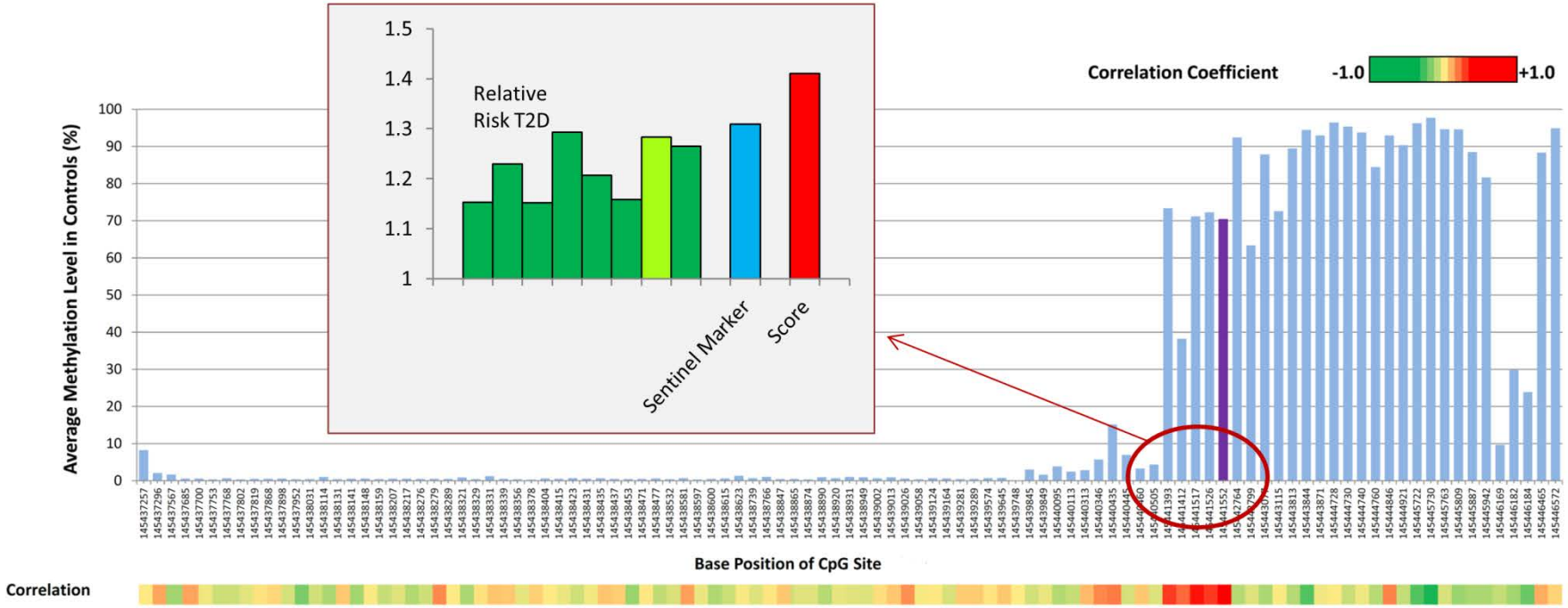
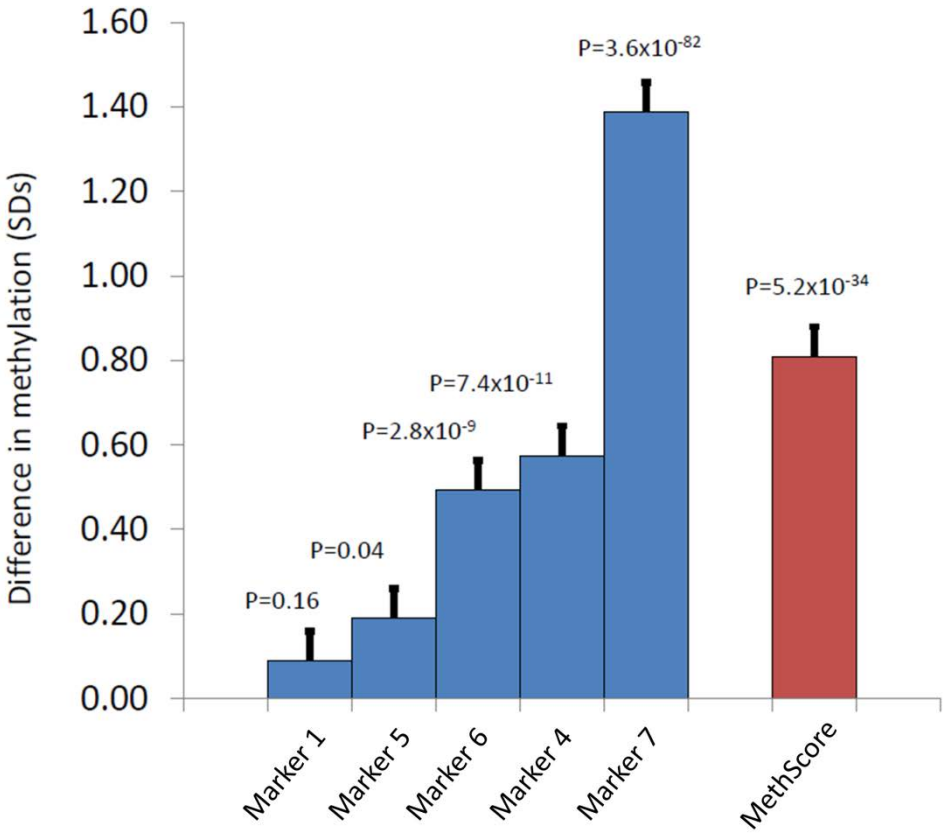


Figure 10. Difference in methylation between South Asians and Europeans without T2D. Results presented as the SD increase in methylation amongst South Asians compared to Europeans, aligned to the direction of effect for increased T2D risk, adjusted for age and gender.



References

1. International Diabetes Federation, *IDF Diabetes Atlas*: Brussels, Belgium.
2. Danaei, G., et al., *National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants.* Lancet, 2011. **378**(9785): p. 31-40.
3. Chambers, J.C., et al., *Plasma homocysteine concentrations and risk of coronary heart disease in UK Indian Asian and European men.* Lancet, 2000. **355**(9203): p. 523-7.
4. Kooner, J.S., et al., *Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci.* Nat Genet, 2011. **43**(10): p. 984-9.
5. Smith, Z.D. and A. Meissner, *DNA methylation: roles in mammalian development.* Nat Rev Genet, 2013. **14**(3): p. 204-20.
6. Dowen, R.H., et al., *Widespread dynamic DNA methylation in response to biotic stress.* Proc Natl Acad Sci U S A, 2012. **109**(32): p. E2183-91.
7. Murgatroyd, C., et al., *Dynamic DNA methylation programs persistent adverse effects of early-life stress.* Nat Neurosci, 2009. **12**(12): p. 1559-66.
8. Wagner, J.R., et al., *The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts.* Genome Biol, 2014. **15**(2): p. R37.
9. Lehne, B., et al., *A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies.* Genome Biol, 2015. **In press**.
10. Shaw, J.E., R.A. Sicree, and P.Z. Zimmet, *Global estimates of the prevalence of diabetes for 2010 and 2030.* Diabetes Res Clin Pract, 2010. **87**(1): p. 4-14.
11. Ebrahim, S., et al., *The effect of rural-to-urban migration on obesity and diabetes in India: a cross-sectional study.* PLoS Med, 2010. **7**(4): p. e1000268.
12. McKeigue, P.M., B. Shah, and M.G. Marmot, *Relation of central obesity and insulin resistance with high diabetes prevalence and cardiovascular risk in South Asians.* Lancet, 1991. **337**(8738): p. 382-6.
