PLOS ONE

Cohort profile: Epigenetics in Pregnancy (EPIPREG) – population-based sample of European and South Asian pregnant women with epigenome-wide DNA methylation (850k) in peripheral blood leukocytes
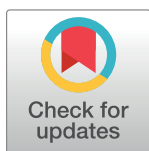
Nicolas Fragoso-Bargas[1,2], Julia O. Opsahl[2], Nadezhda Kiryushchenko[1,3], Yvonne Böttcher[2,4,5], Sindre Lee-Ødegård[2], Elisabeth Qvigstad[1,2], Kåre Rønn Richardsen[6], Christin W. Waage[7,8], Line Sletner[2,9], Anne Karen Jenum[8], Rashmi B. Prasad[10], Leif C. Groop[10], Gunn-Helen Moen[2,11,12,13], Kåre I. Birkeland[1,2], Christine Sommer[1] *

1 Department of Endocrinology, Morbid Obesity and Preventive Medicine, Oslo University Hospital, Oslo, Norway, 2 Institute of Clinical Medicine, University of Oslo, Oslo, Norway, 3 Department of Bioscience, University of Oslo, Oslo, Norway, 4 Department of Clinical Molecular Biology, Akershus University Hospital, Lørenskog, Norway, 5 Helmholtz-Institute for Metabolic, Adiposity and Vascular Research, Leipzig, Germany, 6 Faculty of Health Sciences, Department of Physiotherapy, Oslo Metropolitan University, Oslo, Norway, 7 Faculty of Health Sciences, Department of Physiotherapy, Oslo Metropolitan University, Oslo, Norway, 8 Department of General Practice, General Practice Research Unit (AFE), Institute of Health and Society, University of Oslo, Oslo, Norway, 9 Department of Pediatric and Adolescents Medicine, Akershus University Hospital, Lørenskog, Norway, 10 Lund University Diabetes Centre, Malmö, Sweden, 11 The University of Queensland Diamantina Institute, The University of Queensland, Woolloongabba, QLD, Australia, 12 Department of Public Health and Nursing, K.G. Jebsen Center for Genetic Epidemiology, NTNU, Norwegian University of Science and Technology, Trondheim, Norway, 13 Population Health Science, Bristol Medical School, University of Bristol, Bristol, United Kingdom

* christine.sommer@medisin.uio.no

Abstract

Pregnancy is a valuable model to study the association between DNA methylation and several cardiometabolic traits, due to its direct potential to influence mother's and child's health. Epigenetics in Pregnancy (EPIPREG) is a population-based sample with the aim to study associations between DNA-methylation in pregnancy and cardiometabolic traits in South Asian and European pregnant women and their offspring. This cohort profile paper aims to present our sample with genetic and epigenetic data and invite researchers with similar cohorts to collaborative projects, such as replication of ours or their results and meta-analysis. In EPIPREG we have quantified epigenome-wide DNA methylation in maternal peripheral blood leukocytes in gestational week 28±1 in Europeans (n = 312) and South Asians (n = 168) that participated in the population-based cohort STORK Gror-uddalen, in Norway. DNA methylation was measured with Infinium MethylationEPIC BeadChip (850k sites), with technical validation of four CpG sites using bisulphite pyrose-quencing in a subset (n = 30). The sample is well characterized with few missing data on e.g. genotype, universal screening for gestational diabetes, objectively measured physical activity, bioelectrical impedance, anthropometrics, biochemical measurements, and a

biobank with maternal serum and plasma, urine, placenta tissue. In the offspring, we have repeated ultrasounds during pregnancy, cord blood, and anthropometrics up to 4 years of age. We have quantified DNA methylation in peripheral blood leukocytes in nearly all eligible women from the STORK Groruddalen study, to minimize the risk of selection bias. Genetic principal components distinctly separated Europeans and South Asian women, which fully corresponded with the self-reported ethnicity. Technical validation of 4 CpG sites from the methylation bead chip showed good agreement with bisulfite pyrosequencing. We plan to study associations between DNA methylation and cardiometabolic traits and outcomes.

Introduction

Studies of epigenetic marks have in recent years gained increased interest in the context of human diseases [1]. Such studies may enhance our biological understanding of the aetiology of several diseases, increase our understanding of detrimental or protective mechanisms, or for prognosis and risk prediction [2]. One of the most studied epigenetic mechanisms is DNA methylation, which plays an important role in normal development, chromatin organization and gene expression [3]. Several studies have indicated that DNA methylation is associated with cardiovascular risk factors such as body mass index (BMI) [4–7], gestational diabetes (GDM) [8], type 2 diabetes (T2D) [9–13], lipid levels [14,15], hypertension [16,17], smoking [18–22] and alcohol intake [23–25], suggesting that cardiometabolic diseases have an epigenetic component.

Although scarcely studied, pregnant women provide a unique opportunity to study the association between blood DNA methylation and several phenotypes related to glucose homeostasis and cardiovascular traits. This is because pregnancy has been proposed as a stress test for metabolism in several organs [26], including the pancreatic beta-cells, since insulin resistance increase naturally in all pregnancies [27]. In the third trimester of pregnancy, this insulin resistance in many women reaches a level similar to that observed in type 2 diabetes, requiring the beta-cell to increase its insulin secretion considerably to compensate [28]. Similarly, pregnancy-induced hypertension is associated with increased risk of future cardiovascular disease [29].

In the Epigenetics in Pregnancy (EPIPREG) sample, we have quantified epigenome-wide DNA-methylation in peripheral blood leukocytes in women of European and South Asian origin attending the well-characterized, multi-ethnic and population-based STORK Groruddalen (STORK G) study [30]. The population based design and inclusion of women with Western and South Asian ethnicity allow us to study of a wide range of phenotypes to detect either ethnic-specific and/or common DNA methylation patterns in relation to phenotypes of interest. Furthermore, South Asians are of special interest due to their higher weight retention after pregnancy, increased prevalence of gestational diabetes, and increased risk for later type 2 diabetes compared to Europeans [31,32]. The aim of EPIPREG is to discover novel associations between DNA-methylation in pregnancy and cardiometabolic related traits in South Asian and European pregnant women and their offspring, which may have potential for prevention and treatment. This cohort profile paper aims to present our sample with genetic and epigenetic data and invite researchers with similar cohorts to collaborative projects, such as replication of ours or their results and meta-analysis.

Cohort description

Study population

EPIPREG (n = 480) is a sub-study of the larger STORK Groruddalen (STORK G) study, which is a population-based cohort of 823 healthy women with different ethnic origin (European, South Asian, African, Middle Eastern and South American) attending three public Child Health Clinics for antenatal care in the multi-ethnic area of Groruddalen, Oslo, Norway, 2008–2010 [30]. Briefly, women were eligible if they: 1) Lived in the study districts; 2) planned to give birth at one of two study hospitals; 3) were <20 weeks pregnant; 4) could communicate in Norwegian or any of the eight translated languages; and 5) were able to give an informed consent. Women with pre-gestational diabetes, or in need of intensive hospital follow-up during pregnancy, were excluded. The overall participation rate in STORK G was 74%, 81.5% for Europeans and 73% for South Asians [30].

