This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 774548.

EC Framework Programme for Research and Innovation

Horizon 2020

H2020-SFS-2017-2-RIA-774548-STOP:

Science & Technology in childhood Obesity Policy

Science and Technology in childhood Obesity Policy

Science & Technology in childhood Obesity Policy

Start date of project: 1st June 2018 Duration: 48 months

D3.2: Report on completion of analyses for the molecular signature for obesity and its validation, including microcirculation and telomere length

Author(s): Rossella Alfano, Evangelos Chandakas, Michelle Plusquin, Tim Nawrot, Oliver Robinson, Paolo Vineis, Imperial College London

Version: Final

Preparation date: 15/11/2019

Dissemination Level

PU	Public	☒
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Abbreviation	Definition
BMI	Body Mass Index
WHO	World Health Organisation
FDR	False Discovery Rate
CHOP	Childhood Obesity Programme
ESI	Electrospray Ionization
TPPO	Triphenylphosphine Oxide
SBO	Systolic blood pressure
DBP	Diastolic Blood Pressure
CRAE	Central Retinal Arteriolar Equivalent
AVR	Arterio-Venous Ratio
RCT	Randomised Control Trial

Science and Technology in childhood Obesity Policy

Table of contents

1 Introduction

Adiposity in childhood is related to the risk of morbidity (e.g. including the metabolic syndrome and type 2 diabetes) both during childhood and later in adult life, leading to greater risk of premature mortality. Body size in childhood results from the interaction between genetic make-up, lifestyle factors, obesogenic environments and social determinants. However, it is unclear which molecular processes are involved in the causal network. Here we propose the identification of molecular signatures in relation to childhood obesity.

Overall, in WP3 we take a hierarchical approach to the analysis of pathways to obesity:

1. Environmental context analysis, based on geospatial analysis ("urban exposome") to assess built, natural and food environment, deprivation, facilities, pollutants etc. (D 3.1)
2. Individual level factors including physical/sedentary activity, dietary patterns, socio-economic factors, prenatal and family exposures (D3.1 and D3.2)
3. Internal factors, including "molecular signatures" (epigenetics, metabolomics) and role of gut hormones (D3.2, D3.3, D3.4)

Strong emphasis is placed on the role of socio-economic status in pathways to obesity, and in a final causal assessment based on mediation analysis (D3.5).

In Task 3.2 we aim to identify a molecular signature of child obesity that **is easily measurable and robust, is related to obesogenic environments, is predictive of obesity onset (and possibly clinical gravity) and can be used as a surrogate of the outcome in interventions.** This has been done using already existing measurements obtained from several children cohorts (see Figure 1 below).

Extensive omic measurements (particularly epigenomics and UPLC-MS metabolomics) are available for approximately 2,000 children aged from 0 to 10 years from six different European countries. We will first analyse the association of omic measurements with indicators of obesity, and subsequently explore the intermediate molecular steps and pathways derived from this task in order to connect obesogenic environments to obesity onset (D3.5). The latter task will be performed via mediation analysis. The purpose is to identify pathways that are suitable for intervention; this includes reversibility of biomarker levels after cessation of exposure when multiple samples are available from the same subjects. Particular emphasis will be put on **causal reasoning** including the ability to exclude reverse causation, i.e. whether biomarkers are in the chain connecting obesogenic exposures with obesity or they are a manifestation of obesity itself.

In the present deliverable we aim to answer two practical questions for delivery to the following WPs:

Can we identify a limited number of molecules that are associated with childhood obesity?

Can we link those molecules to prenatal risk factors?

2 Epigenetic signatures

Methylation of DNA is an indicator of the modulation of gene expression, i.e. of the translation of the DNA sequence into proteins. DNA methylation is a manifestation of the interactions between genes and the environment. Previous studies (including our own) have found altered methylation in relation to adult obesity, with some genes identified as being involved in increased BMI, mainly as a consequence rather than a determinant of adiposity (Alfano et al, systematic review in preparation: “A systematic review of epigenetics and childhood obesity”).

Several studies on adiposity in children and methylation have been published recently in particular from the CHOP consortium [1]. In the latter, specific DNAm variants were identified to be associated with children BMI (n=212), fat-mass (n=230), fat-free-mass (n=120), fat-mass-index (n=24) and fat-free-mass-index (n=15). Probes in genes SNED1(IRE-BP1), KLHL6, WDR51A(POC1A), CYTH4-ELFN2, CFLAR, PRDM14, SOS1, ZNF643(ZFP69B), ST6GAL1, C3orf70, CILP2, MLLT4 and ncRNA LOC101929268 remained significantly associated after Bonferroni-correction of P-values. However, the direction of causality was unclear.

In a study based on the NEST cohort [2], DNA methylation was measured in relation to pre-pregnancy obesity in 187 mother-female and 173 mother-male offspring. Maternal pre-pregnancy obesity was associated with 876 CpGs in female and 293 CpGs in male offspring (false discovery rate <5%). Among female offspring, 57 CpG sites, including the top 18, mapped to the TAPBP gene, but none of the TAPBP CpG sites were replicated in the ALSPAC cohort.

Lillycrop et al [3] measured DNA methylation in umbilical cord from children in the Southampton Women's Survey cohort (n = 680), focusing on 5 CpGs within the **SLC6A4** gene previously identified as being associated with adult obesity. Lower methylation of one CpG in particular was associated with higher total fat mass at 4 years, total fat mass at 6-7 years and % fat mass at 6-7 years. They also measured methylation in adolescents in the Western Australian Pregnancy Cohort Study (n = 812), and in adipose tissue from lean and obese adults from the UK BIOCLAIMS cohort (n = 81). Hypomethylation of SLC6A4 was carried on to adiposity in adolescence and adulthood.

Within STOP WP3 we have conducted a systematic review (in preparation for publication: “A systematic review of epigenetics and childhood obesity”), as described below and in **Addendum 1** attached.

Within deliverable 3.2 we aim to identify signatures of cord blood DNA methylation related to obesity onset and child growth using data from six European population-based birth-cohorts (Figure 1). In order to identify a molecular signature of child adiposity that is easily measurable and robust we selected from the systematic review described below 966 CpGs that have been related to childhood body size measurements by previous studies.

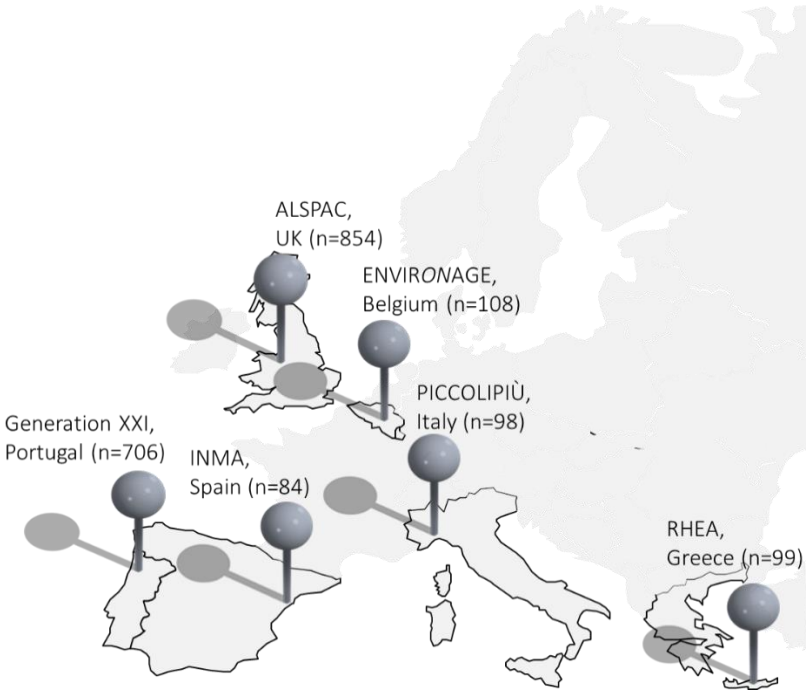


Figure 1 Study population

2.1 Methods

2.1.1 Population

Table 1 shows the general characteristics of the six cohorts under the epigenetics study: the Avon Longitudinal Study of Parents And Children (ALSPAC) cohort [4], the ENVironmental Influence ON early AGEing (ENVIRONAGE) cohort [5], the Generation XXI (GXXI) cohort [6], the INfancia y Medio Ambiente (INMA) cohort [7], the Piccolipiù cohort [8], and the Rhea cohort [9, 10].

Table 1 Population characteristics

	ALSPAC (n=854)	**ENVIRONAGE (n=111)**	**Generation XXI (n=706)**	**INMA (n=84)**	**Piccolipiù (n=98)**	**RHEA (n=92)**	**P value**
Sex (boys), n (%)	415 (48.6)	54 (48.6)	380 (53.8)	41 (48.8)	53 (54.1)	51 (55.4)	0.334
Birthweight (gr), mean (SD)	3491.35 (477.99)	3374.01 (506.55)	3223.20 (457.99)	3300.60 (399.89)	3221.07 (433.28)	3266.63 (438.85)	<0.001
Gestational age (weeks), mean (SD)	39.57 (1.49)	39.02 (1.70)	38.82 (1.44)	39.85 (1.56)	39.58 (1.57)	38.53 (1.28)	<0.001
Maternal age (years), mean (SD)	29.63 (4.40)	29.15 (3.57)	29.56 (5.23)	31.76 (4.07)	33.32 (4.46)	30.04 (4.91)	<0.001
Maternal weight (kg), mean (SD)	62.20 (10.93)	67.13 (13.98)	63.42 (12.70)	62.87 (11.35)	60.88 (11.13)	66.97 (16.14)	<0.001
Maternal education, n (%)							<0.001
Low (primary school)	422 (50.4)	8 (7.5)	344 (49.2)	18 (21.4)	8 (8.2)	8 (8.8)	
Medium (secondary school)	247 (29.5)	29 (27.1)	156 (22.3)	38 (45.2)	41 (41.8)	52 (57.1)	
High (degree or higher)	168 (20.1)	70 (65.4)	199 (28.5)	28 (33.3)	49 (50.0)	31 (34.1)	
Paternal education, n (%)							<0.001
Low (primary school)	368 (44.5)	9 (8.9)	209 (52.8)	22 (26.5)	16 (16.3)	20 (22.2)	
Medium (secondary school)	248 (30.0)	46 (45.5)	88 (22.2)	42 (50.6)	46 (46.9)	53 (58.9)	
High (degree or higher)	211 (25.5)	46 (45.5)	99 (25.0)	19 (22.9)	36 (36.7)	17 (18.9)	

Maternal smoking, n (%)	325 (39.1)	6 (5.4)	150 (21.2)	19 (22.6)	21 (21.4)	18 (19.8)	<0.001

2.1.2 DNA methylation

Cord blood DNA methylation was measured using Illumina HumanMethylation 450K BeadChip arrays for all the cohorts apart GXXI, in which Illumina HumanMethylation 850K BeadChip array was used. The study was performed on a selection of candidate DNA methylation targets identified as we describe below.

Selection of the candidate DNA methylation targets through literature review: methods and results in brief

Selection of the candidate DNA methylation targets was based on a previous systematic review on the association of body size measurements in early life (defined as weight, height, BMI and any change in any of these measurements after birth till 12 years of age) with the epigenome [11].

In addition, as the review coverage was limited to early life period between 0 and 12 years and to articles published until March 2017, we performed a further systematic search in PubMed, using the R package RISmed, aimed at identifying studies on early life DNA methylation changes and body size measurement (defined as weight, height, BMI, any change of these measurements, including categorization into obesity or overweight outcomes) of: (i) adolescents between 12 and 18 years for studies published before March 2017; and (ii) children below 18 years of age for studies published after March 2017 (see Addendum 1). Articles were identified using as search terms the keywords related to child DNA methylation and body sizes combined via boolean operators as reported in Addendum 1. Duplicated, non-original (eg. reviews), and non-English articles were excluded. Two researchers (RA and MP), independently, screened titles and abstracts and then assessed eligibility of full-text articles dealing with body size measurements (age period considered was >0-18 years of age) and DNA methylation in newborns, children or adolescents, and excluded articles in animals or *in vitro*.

Information on body size outcomes along with the associated CpG sites or gene regions, and the direction of effects were extracted from 48 papers and are reported in the Supplementary Table S1. Candidate selection, from papers that reported at least one statistically significant association, resulted in a list of 966 CpGs.

2.1.3 Body size measurements

Measurements of weight and height at birth were obtained from obstetric records while measurements at later time points were self-reported from parents (based on medical records) (for ENVIRONAGE, GXXI, Rhea and Piccolipiù) or measured by trained staff (for ALSPAC, GXXI and INMA cohorts) and were used to generate body size outcomes as described below.

2.1.4 Growth in weight at 6 and 12 months

Using WHO growth reference, each measurement of weight was transformed in standard sex- and age- adjusted deviation (SD) scores. Infancy growth in weight was calculated as the differences between SD-scores of predicted weight at 6 months and actual values of birthweight and dichotomized into rapid (>0.67 SD) or normal <0.67) growth (Table 2) as previously described by[12].

Table 2 Rapid growth outcomes.

	N subjects	Rapid growth at 6 months (%)	Rapid growth at 12 months (%)
ALSPAC	835	214 (26%)	258 (31%)
ENVIRONAGE	108	10 (9%)	32 (30%)
INMA	84	17 (20%)	23 (27%)
Piccolipiù	98	11 (11%)	17 (17%)
RHEA	92	21 (23%)	35 (38%)

Repeated measurement of BMI during infancy and mid-childhood

Repeated measurements of BMI until infancy (2 years of age) and mid-childhood (10 years) were calculated as the ratio of weight on squared height.

Overweight and obesity during infancy and mid-childhood

The repeated measurements of BMI were then categorized into overweight and obese based on BMI WHO SD scores. We used WHO cut off for overweight and obesity of >2SD and >3SD in children under 5 years and >1SD and >2SD in children older than 5 years.

Table 3 Repeated measurements of BMI and corresponding categorization in obese and overweight during infancy (up to 2 years of age).

	N subjects	N obs	Mean N repeated obs	Age range, days	Mean BMI	N (%) Obese obs	N (%) Over Weight obs
ALSPAC	122	427	3.50	105-595	17.64	6 (1%)	34 (8%)
ENVIRONAGE	111	992	8.93	8-728	16.51	0	28 (3%)
INMA	84	310	3.69	4-729	16.13	0	5 (2%)
Piccolipiù	88	603	6.85	12-728	16.22	1 (1%)	18 (3%)
RHEA	92	1227	13.33	3-728	16.16	6 (1%)	40 (3%)

Obs= observations

Table 4 Repeated measurements of BMI and corresponding categorization in obese and overweight during childhood (up to 10 years).

	N subjects	N obs	Mean N repeated obs	Age range, days	Mean BMI	N (%) Obese obs	N (%) Over Weight obs
ALSPAC	854	1940	2.27	105-3318	16.68	79 (4%)	287 (15%)
ENVIRONAGE	111	1025	9.23	8-1505	16.50	0	30 (3%)
GXXI	706	1322	1.87	1430-2809	16.61	116 (9%)	302 (23%)
INMA	84	661	7.86	4-2759	16.24	23 (3%)	58 (9%)
Piccolipiù	97	741	7.63	12-1732	16.20	4 (1%)	23 (3%)
RHEA	92	1731	18.81	3-2781	16.19	30 (2%)	134 (8%)

Obs= observations

2.1.5 Statistical analysis

First, we performed a separate analysis for each cohort under study (GXXI, Rhea, INMA, Piccolipiù, ENVIRONAGE, ALSPAC: Figure 1), using generalized linear mixed models. Analyses of different outcomes were performed according to linear or logistic models including different sets of confounders and at different ages. Outcomes were BMI and height (repeated measurements between ages 0-2 and 0-10), obesity and overweight (repeated measurements between ages 0-2 and 0-10), and rapid growth (0-6 months and 0-12 months). Confounders were age in days and gender of the child and models included fixed effects on child (for repeated measurements), bead array row and bisulfite conversion batch (technical covariates). The results from single cohorts were then combined using a meta-analytical approach. CpGs were considered significant if 5% False Discovery Rate (FDR) adjusted p-values were below 0.05 threshold, using Benjamini-Hochberg procedure.

We focus in particular on continuous weight growth measures because these were the most frequently available outcomes (see Table 2 above). During childhood there were only 13 obese children in all the cohorts, too few to allow any meaningful analysis (see Table 3 above).

2.2 Results and discussion

No FDR-significant CpG was related to BMI, obesity, overweight or height in any of the models tested during infancy or childhood. Full results are presented in **Addendum 2.**

Seven and 12 FDR-significant CpGs were related to rapid growth at 6 and 12 months, respectively. These included CpG sites located on the genes *NFIX, AQP9, HIPK3, LY86, HIPK, MCF2L2, DNMT3B, IL16* and *PCK2* (Figure 2).

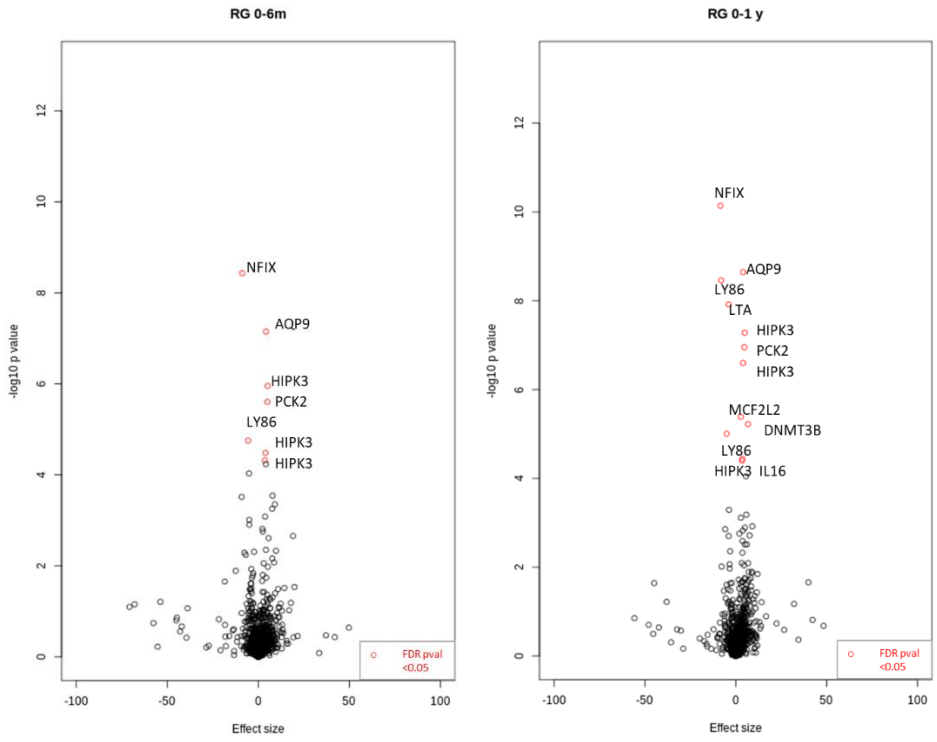


Figure 2 Volcano plots of the associations between candidate CpGs and 0-6 months (on the left) and 0-12 months (on the right) rapid growth.

Cpg site Cg10264529, located on the promoter region of the *PCK2* gene, was in addition associated with (childhood and infancy) height and (childhood) obesity (nominal p-value<0.05) (Figure 3).

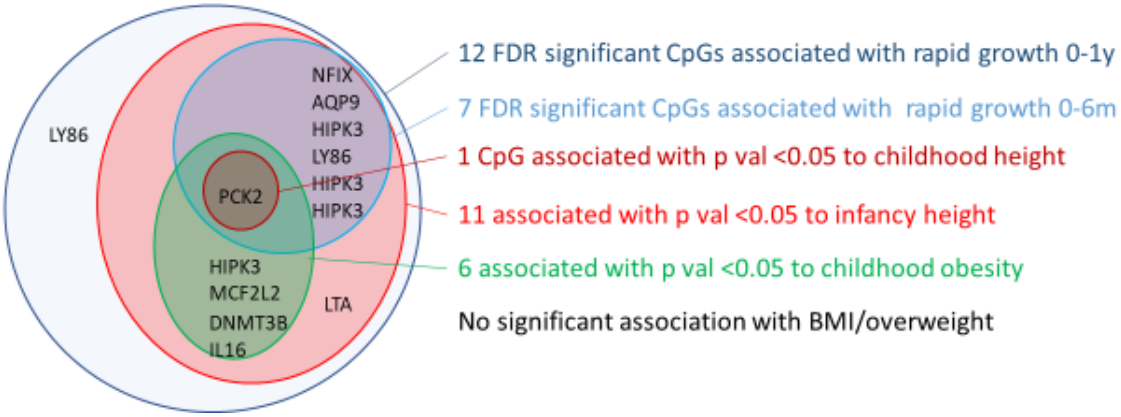


Figure 3 Venn diagram showing overlap of Cpg associations with different outcomes.

PCK2 is particularly relevant both because the consistency of the findings across growth outcomes and because of the biological meaning: phosphoenolpyruvate carboxykinase is an enzyme found in mitochondria that is involved in glucose metabolism [13]. The methylation of ***NFIX*** (cg15783941) has been previously found in ALSPAC as being associated with birthweight (negative estimate) and fat mass at 17yrs (positive estimate) [14]

2.3 Mediation analysis

Since in a previous study *PCK2* was found to mediate the association between childhood deprivation and adverse experiences [15], we performed mediation analysis to investigate the hypothesis that the *PCK2* CpG cg10264529 may mediate the association between paternal education and rapid growth between 0 and 12 months.

We used causal mediation modeling as implemented by the *lanvaan* R package to estimate the direct and indirect (via the methylation) effect paternal education and rapid growth between 0 and 12 months.

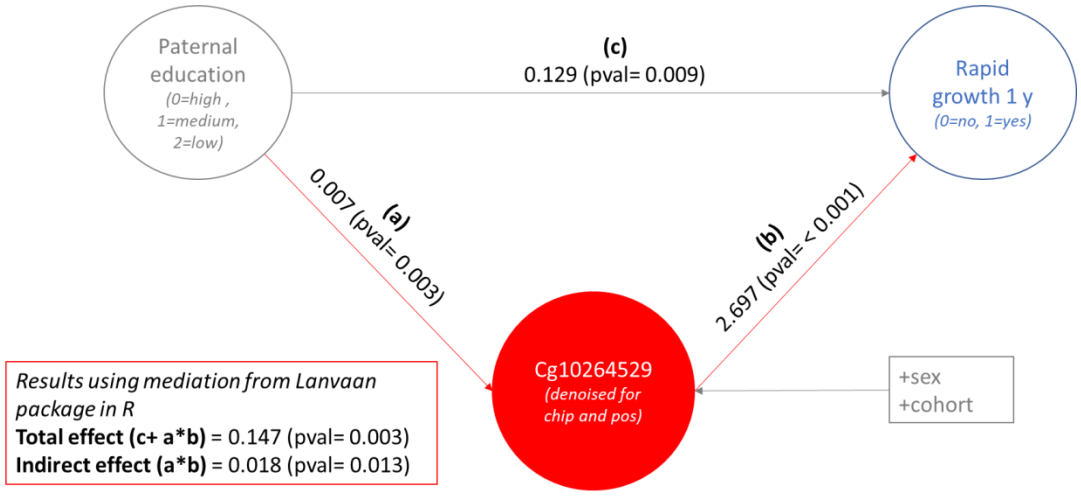


Figure 4 Graphical representation of the mediation analysis

We found that paternal education was significantly associated with *PCK2* DNA methylation (cg10264529) (p-value=0.003) and a stronger association between the latter and rapid growth at 1 year of age (r=2.7, p-value<0.001). We found that the association between paternal education and rapid growth at 1 year of age was mediated by methylation of cg10264529 (r=0.018, p-value=0.013).

In conclusion, we have identified *PCK2* and *NFIX* as the most replicated and strongest associations, both associated with childhood obesity in previous investigations. Further work will continue investigating the role of the epigenome in causal pathways mediating risk factors of childhood obesity.

3 Metabolomics and proteomics

3.1 Background

3.1.1 STOP Review of metabolomics and child obesity

Elucidating the mechanisms of development of childhood obesity at a molecular level may contribute to identifying potential targeted intervention approaches in childhood. “Metabolomics”, the study of small molecules (<1500 kDa) in a biological sample, can improve understanding of biological responses and alterations due to changes at the genetic, epigenetic or protein level, but also due to environmental exposure such as diet, physical activity, microbiome and toxins. Thus, metabolomics can help to define molecular phenotypes and may elucidate mechanisms for obesity and associated conditions such as diabetes. While the literature regarding application of metabolomics to obesity in adults is relatively mature, fewer studies have been conducted specifically in child populations [16]. Metabolic signatures of obesity in children may differ from a signature observed in adults for reasons including a relatively shorter duration of obesity, ongoing linear growth, and pubertal hormones. We therefore conducted a review of the literature related to rapid growth (an early risk factor of obesity risk) or overweight or obesity and metabolomics in children (Chandakas et al: “A systematic review of metabolomics and childhood obesity”).

Rzehak et al.[17] applied LC-MS (liquid-chromatography- mass spectrometry) to profile 168 metabolites using the Biocrates kit to analyse plasma blood samples collected at age 6 months from 726 infants participating a European multi-centre randomized trial (Childhood Obesity Programme, CHOP), randomized to a high- or low-protein content formula and breast-fed infants. Anthropometric data up to the age of 6 years were available. The amino acids tyrosine and citrulline, the diacyl-phosphatidylcholine PCaaC34:4, and the lyso-phosphatidylcholine LPCaC14:0 were associated with rapid growth during the first 6 months of life. 15 other metabolites were associated with less-than-ideal growth (“slow growth”). However, metabolic signatures were significantly different by feeding group and after adjusting for feeding group, only LPCaC14:0 remained significantly associated with rapid weight gain. Intriguingly, only LPCaC14:0 at age 6 months was predictive of subsequent overweight/obesity at age 6 years, suggesting a metabolically programmed effect of infant weight gain on the later obesity risk.

Wahl et al.[18] kit also applied Biocrates LC-MS in serum of 80 obese and 40 normal weight German children between 6 and 15 years of age. Concentrations of two acylcarnitines (C12:1 and C16:1) were significantly increased in obese compared to normal-weight children, while concentrations of glutamine, methionine, proline and nine phospholipids (PC ae C34:1, C34:2, C34:3, C36:2, C36:3 and C38:2, LPC a C18:1, C18 2 and C20:4) were significantly decreased in obese children. The identified metabolite markers are indicative of oxidative stress and of changes in sphingomyelin metabolism, in β-oxidation, and in pathways associated with energy expenditure. Cho et al.[19] analysed urine to compare non-obese ($n = 91$) and obese ($n = 93$) adolescent Koreans, using both untargeted MS analysis and targeted Biocrates based analysis. Untargeted metabolomics identified lowered levels among children who were obese of four metabolites that function in the detoxification of xenobiotics by the microbiome (4-hydroxybenzaldehyde, hippuric acid, 4-sulfobenzyl alcohol and N,N-dimethyl-safingol). Other metabolites associated with inflammation such as docosanoic acid were elevated in the obese group. They found that 45 metabolites were differentially expressed ($P < 0.05$ uncorrected for multiple testing) in urine in the targeted analysis (many overlapping with other Biocrates based

studies), including nine acylcarnitines (including propionylcarnitine, and butyrylcarnitine), 11 amino acids, seven biogenic amines (asymmetric dimethylarginine, carnosine, 3,4-dihydroxyphenylalanine, dopamine, putrescine, serotonin and total dimethylarginine) and 14 glycerophospholipids and four sphingolipids.

Perng et al.[20] applied a combination of LC and GC-MS methods provided by the Metabolon company to assess 345 compounds in 84 obese (35.9%), 28 overweight (20.7%), and 150 normal-weight (57.3%) American children of median age 7.7 years. They applied a data reduction approach to reduce the metabolomic data to 18 principal components or metabolic patterns. A branched-chain amino acid (BCAA)-related pattern and an androgen hormone pattern were higher in obese vs. lean children. Both patterns were associated with adiposity and worse cardiometabolic profiles, each increment in the BCAA and androgen pattern scores corresponded with 6% higher HOMA-IR. Children of obese mothers had higher BCAA score than their counterparts, although this association was somewhat attenuated upon adjustment for child BMI. Butte et al.[21] assessed 304 metabolites also through the Metabolon profiling service in 803 Hispanic children (mean ± SD age: 11.1 ± 3.9 y) of whom 56% (450) were obese. BCAAs (Leu, Ile, and Val) and their catabolites, propionylcarnitine (C3) and butyrylcarnitine (C4), were elevated in obese children, while lysolipids and dicarboxylated fatty acids were lower. Steroid derivatives were markedly higher in obese children as were markers of inflammation and oxidative stress. Tyrosine was the highest-ranked metabolite based on its contribution to the obesity classification. In principal component analysis, the BCAAs/aromatic AAs component and another AA component (asparagine, glycine, and serine) made the largest contributions to BMI, and two acylcarnitine components made the largest contributions to adiposity.

Lau et al [22] assessed metabolome associations between standardised BMI z-score in 1192 children (aged 6–11 years) recruited from birth cohorts in six European countries using an untargeted NMR analysis in urine and targeted LC-MS (Biocrates) in serum. They reported positive associations between urinary 4-deoxyerythronic acid and the BCAA valine and negative associations with urinary p-cresol sulphate (a microbial metabolite) and pantothenate (vitamin B5). Amongst serum metabolites, positive associations with BMI z-score included free carnitine, short-chain acylcarnitines (C3, C5), seven amino acids including glutamate and BCAAs, sphingolipids (SM C16:0, SM C16:1, SM C18:1), multiple phosphatidylcholine species and four lyso-phosphatidylcholines (lysoPC a C14:0, lysoPC a C16:1, lysoPC a C18:1, lysoPC a C18:2). Hellmuth et al [23] similarly examined BMI z-score in meta-analysis of 1,020 children aged 8 to 10 years in five European countries. They employed multiple MS based targeted analyses for amino acids, non-esterified fatty acids, carboxylic acids, acyl-carnitines and phospholipids, with 108 metabolites available in all studies. SM 32:2 was the metabolite that showed the strongest association with BMI z-score followed by tyrosine, valine, PCaa 34:4, and PCaa 38:3. Fewer metabolites were negatively associated with BMI z-score. Only three metabolites, namely PCae 36:2, PCae 36:1, and NEFA 12:0 were significantly and negatively associated with BMI z-score at the Bonferroni corrected significance threshold. The strength of the association with the SM 32:2 was striking, particularly as this is a sphingomyelin not assayed by the other platforms included in this review.

Two studies examined changes in metabolites measured before and after a weight loss intervention program. Leal-Witt et al [24] applied untargeted LC-MS profiling to plasma samples collected at baseline and after a 6-month-long lifestyle intervention program from 35 children (7–10 years) with obesity (body mass index (BMI)>2 standard deviations) in Spain. The

intervention significantly reduced BMI and improved HbA1c (%), total cholesterol levels and adherence to a healthy diet. PCA identified one component (PC1) significantly altered by the intervention (Bonferroni adjusted P=0.008). A sphingolipid metabolism-related signature was identified as the major contributor to PC1. Sphingolipid metabolites were decreased by the intervention, and included multiple sphingomyelin, ceramide, glycosylsphingosine and sulfatide species. However, changes of individual metabolites such as sphingomyelin 23:0 were associated only with changes in HbA1c (%) and not BMI. Reinehr et al.[25] analysed 14 metabolites identified previously as associated with obesity by Wahl et al. [18] in serum samples of 80 obese children with substantial weight loss (BMI-SDS reduction >0.5) and in 80 obese children with stable weight status, all participating in a 1-year lifestyle intervention. In children with substantial weight loss, glutamine, methionine, the lyso-phosphatidylcholines LPCaC18:1, LPCaC18:2, and LPCa20:4, as well as the acyl–alkyl phosphatidylcholine PCaeC36:2 increased significantly, while the acylcarnitines C12:1 and C16:1, proline, PCaeC34:1, PCaeC34:2, PCaeC34:3, PCaeC36:3, and PCaeC38:2 did not change significantly. Since the six significantly increasing metabolites were previously found to be lower in obese children compared with normal-weight children, these findings suggest that the changes of these metabolites in obesity are a consequence of changing diet or exercise habits due to lifestyle intervention.

In summary, these mainly cross-sectional studies have reported consistent metabolic signatures of childhood obesity, despite differences in analytical platform and biological matrix. The obesity-related metabolic signature in children include alterations to levels of amino acids (particularly BCAAs and tyrosine), acylcarnitines (particularly those of shorter chain length), sphingomyelins, phosphatidylcholines and lyso-phosphatidylcholines, steroidal hormones, inflammation related metabolites and gut microbial related metabolites. These signatures are largely concordant with those in adult studies; however, increased steroid metabolites appear unique to specific life stages such as childhood.

