LifeCycle

Report on the associations of DNA methylation loci related to early-life stressors with functional outcomes, including RNA expression

LifeCycle report D8.4

Authors:

Mariona Bustamante (ISGlobal)

Janine F. Felix (ERASMUS)

Version 2.0

Delivery date: Month 66

LIFECYCLE Project
Coordination and leadership:

Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands

Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

LifeCycle

Document Information

Grant Agreement No.	**733206**
Project Title	Early-life stressors and LifeCycle health (The LifeCycle Project)
Project Start Date	01 January 2017
Work package title	WP8 – DNA methylation and gene expression in life course health trajectories
Related task(s)	Task 8.4
Lead Organisation	ISGlobal
Submission date	16 September 2022
Dissemination Level	Public

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

Table of Contents

List of Figures 3
List of Abbreviations 4
Executive summary 6
1. Introduction 7
2. Description of progress and results 9
2.1 Building an expression quantitative trait methylation (eQTM) catalogue 9
2.2 Functional interpretation of EWAS of early-life stressors 11
2.2.1 Environmental pollutants in pregnancy 11
2.2.2 Maternal diet in pregnancy 16
2.2.3 Maternal cardiometabolic traits in pregnancy 18
2.3 Functional interpretation of EWAS of health outcomes 20
2.4 Transcriptome-wide association studies (TWAS) 29
3. General discussion 30
3.1 Building expression quantitative trait methylation (eQTM) catalogues 30
3.2 Functional interpretation of EWAS of early-life stressors and of health outcomes . 31
3.3 Transcriptome-wide association studies (TWAS) 33
4. Conclusion 33
5. Contribution of partners 34
6. Deviations from original plan 35
7. Dissemination activities 35
8. References 35
Appendix 1 38
Appendix 2 39

List of Figures

Figure 1: Distance between CpG and TC's TSS and effect size in child blood autosomal cis-eQTMs
Figure 2: Genetic contribution to autosomal cis-eQTMs in children's blood
Figure 3: Circus plot showing the association between sustained maternal smoking in pregnancy and child blood DNA methylation and transcription along the chromosomes (outer circle)
Figure 4: Regional plot of the AHRR locus (50 kb upstream and downstream of the 5 CpG sites associated with any maternal smoking in pregnancy)
Figure 5: Enriched biological processes (p-value <0.05) in GO analysis of CpG sites with p-value <0.001 (n=644 CpG sites, n=396 genes) ordered by significance
Figure 6: Integrative pathway analysis of all genes annotated to DMRs associated with childhood FEV1
Figure 7: A network is shown for 4 categories of disease and biological functions overlapping between analyses of asthma in relation to either newborn or childhood methylation: immunological disease, cell-mediated immune response, inflammatory disease, and tissue morphology

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

List of Abbreviations

ABOS: Biological Atlas of Severe Obesity
ADHD: Attention Deficit and Hyperactivity Disorder
ALSPAC: Avon Longitudinal Study of Parents and Children
BAMSE: Swedish abbreviation for Children, Allergy, Milieu, Stockholm, Epidemiology
BEAR: Biomarkers of Exposure to ARsenic
BiB: Born in Bradford
BIOS: Biobank-based integrative omics study
BMI: Body mass index
CI: Confidence interval
COPD: Chronic obstructive pulmonary disease
CpG: Cytosine-phosphate-guanine site
DAVID: Database for Annotation, Visualization and Integrated Discovery
DMR: Differently methylated region
DNA: Deoxyribonucleic acid
EARLI: Early Autism Risk Longitudinal Investigation
EDEN: Étude des Déterminants pré et postnatals du développement et de la santé de l'enfant
IoW-3: Isle of Wight 3rd Generation Study
EWAS: Epigenome-wide association study
FDR: False discovery rate
FEV1: Forced expiratory volume in 1 second
FVC: Forced vital capacity
GWG: Gestational weight gain
GO: Gene Ontology
HDP: Hypertensive disorders of pregnancy
HELIX: Human Early Life Exposome project
Hg: Mercury
INMA: INfancia y Medio ambiente
IPA: Ingenuity Pathways Analysis
KEGG: Kyoto Encyclopedia of Genes and Genomes
MeHg: Methylmercury
MoBa: Norwegian Mother and Child Cohort Study
PACE: Pregnancy And Childhood Epigenetics
PE: Preeclampsia
PM: Particulate matter
QC: Quality control
rMED: Relative Mediterranean diet score
rMEDp: Relative Mediterranean diet score in pregnancy, adjusted rMED excluding alcohol
RNA: Ribonucleic acid
Rhea: The Rhea Mother-Child Study in Crete
SD: Standard deviation
SDS: Standard deviation score
SE: Standard error
SHS: Second-hand smoke
TC: Transcript cluster
TSS: Transcriptional start site

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

LIFECYCLE Project
Coordination and leadership:

Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands

Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

Executive summary

Description of the deliverable: Epigenome-wide association studies produce lists of genomic regions with altered DNA methylation patterns as a consequence of environmental factors or disease outcomes, however final functional consequences are difficult to elucidate. This deliverable summarizes the studies conducted in LifeCycle in order to identify biological pathways associated with early life environmental exposures and disease outcomes in children.

Findings: Findings can be structured in three main areas of research:

1) **Development of methodological tools:** Functional characterization of DNA methylation changes requires of gene expression data or *cis*-eQTM catalogues (lists of positions of the genome whose DNA methylation is correlated with the expression of nearby genes). While there are some *cis*-eQTM catalogues, they are limited in sample size or age range. In LifeCycle, we have created the largest peadiatric blood *cis*-eQTM catalogue to date, which is publicly open for biological interpretation of DNA methylation findings (https://helixomics.isglobal.org/).

2) **Functional consequences of DNA methylation changes:** Using the peadiatric blood cis-eQTM catalogue or other approaches, we investigated the functional consequences of DNA methylation changes associated with in utero exposures (environmental stressors, maternal diet, maternal cardiometabolic traits) or health outcomes (reproductive, anthropometric, lung function and behaviour). Altered biological pathways were diverse, and while some of them were specific to the exposure/disease, others covered generic biological functions.

3) **Direct analysis of gene expression changes:** Two of the LifeCycle cohorts dispose of blood transcriptomics data, thus allowing a direct analysis of the effects of in utero exposure or disease outcomes on gene expression. The first analysis conducted with these data aims to identify deregulated genes in relation to child body mass index.

Next steps for research and policy: We have created a peadiatric blood *cis*-eQTM catalogue which represents a remarkable advance for characterization of funcional consequences of DNA methylation changes. However, proper biological interpretation will ultimately require transcriptomic and DNA methylation data measured in the same children. Moreover, these datasets will have to be large enough to identify substantial numbers of input genes for robust pathway enrichment analyses. Finally, cell-type specific effects should be addressed through single cell technolgies, or deconvolution methods if the first are not possible due to biological sample access and cost constrains.

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

1. Introduction

Work package 8 of the LifeCycle project focuses on using DNA methylation and RNA expression data to assess biological pathways underlying associations of early-life stressors and later-life health outcomes. The specific objective of Task 8.4 is to explore the functionality of any differences in DNA methylation.

A functional consequence of differences in DNA methylation are differences in gene expression, which in turn lead to activation or repression of functional pathways. Understanding the biological mechanisms associated with early-life stressors and later health outcomes might help the development of new treatments and preventive strategies.

In order to interpret findings of epigenome-wide association studies (EWAS) there is a need to link CpG sites to genes (functional units). There are two main strategies for this:

1) **Positional annotation**, where CpGs are annotated to the closest gene. Most of the EWAS from LifeCycle did positional annotation using information provided by Illumina (1). This annotation links CpGs to the closest gene when the CpG is located in the gene body, untranslated, or promoter region defined as <1,500 bp upstream of the transcription start site (TSS). A few LifeCycle EWAS complemented Illumina annotation by annotating intergenic CpGs to the closest gene by using other bioinformatics tools such as FDb.InfiniumMethylation.hg19 R package.

2) **Functional annotation (eQTM gene)**, where CpGs are linked to the regulated gene. This was done correlating DNA methylation and RNA expression in LifeCycle cohorts, or using public expression quantitative trait methylation (eQTM) catalogues. Datasets/eQTM catalogues used in LifeCycle include:
   - Early Autism Risk Longitudinal Investigation (EARLI) cohort for cord blood (n~119) (2);
   - Biomarkers of Exposure to ARsenic (BEAR) pregnancy cohort for cord blood (n~38) (3);
   - Isle of Wight 3rd Generation Study (IoW-3) for cord blood (n~157)(4);
   - ENID trial in Gambia for child's blood at age 2 years (n~100) (5);
   - INfancia y Medio Ambiente (INMA) cohort for child's blood at age 4 years (6) (this cohort is part of LifeCycle, n~120 assessed with the 450K at 0y and 4y and with the Affymetrix HTAv2 array at 4y);
   - Human Early Life Exposome (HELIX) project for child's blood at age 9 years (7) (HELIX includes the LifeCycle partners ISGLOBAL, BTHFT, UOC, NIPH, and INSERM, n~900 assessed with the 450K and the Affymetrix HTAv2 arrays);
   - BAMSE cohort for adolescent blood (n~244) (8);
   - Biobank-based integrative omics study (BIOS) for adult blood (n~2,000-3,000) (9);
   - Leipzig Childhood AT Cohort for adipose tissue from 0 to 21 years (n~223) (10);
   - A Biological Atlas of Severe Obesity (ABOS) for liver and muscle tissue from adults (n~71 for muscle, n~319 for liver) (11).

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

Once CpGs are linked to genes, either by positional or eQTM annotation, functional enrichment analyses can be performed. Two tools for functional enrichment were widely used in LifeCycle: ConsensusPathDB (12) and missMethyl R package (13). The second one corrects the analyses for the underlying distribution of CpGs in genes in the 450K and EPIC arrays. Other less frequent tools were Database for Annotation, Visualization and Integrated Discovery (DAVID) (14) and Ingenuity Pathways Analysis (IPA). The public databases of biological pathways queried as reference in most of the LifeCycle EWAS were: Gene Ontology (GO) terms (15) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) (16). Also, some EWAS checked for exposure/traits reported in the EWAS Atlas (17) and gene-disease search in Online Mendelian Inheritance in Man (https://www.omim.org/). Finally, in the most recent EWAS from LifeCycle, it was also explored whether CpGs were enriched for tissue specific regulatory regions (ie. blood promoters) or transcription factors using eFORGE (18), eFORGE-TF (19), or by applying Chi2 tests. These tools incorporate data on chromatin states and other epigenetic marks from the NIH Roadmap Epigenomics Mapping Consortium (http://www.roadmapepigenomics.org/) (20). We refer to this type of enrichment as "molecular enrichment analyses".

Moreover, some EWAS incorporated, in addition to DNA methylation, a direct look-up at the association between environmental factors and gene expression levels. We refer to these analyses conducted systematically along the genome as transcriptome-wide association studies (TWAS). These studies include INMA and HELIX (described above), and the Generation R Study, which has blood DNA methylation and RNA expression assessed with the 450K array and next generation sequencing platforms, respectively, for n=184 children at the age of 9 years (21).

In this report, we first describe the eQTM catalogue generated in LifeCycle to help interpretation of EWAS findings. Then, we review genes and functional pathways identified in 18 EWAS from the LifeCycle project, including EWAS on early-life stressors (9) and on health outcomes (9). Some of these analyses have been conducted in the context of the Pregnancy And Childhood Epigenetics Consortium (PACE). Finally, we introduce on-going TWAS in LifeCycle.

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

2. Description of progress and results

Below, we present the results, either published or ongoing, for this deliverable.

2.1 Building an expression quantitative trait methylation (eQTM) catalogue

HELIX *cis*-eQTM catalogue

Status: Published (22)
Partner(s) involved: ISGLOBAL (leader), BTHFT, UOC, NIPH, INSERM
Summary: The identification of expression quantitative trait methylation (eQTMs), defined as associations between DNA methylation levels and gene expression, might help the biological interpretation of EWAS. We aimed to identify autosomal *cis*-eQTMs in children's blood, using data from 832 European-ancestry children of the HELIX project. Blood DNA methylation and gene expression were measured with the Illumina 450K and the Affymetrix HTA v2 arrays, respectively. The relationship between methylation levels and expression of nearby genes (1 Mb window centered at the transcription start site, TSS) was assessed by fitting 13.6 M linear regressions adjusting for sex, age, cohort, and blood cell composition. We identified 39,749 blood autosomal *cis*-eQTMs, representing 21,966 unique CpGs (eCpGs, 5.7% of total CpGs) and 8,886 unique transcript clusters (eGenes, 15.3% of total transcript clusters (TCs), equivalent to genes). In 87.9% of these *cis*-eQTMs, the eCpG was located at <250 kb from eGene's TSS; and 58.8% of all eQTMs showed an inverse relationship between the methylation and expression levels (**Figure 1**). Only around half of the eGenes could be captured through annotation of the eCpG to the closest gene using the Illumina annotation. eCpGs had less measurement error, and were enriched for active blood regulatory regions and for CpGs reported to be associated with environmental exposures or phenotypic traits. 40.4% of all eQTMs had at least one genetic variant associated with methylation and expression levels (**Figure 2**). The overlap of autosomal *cis*-eQTMs in children's blood with those described in adults was small (13.8%), and age-shared *cis*-eQTMs tended to be proximal to the TSS and enriched for genetic variants.
Conclusion: The *cis*-eQTM catalogue is publicly available at https://helixomics.isglobal.org/. From the webpage, researchers can download the summarized results of the 13.6 M CpG-TC associations or the subset that passed the multiple-testing correction.