Ethical approval

The STORK G study including genetic and epigenetic data is approved by the Norwegian Regional Committee for Medical Health Research Ethics South East (ref.number 2015/1035). We obtained written informed consent from all participants before any study-related procedure.

Data collection

Questionnaire data and anthropometrics. In STORK G, interviewer-administered questionnaires were completed at gestational week 15±3 (visit 1) and 28±2 during pregnancy (visit 2), and 12±2 weeks postpartum (visit 3) [30,33]. Questionnaire data included information on mother's general health, physical activity and a dietary habits, in addition to some information about the father. Details about the questions used and data gathered have been described previously [30], and are available upon request to the corresponding.

Ethnic origin was defined by either the individual's country of birth or their mother's country birth, if the latter was born outside Europe [34]. We have detailed data on parity, pre-pregnant BMI, smoking status, alcohol intake, education, marital status and diet [30,33,35]. At all the three visits, we measured maternal height, body weight, fat mass with bioelectrical impedance (Tania-Weight BC-418 MA), skinfold thickness at three sites (Holtain T/W Skinfold Caliper, Holtain Ltd., Crymych) [31] and systolic and diastolic blood pressure (Omron HEM-7000-E M6 Comfort) [36].

Universal screening for gestational diabetes. All women underwent a 75 g oral glucose tolerance test at gestational week 28 ±2. Fasting and 2-hour glucose were analysed with a point-of-care instrument (HemoCue, Angelholm, Sweden). Women were diagnosed with gestational diabetes based on the WHO 1999 criteria (fasting glucose $\geq$ 7.0 mmol/l or 2-hour glucose $\geq$ 7.8 mmol/l) [37]. Furthermore, in retrospect and exclusively for research purposes, we also classified the samples using a slightly modified version of the WHO 2013 criteria (fasting glucose $\geq$ 5.1–6.9 mmol/l or 2-hour glucose $\geq$ 8.5–11 mmol/l, no data for 1-hour glucose) [38].

Laboratory data. Venous blood was drawn at the three visits into tubes with ethylenediaminetetraacetic acid (EDTA). Subsequently, the samples were aliquoted and biobanked or subject to routine laboratory analyses that were performed continuously during the study period. Fasting glucose, total Cholesterol, LDL-Cholesterol, HDL-Cholesterol and triglycerides levels were measured with a colorimetric method (Vitros 5.1 fs, Ortho clinical

diagnostics) [35], HbA1c levels were assessed in full blood with HPLC (Tosoh G8) [34], fasting C-peptide and insulin were measured at the Hormone Laboratory, Oslo University Hospital, with non-competing immunoflurometric assays (DELFIA, PerkinElmer Life Sciences, Wallac Oy, Turku, Finland) [39]. Serum 25(OH)D was analysed by competitive RIA (Dia-Sorin) at visit 1 and visit 2 [40], and S-leptin was analyzed by HADCYMAG-61K based on Luminex® xMAP® technology [32], at the Hormone Laboratory, Oslo University Hospital. Serum vitamin B12 and folate were measured with Electrochemiluminescence (ECLIA) assays, Roche, at Medical Biochemistry, Oslo University Hospital. HOMA-IR and HOMA-B [39,41] were estimated by Oxford University HOMA Calculator 2.2 using fasting glucose and C-peptide.

Objectively measured physical activity. Physical activity (PA) was objectively measured from visit 1 to 3 using SenseWear™ Pro3 Armband (SWA) (BodyMedia Inc, Pittsbur, PA, USA) [42]. Data from women with at least one valid day (defined as $\geq$ 19.2h) were considered valid [43]. Physical activity was characterized as Sedentary behaviour ($<$ 1.5 metabolic equivalents (METs)) light intensity (1.5 to $<$3 METS) or moderate or intense ($\geq$ 3 METs) [42,44].

Offspring data. The STORK G study also collected abdominal circumference, head circumference, bi-parietal diameter, femur length and estimated fetal weight by ultrasound measured on three different time points during pregnancy, and gestational age at birth derived from the first day of the woman's last menstrual period [45]. We have detailed anthropometric measurements at birth such as birthweight, head circumference, abdominal circumference, crown-heel length and neonatal skinfolds measured with a Holtain T/W Skinfold Caliper (Holtain Ltd., Crymych) [46]. Measurements of weight and length/height were collected during routine follow-up at the Mother—and Child Health Clinics when the children were 6 weeks old and thereafter at the 3, 6, 12, 15, 24 and 48 months visits. Further register-based follow-up is planned.

Venous serum cord blood samples were collected at birth and stored at -80˚C. Several sections from the placenta and umbilical cord have been sampled, and stored as Formalin-Fixed Paraffin-Embedded (FFPE) blocks. Currently, an ongoing pilot study of 80 FFPE placentas (40 South Asian, 40 European) demonstrate that enough DNA can be extracted from the FFPE to successfully run pyrosequencing (Sletner, unpublished). Furthermore, we have frozen placenta biopsies of about 1/3 of the women's offspring.

DNA extraction. In STORK G, at gestational week 28±2, DNA from peripheral blood leukocytes was extracted continuously throughout the data collection, at the Hormone Laboratory, Oslo University Hospital, using a salting out procedure [47], and stored at -80˚C.

Genetic data. The samples were genotyped using the Illumina CoreExome chip, by the Department of Clinical Sciences, Clinical Research Centre, Lund University, Malmö, Sweden [48]. Of the 664 genotyped samples those with low call rate (i.e. $<$ 95%, n = 0), extreme heterozygosity ($>$ |mean± (3xSD)|, n = 1), mismatched gender (n = 24) or cryptic relatedness (i.e. one individual (chosen at random) from each related pair, defined as genome-wide Identity by descent (IBD) $>$ 0.185 (n = 6) were excluded from analyses. Genetic ethnic origin was defined by ancestry informative principal component analysis based on the variance-standardized relationship matrix generated in PLINK 1.9 software package [49] (https://www.cog-genomics. org/plink/1.9/)).

Variants with call rate <95% (10081 SNPs), out of Hardy-Weinberg equilibrium (exact $p<10^{-6}$, 1971 SNPs) or with low minor allele frequency (MAF) <1% (245221 SNPs) were removed before imputation. Quality control was performed using the PLINK 1.9 software package [49]. After quality control, 293914 variants were left for imputation.

Imputation in European and South Asian samples was performed as follows: The GWAS scaffold was mapped to NCBI build 37 of the human genome. Imputation to the 1000G reference panel (Phase3- http://www.well.ox.ac.uk/~wrayner/tools/) was performed separately in Europeans and South Asians using their respective panels from the 1000G. The populations that are included in the European panel are Utah residents with Northern and Western European ancestry, Iberian populations in Spain, Finish in Finland, British in England and Scotland, and Tuscany in Italy. In the South Asian panel the included populations are Bengali in Bangladesh, Gujarati Indian in Houston, Texas, Indian Telugu in the UK, Punjabi in Lahore, Pakistan and Sri Lankan Tamil in the UK. The software used for imputation was IMPUTE (version 2.3.2) [50].