Only two studies have cord blood to predict child weight status in later life. Isganatis et al. [26] compared the cord blood metabolomic profiles, measured by LC-MS (Metabolon, 415 metabolites), of cases (n=26) based on top quartile of change in weight-for-age 0-6 mo and BMI >85th percentile in mid-childhood (median 7.7 years) to age and sex matched controls (n=26) in an American cohort. Tryptophan metabolites serotonin, tryptophan betaine, and tryptophyl leucine were 46%, 48% and 26% lower in cases respectively, as were two methyl donors, dimethylglycine and N-acetylmethionine (18% and 16% lower respectively). While nominally significant, these changes did not pass FDR correction. Pathway analysis identified enrichment in “Tryptophan Metabolism” and “Excitatory neural signaling through 5HTR4/6/7 and serotonin” pathways. Hellmuth et al [27] applied a range of targeted LC-MS assays (209 metabolites) to profile cord blood of 700 German children to predict rapid weight gain, and BMI at 2 and 7 years. While many metabolites were associated with weight at birth, no associations with post-natal measures survived multiple testing correction. Cord blood metabolites that were associated with increased birth weight showed a tendency to be associated with lower postnatal weight gain z-scores and zBMI at ages 2 and 15 years, while the converse was observed for metabolites with birth weight-lowering effects. Interestingly, associations of cord blood metabolites with later zBMI showed similar effect sizes as with birth weight, although significance level were clearly stronger with birth weight.

3.1.2 Objectives of current study

To investigate the association of cord blood metabolites, measured through untargeted UPLC-MS metabolomics, and inflammatory proteins, measured through multiplex assays, with growth outcomes in childhood including rapid infant weight gain in the first 12 months, childhood overweight and obesity as well as height and BMI from birth to the age of 2 and 7 years.

3.2 Methods

3.2.1 Study population

The study population included participants from four different cohorts of the EXPOsOMICs project (ENVIRONAGE, Rhea, INMA and PICCOLI+). The population consists of white-European background children. The study population demographic characteristics are presented in Table 5 Cohort characteristics. A break-down of children with overweight and obesity at different ages is given in supplementary tables s1 and s2, and height, weight and BMI measures over time is give in supplementary figure s1 (**Addendum 3)**

Table 5 Cohort characteristics

	ENVIRONAGE	INMA	PICCOLI+	Rhea
Start recruitment	2012	2004-2006	2011 - 2015	2007-2008
Country	Dutch	Spain	Italy	Greece
Growth trajectories until	from 0 to 2 years	Repeat measures at 0, 6m, 14 m, 4 yr, 7 yr	Repeat measures up to 5 yr	0, 1m, 18 m, 4 yr, 7 yr
Available samples for the presented study	107	93	99	100
Available observations	≈1850	≈1700	≈1550	≈3000
Cohort reference	28	29	30	31

3.2.2 Outcome assessment

The selected outcomes were rapid growth at 6 months, z-BMI, z-weight and z-height at 2, 5, 6,7 years of age and repeated BMI, weight and height until the age of 7 years. The z-scores of anthropometric indicators were calculated using WHO child growth standards for sex and age [32, 33] in order to allow a comparison with other studies. The classification for healthy, overweight and obese (over and under 5 years of age) was done based on WHO adjusted z-scores. In particular, for the ages under 5 years, overweight and obesity cut-offs for the BMI z score were set at +1 SD and +2 SD [34] and for ages over 5 at +2 SD and +3 SD respectively [35].

Rapid growth and normal-growth infants at 6 months were categorized based on the definition of Ong et al. [36]. According to this definition, a clinically significant increment that indicates rapid growth occurs when between different target ages an increment of at least 0.67 standard deviations occurs (SDS). Slow growth, in turn, is defined as a decrease in SDS< -0.67 and normal growth as the increment or decrement between −0.67 and +0.67 SD. In this study we examined the rapid growth of the infants at 6 months. Because of the lack of height measurements at birth, we were unable to calculate the participants z-BMI score at birth, since rapid growth was defined as the z score change of greater than 0.67 SD between birth and 6 months of age. While repeated anthropometric measurements were available for all cohorts,

data at exactly 6 months of age were lacking. For this reason, a prediction model based on the repeated measurements was used for calculating weight at exactly 6 months (termed “predicted data” below, as opposed to “real data”).

3.2.3 UPLC-MS metabolomics

UPLC-MS spectra were acquired in a single batch from cord blood using 1290 Binary LC system, a Jet Stream electrospray ionization (ESI) source, and a 6550 QTOF mass spectrometer (Agilent Technologies) at IARC, France. Metabolic features present in <60% of the samples were removed, data were log-transformed, and missing values were inputed leaving 4714 features for analysis.

3.2.4 Inflammatory proteins

In total a set of 16 inflammatory proteins were analysed at the Institute for Risk Assessment Sciences in Utrecht, in The Netherlands. These included: interleukin 1 receptor antagonist (IL-1ra), interleukin (IL)-6, IL-8, IL-17, interferon gamma-induced protein 10 (CXCL10), epidermal growth factor 2 (EGF2), granulocyte colony-stimulating factor (GCSF), melanoma growth stimulatory activity/growth-related oncogene alpha (GRO-α), macrophage-derived chemokine (CCL22), macrophage inflammatory protein-1 beta (MIP-1β), monocyte chemotactic protein-1 (MCP-1), myeloperoxidase (MPO), periostin, tumor necrosis factor alpha (TNF-α), and vascular endothelial growth factor (VEGF) that were assessed using an R & D Systems Luminex screening assay according to the protocol described by the manufacturer, and C reactive protein (CRP) that was assessed using an R & D Systems Solid Phase Sandwich ELISA. Inflammatory proteins that had at least 40% detectable measurements were imputed based on a maximum likelihood estimation procedure [37]. 490 samples had protein levels available for our analyses.

3.2.5 Statistical Analysis

A univariate approach combined with Linear mixed-effect regression models was applied to examine the association of metabolomic and proteomic cord blood signatures with longitudinal changes in outcomes. The model consisted of participant-level (repeated measures within person) and cohort-level (participants clustered within 4 cohorts) random intercepts with an unstructured covariance matrix, as well as time (in days) and gender as continuous and binary variables, respectively. In proteomics statistical analyses, an additional variable was added to incorporate the random effect of the plate.

Multiple testing was applied using the Benjamini-Hochberg method [38], considering the following sets for analysis: 1) whole metabolome (over 4000 untargeted features), 2) metabolome features previously annotated identified as associated with birthweight [39] (Alfano et al *under review*) (95 metabolites, see table s3); 3) 16 inflammatory proteins. False discovery rate adjusted p-values were considered as statistically significant at a level of 5% (Q-value<0.05).

To check whether any loss to follow-up has been differential with respect to relevant demographic and other characteristics, we conducted sensitivity analyses by limiting models to those with complete follow-up data, to assess robustness of our key findings. Linear mixed-effect regression models were applied to examine the effect of metabolomic and proteomic cord blood signatures at the exact ages of 2, 5, 6 and 7 years. These models consisted of cohort-level (participants clustered within 4 cohorts) random intercepts with an unstructured covariance matrix, and time (in days) and gender as continuous and binary variables respectively. As follow-up measurement at the exact ages we defined the closest observation to the given age with a

tolerance of 30, 60, 90 days for the classes of 0.5, 1 and 1.5 years respectively and +/-180 days for the rest.

A leave-one-out cross-validation approach was applied to assess the generalization of results. This step provides the opportunity of exploring model uncertainties due to limited sample size and revealing instability in model fitting.

We used Manhattan plots to illustrate the results after applying the FDR correction. Manhattan plots express the −log10 p-value multiplied by the sign of the corresponding regression coefficient.

All statistical analyses were performed using R ('The R Project for Statistical Computing') software environment (v3.5.2). lme4 package was used for linear mixed-effect regression models [40]. The ggplot package was used for plotting the results [41].

3.2.6 Metabolic pathway enrichment analysis

We performed Pathway identification for metabolic features nominally significant ($p < 0.05$). Mummichog (v. 1.0.5) was used for the identification. Mummichog is a state of art bioinformatics Python-based platform that infers and categorizes functional biological activity using directly the output from mass spectrometry [42]. The algorithm searches tentative compound lists from metabolite reference databases against an integrated model of human metabolism to identify functional activity. Fisher's exact tests are used to infer p-values, which are adjusted for type I error through a pathway permutation procedure (43). Likelihood of pathway enrichment across significant features is compared to pathways identified across the entire compound set in a reference list (the entire metabolome dataset), considering the probability of mapping the significant m/z features to pathways. Mummichog parameters were set to match against ions included in the 'positive mode' setting at ± 5 ppm mass tolerance. Visualisation of enriched pathways on the KEGGscape network was performed through the MetaboAnalyst platform (Chong et al., 2018)

3.2.7 General workflow

Firstly, the data was collected, harmonized and merged in one database. Then, we checked for completeness, consistency, validity, integrity, uniformity and linkage [43] applying a wide range of queries and tests. After filtering and cleaning the data based on the above-mentioned criteria, the initial number of participants was reduced from 403 to 399 and the number of total observations was more than 4500 (i.e. including repeat observations). The total predicted observations at the age of 6, 12, 18, 24, 36, 48, 60, 72 and 84 were 3420.

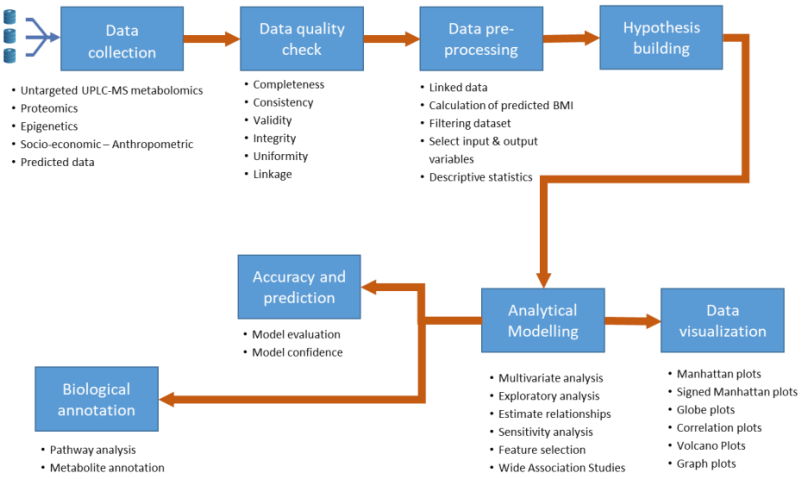


Figure 5 Analysis workflow

3.3 Results

3.3.1 Untargeted Metabolomics

In the untargeted MWAS of cold blood (4712 features), multiple features were associated with BMI, height, overweight, obesity and rapid growth at a nominal p-value of 5%. The proportion of the statistically significant features ranged from 2% to 8% (Table 6). More associations were observed with rapid growth, than other outcomes. At 2 years most associations were observed with obesity and at 7 years most associations were observed with overweight. Correcting for 5% false-discovery rate in all the models greatly reduced the total number of associations, with 12 significant features associated with rapid growth (Figure 7) and only one significant feature associated with obesity at 7 yrs and overweight at 2 and 7 yrs (Table 6). Figure 7 and Figure 7 show metabolome wide associations with rapid growth and overweight up to 7 years respectively.

Table 6 Number of associated features for the mixed effect models for repeated measurements up 2 and 7 years by outcome (BMI, height, obese and overweight) as well as rapid growth at 1 year old.

	12 months		up to 2 yr		up to 7 yr	
Model outcome	number of statistically significant features (total)	number of statistically significant features after applying FDR adjustment (total)	number of statistically significant features (total)	number of statistically significant features after applying FDR adjustment (total)	number of statistically significant features (total)	number of statistically significant features after applying FDR adjustment (total)
BMI	-	-	212 (4712)	0 (4712)	312 (4712)	0 (4712)
height	-	-	79 (4712)	0 (4712)	265 (4712)	0 (4712)
obese	-	-	353 (4712)	1 (4712)	232 (4712)	0 (4712)
overweight	-	-	157 (4712)	1 (4712)	376 (4712)	1 (4712)
rapid growth	380 (4712)	12 (4712)	-	-	-	-

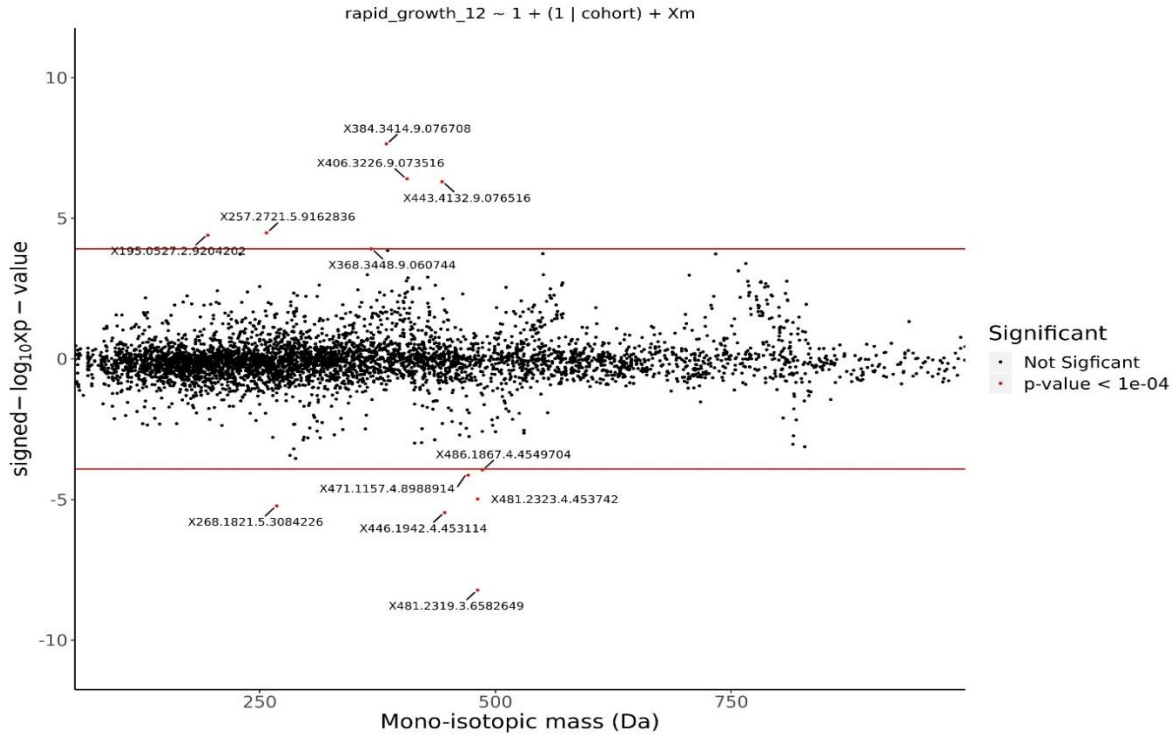


Figure 6 Signed Manhattan plot showing metabolome wide associations with rapid growth up to 1 year. Red line show FDR corrected significance level. Sign of p value shows direction of effect.

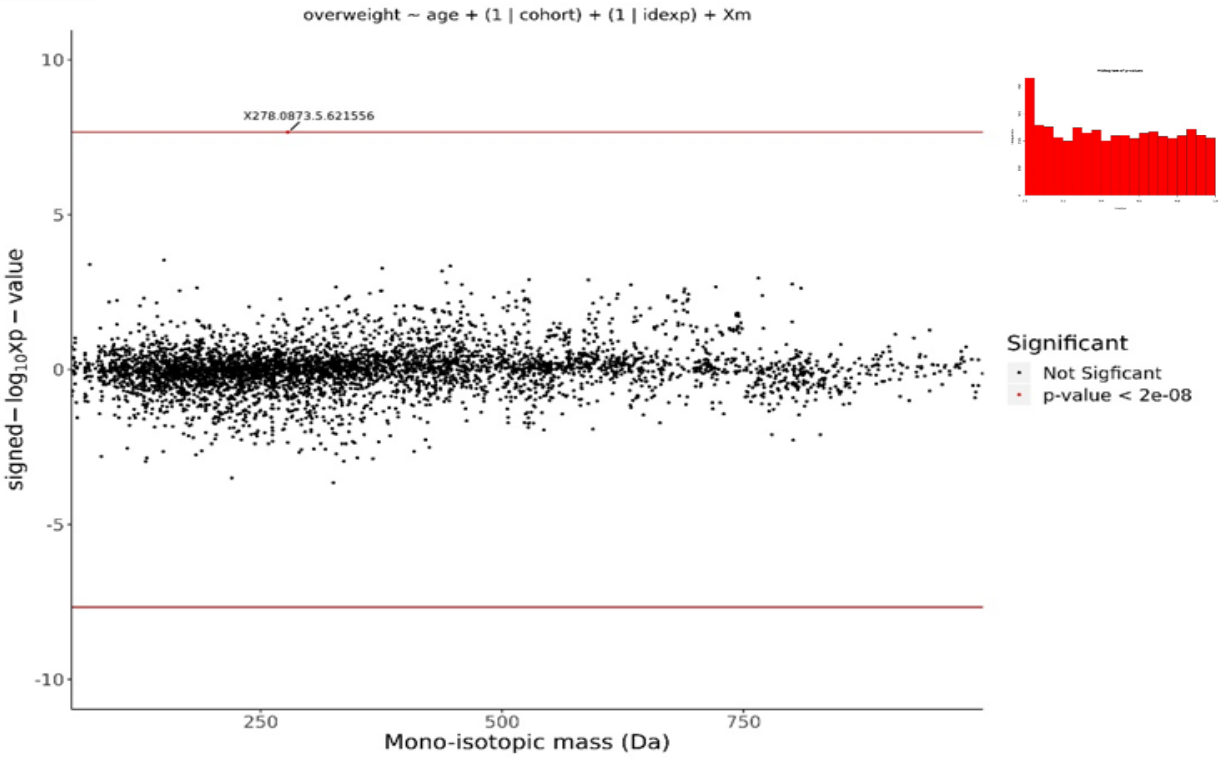


Figure 7 Signed Manhattan plot showing metabolome wide associations with repeated measurements of overweight up to 7 years. Red line show FDR corrected significance level. Sign of p value shows direction of effect. Inset shows distribution of p-values with greater frequency of lower p-values.

The top 20 most significant features with each outcome are shown in supplementary tables 4-8 (**Addendum 3**). There was little overlap in top 20 most significant features with rapid growth up to 1 year and growth outcomes up to 7 years (Figure 8). 19 of the top 20 features associated with height up to 7 years were not associated with the adiposity related outcomes up to 7 years. One feature (m/z 278.0873, retention time 5.6216), was associated with both overweight and obesity and was also significant after FDR adjustment. The closest annotation match for this feature, through exact mass search of the METLIN database, is **triphenylphosphine oxide (TPPO**), a food contaminant that has known function in the regulation of energy intake in animal models [30]. However, full laboratory confirmation of hits in the untargeted analysis is currently ongoing.

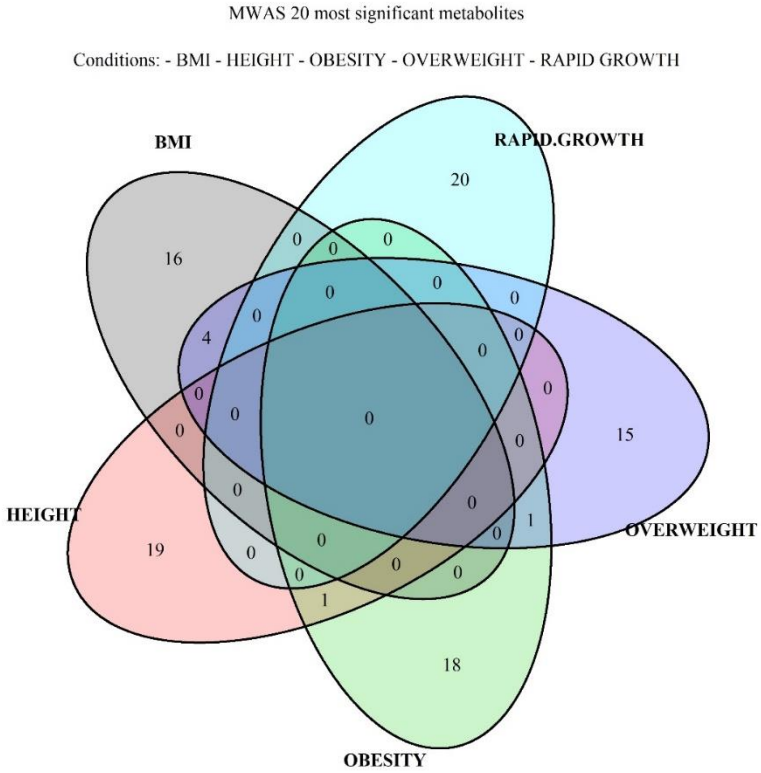


Figure 8. Venn diagram showing overlap of top 20 most significant associations for rapid growth up to 1 year and other outcomes up to 7 years.

Pathway enrichment analysis, using the Mummichog algorithm with overweight up to 7 yrs identified enrichment in three metabolic pathways (Table 7, Figure 9): Fatty acid activation, Fatty Acid Metabolism and Linoleate metabolism.

Table 7 Top 10 enriched metabolic pathways among predictors present of overweight up to 7 years.

Pathways	overlap size[a]	pathway size[b]	p-value[c]
Fatty acid activation	8	15	0.040
Fatty Acid Metabolism	6	12	0.043
Linoleate metabolism	7	19	0.048
Tryptophan metabolism	13	50	0.058
Vitamin B6 (pyridoxine) metabolism	3	6	0.060
Prostaglandin formation from dihomo gama-linoleic acid	3	6	0.060
De novo fatty acid biosynthesis	5	15	0.061
Alkaloid biosynthesis II	3	7	0.066
Propanoate metabolism	3	7	0.066

[a] *The number of model predictors matched to each pathway*

[b] *The number of metabolites in the whole metabolome dataset matched to each pathway*

[c] *P-values adjusted for type 1 error through Gamma-based permutation procedure*

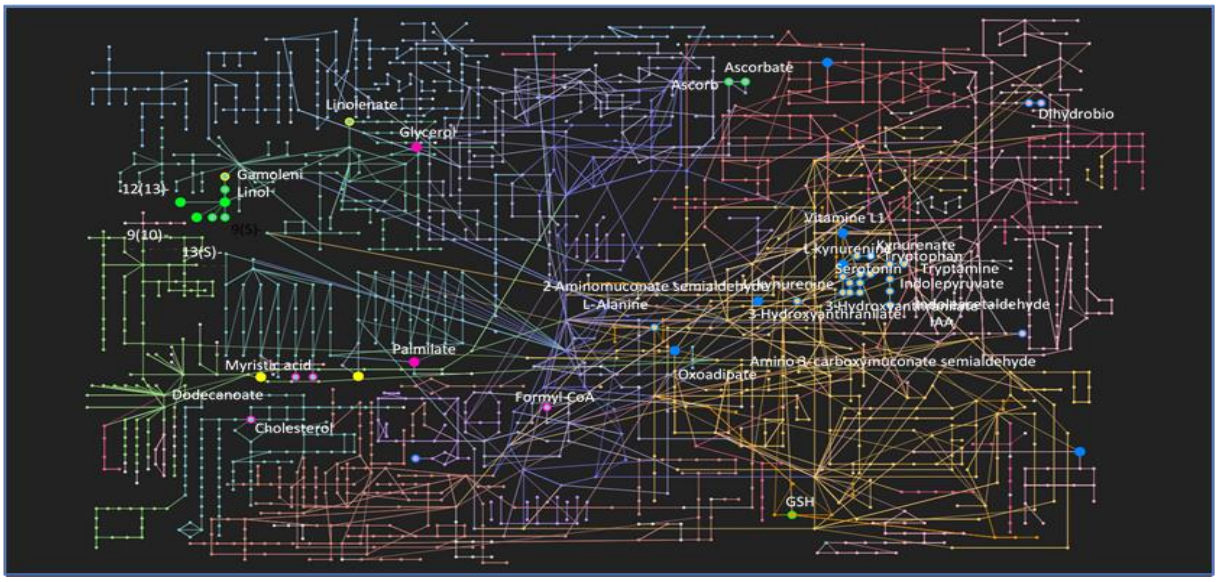


Figure 9 Metabolic network visualization of significantly enriched pathways based on the manually curated KEGG global metabolic network [44]. The metabolites of significantly enriched pathways are represented as nodes on the network. Empty nodes represent compounds identified from the feature list by Mummichog but not significant, while solid nodes represent significantly enriched features. Note not all metabolites from the KEGG global network are displayed. Fatty acid activation pathway nodes are coloured green, Fatty Acid Metabolism nodes are coloured pink, Linoleate metabolism nodes are coloured yellow, Tryptophan metabolism pathway nodes are coloured blue.

3.3.2 Birthweight-related metabolites in cord blood metabolomics analysis

In this analysis we analysed 95 metabolites related to birthweight and pre-annotated based on previous work (supplementary s3 for list). Multiple features were associated with BMI, height, overweight, obesity and rapid growth at a nominal p value of 5%. The proportion of the statistically significant features ranged from 2% to 14% (Table 8). More associations were observed with rapid growth, than other outcomes. At 2 years most associations were observed with obesity and at 7 years most associations were observed with overweight. Correcting for 5% false-discovery rate in all the models greatly reduced the total number of associations, with 19 significant features associated with rapid growth (Figure 6), only two significant features associated with obesity at 2 yrs and one feature associated with height up to 7 years.

Table 8 Number of Associated features for the mixed effect models for repeated measurements up 2 and 7 years by outcome (BMI, height, obese, overweight) as well as rapid growth at 1 year old.

	12 months		up to 2 yr		up to 7 yr	
Model outcome	number of statistically significant features (total)	number of statistically significant features after applying FDR adjustment (total)	number of statistically significant features (total)	number of statistically significant features after applying FDR adjustment (total)	number of statistically significant features (total)	number of statistically significant features after applying FDR adjustment (total)
BMI	-	-	12 (95)	0 (95)	9 (95)	0 (95)
height	-	-	4 (95)	0 (95)	8 (95)	1 (95)
obese	-	-	10 (95)	2 (95)	2 (95)	0 (95)
overweight	-	-	13 (95)	0 (95)	10 (95)	0 (95)
rapid growth	28 (95)	19 (95)	-	-	-	-

The top 20 most significant features with each outcome are shown in **Error! Reference source not found.** to Table *13*. Progesterone and Trans-2-Dodecenoylcarnitine(C12:1) were associated after FDR-correction with obesity up to 2 years and Tetradecanoylcarnitine(C14:0) was associated after FDR-correction with height up to 7 years. The omega-3 fatty acid docosahexaenoic acid was the most strongly associated feature (p = 0.0013) with overweight up to 7 years, almost reaching FDR level significance. FDR adjusted features associated with rapid growth up to 12 years included cholestenone, cholesterol, PC(C32:0), Plasmalogen (PC(36:4)), PC(C34:2), PC(36:4), PC(30:0), Plasmalogen (PC(36:3)), PC(C36:4), Indolelactic acid, Decenoylcarnitine_2(C10:1), Plasmalogen (PC(38:4)) and Butyrylcarnitine/Isobutyryl-L-carnitine(C4:0).

Table 9 Top 20 strongest associations with BMI using repeated measurements until the age of 2 (left) and 7 (right) years, among birthweight-associated metabolites.

id	Metabolite	Coefficient	p-value	q-value
1	Tetradecenoylcarnitine(C14:1)	-4.58E-01	0.001	0.0719
2	Decanoylcarnitine(C10:0)	-4.44E-01	0.0018	0.0719
3	3-Hydroxy-cis-5-tetradecenoylcarnitine(C14:1)	-3.28E-01	0.0052	0.0947
4	Dodecanoylcarnitine(C12:0)	-4.10E-01	0.0057	0.0947
5	L-Octanoylcarnitine(C8:0)	-3.79E-01	0.0064	0.0947
6	35-Tetradecadiencarnitine(C14:2)	-4.30E-01	0.007	0.0947
7	Trans-2-Dodecenoylcarnitine(C12:1)	-3.77E-01	0.0099	0.1144
8	Unassigned_7	-1.85E-01	0.0146	0.1475
9	Docosahexaenoicacid	-3.78E-01	0.0309	0.2785
10	Unassigned_14	-2.98E-01	0.0363	0.2939
11	Tetradecanoylcarnitine(C14:0)	-3.15E-01	0.044	0.3132
12	Decenoylcarnitine_2(C10:1)	-3.72E-01	0.0464	0.3132
13	Hexadecenoylcarnitine(C16:1)	-3.07E-01	0.0519	0.3234
14	Unassigned_57	-2.51E-01	0.0693	0.401
15	Unassigned_33	-4.19E-01	0.0917	0.4394
16	2-Hexenoylcarnitine(C6:1)	-1.95E-01	0.0955	0.4394
17	LysoPC(C18:3)	3.21E-01	0.0973	0.4394
18	912-Hexadecadienoylcarnitine(C14:1)	-2.66E-01	0.1014	0.4394
19	Unassigned_31	-3.08E-01	0.1032	0.4394
20	Unassigned_36	-2.46E-01	0.1108	0.4394

id	Metabolite	Coefficient	p-value	q-value
1	Unassigned_7	-1.93E-01	0.0066	0.3689
2	Unassigned_31	-4.29E-01	0.014	0.3689
3	Docosahexaenoicacid	-3.84E-01	0.0166	0.3689
4	Tetradecenoylcarnitine(C14:1)	-3.03E-01	0.0216	0.3689
5	Decanoylcarnitine(C10:0)	-2.96E-01	0.0285	0.3689
6	LysoPC(C18:3)	3.73E-01	0.0379	0.3689
7	Leucine	-6.01E-01	0.0399	0.3689
8	35-Tetradecadiencarnitine(C14:2)	-3.09E-01	0.0406	0.3689
9	Unassigned_58	-1.97E-01	0.041	0.3689
10	Trans-2-Dodecenoylcarnitine(C12:1)	-2.68E-01	0.0529	0.3854
11	Unassigned_2	-2.33E-01	0.0567	0.3854
12	L-Octanoylcarnitine(C8:0)	-2.51E-01	0.0571	0.3854
13	Unassigned_14	-2.45E-01	0.0664	0.4134
14	Dodecanoylcarnitine(C12:0)	-2.50E-01	0.0756	0.4374
15	Unassigned_49	2.76E-01	0.0884	0.4771
16	Unassigned_43	2.64E-01	0.1045	0.4983
17	Unassigned_33	-3.80E-01	0.1053	0.4983
18	LysoPC(C16:1)	3.21E-01	0.1107	0.4983
19	912-Hexadecadienoylcarnitine(C14:1)	-2.37E-01	0.1237	0.5012
20	Hexadecenoylcarnitine(C16:1)	-2.26E-01	0.1295	0.5012

Table 10 Top 20 strongest associations with overweight using repeated measurements until the age of 2 (left) and 7 (right) years, among birthweight-associated metabolites.