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

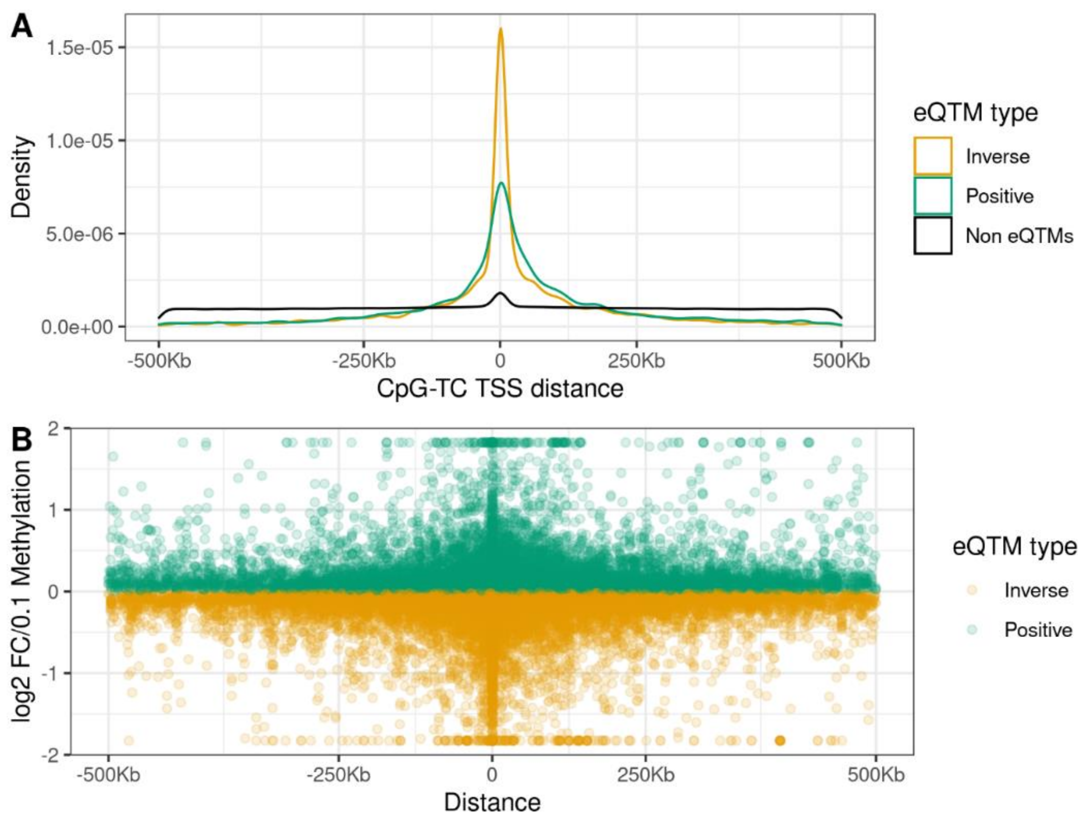

Figure 1. Distance between CpG and TC's TSS and effect size in child blood autosomal *cis*-eQTMs. A) Distribution of the distance between CpG and TC's TSS by eQTM type. CpG-TC pairs were classified in non eQTMs (black); inverse eQTMs (yellow); and positive eQTMs (green). The x-axis represents the distance between the CpG and the TC's TSS (kb). Non eQTMs median distance: -0.013 kb (interquartile range - IQR = -237; 236). Positive eQTMs median distance: -4.9 kb (IQR = -38; 79). Inverse eQTMs median distance: -0.7 kb (IQR = -29; 54). B) Effect size versus eCpG-Gene's TSS distance in eQTMs. The x-axis represents the distance between the eCpG and the eGene's TSS (kb). The y-axis represents the effect size as the $\log_2$ fold change in gene expression produced by a 0.1 increase in DNA methylation (or 10 percentile increase). In order to improve visualization, a 99% winsorization has been applied to $\log_2$ fold change values: values more extreme than the 99% percentile (in absolute value) have been changed for the 99% quantile value (in absolute value). eQTMs are classified in inverse (yellow) and positive (green). Each eQTM is represented by one dot. The darker the colour, the more dots overlapping, and so the higher the number of eQTMs with the same effect size and eCpG-eGene's TSS distance.

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

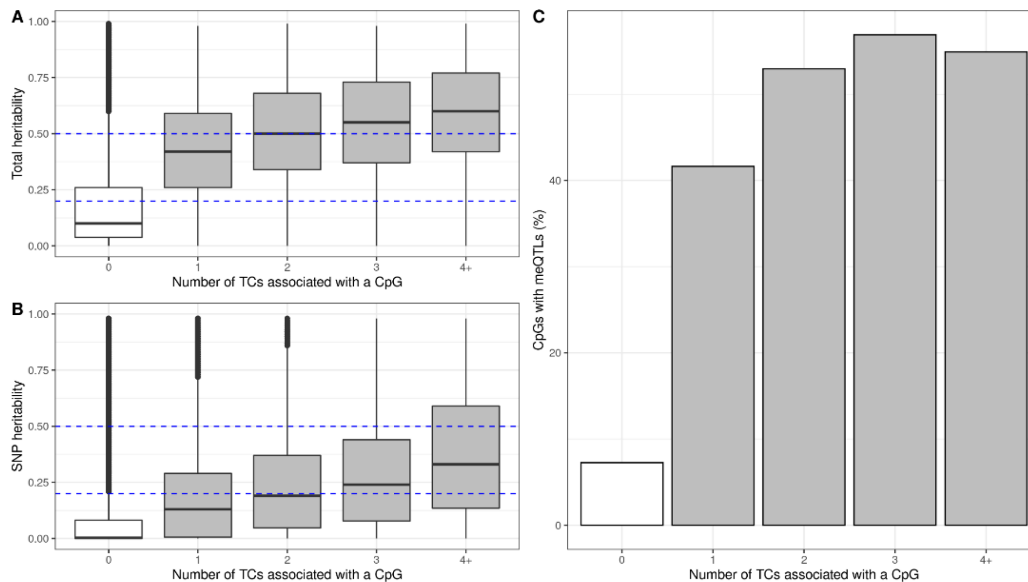

Figure 2. Genetic contribution to autosomal *cis*-eQTMs in children's blood. CpGs were grouped by the number of TCs they were associated with, where 0 means that a CpG was not associated with any TC (non eCpG). A) Total additive heritability and B) SNP heritability as inferred by Van Dongen and colleagues (23). The y-axis represents heritability and the x-axis each group of CpGs associated with a given number of TCs. C) Proportion of CpGs having a meQTL (methylation quantitative trait locus), by each group of CpGs associated with a given number of TCs.

2.2 Functional interpretation of EWAS of early-life stressors

We have selected a total of 9 EWAS of early-life stressors from LifeCycle, which use gene expression data to interpret the effect of CpGs, and/or conduct functional enrichment analyses. Early-life stressors have been classified in several groups: environmental pollutants during pregnancy (maternal tobacco smoking, air pollution, and to mercury), maternal diet during pregnancy (vitamin B12 concentrations, glycemic index and load, and Mediterranean diet), and maternal cardiometabolic traits during pregnancy (body mass index, hypertensive disorders and glycemic dysregulation). Main results, including identified genes and functional pathways, are listed in Appendix 1.

2.2.1 Environmental pollutants in pregnancy

In utero and childhood exposure to tobacco smoke and multi-layer molecular signatures in children

Status: Published (24)

Partner(s) involved: ISGLOBAL (leader), BTHFT, UOC, NIPH, INSERM

Summary: The adverse health effects of early-life exposure to tobacco smoking have been widely reported. In spite of this, the underlying molecular mechanisms of *in utero* and postnatal exposure to tobacco smoke are only partially understood. Here, we aimed to identify multi-layer molecular signatures associated with exposure to tobacco smoke in these two exposure windows. We investigated the associations of maternal smoking during pregnancy and childhood second-hand smoke (SHS) exposure with molecular features

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

measured in 1,203 European children (mean age 8.1 years) from the HELIX project. Molecular features, covering 4 layers, included blood DNA methylation and gene and miRNA transcription, plasma proteins, and serum and urinary metabolites.

Maternal smoking during pregnancy was associated with DNA methylation changes at 41 unique CpGs, but not with other molecular layers. 24/41 CpGs were associated with both smoking definitions (any and sustained), 3 with any smoking, and 14 with sustained smoking, although all of them were at least nominally significant in both models. These 41 CpGs were located in 18 loci, defined as regions of <2 Mb, and were distributed along the genome. Conversely, childhood SHS was not associated with blood DNA methylation or transcription patterns, but with reduced levels of several serum metabolites and with increased plasma PAI1 (plasminogen activator inhibitor-1), a protein that inhibits fibrinolysis. Some of the *in utero* and childhood smoking-related molecular marks showed dose-response trends, with stronger effects with higher dose or longer duration of the exposure. The persistent and dose-dependent changes in the methylome make CpGs good candidates to develop biomarkers of past exposure.

Results specifically related to Task 8.4: We examined whether the 41 CpGs might be eQTMs. A total of 480 unique TCs (equivalent to known or putative genes) with their TSS within ± 500 kb of the 41 CpGs were identified. At 5% False Discovery Rate (FDR), 15 methylation to expression relationships were found, which included 12 unique CpGs in 5 loci and 7 unique TCs (**Figure 3**). All eQTMs, except cg21161138 (*AHRR*)-TC05002792.hg.1, occurred between CpGs and genes located at <160 kb. None of the eQTMs genes corresponded to the most proximal gene to the CpG site. Two out of the 15 eQTM relationships were inverse, meaning higher methylation-lower gene expression, while the others were positive. The inverse associations were between *PNOC* and a CpG (cg17199018) located downstream of the gene, and between *EXOC3* and a CpG (cg11902777) located upstream.

With a nominal p-value <0.05, 65 other eQTMs were detected, in total, involving 33 unique CpGs in 14 out of the 18 loci. Only the methylation levels of CpGs at the gene body of *GFI1* and *AHRR* were associated with the expression of the same genes. *AHRR*, associated with maternal smoking in pregnancy and current smoking in adults, is an interesting example. Five CpGs in the *AHRR* locus were associated with maternal smoking in pregnancy: 2 hyper-methylated located at intron 1 (cg17924476, cg23067299), and 3 hypo-methylated at other introns (cg11902777, cg05575921, cg21161138) (**Figure 4**). Hyper-methylated CpGs in relation to maternal smoking in pregnancy were positive eQTMs for *AHRR*, *PDCD6*, and *EXOC3* genes, while hypo-methylated CpGs were inverse eQTMs for the same genes (in both cases implying higher expression of the genes).

Overall, DNA methylation at 5/18 loci associated with maternal smoking was related to expression of nearby genes (**Figure 3**). However, the expression of these genes themselves was only weakly associated with maternal smoking. Therefore, compared to methylation, the weak association of maternal smoking in pregnancy with gene expression suggests different reversal rates and a methylation-based memory to past exposures.

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

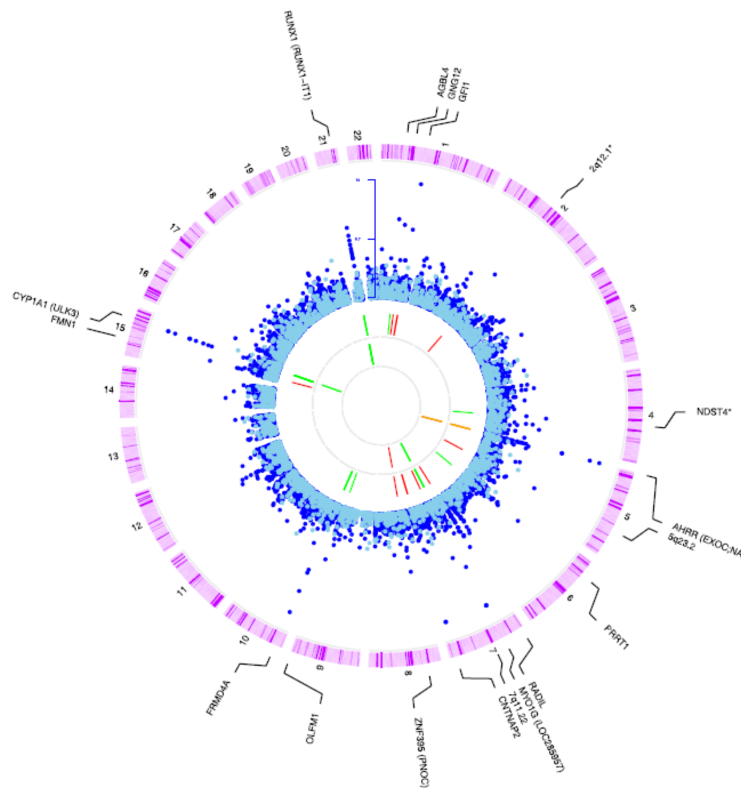

Figure 3. Circus plot showing the association between sustained maternal smoking in pregnancy and child blood DNA methylation and transcription along the chromosomes (outer circle). Second circus shows the statistical significance (−log10(p value)) for DNA methylation (dark blue) and transcription (light blue). Only the 18 loci significant at 5% FDR in the methylation analysis are annotated. Next circus shows the direction of the association of the CpGs in these 18 loci with maternal smoking in pregnancy (green, positive; red, inverse; and orange, loci with CpGs associated in both directions). The inner circus shows the 5 loci for which cis eQTMs at 5% FDR were identified (green, positive, meaning higher DNA methylation–higher gene expression; red, inverse; and orange, both). Genes annotated in parenthesis are significant eQTM genes, and none of them corresponds to the closest gene to the CpG site. Loci annotated with an asterisk are those surviving multiple-testing correction only in the any maternal smoking in pregnancy models. To gain graphical resolution, only associations with p value < 0.05 are shown, and p values < 1E−10 are truncated to 1E−10.

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

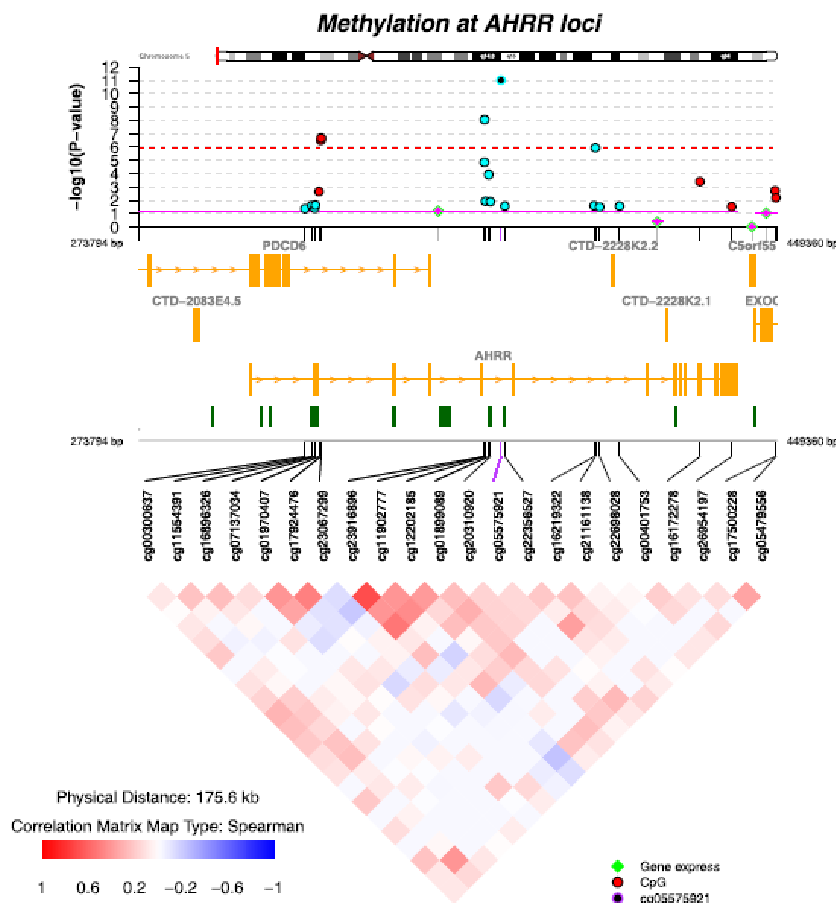

Figure 4. Regional plot of the AHRR locus (50 kb upstream and downstream of the 5 CpG sites associated with any maternal smoking in pregnancy). The y-axis of the top panel shows the −log (10) p value of the associations between any maternal smoking in pregnancy and methylation levels at CpG sites (circles) and gene expression levels (lines). Only CpGs nominally associated (p value < 0.05) are shown, hyper-methylated in red and hypo-methylated in blue. The top CpG, cg05575921, is shown in purple. Five of the CpGs survived multiple-testing correction (dashed red line): 2 hyper-methylated in intron 1 (cg17924476 and cg23067299), and 3 hypo-methylated at other introns (cg11902777, cg05575921, and cg21161138). The correlation of methylation levels among CpGs is shown at the bottom panel. The middle panel shows the annotation of genes (yellow) and CpG islands (green). The expression of none of the genes in the locus for which there were probes in the gene expression array was associated with any maternal smoking in pregnancy (p value < 0.05). All of them showed negative coefficients of the association (indicated as pink lines). List of TCs and gene annotation: TC05001094.hg.1 annotated to EXOC3-AS1, TC05000006.hg.1 annotated to EXOC3, TC05000005.hg.1 annotated to both AHRR and PDCD6, and TC05002795.hg.1 not annotated.