Epigenome-wide DNA methylation. Europeans and South Asians were the largest ethnic groups in STORK G, and South Asians of special interest due to their higher weight retention after pregnancy, increased prevalence of gestational diabetes, and increased risk for later type 2 diabetes compared to Europeans [32]. In EPIPREG, we quantified DNA methylation in maternal peripheral blood leukocytes in gestational week 28±1.2 in all Europeans (n = 312) and South Asians (n = 168) participating in STORK G who were genotyped and had fasting glucose data recorded. DNA samples were bisulfite converted using EZ DNA MethylationTM Kit (Zymo Research, Tustin, CA, USA) before added onto Infinium MethylationEPIC BeadChip (Illumina, San Diego, CA, USA) at the Department of Clinical Sciences, Clinical Research Centre, Lund University, Malmö, Sweden. Raw signal intensities of each probe were extracted using Illumina's GenomeStudio Software. The methylation level at each site was represented as a beta (β) value of the fluorescent intensity radio ranging from 0 (not methylated) to 1 (completely methylated). Meffil R package [51] (https://cran.r-project.org/) was used for quality control, normalization and reporting of beta values. During QC, we removed 8 samples: 1 due to sex mismatch (predicted sex outliers > 5SD), 1 outlier in control probes bisulfite 1 and bisulfite 2 (>5 SD), and 6 outliers from the methylated/unmethylated ratio comparison (>3 SD). Furthermore, 1299 probes with a detection p-value <0.01, and bead count <3 were removed. We used functional normalization, adjusting for effects of different batches, plates, columns and rows. A total of 307 European and 165 South Asian women and 864 560 probes of the array passed the QC.

Pyrosequencing. Random samples of 30 women were selected for technical validation of four CpGs sites by bisulfite pyrosequencing. The four CpG sites were chosen from preliminary top associations with fasting glucose (cg08098128, cg14120215), 2-hour glucose (cg19327414) and BMI (cg17148978). DNA samples were first bisulphite converted per QIAGEN Bisulfite conversion protocol [52] using 500 ng of DNA. Short DNA sequences that contained the CpG site of interest were amplified by PCR using PyroMark PCR kit from QIAGEN following the manufacturer instructions [53]. The PCR was performed on the 30 samples in duplicates with two positive controls, one was unmethylated converted DNA and the other was fully methylated converted DNA, both controls were commercially available from the EpiTect PCR Control DNA Set (100) by QIAGEN, and were used per the provider instructions. Also, a negative control only containing RNAse free water was used. Pyrosequecing was performed using the PyroMark Q48 Autoprep per the instructions provided in the user manual [54]. We used Bland-Altman plots to evaluate the agreement between pyrosequencing and the Infinium MethylationEPIC BeadChip, which have been used previously to assess agreement for this type of data [55,56]. To asses if there were proportional bias, we regressed the mean difference with the mean between methods. We also performed Pearson correlations between methods per CpG site and we followed a

previous published approach which consist in pooling all the CpG sites for the correlation analysis [57].

Follow- up study of the women 10–12 years after delivery. A 10-year follow-up of the women who attended STORK G is currently ongoing and expected to finish in 2021. The main aims are to assess the incidence of prediabetes and T2D and explore changes in risk factors for T2D (and CVD) over the last 10 years. We measure weight, height, physical inactivity, and blood pressure, and collect data on self-reported smoking, and chronic diseases/conditions. Currently, dried blood spots are biobanked and about 60% meet for fasting blood samples, and DNA from buccal swabs is being collected as well. We estimate to reach a sample of 350 women– 50% of those eligible.

Results to date

In the EPIPREG sample, we excluded women without fasting plasma glucose available in week 28±2, and those without genotype data due to low DNA concentrations or problems with DNA extraction. For Europeans, we were able to quantify DNA methylation in 99% of the eligible samples (empty wells = 2 and full plates = 1), 87.2% of the total number of Europeans participating in STORK G at week 28±2. For South Asians, methylation status could be determined in 100% of the eligible subjects, representing 87.5% of total South Asians participating in STORK G at week 28±2 (Fig 1). Hence, EPIPREG resulted in DNA methylation data of 312 Europeans and 168 South Asians.

Some characteristics of the mothers and offspring of the EPIPREG sample are shown in Table 1. When comparing the clinical characteristics of the women and their offspring with and without DNA methylation data, the average of sedentary hours and percentage of truncal fat were significantly higher in the European subjects included in EPIPREG, whereas gestational week, hours of light and moderate-intense physical activity, and 25-hydroxyvitamin D levels were higher in the excluded samples (S1 Table). The differences were small and generally followed the same trend as the overall STORK G missing data analysis [34]. In South Asians we did not detect any significant differences between the women and their offspring included in EPIPREG vs the excluded individuals (S2 Table).

Genetic principal components distinctly separated Europeans and South Asian women (Fig 2), which fully corresponded with the self-reported ethnicity. South Asians were separated in two groups, the largest cluster mainly consisted of Pakistani women and the smaller group of Sri-Lankan women.