id	Metabolite	Coefficient	p-value	q-value
1	Trans-2-Dodecenoylcarnitine(C12:1)	-5.55E-02	0.0006	0.0515
2	3-Hydroxy-cis-5-tetradecenoylcarnitine(C14:1)	-4.11E-02	0.0016	0.0638
3	Tetradecenoylcarnitine(C14:1)	-4.36E-02	0.0047	0.0968
4	Decanoylcarnitine(C10:0)	-4.51E-02	0.0048	0.0968
5	35-Tetradecadiencarnitine(C14:2)	-4.92E-02	0.006	0.0968
6	Unassigned_33	-6.74E-02	0.0112	0.1506
7	PC(C34:2)	-6.85E-02	0.0136	0.1572
8	Unassigned_7	-1.97E-02	0.0165	0.1608
9	Dodecanoylcarnitine(C12:0)	-3.87E-02	0.0196	0.1608
10	L-Octanoylcarnitine(C8:0)	-3.54E-02	0.0217	0.1608
11	Unassigned_61	-9.28E-02	0.0218	0.1608
12	Decenoylcarnitine_1(C10:1)	-3.84E-02	0.0391	0.2566
13	Docosahexaenoicacid	-3.46E-02	0.0412	0.2566
14	Hexadecenoylcarnitine(C16:1)	-3.29E-02	0.0618	0.3577
15	Unassigned_58	-1.82E-02	0.0819	0.4423
16	Tetradecanoylcarnitine(C14:0)	-2.85E-02	0.0958	0.485
17	Unassigned_63	-4.97E-02	0.1192	0.568
18	Diacylglycerol(C36:4)	-1.99E-02	0.1278	0.568
19	Unassigned_17	-3.77E-02	0.1332	0.568
20	Progesterone	-2.16E-02	0.1412	0.572

id	Metabolite	Coefficient	p-value	q-value
1	Unassigned_7	-2.30E-02	0.0011	0.0521
2	Docosahexaenoic acid	-4.69E-02	0.0013	0.0521
3	Unassigned_58	-2.61E-02	0.0036	0.0811
4	Trans-2-Dodecenoylcarnitine(C12:1)	-4.05E-02	0.004	0.0811
5	Tetradecenoylcarnitine(C14:1)	-3.09E-02	0.0216	0.3465
6	Unassigned_63	-6.12E-02	0.0283	0.3465
7	Decanoylcarnitine(C10:0)	-3.00E-02	0.0299	0.3465
8	Leucine	-5.86E-02	0.0392	0.3701
9	Unassigned_61	-7.06E-02	0.0435	0.3701
10	Tryptophan	-5.36E-02	0.0483	0.3701
11	Diacylglycerol(C36:4)	-2.13E-02	0.0605	0.3701
12	Diacylglycerol(C36:3))	-1.89E-02	0.0694	0.3701
13	Unassigned_31	-3.10E-02	0.0707	0.3701
14	Hexadecenoylcarnitine(C16:1)	-2.71E-02	0.0758	0.3701
15	PC(C34:2)	-4.28E-02	0.0781	0.3701
16	3-Hydroxy-cis-5-tetradecenoylcarnitine(C14:1)	-1.99E-02	0.0795	0.3701
17	Unassigned_57	-2.23E-02	0.0838	0.3701
18	35-Tetradecadiencarnitine(C14:2)	-2.62E-02	0.0909	0.3701
19	L-Octanoylcarnitine(C8:0)	-2.23E-02	0.0945	0.3701
20	Unassigned_49	2.51E-02	0.0947	0.3701

Table 11 Top 20 strongest associations with obesity using repeated measurements until the age of 2 (left) and 7 (right) years, among birthweight-associated metabolites.

id	Metabolite	Coefficient	p-value	q-value
1	Progesterone	-1.04E-02	0.0010	0.0425
2	Trans-2-Dodecenoylcarnitine(C12:1)	-1.16E-02	0.0010	0.0425
3	Unassigned_57	-9.20E-03	0.0047	0.1206
4	Unassigned_19	-1.09E-02	0.0060	0.1206
5	Unassigned_58	-5.70E-03	0.0142	0.2307
6	PC(C34:2)	-1.36E-02	0.0241	0.2924
7	35-Tetradecadiencarnitine (C14:2)	-8.70E-03	0.0261	0.2924
8	Docosahexaenoicacid	-8.00E-03	0.0289	0.2924
9	Unassigned_7	-3.60E-03	0.0413	0.3603
10	Diacylglycerol(C36:4)	-5.70E-03	0.0445	0.3603
11	LysoPC(C18:3)	8.70E-03	0.0546	0.3771
12	Tetradecenoylcarnitine(C14:1)	-6.50E-03	0.0559	0.3771
13	Diacylglycerol(C36:3))	-5.00E-03	0.0639	0.3984
14	PC(36:4)	1.13E-02	0.0829	0.4701
15	Decenoylcarnitine_1(C10:1)	-6.80E-03	0.0909	0.4701
16	Unassigned_52	-5.40E-03	0.1015	0.4701
17	Unassigned_26	-4.60E-03	0.1088	0.4701
18	3-Hydroxy-cis-5-tetradecenoylcarnitine(C14:1)	-4.40E-03	0.1179	0.4701
19	LysoPC(C22:5)	5.10E-03	0.1223	0.4701
20	Cholesterol	1.76E-02	0.1225	0.4701

id	Metabolite	Coefficient	p-value	q-value
1	Unassigned_63	-2.84E-02	0.0110	0.8879
2	Decenoylcarnitine_2(C10:1)	1.68E-02	0.0224	0.9061
3	Unassigned_10	1.57E-02	0.0754	0.9941
4	Progesterone	-8.40E-03	0.1054	0.9941
5	Docosahexaenoicacid	-9.30E-03	0.1251	0.9941
6	Decenoylcarnitine_1(C10:1)	1.02E-02	0.1264	0.9941
7	LysoPC(C22:5)	8.00E-03	0.1539	0.9941
8	3-Hydroxyhexadecadienoylcarnitine(C16:1)	-8.80E-03	0.1752	0.9941
9	Unassigned_17	1.17E-02	0.1871	0.9941
10	Unassigned_46	-4.90E-03	0.1927	0.9941
11	Unassigned_19	-8.30E-03	0.1980	0.9941
12	LysoPC(C18:1)	-1.52E-02	0.2236	0.9941
13	LysoPC(C20:2)	7.10E-03	0.2571	0.9941
14	Unassigned_39	-5.40E-03	0.2584	0.9941
15	Unassigned_44	1.08E-02	0.2598	0.9941
16	2-Hexenoylcarnitine(C6:1)	-4.80E-03	0.2629	0.9941
17	35-Tetradecadiencarnitine(C14:2)	7.00E-03	0.2730	0.9941
18	Unassigned_58	-3.80E-03	0.3062	0.9941
19	Unassigned_48	6.70E-03	0.3349	0.9941
20	Trans-2-Dodecenoylcarnitine(C12:1)	-5.50E-03	0.3507	0.9941

Table 12 Top 20 strongest associations with height using repeated measurements until the age of 2 (left) and 7 (right) years, among birthweight-associated metabolites.

id	Metabolite	Coefficient	p-value	q-value
1	912-Hexadecadienoylcarnitine(C14:1)	-9.59E-01	0.0076	0.5525
2	Tetradecanoylcarnitine(C14:0)	-8.61E-01	0.0136	0.5525
3	Dodecanoylcarnitine(C12:0)	-7.04E-01	0.0333	0.8763
4	Hexadecenoylcarnitine(C16:1)	-7.01E-01	0.0463	0.8763
5	Unassigned_11	4.92E-01	0.0633	0.8763
6	Unknown(U88)	5.24E-01	0.0758	0.8763
7	Tetradecenoylcarnitine(C14:1)	-5.49E-01	0.0794	0.8763
8	Unassigned_1	-4.65E-01	0.0994	0.8763
9	Unassigned_63	-1.08E+00	0.1005	0.8763
10	Progesterone	-4.99E-01	0.1122	0.8763
11	35-Tetradecadiencarnitine(C14:2)	-5.27E-01	0.14	0.8763
12	Unassigned_57	4.49E-01	0.1469	0.8763
13	Retinol	6.37E-01	0.1668	0.8763
14	Unassigned_44	-7.78E-01	0.1772	0.8763
15	Methoxykynurenate(C11H9NO4)	-5.30E-01	0.1824	0.8763
16	Sphingosine	-2.87E-01	0.1834	0.8763
17	Indolelacticacid-Indolelactate	-6.95E-01	0.1852	0.8763
18	PC(C36:4)	-1.16E+00	0.2133	0.8763
19	Trans-2-Dodecenoylcarnitine(C12:1)	-4.00E-01	0.2212	0.8763
20	PC(C32:0)	-4.66E-01	0.2426	0.8763

id	Metabolite	Coefficient	p-value	q-value
1	Tetradecanoylcarnitine(C14:0)	-1.27E+00	0.0005	0.0427
2	912-Hexadecadienoylcarnitine(C14:1)	-1.19E+00	0.0016	0.0664
3	Dodecanoylcarnitine(C12:0)	-9.59E-01	0.0058	0.1382
4	Hexadecenoylcarnitine(C16:1)	-1.00E+00	0.0068	0.1382
5	Progesterone	-7.91E-01	0.0166	0.2682
6	Tetradecenoylcarnitine(C14:1)	-7.26E-01	0.0265	0.3575
7	35-Tetradecadiencarnitine(C14:2)	-7.95E-01	0.0329	0.3812
8	3-Hydroxyhexadecadienoylcarnitine(C16:1)	-7.79E-01	0.0396	0.4013
9	Unassigned_6	-5.44E-01	0.0585	0.4705
10	Unassigned_45	1.23E+00	0.0675	0.4705
11	Methoxykynurenate(C11H9NO4)	-7.38E-01	0.0748	0.4705
12	2-Hexenoylcarnitine(C6:1)	-4.92E-01	0.0796	0.4705
13	Trans-2-Dodecenoylcarnitine(C12:1)	-5.89E-01	0.0854	0.4705
14	Unassigned_4	-5.92E-01	0.0875	0.4705
15	PC(C36:4)	-1.63E+00	0.0912	0.4705
16	PC(C38:4)	-1.07E+00	0.0929	0.4705
17	Unassigned_46	-3.95E-01	0.1089	0.4739
18	Unassigned_7	-2.82E-01	0.1136	0.4739
19	Indolelactic acid	-8.57E-01	0.1269	0.4739
20	Unassigned_11	4.05E-01	0.1349	0.4739

Table 13 Top 20 strongest associations with rapid growth, among birthweight-associated metabolites.

Id	Metabolite	Coefficient	p-value	q-value
1	Cholestenone	0.54998	3.18E-08	2.58E-06
2	Unknown(U88)	-0.22354	4.68E-06	1.90E-04
3	Cholesterol	0.68047	0.00020	0.00544
4	PC(C32:0)	0.24883	0.00033	0.00544
5	PlasmalogenPC(36:4)orPC(O-36:5)orPC-O(C36:4)(C44H80NO7P)	0.35125	0.00034	0.00544
6	PC(C34:2)	0.33848	0.00067	0.00906
7	Unassigned_61	0.46091	0.00106	0.01230
8	PC(36:4)	0.32456	0.00187	0.01651
9	PC(30:0)	0.24857	0.00201	0.01651
10	PlasmalogenPC(36:3)orPC(O-36:4)	0.31035	0.00216	0.01651
11	PC(C36:4)	0.48891	0.00252	0.01651
12	Indolelactic acid	0.26352	0.00269	0.01651
13	Decenoylcarnitine_2(C10:1)	0.21469	0.00285	0.01651
14	Unassigned_63	0.33461	0.00301	0.01651
15	UnidentifiablePC/LysoPC	0.17535	0.00306	0.01651
16	Unassigned_26	0.12833	0.00533	0.02699
17	PlasmalogenPC(38:4)orPC(O-38:5)	0.28759	0.00582	0.02774
18	Butyrylcarnitine/Isobutyryl-L-carnitine(C4:0)	0.12774	0.00703	0.03165
19	Unassigned_12	0.17093	0.01010	0.04304
20	Unassigned_10	0.21687	0.01571	0.06363

3.3.3 Proteomics

The Proteomics analysis shows that several inflammatory proteins were associated with growth, overweight, obesity, height, weight and BMI (Table 14).

In particular, Interleukin 8 was significantly associated (after FDR correction) with BMI, overweight and obesity both at up to 2 years and up to 7 years (

*T*able 15*,Table 16 and Table 17). Periostin was associated with rapid growth up to 1 year, after FDR correction (*

Table 18).

Table 14 Number of associated features for the mixed effect models for repeated measurements up 2 and 7 years by outcome (BMI, height, obese, overweight) as well as rapid growth at 1 year old.

Model outcome	12 months		up to 2 yr		up to 7 yr	
	number of statistically significant features (total)	number of statistically significant features after applying FDR adjustment (total)	number of statistically significant features (total)	number of statistically significant features after applying FDR adjustment (total)	number of statistically significant features (total)	number of statistically significant features after applying FDR adjustment (total)
BMI	-	-	1 (16)	0 (16)	1 (16)	1 (16)
height	-	-	0 (16)	0 (16)	0 (16)	0 (16)
obese	-	-	2 (16)	2 (16)	3 (16)	2 (16)
overweight	-	-	2 (16)	1 (16)	1 (16)	1 (16)
rapid growth	3 (16)	1 (16)	13 (16)	0 (16)	10 (16)	0 (16)

*T*able 15 Associations of proteins with BMI using repeated measurements until the age of 2 (left) and 7 (right) years.

id	Inflammatory protein	Coefficient	p-value	q-value
1	IL.8.C	1.39E-03	0.0013	0.0214
2	GRO.a	-2.82E-04	0.4227	0.9896
3	IL.17	1.25E-02	0.6233	0.9896
4	TNF.a	2.70E-02	0.6761	0.9896
5	MIP1.B	2.67E-05	0.6897	0.9896
6	G.CSF	-9.07E-05	0.7178	0.9896
7	MDC.CC	-7.48E-05	0.7458	0.9896
8	Perios	-2.20E-07	0.7487	0.9896
9	CRP	-5.39E-05	0.8086	0.9896
10	EGF.2	-7.08E-05	0.8832	0.9896
11	IP.10	-2.49E-04	0.8846	0.9896
12	IL.1ra	-1.60E-06	0.9029	0.9896
13	IL.6	5.86E-05	0.9486	0.9896
14	VEGF	-5.66E-05	0.9584	0.9896
15	MCP.1	8.06E-06	0.9773	0.9896
16	MPO.5	-2.71E-08	0.9896	0.9896

id	Inflammatory protein	Coefficient	p-value	q-value
1	IL.8.C	1.25E-03	0.0024	0.0382
2	TNF.a	7.67E-02	0.2086	0.9883
3	GRO.a	-4.06E-04	0.2262	0.9883
4	EGF.2	-2.92E-04	0.5191	0.9883
5	VEGF	-6.34E-04	0.5315	0.9883
6	G.CSF	-1.20E-04	0.6232	0.9883
7	IL.17	-1.09E-02	0.6454	0.9883
8	MCP.1	-9.87E-05	0.7188	0.9883
9	IP.10	4.87E-04	0.7664	0.9883
10	MDC.CC	4.95E-05	0.8195	0.9883
11	IL.6	1.81E-04	0.8375	0.9883
12	CRP	-4.41E-05	0.8424	0.9883
13	MPO.5	3.10E-07	0.8737	0.9883
14	IL.1ra	-1.88E-06	0.8802	0.9883
15	MIP1.B	4.17E-06	0.945	0.9883
16	Perios	9.55E-09	0.9883	0.9883

Table 16 Associations of Inflammatory proteins with overweight using repeated measurements until the age of 2 (left) and 7 (right) years

id	Inflammatory protein	Coefficient	p-value	p-value
1	IL.8.C	3.12E-04	0.0000	0.0000

id	Inflammatory protein	Coefficient	p-value	p-value
1	IL.8.C	2.61E-04	0.0000	0.0000

2	TNF.a	1.23E-02	0.0837	0.6694
3	MPO.5	1.76E-07	0.3684	0.8705
4	IL.17	1.95E-03	0.4442	0.8705
5	MCP.1	-2.26E-05	0.4639	0.8705
6	IL.1ra	-1.03E-06	0.4759	0.8705
7	CRP	-1.64E-05	0.5144	0.8705
8	IP.10	1.17E-04	0.5484	0.8705
9	MDC.CC	-1.30E-05	0.6086	0.8705
10	G.CSF	-1.21E-05	0.6624	0.8705
11	EGF.2	2.15E-05	0.6654	0.8705
12	IL.6	-4.08E-05	0.6852	0.8705
13	Perios	1.81E-08	0.8026	0.8705
14	VEGF	2.26E-05	0.8482	0.8705
15	MIP1.B	-1.38E-06	0.8588	0.8705
16	GRO.a	6.17E-06	0.8705	0.8705

2	TNF.a	1.58E-02	0.0100	0.0804
3	Perios	7.10E-08	0.2688	0.9904
4	MCP.1	-2.42E-05	0.3875	0.9904
5	CRP	-1.81E-05	0.4382	0.9904
6	IL.17	1.64E-03	0.4750	0.9904
7	MPO.5	1.22E-07	0.5060	0.9904
8	VEGF	-5.68E-05	0.5782	0.9904
9	IL.6	4.90E-05	0.5902	0.9904
10	EGF.2	-1.60E-05	0.7174	0.9904
11	MDC.CC	-6.27E-06	0.7752	0.9904
12	MIP1.B	-1.61E-06	0.7912	0.9904
13	G.CSF	-4.55E-06	0.8564	0.9904
14	IP.10	2.65E-05	0.8776	0.9904
15	IL.1ra	-1.73E-08	0.9891	0.9904
16	GRO.a	-4.05E-07	0.9904	0.9904

Table 17 Associations of proteins with obesity using repeated measurements until the age of 2 (left) and 7 (right) years.

Id	Inflammatory protein	Coefficient	p-value	q-value
1	IL.8.C	0.00E+00	8.41E-05	0.0000
2	TNF.a	6.00E-04	5.22E-03	0.0046
3	VEGF	1.00E-02	6.54E-05	0.0531
4	MPO.5	9.43E-02	7.03E-08	0.3773
5	EGF.2	1.58E-01	1.45E-05	0.5045
6	MDC.CC	2.28E-01	-6.69E-06	0.6080
7	Perios	3.12E-01	1.64E-08	0.7127
8	IL.17	4.69E-01	4.10E-04	0.9387
9	IP.10	5.47E-01	-2.76E-05	0.9718
10	CRP	6.30E-01	-2.73E-06	0.9740
11	IL.1ra	8.13E-01	-7.39E-08	0.9740
12	G.CSF	8.57E-01	-1.09E-06	0.9740
13	GRO.a	8.67E-01	-1.38E-06	0.9740
14	IL.6	8.70E-01	-3.58E-06	0.9740
15	MCP.1	9.57E-01	-3.59E-07	0.9740
16	MIP1.B	9.74E-01	-6.13E-08	0.9740

id	Inflammatory protein	Coefficient	p-value	p-value
1	IL.8.C	9.99E-05	0.0000	0.0000
2	TNF.a	9.08E-03	0.0002	0.0020
3	IL.6	5.28E-05	0.1675	0.8766
4	MCP.1	-1.13E-05	0.3305	0.8766
5	Perios	2.48E-08	0.3497	0.8766
6	VEGF	3.13E-05	0.4512	0.8766
7	EGF.2	-1.26E-05	0.4629	0.8766
8	MDC.CC	-6.31E-06	0.4850	0.8766
9	CRP	-6.45E-06	0.5084	0.8766
10	MPO.5	3.96E-08	0.5713	0.8766
11	IP.10	-3.82E-05	0.6050	0.8766
12	GRO.a	-6.13E-06	0.6574	0.8766
13	IL.1ra	1.78E-07	0.7296	0.8910
14	IL.17	2.02E-04	0.8229	0.8910
15	G.CSF	2.20E-06	0.8354	0.8910
16	MIP1.B	-2.02E-07	0.9342	0.9342

Table 18 Associations of inflammatory proteins with rapid growth at 12 months

Id	Inflammatory protein	Coefficient	p-value	q-value
1	Perios	8.57E-07	0.0004	0.0069
2	IP.10	0.0015	0.0111	0.0886
3	MDC.CC	-0.0002	0.0438	0.2279
4	TNF.a	0.0476	0.0570	0.2279
5	IL.1ra	0.0000	0.1531	0.4510
6	MIP1.B	0.0000	0.1691	0.4510
7	CRP	-0.0001	0.3070	0.5925
8	GRO.a	-0.0001	0.3318	0.5925
9	G.CSF	-0.0001	0.3333	0.5925
10	IL.6	-0.0003	0.4274	0.6838
11	IL.8.C	0.0001	0.4907	0.7137
12	VEGF	-0.0002	0.6013	0.7410
13	MPO.5	-3.95E-07	0.6021	0.7410
14	MCP.1	-2.89E-05	0.7883	0.8485
15	IL.17	-2.43E-03	0.7955	0.8485
16	EGF.2	0.0000	0.9877	0.9877

Table 19 Associations of proteins with height using repeated measurements until the age of 2 (left) and 7 (right) years.

Id	Inflammatory protein	Coefficient	p-value	q-value
1	MCP.1	-1.21E-03	0.0549	0.7740
2	IL.6	-2.71E-03	0.1794	0.7740
3	G.CSF	-7.30E-04	0.1914	0.7740
4	MPO.5	-5.12E-06	0.2927	0.7740
5	GRO.a	-8.25E-04	0.2957	0.7740
6	IL.1ra	-3.03E-05	0.299	0.7740
7	VEGF	2.32E-03	0.3386	0.7740
8	MDC.CC	-4.11E-04	0.4248	0.8495
9	IP.10	-2.54E-03	0.5077	0.8610
10	EGF.2	-4.23E-04	0.699	0.8610
11	IL.17	2.13E-02	0.7169	0.8610
12	Perios	-5.63E-07	0.7175	0.8610
13	IL.8.C	-3.42E-04	0.7276	0.8610
14	TNF.a	-3.69E-02	0.7975	0.8610
15	CRP	1.21E-04	0.8072	0.8610
16	MIP1.B	1.41E-05	0.9239	0.9239

id	Inflammatory proteins	Estimate	p-value	q-value
1	MDC.CC	-1.01E-03	0.0572	0.8316
2	IP.10	-6.68E-03	0.1234	0.8316
3	MCP.1	-8.58E-04	0.2091	0.8316
4	GRO.a	-9.61E-04	0.2499	0.8316
5	MPO.5	-5.20E-06	0.3033	0.8316
6	G.CSF	-4.97E-04	0.4129	0.8316
7	CRP	3.41E-04	0.5451	0.8316
8	EGF.2	-6.62E-04	0.5471	0.8316
9	Perios	-8.78E-07	0.5951	0.8316
10	TNF.a	7.71E-02	0.6029	0.8316
11	IL.1ra	-1.54E-05	0.6076	0.8316
12	IL.6	-1.08E-03	0.6237	0.8316
13	VEGF	7.32E-04	0.7673	0.8963
14	IL.8.C	-2.68E-04	0.7843	0.8963
15	MIP1.B	-2.84E-05	0.8439	0.9002
16	IL.17	2.27E-03	0.9701	0.9701

3.4 Discussion

We have taken a three-pronged approach to assess the role of circulating metabolites and proteins in cord blood in predicting onset of obesogenic growth: An untargeted metabolome-wide analysis that gives a less-biased overview of the whole metabolome at the expense of lower statistical power due its high dimensionality and the need for subsequent downstream annotation work; a targeted analysis of pre-annotated metabolites known to be associated with weight at birth; and a targeted analysis of proteins related to inflammation, which is a known associate of adiposity.

In the metabolome-wide analysis, we identified multiple metabolomic features associated with rapid growth and with overweight, even up to the age of 7 years. The increased number of associations we observed with rapid growth may not be surprising, since this outcome is the closest in time to the point of metabolome measurement, reducing the influence of the post-natal environment. A suggestive finding was the relatively strong association of feature 278.0873_ 5.6216 (m/z_rt) with subsequent child overweight/obesity. This feature was tentatively annotated as TPPO, a food contaminant that is a potent inhibitor of TRPM5[45], which has been implicated in modifying the risk of diabetes and obesity through distinct functional roles in taste perception. While intriguing, this finding needs to be confirmed with further laboratory experiments. We also conducted a pathway enrichment analysis using *Mummichog*, a probabilistic algorithm that bypasses laboratory annotation bottlenecks, to assess effects of the metabolome overall. The most significant three pathways were related to fatty acid metabolism, which contribute to the regulation of energy balance and metabolic homeostasis by distinct mechanisms in the hypothalamus and are of known importance in the aetiology of obesity and type 2 diabetes mellitus [46]. The next most enriched pathway was tryptophan metabolism, which was also reported to be enriched in cord blood of children who are obese by[26]. Together these pathways suggest the role of metabolism at birth in influencing post-natal growth trajectories.

In the analysis of pre-annotated birthweight-related metabolites, we identified multiple metabolites predictive of rapid growth in the first year including cholesterol, and phosphatidylcholines, short chain acylcarnitines and idolactic acid (a tryptophan metabolite) that have been associated with obesity in cross-sectional studies. Lowered levels of progesterone, the most important pregnancy hormone, and trans-2-dodecenoylcarnitine(C12:1) were associated after FDR correction only with obesity at two years. Furthermore, there was evidence for an inverse association between docosahexaenoic acid (DHA), an omega-3 fatty acid with important anti-inflammatory properties, and overweight/BMI at 7 years.

BMI/overweight/obesity were associated with raised levels of Interleukin 8 (IL8), a cytokine involved in the inflammatory response. While IL8 has been reported to be raised among children who are obese in cross-sectional studies[47], this is the first time to the best of our knowledge, that cord blood levels have been associated with subsequent overweight/obesity. This result was consistent across cohorts and the cytokine is likely reflective of a sub-optimal intrauterine environment[48]. IL8 therefore provides a candidate molecular marker to trace causal pathways between prenatal risk-factors and child obesity. Rapid growth (up to 6 months and 12 months) was uniquely associated with Periostin, a protein also linked to obesity, asthma and other metabolic conditions[49].

In multivariate analysis using Random Forests we examined the prenatal risk factors [50] associated with overweight at 5 and 7 years (see supplementary materials for details of this analysis). Paternal BMI, parity, maternal weight gain at the end of pregnancy (kg), maternal pre-pregnancy BMI (kg/m2) and maternal passive smoking at any time during pregnancy were the 5 most important covariates in predicting overweight in our dataset. We will explore the role of these risk factors and others in influencing levels of molecular features identified here in the next steps. Preliminary work is reported in the second part of Addendum 3.

4 Retinal microcirculation

Blood pressure and retinal microvasculature were measured in 4- to 6-year old children in association with maternal pre-pregnancy BMI (Cox et al, submitted).

The risk of cardiovascular disease, the leading cause of death worldwide, may be set in prenatal life. Studies have reported positive associations between maternal pre-pregnancy BMI and offspring cardio-metabolic risk factors such as fat mass, glucose and insulin levels, and blood pressure, but these associations appear to be largely explained by offspring BMI. No studies so far have assessed potential alterations in the retinal microvasculature in association with maternal pre-pregnancy BMI.

Using data from 240 mother-newborn pairs that participated in the follow-up examination of the ENVIR*ON*AGE birth cohort study, we investigated the association between maternal pre-pregnancy BMI and anthropometric, blood pressure, and retinal microcirculation parameters in 4- to 6-year old children.

4.1 Methods

Within STOP, blood pressure and retinal microvasculature were measured in 4- to 6-year old children in association with maternal pre-pregnancy BMI. In short, anthropometric, blood pressure, and retinal microcirculation measurements were performed by a trained examiner in a quiet environment. Children's blood pressure was measured using the automated Omron 750IT (HEM-759-E) oscillometric device previously validated in children. According to a standard protocol, blood pressure measurements were taken on the right arm while children were seated with the upper arm at heart level. To ensure an accurate measurement, appropriate cuffs for children, size small (15-22 cm) or extra small (9-14 cm), were used depending on the child's arm circumference. Five consecutive measurements were taken with one-minute intervals. Systolic blood pressure (SBP) and diastolic blood pressure (DBP) were calculated by averaging the last three measurements. Images of the fundus of both the left and right eye of the child were taken with a Canon CR-2 plus 45° 6.3 megapixels digital nonmydriatic retinal camera (Hospithera, Brussels, Belgium). The central retinal arteriolar equivalent (CRAE) and the central retinal venular equivalent (CRVE), were calculated. The arterio-venous ratio (AVR) is the ratio between CRAE and CRVE. The tortuosity index (TI) is a measure for the curvature of the retinal vasculature. Average CRAE, CRVE, and TI values of both eyes were used in further analyses if both pictures were available.

4.2 Results

Maternal pre-pregnancy BMI was significantly and positively associated with child birth weight, BMI, waist circumference, blood pressure and retinal tortuosity. A 1 kg/m2 higher maternal pre-pregnancy BMI was associated with a 0.26 mm Hg (95% CI: 0.08, 0.44) higher mean arterial pressure, with similar estimates for systolic and diastolic blood pressure. Independent from the association with blood pressure, a 1 kg/m2 higher maternal pre-pregnancy BMI was associated with a 0.40 (95% CI: 0.01, 0.80) higher retinal tortuosity index (x 10^3). Evidence for direct intrauterine effects is supported by the following observations: associations for blood pressure and retinal tortuosity were independent from current child BMI, effect estimates for maternal pre-pregnancy BMI were mostly larger than those for maternal and paternal BMI at the follow-up visit, and estimates for paternal BMI did not reach significance (except for the association with child BMI). Retinal vessel diameter parameters (CRVE and AVR), however, were only significantly associated with maternal BMI at the follow-up visit.

Considering that blood pressure tracks from childhood into adulthood and that microvascular changes may be early markers of cardio-metabolic disease development, our results suggest that maternal pre-pregnancy is an important modifiable risk factor for later-life health of the offspring.

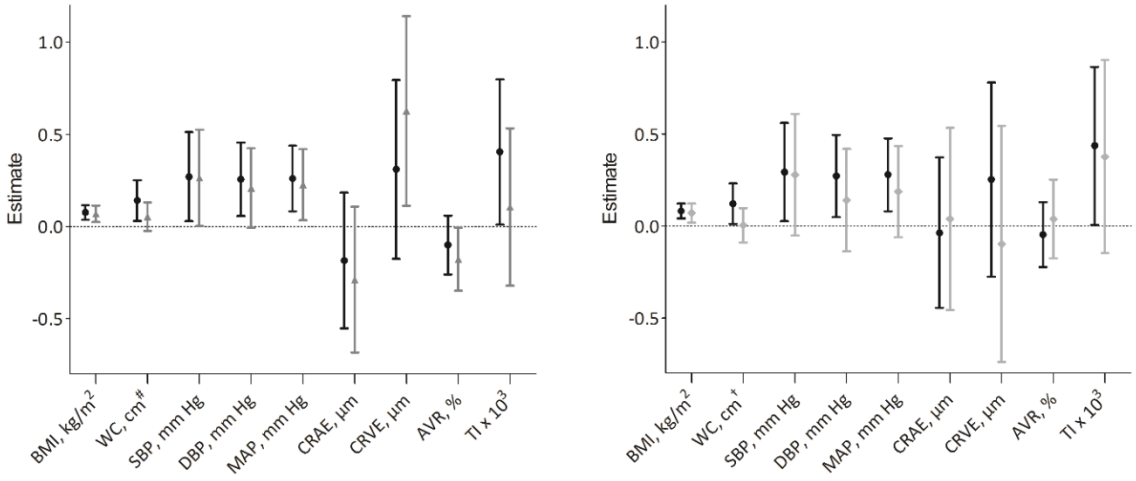


Figure 10 A comparison of estimates for maternal pre-pregnancy BMI and maternal BMI at the follow-up visit (left panel, n = 240), and a comparison of estimates for maternal pre-pregnancy BMI and paternal BMI at the follow-up visit (right panel, n = 214). Abbreviations: BMI, body mass index; WC, waist circumference; SBP, systolic blood pressure; DBP, diastolic blood pressure; MAP, mean arterial pressure; CRAE, central retinal arteriolar equivalent; CRVE, central retinal venular equivalent; AVR, arteriole-to-venule ratio; TI, tortuosity index. # n = 237.† n = 211.

Estimates represent the difference (with 95% CI) in child cardio-metabolic outcomes associated with a 1 kg/m^2 increase in maternal pre-pregnancy BMI (both panels, black circles) or a 1 kg/m^2 increase in maternal BMI at the follow-up visit (left panel, dark grey triangles) or a 1 kg/m^2 increase in paternal BMI at the follow-up visit (right panel, light grey diamonds). Estimates were adjusted for sex, gestational age, parity, newborn ethnicity, maternal age, maternal education, maternal smoking, gestational weight gain, date and season of follow-up visit, child age, and birth weight. Models for blood pressure and retinal microcirculation parameters were additionally adjusted for child BMI, models for retinal parameters were additionally adjusted for MAP, and models for CRAE were additionally adjusted for CRVE and vice versa. Child birth weight and BMI were modelled using a natural cubic spline with 3 degrees of freedom.

5 Telomere length tracking early in life in association with growth

Children telomere length is currently being measured in the same population as above, in triplicates in a 384-well format using the 7900HT. Fast Real-Time PCR System (Applied Biosystems). Relative average TLs are expressed as the ratio of telomere copy number to single copy gene number (T/S) relative to the average T/S ratio of the entire sample set. To reduce inter-run variability different inter-run calibrators are used. Control and non-template samples are run on each plate as well as reaction efficiency is evaluated on each reaction plate. Our method was validated with a U.S. based reference lab (laboratory of Prof. Dr. Immaculata de Vivo, Harvard T.H. Chan School of Public Health), and we were able to achieve inter-assay and intra-assay coefficients of variation between 6- 7%. In the future analysis we will study the growth curves in relation to telomere length at age of 4. Uniquely, as we recruited the children at birth and telomere length at birth is also measured, we will be able to study the changes over the early life course in association with growth.