Prenatal particulate air pollution and DNA methylation in newborns: an epigenome-wide meta-analysis

Status: Published (25)

Partner(s) involved: ERASMUS, ISGLOBAL, UNITO, UNIVBRIS, UOC, INSERM

Summary: Prenatal exposure to air pollution has been associated with childhood respiratory disease and other adverse outcomes. Epigenetics is a suggested link between exposures and health outcomes. We investigated associations between prenatal exposure to particulate matter (PM) with diameter <10 (PM_{10}) or < 2.5 µm ($PM_{2.5}$) and DNA methylation in newborns and children. We meta-analyzed associations between exposure to particulate matter (PM) with diameter <10 (PM_{10}) (n = 1,949) and < 2.5 µm ($PM_{2.5}$) (n = 1,551) at

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

maternal home addresses during pregnancy and newborn DNA methylation assessed by Illumina 450K array. Six CpGs were significantly associated (FDR <0.05) with prenatal PM_{10} and 14 with $PM_{2.5}$ exposure. Two of the PM_{10}-related CpGs mapped to *FAM13A* (cg00905156) and *NOTCH4* (cg06849931) previously associated with lung function and asthma. Although these associations did not replicate in the smaller newborn sample, both CpGs were significant (p-value <0.05) in 7-to 9-y-olds. For cg06849931, however, the direction of the association was inconsistent. We also identified several differently methylated regions (DMRs) associated with either prenatal PM_{10} and or $PM_{2.5}$ exposure of which two PM_{10}-related DMRs, including *H19* and *MARCH11*, replicated in newborns.
Results specifically related to Task 8.4: The relationship between DNA methylation and expression of nearby genes was investigated in several databases. In adult blood (BIOS, n=2000), the top 3 PM_{10}-related CpGs as well as 6 out of 14 $PM_{2.5}$-associated CpGs, were significantly associated with gene expression. In newborns' blood (EARLI, n=119), no significant association of *in utero* PM_{10} exposure with expression of genes annotated to the respective CpG was detected, whereas $PM_{2.5}$ exposure was associated with expression of *ZNF695*. In adolescent's blood (BAMSE, n= 244), current PM_{10} exposure at 16y was associated with *NOTCH4* and *USP43* expression levels in peripheral blood cells. Among the $PM_{2.5}$ associated genes, *C7orf50* was significantly differentially expressed in relation to current $PM_{2.5}$ exposure. Positional annotation of CpGs identified in cord blood was used to generate a list of genes for functional enrichment analyses. Using ConsensusPathDB and after multiple-testing correction the following pathways were identified for PM_{10}: Notch signaling pathway, Rho GTPase cycle, Neurotransmitter release cycle, and GABA synthesis, release, reuptake and degradation. missMethyl identified similar pathways, but these did not remain statistically significant after multiple-testing correction. However, top pathways were equivalent. No statistically significantly enriched pathways were identified for $PM_{2.5}$.

Exposure to mercury (Hg) during pregnancy and offspring DNA methylation

Status: Published (26)
Partner(s) involved: ISGLOBAL, UNIVBRIS, UOC, NIPH
Summary: Mercury (Hg) is a ubiquitous heavy metal that originates from both natural and anthropogenic sources and is transformed in the environment to its most toxicant form, methylmercury (MeHg). Recent studies suggest that MeHg exposure can alter epigenetic modifications during embryogenesis. In this study, we examined associations between prenatal MeHg exposure and levels of cord blood DNA methylation by meta-analysis in up to seven independent studies (n = 1,462) as well as persistence of those relationships in blood from 7 to 8 year-old children (n = 794). In cord blood, we found limited evidence of differential methylation at cg24184221 in *MED31* (β = 2.2E-04, p-value = 5.9E-05) in relation to prenatal MeHg exposure. In child blood, we identified differential methylation at cg15288800 (β = 0.004, p-value = 5E-05), also located in *MED31*. This repeated link to *MED31*, a gene involved in lipid metabolism and RNA Polymerase II transcription function, may suggest a methylation perturbation related to MeHg exposure that persists into early childhood. Further, we found evidence for association between prenatal MeHg exposure and child blood methylation levels at two additional CpGs: cg12204245 (β = 0.002, p-value = 4.8E-07) in *GRK1* and cg02212000 (β = -0.001, p-value = 8.1E-07) in *GGH*. Prenatal MeHg

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu
This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

exposure was associated with methylation modifications that may influence health outcomes, such as cognitive or anthropometric development, in different populations.
Results specifically related to Task 8.4: No statistically significant GO terms were found considering the positional genes at p-value <1E-05. However, top GO terms included vascular permeability to maintain physiological tissue homeostasis, astrocyte activation and microtubule nucleation that play a crucial role in orchestrating neural development by coordinating synapse formation and function, as well as titin binding, involved in muscular development.

2.2.2 Maternal diet in pregnancy

A meta-analysis of epigenome-wide association studies on pregnancy vitamin B12 concentrations and offspring DNA methylation
Status: Submitted
Partner(s) involved: ERASMUS (leader), ISGLOBAL, UNIVBRIS, NIPH
Summary: Circulating vitamin B12 concentrations during pregnancy are associated with offspring health. Fetal DNA methylation changes could underlie these associations. Within the PACE Consortium, we meta-analyzed epigenome-wide associations of circulating vitamin B12 concentrations in either mothers during pregnancy (n=2,420) or cord blood (n=1,029) with cord blood DNA methylation. We observed associations of maternal and newborn vitamin B12 concentrations with DNA methylation at 109 and 7 CpGs, respectively (FDR p-value <0.05). Persistent associations with DNA methylation in peripheral blood of children aged 5-10 years (n=482, largely overlapping subjects) were observed for 40.7% of maternal and 57.1% of newborn vitamin B12-related CpGs. Of the CpGs identified in the maternal and newborn meta-analyses, 4.6% and 14.3%, respectively, were also related to birth weight or gestational age in previous work.
Results specifically related to Task 8.4: For the 109 CpGs from the maternal meta-analysis, we identified 57 unique CpG-gene expression pairs using the HELIX eQTM catalogue. These *cis*-eQTMs involved 18 unique TCs (equivalent to putative genes) with TSS within ±500 kb of any of 20/109 (18.3%) CpGs. Most associations (41/57, 71.9%) were negative, indicating that higher methylation was associated with lower gene expression. We observed the most significant association between methylation at cg21482265 and gene expression of *PAX8* with a $\log_2$ fold change in expression per 10% increase in DNA methylation of -0.096 (SE 0.025; P-value 9.48E-185).
Using eFORGE we found that the 109 CpGs of the maternal meta-analysis were enriched for certain chromatin states and histone marks, but not for DNAseI hypersensitive sites. Moreover, we observed evidence for enrichment of the transcription factors Methyl CpG Binding Protein 2 (MECP2, maternal meta-analysis prioritized CpGs) and Deformed Epidermal Autoregulatory Factor 1 (DEAF1, newborn meta-analysis prioritized CpGs) (eFORGE-TF). Finally, the CpGs from both maternal and newborn meta-analysis showed little evidence for functional enrichment of GO or KEGG terms.

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

Maternal Mediterranean diet in pregnancy and newborn DNA methylation: a meta-analysis in the PACE Consortium

Status: Published (27)

Partner(s) involved: ERASMUS (leader), ISGLOBAL, UNIVBRIS

Summary: Higher adherence to the Mediterranean diet during pregnancy is related to a lower risk of preterm birth and to better offspring cardio-metabolic health. DNA methylation may be an underlying biological mechanism. We evaluated whether maternal adherence to the Mediterranean diet was associated with offspring cord blood DNA methylation. We meta-analyzed EWAS of maternal adherence to the Mediterranean diet during pregnancy and offspring cord blood DNA methylation in 2,802 mother-child pairs from 5 cohorts. We calculated the relative Mediterranean diet (rMED) score with range 0-18 and an adjusted rMED excluding alcohol (rMEDp, range 0-16). DNA methylation was measured using Illumina 450K arrays. We used robust linear regression modelling adjusted for child sex, maternal education, age, smoking, body mass index, energy intake, batch and cell types. We performed several functional analyses and examined persistence of differential DNA methylation into childhood (4.5-7.8y). rMEDp was associated with cord blood DNA methylation at cg23757341 (0.064% increase in DNA methylation per 1-point increase in the rMEDp score, SE=0.011, P-VALUE=2.41E-08). This CpG site maps to *WNT5B*, associated with adipogenesis and glycemic phenotypes. The association did not persist into childhood.

Results specifically related to Task 8.4: We did not identify associations with childhood gene expression, nor did we find enriched biological pathways.

Maternal dietary glycemic index and glycemic load in pregnancy and offspring cord blood DNA methylation

Status: Accepted in Diabetes Care, 2022

Partner(s) involved: ERASMUS (leader), ISGLOBAL, UNIVBRIS

Summary: Suboptimal nutrition in pregnancy is associated with worse offspring cardio-metabolic health. DNA methylation may be an underlying mechanism. We meta-analyzed EWAS of maternal dietary glycemic index and load with cord blood DNA methylation. We calculated maternal glycemic index and load from food frequency questionnaires, and ran EWAS on cord blood DNA methylation in 2,003 mother-offspring pairs from three cohorts. Analyses were additionally stratified by maternal BMI categories. Maternal glycemic index and load were associated with cord blood DNA methylation at 41 CpGs, mostly in mothers with overweight/obesity. We did not observe overlap with CpGs associated with maternal glycemic traits, BMI or child birthweight or BMI. Additional studies are required to further explore functionality, uncover causality, and study pathways to offspring health.

Results specifically related to Task 8.4: In the *cis*-eQTM analyses using HELIX catalogue, we found three CpG–TC associations in childhood blood for cg24458009, with the *PCED1B* gene annotated to this TC. We further found one CpG–TC association for cg23347399, with the Protocadherin Gamma gene family (*PCDHGA*, *PCDHGB*, *PCDHGC*) annotated to this TC. These two CpGs were both found in the EWAS for associations with glycemic index in mothers with overweight or obesity. In the analyses of associations between DNA

LIFECYCLE Project
Coordination and leadership: Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

methylation and RNA transcript levels in adipose tissue (10), 4 of the 561 CpG-transcript pairs tested reached statistical significance. These all involved cg27193519, showing a direct association between DNA methylation and expression levels of *TFAP4*, *ZNF500*, *PPL* and *ANKS3*.
Functional enrichment analyses on all 41 CpGs from the full group or stratified meta-analyses, did not result in FDR significant GO or KEGG pathways using missMethyl R package. We further found no functional enrichment for the 76 and 231 CpGs with p-value <0.0001 from the full group meta-analyses on glycemic index and load, respectively. Similarly, there was no functional enrichment for the 116 or 1,176 CpGs with p-value <0.0001 from the stratified analyses for glycemic index in mothers with normal weight or overweight or obesity, respectively, nor for the 222 or 265 CpGs with p-value <0.0001 from the stratified analyses for glycemic load, respectively. Finally, using the set of 41 CpGs, we found no evidence of enrichment for tissue-specific DNaseI hypersensitivity regions using the e-FORGE application.

2.2.3 Maternal cardiometabolic traits in pregnancy

Maternal BMI at the start of pregnancy and offspring epigenome-wide DNA methylation: findings from the pregnancy and childhood epigenetics (PACE) Consortium
Status: Published (28)
Partner(s) involved: UNIVBRIS (leader), ERASMUS, ISGLOBAL, UMCG, NIPH, INSERM, UWA
Summary: Pre-pregnancy maternal obesity is associated with adverse offspring outcomes at birth and later in life. Individual studies have shown that epigenetic modifications such as DNA methylation could contribute. Within the PACE Consortium, we meta-analysed the association between pre-pregnancy maternal body mass index (BMI) and methylation at over 450,000 sites in newborn blood DNA, across 19 cohorts (9,340 mother-newborn pairs). In four additional cohorts (1,817 mother-child pairs), we meta-analysed the association between maternal BMI at the start of pregnancy and blood methylation in adolescents. In newborns, maternal BMI was associated with small (<0.2% per BMI unit (1 kg/m^2), p-value <1.06E-07) methylation variation at 9,044 sites throughout the genome. Adjustment for estimated cell proportions greatly attenuated the number of significant CpGs to 104, including 86 sites common to the unadjusted model. At 72/86 sites, the direction of the association was the same in newborns and adolescents, suggesting persistence of signals. However, we found evidence for a causal intrauterine effect of maternal BMI on newborn methylation at just 8/86 sites. In conclusion, this well-powered analysis identified robust associations between maternal adiposity and variations in newborn blood DNA methylation, but these small effects may be better explained by genetic or lifestyle factors than a causal intrauterine mechanism. This highlights the need for large-scale collaborative approaches and the application of causal inference techniques in epigenetic epidemiology.
Results specifically related to Task 8.4: The 86 maternal BMI associated CpGs were near 77 genes, and there were several instances where multiple sites mapped to the same gene: *RBMS1* (3 sites), *POM121L1P* (3 sites), *VIPR2* (2 sites), *SQLE* (2 sites), *RASA3* (2 sites), *MIR200B* (2 sites), *KAT6B* (2 sites). The list of these 77 genes was not enriched for any GO term of KEGG pathway after multiple-testing correction using missMethyl R package.