To assess the agreement between the Infinium MethylationEPIC BeadChip and the technical validation by pyrosequencing, we used Bland-Altman plots to illustrate agreement per CpG site (Fig 3). In Fig 3, we can see that most of the samples in cg17148978, cg14120215 and cg02098128 were within the limits of agreement (LAG). For cg17148978 the average mean difference was 0.74% (lower LAG:-0.56, upper LAG: 2.05%), for cg14120215 7.56% (lower LAG: -6.95, upper LAG: -22.09%) and for cg02098128–0.66% (lower LAG:-2.09, upper LAG 0.78%). We found no evidence of proportional bias in cg17148978, cg14120215 or cg02098128, meaning that the agreement in these CpG sites was good. In cg19327414, the mean difference was 7.79% (lower LAG: -12.32, upper LAG: 27.89%), and there was a significant proportional bias (Beta = -1.85, pval = <0.001). All the samples but one were within the lines of agreement. In the correlation analysis, the overall correlation when pooling all 4 CpG sites was high (r = 0.98, p<0.001) (Fig 4), while site specific correlations for cg17148978 (R = -0.23, p = 0.22), cg19327414 (R = 0.18, p = 0.40), cg14120215 (R = 0.18, p = 0.33) and cg02098128 (R = 0.30, p = 0.11) were weak.
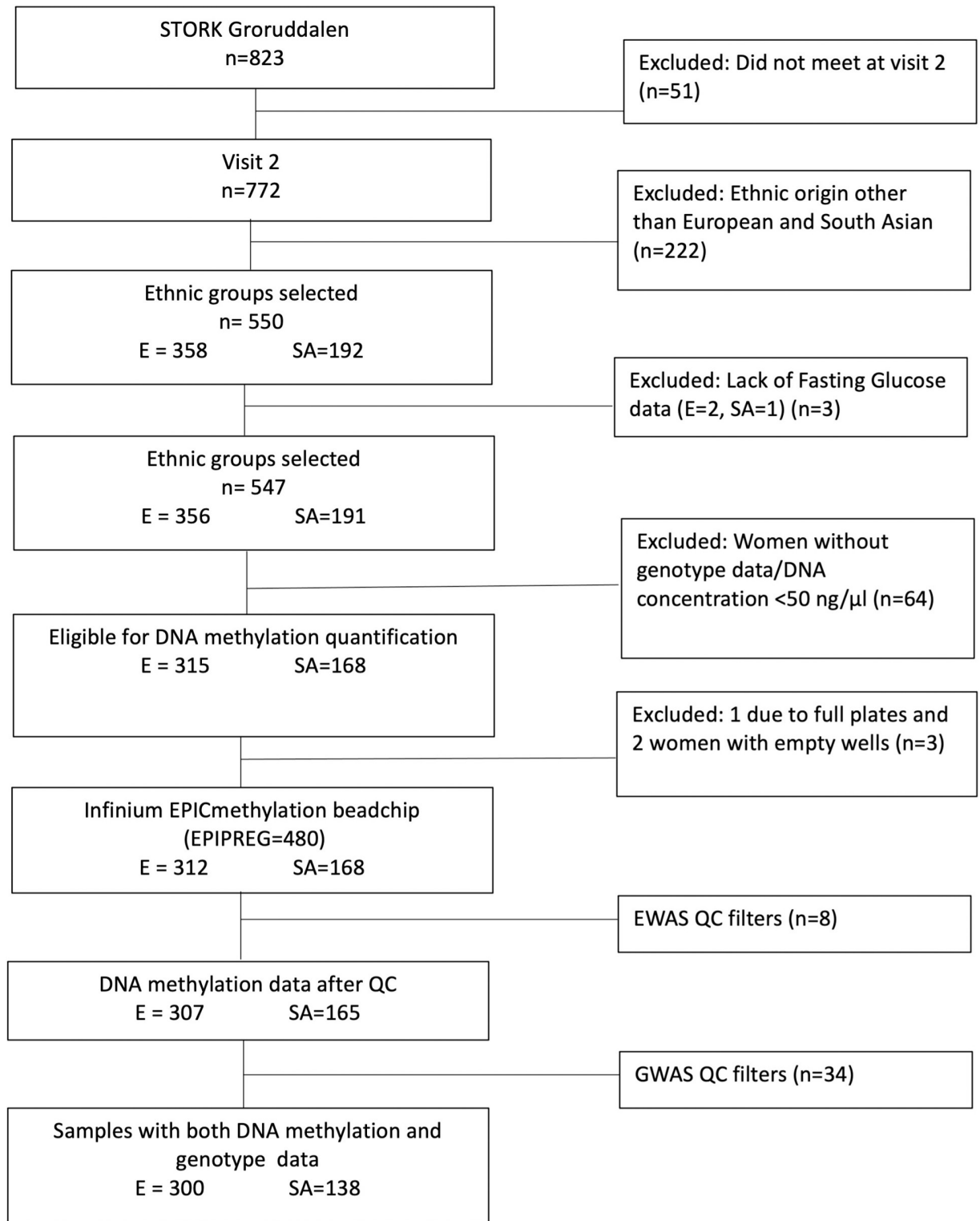
Fig 1. Flow chart of the EPIPREG sample. E = European, SA = South Asian.

https://doi.org/10.1371/journal.pone.0256158.g001

Table 1. Characteristics of the EPIPREG sample.

Variable	N	Europeans, n = 312	South Asians, n = 168
Age	480	30.1 (4.6)	28.2 (4.6)
Weeks' gestation	480	28.1 (1.2)	28.2 (1.2)
Height (cm)	480	167.4 (5.8)	159.9 (5.7)
Smoking status	476		
Current		20 (6.5)	1 (0.6)
3 months pre-pregnancy		80 (25.9)	2 (1.2)
Former*		87 (28.2)	9 (5.4)
Never		122 (39.5)	155 (92.8)
Pre-pregnancy alcohol intake, n (%)	472	234 (76.2)	5 (3.0)
Pre-pregnancy BMI (kg/m^2)	474	24.6 (4.9)	23.8 (4.1)
Actual BMI (kg/m^2)	478	27.8 (4.7)	26.8 (4.1)
Total fat (%)	465	29.6 (9.7)	26.2 (8.2)
Truncal fat (%)	465	15.9 (5.5)	14.0 (5.4)
Systolic blood pressure (mmHg)	480	107.0 (9.6)	101.1 (8.7)
Diastolic blood pressure (mmHg)	480	68.4 (7.1)	66.1 (6.9)
First degree relative with diabetes	474	43 (14.0)	79 (47.3)
GDM (WHO$_{2013}$), n (%)	480	76 (24.4)	70 (41.7)
GDM (WHO$_{1999}$), n (%)	478	37 (11.9)	25 (15.1)
Fasting glucose (mmol/L)	480	4.7 (0.6)	5.0 (0.6)
2 hour glucose (mmol/L)	477	6.0 (1.4)	6.4 (1.5)
HbA1c (%)	475	5.1 (0.3)	5.3 (0.3)
Fasting Insulin (pmol/L)	474	48.0 [33.0, 70.2]	71.0 [57.0, 100.5]
Fasting C-peptide (pmol/L)	474	712.0 [560.2, 901.8]	855.5 [688.0, 1067.5]
HOMA-B	474	173.2 [151.3, 199.7]	179.1 [154.9, 207.9]
HOMA-IR	474	1.5 [1.2, 1.9]	1.8 [1.4, 2.3]
Total cholesterol (mmol/L)	480	6.4 (1.1)	6.0 (1.0)
HDL-cholesterol (mmol/L)	480	1.9 (0.4)	1.9 (0.4)
LDL-cholesterol (mmol/L)	473	3.7 (1.0)	3.3 (0.9)
Triglycerides (mmol/L)	480	2.0 (0.7)	2.0 (0.
Folate (nmol/l)	471	13.0 [9.9, 18.0]	12.0 [9.4, 16.0]
Vitamin D (µmol/L)	472	69.9 (27.8)	45.6 (22.2)
Leptin (µmol/L)	476	1599.5 [966.9, 2497.7]	2276.9 [1532.8, 3295.8]
Folate (nmol/l)	471	13.0 [9.9, 18.0]	12.0 [9.4, 16.0]
Vitamin B12 (µmol/L)	472	211.0 [169.2, 253.0]	186.0 [150.0, 232.2]
Sedentary time (hours/day)**	382	18 (1.6)	17.8 (1.7)
Light physical activity (hours/day)**	382	4.3 (1.3)	4.6 (1.3)
Moderate-intense physical activity (hours/day)**	382	1.0 [0.7, 1.5]	0.8 [0.5, 1.2]
Offspring data			
Gestational age (days)	475	280.9 (11.7)	277.1 (12.6)
Female sex, n (%)	462	147 (49.0)	79 (48.8)
Birth weight (g)	472	3582.2 (529.4)	3211.7 (512.2)
Birth length (cm)	432	50.1 (2.3)	49.3 (2.2)
Neonatal sum of skinfolds (mm)	359	18.8 (4.0)	17.0 (3.6)

Data from gestational visit 2 for cross-sectional associations with DNA methylation data, otherwise specified. Data are presented in mean (SD) for normally distributed variables and median [IQR] for non-normal variables. Categorical variables are presented by frequency (%). Despite 8 individuals did not pass the QC procedure, the data of the 480 individuals are presented for informative purposes. WHO = World health organization.

GDM = gestational diabetes mellitus. BMI = body mass index. HDL = high-density lipoproteins. LDL = low-density lipoproteins.