6 Overall conclusions

Our suggestion is to consider three molecules for carry-over to the next steps, that may include replication in larger cohorts, testing in animals and testing in the context of RCTs in humans as short-term markers of the effectiveness of interventions. These are IL-8 (an inflammatory marker), CPK2 (a mitochondrial gene involved in glucose metabolism) and potentially triphenylphosphine oxide, pending laboratory confirmation.

Microcirculation of the retina is a very promising marker associated with metabolic disorders in infancy and also with maternal pre-conception weight, but it requires further replication and is not ready for inclusion e.g. in RCTs. Work on telomere length is still on-going.

7 References

1. Rzehak P, Covic M, Saffery R, et al. DNA-Methylation and Body Composition in Preschool Children: Epigenome-Wide-Analysis in the European Childhood Obesity Project (CHOP)-Study. *Sci Rep* 2017; **7**: 14349.
2. Martin CL, Jima D, Sharp GC, et al. Maternal pre-pregnancy obesity, offspring cord blood DNA methylation, and offspring cardiometabolic health in early childhood: an epigenome-wide association study. *Epigenetics* 2019; **14**: 325-40.
3. Lillycrop KA, Garratt ES, Titcombe P, et al. Differential SLC6A4 methylation: a predictive epigenetic marker of adiposity from birth to adulthood. *International Journal of Obesity* 2019; **43**: 974-88.
4. Fraser A, Macdonald-Wallis C, Tilling K, et al. Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol* 2013; **42**: 97-110.
5. Janssen BG, Madlhoum N, Gyselaers W, et al. Cohort Profile: The ENVIRonmental influence ON early AGEing (ENVIRONAGE): a birth cohort study. *Int J Epidemiol* 2017.
6. Kana MA, Rodrigues C, Fonseca MJ, Santos AC, Barros H. Effect of maternal country of birth on breastfeeding practices: results from Portuguese GXXI birth cohort. *Int Breastfeed J* 2018; **13**: 15.
7. Guxens M, Ballester F, Espada M, et al. Cohort Profile: the INMA--INfancia y Medio Ambiente--(Environment and Childhood) Project. *Int J Epidemiol* 2012; **41**: 930-40.
8. Farchi S, Forastiere F, Vecchi Brumatti L, et al. Piccolipiu, a multicenter birth cohort in Italy: protocol of the study. *BMC Pediatr* 2014; **14**: 36.
9. Chatzi LL, V.; Vafeiadi,M.; Koutra, K.; Roumeliotaki, T., Chalkiadaki, G., Karachaliou, M.; Daraki, V.; Kyriklaki, A.; Kampouri, M.; Fthenou, E.; Sarri K.; Vassilaki, M.; Fasoulaki, M; Bitsios, P.; Koutis, A.; Stephanou, E.G.; Kogevinas, M. Cohort Profile: The Mother-Child Cohort in Crete, Greece (Rhea Study). *Int J Epidemiol* 2017.
10. Chatzi L, Plana E, Daraki V, et al. Metabolic syndrome in early pregnancy and risk of preterm birth. *Am J Epidemiol* 2009; **170**: 829-36.
11. Maddock J, Wulaningsih W, Fernandez JC, et al. Associations between body size, nutrition and socioeconomic position in early life and the epigenome: A systematic review. *PLoS One* 2018; **13**: e0201672.
12. Ong KK, Ahmed ML, Emmett PM, Preece MA, Dunger DB. Association between postnatal catch-up growth and obesity in childhood: prospective cohort study. *BMJ (Clinical research ed)* 2000; **320**: 967-71.
13. Stark R, Guebre-Egziabher F, Zhao X, et al. A role for mitochondrial phosphoenolpyruvate carboxykinase (PEPCK-M) in the regulation of hepatic gluconeogenesis. *J Biol Chem* 2014; **289**: 7257-63.
14. Simpkin AJ, Suderman M, Gaunt TR, et al. Longitudinal analysis of DNA methylation associated with birth weight and gestational age. *Hum Mol Genet* 2015; **24**: 3752-63.
15. Kaufman J, Montalvo-Ortiz JL, Holbrook H, et al. Adverse Childhood Experiences, Epigenetic Measures, and Obesity in Youth. *J Pediatr* 2018; **202**: 150-6.e3.
16. Rangel-Huerta OD, Pastor-Villaescusa B, Gil A. Are we close to defining a metabolomic signature of human obesity? A systematic review of metabolomics studies. *Metabolomics : Official journal of the Metabolomic Society* 2019; **15**: 93-.
17. Rzehak P, Hellmuth C, Uhl O, et al. Rapid Growth and Childhood Obesity Are Strongly Associated with LysoPC(14:0). *Annals of Nutrition and Metabolism* 2014; **64**: 294-303.

18. Wahl S, Yu Z, Kleber M, et al. Childhood Obesity Is Associated with Changes in the Serum Metabolite Profile. *Obesity Facts* 2012; **5**: 660-70.
19. Cho K, Moon JS, Kang J-H, et al. Combined untargeted and targeted metabolomic profiling reveals urinary biomarkers for discriminating obese from normal-weight adolescents. *Pediatric Obesity* 2017; **12**: 93-101.
20. Perng W, Gillman MW, Fleisch AF, et al. Metabolomic profiles and childhood obesity. *Obesity* 2014; **22**: 2570-8.
21. Butte NF, Liu Y, Zakeri IF, et al. Global metabolomic profiling targeting childhood obesity in the Hispanic population. *The American Journal of Clinical Nutrition* 2015; **102**: 256-67.
22. Lau C-HE, Siskos AP, Maitre L, et al. Determinants of the urinary and serum metabolome in children from six European populations. *BMC Medicine* 2018; **16**: 202.
23. Hellmuth C, Kirchberg FF, Brandt S, et al. An individual participant data meta-analysis on metabolomics profiles for obesity and insulin resistance in European children. *Scientific Reports* 2019; **9**: 5053.
24. Leal-Witt MJ, Ramon-Krauel M, Samino S, et al. Untargeted metabolomics identifies a plasma sphingolipid-related signature associated with lifestyle intervention in prepubertal children with obesity. *International Journal of Obesity* 2018; **42**: 72-8.
25. Reinehr T, Wolters B, Knop C, et al. Changes in the serum metabolite profile in obese children with weight loss. *European Journal of Nutrition* 2015; **54**: 173-81.
26. Isganaitis E, Rifas-Shiman SL, Oken E, et al. Associations of cord blood metabolites with early childhood obesity risk. *Int J Obes (Lond)* 2015; **39**: 1041-8.
27. Hellmuth C, Uhl O, Standl M, et al. Cord Blood Metabolome Is Highly Associated with Birth Weight, but Less Predictive for Later Weight Development. *Obesity Facts* 2017; **10**: 85-100.
28. Janssen BG, Madhloum N, Gyselaers W, et al. Cohort profile: the ENVIRonmental influence ON early AGEing (ENVIR ON AGE): a birth cohort study. 2017; **46**: 1386-7m.
29. Guxens M, Ballester F, Espada M, et al. Cohort profile: the INMA—INfancia y Medio Ambiente—(environment and childhood) project. 2011; **41**: 930-40.
30. Felix JF, Joubert BR, Baccarelli AA, et al. Cohort profile: pregnancy and childhood epigenetics (PACE) consortium. *International Journal of Epidemiology* 2017; **47**: 22-3u.
31. Chatzi L, Leventakou V, Vafeiadi M, et al. Cohort profile: the mother-child cohort in Crete, Greece (Rhea study). 2017; **46**: 1392-3k.
32. World Health Organization W. WHO Anthro (version 3.2. 2, January 2011) and macros. *World Health Organization, Geneva, Switzerland http://wwwwhoint/childgrowth/software/en* 2011.
33. World Health Organization %J World Health Organization G, Switzerland. http://www. who. int/childgrowth/software/en. WHO Anthro (version 3.2. 2, January 2011) and macros. 2011.
34. WHO WHOJG. Training course on child growth assessment. 2008: p17-25.
35. De Onis M, Lobstein T. Defining obesity risk status in the general childhood population: which cut-offs should we use? : Taylor & Francis; 2010.
36. Ong KK, Ahmed ML, Emmett PM, Preece MA, Dunger DBJB. Association between postnatal catch-up growth and obesity in childhood: prospective cohort study. 2000; **320**: 967-71.
37. Lubin JH, Colt JS, Camann D, et al. Epidemiologic evaluation of measurement data in the presence of detection limits. *Environmental health perspectives* 2004; **112**: 1691-6.
38. Benjamini Y, Hochberg YJJotRsssB. Controlling the false discovery rate: a practical and powerful approach to multiple testing. 1995; **57**: 289-300.
39. Robinson O, Keski-Rahkonen P, Chatzi L, et al. Cord Blood Metabolic Signatures of Birth Weight: A Population-Based Study. *Journal of proteome research* 2018; **17**: 1235-47.
40. Bates D, Maechler M, Bolker B, Walker S. lme4: Linear mixed-effects models using Eigen and S4. *R package version* 2014; **1**: 1-23.
41. Wickham H. *ggplot2: elegant graphics for data analysis*: Springer; 2016.
42. Li S, Park Y, Duraisingham S, et al. Predicting network activity from high throughput metabolomics. *PLoS computational biology* 2013; **9**: e1003123.

43. Cappiello C, Francalanci C, Pernici B. *Data quality assessment from the user's perspective*. Proceedings of the 2004 international workshop on Information quality in information systems; 2004: ACM; 2004. p. 68-73.
44. Chong J, Soufan O, Li C, et al. MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Research* 2018; **46**: W486-W94.
45. Vennekens R, Mesuere M, Philippaert K. TRPM5 in the battle against diabetes and obesity. *Acta Physiologica* 2018; **222**: e12949.
46. Blaak EE. Fatty acid metabolism in obesity and type 2 diabetes mellitus. *Proceedings of the Nutrition Society* 2007; **62**: 753-60.
47. Tam CS, Garnett SP, Cowell CT, et al. IL-6, IL-8 and IL-10 levels in healthy weight and overweight children. *Hormone research in paediatrics* 2010; **73**: 128-34.
48. Otsubo Y, Hashimoto K, Kanbe T, Sumi M, Moriuchi H. Association of cord blood chemokines and other biomarkers with neonatal complications following intrauterine inflammation. *PLOS ONE* 2017; **12**: e0175082.
49. Kimura H, Konno S, Makita H, et al. Serum periostin is associated with body mass index and allergic rhinitis in healthy and asthmatic subjects. *Allergology International* 2018; **67**: 357-63.
50. Maheswaran T, Abikshyeet P, Sitra G, Gokulanathan S, Vaithiyanadane V, Jeelani S. Gustatory dysfunction. *Journal of pharmacy & bioallied sciences* 2014; **6**: S30.

Addendum 1

Systematic review of the literature on epigenetics and mediation analysis

Strategy for selection of the DNA methylation targets through literature review

Previous systematic literature review (2017)

DNA methylation targets have been selected from the literature exploring growth outcomes and DNA methylation of newborns, children or adolescents. In particular, we used 56 papers identified from a systematic review of the association of body size in early life (defined as weight, height, BMI or change in any change of these measurements from birth till 12 years of age) and the epigenome. We excluded papers dealing with any body size measurement other than BMI, weight, height (eg. birthweight, birthlength, head circumference, subscapular triceps skinfold etc.). In total 6 papers identified at least one CpG or gene region in association with body size measurements.

Update of the previous systematic literature review

As the review coverage was limited to March 2017, we updated the search to Pubmed engine from March 2017 to September 2019 via the R package RISmed(45) using the following terms:

[1] "neonat* AND obesity AND methylation"

[2] "child* AND obesity AND methylation"

[3] "infant AND obesity AND methylation"

[4] "adolescent* AND obesity AND methylation"

[5] "neonat* AND overweight AND methylation"

[6] "child* AND overweight AND methylation"

[7] "infant AND overweight AND methylation"

[8] "adolescent* AND overweight AND methylation"

[9] "neonat* AND BMI AND methylation"

[10] "child* AND BMI AND methylation"

[11] "infant AND BMI AND methylation"

[12] "adolescent* AND BMI AND methylation"

[13] "neonat* AND body mass index AND methylation"

[14] "child* AND body mass index AND methylation"

[15] "infant AND body mass index AND methylation"

[16] "adolescent* AND body mass index AND methylation"

[17] "neonat* AND weight AND methylation"

[18] "child* AND weight AND methylation"

[19] "infant AND weight AND methylation"

[20] "adolescent* AND weight AND methylation"

[21] "neonat* AND height AND methylation"

[22] "child* AND height AND methylation"

[23] "infant AND height AND methylation"

[24] "adolescent* AND height AND methylation"

[25] "neonat* AND length AND methylation"

[26] "child* AND length AND methylation"

[27] "infant AND length AND methylation"

[28] "adolescent* AND length AND methylation"

Results were restricted to papers published from March 2017 onward.

In addition, as the review coverage was limited to 0-12 years of age, we expanded the search in the Pubmed engine till March 2017 via the R package RISmed(45) using the following terms:

[1] adolescent* AND obesity AND methylation"

[2] "adolescent* AND overweight AND methylation"

[3] "adolescent* AND BMI AND methylation"

[4] "adolescent* AND body mass index AND methylation"

[5] "adolescent* AND weight AND methylation"

[6] "adolescent* AND height AND methylation"

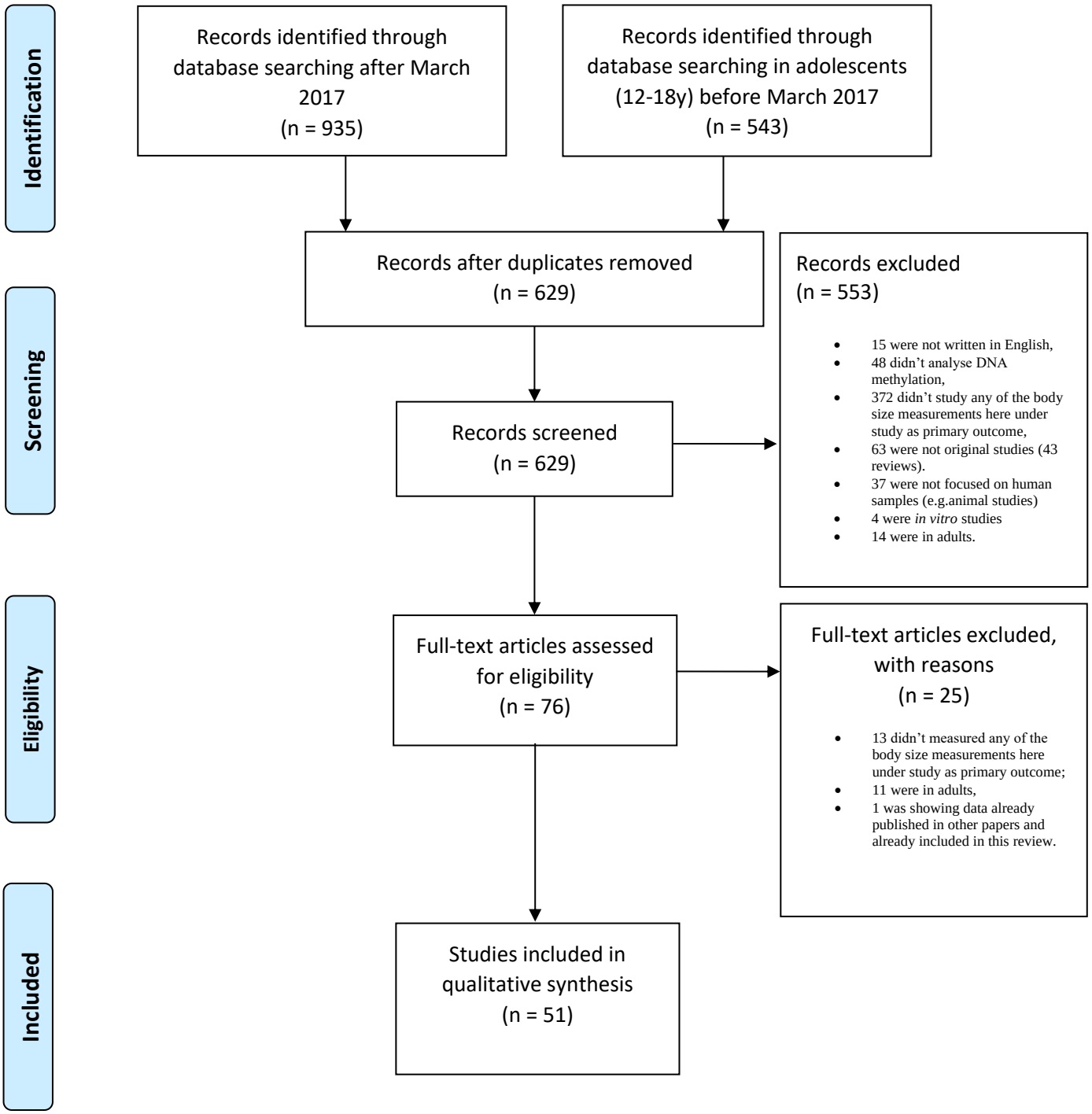
[7] "adolescent* AND length AND methylation"

As reported in the **Figure S1**, the search resulted in 1478 references. Articles that were duplicates, non-English, non-original, *in vitro,* in animals, in human adults, or not dealing with body size in early life (from birth up to 18 years of age) or DNA methylation were excluded. In total, 51 articles were included. Among these, 39 identified at least one CpG or gene region as associated with childhood body size measurements. Additionally, 3 articles were added via search of the references as they identified additional CpGs or gene regions in association with childhood body size measurements. As some findings referred to gene regions or either single CpGs measured with detection methods different from Infinium Human Methylation 450K BeadChip and for which we weren't able to identify an exact corresponding CpG, we considered the entire gene regions as associated with the body size measurements.

Identification of DNA methylation targets

In total, we identified methylation of 438 CpGs associated to body size measurements in children and adolescents in addition to 28 gene regions. 365 CpGs out of the 438 were available from Infinium Human Methylation 450K BeadChip. These, in addition to the 28 gene regions were mapped to 604 CpG. In total, the candidate list of CpGs used for the present study counts 966 CpGs (**Supplemental Table 1**).

Figure S1. PRISMA flow diagram of the study selection.



Supplementary Table 1. Candidate CpGs used in analysis

growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
BMI in chidhood	cg23415434	ARHGEF4	- /+	Clark (2019)	Y
BMI in chidhood	cg10895452	EN1	+	Clark (2019)	Y
BMI in chidhood	cg11097064	HDAC4	+	Clark (2019)	N
BMI in chidhood	cg02447542		- /+	Clark (2019)	N
BMI in chidhood	cg23819411	MCF2L2	- /+	Clark (2019)	Y
BMI in chidhood	cg20627174		- /+	Clark (2019)	Y
BMI in chidhood	cg18670801		-	Clark (2019)	N
BMI in chidhood	cg13211302		-/ +	Clark (2019)	Y
BMI in chidhood	cg04786207	MEST/MESTIT 1	+	Clark (2019)	Y
BMI in chidhood	cg00702231	PLAGL/HYMAI	-	Clark (2019)	Y
BMI in chidhood	cg14886255	MGAT4B	+	Clark (2019)	N
BMI in chidhood	cg26573334	TPCN1	+	Clark (2019)	Y
BMI in chidhood	cg26328335	AQP5	+	Clark (2019)	Y
BMI in chidhood	cg00797058		-/ +	Clark (2019)	Y
BMI in chidhood	cg20088245		-	Clark (2019)	Y
BMI in chidhood	cg04834608	DNASE1	-	Clark (2019)	Y
BMI in chidhood	cg04478251	ABR	+	Clark (2019)	Y
BMI in chidhood	cg26536949		-/ +	Clark (2019)	Y
BMI in chidhood	cg05528899		-	Clark (2019)	Y
BMI in chidhood	cg06417478	HOOK2	-/ +	Clark (2019)	Y
BMI in chidhood	cg00130086		-/ +	Clark (2019)	N
BMI in chidhood	cg10296238	SPATC1L	+	Clark (2019)	Y
BMI in chidhood	cg13012494	SPATC1L	-/ +	Clark (2019)	Y
BMI in chidhood	cg11324748	ANKRD20A11P	-	Clark (2019)	Y
BMI in chidhood	cg00566515	ANKRD20A11P	-	Clark (2019)	Y
BMI in chidhood	cg27306467	ANKRD20A11P	-/ +	Clark (2019)	Y
BMI in chidhood	cg10386298		+	Clark (2019)	Y
BMI in chidhood	cg19633822	MGC13005	-	Clark (2019)	Y
BMI in chidhood	cg01101448		-	Clark (2019)	Y
BMI in chidhood	cg25461508	RGS14	-	Clark (2019)	Y
BMI in chidhood	cg17597639	SDHAP3	+	Clark (2019)	Y
BMI in chidhood	cg12301216		+	Clark (2019)	N
BMI in chidhood	cg12757684	PLAGL1/HYMAI	-	Clark (2019)	Y
BMI in chidhood	cg16807687	PCDH21	+	Clark (2019)	Y
BMI in chidhood	cg02276826		+	Clark (2019)	Y
BMI in chidhood	cg04824767	FADS3	+	Clark (2019)	N
BMI in chidhood	cg24213572	NDUFS8/MIR4691	-	Clark (2019)	N
BMI in chidhood	cg13906860		+	Clark (2019)	Y

growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
BMI in chidhood	cg03972466	C16orf72	-	Clark (2019)	Y
BMI in chidhood	cg20505444	BCL2	+	Clark (2019)	N
BMI in chidhood	cg13732083	SPATC1L	-/ +	Clark (2019)	Y
BMI in chidhood	cg06019884	ARHGEF4	+	Clark (2019)	N
BMI in chidhood	cg02366959	NCRNA00164	-	Clark (2019)	Y
BMI in chidhood	cg17228698	CHST12	+	Clark (2019)	Y
BMI in chidhood	cg16247391	ZNF890P	+	Clark (2019)	N
BMI in chidhood	cg26507521		-	Clark (2019)	Y
BMI in chidhood	cg15490703	EFNA2	+	Clark (2019)	Y
obesity in adolescence	cg16507569	COL11A2	-	Cao-Lei (2019)	Y
obesity in adolescence	cg27289463		+	Cao-Lei (2019)	Y
obesity in adolescence	cg25013753	ARHGAP22	-	Cao-Lei (2019)	Y
obesity in adolescence	cg19697575		-	Cao-Lei (2019)	Y
obesity in adolescence	cg26935333		-	Cao-Lei (2019)	Y
obesity in adolescence	cg14007688	DBH	+	Cao-Lei (2019)	Y
obesity in adolescence	cg22459517	EPS8L1	+	Cao-Lei (2019)	Y
obesity in adolescence	cg25813936	GNE	+	Cao-Lei (2019)	Y
obesity in adolescence	cg00320354	TSPAN5	+	Cao-Lei (2019)	Y
obesity in adolescence	cg08282428	RBM46	-	Cao-Lei (2019)	Y
obesity in adolescence	cg10977910	TTLL7	-	Cao-Lei (2019)	Y
obesity in adolescence	cg27413643	ANKRD27;RGS9BP	-	Cao-Lei (2019)	Y
obesity in adolescence	cg06549275	ASAP2	+	Cao-Lei (2019)	Y
obesity in adolescence	cg26658743		-	Cao-Lei (2019)	Y
obesity in adolescence	cg05248234	FSCN2	-	Cao-Lei (2019)	Y
obesity in adolescence	cg00340052	AGAP1	-	Cao-Lei (2019)	Y
obesity in adolescence	cg19181528		+	Cao-Lei (2019)	Y
obesity in adolescence	cg01835922		-	Cao-Lei (2019)	Y
obesity in adolescence	cg03877767	GREB1	-	Cao-Lei (2019)	Y

growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
obesity in adolescence	cg07248242	TSPAN9	-	Cao-Lei (2019)	Y
obesity in adolescence	cg00152799	TXNL1	-	Cao-Lei (2019)	Y
obesity in adolescence	cg06916011	EPS8L1	+	Cao-Lei (2019)	Y
obesity in adolescence	cg21158163		+	Cao-Lei (2019)	Y
obesity in adolescence	cg21040096	RPH3AL	+	Cao-Lei (2019)	Y
obesity in adolescence	cg09949906	HLA-DRB1	+	Cao-Lei (2019)	Y
obesity in adolescence	cg14377951	GSTM5	+	Cao-Lei (2019)	Y
obesity in adolescence	cg01520956	FAM47E	+	Cao-Lei (2019)	Y
obesity in adolescence	cg23336892		-	Cao-Lei (2019)	Y
obesity in adolescence	cg05808124		+	Cao-Lei (2019)	Y
obesity in adolescence	cg10764750		+	Cao-Lei (2019)	Y
obesity in adolescence	cg06508867		-	Cao-Lei (2019)	Y
obesity in adolescence	cg26590106	HLA-DRB1	-	Cao-Lei (2019)	Y
obesity in adolescence	cg27112983		+	Cao-Lei (2019)	Y
obesity in adolescence	cg20756026	HEXDC	+	Cao-Lei (2019)	Y
obesity in adolescence	cg07618759	COL4A2	+	Cao-Lei (2019)	Y
obesity in adolescence	cg06052372		+	Cao-Lei (2019)	Y
obesity in adolescence	cg11317459		-	Cao-Lei (2019)	Y
obesity in adolescence	cg11251367	FMN2	-	Cao-Lei (2019)	Y
obesity in adolescence	cg07437923	DNAJA3	-	Cao-Lei (2019)	Y
obesity in adolescence	cg05834845	MUC4	+	Cao-Lei (2019)	Y
obesity in adolescence	cg27362989	HLA-DRB5	-	Cao-Lei (2019)	Y
obesity in adolescence	cg23477460		-	Cao-Lei (2019)	Y
obesity in adolescence	cg05545777		+	Cao-Lei (2019)	Y

growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
obesity in adolescence	cg15586842	ANXA8L2	+	Cao-Lei (2019)	Y
obesity in adolescence	cg158269415	PTPRN2	+	Lee (2019)	N
BMI in adolescence	Chr6:100903612	SIM1	+	He(2019)	N
BMI in adolescence	CpG5	SLC6A4	-	Lillycrop (2019)	N
BMI childhood	cg10264529	PCK2		Kaufman (2018)	Y
BMI childhood	cg14929207	DHRS13		Kaufman (2018)	Y
BMI childhood	cg16110788			Kaufman (2018)	Y
BMI childhood	cg14855841	CXCL10		Kaufman (2018)	Y
BMI childhood	cg26103104			Kaufman (2018)	Y
BMI childhood	cg01555853	KCNS3		Kaufman (2018)	Y
BMI childhood	gene region	PPARA	+	Kochmanski (2019)	
WFL in childhood	gene region	MEST		Nahm (2018)	
WFL in childhood	gene region	PEG10		Nahm (2018)	
WFL in childhood	gene region	IGF2		Nahm (2018)	
WFL in childhood	gene region	PLAGL1		Nahm (2018)	
WFL in childhood	gene region	PEG3		Nahm (2018)	
BMI in chidhood	cg05995465	HDAC4	+	Li(2018)	Y
BMI in chidhood	cg27288829	RAX2	+	Li(2018)	Y
BMI in chidhood	cg12556569	APOA5	-	Li(2018)	Y
BMI in chidhood	cg02301920	CES1	-	Li(2018)	Y
BMI in chidhood	cg10059324	PER3	-	Samblas (2018)	Y
BMI in chidhood	cg04757389	PTPRS	-	Samblas (2018)	Y
midchildhood weight	gene region	LPL	+	Gagne-oullet (2017)	
midchildhood fat mass	gene region	LPL	+	Gagne-oullet (2017)	
midchildhood lean mass	gene region	LPL	-	Gagne-oullet (2017)	
childhood fat mass	gene region	CDKN2A	-	Lillycrop (2017)	
BMI in chidhood	gene region	IL16	+	Wu (2018)	
BMI in chidhood	gene region	CYP27B1	+	Acs (2017)	
BMI in chidhood	gene region	IGF2	-	Acs (2017)	
BMI in chidhood	cg07814318	FLK13	+	Koh (2018)	Y
childhood fat mass	gene region	MEG3	-	Prats (2017)	
lenght increase during the 1st year	gene region	MEG3	-	Prats (2017)	
BMI in chidhood	cg23314826		-	Fradlin (2017)	Y

growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
BMI in chidhood	cg11324650		-	Fradlin (2017)	Y
BMI in chidhood	cg08468401		-	Fradlin (2017)	Y
BMI in chidhood	cg04052934		-	Fradlin (2017)	Y
BMI in chidhood	cg06769708	DLGAP4	-	Fradlin (2017)	Y
BMI in chidhood	cg19835478	COL6A1	+	Fradlin (2017)	Y
BMI in chidhood	cg26645655		-	Fradlin (2017)	Y
BMI in chidhood	cg00562180		+	Fradlin (2017)	Y
BMI in chidhood	cg07060864		+	Fradlin (2017)	Y
BMI in chidhood	cg13844049		-	Fradlin (2017)	Y
BMI in chidhood	cg26834418	CHORDC1	-	Fradlin (2017)	Y
BMI in chidhood	cg16427670	ARHGEF7	+	Fradlin (2017)	Y
BMI in chidhood	cg16705578		-	Fradlin (2017)	Y
BMI in chidhood	cg26138821	CHORDC1	-	Fradlin (2017)	Y
BMI in chidhood	cg14720024		-	Fradlin (2017)	Y
BMI in chidhood	cg00103778		-	Fradlin (2017)	Y
BMI in chidhood	cg06686058	ADM	+	Fradlin (2017)	Y
BMI in chidhood	cg08025415	TLE2	-	Fradlin (2017)	Y
BMI in chidhood	cg16906712	HIST1H2BG	-	Fradlin (2017)	Y
BMI in chidhood	cg23043119	STARD4	+	Fradlin (2017)	Y
BMI in chidhood	cg07607077	C7orf20	-	Fradlin (2017)	Y
BMI in chidhood	cg02399233	NADSYN1	-	Fradlin (2017)	Y
BMI in chidhood	cg08515072		-	Fradlin (2017)	Y
BMI in chidhood	cg12093371		+	Fradlin (2017)	Y
BMI in chidhood	cg06918474	MARCH3	+	Fradlin (2017)	Y
BMI in chidhood	cg16331674		+	Fradlin (2017)	Y
BMI in chidhood	cg15991546	DPYD	-	Fradlin (2017)	Y
BMI in chidhood	cg17128312	FBXW9	+	Fradlin (2017)	Y
BMI in chidhood	cg16334849		-	Fradlin (2017)	Y
BMI in chidhood	cg03607644		-	Fradlin (2017)	Y
BMI in chidhood	cg24317742		-	Fradlin (2017)	Y
BMI in adolescence	gene region	LEP	-	Dunstan (2017)	
weight in childhood	cg00510507	ANK3	+	Lin (2017)	Y
weight in childhood	cg08390209	CDKN2B	+	Lin (2017)	Y
weight in childhood	cg23671997	IGDCC4	+	Lin (2017)	Y
weight in childhood	cg14300531	P4HA3	-	Lin (2017)	Y
weight in childhood	cg25685359	MIRLET7BHG	-	Lin (2017)	Y
weight in childhood	cg22383874	CACNA1G	+	Lin (2017)	Y
weight in childhood	cg02729344	ZNF423	+	Lin (2017)	Y
weight in childhood	cg25487405	MIRLET7BHG	-	Lin (2017)	Y
BMI in childhood	gene region	POMC	+	Kühnen (2016)	
obesity in adolescence	cg26846943	FYN	+	Huang (2015)	Y

growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
obesity in adolescence	Chr6: 112165053	FYN	+	Huang (2015)	N
obesity in adolescence	Chr6: 112165057	FYN	+	Huang (2015)	N
obesity in adolescence	cg16436762	PIWIL4	-	Huang (2015)	Y
obesity in adolescence	cg17627898	TAOK3	-	Huang (2015)	Y
BMI in childhood	46801699	HIF3A	+	Wang (2015)	N
obesity in childhood	gene region	FAIM2	-/ +	Wu (2015)	
BMI in adolescence	gene region	LY86	+	Su (2015)	
BMI in adolescence	gene region	ADIPOQ	-	García-Cardona (2014)	
BMI in adolescence	gene region	LEP	-	García-Cardona (2014)	
BMI and weight in adolescence	gene region	*AQP9*	-/ +	Moleres (2013)	
BMI and weight in adolescence	gene region	DUSP22	+	Moleres (2013)	
BMI and weight in adolescence	gene region	HIPK3	-	Moleres (2013)	
BMI and weight in adolescence	gene region	TNNI3	-	Moleres (2013)	
BMI and weight in adolescence	gene region	TNNT1	+	Moleres (2013)	
obesity in childhood	gene region	POMC	+	Kühnen (2012)	
obesity in adolescence	13517	UBASH3A	+	Wang (2012)	N
obesity in adolescence	6160	CTNND1	-	Wang (2012)	N
obesity in adolescence	17029	TRIM3	-	Wang (2012)	N
BMI in childhood	cg27146050	HIF3A	+	Richmond (2016)	Y
Height in adolescence	cg14597739	LTA	-	Simpkin (2015)	Y
weight in adolescence	cg15783941	NFIX	-	Simpkin (2015)	Y
Height in adolescence	cg15783941	NFIX	-	Simpkin (2015)	Y
growth	gene region	*IGF2*	-	Bouwland-Both (2013)	
BMI in childhood	chr5:135414858	VTRNA2-1	-	van Dijk (2017)	cg11852404
BMI in childhood	chr11:2889602	KCNQ1DN		van Dijk (2017)	cg26690742
BMI in childhood	chr6:32063394	TNXB		van Dijk (2017)	cg24475062
BMI trajectory	cg23381058		+	Sherwood (2019)	N
BMI trajectory	cg05091920		-	Sherwood (2019)	N
BMI in chidhood	cg13850887	*SNED1*	-	Rzehak (2017)	Y

growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
BMI in chidhood	cg01706498	*KLHL6*	+	Rzehak (2017)	Y
BMI in chidhood	cg26401512	*ZNF643*	-	Rzehak (2017)	Y
BMI in chidhood	cg21525627	*(ZDHHC17)*	+	Rzehak (2017)	Y
BMI in chidhood	cg26867987	*COL11A2*	-	Rzehak (2017)	Y
BMI in chidhood	cg17810765	*ANO7*	+	Rzehak (2017)	Y
BMI in chidhood	cg14518658	*(CYTH4-ELFN2)*	+	Rzehak (2017)	Y
BMI in chidhood	cg21126338	*FARP1*	+	Rzehak (2017)	Y
BMI in chidhood	cg27285599	*TBCD;ZNF750*	+	Rzehak (2017)	Y
BMI in chidhood	cg17935297	*CILP2*	+	Rzehak (2017)	Y
BMI in chidhood	cg18704658		+	Rzehak (2017)	Y
BMI in chidhood	cg15026574	*ST6GAL1*	-	Rzehak (2017)	Y
BMI in chidhood	cg26917480	*ADAP2*	+	Rzehak (2017)	Y
BMI in chidhood	cg08074767		-	Rzehak (2017)	Y
BMI in chidhood	cg14391016	*CCR6*	+	Rzehak (2017)	Y
BMI in chidhood	cg06322432		+	Rzehak (2017)	Y
BMI in chidhood	cg14401837	*NPSR1*	-	Rzehak (2017)	Y
BMI in chidhood	cg17989572	*RAB5C*	-	Rzehak (2017)	Y
BMI in chidhood	cg06369443	*KCNQ4*	+	Rzehak (2017)	Y
BMI in chidhood	cg06594770	*TRIOBP*	+	Rzehak (2017)	Y
BMI in chidhood	cg26995653		+	Rzehak (2017)	Y
BMI in chidhood	cg06437396		+	Rzehak (2017)	Y
BMI in chidhood	cg02682525	*ANKK1*	+	Rzehak (2017)	Y
BMI in chidhood	cg24751284	*APEX1*	+	Rzehak (2017)	Y
BMI in chidhood	cg06376715	*TP73*	+	Rzehak (2017)	Y

growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
BMI in chidhood	cg14950834	*SPTBN4;BLVRB*	-	Rzehak (2017)	Y
BMI in chidhood	cg26618041	*CILP2*	+	Rzehak (2017)	Y
BMI in chidhood	cg14795409		-	Rzehak (2017)	Y
BMI in chidhood	cg01257171		+	Rzehak (2017)	Y
BMI in chidhood	cg23416307	*GAK*	+	Rzehak (2017)	Y
BMI in chidhood	cg15941159	*GRAMD4*	+	Rzehak (2017)	Y
BMI in chidhood	cg15650694	*SFRS12*	+	Rzehak (2017)	Y
BMI in chidhood	cg21332304	*NPEPPS*	-	Rzehak (2017)	Y
BMI in chidhood	cg24332767	*C3orf70*	+	Rzehak (2017)	Y
BMI in chidhood	cg06217450		+	Rzehak (2017)	Y
BMI in chidhood	cg08568550	*C11orf16*	+	Rzehak (2017)	Y
BMI in chidhood	cg04813904	*CD180*	+	Rzehak (2017)	Y
BMI in chidhood	cg04979599	*AFAP1L1*	+	Rzehak (2017)	Y
BMI in chidhood	cg13641993	*FBXO10*	+	Rzehak (2017)	Y
BMI in chidhood	cg26666886	*ANKRD11*	-	Rzehak (2017)	Y
BMI in chidhood	cg21960828	*TADA2B*	+	Rzehak (2017)	Y
BMI in chidhood	cg03012642		-	Rzehak (2017)	Y
BMI in chidhood	ch.8.1234861F	*XKR4*	-	Rzehak (2017)	ch.8.1234861F
BMI in chidhood	cg10861407	*FSD1L*	-	Rzehak (2017)	Y
BMI in chidhood	cg24665113	*DTNB*	+	Rzehak (2017)	Y
BMI in chidhood	cg13298389		+	Rzehak (2017)	Y
BMI in chidhood	cg05945266	*FOXK2*	+	Rzehak (2017)	Y
BMI in chidhood	cg17287326	*AVPI1*	+	Rzehak (2017)	Y
BMI in chidhood	cg19864468	*CHCHD5*	+	Rzehak (2017)	Y

growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
BMI in chidhood	cg00384539	*PRDM14*	+	Rzehak (2017)	Y
BMI in chidhood	cg10802005		+	Rzehak (2017)	Y
BMI in chidhood	cg26627956	*CFLAR*	-	Rzehak (2017)	Y
BMI in chidhood	cg01936370		+	Rzehak (2017)	Y
BMI in chidhood	cg25048701	*FOLR1*	+	Rzehak (2017)	Y
BMI in chidhood	cg18051668		+	Rzehak (2017)	Y
BMI in chidhood	cg12717591	*SFMBT2*	+	Rzehak (2017)	Y
BMI in chidhood	cg11245333	*MOSC1*	+	Rzehak (2017)	Y
BMI in chidhood	cg17276103	*DDAH1*	+	Rzehak (2017)	Y
BMI in chidhood	cg27267258	*ANKRD11*	+	Rzehak (2017)	Y
BMI in chidhood	cg08692210	*WDR51A*	+	Rzehak (2017)	Y
BMI in chidhood	cg14699734		+	Rzehak (2017)	Y
BMI in chidhood	cg15660740		-	Rzehak (2017)	Y
BMI in chidhood	cg00193021	*SNCA*	-	Rzehak (2017)	Y
BMI in chidhood	cg03862225	*CTBP2*	+	Rzehak (2017)	Y
BMI in chidhood	cg11027140	*GPR144*	-	Rzehak (2017)	Y
BMI in chidhood	cg04292672	*MORN1*	+	Rzehak (2017)	Y
BMI in chidhood	cg13345817		+	Rzehak (2017)	Y
BMI in chidhood	cg02259997	*FGF9*	-	Rzehak (2017)	Y
BMI in chidhood	cg07664183	*HAUS6*	+	Rzehak (2017)	Y
BMI in chidhood	cg13931565	*NCDN*	+	Rzehak (2017)	Y
BMI in chidhood	cg14658149		+	Rzehak (2017)	Y
BMI in chidhood	cg13367929	*TSPAN17*	-	Rzehak (2017)	Y
BMI in chidhood	cg27038634		+	Rzehak (2017)	Y

growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
BMI in chidhood	cg10738025	*PBX4*	+	Rzehak (2017)	Y
BMI in chidhood	cg02381820		+	Rzehak (2017)	Y
BMI in chidhood	cg17644856		+	Rzehak (2017)	Y
BMI in chidhood	cg25371332		+	Rzehak (2017)	Y
BMI in chidhood	cg14270612	*ABL1*	+	Rzehak (2017)	Y
BMI in chidhood	cg05030574	*NEAT1*	-	Rzehak (2017)	Y
BMI in chidhood	cg04010122	*SOS1*	+	Rzehak (2017)	Y
BMI in chidhood	cg26360930	*DLC1*	-	Rzehak (2017)	Y
BMI in chidhood	cg13274254	*GULP1*	+	Rzehak (2017)	Y
BMI in chidhood	cg11504355	*KIRREL*	+	Rzehak (2017)	Y
BMI in chidhood	cg25734624	*CNGA3*	+	Rzehak (2017)	Y
BMI in chidhood	cg16658008		-	Rzehak (2017)	Y
BMI in chidhood	cg13282068	*GRB2*	+	Rzehak (2017)	Y
BMI in chidhood	cg06410191	*LVRN*	+	Rzehak (2017)	Y
BMI in chidhood	cg01962821		+	Rzehak (2017)	Y
BMI in chidhood	cg26662102	*AJAP1*	-	Rzehak (2017)	Y
BMI in chidhood	cg01338599	*TM9SF4*	-	Rzehak (2017)	Y
BMI in chidhood	cg24074783		+	Rzehak (2017)	Y
BMI in chidhood	cg08008403	*LHX9*	+	Rzehak (2017)	Y
BMI in chidhood	cg22726039	*SLC30A6*	-	Rzehak (2017)	Y
BMI in chidhood	cg01363869	*TRIM39*	-	Rzehak (2017)	Y
BMI in chidhood	cg00905156	*FAM13A*	-	Rzehak (2017)	Y
BMI in chidhood	cg26393837	*PCSK7*	+	Rzehak (2017)	Y
BMI in chidhood	cg14240060		+	Rzehak (2017)	Y

growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
BMI in chidhood	cg13718870	*BRD3*	-	Rzehak (2017)	Y
BMI in chidhood	cg03174228	*TTLL11*	+	Rzehak (2017)	Y
BMI in chidhood	cg20219100	*MACROD1*	+	Rzehak (2017)	Y
BMI in chidhood	cg14524775	*ARHGEF2*	+	Rzehak (2017)	Y
BMI in chidhood	cg27247382	*PLEKHM3*	+	Rzehak (2017)	Y
BMI in chidhood	cg17111837	*NCRNA00176*	+	Rzehak (2017)	Y
BMI in chidhood	cg08352032	*MARK4*	-	Rzehak (2017)	Y
BMI in chidhood	cg22078907	*USP22*	+	Rzehak (2017)	Y
BMI in chidhood	cg22318872	*GNAL*	+	Rzehak (2017)	Y
BMI in chidhood	cg00119053		+	Rzehak (2017)	Y
BMI in chidhood	cg01548142		-	Rzehak (2017)	Y
BMI in chidhood	cg09936845	*NTRK1*	-	Rzehak (2017)	Y
BMI in chidhood	cg16465882	*IQCD*	+	Rzehak (2017)	Y
BMI in chidhood	cg24067118		+	Rzehak (2017)	Y
BMI in chidhood	cg19992450	*EIF2C2*	+	Rzehak (2017)	Y
BMI in chidhood	cg09204618	*TSSK2*	+	Rzehak (2017)	Y
BMI in chidhood	cg04880874	*CCNL2*	+	Rzehak (2017)	Y
BMI in chidhood	cg23827531	*FAM107A*	+	Rzehak (2017)	Y
BMI in chidhood	cg24673769	*CHCHD6*	-	Rzehak (2017)	Y
BMI in chidhood	cg03520802	*HIPK2*	+	Rzehak (2017)	Y
BMI in chidhood	cg09927637	*KIAA0664*	+	Rzehak (2017)	Y
BMI in chidhood	cg15025089	*TAPBP*	+	Rzehak (2017)	Y
BMI in chidhood	cg02393428		+	Rzehak (2017)	Y
BMI in chidhood	cg22243918		+	Rzehak (2017)	Y

growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
BMI in chidhood	cg09423599	*ABCB9*	-	Rzehak (2017)	Y
BMI in chidhood	cg07831312	*FAM188B*	-	Rzehak (2017)	Y
BMI in chidhood	cg22503047	*BAT2*	-	Rzehak (2017)	Y
BMI in chidhood	cg13153009	*ZC3H3*	-	Rzehak (2017)	Y
BMI in chidhood	cg17167536	*XKR6*	+	Rzehak (2017)	Y
BMI in chidhood	cg10802680	*DIABLO*	+	Rzehak (2017)	Y
BMI in chidhood	cg14834653	*FGFR2*	-	Rzehak (2017)	Y
BMI in chidhood	cg19821475		+	Rzehak (2017)	Y
BMI in chidhood	cg19104196	*SFMBT2*	+	Rzehak (2017)	Y
BMI in chidhood	cg03356833		+	Rzehak (2017)	Y
BMI in chidhood	cg13976853	*ZNF77*	-	Rzehak (2017)	Y
BMI in chidhood	cg25639557	*FURIN*	+	Rzehak (2017)	Y
BMI in chidhood	cg09121516	*TFAP4*	-	Rzehak (2017)	Y
BMI in chidhood	cg24102266	*KIAA1522*	-	Rzehak (2017)	Y
BMI in chidhood	cg02081925	*DOK3*	-	Rzehak (2017)	Y
BMI in chidhood	cg23687103	*MIIP*	+	Rzehak (2017)	Y
BMI in chidhood	cg19358373		+	Rzehak (2017)	Y
BMI in chidhood	cg18699524	*TBCD*	+	Rzehak (2017)	Y
BMI in chidhood	cg19382175	*PDE6A*	-	Rzehak (2017)	Y
BMI in chidhood	cg15914340		+	Rzehak (2017)	Y
BMI in chidhood	cg26879644		+	Rzehak (2017)	Y
BMI in chidhood	cg04556432	*SETD3*	+	Rzehak (2017)	Y
BMI in chidhood	cg24035595	*DLGAP2*	+	Rzehak (2017)	Y
BMI in chidhood	cg02942594	*LOC148189*	-	Rzehak (2017)	Y

growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
BMI in chidhood	cg00577950	*PDXK*	-	Rzehak (2017)	Y
BMI in chidhood	cg09290175		-	Rzehak (2017)	Y
BMI in chidhood	cg19922435	*LOC285419*	+	Rzehak (2017)	Y
BMI in chidhood	cg21678312	*PTPRR*	+	Rzehak (2017)	Y
BMI in chidhood	cg23483886	*LPCAT1*	-	Rzehak (2017)	Y
BMI in chidhood	cg10822034		-	Rzehak (2017)	Y
BMI in chidhood	cg02695873		+	Rzehak (2017)	Y
BMI in chidhood	cg11276053	*RSPH6A*	-	Rzehak (2017)	Y
BMI in chidhood	cg19615711		+	Rzehak (2017)	Y
BMI in chidhood	cg21584983	*ECSIT*	-	Rzehak (2017)	Y
BMI in chidhood	cg01806722	*MAML2*	+	Rzehak (2017)	Y
BMI in chidhood	cg14085500	*TRIM32;ASTN 2*	+	Rzehak (2017)	Y
BMI in chidhood	cg02518775	*CNIH2*	+	Rzehak (2017)	Y
BMI in chidhood	cg16025035	*TFDP1*	-	Rzehak (2017)	Y
BMI in chidhood	cg17459290	*LGALS8*	+	Rzehak (2017)	Y
BMI in chidhood	cg01565985	*CHD5*	+	Rzehak (2017)	Y
BMI in chidhood	cg18321548	*ALX4*	+	Rzehak (2017)	Y
BMI in chidhood	cg14002226	*KCNT1*	-	Rzehak (2017)	Y
BMI in chidhood	cg19419146	*STK32A*	+	Rzehak (2017)	Y
BMI in chidhood	cg25554998		-	Rzehak (2017)	Y
BMI in chidhood	cg16174341	*PFDN5*	+	Rzehak (2017)	Y
BMI in chidhood	cg04219544	*KRT24*	-	Rzehak (2017)	Y
BMI in chidhood	cg02431597	*CACNA1E*	+	Rzehak (2017)	Y
BMI in chidhood	cg13023210	*PAX6*	+	Rzehak (2017)	Y

growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
BMI in chidhood	cg14398659	*INTS4*	-	Rzehak (2017)	Y
BMI in chidhood	cg20380470		+	Rzehak (2017)	Y
BMI in chidhood	cg02283735	*PTPRF*	+	Rzehak (2017)	Y
BMI in chidhood	cg11434973	*GUCY2D*	-	Rzehak (2017)	Y
BMI in chidhood	cg15423797	*MED16*	-	Rzehak (2017)	Y
BMI in chidhood	cg21209395	*LOC158381*	+	Rzehak (2017)	Y
BMI in chidhood	cg05895868		+	Rzehak (2017)	Y
BMI in chidhood	cg05931272		-	Rzehak (2017)	Y
BMI in chidhood	cg04094751	*HSPA12B*	+	Rzehak (2017)	Y
BMI in chidhood	cg24824703	*GNA12*	+	Rzehak (2017)	Y
BMI in chidhood	cg17936495	*SCAMP3;CLK2*	-	Rzehak (2017)	Y
BMI in chidhood	cg01624768		+	Rzehak (2017)	Y
BMI in chidhood	cg00049616	*CLYBL*	+	Rzehak (2017)	Y
BMI in chidhood	cg03904209		+	Rzehak (2017)	Y
BMI in chidhood	cg14282798	*PRDM15*	+	Rzehak (2017)	Y
BMI in chidhood	cg02162339	*ING5*	+	Rzehak (2017)	Y
BMI in chidhood	cg11759917	*HGS*	-	Rzehak (2017)	Y
BMI in chidhood	cg02752882	*KCNH1*	+	Rzehak (2017)	Y
BMI in chidhood	cg21129976	*EPHB4*	-	Rzehak (2017)	Y
BMI in chidhood	cg26823535		-	Rzehak (2017)	Y
BMI in chidhood	cg11841546	*MEGF11*	-	Rzehak (2017)	Y
BMI in chidhood	cg19196862	*IGF2R*	-	Rzehak (2017)	Y
BMI in chidhood	cg17491360		-	Rzehak (2017)	Y
BMI in chidhood	cg04305494	*LRFN1*	-	Rzehak (2017)	Y

growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
BMI in chidhood	cg01360054	*PRDM6*	+	Rzehak (2017)	Y
BMI in chidhood	cg09034753	*C2orf85*	+	Rzehak (2017)	Y
BMI in chidhood	cg08911048		+	Rzehak (2017)	Y
BMI in chidhood	cg16777782	*CDH13*	-	Rzehak (2017)	Y
BMI in chidhood	cg18022926	*PBLD;HNRNPH3*	+	Rzehak (2017)	Y
BMI in chidhood	cg08552167	*INPP5A*	-	Rzehak (2017)	Y
BMI in chidhood	cg04847478		+	Rzehak (2017)	Y
BMI in chidhood	cg10161743	*GRIN2D*	+	Rzehak (2017)	Y
BMI in chidhood	cg02256198	*C4orf22*	+	Rzehak (2017)	Y
BMI in chidhood	cg04501840	*C1orf95*	+	Rzehak (2017)	Y
BMI in chidhood	cg12921275	*MRPL23*	+	Rzehak (2017)	Y
BMI in chidhood	cg25535435	*TREML4*	-	Rzehak (2017)	Y
BMI in chidhood	cg13008301	*CSTF1*	-	Rzehak (2017)	Y
BMI in chidhood	cg06843240		-	Rzehak (2017)	Y
BMI in chidhood	cg17481047		+	Rzehak (2017)	Y
BMI in chidhood	cg10108042	*MAPK8IP3*	+	Rzehak (2017)	Y
BMI in chidhood	cg27583815	*TRIM15*	-	Rzehak (2017)	Y
BMI in chidhood	cg19274401	*PEPD*	+	Rzehak (2017)	Y
BMI in chidhood	cg00852675	*LOC728743*	+	Rzehak (2017)	Y
obesity in childhood	cg46801562	*HIF3A*	+	Lee (2017)	N
obesity in childhood	cg16672562	*HIF3A*	+	Lee (2017)	Y
BMI in chidhood	gene region	*H19*	+	Bowman (2019)	
Height in children	cg11908057	*HOXA4*	+	Muurinen (2017)	Y
Height in children	cg04317399	*HOXA5*	+	Muurinen (2017)	Y
Height in children	cg19142026	*HOXA6*	+	Muurinen (2017)	Y

growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
Height in children	cg04321618	*HOXA7*	+	Muurinen (2017)	Y
Height in children	cg14359292	*HOXA8*	+	Muurinen (2017)	Y
Height in children	cg25952581	*HOXA9*	+	Muurinen (2017)	Y
BMI in childhood	cg04413090	SBK1	+	Chen (2017)	Y
BMI in childhood	cg12140144	*PRDM16*	+	Chen (2017)	Y
BMI in childhood	cg10772621		+	Chen (2017)	Y
BMI in childhood	cg07464358		+	Chen (2017)	Y
BMI in adolescence	gene region	*TFAM*	-	Gemma (2010)	
adiposity in childhood	chr7:150315553+	eNOS	+	Godfrey (2011)	N
adiposity in childhood	chr9:136355885+	RXRA	+	Godfrey (2011)	N
fat in childhood and adolescence	gene region	PPARGC1A	+	Clarke-Harris (2014	
BMI at 9 years	CpG sites	CASP10_P2	-	Relton (2012)	N
BMI at 9 years	CpG sites	CDKN1C_P2	+	Relton (2012)	N
BMI at 9 years	CpG sites	EPHA1_P	+	Relton (2012)	N
Fat mass at 9 years	CpG sites	CDKN1C_P2	+	Relton (2012)	N
Fat mass at 9 years	CpG sites	EPHA1_P	+	Relton (2012)	N
Fat mass at 9 years	CpG sites	HLA_DOB3	-	Relton (2012)	N
Fat mass at 9 years	CpG sites	NID_1P	-	Relton (2012)	N
Lean mass at 9 years	CpG sites	CDKN1C_P2	+	Relton (2012)	N
Lean mass at 9 years	CpG sites	MMP9_P	+	Relton (2012)	N
Lean mass at 9 years	CpG sites	MPL_P	+	Relton (2012)	N
height at 9 years	CpG sites	ALPL_P	-	Relton (2012)	N
height at 9 years	CpG sites	IRF5_P	-	Relton (2012)	N
height at 9 years	CpG sites	IRF5_E	-	Relton (2012)	N
weight at 12 and 18 months	cg27146050	HIF3A	+	Pan (2015)	Y
weight at 12 and 18 months	cg16672562	HIF3A	+	Pan (2015)	Y
BMI at 12 and 18 months	cg22891070	HIF3A	+	Pan (2015)	Y
childhood obesity	Chr9:130222317	CERCAM	-	Almen (2012)	N
childhood obesity	Chr1:98158797	DPYD	-	Almen (2012)	N
childhood obesity	Chr3:161189517	IL12A	+	Almen (2012)	N
childhood obesity	Chr3:446650023	ZNF35	-	Almen (2012)	N
childhood obesity	Chr1:33494712	ZNF362	-	Almen (2012)	N
childhood obesity	Chr3:151609841	TSC22D2	+	Almen (2012)	N

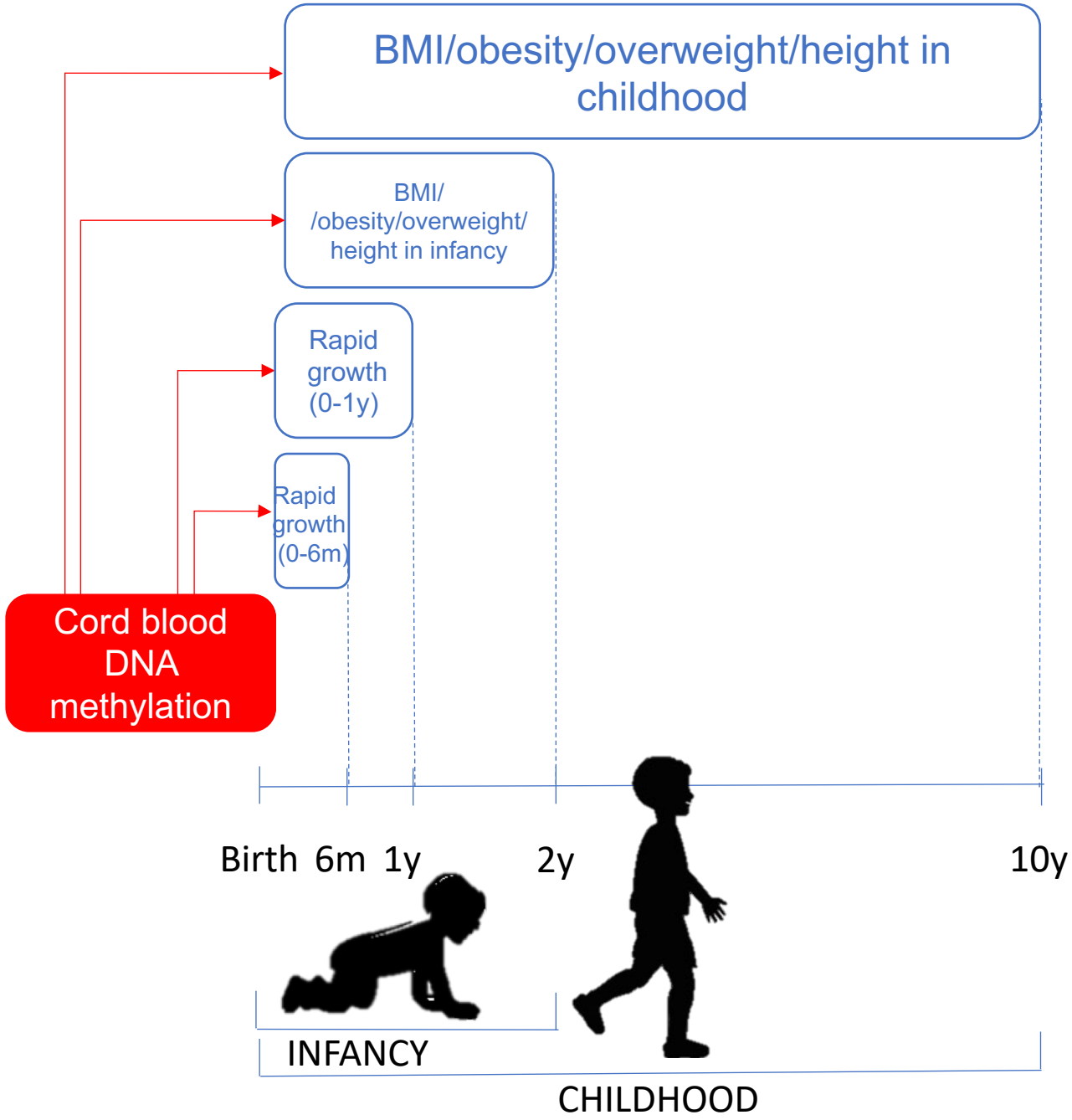
growth outcome	CpG	closest gene	direction of association	author (year)	CpG in 450k?
childhood obesity	Chr22:37598108	CBX6	-	Almen (2012)	N
childhood obesity	Chr16:85101840	FOXF1	-	Almen (2012)	N
childhood obesity	Chr16:72888397	PSMD7	-	Almen (2012)	N
childhood obesity	Chr3:130517349	H1FX	+	Almen (2012)	N
childhood obesity	Chr1:169721292	PRRC2C	-	Almen (2012)	N
childhood obesity	Chr12:119292001	MSI1	-	Almen (2012)	N
childhood obesity	Chr13:109758453	COL4A1	+	Almen (2012)	N
childhood obesity	Chr1:216393199	NBPF3	-	Almen (2012)	N
childhood obesity	Chr12:683135	USP5	-	Almen (2012)	N
childhood obesity	Chr3:147362332	PLOD2	-	Almen (2012)	N
childhood obesity	Chr15:68177287	TLE3	-	Almen (2012)	N
childhood obesity	Chr10:79463659	RPS24	-	Almen (2012)	N
childhood obesity	Chr3:185356123	DVL3	+	Almen (2012)	N
childhood obesity	Chr11:73981434	POLD3	+	Almen (2012)	N
childhood height	promotor	IGF1	+	Ouni (2016)	
childhood height	promotor	IGF1	-	Ouni (2015)	
childhood obesity	promotor	DNMT3B	-	Gardner (2015)	
postnatal growth (3 months)	gene region	TACSTD2	-	Groom (2012)	

Addendum 2

Identification of an epigenetic signature of childhood obesity
Supplementary results

Science and Technology in childhood Obesity Policy

Aim


BMI/obesity/overweight/height in childhood
BMI/ /obesity/overweight/ height in infancy
Rapid growth (0-1y)
Rapid growth (0-6m)
Cord blood DNA methylation
Birth
6m
1y
2y
10y
INFANCY
CHILDHOOD

Descriptive statistics:

Growth outcomes

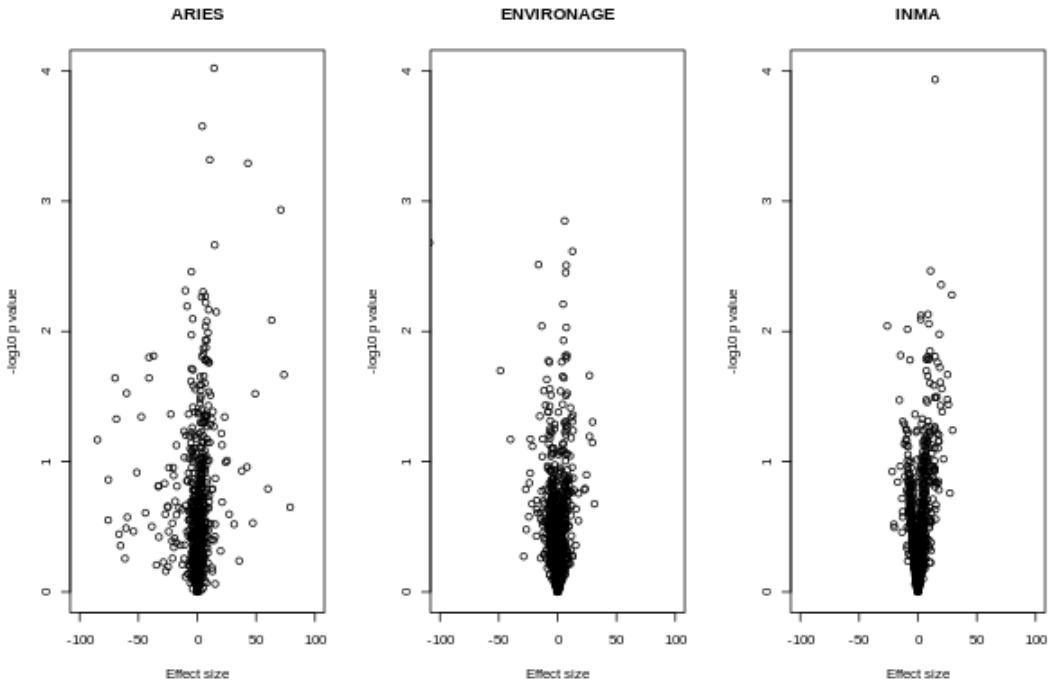
INFANCY (0-2y)

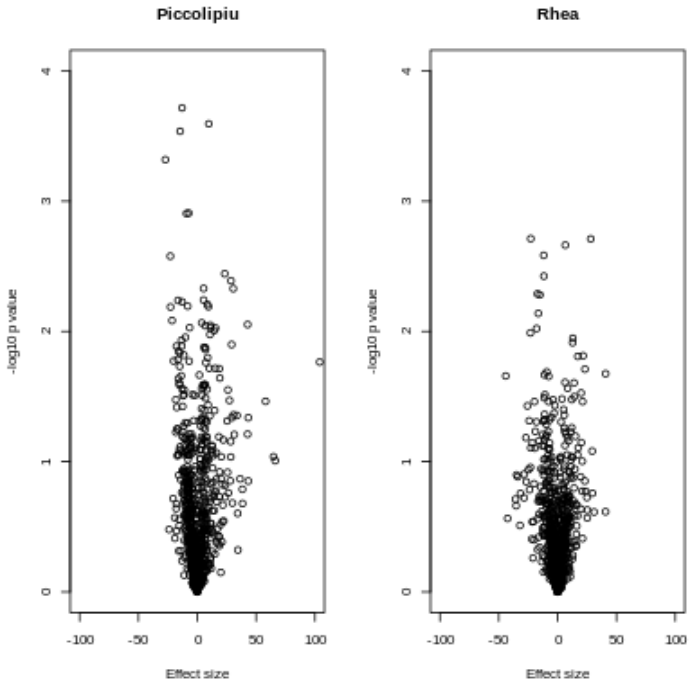
	N id	N obs	Mean N repeated obs	Age range	Mean BMI	N (%) Obese	N (%) Over weight	N id	N obs	Mean N repeated obs	Age range	Mean height
ALSPAC	122	427	3.50	105-595	17.64	6 (1%)	34 (8%)	122	427	3.50	105-595	73.67
ENVIRONAGE	111	992	8.93	8-728	16.51	0	28 (3%)	111	9942	8.95	8-728	66.03
INMA	84	310	3.69	4-729	16.13	0	5 (2%)	84	310	3.69	4-729	65.82
Piccolipiù	88	603	6.85	12-728	16.22	1 (1%)	18 (3%)	89	608	6.83	12-728	68.94
RHEA	92	1227	13.33	3-728	16.16	6 (1%)	40 (3%)	92	1231	13.38	3-728	70.39