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

LifeCycle

Hypertensive disorders of pregnancy and DNA methylation in newborns

Status: Published (29)
Partner(s) involved: UNIVBRIS (leader), ERASMUS, NIPH
Summary: Hypertensive disorders of pregnancy (HDP) are associated with low birth weight, shorter gestational age, and increased risk of maternal and offspring cardiovascular diseases later in life. The mechanisms involved are poorly understood, but epigenetic regulation of gene expression may play a part. We performed meta-analyses in the PACE Consortium to test the association between either maternal HDP (10 cohorts; n=5,242 [cases=476]) or preeclampsia (3 cohorts; n=2,219 [cases=135]) and epigenome-wide DNA methylation in cord blood using the Illumina 450K array. In models adjusted for confounders, and with Bonferroni correction, HDP and preeclampsia were associated with DNA methylation at 43 and 26 CpG sites, respectively. HDP was associated with higher methylation at 27 (63%) of the 43 sites, and across all 43 sites, the mean absolute difference in methylation was between 0.6% and 2.6%. Epigenome-wide associations of HDP with offspring DNA methylation were modestly consistent with the equivalent epigenome-wide associations of preeclampsia with offspring DNA methylation (R^2=0.26). In longitudinal analyses conducted in 1 study (n=108 HDP cases and 550 controls), there were similar changes in DNA methylation in offspring of those with and without HDP up to adolescence.
Results specifically related to Task 8.4: Pathway analysis suggested that genes located at/near HDP-associated sites may be involved in developmental, embryogenesis, or neurological pathways.

Maternal glycemic dysregulation during pregnancy and neonatal blood DNA methylation: meta-analyses of epigenome wide association studies

Status: Published (30)
Partner(s) involved: ERASMUS, INSERM, UOULU
Summary: Maternal glycemic dysregulation during pregnancy increases the risk of adverse health outcomes in the offspring. The risk of adverse offspring outcomes is linearly associated with maternal hyperglycemia. It has been hypothesized that changes in offspring DNA methylation may explain these associations. We conducted fixed-effect meta-analyses of EWAS results from eight birth cohorts investigating relationships between cord-blood DNA methylation and fetal exposure to maternal glucose (n max= 3,503), insulin (n max= 2,062), and the area under the curve of glucose (AUCgluc) following oral glucose tolerance tests (OGTT, n max= 1,505). We performed look-ups on identified CpG dinucleotides (CpGs) in existing datasets to examine associations between DNA methylation and cardiometabolic traits and with tissue-specific expression.
Greater maternal AUCgluc was associated with lower cord blood DNA methylation at neighboring CpGs cg26974062 and cg02988288 in *TXNIP*. The association was mitigated by GDM treatment. Lower blood DNA methylation at *TXNIP* was associated multiple metabolic traits later in life including type 2 diabetes, and in liver biopsies, DNA methylation at cg26974062 was associated with hepatic gene expression of *TXNIP*. We observed little evidence of associations between either maternal glucose or insulin and cord-blood DNA methylation.

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

Maternal hyperglycemia, as reflected by AUCgluc, was associated with lower cord blood DNA methylation at *TXNIP*, a liver glucose transport gene. In independent studies, DNA methylation at *TXNIP* was associated with cardiometabolic traits in later life, suggesting that this epigenetic mark may have a role in glycemic regulation and metabolic health across the life course.

Results specifically related to Task 8.4: We used data from muscle and liver biopsies of women with obesity from the ABOS atlas project (11) to examine associations of DNA methylation levels at cg26974062 and cg02988288 and *TXNIP* expression. In liver, but not muscle, tissue, DNA methylation at cg02988288 was associated with *TXNIP* gene expression (β= -0.76 [SE= 0.34], p-value = 0.024, n=322). DNA methylation at cg26974062 was not significantly associated with *TXNIP* expression in either liver or muscle tissue.

2.3 Functional interpretation of EWAS of health outcomes

There are 9 EWAS of health outcomes from LifeCycle, which incorporate gene expression data and/or functional enrichment analyses: birth weight, gestational age, child BMI, epigenetic age acceleration, ADHD, bullying, aggression, lung function and asthma. Main results, including identified genes and functional pathways are listed in Appendix 2.

Epigenome-wide meta-analysis of blood DNA methylation in newborns and children identifies numerous loci related to gestational age

Status: Published (31)

Partner(s) involved: ERASMUS, ISGLOBAL, UNIVBRIS, UMCG, UOC, NIPH, INSERM, UOULU, UWA

Summary: Preterm birth and shorter duration of pregnancy are associated with increased morbidity in neonatal and later life. As the epigenome is known to have an important role during fetal development, we investigated associations between gestational age and blood DNA methylation in children. We performed meta-analysis of 450K array associations between gestational age and cord blood DNA methylation in 3,648 newborns from 17 cohorts without common pregnancy complications, induced delivery or caesarean section. We identified 8,899 CpGs in cord blood that were associated with gestational age (range 27-42 weeks), at Bonferroni corrected significance, p-value <1.06E.07, of which 3,343 were novel. After restricting findings to at least three significant adjacent CpGs, we identified 1,276 CpGs. We also explored associations of gestational age with DNA methylation measured at different ages (longitudinal analyses) identifying 222 CpGs.

Results specifically related to Task 8.4: Using positional annotation, the 8,899 CpGs associated with gestational age in cord blood were linked to 4,966 genes; while using eQTM annoation with a cord blood reference database (3), only 200 genes were identified. Functional enrichment analyses conducted with the list of closest genes, highlighted 1,784 GO terms, including regulation of cellular and biological processes, system development, different signalling pathways and organ development; and 124 KEGG pathways, including various cancers, viral infections, metabolic processes and immune-related disorders. The same analyses conducted only with the 222 CpGs (139 annotated genes) associated with gestational age longitudinal data, retried 13 KEGG pathways specially related to infection- and immune-related disorders.

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

LifeCycle

Meta-analysis of epigenome-wide association studies in neonates reveals widespread differential DNA methylation associated with birthweight

Status: Published (32)
Partner(s) involved: ERASMUS (leader), UNIVBRIS (co-leader), UMCG (co-leader), ISGLOBAL, UOC, NIPH, UOULU, UWA
Summary: Birthweight is associated with health outcomes across the life course, DNA methylation may be an underlying mechanism. In this meta-analysis of EWAS of 8,825 neonates from 24 birth cohorts in the PACE Consortium, we find that DNA methylation in neonatal blood was associated with birthweight at 914 sites, with a difference in birthweight ranging from −183 to 178 grams per 10% increase in methylation (p-value <1.06E−07). In additional analyses in 7,278 participants, <1.3% of birthweight-associated differential methylation was also observed in childhood and adolescence, but not adulthood. Birthweight-related CpGs overlapped with some Bonferroni-significant CpGs that were previously reported to be related to maternal smoking (55/914, p-value = 6.12E−74) and BMI in pregnancy (3/914, p-value = 1.13E−03), but not with those related to folate levels in pregnancy. Whether the associations are causal or explained by confounding or fetal growth influencing DNA methylation (i.e. reverse causality) requires further research.
Results specifically related to Task 8.4: The 914 BW-related CpGs identified in cord blood were annotated to 729 genes using the Illumina annotation, to 85 genes using an eQTM database of adult blood (BIOS, n~2,000)(9), to 17 genes using an eQTM database of child blood at age 4y (INMA, n~150) , and to 2 genes using an eQTM database of child blood at age 2y (Gambian cohort, n~100) (5). The overlap among eQTM genes identified in the different databases was very low. After adjusting for multiple-testing correction, the list of annotated genes with Illumina did not retrieve and functional pathway (KEGG and GO).

DNA methylation and body mass index from birth to adolescence: meta-analyses of epigenome-wide association studies

Status: Published (33)
Partner(s) involved: ERASMUS (leader), UNIVBRIS (co-leader), ISGLOBAL, BTHFT, UMCG, UOC, NIPH, INSERM, UOULU, LMU, UWA
Summary: DNA methylation has been shown to be associated with adiposity in adulthood. However, whether similar DNA methylation patterns are associated with childhood and adolescent BMI is largely unknown. We examined whether DNA methylation in cord blood and whole blood in childhood and adolescence was associated with BMI in the age range from 2 to 18 years using both cross-sectional and longitudinal models. We performed meta-analyses of epigenome-wide association studies including up to 4,133 children from 23 studies. DNA methylation at 3 CpGs (cg05937453, cg25212453, and cg10040131), each in a different age range, was associated with BMI at Bonferroni significance, with a 0.96 standard deviation score (SDS) (standard error (SE) 0.17), 0.32 SDS (SE 0.06), and 0.32 BMI SDS (SE 0.06) higher BMI per 10% increase in methylation, respectively. DNA methylation at nine additional CpGs in the cross-sectional childhood model was associated with BMI at FDR significance. The strength of the associations of DNA methylation at the 187 CpGs previously identified to be associated with adult BMI, increased with advancing age across childhood

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

and adolescence in our analyses. In addition, correlation coefficients between effect estimates for those CpGs in adults and in children and adolescents also increased. With the advancing age of the participants across childhood and adolescence, we observed increasing overlap with altered DNA methylation loci reported in association with adult BMI. These findings may be compatible with the hypothesis that DNA methylation differences are mostly a consequence rather than a cause of obesity.
Results specifically related to Task 8.4: Functional interpretation of the most significantly associated CpGs (p-value <1E-04) was investigated with the missMethyl R package and the GO and KEGG databases. No pathway or GO term passed the multiple-testing correction threshold.

The early-life exposome and epigenetic age acceleration in children

Status: Published (34)
Partner(s) involved: ISGLOBAL (leader), BTHFT, UOC, NIPH, INSERM, UOC
Summary: The early-life exposome influences future health and accelerated biological aging has been proposed as one of the underlying biological mechanisms. We investigated the association between more than 100 exposures assessed during pregnancy and in childhood (including indoor and outdoor air pollutants, built environment, green environments, tobacco smoking, lifestyle exposures, and biomarkers of chemical pollutants), and epigenetic age acceleration in 1,173 children aged 7 years old from the HELIX project. Age acceleration was calculated based on Horvath's Skin and Blood clock using child blood DNA methylation measured by 450K array. We performed an exposure-wide association study between prenatal and childhood exposome and age acceleration. Maternal tobacco smoking during pregnancy was nominally associated with increased age acceleration. For childhood exposures, indoor particulate matter absorbance (PM_{abs}) and parental smoking were nominally associated with an increase in age acceleration. Exposure to the organic pesticide dimethyl dithiophosphate and the persistent pollutant polychlorinated biphenyl-138 (inversely associated with child body mass index) were protective for age acceleration. None of the associations remained significant after multiple-testing correction. Pregnancy and childhood exposure to tobacco smoke and childhood exposure to indoor PM_{abs} may accelerate epigenetic aging from an early age.
Results specifically related to Task 8.4: To interpret the biological meaning of epigenetic age, we searched the genes whose expression was associated with the methylation levels of the CpGs included in the "Horvath's Skin and Blood clock" in HELIX eQTM catalogue. 72 CpGs out of the 391 (18.41%) in "Horvath's Skin and Blood clock" were associated with the expression of 151 unique TCs (or genes), which were annotated to 129 unique gene symbols. 122 out of 129 genes were identified in ConsensusPathDB (KEGG, Reactome or Biocarta) and were enriched for the following biological pathways (q-value < 0.025): adaptive and innate immune system, apoptosis, cell cycle and cancer, and detoxification of xenoestrogens.

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

LifeCycle

Association between DNA methylation and ADHD symptoms from birth to school age: a prospective meta-analysis

Status: Published (35)
Partner(s) involved: ERASMUS (leader), ISGLOBAL, UNIVBRIS
Summary: Attention-deficit and hyperactivity disorder (ADHD) is a common childhood disorder with a substantial genetic component. However, the extent to which epigenetic mechanisms play a role in the etiology of the disorder is unknown. We performed EWAS in collaboration with the PACE Consortium to identify DNA methylation sites associated with ADHD symptoms at two methylation assessment periods: birth and school age. We examined associations of both DNA methylation in cord blood with repeatedly assessed ADHD symptoms (age 4-15 years) in 2477 children from 5 cohorts and of DNA methylation at school age with concurrent ADHD symptoms (age 7-11 years) in 2,374 children from 9 cohorts, with 3 cohorts participating at both timepoints. CpGs identified with nominal significance (p-value <0.05) in either of the EWAS were correlated between timepoints (ρ = 0.30), suggesting overlap in associations; however, top signals were very different. At birth, we identified nine CpGs that predicted later ADHD symptoms (p-value <1E-07), including *ERC2* and *CREB5*. Peripheral blood DNA methylation at one of these CpGs (cg01271805 in the promoter region of *ERC2*, which regulates neurotransmitter release) was previously associated with brain methylation. Another (cg25520701) lies within the gene body of *CREB5*, which previously was associated with neurite outgrowth and an ADHD diagnosis. In contrast, at school age, no CpGs were associated with ADHD with p-value <1E-07. In conclusion, we found evidence in this study that DNA methylation at birth is associated with ADHD. Future studies are needed to confirm the utility of methylation variation as biomarker and its involvement in causal pathways.
Results specifically related to Task 8.4: Two-hundred forty-nine probes showed suggestive (p-value <1E-05) associations and were annotated to 182 unique genes. In gene-based analyses, no pathway survived multiple-testing correction. The 248 suggestive CpGs were enriched in intergenic regions. Of these, hypomethylated CpGs were enriched for 3'-untranslated regions and depleted for TSS200 and first exon regions, open sea, north shelf and south shelf regions, south shore, and islands. Regarding chromatin states, hypomethylated probes showed an enrichment for transcription (Tx and TxWk), quiescent positions, and depletion for transcription start site positions (TSSA, TxFlnk, TxFlnk), bivalent (EnhBiv), and repressor (ReprPC) positions. Hypermethylated probes showed the opposite enrichment/depletion patterns.

Epigenomics of being bullied: changes in DNA methylation following bullying exposure

Status: Published (36)
Partner(s) involved: ERASMUS (leader), UNIVBRIS (co-leader)
Summary: Bullying among children is ubiquitous and associated with pervasive mental health problems. However, little is known about the biological pathways that change after exposure to bullying. Epigenome-wide changes in DNA methylation in peripheral blood were studied from pre- to post measurement of bullying exposure, in a longitudinal study of the population-based Generation R Study and Avon Longitudinal Study of Parents and Children

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu
This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

(ALSPAC) (combined n = 1,352). Linear mixed-model results were meta-analysed to estimate how DNA methylation changed as a function of exposure to bullying. One site, cg17312179, showed small changes in DNA methylation associated to bullying exposure (b = −2.67e-03, SE = 4.97e-04, p-value = 7.17e-08). This site is annotated to *RAB14*, an oncogene related to Golgi apparatus functioning, and its methylation levels decreased for exposed but increased for non-exposed. This result was consistent across sensitivity analyses including co-occurring child characteristics and risks. Top CpG sites tended to have overall low levels of DNA methylation, decreasing in exposed, increasing in non-exposed individuals. A candidate follow-up was employed for CpG sites annotated to *5-HTT* and *NR3C1*, but not relevant findings were found. This is the first study to identify changes in DNA methylation associated with bullying exposure at the epigenome-wide significance level. Consistent with other population-based studies, we do not find evidence for strong associations between bullying exposure and DNA methylation.