* Ex-smokers and occasional smokers who did not smoke 3 months before pregnancy,

**Women with at least one valid day of registered physical activity (Armband).
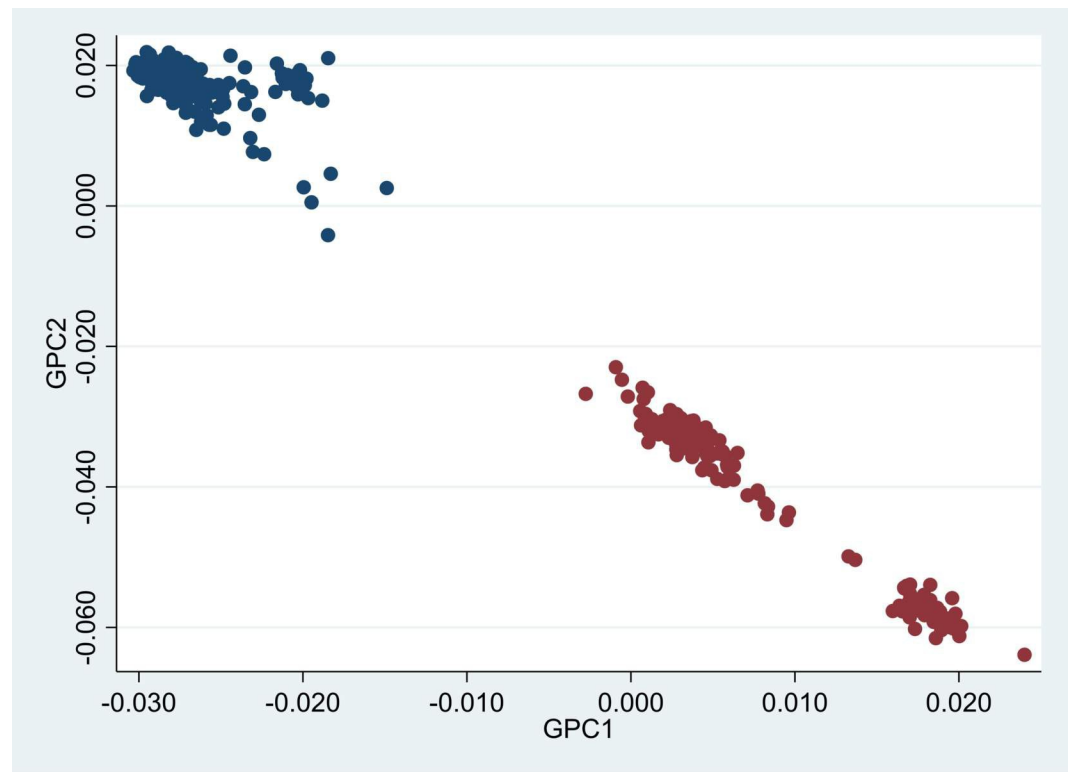
Fig 2. Scatter dot plot of genetic PC1 (GPC1) and PC2 (GPC2) (n = 438). Blue dots are Europeans, red dots are South Asians, based on self-reported ethnicity. It can be noticed that South Asians were separated in two groups, being the upper largest group mainly composed of Pakistanis and the smaller lower group of Sir Lankans.

https://doi.org/10.1371/journal.pone.0256158.g002

Strengths and limitations

EPIPREG is a population-based sample of European and South Asian pregnant women with epigenome-wide DNA methylation in maternal peripheral blood leukocytes. EPIPREG has detailed phenotype data from both the mother and the offspring, as well as genotype data. DNA methylation measured with the epigenome-wide chip showed high agreement with bisulphite pyrosequencing.

The inclusion of women with both European and South Asian ethnic background enables interesting studies into the role of DNA methylation in ethnic disparities in health. Furthermore, EPIPREG has both genome-wide genotype and DNA methylation data also allowing for methylation quantitative trait loci (mQTL) analysis.

Regarding the technical validation, correlations were low for each CpG site separately. However, correlations could be misleading for agreement analyses as they mainly measure the linearity of the variables irrespective of the data's shape [58], and are sensitive to the range of values—the broader the range, the higher the correlation coefficient [59]. Bland-Altman plots are considered a better test of agreement between methods [59], especially when the range of values is low as for three of our four CpG sites. Therefore although our individual sites showed low correlation between the two methods, they had high agreement.

EPIPREG's population-based design and comprehensive phenotyping allows for gaining representative data about the associations between DNA methylation and a wide range of phenotypic traits, exposures and outcomes. Hence, a major advantage of the cohort is the availability of several maternal phenotypes collected in gestational weeks 15 and 28, and 3 months
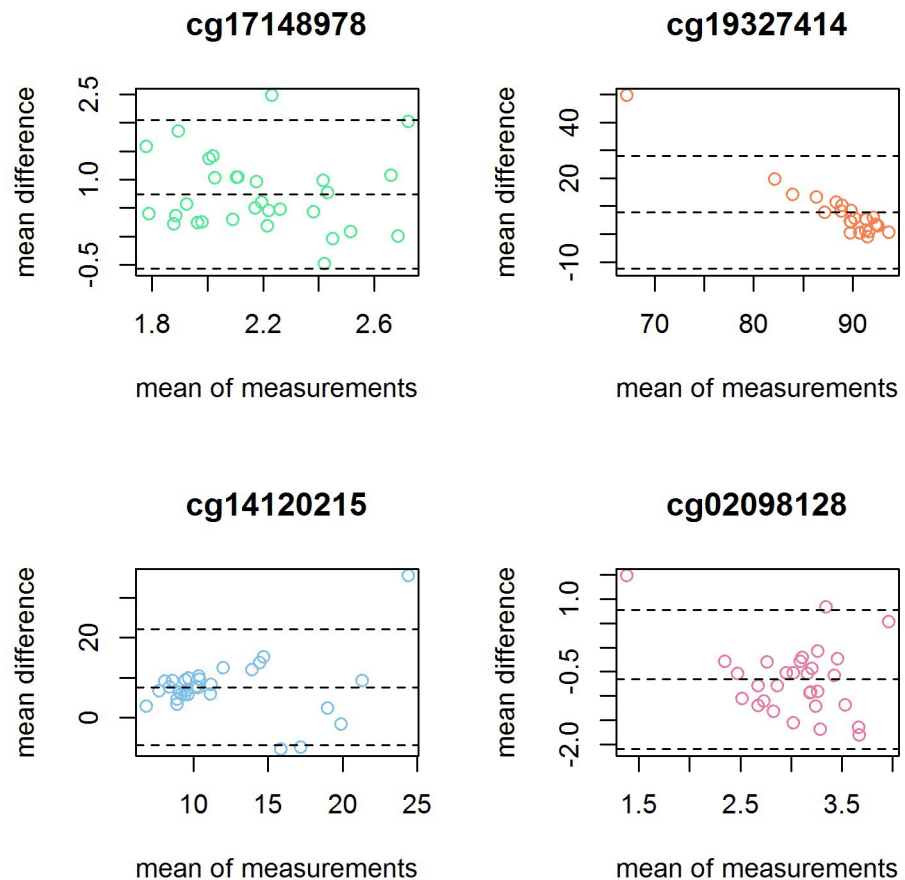
Fig 3. Bland-Altman plots showing the mean difference between methods (y-axis) versus the mean between methods (x-axis) for each CpG site tested for technical validation.

https://doi.org/10.1371/journal.pone.0256158.g003

after delivery. Furthermore, there is an ongoing follow-up in some women 10 to 12 years after pregnancy. In the offspring we have anthropometric data recorded in utero, at birth and during the first four years of life, as well as serum cord blood and placental tissue biobanked. Lastly, we have permission for linkage with Norwegian national registries using the personal identification number.