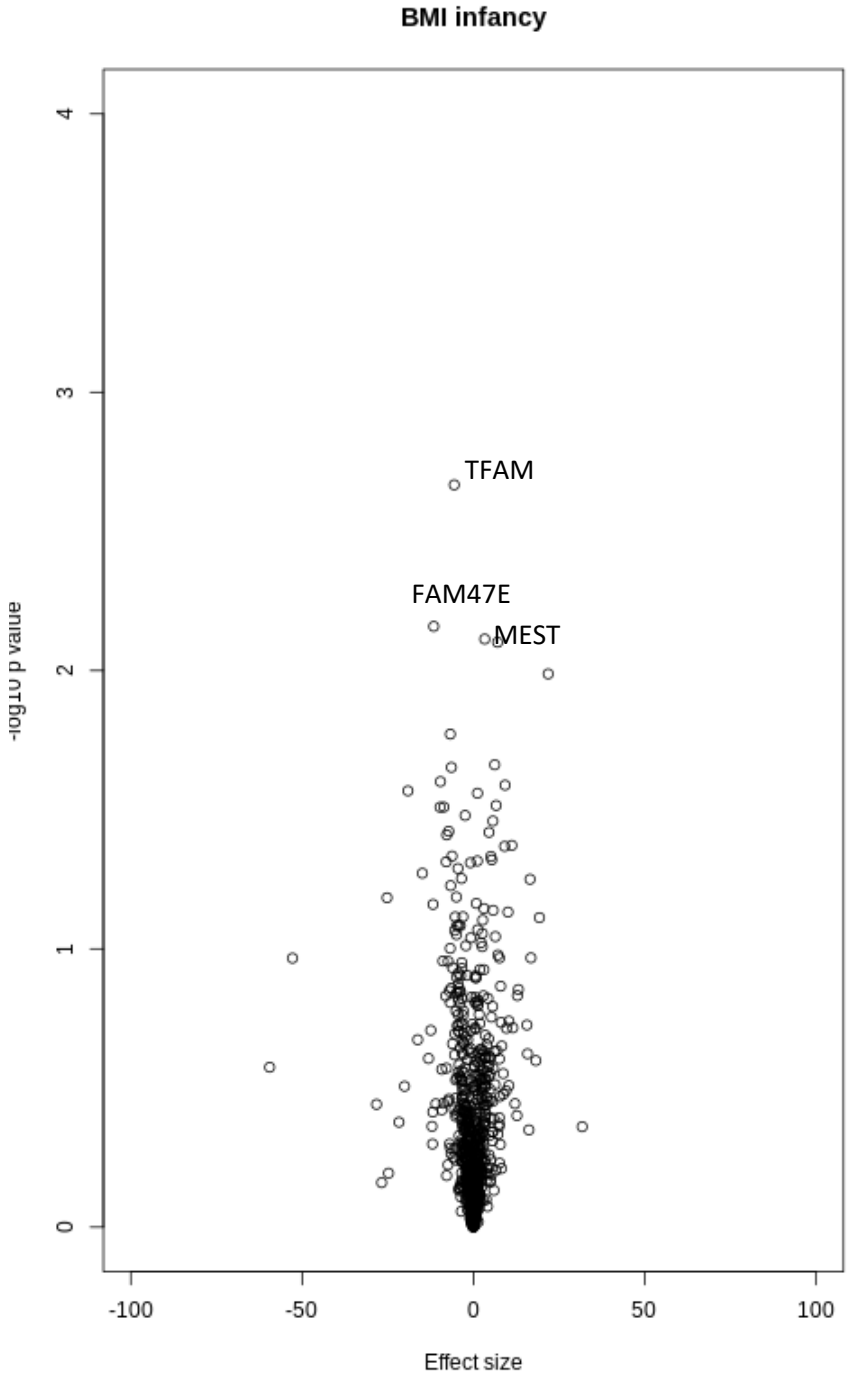
Results

INFANCY (0-2y)

- BMI





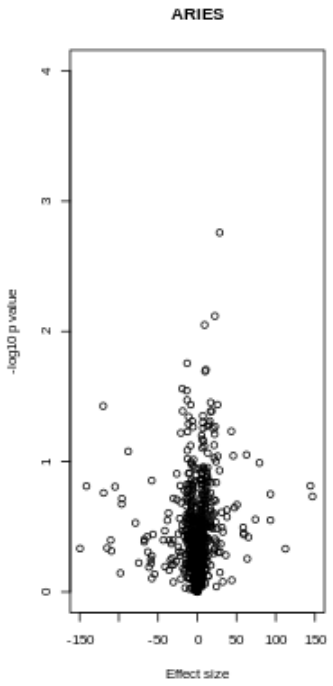


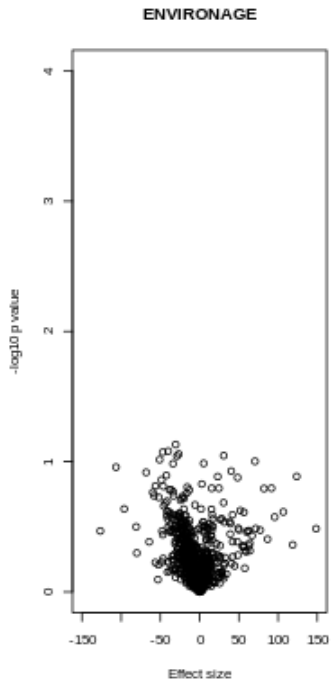
Results

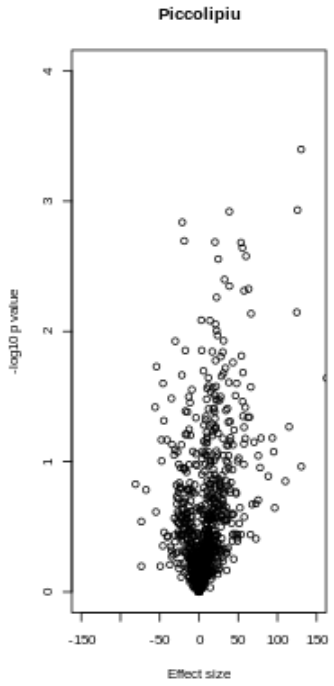
INFANCY

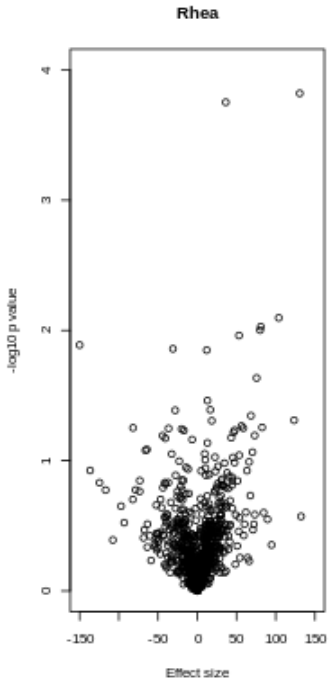
(0-2y)

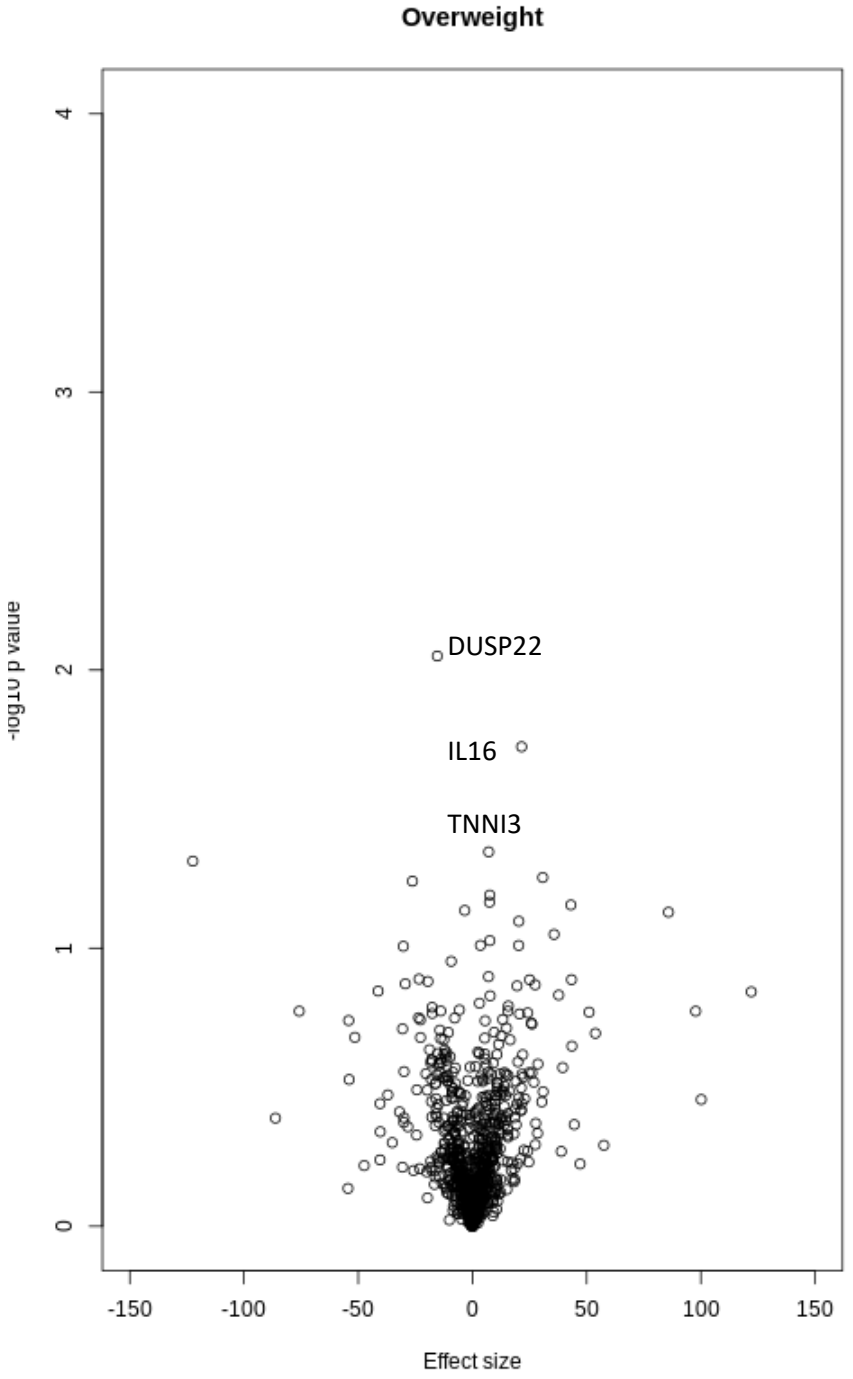
- Overweight







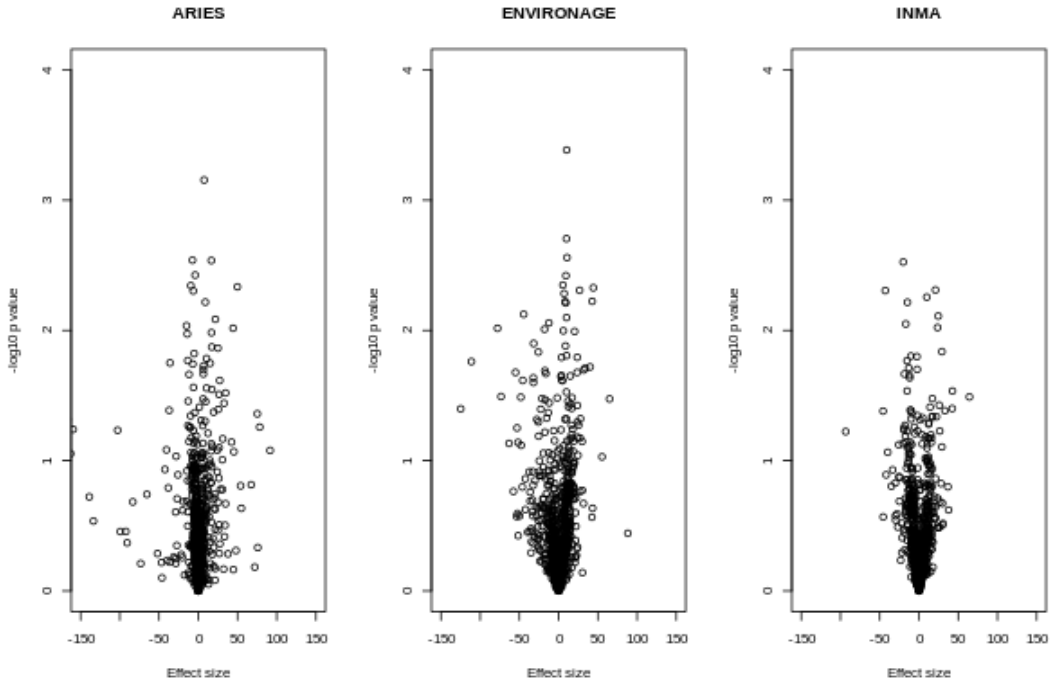


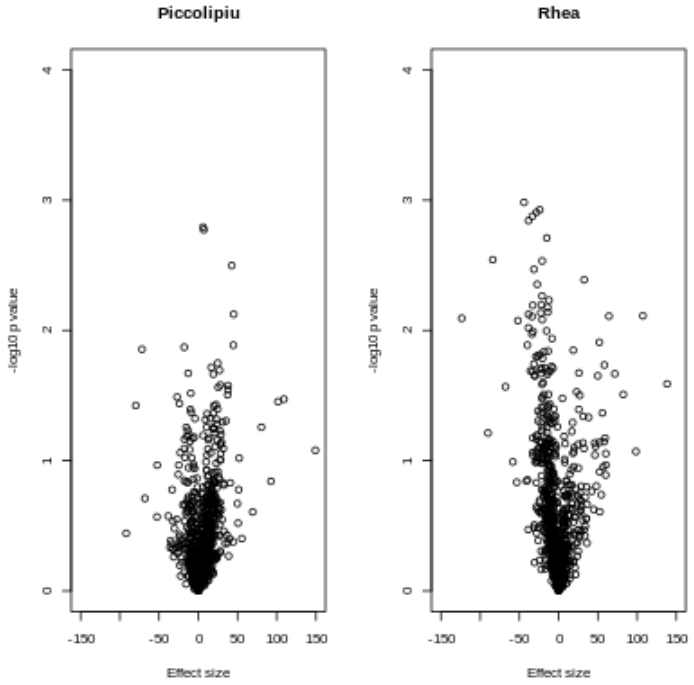


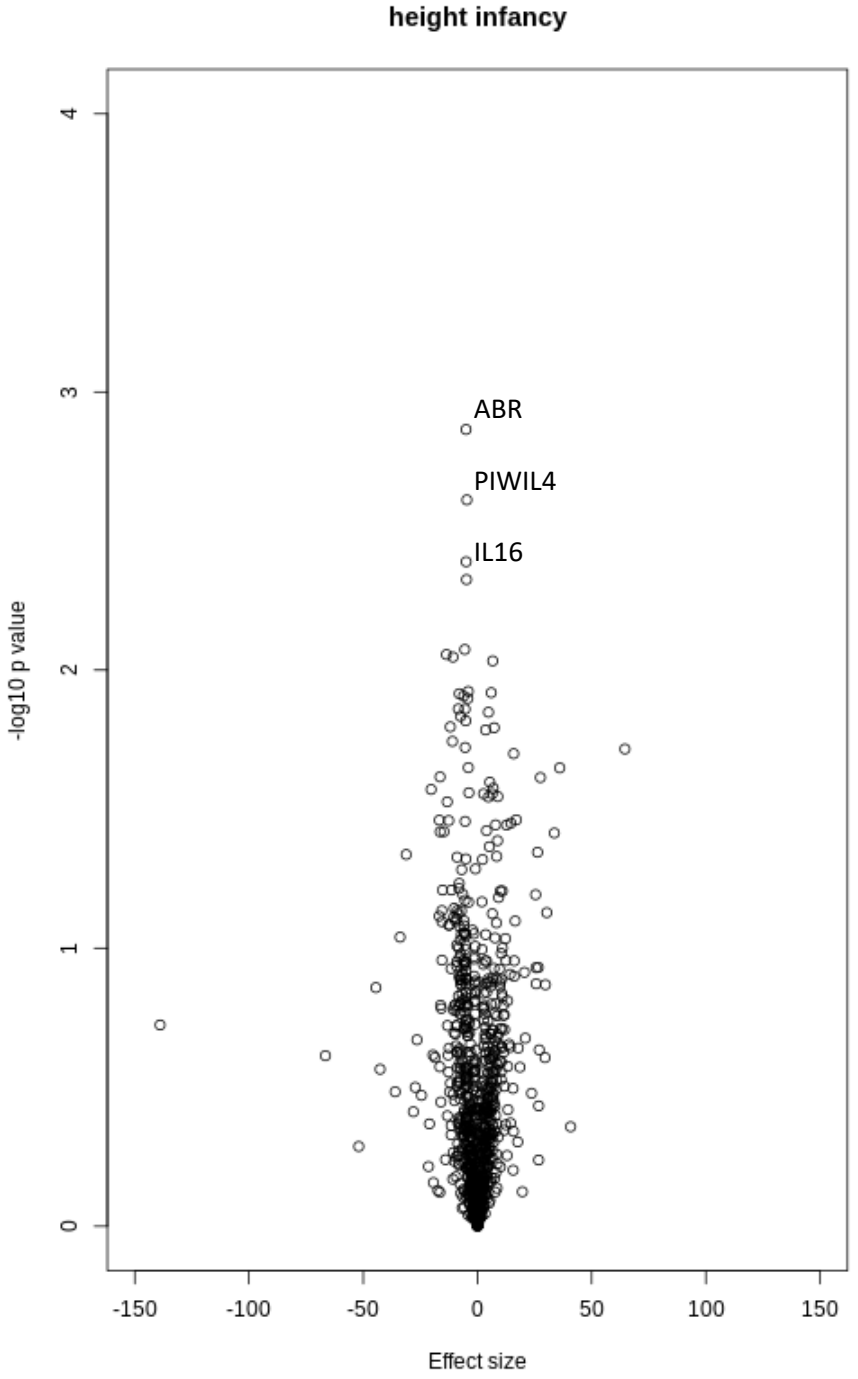
Results

INFANCY (0-2y)

- Height





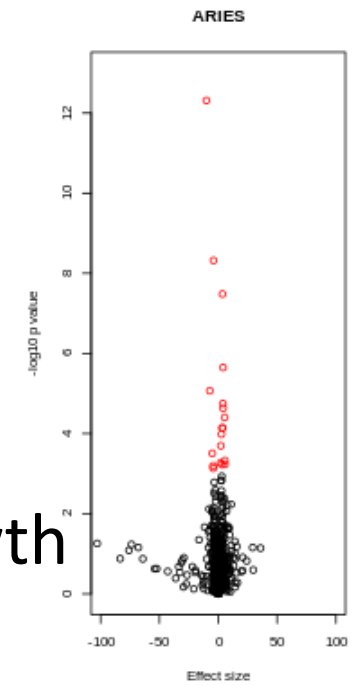


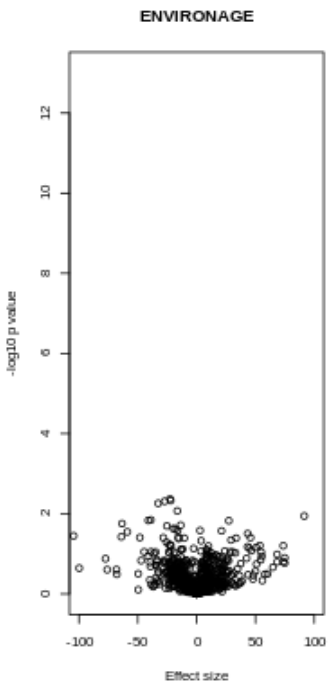
Results

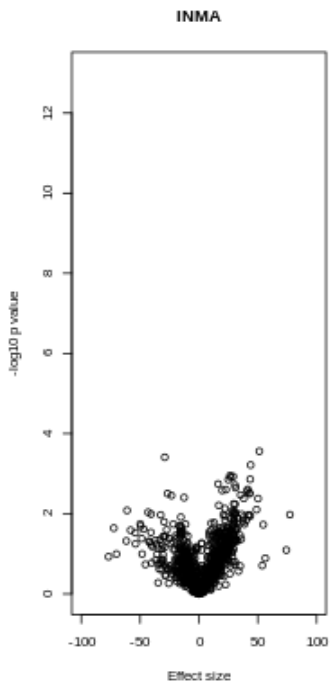
INFANCY

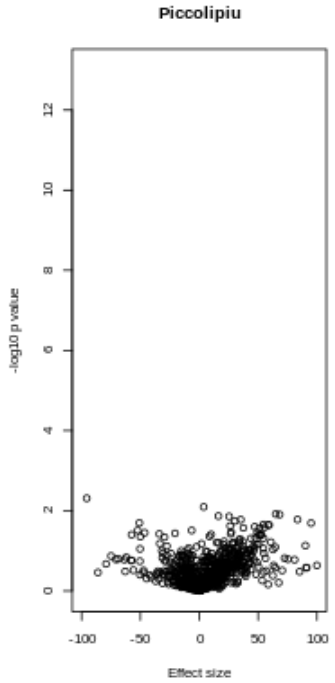
(0-2y)

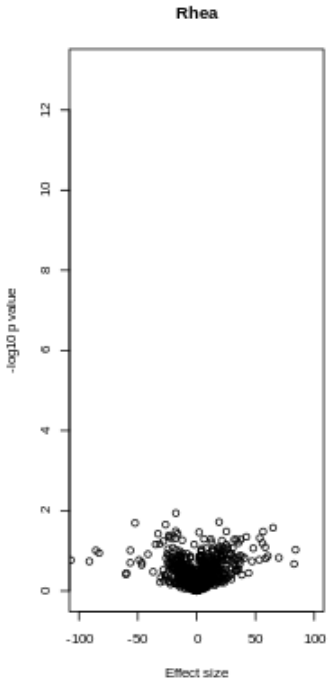
- Rapid growth (0-6m)

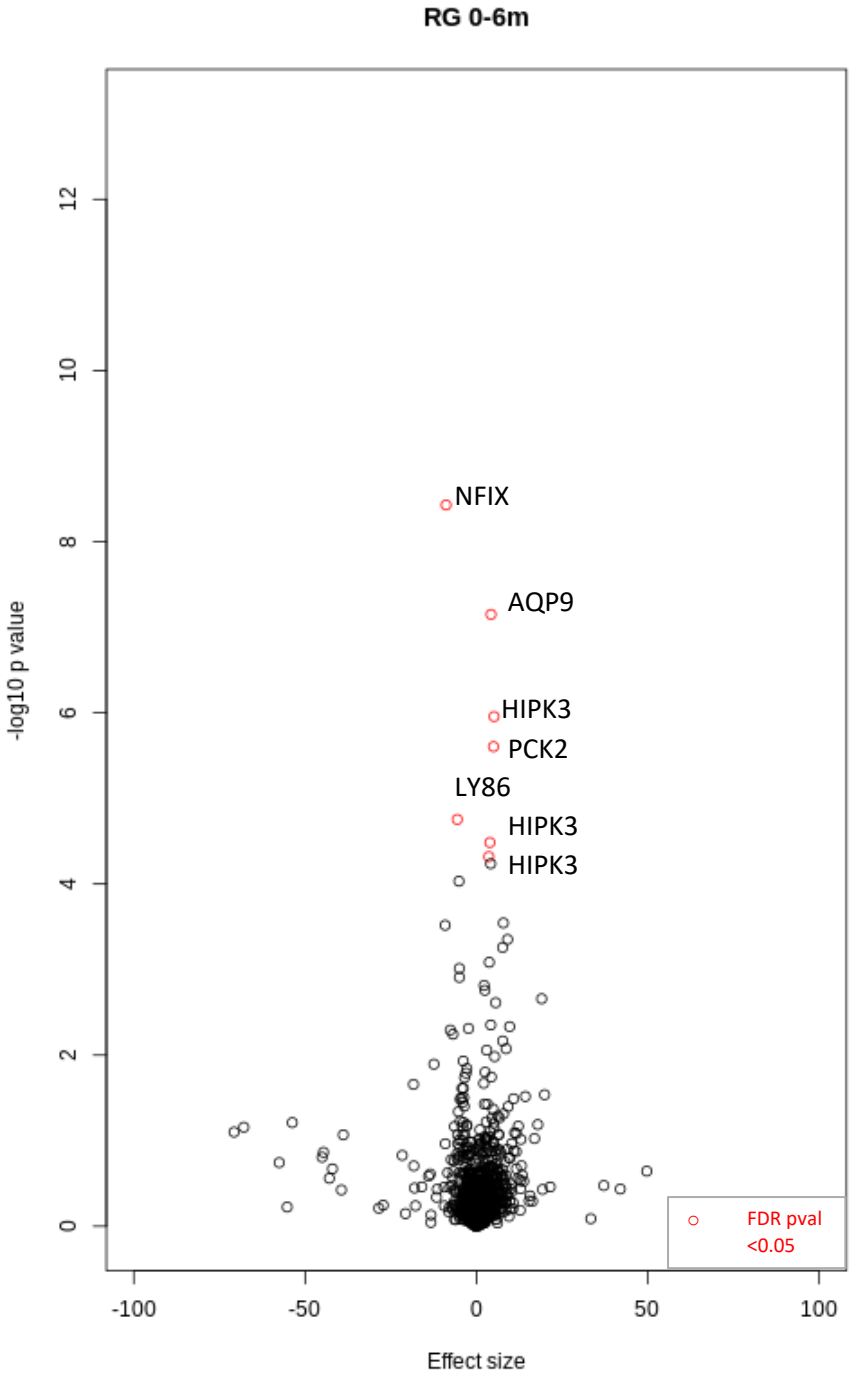










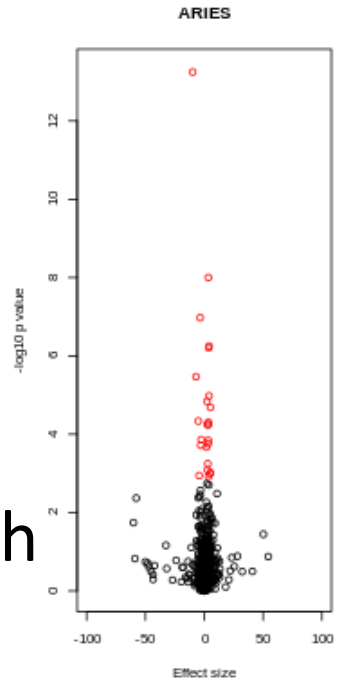


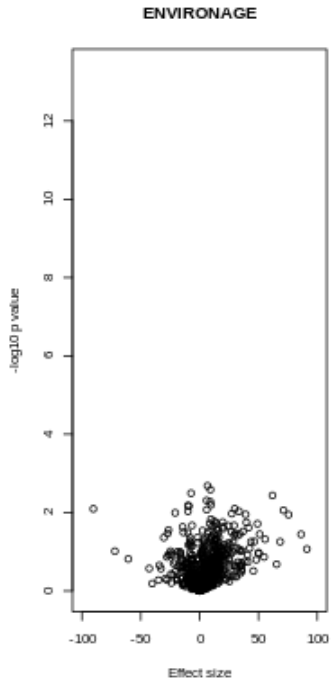
Results

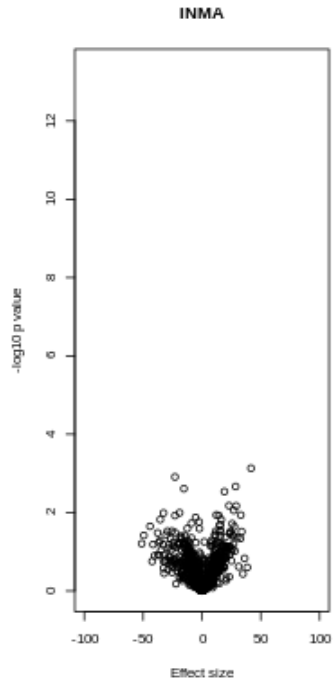
INFANCY

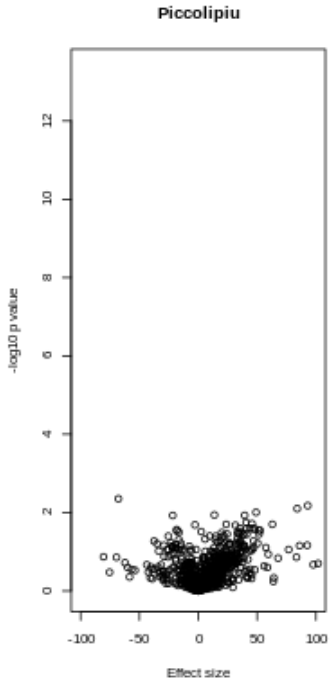
(0-2y)

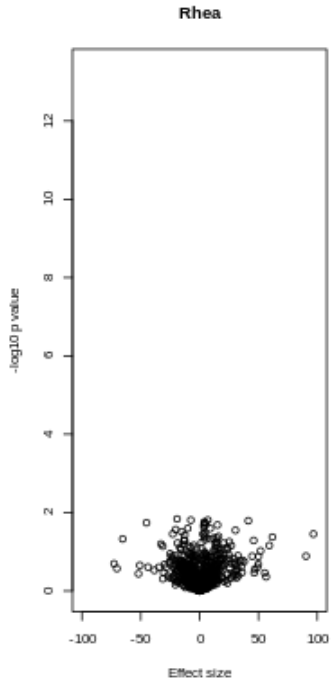
- Rapid growth (0-1y)

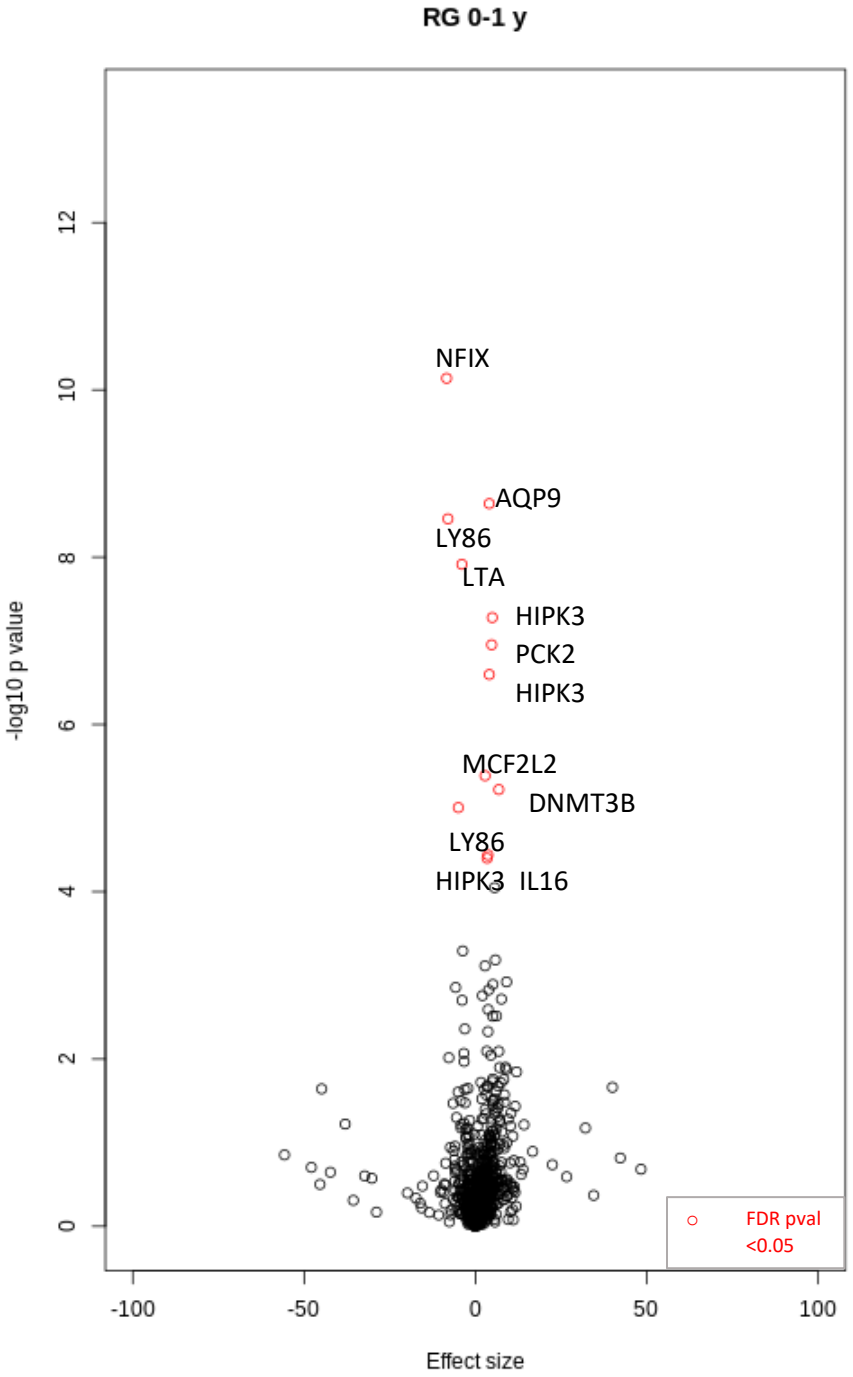












Descriptive statistics:

Growth outcomes

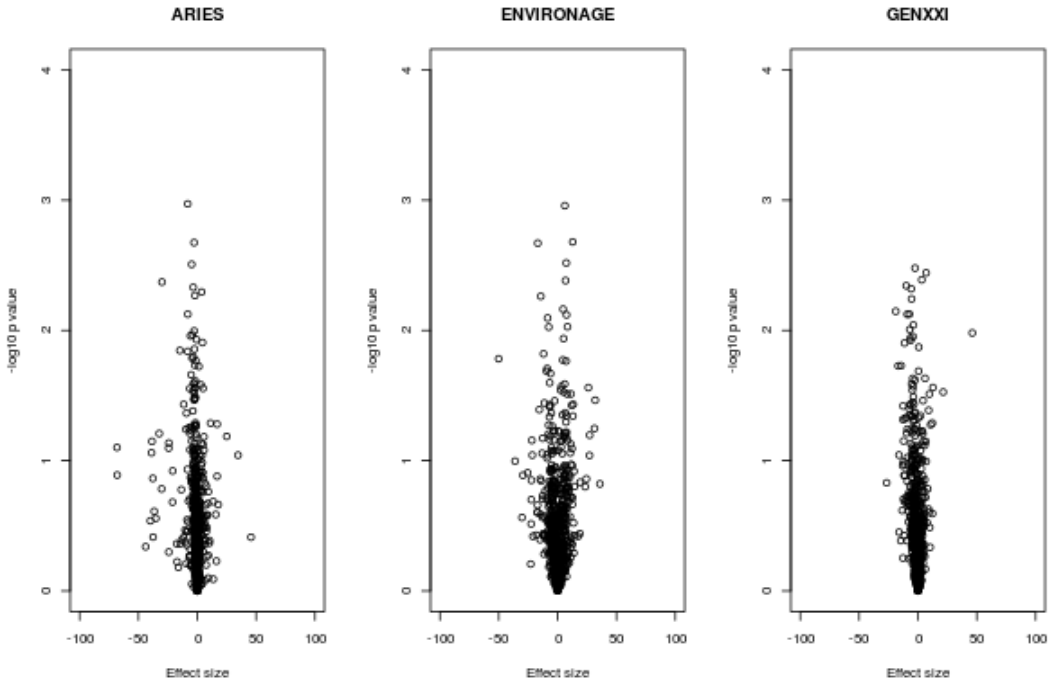
CHILDHOOD (0-10y)

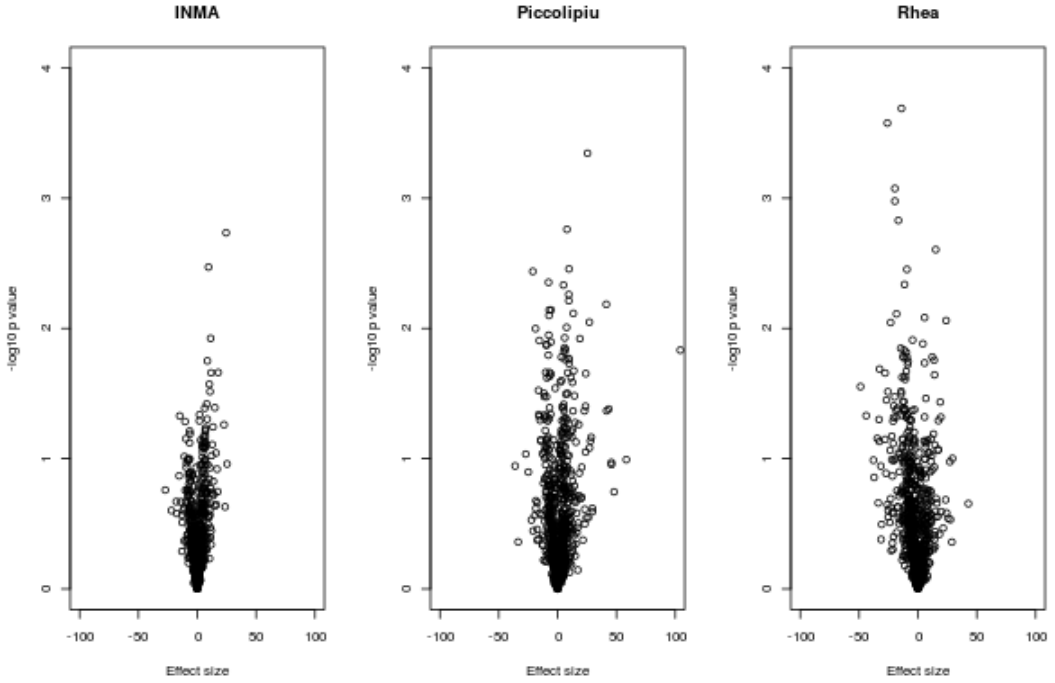
	N id	N obs	Mean N repeated obs	Age range	Mean BMI	N (%) Obese	N (%) Over weight	N id	N obs	Mean N repeated obs	Age range	Mean height
ALSPAC	854	1940	2.27	105-3318	16.68	79 (4%)	287 (15%)	854	1940	2.27	105-3318	105.34
ENVIRONAGE	111	1025	9.23	8-1505	16.50	0	30 (3%)	111	1027	9.25	8-1505	66.92
Generation XXI	706	1322	1.87	1430-2809	16.61	116 (9%)	302 (23%)	706	1322	1.87	1430-2809	114.65
INMA	84	661	7.86	4-2759	16.24	23 (3%)	58 (9%)	84	661	7.86	4-2759	87.19
Piccolipiù	97	741	7.63	12-1732	16.20	4 (1%)	23 (3%)	97	746	7.69	12-1732	74.30
RHEA	92	1731	18.81	3-2781	16.19	30 (2%)	134 (8%)	92	1739	18.90	3-2781	80.29

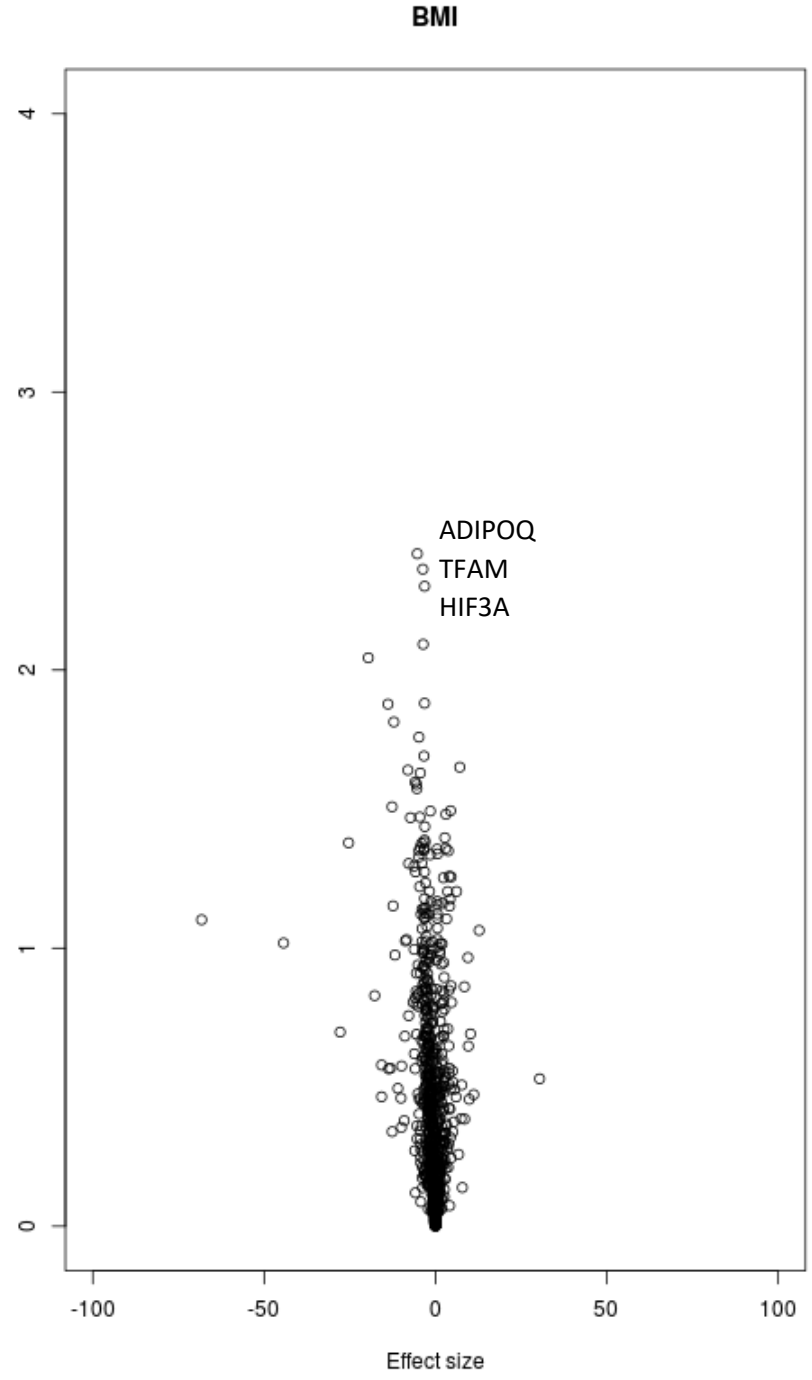
Results

CHILDHOOD (0-10y)

- BMI



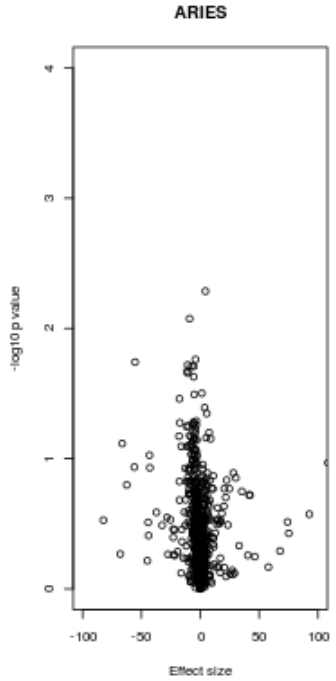


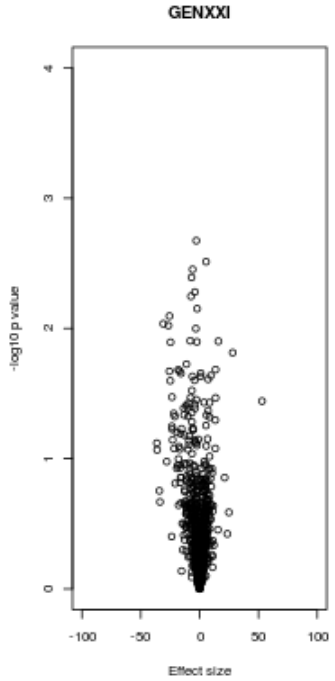


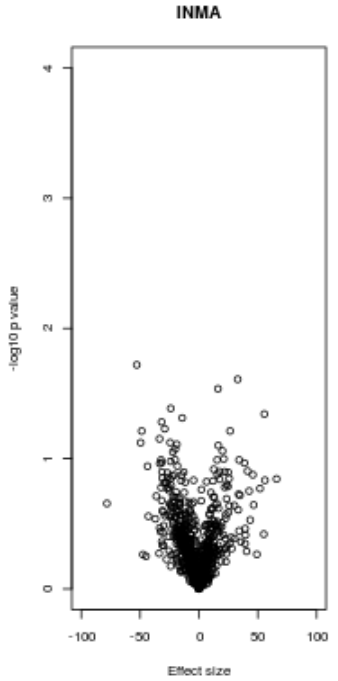
Results

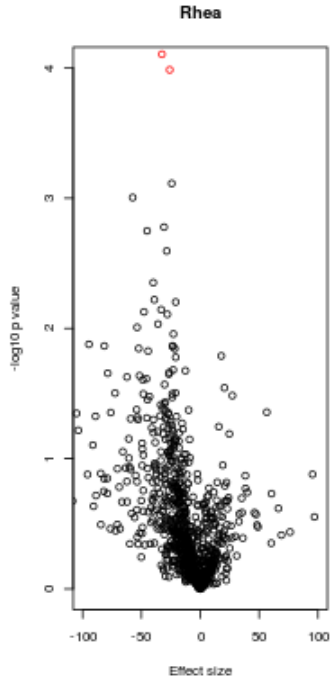
CHILDHOOD (0-10y)

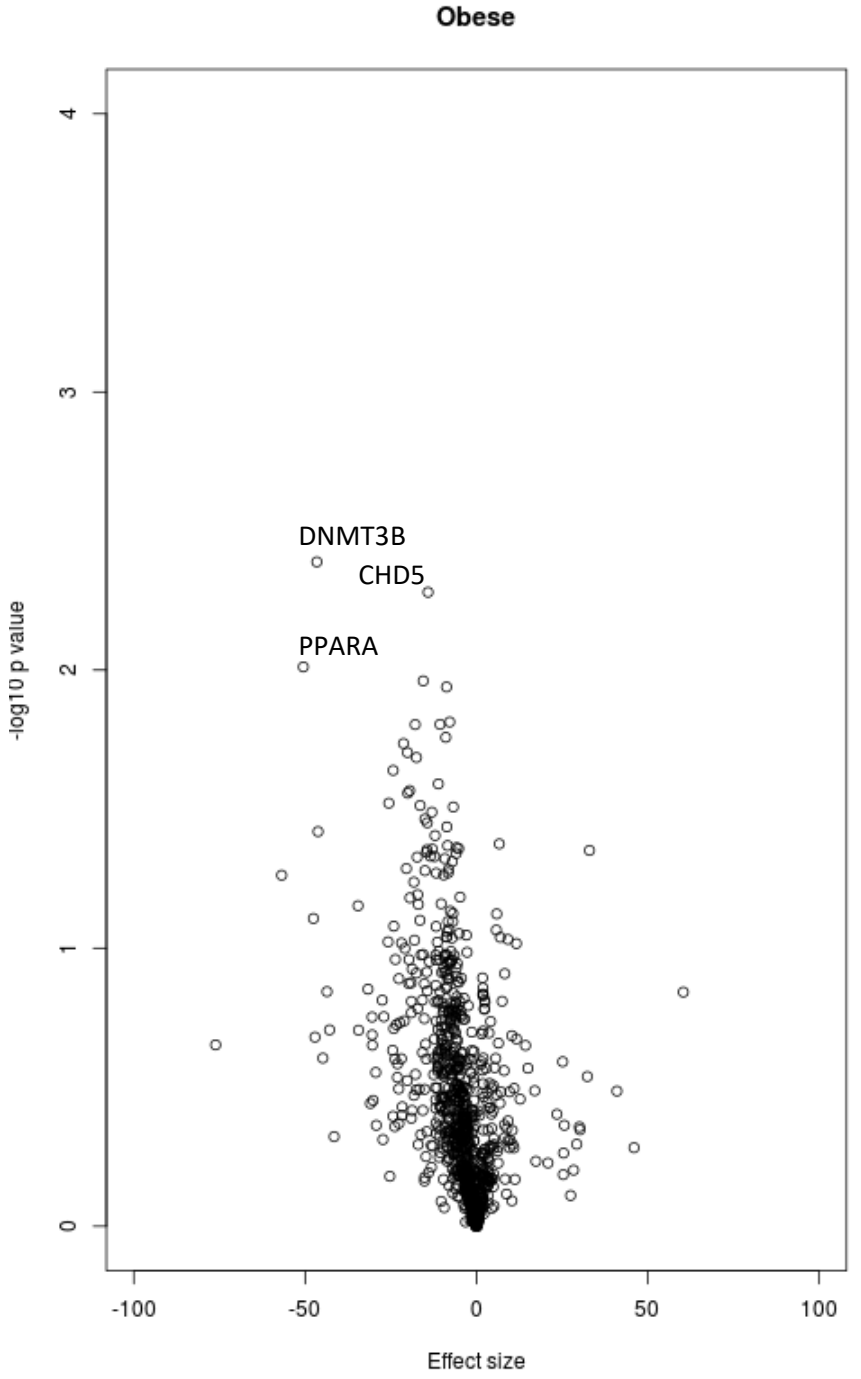
- Obesity







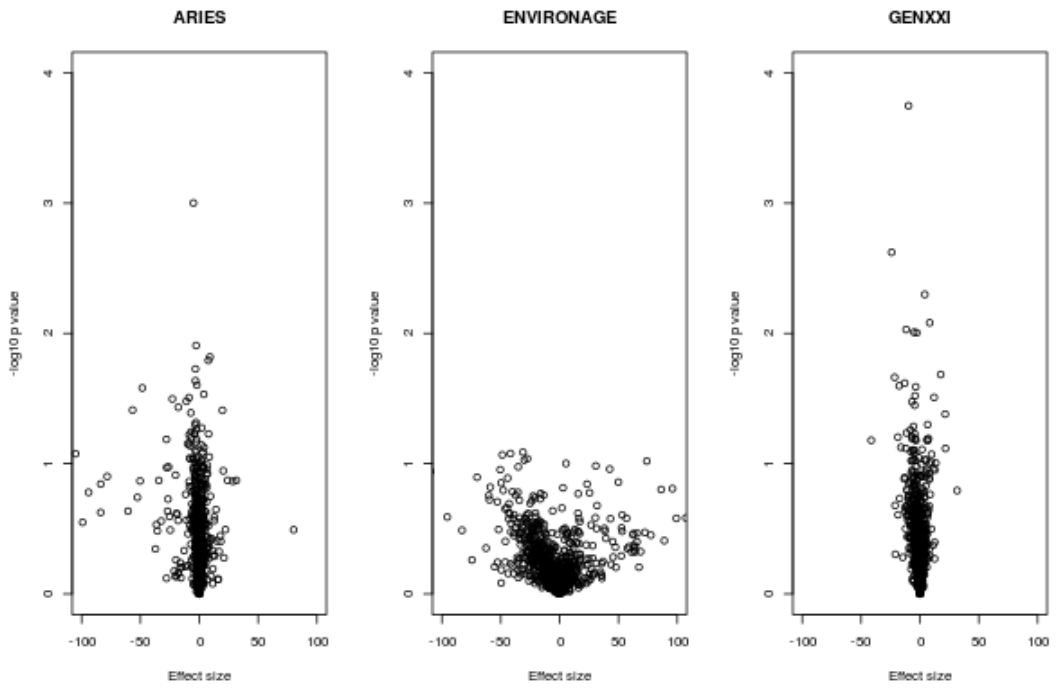


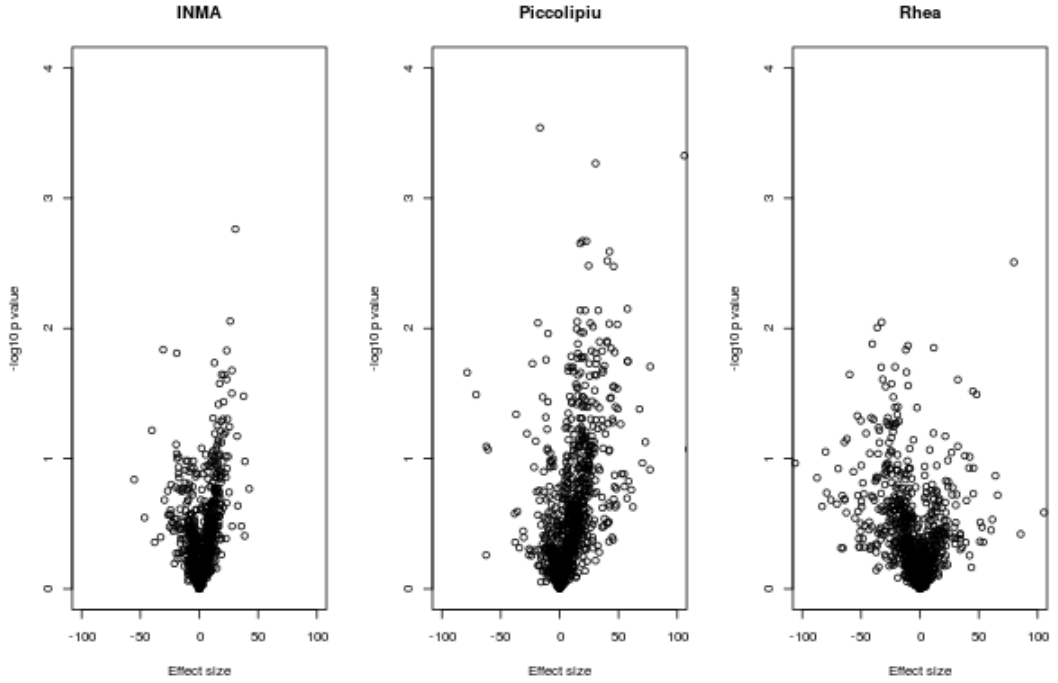


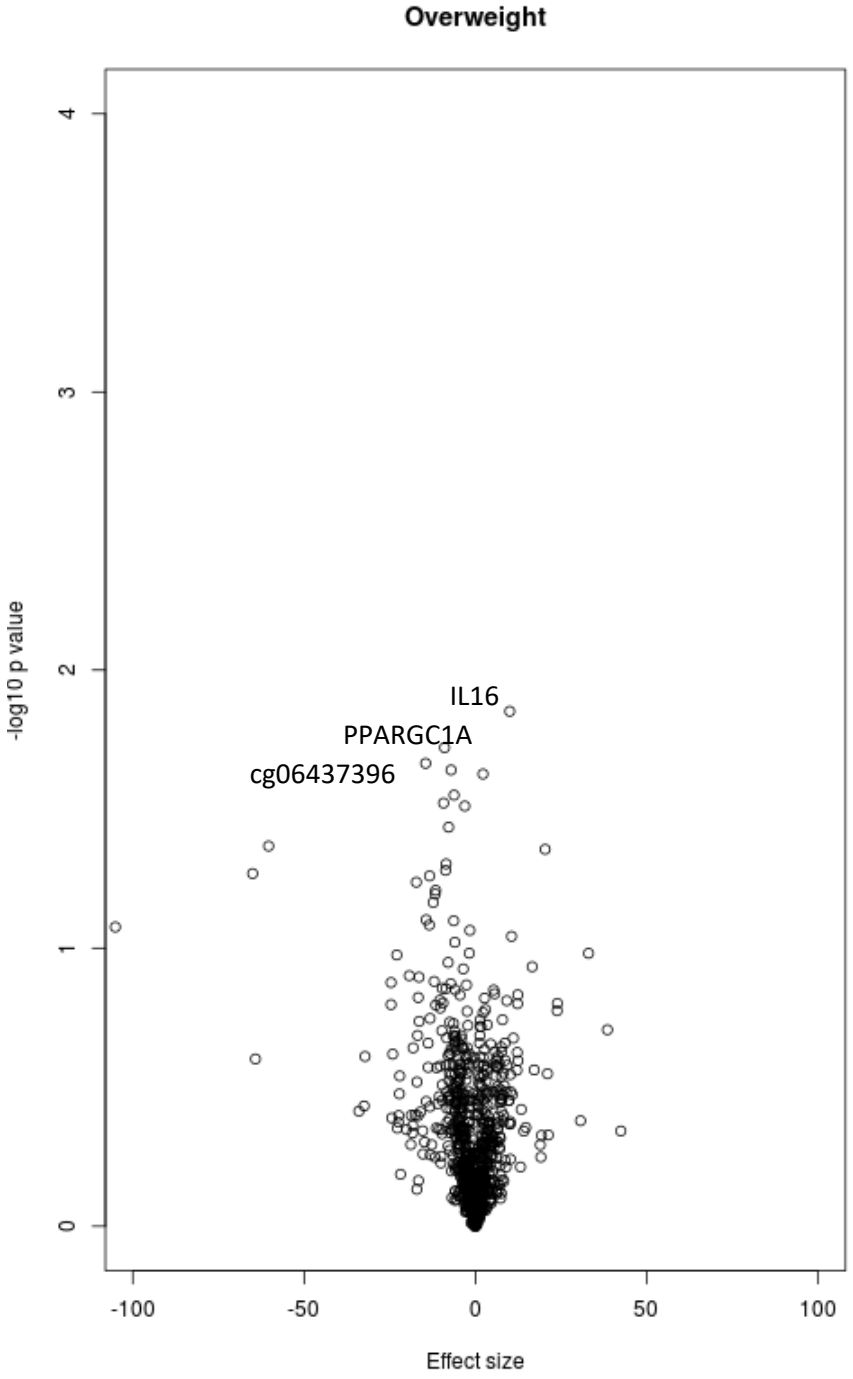
Results

CHILDHOOD (0-10y)

- Overweight



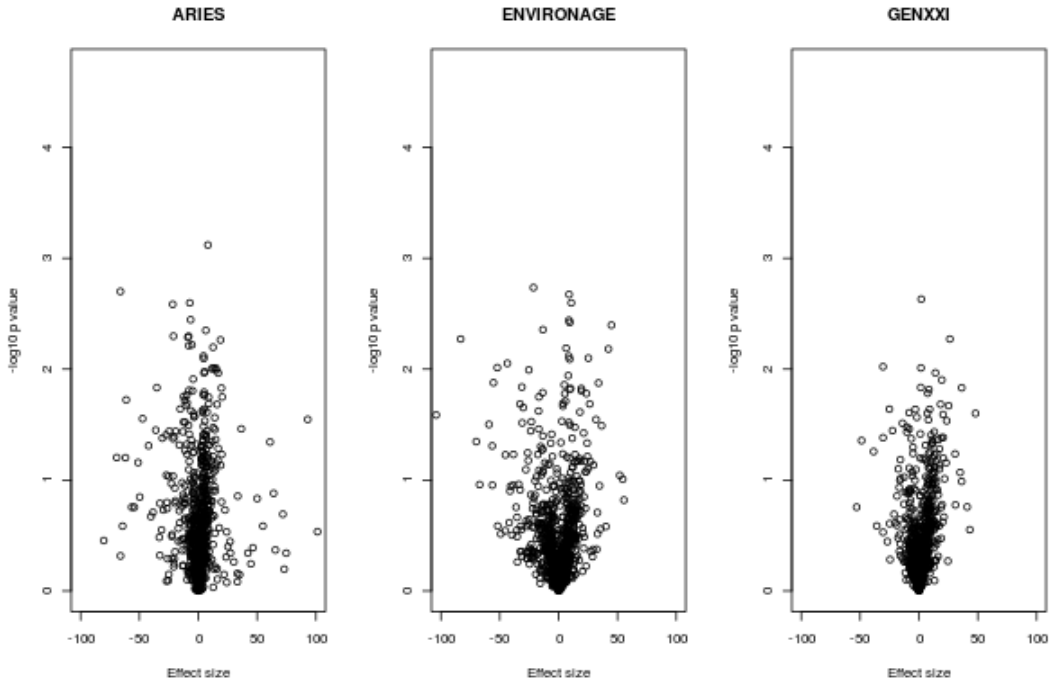


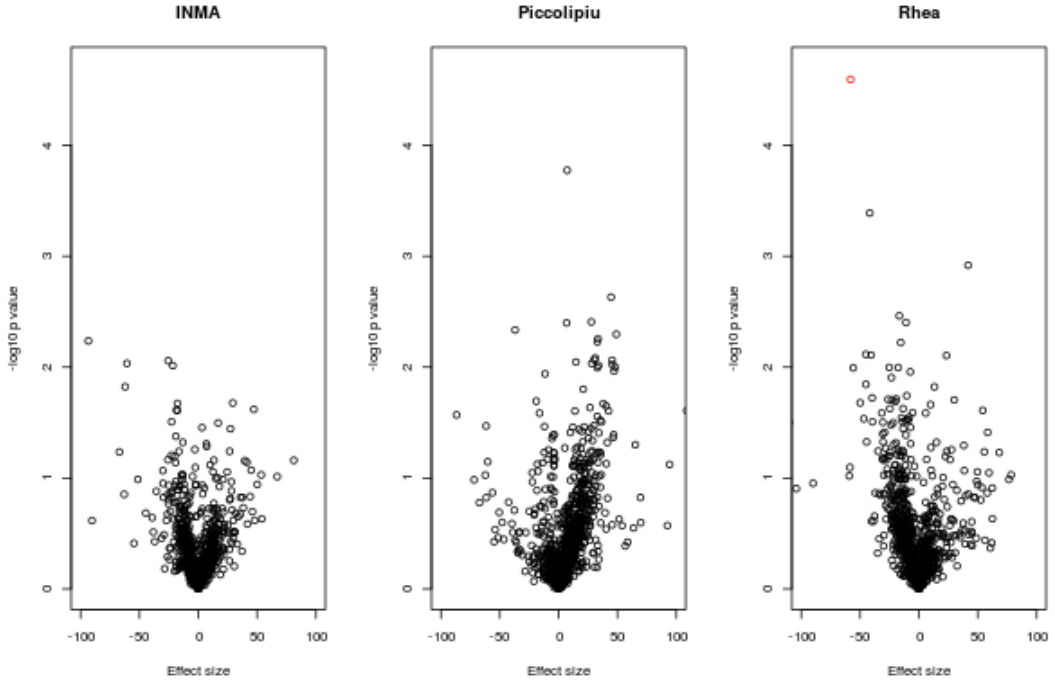


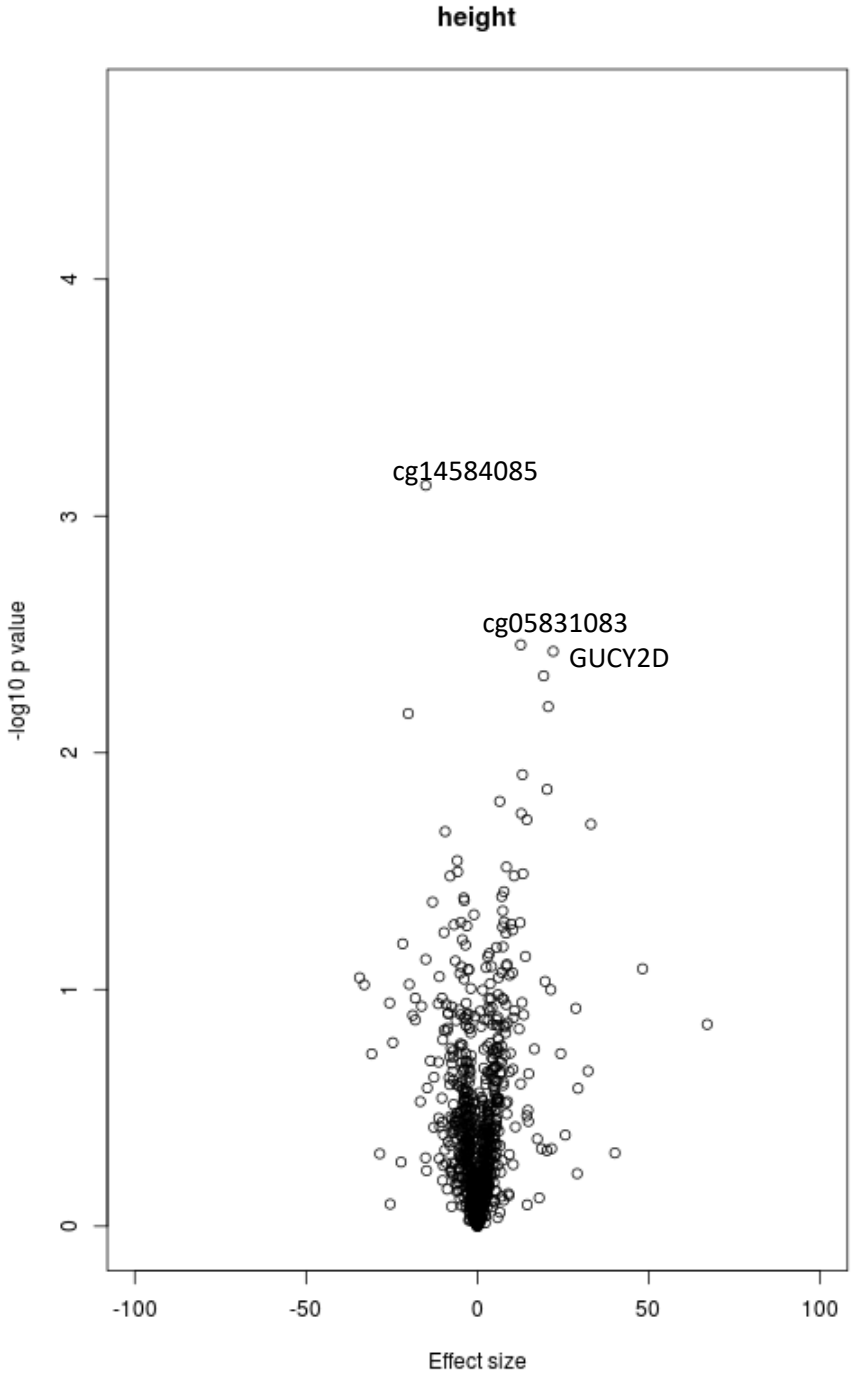
Results

CHILDHOOD (0-10y)