Results specifically related to Task 8.4: GO terms analysis on CpGs with p <0.001 (n = 644 CpGs, n = 396 genes) yielded 126 pathways, 25 of which were confirmed by a GO analysis on CpGs with p <0.01 (n = 5,997 CpGs, n = 3,722 genes) and 43 of which were confirmed by a GO analysis on CpGs with p < 0.0001 (n = 66 CpGs, n = 53 genes). Ryanodine-sensitive calcium-release channel activity as the most enriched biological process (p-value = 9.99E−08) (**Figure 5**). Other enriched terms for biological processes involve various neurodevelopmental processes, such as astrocyte differentiation and action potential regulation, as well as processes such as muscle fibre development.

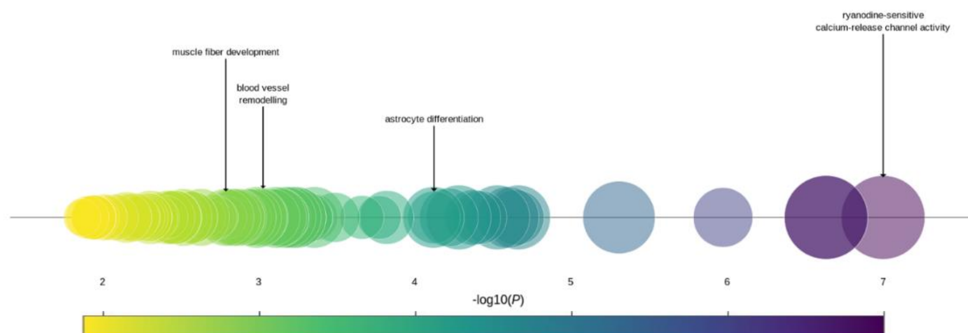

Figure 5. Enriched biological processes (p-value <0.05) in GO analysis of CpG sites with p-value <0.001 (n=644 CpG sites, n=396 genes) ordered by significance. Circle size represents percentage of genes represented in pathway versus all genes in that pathway.

DNA methylation signatures of aggression and closely related constructs: a meta-analysis of epigenome-wide studies across the lifespan

Status: Published (37)

Partner(s) involved: ERASMUS, ISGLOBAL, UNIVBRIS, UMCG, UOC, UOULU

Summary: DNA methylation profiles of aggressive behavior may capture lifetime cumulative effects of genetic, stochastic, and environmental influences associated with aggression. Here, we report the first large meta-analysis of EWAS of aggressive behavior (N = 15,324 participants). In peripheral blood samples of 14,434 participants from 18 cohorts with mean

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu
This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

ages ranging from 7 to 68 years, 13 methylation sites were significantly associated with aggression (alpha = 1.2E−07; Bonferroni correction). In cord blood samples of 2,425 children from 5 cohorts with aggression assessed at mean ages ranging from 4 to 7 years, 83% of these sites showed the same direction of association with childhood aggression (r = 0.74, p-value = 0.006) but no epigenome-wide significant sites were found. Top-sites (48 at a FDR of 5% in the peripheral blood meta-analysis or in a combined meta-analysis of peripheral blood and cord blood) have been associated with chemical exposures, smoking, cognition, metabolic traits, and genetic variation (mQTLs). Three genes whose expression levels were associated with top-sites were previously linked to schizophrenia and general risk tolerance. At six CpGs, DNA methylation variation in blood mirrors variation in the brain. On average 44% (range = 3–82%) of the aggression–methylation association was explained by current and former smoking and BMI. These findings point at loci that are sensitive to chemical exposures with potential implications for neuronal functions. We hope these results to be a starting point for studies leading to applications as peripheral biomarkers and to reveal causal relationships with aggression and related traits.
Results specifically related to Task 8.4: Based on peripheral blood RNA-seq and DNA methylation data (N = 2101) (9), 17 significant DNA methylation-gene expression associations were identified among 15 CpGs and ten transcripts. For most transcripts, a higher methylation level at a CpG site in *cis* correlated with lower expression (82.4%): cg03935116 and *FAM60A*, cg00310412 and *SEMA7A*, cg03707168 and *PPP1R15A*, cg03636183 and *F2RL3*, two intergenic CpGs on chromosome 6, where methylation level correlated negatively with expression levels of *FLOT1*, *TUBB*, and *LINC00243*, and six CpGs annotated to *AHRR* were negatively associated with *EXOC3* expression level. Positive correlations were observed between methylation levels at 2 CpGs on chromosome 7 and levels of *RP4-647J21.1* (novel transcript, overlapping *MYO1G*) and between cg02895948 and *PLXNA2*.

Newborn DNA-methylation, childhood lung function, and the risks of asthma and COPD across the life course

Status: Published (38)
Partner(s) involved: ERASMUS (leader), ISGLOBAL, UNIVBRIS
Summary: We aimed to identify differentially methylated regions (DMRs) in cord blood DNA associated with childhood lung function, asthma and chronic obstructive pulmonary disease (COPD) across the life course. We meta-analysed epigenome-wide data of 1,688 children from five cohorts to identify cord blood DMRs and their annotated genes, in relation to forced expiratory volume in 1 s (FEV_1), FEV_1/forced vital capacity (FVC) ratio and forced expiratory flow at 75% of FVC at ages 7–13 years. Identified DMRs were explored for associations with childhood asthma, adult lung function and COPD, gene expression and involvement in biological processes.
We identified 59 DMRs associated with childhood lung function, of which 18 were associated with childhood asthma and nine with COPD in adulthood. Genes annotated to the top 10 identified DMRs were *HOXA5*, *PAOX*, *LINC00602*, *ABCA7*, *PER3*, *CLCA1*, *VENTX*, *NUDT12*, *PTPRN2* and *TCL1A*. Our findings suggest that the epigenetic status of the newborn affects respiratory health and disease across the life course.

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

Results specifically related to Task 8.4: Of the 59 identified DMRs, 32 (54%) DMRs at birth were associated with gene expression at age 4 years, and 18 (31%) DMRs with gene expression in adulthood. The DMR annotated to *HOXA5* was associated with differential expression of several genes of the HOX-family. The DMRs annotated to *PER3*, *VENTX*, *NUDT12* and *TCL1A* were associated with differential expression of their respective genes. The DMRs annotated to *PAOX*, *LINC00602*, *ABCA7*, *CLCA1* and *PTPRN2* were not associated with expression of their corresponding genes. Genes annotated to 28 (47%) of all identified DMRs were expressed in adult lung tissue, including the top significant DMRs annotated to *PAOX*, *ABCA7*, *CLCA1*, *VENTX* and *NUDT12*.
Integrative pathway analysis showed that the majority of genes annotated to DMRs associated with childhood lung function were in co-expression and are known for physical and genetic interactions (**Figure 6**). Genes related to the identified DMRs, including *HOXA5*, *PER3*, *CLCA1*, *NUDT12* and *PTPRN2* were located in pathways related to regionalisation, DNA- and RNA-regulation and embryonic development.

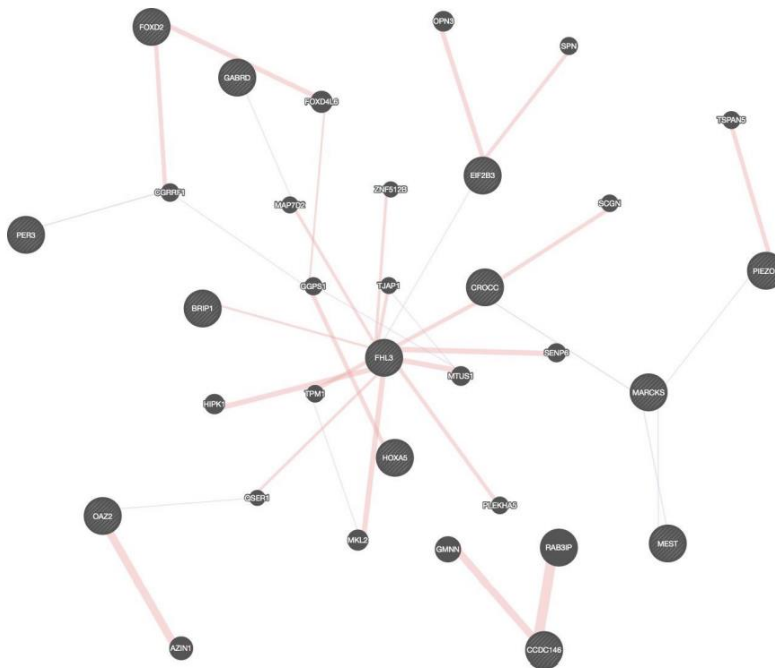

Figure 6. Integrative pathway analysis of all genes annotated to DMRs associated with childhood FEV1. Interpretation of line colours: purple: co-expression; pink: physical interaction; blue: co-localisation.

Epigenome-wide meta-analysis of DNA methylation and childhood asthma

Status: Published (39)
Partner(s) involved: ERASMUS, ISGLOBAL, UNIVBRIS, UMCG, NIPH, INSERM, UOULU, LMU, UWA
Summary: Epigenetic mechanisms, including methylation, can contribute to childhood asthma. Identifying DNA methylation profiles in asthmatic patients can inform disease

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

pathogenesis. We sought to identify differential DNA methylation in newborns and children related to childhood asthma. Within the PACE Consortium, we performed epigenome-wide meta-analyses of school-age asthma in relation to CpG methylation
(450K) in blood measured either in newborns, in prospective analyses, or cross-sectionally in school-aged children. We also identified differentially methylated regions.
In newborns (8 cohorts, 668 cases), 9 CpGs (and 35 regions) were differentially methylated (epigenome-wide significance, FDR <0.05) in relation to asthma development. In a cross-sectional meta-analysis of asthma and methylation in children (9 cohorts, 631 cases), we identified 179 CpGs (FDR <0.05) and 36 differentially methylated regions. In replication studies of methylation in other tissues, most of the 179 CpGs discovered in blood replicated, despite smaller sample sizes, in studies of nasal respiratory epithelium or eosinophils. Several implicated genes are targets for approved or experimental drugs, including *IL5RA* and *KCNH2*. Novel loci differentially methylated in newborns represent potential biomarkers of risk of asthma by school age. Cross-sectional associations in children can reflect both risk for and effects of disease. Asthma-related differential methylation in blood in children was substantially replicated in eosinophils and respiratory epithelium.
Results specifically related to Task 8.4: For the newborn analysis, all 7 significant CpGs were near a transcription factor binding site, and 6 were in a DNase hypersensitivity site identified in at least 1 ENCODE cell line, supporting a potential functional relevance to transcriptional activity. Among the 179 CpGs significantly differentially methylated in childhood in relation to asthma, there was significant depletion of localization to CpG islands (17 CpGs, 9.5%, p-value = 1.09E-11) and promoters (34 CpGs, 19.0%, p-value = 1.10E-04). Among the 179 CpGs, 113 were in DNAse hypersensitivity sites. Using eFORGE to examine enrichment of all 179 significant CpGs for histone marks, we found significant enrichment for H3K4me1 in blood and lung tissue and H3K36me3 in blood.
For the CpGs and regions we identified as differentially methylated in either newborns or children in relation to asthma, we assessed association between paired levels of blood DNA methylation and whole-blood gene expression for nearby transcripts defined as within a 500-kb window of the significant CpG or DMR in newborns (Gene Expression Omnibus, n = 38; INMA, n = 113; IoW-3, n = 157), children (4-year-olds in INMA, n = 112;
16-year-olds in BAMSE, n = 248), and adults (BIOS Consortium, n = 3096). Among 9 CpGs differentially methylated in newborns in relation to asthma, 3 were associated with expression of a nearby transcript in 3 data sets (cg17333211 in newborns, 4-year-olds, and adults and cg02331902 and cg07156990 in 2 newborn data sets and 4-year-olds), and an additional 3 CpGs were associated with expression in 2 data sets (cg13427149 in 16-year-olds and adults and cg13289553 and cg21486411 in newborns and 4-year-olds. All regions differentially methylated in newborns in relation to asthma were related to expression in at least 1 data set. For methylation in childhood, nearly all (176/179) CpGs related to asthma also associated with expression in at least 1 data set. CpGs annotated to *IL5RA* were significantly associated with expression in 4 cohorts (BIOS Consortium, INMA, IoW, and BAMSE). All 36 regions differentially methylated in childhood were associated with expression of a nearby transcript in at least 1 data set.
Using Ingenuity Pathway Analysis (IPA), we identified pathways, as well as disease processes and biological functions, significantly enriched (p-value <0.05) for the genes to which

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

significant individual CpGs or DMRs annotated in the meta-analysis of asthma in relation to newborn or childhood methylation. Genes to which the 7 significant CpGs and 35 significant DMRs in newborn methylation analysis were annotated were significantly enriched (p-value <0.05) for canonical pathways relevant to immune function in asthmatic patients, including endothelial nitric oxide synthase (eNOS) signaling, the inflammasome, and nuclear factor kB (NF-kB) signaling. Enriched disease processes and biologic functions included several involving immune function and others involving immune and organ development. Given the larger number of implicated genes for childhood methylation, many more pathways, disease processes, and biological functions were enriched. There was substantial overlap in newborns and children in the significantly enriched pathways and diseases and biological function relevant to immune function, immunologic disease, and development. As an example, **Figure 7** shows the network of 4 overlapping disease and biological processes between newborns and children: tissue morphology, immunological disease, inflammatory disease, and cell-mediated immune response.

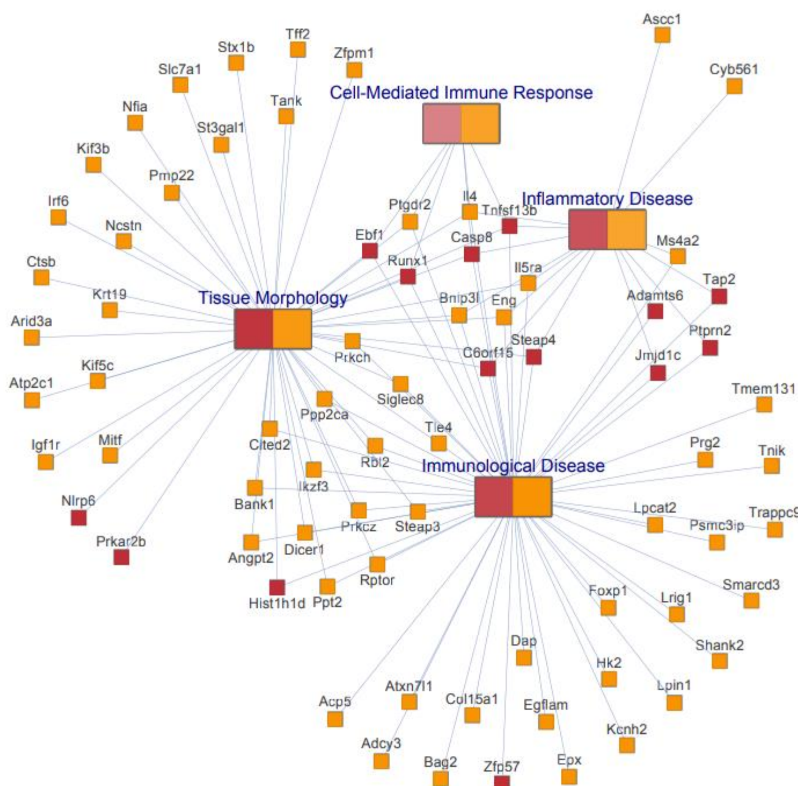

Figure 7. A network is shown for 4 categories of disease and biological functions overlapping between analyses of asthma in relation to either newborn or childhood methylation: immunological disease, cell-mediated immune response, inflammatory disease, and tissue morphology. A gene is connected to a disease or function if it has been previously shown to be involved in it. All genes marked in red are implicated from newborn methylation analyses, and those marked in in orange are implicated from childhood methylation analyses.

Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu
This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

2.4 Transcriptome-wide association studies (TWAS)

Early-life exposures and childhood gene expression in the Generation R Study

Status: Ongoing
Partner(s) involved: ERASMUS
Background: Several early-life exposures have been associated with DNA methylation at birth or in childhood (28,32,40). From the HELIX project (https://www.biorxiv.org/content/10.1101/2020.11.05.368076v2.external-links.html), it is known that DNA methylation is associated with gene expression in childhood blood. There is limited knowledge on associations of early-life exposures with RNA expression in childhood. We therefore aimed to explore if well-established early-life exposures that are known to be associated with cord blood DNA methylation and partially also with DNA methylation in childhood, are also associated with gene expression in childhood.
Methods: In the Generation R Study, RNA sequencing was performed in a subgroup of 184 children at a mean age of 9y. Total RNA for 62,446 transcripts was sequenced using whole blood. Total RNA was depleted of globin transcripts using the Ambion GLOBINclear kit and subsequently processed for sequencing using the Illumina TruSeq RNA Library Preparation Kit. Paired-end sequencing of 2 × 150-bp reads was performed using the Illumina HiSeq 4000 platform. The expression matrix was normalized using *edgeRs* trimmed mean of M-values (TMM) methodology. After QC, information on 20,011 transcripts was available, we applied Bonferroni correction, resulting in a p-value threshold of 2.5E-6.
We used robust linear regression modelling for all analyses and for each transcript individually. We were interested in three early-life exposures: Maternal smoking during pregnancy, maternal pre-pregnancy BMI, and birthweight. Models were adjusted for child sex, age at gene expression measurement, maternal education, maternal age at inclusion, parity, batch, 5' to 3' bias, and cell type proportions. The model for maternal smoking was adjusted for maternal BMI, and vice versa. The model for birthweight was adjusted for both maternal smoking and maternal BMI.

Transcriptome-wide association study of body mass index and related traits in children

Status: Ongoing
Partner(s) involved: ISGLOBAL and ERASMUS (leaders), BTHFT, UOC, NIPH, INSERM
Background: Transcriptome-wide analysis (TWAS), especially in children, are a developing field, timely relevant to shed light on biological pathways of common phenotypes in childhood. We will conduct a cross-sectional TWAS of BMI and BMI-related phenotypes in late childhood (age 8-10y) using data from the HELIX project and Generation R.
Methods: HELIX consists of 1,300 children from 6 European birth cohort studies, 5 of which are also part of LifeCycle project: BiB, EDEN, INMA, MoBa, and RHEA. Blood genome-wide gene expression data was obtained with the Affymetrix HTAv2 array at the age of 8y-10y and BMI and related traits (waist circumference and fat mass) were assessed at the same age using a harmonized protocol across the cohorts. First, we will test the association between BMI and main covariates, including cell type proportions. Then, we will run multivariable linear regression models, adjusted for child sex, age, cohort, genome-wide genetic principal components, batch and cell type proportions, modelling BMI and related

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

phenotypes as the exposure and gene expression as the outcome. Multiple testing correction will be applied using the FDR method. *In silico* functional enrichment analyses will be conducted with top genes in order to understand which pathways are affected. Finally, the findings will be replicated in the Generation R Study, where blood mRNA sequencing was performed at the 9 year-follow-up in a total of 184 children.
This project will be co-led by ISGlobal and Generation R, thanks to a LifeCycle fellowship.

3. General discussion

This general discussion will start by commenting the HELIX *cis*-eQTM catalogue in child blood, and then, the results of the biological interpretation of the 18 selected EWAS from LifeCycle.

3.1 Building expression quantitative trait methylation (eQTM) catalogues

We have created a catalogue of *cis*-eQTMs in child blood using HELIX DNA methylation and gene expression data, which is publicly available (https://helixomics.isglobal.org/). The catalogue includes the results of 13 M CpG-TC (gene) associations, 39,749 of which pass the multiple-testing correction threshold, representing 21,966 unique CpGs and 8,886 unique TCs. In the HELIX eQTM catalogue, only around half of the genes could be captured through annotation of the CpG to the closest gene using the Illumina annotation, highlighting the importance of incorporating eQTM gene identification in EWAS downstream analyses. Therefore, the HELIX eQTM catalogue will become an essential tool for biological interpretation of EWAS conducted in children, and indeed, it has already been used in 5/18 EWAS included in this report (see below).

Two aspects of the HELIX eQTMs that are relevant for biological interpretation are:

1) **Genetic contribution to eQTMs:** We found that a high proportion of eQTMs were influenced by genetic variants indicating the importance of investigating the effect of genetics in combination with early-life stressors in EWAS studies.
2) **eQTM age specificity:** The overlap between *cis*-eQTMs described in child blood (HELIX catalogue) and in adult blood (Kennedy catalogue) was small (41). This can be explained by methodological differences between the two studies, such as gene expression platforms with low overlap; statistical methods and statistical power; cohort-specific environmental exposures; cellular composition; as well as by real differences in the relationship between DNA methylation and gene with age. The short list of age-shared eQTMs tended to encompass CpGs located in promoters and regulated by genetic variants. This could represent a specific characteristic of eQTMs that are persistent over time or it also could be that this kind of eQTMs (genetically regulated and close to the TSS) are easier to be detected and shared among any two studies because they show stronger effects, regardless of age.

Despite the tremendous utility of the HELIX eQTM catalogue, it also has some limitations which have to be considered:

1) **Array coverage:** It only covers a fraction of all CpG-TC pairs, as both the methylation and gene expression arrays have limited resolution. Nonetheless, the catalogue will be useful

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

for most researchers as the 450K methylation array is widely used, and the gene expression array covers almost all the coding genes.

2) **Sex chromosomes:** The catalogue does not include sex chromosomes which require more complex analyses to address X-inactivation and sex-specific effects that will be investigated in future studies.
3) **Only *cis* effects:** Due to statistical power limitations, only *cis* effects were tested. Despite that, we observed that CpGs tended to be close to the gene they regulate, so the catalogue is expected to cover most of the CpG-TC associations.
4) **Cellular composition:** Models were adjusted for blood cell type composition and, while this has allowed us to control for major differences in methylation and gene expression among blood cell types, it might also have resulted in over-adjustment in some CpG-TC pairs. Moreover, the analysis of bulk data might have limited the identification of eQTMs specific to a subset of blood cell types, the identification of which would need more sophisticated statistical and/or experimental methods.

3.2 Functional interpretation of EWAS of early-life stressors and of health outcomes

We have screened 18 LifeCycle EWAS projects that have investigated biological pathways. Nine of them are related to early life stressors and nine to offspring health outcomes. Below we comment on the main observations:

1) **eQTM annotation:** 11/18 EWAS conducted eQTM annotation using different catalogues or datasets (listed in Section 2. Introduction and Methodology).
   - **Positional versus eQTM annotation:** As expected from the results of the HELIX catalogue, there was limited overlap between positional and eQTM annotated genes.
   - **eQTM age specificity:** Consistent with HELIX observations, no much overlap was found across eQTM catalogues of different age ranges. This was the case for the EWAS of birth weight (32) and asthma (39) that examined blood eQTM catalogues built with samples obtained at birth, childhood, adolescence and adulthood. However, we have to acknowledge that a fair comparison was not possible due to the different sample size and thus statistical power of each eQTM catalogue. However there were eQTM genes consistently identified at all age ranges, such as the *Interleukin 5 Receptor Subunit Alpha* (*IL5RA*) gene in the EWAS of asthma (39). Interestingly, *IL5RA*, already has an approved asthma drug that inhibits its product.
   - **eQTM tissue specificity:** It is well documented that DNA methylation and gene expression are tissue specific, and thus only partial overlap of eQTMs across tissues is expected. For instance, there was no overlap between the genes related to maternal glycemic load identified through eQTM annotation using datasets from two different tissues: blood (HELIX catalogue at 8y) and adipose tissue (Leipzig Childhood AT Cohort with samples collected at the age of 0-21y).

   Overall, this suggests the importance of using eQTM catalogues built in similar tissues and with similar age ranges as the samples of the target EWAS. In addition, these

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu
This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

catalogues should be built with a relatively large number of samples to provide robust CpG-gene associations.

2) **CpG gene relative position and gene expression:** The relationship between DNA methylation and gene expression is complex and dependent on the genomic context. In other words, hyper-methylated states cannot be assumed to be repressive of gene expression and hypo-methylated to be active. For instance, maternal smoking in pregnancy was associated with hyper- and hypo-methylation of CpGs in the *AHRR* locus (24). Hyper-methylated CpGs, located in intron 1, were positive eQTMs for *AHRR*, *PDCD6*, and *EXOC3* gene; while hypo-methylated CpGs, located in other introns, were inverse eQTMs for the same genes (in both cases implying higher expression of the genes in response to tobacco smoking).

3) **Functional enrichment with genes identified through positional annotation:** 15/18 EWAS performed functional enrichment analyses. 8/18 EWAS were not able to identify any pathway surviving multiple-testing correction. Diverse non-overlapping pathways were identified in the other EWAS (7/18). The number of pathway identified tended to be dependent on the number of genes used as input in the enrichment analyses (thus the of statistically significant CpGs). Moreover, missMethyl R package (13), which corrects for array background, retrieved less pathways than ConsensusPathDB (12). KEGG pathways (16) seemed to be more specific and easier to interpret than GO terms (15). Identified pathways were diverse, and while some of them were specific and seemed relevant for the exposure or trait under investigation (e.g. nuclear factor kB (NF-kB) signaling in asthma (39)); others were less informative generic biological functions (e.g. DNA- and RNA-regulation identified in lung function (38)).

4) **Functional enrichment with genes identified through eQTM annotation:** The study of the exposome and child epigenetic age was the only one to use eQTM genes as input in the pathway analyses (34). The EWAS of lung function used a combination of both: positional and eQTM genes (38). The rest of EWAS conducted functional enrichment analysis using only positional annotated genes. It would be interesting that future LifeCycle projects run enrichment analysis with both list of annotated genes.

5) **Molecular enrichment analyses:** 6/18 EWAS conducted enrichment for tissue specific regulatory regions using the eFORGE (18), eFORGE-TF (19) tools or Chi2 tests with ROADMAP annotation of chromatin states. This is a good strategy to complement functional enrichment analyses.

6) **Overlap with the literature:** The EWAS of aggressive behaviour (37) searched for information of the top CpGs in the EWAS Atlas (17). Although this is not a formal enrichment analysis it provided interesting results: CpGs associated with aggressive behaviour had previously been related to smoking, other chemical exposures (PCBs and PCDFs), alcohol consumption, cognitive function, educational attainment, ageing, and metabolic traits. More research is needed to see if there is a causal link between these

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

exposures and aggression or if they reflect residual confounding. Given the amount of EWAS uploaded in the EWAS Atlas and Catalog (http://ewascatalog.org/), it would be interesting to systematically compare findings with these databases.

3.3 Transcriptome-wide association studies (TWAS)

Only 2/18 EWAS explored the association between the exposure/trait of interest and gene expression in the same children where they investigated DNA methylation changes. Gruzieva et al. studied the association between exposure to air pollution and gene expression in cord blood and adolescent's blood, and found some genes differently expressed in each dataset (cross-sectional association in adolescents) (25).
Vives-Usano et al. did not find any gene differently expressed in child's blood in response to maternal smoking after adjusting for FDR, in contrast to the 18 differently methylated loci that were identified (24). DNA methylation in these loci was related to the expression of nearby genes. The weak association of maternal smoking in pregnancy with child gene expression versus the stronger and persistent effects on DNA methylation suggests different reversal rates and a methylation-based memory to past exposures. Therefore, it will be advisable to acquire both DNA methylation and RNA expression data from the same children to be able to investigate the complex correlations between them and with early-life stressors and health outcomes.

Contrary to the HELIX null results, preliminary results from Generation R suggest an association between maternal smoking and *RP11-19J3.7* gene. Moreover, maternal pre- or early-pregnancy BMI was associated with *SPRYD4* and *NLRP2* (p-value = 2.4E-6).
During the final period of LifeCycle, we will exploit available gene expression data in LifeCycle cohorts (INMA, HELIX and Generation R). We will start by conducting a TWAS of BMI-related traits in HELIX and Generation R. Given that gene expression was quantified using different technologies in these two cohorts (arrays versus next generation sequencing), the project will also enhance the knowledge on how to best combine transcriptomic data in multi-cohort settings.

4. Conclusion

In summary, 18 EWAS of LifeCycle investigated functional consequences of DNA methylation changes in response to early-life stressors and health outcomes. CpGs were annotated to genes by proximity or using public eQTM catalogues. Around half of the EWAS identified biological pathways: some of them of relevance for the trait/exposure under investigation and others less informative. Not much overlap was observed among eQTM genes retrieved from different catalogues, in part due to limited sample size and different age range and tissue type. To overcome this, we developed a blood cis-eQTM catalogue using data from ~900 HELIX children, which is publicly available. Finally, there were only few TWAS conducted in LifeCycle and with limited sample sizes. Next steps will explore how to combine gene expression data obtained from different platforms in LifeCycle cohorts.