Since EPIPREG has a moderate sample size, our study has limited statistical power for EWAS, nevertheless our broad availability of phenotypes will allow us to perform several DNA methylation-phenotype association analyses. Also, our sample is well suited for meta-analysis efforts, or to serve as a replication cohort. Another limitation is that DNA methylation is only measured in gestational week 28±2.

Collaboration

EPIPREG may serve as a useful sample for generation of new hypotheses about associations between DNA-methylation and phenotypic traits relating to GDM, for replication of findings from other studies, and for meta-analysis efforts. We are currently welcoming collaborations with cohorts with similar data and researchers interested in collaboration are welcome to contact Christine Sommer, or visit our website: www.epipreg.no.
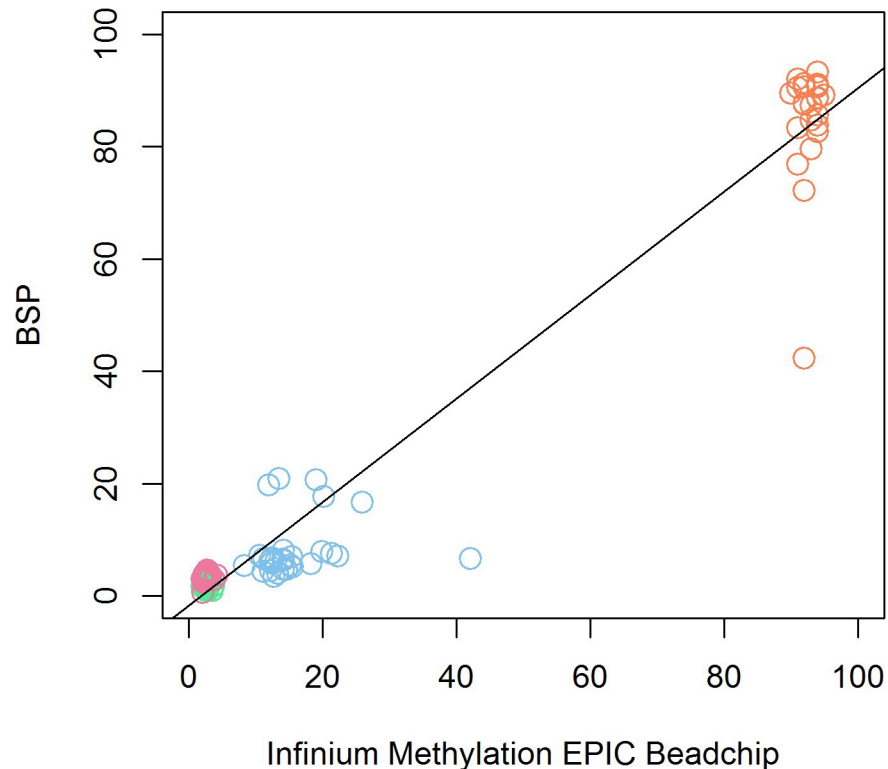
Fig 4. Scatter plot showing the relationship between the DNAm values of the four selected CpG sites quantified with the Infinium MethylationEPIC BeadChip (x-axis) and with bisulphite pyrosequencing (BSP) (y-axis). Dots colour code: Pink: cg02098128, green: cg17148978, blue: cg14120215, orange: cg19327414.

https://doi.org/10.1371/journal.pone.0256158.g004

Supporting information

S1 Table. Mean comparison between the European samples included in the Infinium MethylationEPIC BeadChip versus their respective excluded samples.
(XLSX)

S2 Table. Mean comparison between the South Asian samples included in the Infinium MethylationEPIC BeadChip versus their respective excluded samples.
(XLSX)

Acknowledgments

We would like to thank the women who participated in the STORK Groruddalen study, Maria Sterner, Malin Neptin, and Gabriella Gremsperger at the Genomics Diabetes and Endocrinology CRC, Malmö, for experiments.

Author Contributions

Conceptualization: Anne Karen Jenum, Kåre I. Birkeland, Christine Sommer.

Data curation: Sindre Lee-Ødegård, Kåre Rønn Richardsen, Line Sletner.

Formal analysis: Nicolas Fragoso-Bargas, Julia O. Opsahl, Nadezhda Kiryushchenko, Gunn-Helen Moen, Christine Sommer.

References

1. Conerly M, Grady WM. Insights into the role of DNA methylation in disease through the use of mouse models. Dis Model Mech. 2010; 3(5–6):290–7. https://doi.org/10.1242/dmm.004812 PMID: 20427558

2. Yong WS, Hsu FM, Chen PY. Profiling genome-wide DNA methylation. Epigenet Chromatin. 2016; 9. https://doi.org/10.1186/s13072-016-0075-3 PMID: 27358654

3. Moore LD, Le T, Fan GP. DNA Methylation and Its Basic Function. Neuropsychopharmacol. 2013; 38 (1):23–38. https://doi.org/10.1038/npp.2012.112 PMID: 22781841

4. Demerath EW, Guan WH, Grove ML, Aslibekyan S, Mendelson M, Zhou YH, et al. Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. Hum Mol Genet. 2015; 24(15):4464–79. https://doi.org/10.1093/hmg/ddv161 PMID: 25935004

5. Opsahl JO, Moen G-H, Qvigstad E, Böttcher Y, Birkeland KI, Sommer C. Epigenetic signatures associated with maternal body mass index or gestational weight gain: a systematic review. Journal of Developmental Origins of Health and Disease. 2020:1–11. https://doi.org/10.1017/S2040174420000811 PMID: 32873364

6. Sun DJY, Zhang T, Su SY, Hao G, Chen T, Li QZ, et al. Body Mass Index Drives Changes in DNA Methylation A Longitudinal Study. Circ Res. 2019; 125(9):824–33. https://doi.org/10.1161/CIRCRESAHA.119.315397 PMID: 31510868

7. Wahl S, Drong A, Lehne B, Loh M, Scott WR, Kunze S, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. Nature. 2017; 541(7635):81–+. https://doi.org/10.1038/nature20784 PMID: 28002404

8. Canouil M, Khamis A, Keikkala E, Hummel S, Lobbens S, Bonnefond A, et al. Epigenome-Wide Association Study Reveals Methylation Loci Associated With Offspring Gestational Diabetes Mellitus Exposure and Maternal Methylome. Diabetes Care. 2021. https://doi.org/10.2337/dc20-2960 PMID: 34116986

9. Florath I, Butterbach K, Heiss J, Bewerunge-Hudler M, Zhang Y, Schottker B, et al. Type 2 diabetes and leucocyte DNA methylation: an epigenome-wide association study in over 1,500 older adults. Diabetologia. 2016; 59(1):130–8. https://doi.org/10.1007/s00125-015-3773-7 PMID: 26433941

10. Cardona A, Day FR, Perry JRB, Loh M, Chu ARY, Lehne B, et al. Epigenome-Wide Association Study of Incident Type 2 Diabetes in a British Population: EPIC-Norfolk Study. Diabetes. 2019; 68(12):2315–26. https://doi.org/10.2337/db18-0290 PMID: 31506343

11. Al Muftah WA, Al-Shafai M, Zaghlool SB, Visconti A, Tsai PC, Kumar P, et al. Epigenetic associations of type 2 diabetes and BMI in an Arab population. Clin Epigenetics. 2016; 8.