13. Sicree, R., J. Shaw, and P. Zimmet *The global burden of diabetes*.
14. Ramachandran, A., et al., *Increasing expenditure on health care incurred by diabetic subjects in a developing country: a study from India.* Diabetes Care, 2007. **30**(2): p. 252-6.
15. *Eurostat - Healthcare statistics.* . [cited 2014; Available from: http://epp.eurostat.ec.europa.eu.
16. Nyamdorj, R., et al., *Ethnic comparison of the association of undiagnosed diabetes with obesity.* Int J Obes (Lond), 2010. **34**(2): p. 332-9.
17. Ma, R.C. and J.C. Chan, *Type 2 diabetes in East Asians: similarities and differences with populations in Europe and the United States.* Ann N Y Acad Sci, 2013. **1281**: p. 64-91.
18. Chambers, J.C. and J.S. Kooner, *Diabetes, insulin resistance and vascular disease among Indian Asians and Europeans.* Semin Vasc Med, 2002. **2**(2): p. 199-214.
19. Kooner, J.S., et al., *Abdominal obesity, impaired nonesterified fatty acid suppression, and insulin-mediated glucose disposal are early metabolic abnormalities in families with premature myocardial infarction.* Arterioscler Thromb Vasc Biol, 1998. **18**(7): p. 1021-6.
20. Sevak, L., P.M. McKeigue, and M.G. Marmot, *Relationship of hyperinsulinemia to dietary intake in south Asian and European men.* Am J Clin Nutr, 1994. **59**(5): p. 1069-74.

21. McKeigue, P.M., et al., *Association of early-onset coronary heart disease in South Asian men with glucose intolerance and hyperinsulinemia.* Circulation, 1993. **87**(1): p. 152-61.
22. Barnett, A.H., et al., *Type 2 diabetes and cardiovascular risk in the UK south Asian community.* Diabetologia, 2006. **49**(10): p. 2234-46.
23. Chambers, J.C., et al., *Common genetic variation near MC4R is associated with waist circumference and insulin resistance.* Nat Genet, 2008. **40**(6): p. 716-8.
24. Chambers, J.C., et al., *Common genetic variation near melatonin receptor MTNR1B contributes to raised plasma glucose and increased risk of type 2 diabetes among Indian Asians and European Caucasians.* Diabetes, 2009. **58**(11): p. 2703-8.
25. Kooner, J.S., et al., *Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides.* Nat Genet, 2008. **40**(2): p. 149-51.
26. Petronis, A., *Epigenetics as a unifying principle in the aetiology of complex traits and diseases.* Nature, 2010. **465**(7299): p. 721-7.
27. Schaefer, C.B., et al., *Epigenetic decisions in mammalian germ cells.* Science, 2007. **316**(5823): p. 398-9.
28. Bhaumik, S.R., E. Smith, and A. Shilatifard, *Covalent modifications of histones during development and disease pathogenesis.* Nat Struct Mol Biol, 2007. **14**(11): p. 1008-16.
29. Bronner, C., *Control of DNMT1 abundance in epigenetic inheritance by acetylation, ubiquitylation, and the histone code.* Sci Signal, 2011. **4**(157): p. pe3.
30. Youngson, N.A. and E. Whitelaw, *Transgenerational epigenetic effects.* Annu Rev Genomics Hum Genet, 2008. **9**: p. 233-57.
31. Gluckman, P.D., et al., *Epigenetic mechanisms that underpin metabolic and cardiovascular diseases.* Nat Rev Endocrinol, 2009. **5**(7): p. 401-8.
32. Bell, C.G., et al., *Integrated genetic and epigenetic analysis identifies haplotype-specific methylation in the FTO type 2 diabetes and obesity susceptibility locus.* PLoS One, 2010. **5**(11): p. e14040.
33. Dolinoy, D.C., D. Huang, and R.L. Jirtle, *Maternal nutrient supplementation counteracts bisphenol A-induced DNA hypomethylation in early development.* Proc Natl Acad Sci U S A, 2007. **104**(32): p. 13056-61.