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

5. Contribution of partners

All partners have contributed producing results for this deliverable:

- **ISGLOBAL:** Leader of the cis-eQTM catalogue and of the EWAS of maternal smoking and child epigenetic age. Participated in the following EWAS: air pollution, mercury, vitamin B12, Mediterranean diet, glycemic load, maternal BMI, GA, BW, child BMI, ADHD, aggression lung function and asthma.
- **BTHFT:** Participated in the cis-eQTM cataloguen and in the following EWAS: maternal smoking, child BMI and epigenetic age.
- **UOC:** Participated in the cis-eQTM cataloguen and in the following EWAS: maternal smoking, air pollution, mercury, GA, BW, child BMI, epigenetic age, and aggression.
- **NIPH:** Participated in the cis-eQTM cataloguen and in the following EWAS: maternal smoking, mercury, vitamin B12, maternal BMI, HPD, GA, BW, child BMI, epigenetic age, and asthma.
- **INSERM:** Participated in the cis-eQTM cataloguen and in the following EWAS: maternal smoking, air pollution, maternal BMI, glycemic metabolims, GA, child BMI, epigenetic age, and asthma.
- **ERASMUS:** Leader of the following EWAS: vitamin B12, Mediterranean diet, glycemic load, glycemic metabolism, BW, child BMI, ADHD, bulling, and lung function. Participated in the following EWAS: air pollution, maternal BMI, HPD, GA, aggression, and asthma.
- **UNITO:** Participation in the EWAS of air pollution.
- **UNIVRIS:** Leader of the following EWAS: maternal BMI, HPD, BW, child BMI, and bulling. Participated in the following EWAS: air pollution, mercury, vitamin B12, Mediterranean diet, glycemic load, GA, ADHD, aggression, lung function, and asthma.
- **UMCG:** Leader of the EWAS of BW and participated in the following EWAS: maternal BMI, GA, aggression, and asthma.
- **UWA:** Participated in the following EWAS: maternal BMI, GA, BW, child BMI, and asthma.
- **UOULU:** Participated in the following EWAS: glycemic metabolism, GA, BW, child BMI, aggression, and asthma.
- **LMU:** Participated in the following EWAS: child BMI and asthma.

The report was written by ISGLOBAL and ERASMUS and reviewed by all the partners.

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu
This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

6. Deviations from original plan

This deliverable has been fulfilled fully in line with the original plan. However, in the last 2 years of the work, COVID-19 related issues affected the ability of LifeCycle members to meet face to face, and in some cases to dedicate as much time to usual research activities. Nevertheless, much of the work continued by telecommunication.

7. Dissemination activities

We presented some of the work in this deliverable as poster presentations at the:

Epigenomics of Common Diseases Conference in November 2020

- Identification of blood autosomal cis-expression quantitative trait methylation (cis-eQTM) in children. Poster presented by Mariona Bustamante (ISGLOBAL)
- The early-life exposome and epigenetic age acceleration in children. Poster presented by Paula de Prado (ISGLOBAL)

Epigenomics of Common Diseases Conference in November 2021

- Maternal Mediterranean diet in pregnancy and newborn DNA methylation: a meta-analysis in the PACE Consortium. Poster presented by Leanne Küpers (ERASMUS).
- Maternal dietary glycemic index and glycemic load in pregnancy and offspring cord blood DNA methylation. Poster presented by Leanne Küpers (ERASMUS).
- Pregnancy vitamin B12 concentrations and offspring cord blood DNA methylation. Poster presented by Giulietta Monasso (ERASMUS).

8. References

1. Hansen KD. IlluminaHumanMethylation450kmanifest: Annotation for Illumina's 450k methylation arrays. R. Package Version 0.4.0. 2012.
2. Heavner K, Newschaffer C, Hertz-Picciotto I, Bennett D, Burstyn I. Quantifying the potential impact of measurement error in an investigation of autism spectrum disorder (ASD). J Epidemiol Community Health. 2014;68(5):438-45.
3. Rojas D, Rager JE, Smeester L, Bailey KA, Drobná Z, Rubio-Andrade M, et al. Prenatal arsenic exposure and the epigenome: identifying sites of 5-methylcytosine alterations that predict functional changes in gene expression in newborn cord blood and subsequent birth outcomes. Toxicol Sci. 2015;143(1):97-106.
4. Arshad SH, Karmaus W, Zhang H, Holloway JW. Multigenerational cohorts in patients with asthma and allergy. J Allergy Clin Immunol. 2017;139(2):415-21.
5. Moore SE, Fulford AJC, Darboe MK, Jobarteh ML, Jarjou LM, Prentice AM. A randomized trial to investigate the effects of pre-natal and infant nutritional supplementation on infant immune development in rural Gambia: The ENID trial: Early Nutrition and Immune Development. BMC Pregnancy Childbirth. 2012;12:107.
6. Guxens M, Ballester F, Espada M, Fernández MF, Grimalt JO, Ibarluzea J, et al. Cohort profile: The INMA-INfancia y Medio Ambiente-(environment and childhood) project. Int J Epidemiol. 2012;41(4):930–40.
7. Maitre L, de Bont J, Casas M, Robinson O, Aasvang GM, Agier L, et al. Human Early Life Exposome (HELIX) study: a European population-based exposome cohort. BMJ Open. 2018;8(9):e021311.

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

8. Wickman M, Kull I, Pershagen G, Nordvall SL. The BAMSE project: Presentation of a prospective longitudinal birth cohort study. Pediatric Allergy and Immunology. 2002;13(s15):11-3.
9. Bonder MJ, Luijk R, Zhernakova D V., Moed M, Deelen P, Vermaat M, et al. Disease variants alter transcription factor levels and methylation of their binding sites. Nat Genet. 2017;49(1):131-8.
10. Landgraf K, Rockstroh D, Wagner I V., Weise S, Tauscher R, Schwartze JT, et al. Evidence of early alterations in adipose tissue biology and function and its association with obesity-related inflammation and insulin resistance in children. Diabetes. 2015;64(4):1249-61.
11. Margerie D, Lefebvre P, Raverdy V, Schwahn U, Ruetten H, Larsen P, et al. Hepatic transcriptomic signatures of statin treatment are associated with impaired glucose homeostasis in severely obese patients. BMC Med Genomics. 2019;12(1):80.
12. Kamburov A, Stelzl U, Lehrach H, Herwig R. The ConsensusPathDB interaction database: 2013 Update. Nucleic Acids Res. 2013;41(D1).
13. Phipson B, Maksimovic J, Oshlack A. MissMethyl: An R package for analyzing data from Illumina's HumanMethylation450 platform. Bioinformatics. 2016;32(2):286-8.
14. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc [Internet]. 2009;4(1):44–57.
15. Carbon S, Douglass E, Dunn N, Good B, Harris NL, Lewis SE, et al. The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res. 2019;47(D1):D330-8.
16. Tanabe M, Kanehisa M. Using the KEGG database resource. Curr Protoc Bioinforma. 2012;Chapter 1:Unit 1.12.
17. Li M, Zou D, Li Z, Gao R, Sang J, Zhang Y, et al. EWAS Atlas: A curated knowledgebase of epigenome-wide association studies. Nucleic Acids Res. 2019;47(D1):D983-8.
18. Breeze CE, Paul DS, van Dongen J, Butcher LM, Ambrose JC, Barrett JE, et al. eFORGE: A Tool for Identifying Cell Type-Specific Signal in Epigenomic Data. Cell Rep. 2016;17(8):2137–50.
19. Breeze CE, Reynolds AP, Van Dongen J, Dunham I, Lazar J, Neph S, et al. EFORGE v2.0: Updated analysis of cell type-specific signal in epigenomic data. Bioinformatics. 2019;35(22):4767-9.
20. Ernst J, Kellis M. Chromatin-state discovery and genome annotation with ChromHMM. Nat Protoc. 2017;12(12):2478-92.
21. Kooijman MN, Kruithof CJ, Van Duijn CM, Duijts L, Franco OH, Van IJzendoorn MH, et al. The Generation R Study: design and cohort update 2017. Eur J Epidemiol 2016;31(12):1243-64.
22. Ruiz-Arenas C, Hernandez-Ferrer C, Vives-Usano M, Mari S, Quintela I, Mason D, et al. Identification of autosomal cis expression quantitative trait methylation (cis eQTMs) in children's blood. Elife. 2022;11:e65310.
23. van Dongen J, Nivard MG, Willemsen G, Hottenga J-J, Helmer Q, Dolan C V, et al. Genetic and environmental influences interact with age and sex in shaping the human methylome. Nat Commun. 2016;7:11115.
24. Vives-Usano M, Hernandez-Ferrer C, Maitre L, Ruiz-Arenas C, Andrusaityte S, Borràs E, et al. In utero and childhood exposure to tobacco smoke and multi-layer molecular signatures in children. BMC Med. 2020;18(1):243.
25. Gruzieva O, Xu C-J, Yousefi P, Relton C, Merid SK, Breton C V, et al. Prenatal particulate air pollution and DNA methylation in newborns: An epigenome-wide meta-analysis. Environ Health Perspect. 2019;127(5):57012.
26. Lozano M, Yousefi P, Broberg K, Soler-Blasco R, Miyashita C, Pesce G, et al. DNA methylation changes associated with prenatal mercury exposure: A meta-analysis of prospective cohort studies from PACE consortium. Environ Res. 2022;204(Pt B):112093.

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

27. Küpers LK, Fernández-Barrés S, Nounu A, Friedman C, Fore R, Mancano G, et al. Maternal Mediterranean diet in pregnancy and newborn DNA methylation: a meta-analysis in the PACE Consortium. Epigenetics. 2022;1-3.
28. Sharp GC, Salas LA, Monnereau C, Allard C, Yousefi P, Everson TM, et al. Maternal BMI at the start of pregnancy and offspring epigenome-wide DNA methylation: Findings from the pregnancy and childhood epigenetics (PACE) consortium. Hum Mol Genet. 2017;26(20):4067–85.
29. Kazmi N, Sharp GC, Reese SE, Vehmeijer FO, Lahti J, Page CM, et al. Hypertensive Disorders of Pregnancy and DNA Methylation in Newborns: Findings from the Pregnancy and Childhood Epigenetics Consortium. Hypertension. 2019;74(2):375-83.
30. Tobi EW, Juvinao-Quintero DL, Ronkainen J, Ott R, Alfano R, Canouil M, et al. Maternal Glycemic Dysregulation During Pregnancy and Neonatal Blood DNA Methylation: Meta-analyses of Epigenome-Wide Association Studies. Diabetes Care. 2022;45(3):614-23.
31. Merid SK, Novoloaca A, Sharp GC, Küpers LK, Kho AT, Roy R, et al. Epigenome-wide meta-analysis of blood DNA methylation in newborns and children identifies numerous loci related to gestational age. Genome Med. 2020;12(1):1–17.
32. Küpers LK, Monnereau C, Sharp GC, Yousefi P, Salas LA, Ghantous A, et al. Meta-analysis of epigenome-wide association studies in neonates reveals widespread differential DNA methylation associated with birthweight. Nat Commun. 2019;10(1):1893.
33. Vehmeijer FOL, Küpers LK, Sharp GC, Salas LA, Lent S, Jima DD, et al. DNA methylation and body mass index from birth to adolescence: meta-analyses of epigenome-wide association studies. Genome Med. 2020;12(1):105.
34. de Prado-Bert P, Ruiz-Arenas C, Vives-Usano M, Andrusaityte S, Cadiou S, Carracedo Á, et al. The early-life exposome and epigenetic age acceleration in children. Environ Int. 2021;155:106683.
35. Neumann A, Walton E, Alemany S, Cecil C, González JR, Jima DD, et al. Association between DNA methylation and ADHD symptoms from birth to school age: a prospective meta-analysis. Transl Psychiatry. 2020;10(1):398.
36. Mulder RH, Walton E, Neumann A, Houtepen LC, Felix JF, Bakermans-Kranenburg MJ, et al. Epigenomics of being bullied: changes in DNA methylation following bullying exposure. Epigenetics. 2020;15(6-7):750-64.
37. van Dongen J, Hagenbeek FA, Suderman M, Roetman PJ, Sugden K, Chiocchetti AG, et al. DNA methylation signatures of aggression and closely related constructs: A meta-analysis of epigenome-wide studies across the lifespan. Mol Psychiatry. 2021;26(6):2148-62.
38. den Dekker HT, Burrows K, Felix JF, Salas LA, Nedeljkovic I, Yao J, et al. Newborn DNA-methylation, childhood lung function, and the risks of asthma and COPD across the life course. Eur Respir J. 2019;53(4):1801795.
39. Reese SE, Xu C-J, den Dekker HT, Lee MK, Sikdar S, Ruiz-Arenas C, et al. Epigenome-wide meta-analysis of DNA methylation and childhood asthma. J Allergy Clin Immunol. 2019;143(6):2062–74.
40. Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, et al. DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. Am J Hum Genet. 2016 Apr;98(4):680–96.
41. Kennedy EM, Goehring GN, Nichols MH, Robins C, Mehta D, Klengel T, et al. An integrated -omics analysis of the epigenetic landscape of gene expression in human blood cells. BMC Genomics. 2018;19(1):476.