12. Juvinao-Quintero DL, Marioni RE, Ochoa-Rosales C, Russ TC, Deary IJ, van Meurs JBJ, et al. DNA methylation of blood cells is associated with prevalent type 2 diabetes in a meta-analysis of four European cohorts. Clin Epigenetics. 2021; 13(1).

13. Chambers JC, Loh M, Lehne B, Drong A, Kriebel J, Motta V, et al. Epigenome-wide association of DNA methylation markers in peripheral blood from Indian Asians and Europeans with incident type 2 diabetes: a nested case-control study. Lancet Diabetes Endo. 2015; 3(7):526–34. https://doi.org/10.1016/S2213-8587(15)00127-8 PMID: 26095709

14. Xie T, Gorenjak V, Stathopoulou MG, Dade S, Marouli E, Masson C, et al. Epigenome-Wide Association Study (EWAS) of Blood Lipids in Healthy Population from STANISLAS Family Study (SFS). Int J Mol Sci. 2019; 20(5).

15. Braun KVE, Dhana K, de Vries PS, Voortman T, van Meurs JBJ, Uitterlinden AG, et al. Epigenome-wide association study (EWAS) on lipids: the Rotterdam Study. Clin Epigenetics. 2017; 9. https://doi.org/10.1186/s13148-016-0304-4 PMID: 28194238

16. Kazmi N, Elliott HR, Burrows K, Tillin T, Hughes AD, Chaturvedi N, et al. Associations between high blood pressure and DNA methylation. Plos One. 2020; 15(1). https://doi.org/10.1371/journal.pone.0227728 PMID: 31999706

17. Richard MA, Huan TX, Ligthart S, Gondalia R, Jhun MA, Brody JA, et al. DNA Methylation Analysis Identifies Loci for Blood Pressure Regulation. Am J Hum Genet. 2017; 101(6):888–902. https://doi.org/10.1016/j.ajhg.2017.09.028 PMID: 29198723

18. Ambatipudi S, Cuenin C, Hernandez-Vargas H, Ghantous A, Le Calvez-Kelm F, Kaaks R, et al. Tobacco smoking-associated genome-wide DNA methylation changes in the EPIC study. Epigenomics-Uk. 2016; 8(5):599–618. https://doi.org/10.2217/epi-2016-0001 PMID: 26864933

19. Dugue PA, Jung CH, Joo JE, Wang XC, Wong EM, Makalic E, et al. Smoking and blood DNA methylation: an epigenome-wide association study and assessment of reversibility. Epigenetics-Us. 2020; 15 (4). https://doi.org/10.1080/15592294.2019.1668739 PMID: 31552803

20. Wiklund P, Karhunen V, Richmond RC, Parmar P, Rodriguez A, De Silva M, et al. DNA methylation links prenatal smoking exposure to later life health outcomes in offspring. Clin Epigenetics. 2019;11. https://doi.org/10.1186/s13148-018-0598-5 PMID: 30654849

21. Joubert BR. 450K Epigenome-Wide Scan Identifies Differential DNA Methylation in Newborns Related to Maternal Smoking During Pregnancy (vol 120, pg 1425, 2012). Environ Health Persp. 2012; 120(12): A455–A.

22. Joehanes R, Just AC, Marioni RE, Pilling LC, Reynolds LM, Mandaviya PR, et al. Epigenetic Signatures of Cigarette Smoking. Circ-Cardiovasc Gene. 2016; 9(5):436–47. https://doi.org/10.1161/CIRCGENETICS.116.001506 PMID: 27651444

23. Bruckmann C, Islam SA, MacIsaac JL, Morin AM, Karle KN, Di Santo A, et al. DNA methylation signatures of chronic alcohol dependence in purified CD3(+) T-cells of patients undergoing alcohol treatment. Sci Rep-Uk. 2017; 7. https://doi.org/10.1038/s41598-017-06847-z PMID: 28747766

24. Christensen BC, Kelsey KT, Zheng SC, Houseman EA, Marsit CJ, Wrensch MR, et al. Breast Cancer DNA Methylation Profiles Are Associated with Tumor Size and Alcohol and Folate Intake. Plos Genet. 2010; 6(7). https://doi.org/10.1371/journal.pgen.1001043 PMID: 20686660

25. Dugue PA, Wilson R, Lehne B, Jayasekara H, Wang XC, Jung CH, et al. Alcohol consumption is associated with widespread changes in blood DNA methylation: Analysis of cross-sectional and longitudinal data. Addict Biol. 2021; 26(1). https://doi.org/10.1111/adb.12855 PMID: 31789449

26. Williams D. Pregnancy: a stress test for life. Curr Opin Obstet Gyn. 2003; 15(6):465–71. https://doi.org/10.1097/00001703-200312000-00002 PMID: 14624211

27. Catalano PM, Tyzbir ED, Roman NM, Amini SB, Sims EAH. Longitudinal Changes in Insulin Release and Insulin Resistance in Nonobese Pregnant-Women. Am J Obstet Gynecol. 1991; 165(6):1667–72. https://doi.org/10.1016/0002-9378(91)90012-g PMID: 1750458

28. Buchanan TA, Xiang AH. Gestational diabetes mellitus. J Clin Invest. 2005; 115(3):485–91. https://doi.org/10.1172/JCI24531 PMID: 15765129

29. Bilhartz TD, Bilhartz PA, Bilhartz TN, Bilhartz RD. Making Use of a Natural Stress Test: Pregnancy and Cardiovascular Risk. J Womens Health. 2011; 20(5):695–701. https://doi.org/10.1089/jwh.2010.2291 PMID: 21453034

30. Jenum AK, Sletner L, Voldner N, Vangen S, Morkrid K, Andersen LF, et al. The STORK Groruddalen research programme: A population-based cohort study of gestational diabetes, physical activity, and obesity in pregnancy in a multiethnic population. Rationale, methods, study population, and participation rates. Scand J Public Health. 2010; 38:60–70.

31. Sommer C, Mørkrid K, Jenum AK, Sletner L, Mosdøl A, Birkeland KI. Weight gain, total fat gain and regional fat gain during pregnancy and the association with gestational diabetes: a population-based cohort study. Int J Obes (Lond). 2014; 38(1):76–81. https://doi.org/10.1038/ijo.2013.185 PMID: 24051503

32. Sommer C, Jenum AK, Waage CW, Morkrid K, Sletner L, Birkeland KI. Ethnic differences in BMI, subcutaneous fat, and serum leptin levels during and after pregnancy and risk of gestational diabetes. Eur J Endocrinol. 2015; 172(6):649–56. https://doi.org/10.1530/EJE-15-0060 PMID: 25740849

33. Sommer C, Sletner L, Jenum AK, Morkrid K, Andersen LF, Birkeland KI, et al. Ethnic differences in maternal dietary patterns are largely explained by socio-economic score and integration score: a population-based study. Food Nutr Res. 2013; 57. https://doi.org/10.3402/fnr.v57i0.21164 PMID: 23843779

34. Jenum AK, Morkrid K, Sletner L, Vangen S, Torper JL, Nakstad B, et al. Impact of ethnicity on gestational diabetes identified with the WHO and the modified International Association of Diabetes and Pregnancy Study Groups criteria: a population-based cohort study (vol 166, pg 317, 2012). Eur J Endocrinol. 2012; 166(3):561.