34. Raychaudhuri, N., et al., *Histone code modifications repress glucose transporter 4 expression in the intrauterine growth-restricted offspring.* J Biol Chem, 2008. **283**(20): p. 13611-26.
35. Yajnik, C.S., et al., *Neonatal anthropometry: the thin-fat Indian baby. The Pune Maternal Nutrition Study.* Int J Obes Relat Metab Disord, 2003. **27**(2): p. 173-80.
36. Whincup, P.H., et al., *Birth weight and risk of type 2 diabetes: a systematic review.* JAMA, 2008. **300**(24): p. 2886-97.
37. Chamson-Reig, A., et al., *Altered pancreatic morphology in the offspring of pregnant rats given reduced dietary protein is time and gender specific.* J Endocrinol, 2006. **191**(1): p. 83-92.
38. Einstein, F., et al., *Cytosine methylation dysregulation in neonates following intrauterine growth restriction.* PLoS One, 2010. **5**(1): p. e8887.
39. Heijmans, B.T., et al., *Persistent epigenetic differences associated with prenatal exposure to famine in humans.* Proc Natl Acad Sci U S A, 2008. **105**(44): p. 17046-9.
40. Ravelli, A.C., et al., *Glucose tolerance in adults after prenatal exposure to famine.* Lancet, 1998. **351**(9097): p. 173-7.
41. [cited 2010; Available from: http://www.unicef.org/progressforchildren/2006n4/index_lowbirthweight.html. .

42. Raghupathy, P., et al., *Glucose tolerance, insulin resistance and insulin secretion in young south Indian adults: Relationships to parental size, neonatal size and childhood body mass index.* Diabetes Res Clin Pract, 2010. **87**(2): p. 283-92.
43. Saleheen, D., et al., *The Pakistan Risk of Myocardial Infarction Study: a resource for the study of genetic, lifestyle and other determinants of myocardial infarction in South Asia.* Eur J Epidemiol, 2009. **24**(6): p. 329-38.
44. Tapp, R.J., et al., *Six year incidence and progression of diabetic retinopathy: results from the Mauritius diabetes complication study.* Diabetes Res Clin Pract, 2006. **73**(3): p. 298-303.
45. Tan, J.T., et al., *Polymorphisms identified through genome-wide association studies and their associations with type 2 diabetes in Chinese, Malays, and Asian-Indians in Singapore.* J Clin Endocrinol Metab, 2010. **95**(1): p. 390-7.
46. Takeuchi, F., et al., *Common variants at the GCK, GCKR, G6PC2-ABCB11 and MTNR1B loci are associated with fasting glucose in two Asian populations.* Diabetologia, 2010. **53**(2): p. 299-308.
47. International Expert, C., *International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes.* Diabetes Care, 2009. **32**(7): p. 1327-34.
48. Wurtz, P., et al., *Branched-chain and aromatic amino acids are predictors of insulin resistance in young adults.* Diabetes Care, 2013. **36**(3): p. 648-55.
49. Wallace, T.M. and D.R. Matthews, *The assessment of insulin resistance in man.* Diabet Med, 2002. **19**(7): p. 527-34.
50. Holle, R., et al., *KORA--a research platform for population based health research.* Gesundheitswesen, 2005. **67 Suppl 1**: p. S19-25.
51. Bibikova, M., et al., *High density DNA methylation array with single CpG site resolution.* Genomics, 2011. **98**(4): p. 288-95.
52. Dedeurwaerder, S., et al., *Evaluation of the Infinium Methylation 450K technology.* Epigenomics, 2011. **3**(6): p. 771-84.
53. Wang, D., et al., *IMA: an R package for high-throughput analysis of Illumina's 450K Infinium methylation data.* Bioinformatics, 2012. **28**(5): p. 729-30.