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

Appendix 1

Biological interpretation of EWAS of early-life stressors

Class of early life factor	Early life factor	Leader	Publication	Cohorts	Tissue and age	Number of CpGs	Gene positional mappping	Gene eQTM mapping	Functional pathways (positional genes)	Molecular pathways (positional genes)	Functional pathways (eQTM genes)	TWAS
Environmental pollutants	Air particles (PMs) in pregnancy	Gruzieva et al.	PMID: 31148503	INMA, Generation R, CHS, EARLI, PRISM, Project Viva, ENVIRONAGE, Piccolipiù, Rhea	Cord/child blood	**Cord blood:** 6 (PM10); 14 (PM2.5)	**Cord blood:** GNB2L1, SNORD96A, USP43, FAM13A, SRPRB, NOTCH4, P4HA2 (for PM10); PLXNA4, ZNF705A, FNIP1, COL22A1, TMCO3, SFRS8, MRI1, PSG5, C7orf50, PLAT, ZNF695 (for PM2.5)	**BIOS eQTM catalogue (adult blood):** AP3M2, USP43, MRI1, FAM13A, PLAT, RER1, TRIM52-AS1, TIFAB, SKI, ACSL6	**Cord blood** (CpGs at p-value<1E-05, ConsensusPathDB, GO)**:** Notch signaling pathway, Rho GTPase cycle, Neurotransmitter release cycle, GABA synthesis, release, reuptake and degradation (for PM10); none (for PM2.5). No replication with missMethyl.	-	-	**Cord blood (EARLI):** ZNF695 (for PM2.5); none (for PM10); **Adolescent's blood (BAMSE)** (cross-sectional exposure): NOTCH4, USP43 (for PM10); C7orf50 (for PM2.5)
Environmental pollutants	Maternal tobbaco smoke	Vives-Usano et al.	PMID: 32811491	HELIX	Child blood	**Child blood:** 41 (18 loci)	**Child blood:** AGBL4, GNG12, GFI1, NDST4, AHRR, PRRT1, RADIL, MYO1G, CNTNAP2, ZNF395, OLFM1, FRMD4A, FMN1, CYP1A1, RUNX1	**HELIX eQTM catalogue (child blood, FDR):** PNOC, LOC285957, RUNX1-IT1, ULK3, EXOC3; **HELIX eQTM catalogue (child blood, p-value<0.05):** 65 other eQTMs.	-	-	-	**Child blood (HELIX):** none
Environmental pollutants	Exposure to mercury (Hg) in pregnancy	Lozano et al.	PMID: 34562483	INMA, KOREA, VIVA, SAPPORO-HOKKAIDO, ALSPAC, MoBa, RHEA	Cord/child blood	**Cord blood** (suggestive p-value): 1; **child blood** (suggestive p-value): 2	**Cord blood:** MED31; **child blood**: GGH, GRK1	-	**Cord blood** (CpGs at p-value 1E-05, missMethyl, GO terms): none, however, top GO terms included vascular permeability, astrocyte activation and microtubule nucleation, and titin binding.	-	-	-
Diet	Maternal mediterraean diet in pregnancy	Küpers et al.	Submitted	ALSPAC, Generation R, INMA, Healthy Start (two sub cohorts: Hispanic and 158 non-Hispanic white participants), and Project Viva.	Cord/child blood	**Cord blood:** cg23757341; **child blood:** none	**Cord blood:** WNT5B	**HELIX eQTM catalogue (child blood):** none	**Cord blood** (CpGs at P<0.0001, missMethyl, KEGG and GO): none	**Cord blood** (CpGs at P<0.0001, eForge-tissue-specific DNaseI hypersensitivity regions): none	-	-
Diet	Maternal vitamin B12 in pregnancy	Monasso et al.	Submitted	ALSPAC, GENR, INMA, MARBLES, MoBa1 and MoBa2	Cord/child blood	**Cord blood:** 109 (for maternal B12) and 7 (for newborn's B12); **child blood:** 40.7% persitent (for maternal B12) and 57.1% persistent (for newborn's B12)	**Cord blood:** several genes	**HELIX eQTM catalogue (child blood):** 18 genes linked to 20/109 CpGs, PAX8 being the top **one** (maternal B12)	**Cord blood** (missMethyl, KEGG and GO): none	**Cord blood** (eForge-tissue-specific DNaseI hypersensitivity regions and chromatin states): some states; (eForge-TF transcription factors): MECP2 and DEAF1	-	-
Metabolic	Maternal glycemic index/load in pregnancy	Küpers et al.	Submitted	ALSPAC, Generation R, INMA	Cord/child/adolescent blood	**Cord blood:** 41 CpGs, mostly in mothers with overweight/obesity; **Child/adolescent blood:** none	**Cord blood:** C3orf70, SCARNA2 SORCS3, MYH6, C2orf69, MFHAS1, AMIGO2, ARHGEF26, JPH4, EXOSC9, MOK. SOX5, NDUFA5, C14orf13, ST3GAL6-AS1, GPR108, GPHN, ACOT13, ESPL1, CCT6B, DPPA3, HOMER1, LRPPRC, TSPAN15, RAP1GAP2, MRPS26, MED15, ZNF830, LRCH3, FARSA, PCDHGA1, MGRN1, C1QBP, AP3M2, AHDC1, KPNA6, HBQ1, CDH4, KATNAL2, LOC154449, ZNF664-RFLNA	**HELIX eQTM catalogue (child blood):** PCED1B and PCDHG; **Leipzig Childhood AT Cohort eQTM catalogue (adipose tissue from 0-21y):** TFAP4, ZNF500, PPL and ANKS3	**Cord blood** (CpGs at different p-values, missMethyl, KEGG and GO): none	**Cord blood** (eForge-tissue-specific DNaseI hypersensitivity regions): none	-	-
Metabolic	Maternal BMI at the start of pregnancy	Sharp et al.	PMID: 29016858	ALSPAC, CBC HISPANICS, CBC CAUCASIANS, CHAMACOS, EARLI, GOYA, GECKO, Generation R, GEN3G, IOW F2, INMA, EDEN, MoBa1, MoBa2, MoBa3, NEST, NFCS, NHBCS, RICHS, Project Viva	Cord/child/adolescent blood	**Cord blood:** 104 (86 common with cell unadjusted model); **child blood:** 72/86 CpGs pesistent effects	**Cord blood:** The 86 maternal BMI associated CpGs were near 77 gene regions, and there were several instances where multiple sites mapped to the same gene: RBMS1 (3 sites), POM121L1P (3 sites), VIPR2 (2 sites), SQLE (2 sites), RASA3 (2 sites), MIR200B (2 sites), KAT6B (2 sites)	-	**Cord blood** (77 genes, missMethyl, GO and KEGG: none	-	None.	-
Metabolic	Hypertensive disorders of pregnancy	Kazmi et al.	PMID: 31230546	ALSPAC, Generation R, GOYA, Healthy start (Hispanic and non-Hispanic), IOW, MoBa1 and MoBa2, PREDO and Project Viva	Cord blood	**Cord blood:** 43 (for hypertenstion); 26 (for pre-eclampsia)	**Cord blood:** GLIS1, DHCR24, SMG7, APOB, CAND2, IQSEC1, BSN, MYLK, PLCH1, HTR3D, SC4MOL, CSF2, SPARC, ZC3H12D, TTLL11, ARID5B, CHST15, PRR5L, ARHGAP1, RBM19, NCOR2, DLEU7, STK24, CCDC33, SLC9A3R2, ABCC1, CACNB1, RARA, IGF2BP1, ARID3A, SBNO2, TCF3, AVP, PLCB4	-	**Cord blood** (GO terms or biological pathways): Organ and system development, regulation of cell communication, and cell differentiation	-	-	-
Metabolic	Glycemic disregulation	Tobi et al	Accepted, Diabetes Care	GUSTO, Gen3G, Healthy Start, FinnGeDi, PREDO, EDEN, ENVIRONAGE, Generation R	Cord blood	**Cord blood:** 1	**Cord blood:** TXNIP	**ABOS eQTM catalogue** (liver and muscle tissue): TXNIP (liver, but not muscle)	-	-	-	-

*If not stated, CpGs, eQTMs and pathways refer to the list corrected for multiple-testing

LIFECYCLE Project
Coordination and leadership:
Erasmus MC
Generation R (Na-2918)
PO box 2040
3000 CA Rotterdam
The Netherlands
Phone: +31 (0)10 704 3405
Email: lifecycle@erasmusmc.nl
Web: lifecycle-project.eu

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle).

Appendix 2

Biological interpretation of EWAS of health outcomes

Class of early life factor	Early life factor	Leader	Publication	Cohort	Tissue and age	Number of CpGs	Gene positional mappping	Gene eQTM mapping	Functional pathways (positional genes)	Molecular pathways (positional genes)	Functional pathways (eQTM genes)	TWAS
Reproductive	Gestational age	Merid et al.	PMID: 32114984	ALSPAC, CBC (Hispanic), CBC (White), CHAMACOS, EDEN, ENVIRONAGE, PiccoliPlus, RHEA, GECKO, Generation R, GEN3G, GOYA, INMA, IOWF2, MoBa1, MoBa2, MoBa3, PREDO, Project Viva, BAMSE, BAMSE-EpiGene, IOWF1, NFBC86, PIAMA, and RAINE	Cord/child/adolescent blood	**Cord blood:** 8899; **Longitudinally stable:** 222	**(Nearest genes within 10 Mb of each site) Cord blood:** 4966 genes; **Longitudinally stable:** 139 genes	**BEAR eQTM analyses (cord blood):** 200 genes	**Cord blood** (ConsensusPathDB; GO and KEGG): 1784 GO terms, regulation of cellular and biological processes, system development, different signalling pathways and organ development; 124 KEGG terms, various cancers, viral infections, metabolic processes and immune-related disorders; **Longitudinally stable:** 13 KEGG terms, infection- and immune-related disorders	-	-	-
Reproductive	Birth weight	Küpers et al.	PMID: 31015461	ALSPAC, CBC (Hispanic), CBC (Caucasian), CHAMACOS, CHS, EARLI, ENVIRONAGE, Piccolipiu, RHEA, GECKO, Generation R, GEN3G, GOYA, Healthy Start (Hispanic), Healthy Strat (Caucasian), Healthy Start (African American), INMA, IOWF2, MoBa1, MoBa2, MoBa3, NCL, NEST (African American), NEST (Caucasian), NHBCS, PREDO, PRISM, PROGRESS, RICHS, and Project Viva	Cord/child/adolescent blood	**Cord blood:** 914; **Child/adolescent blood:** <1.2% of 914	**Cord blood:** 729 genes	**BIOS eQTM catalogue (adult blood):** 85 genes; **INMA eQTM analyses (4y blood):** 17; **ENID trial in Gambia eQTM analyses (2y blood):** 2	**Cord blood** (914 CpGs, missMethyl, GO and KEGG): none	-	-	-
Metabolic	Body mass index	Vehmeijer et al.	PMID: 33239103	ALSPAC, BAMSE, CHAMACOS, CHOP, CHS, DOMInO Trial, GECKO, Generation R, GOYA, Healthy Start Study, HELIX, INMA, IOW F1, IOW F2, MoBa1, MoBa2, NEST, NFBC 1986, PIAMA, PREDO, Project Viva, RAINE and STOPPA	Cord/child/adolescent blood	**Cord blood - early BMI:** 0; **Cord blood - later BMI:** 1; **Child blood - BMI:** 1 (and 9 at FDR); **Adolescent blood - BMI:** 1	**Cord blood - early BMI:** 0; **Cord blood - later BMI:** SFRP5; **Child blood - BMI:** SLC43A2; **Adolescent blood - BMI:** SFXN5	-	**Cord blood** (CpGs at pvalue 1E-04, missMethyl, GO and KEGG): none	-	-	-
Aging	Pregnancy exposome vs epigenetic age acceleration	de Prado et al.	PMID: 34144479	HELIX	Child blood	Epigenetic age acceleration (Horvath clock): 391	-	**HELIX eQTM catalogue (child blood):** 129 genes associated with the methylation of 72/391 CpGs	-		**Child blood** (ConsensusPathDB; KEGG, Reactome and Biocarta): adaptive and innate immune system; apoptosis, cell cycle and cancer; and detoxification of xenoestrogens.	-
Neurobehaviour	Attention Deficit and Hyperactivity Disorder (ADHD)	Neumann et al.	PMID: 33184255	ALSPAC, Generation R, INMA, NEST, PREDO, GLAKU	Cord and child blood	**Cord blood**: 9; **Child blood:** 0	**Cord blood:** CREB5, ZBTB38, PPIL1, ERC2, TRERF1; **Child blood:** 0	-	**Cord blood** (missMethyl, KEGG and curated gene sets from the Molecular signatures database, CpGs at p-value <1E-05): none.	**Cord blood** (Chi2, ROADMAP blood chromatin states, regulatory regions Illumina): different patterns for hypo and hypermethylated CpGs.	-	-
Neurobehaviour	Bullying	Mulder et al.	PMID: 31992121	Generation R and ALSPAC	Cord blood and child/adolescent blood	**Longitudinal EWAS:** 1	**Longitudinal EWAS:** RAB14	-	**GO terms** (CpGs p-value < 0.001): several, including Ryanodine-sensitive calcium-release channel activity astrocyte differentiation, action potential regulation, and muscle fibre development; **Literature overlap:** No RAB14-associated probes were reported.	-	-	-
Neurobehaviour	Aggression	van Dongen et al.	PMID: 33420481	ALSPAC, Dunedin, E-Risk, FinnTwin12, GS:SFHS, GLAKU, HELIX, LLD, NFBC1966, NFBC1986, NTR, SATSA, YFS, GECKO, Generation R, INMA, Poseidon	Cord blood and peripheral blood (child/adolescent /adult blood)	**Peripheral blood:** 13 (BN) and 35 (FDR); **Cord blood:** 0; **Combined:** 13 (BN) and 43 (FDR).	**Peripheral blood:** several; **Cord blood:** 0; **Combined:** several.	**BIOS eQTM catalogue (adult blood):** 17, 15 CpGs and ten transcripts (FAM60A, SEMA7A, PPP1R15A, F2RL3, FLOT1, TUBB, LINC00243, AHRR, EXOC3, RP4-647J21.1 (novel transcript, overlapping MYO1G) and PLXNA2.	**EWAS atlas:** smoking, other chemical exposures (PCBs and PCDFs), alcohol consumption, cognitive function, educational attainment, ageing, and metabolic traits.	-	-	-
Respiratory	Lung function	van Dekker et al.	PMID: 30765504	ALSPAC, Generation R, INMA, CHS, and Project Viva.	Cord blood	**Cord blood:** 22, 15 and 22 DMRs with FEV1, FEV1/FVC ratio and FEF75%, respectively	**Cord blood:** HOXA, PAOX, LINC00602, ABCA7, PER3, CLCA1, VENTX, NUDT12, PTPRN2, and TCL1A.	**Rotterdam eQTM analyses (**adult blood): RP11-431K24.1, HOXA1, HOTTIP, SPRN, ZNF511; **INMA eQTM analyses** (4y blood): NUDT12, PER3, RP3-467I1.4, RNA5SP23, RP4-726F1.1, HOXA1, HOTTIP, CMBL, EVX1, HOXA4, HOXA7, TUBGCP2, RP11-122K13.12, VENTX, ECHS1, SPRN, ZNF511, TCL1A, CCDC85C	**Cord blood** (DAVID, KEGG and GO): regionalisation, DNA- and RNA-regulation and embryonic development, asthma. **Cord blood** (Genemania): physical and genetic interactions; **Cord blood** (OMIM, UniProt): respiratory development including alveogenesis, respiratory diseases and cellular immunity	**Cord blood:** CTCF-binding sites	Both position and functional genes were used in the functional enrichment.	-
Respiratory	Asthma	Reese et al.	PMID: 30579849	ALSPAC, CHS, EDEN, Generation R, GOYA, MoBa1, MoBa2, NEST, BAMSE-EpiGene, BAMSE-MeDALL, CHOP, GALA II; ICAC, NFBC-1986, PIAMA, Raine and STOPPA	Cord/child blood	**Cord blood:** 9; **Child blood:** 179	**Cord blood:** CLNS1A, MAML2/ Mir_548, GPATCH2/SPATA17, SCOC/LOC100129858, AK91866, SUB1, and WDR20; **Child blood:** several	**eQTM analyses in BEAR (cord blood), IoW-3 (cord blood), INMA (4y), BAMSE (16y), and BIOS (adults). Cord blood:** 6 CpGs related to expression**; Child blood:** 176 CpGs related to gene expression. **Top gene:** IL5RA.	**Cord blood** (IPA): endothelial nitric oxide synthase (eNOS) signaling, the inflammasome, nuclear factor kB (NF-kB) signaling, immune function and immune and organ development; **Child blood** (IPA): tissue morphology, immunological disease, inflammatory disease, and cell-mediated immune response (among many others).	**Cord blood** (eForge-tissue-specific DNaseI hypersensitivity regions): 7; **Child blood** (eForge-tissue-specific DNaseI hypersensitivity regions): 113	-	-

*If not stated, CpGs, eQTMs and pathways refer to the list corrected for multiple-testing