35. Sommer C, Sletner L, Morkrid K, Jenum AK, Birkeland KI. Effects of early pregnancy BMI, mid-gestational weight gain, glucose and lipid levels in pregnancy on offspring's birth weight and subcutaneous fat: a population-based cohort study. Bmc Pregnancy Childb. 2015;15. https://doi.org/10.1186/s12884-015-0512-5 PMID: 25879215

36. Waage CW, Mdala I, Jenum AK, Michelsen TM, Birkeland KI, Sletner L. Ethnic differences in blood pressure from early pregnancy to postpartum: a Norwegian cohort study. J Hypertens. 2016; 34 (6):1151–9. https://doi.org/10.1097/HJH.0000000000000918 PMID: 27088633

37. WHO. Definition and classification of diabetes mellitus and its complications. Report of a WHO consultation. Part 1: Diagnosis and classification of diabetes mellitus. WHO. 1999.

38. Agarwal MM, Boulvain M, Coetzee E, Colagiuri S, Falavigna M, Hod M, et al. Diagnostic criteria and classification of hyperglycaemia first detected in pregnancy: A World Health Organization Guideline. Diabetes Res Clin Pr. 2014; 103(3):341–63.

39. Morkrid K, Jenum AK, Sletner L, Vardal MH, Waage CW, Nakstad B, et al. Failure to increase insulin secretory capacity during pregnancy-induced insulin resistance is associated with ethnicity and gestational diabetes. Eur J Endocrinol. 2012; 167(4):579–88. https://doi.org/10.1530/EJE-12-0452 PMID: 22889687

40. Eggemoen AR, Jenum AK, Mdala I, Knutsen KV, Lagerlov P, Sletner L. Vitamin D levels during pregnancy and associations with birth weight and body composition of the newborn: a longitudinal multiethnic population-based study. Brit J Nutr. 2017; 117(7):985–93. https://doi.org/10.1017/S000711451700068X PMID: 28468694

41. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis Model Assessment—Insulin Resistance and Beta-Cell Function from Fasting Plasma-Glucose and Insulin Concentrations in Man. Diabetologia. 1985; 28(7):412–9. https://doi.org/10.1007/BF00280883 PMID: 3899825

42. Berntsen S, Richardsen KR, Morkrid K, Sletner L, Birkeland KI, Jenum AK. Objectively recorded physical activity in early pregnancy: A multiethnic population-based study. Scand J Med Sci Spor. 2014; 24 (3):594–601. https://doi.org/10.1111/sms.12034 PMID: 23278771

43. Morkrid K, Jenum AK, Berntsen S, Sletner L, Richardsen KR, Vangen S, et al. Objectively recorded physical activity and the association with gestational diabetes. Scand J Med Sci Spor. 2014; 24(5): E389–E97. https://doi.org/10.1111/sms.12183 PMID: 24894027

44. Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee IM, et al. Quantity and Quality of Exercise for Developing and Maintaining Cardiorespiratory, Musculoskeletal, and Neuromotor Fitness in Apparently Healthy Adults: Guidance for Prescribing Exercise. Med Sci Sport Exer. 2011; 43 (7):1334–59.

45. Sletner L, Rasmussen S, Jenum AK, Nakstad B, Jensen OHR, Vangen S. Ethnic differences in fetal size and growth in a multi-ethnic population. Early Human Development. 2015; 91(9):547–54. https://doi.org/10.1016/j.earlhumdev.2015.07.002 PMID: 26197025

46. Sletner L, Nakstad B, Yajnik CS, Mørkrid K, Vangen S, Vårdal MH, et al. Ethnic differences in neonatal body composition in a multi-ethnic population and the impact of parental factors: a population-based cohort study. Plos One. 2013; 8(8):e73058. https://doi.org/10.1371/journal.pone.0073058 PMID: 24009731

47. Miller SA, Dykes DD, Polesky HF. A Simple Salting out Procedure for Extracting DNA from Human Nucleated Cells. Nucleic Acids Res. 1988; 16(3):1215. https://doi.org/10.1093/nar/16.3.1215 PMID: 3344216

48. Moen GH, LeBlanc M, Sommer C, Prasad RB, Lekvas T, Normann KR, et al. Genetic determinants of glucose levels in pregnancy: genetic risk scores analysis and GWAS in the Norwegian STORK cohort. Eur J Endocrinol. 2018; 179(6):363–72. https://doi.org/10.1530/EJE-18-0478 PMID: 30324795

49. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira Manuel A R, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. American Journal of Human Genetics. 2007; 81(3):559–75. https://doi.org/10.1086/519795 PMID: 17701901

50. Howie BN, Donnelly P, Marchini J. A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. PLoS genetics. 2009; 5(6):e1000529. https://doi.org/10.1371/journal.pgen.1000529 PMID: 19543373

51. Min JL, Hemani G, Smith GD, Relton C, Suderman M. Meffil: efficient normalization and analysis of very large DNA methylation datasets. Bioinformatics. 2018; 34(23):3983–9. https://doi.org/10.1093/bioinformatics/bty476 PMID: 29931280

52. Qiagen. Epitec (R) Bisulfite handbook. 2014.

53. Qiagen. PyroMark PCR Handbook. 2009.

54. Qiagen. PyroMark (R) Q48 Autoprep User Manual. 2016.

55. Mamtani M, Kulkarni H, Dyer TD, Goring HHH, Neary JL, Cole SA, et al. Genome- and epigenome-wide association study of hypertriglyceridemic waist in Mexican American families. Clin Epigenetics. 2016;8. https://doi.org/10.1186/s13148-016-0174-9 PMID: 26807160

56. Roessler J, Ammerpohl O, Gutwein J, Hasemeier B, Anwar SL, Kreipe H, et al. Quantitative cross-validation and content analysis of the 450k DNA methylation array from Illumina, Inc. BMC Res Notes. 2012; 5:210. https://doi.org/10.1186/1756-0500-5-210 PMID: 22546179

57. Ronn T, Volkov P, Davegardh C, Dayeh T, Hall E, Olsson AH, et al. A Six Months Exercise Intervention Influences the Genome-wide DNA Methylation Pattern in Human Adipose Tissue. Plos Genet. 2013; 9(6). https://doi.org/10.1371/journal.pgen.1003572 PMID: 23825961

58. van Stralen KJ, Dekker FW, Zoccali C, Jager KJ. Measuring Agreement, More Complicated Than It Seems. Nephron Clin Pract. 2012; 120(3):C162–C7. https://doi.org/10.1159/000337798 PMID: 22722148

59. van Stralen KJ, Jager KJ, Zoccali C, Dekker FW. Agreement between methods. Kidney Int. 2008; 74(9):1116–20. https://doi.org/10.1038/ki.2008.306 PMID: 18596728