54. Du, P., W.A. Kibbe, and S.M. Lin, *lumi: a pipeline for processing Illumina microarray.* Bioinformatics, 2008. **24**(13): p. 1547-8.
55. Touleimat, N. and J. Tost, *Complete pipeline for Infinium((R)) Human Methylation 450K BeadChip data processing using subset quantile normalization for accurate DNA methylation estimation.* Epigenomics, 2012. **4**(3): p. 325-41.
56. Hansen, K.D. and M. Aryee, *minfi*, 2012.
57. Morris, T., et al., *ChAMP*, 2013.
58. Illumina, *Illumina Genome Studio*, 2013.
59. Price, M.E., et al., *Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array.* Epigenetics Chromatin, 2013. **6**(1): p. 4.
60. Teschendorff, A.E., et al., *A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data.* Bioinformatics, 2013. **29**(2): p. 189-96.
61. Maksimovic, J., L. Gordon, and A. Oshlack, *SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips.* Genome Biol, 2012. **13**(6): p. R44.
62. Chen, Y.A., et al., *Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray.* Epigenetics, 2013. **8**(2): p. 203-9.
63. Pidsley, R., et al., *A data-driven approach to preprocessing Illumina 450K methylation array data.* BMC Genomics, 2013. **14**: p. 293.

64. Wu, M.C., et al., *A systematic assessment of normalization approaches for the Infinium 450K methylation platform.* Epigenetics, 2014. **9**(2): p. 318-29.
65. Johnson, W.E., C. Li, and A. Rabinovic, *Adjusting batch effects in microarray expression data using empirical Bayes methods.* Biostatistics, 2007. **8**(1): p. 118-27.
66. Houseman, E.A., et al., *DNA methylation arrays as surrogate measures of cell mixture distribution.* BMC Bioinformatics, 2012. **13**: p. 86.
67. Eckhardt, F., et al., *DNA methylation profiling of human chromosomes 6, 20 and 22.* Nat Genet, 2006. **38**(12): p. 1378-85.
68. Bell, J.T., et al., *DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines.* Genome Biol, 2011. **12**(1): p. R10.
69. Ritchie, M.E., et al., *A comparison of background correction methods for two-colour microarrays.* Bioinformatics, 2007. **23**(20): p. 2700-7.
70. Smyth, G.K., *Limma: linear models for microarray data.*, in *Bioinformatics and Computational Biology Solutions using R and Bioconductor.* 2005, Springer: New York. p. 397-420.
71. Schurmann, C., et al., *Analyzing illumina gene expression microarray data from different tissues: methodological aspects of data analysis in the metaxpress consortium.* PLoS One, 2012. **7**(12): p. e50938.
72. Kelly, R.S., et al., *Blood erythrocyte concentrations of cadmium and lead and the risk of B-cell non-Hodgkin's lymphoma and multiple myeloma: a nested case-control study.* PLoS One, 2013. **8**(11): p. e81892.
73. Slieker, R.C., et al., *Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k array.* Epigenetics Chromatin, 2013. **6**(1): p. 26.
74. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform.* Bioinformatics, 2009. **25**(14): p. 1754-60.
75. Feil, R. and M.F. Fraga, *Epigenetics and the environment: emerging patterns and implications.* Nat Rev Genet, 2011. **13**(2): p. 97-109.
76. Dick, K.J., et al., *DNA methylation and body-mass index: a genome-wide analysis.* Lancet, 2014. **383**(9933): p. 1990-8.
77. Evans-Galea, M.V., et al., *FXN methylation predicts expression and clinical outcome in Friedreich ataxia.* Ann Neurol, 2012. **71**(4): p. 487-97.
78. Morris, A.P., et al., *Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes.* Nat Genet, 2012. **44**(9): p. 981-90.
79. Sanghera, D.K. and P.R. Blackett, *Type 2 Diabetes Genetics: Beyond GWAS.* J Diabetes Metab, 2012. **3**(198).