- Height

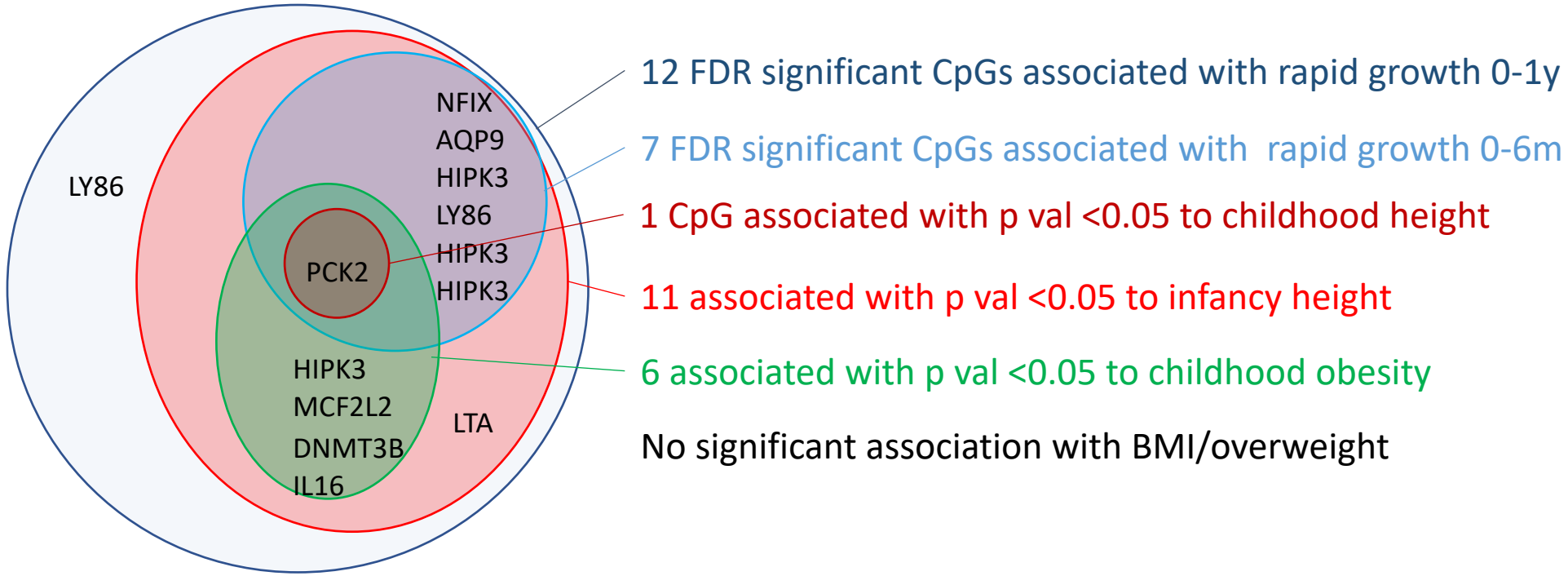




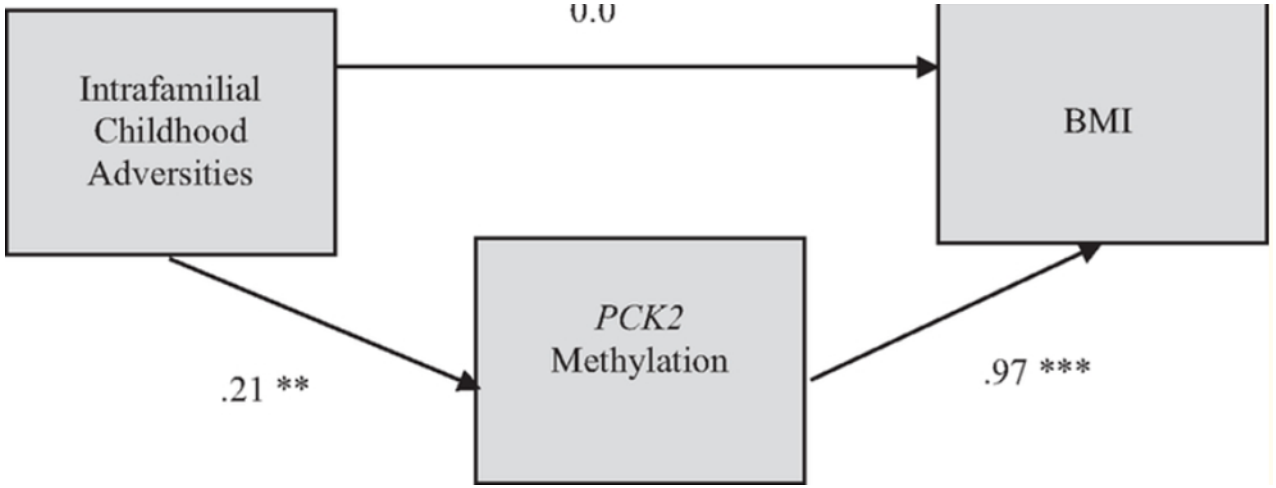


Summary of the results

- No FDR significant CpG related to BMI, obesity, overweight or height in any model tested during infancy or childhood
- 12 and 7 (including in the previous analyses) FDR significant CpGs related to rapid growth at 6 and 12 months
- Look up of the 12 FDR significant found 1 CpG associated with (childhood and infancy) height and (childhood) obesity



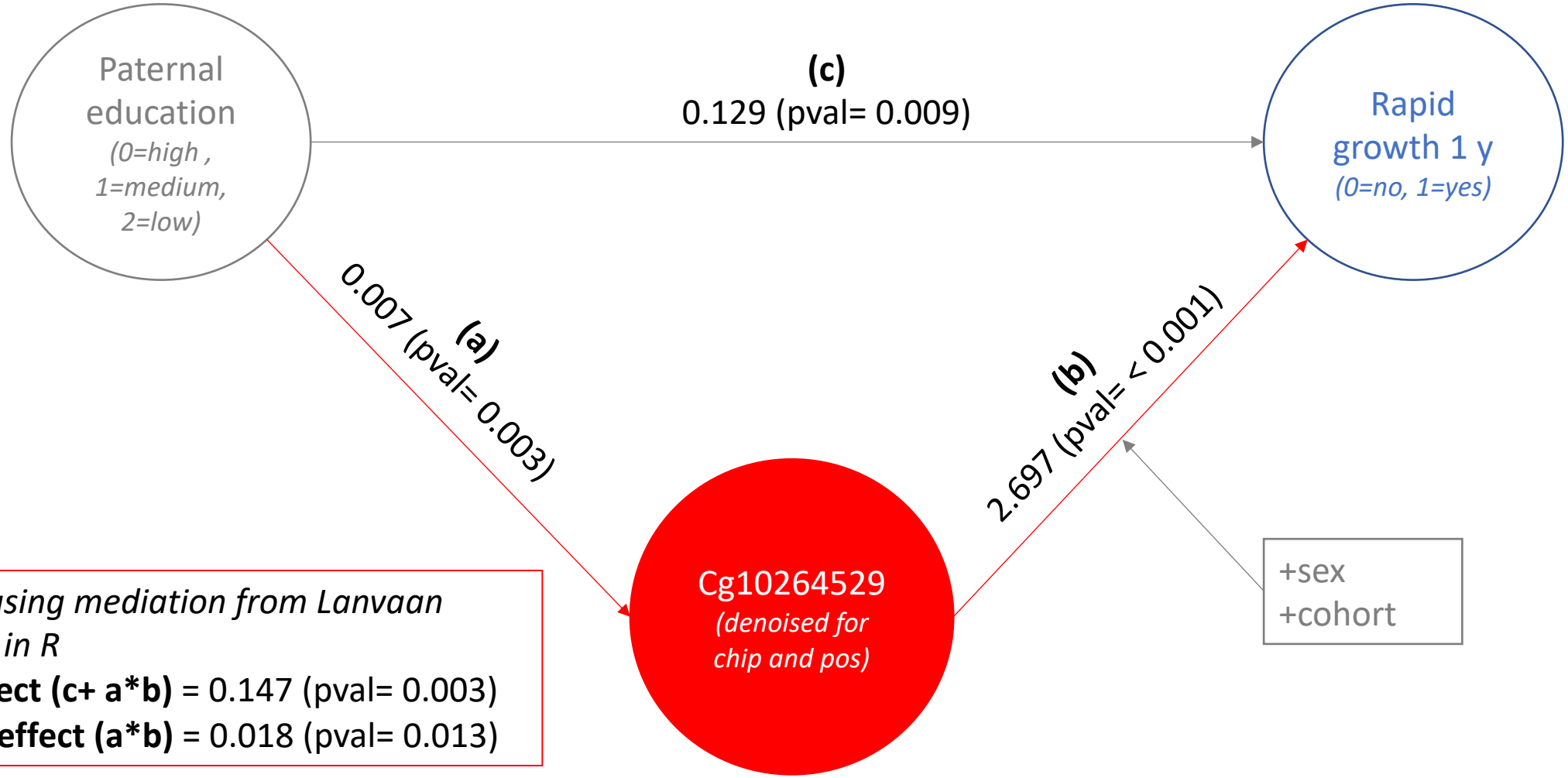
Mediation analysis
Hypothesis



- Study on resilience to maltreatment (37% of the children had an out-of-home placement because of reports of abuse or neglect, another 15% of the children had prior allegations of maltreatment, and 48% of the sample was never referred to protective services)
- Saliva DNA methylation using 850K in US children of age 10y

Kaufman, J Pediatr, 2018

Mediation analysis
Preliminary results



Results using mediation from Lanvaan package in R

Total effect (c+ a*b) = 0.147 (pval= 0.003)

Indirect effect (a*b) = 0.018 (pval= 0.013)

Addendum 3 – Supplementary material metabolomics

1. Supplementary Material

Table s1 Overweight population levels in different age classes

Age (yr)		ENVIRONAGE	INMA-Sabadell	Piccolipiu	RHEA	Total
	Non overweight	81	87	47	84	299
0.5	Overweight	3	1		4	8
	Total	84	88	47	88	307
	Non overweight	95	81	54	85	315
1	Overweight	4	3	4	2	13
	Total	99	84	58	87	328
	Non overweight	93	14	56	76	239
1.5	Overweight	2		1	2	5
	Total	95	14	57	78	244
	Non overweight	51	80	62	64	257
2	Overweight	2	3	5	4	14
	Total	53	83	67	68	271
	Non overweight	1	61	2	49	113
3	Overweight		3		6	9
	Total	1	64	2	55	122
	Non overweight	7	82	53	79	221
4	Overweight		8	4	15	27
	Total	7	90	57	94	248
	Non overweight		38	16	24	78
5	Overweight		6		7	13
	Total		44	16	31	91
	Non overweight		37		24	61
6	Overweight		19		22	41
	Total		56		46	102
	Non overweight		49		20	69
7	Overweight		21		13	34
	Total		70		33	103

Table s2 Obese population levels in different age classes

Age (yr)		ENVIRONAGE	INMA-Sabadell	Piccolipiu	RHEA	Total
	Non obesity	84	88	47	87	306
0.5	Obesity	0	0	0	1	1
	Total	84	88	47	88	307
	Non Obesity	99	84	58	87	328
1	Obesity	0	0	0	0	0
	Total	99	84	58	87	328
	Non Obesity	95	14	57	78	244
1.5	Obesity	0	0	0	0	0
	Total	95	14	57	78	244
	Non Obesity	53	83	66	67	269
2	Obesity	0	0	1	1	2
	Total	53	83	67	68	271
	Non Obesity	1	63	2	53	119
3	Obesity	0	1	0	2	3
	Total	1	64	2	55	122
	Non Obesity	7	89	55	90	241
4	Obesity	0	1	2	4	7
	Total	7	90	57	94	248
	Non Obesity		41	16	30	87
5	Obesity		3	0	1	4
	Total		44	16	31	91
	Non Obesity		48		40	88
6	Obesity		8		6	14
	Total		56		46	102
	Non Obesity		58		29	87
7	Obesity		12		4	16
	Total		70		33	103

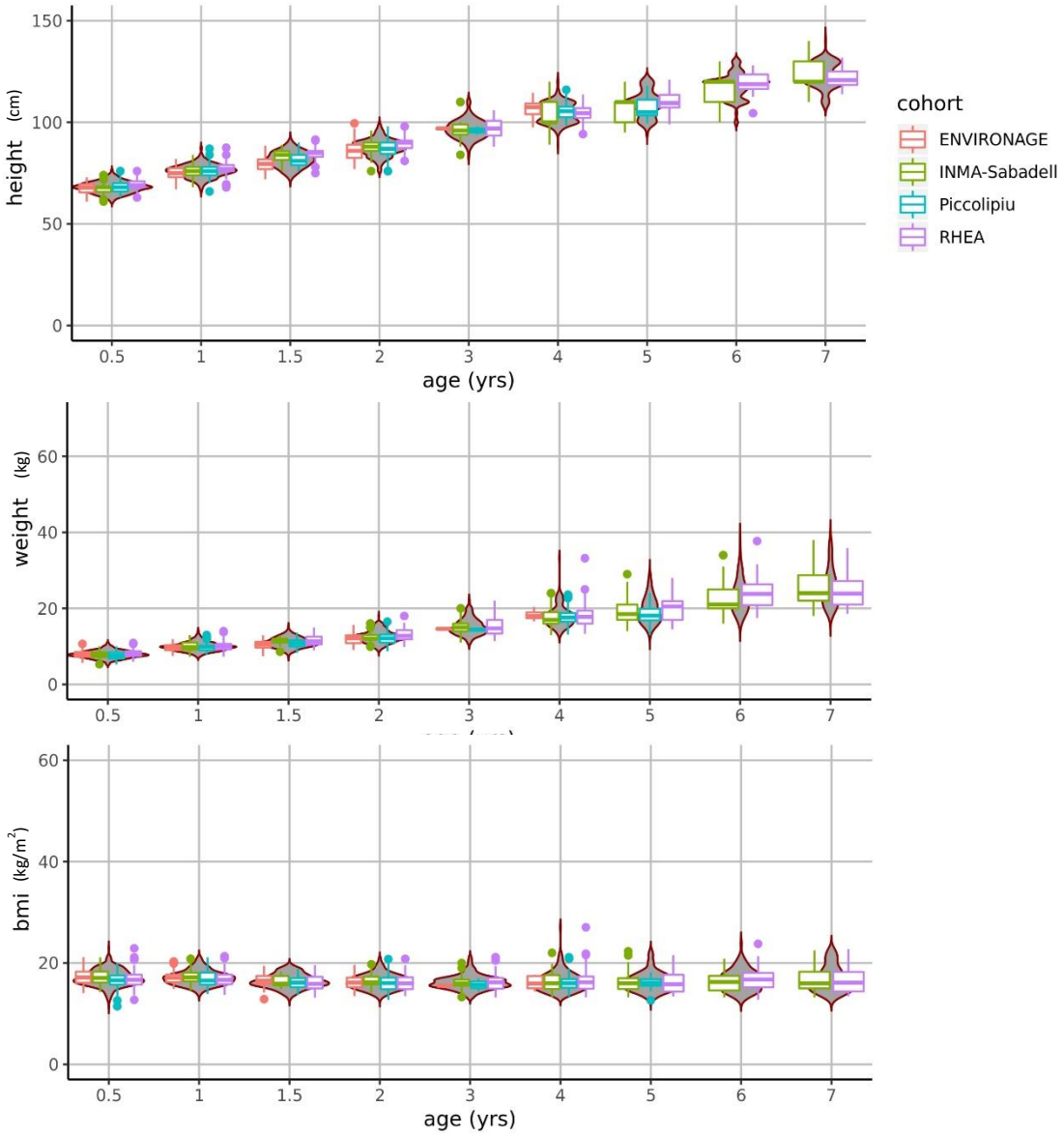


Figure s1. Summary of height, weight and BMI over time by cohort

Table 1 birthweight-related metabolites in cord blood

ID	Metabolite name	m/z	retention time
1	2-Hexenoylcarnitine(C6:1)	X257.1626	2.8306587
2	35-Tetradecadiencarnitine(C14:2)	X367.272	5.631012
3	3-Hydroxy-cis-5-tetradecenoylcarnitine(C14:1)	X385.2826	5.568466
4	3-Hydroxyhexadecadienoylcarnitine(C16:1)	X411.2972	5.749766
5	912-Hexadecadienoylcarnitine(C14:1)	X395.3027	5.9437513
6	Butyrylcarnitine/Isobutyryl-L-carnitine(C4:0)	X231.1464	1.9274178
7	Cholestenone	X384.3414	9.76708
8	Cholestenone	X406.3226	9.73516
9	Cholesterol	X368.3448	9.60744
10	Decanoylcarnitine(C10:0)	X315.2416	5.1387444

11	Decenoylcarnitine_1(C10:1)	X313.2248	4.8776007
12	Decenoylcarnitine_2(C10:1)	X313.2245	4.9669886
13	Diacylglycerol(C34:2)	X614.4886	9.762555
14	Diacylglycerol(C36:3))	X640.5039	9.9381895
15	Diacylglycerol(C36:4)	X638.4873	9.408274
16	Docosahexaenoicacid	X328.2409	7.2322
17	Dodecanoylcarnitine(C12:0)	X343.2724	5.647444
18	Hexadecenoylcarnitine(C16:1)	X397.3191	6.1093335
19	Indolelacticacid-Indolelactate	X205.749	3.8289883
20	Leucine	X131.948	1.4519173
21	L-Octanoylcarnitine(C8:0)	X287.2098	4.4222255
22	LysoPC(20:4)	X543.3342	6.894307
23	LysoPC(20:4)	X562.3068	6.8930106
24	LysoPC(22:6/0:0)	X567.3336	6.88448
25	LysoPC(C16:1)	X493.3177	6.817081
26	LysoPC(C18:1)	X521.3482	6.979925
27	LysoPC(C18:1)	X521.3492	7.580955
28	LysoPC(C18:3)	X517.3143	6.7819257
29	LysoPC(C20:2)	X547.3608	7.141414
30	LysoPC(C22:5)	X569.3478	7.206504
31	Methoxykynurenate(C11H9NO4)	X219.532	3.6709497
32	PC(30:0)	X705.5337	8.492703
33	PC(34:2)	X389.7649	8.686479
34	PC(34:2)	X769.54	8.688093
35	PC(34:2)	X776.5372	8.688971
36	PC(34:2)	X777.398	8.68937
37	PC(34:2)	X789.423	8.67179
38	PC(36:4)	X781.5649	9.57233
39	PC(C32:0)	X733.5627	8.960004
40	PC(C34:2)	X757.5674	8.684198
41	PC(C36:4)	X792.5541	8.628368
42	PC(C38:4)	X809.598	9.168946
43	PlasmalogenPC(36:3)orPC(O-36:4)	X767.581	9.189
44	PlasmalogenPC(36:3)orPC(O-36:4)	X768.585	9.40508
45	PlasmalogenPC(36:4)orPC(O-36:5)orPC-O(C36:4)(C44H80NO7P)	X765.5742	8.858829
46	PlasmalogenPC(38:4)orPC(O-38:5)	X793.5973	9.77853
47	Progesterone	X314.2247	6.3944817
48	Retinol	X268.2205	7.2190323
49	Serotonin	X159.68	1.4937996
50	Sphingosine	X299.2832	6.4203
51	Tetradecanoylcarnitine(C14:0)	X371.3039	6.56033
52	Tetradecenoylcarnitine(C14:1)	X369.2882	5.840157
53	Trans-2-Dodecenoylcarnitine(C12:1)	X341.2568	5.422301
54	Tryptophan	X204.892	2.4842238

55	Unassigned_1	X833.738	0.5795376
56	Unassigned_10	X266.98	2.6798394
57	Unassigned_11	X114.681	2.6996837
58	Unassigned_12	X219.9888	2.7254777
59	Unassigned_14	X142.99	3.423706
60	Unassigned_17	X250.158	3.8512483
61	Unassigned_19	X233.491	4.8377566
62	Unassigned_2	X197.822	0.6991766
63	Unassigned_21	X212.821	4.916926
64	Unassigned_26	X444.2869	5.591346
65	Unassigned_3	X164.686	0.7134448
66	Unassigned_31	X445.2468	5.8538017
67	Unassigned_33	X442.2717	5.996106
68	Unassigned_36	X344.2004	6.265788
69	Unassigned_39	X408.7401	6.8039536
70	Unassigned_4	X204.218	0.85186327
71	Unassigned_41	X294.218	6.8192997
72	Unassigned_43	X537.2978	6.8364196
73	Unassigned_44	X288.1258	6.8376226
74	Unassigned_45	X288.1292	6.9051366
75	Unassigned_46	X440.3215	6.9641623
76	Unassigned_48	X545.3498	6.9990945
77	Unassigned_49	X503.3011	7.445
78	Unassigned_5	X274.21	0.87270516
79	Unassigned_52	X371.303	7.104977
80	Unassigned_53	X376.2604	7.1117277
81	Unassigned_57	X598.4285	7.278556
82	Unassigned_58	X347.286	7.3553934
83	Unassigned_59	X377.7844	7.468393
84	Unassigned_6	X218.112	1.483967
85	Unassigned_61	X781.544	8.670682
86	Unassigned_63	X807.5752	8.730498
87	Unassigned_7	X292.47	1.7388779
88	Unassigned_9	X145.1105	2.5181222
89	UnidentifiablePC/LysoPC	X511.3289	6.1962085
90	Unknown(U88)	X268.1821	5.3084226
91	Unknown(U88)	X286.1932	5.309032
92	Unknown(U88)	X454.1352	5.308898
93	Unknown(U88)	X469.989	5.308984
94	Unknown(U88)	X472.1083	5.3095765
95	Unknown(U88)	X502.3106	5.3096886

Table s4. Top 20 strongest associations in metabolome wide analysis with BMI using repeated measurements until the age of 2 (left) and 7 (right) years.

id	MR/RT	Coefficient	q-value	p-value
1	X336.2518.7.4757833	-0.4823	0.0007	0.8821
2	X639.4683.7.431191	0.8297	0.0009	0.8821
3	X131.0948.0.8659426	-1.1130	0.0010	0.8821
4	X316.2393.5.9756694	-0.3923	0.0018	0.8821
5	X85.089.0.86572987	-0.8401	0.0020	0.8821
6	X331.2359.4.602693	-0.4998	0.0020	0.8821
7	X312.1478.3.6270492	-0.1993	0.0020	0.8821
8	X369.2882.5.840157	-0.4374	0.0021	0.8821
9	X259.1785.3.3326857	-0.2881	0.0026	0.8821
10	X278.0873.5.621556	2.7304	0.0027	0.8821
11	X129.9579.0.48668832	0.4017	0.0028	0.8821
12	X357.2874.5.78511	-0.3188	0.0033	0.8821
13	X157.1104.0.82387877	-0.1722	0.0036	0.8821
14	X349.1951.3.2151089	-0.1465	0.0037	0.8821
15	X275.9929.2.7263856	-0.2364	0.0037	0.8821
16	X590.4257.7.1836786	0.6171	0.0041	0.8821
17	X315.2416.5.1387444	-0.4161	0.0044	0.8821
18	X218.0883.6.0474076	-1.8232	0.0044	0.8821
19	X505.4132.8.112494	0.5862	0.0046	0.8821
20	X219.1113.2.3442717	-0.3708	0.0047	0.8821

id	MR/RT	Coefficient	p-value	p-value
1	X131.0948.0.8659426	-1.2427	0.0001	0.1798
2	X129.9579.0.48668832	0.4963	0.0001	0.1798
3	X336.2518.7.4757833	-0.4372	0.0010	0.7256
4	X85.089.0.86572987	-0.8383	0.0010	0.7256
5	X457.4132.6.647118	0.6385	0.0011	0.7256
6	X375.7688.7.2666116	0.7640	0.0014	0.7256
7	X312.1478.3.6270492	-0.1573	0.0015	0.7256
8	X316.2393.5.9756694	-0.3765	0.0015	0.7256
9	X185.1056.0.7420742	-0.2929	0.0016	0.7256
10	X302.1868.6.4020286	-0.1948	0.0020	0.7256
11	X815.5398.8.588964	-0.6703	0.0023	0.7256
12	X201.174.0.8535159	-0.0893	0.0024	0.7256
13	X226.0755.5.852551	-0.6162	0.0024	0.7256
14	X457.413.6.436681	0.6119	0.0026	0.7256
15	X268.2058.6.6735897	-0.5237	0.0026	0.7256
16	X260.9006.0.49199072	0.3434	0.0027	0.7256
17	X168.089.1.3543156	0.2021	0.0028	0.7256
18	X275.121.6.8170886	0.6267	0.0029	0.7256
19	X112.0267.0.8771368	-0.3975	0.0032	0.7256
20	X275.9929.2.7263856	-0.2239	0.0035	0.7256

Table s5. Top 20 strongest associations in metabolome wide analysis with height using repeated measurements until the age of 2 (left) and 7 (right) years.

id	MR/RT	Coefficient	q-value	p-value
1	X308.2036.7.524293	0.7638	0.0014	0.9999
2	X262.1543.0.8634436	-1.5525	0.0026	0.9999
3	X341.2627.7.1819553	1.1144	0.0052	0.9999
4	X227.1894.7.3115435	-0.4716	0.0074	0.9999
5	X229.1757.0.8583179	-1.0062	0.0085	0.9999
6	X312.1378.6.526101	-1.5430	0.0087	0.9999
7	X121.0893.3.747155	-0.3458	0.0099	0.9999
8	X395.3027.5.9437513	-0.9488	0.0106	0.9999
9	X206.13.5.009051	-1.0108	0.0114	0.9999
10	X129.0578.4.5602818	-0.5898	0.0114	0.9999
11	X300.2458.5.460584	-0.8065	0.0117	0.9999
12	X328.1432.7.9358845	-1.0533	0.0139	0.9999
13	X187.1691.0.5113096	-0.7676	0.0143	0.9999
14	X274.0173.4.5597873	-0.5092	0.0151	0.9999
15	X227.1887.7.347967	-0.5178	0.0159	0.9999
16	X206.1436.2.6781414	-0.2194	0.0161	0.9999
17	X211.0612.4.5604362	-0.4982	0.0173	0.9999
18	X581.2393.6.3809404	-0.8010	0.0174	0.9999
19	X153.0191.0.68496245	-0.8941	0.0181	0.9999
20	X506.3232.6.944828	-0.7723	0.0181	0.9999

id	MR/RT	Coefficient	p-value	p-value
1	X203.0952.4.9076886	-0.9779	0.0008	0.8544
2	X371.3039.6.0056033	-1.2404	0.0009	0.8544
3	X341.2627.7.1819553	1.4112	0.0010	0.8544
4	X262.1918.6.4035983	-0.9171	0.0011	0.8544
5	X262.1543.0.8634436	-1.7428	0.0012	0.8544
6	X743.0785.7.436957	2.6639	0.0014	0.8544
7	X427.3659.6.4829483	-1.4862	0.0015	0.8544
8	X300.2458.5.460584	-1.0289	0.0018	0.8544
9	X308.2036.7.524293	0.7547	0.0021	0.8544
10	X160.0885.6.33853	1.3549	0.0023	0.8544
11	X395.3027.5.9437513	-1.1831	0.0023	0.8544
12	X121.0893.3.747155	-0.4170	0.0025	0.8544
13	X458.2667.5.8580813	-1.7906	0.0030	0.8544
14	X399.3359.6.2763915	-1.1902	0.0035	0.8544
15	X582.2673.5.8608766	-1.2043	0.0038	0.8544
16	X232.0632.3.5262477	-1.2098	0.0041	0.8544
17	X529.3255.5.461062	-0.9855	0.0045	0.8544
18	X227.1887.7.347967	-0.6295	0.0045	0.8544
19	X435.7684.9.072139	-1.2760	0.0047	0.8544
20	X148.1109.3.5258794	-0.9404	0.0051	0.8544

Table s6. Top 20 strongest associations in metabolome wide analysis with obesity using repeated measurements until the age of 2 (left) and 7 (right) years.

id	MR/RT	Coefficient	q-value	p-value
1	X278.0873.5.621556	0.1367	4.21E-11	1.98E-07
2	X690.452.7.2734523	-0.0178	0.0001	0.1371
3	X331.2359.4.602693	-0.0151	0.0001	0.1864
4	X342.1545.7.2798567	-0.0171	0.0002	0.1864
5	X382.1756.5.0068326	-0.0156	0.0003	0.1864
6	X284.1453.6.103807	-0.0258	0.0003	0.1864
7	X364.1913.7.2805057	-0.0151	0.0003	0.1864
8	X334.1598.7.2784166	-0.0140	0.0004	0.1864
9	X179.1293.5.580873	0.0484	0.0004	0.1864
10	X266.1502.5.9844193	0.0203	0.0004	0.1864
11	X373.2975.6.394354	-0.0137	0.0004	0.1864
12	X357.2514.5.0017495	-0.0091	0.0005	0.2051
13	X157.1468.5.5808787	0.0487	0.0007	0.2084
14	X346.2145.5.531753	-0.0127	0.0007	0.2084
15	X406.2447.8.68883	-0.0246	0.0007	0.2084
16	X446.1754.6.395019	-0.0107	0.0008	0.2084
17	X341.2568.5.422301	-0.0122	0.0008	0.2084
18	X314.2247.6.3944817	-0.0107	0.0009	0.2084
19	X360.1948.5.006783	-0.0088	0.0009	0.2084
20	X424.7867.7.2484627	-0.0093	0.0009	0.2084

id	MR/RT	Coefficient	p-value	p-value
1	X278.0873.5.621556	0.1365	0.0001	0.2747
2	X424.208.6.769028	-0.0083	0.0006	0.9744
3	X210.1367.3.863549	-0.0151	0.0015	0.9744
4	X296.0753.7.126658	0.0067	0.0018	0.9744
5	X140.9521.0.4823698	-0.0223	0.0022	0.9744
6	X83.9529.0.4944189	0.0094	0.0024	0.9744
7	X646.4503.7.7784834	-0.0181	0.0025	0.9744
8	X129.0003.0.9498402	0.1045	0.0030	0.9744
9	X287.9154.0.52467614	0.0280	0.0031	0.9744
10	X218.1675.6.096486	-0.0297	0.0033	0.9744
11	X296.076.6.840447	0.0096	0.0040	0.9744
12	X442.3286.7.1370683	-0.0253	0.0046	0.9744
13	X345.1295.6.2690344	0.0102	0.0048	0.9744
14	X314.188.6.313442	-0.0270	0.0051	0.9744
15	X248.0336.3.525237	0.0166	0.0053	0.9744
16	X386.2695.7.452902	-0.0231	0.0054	0.9744
17	X610.9355.7.181809	0.0212	0.0056	0.9744
18	X143.9747.0.49295908	0.0132	0.0056	0.9744
19	X581.2393.6.3809404	-0.0158	0.0062	0.9744
20	X619.9198.7.181728	0.0249	0.0063	0.9744

Table s7. Top 20 strongest associations in metabolome wide analysis with overweight using repeated measurements until the age of 2 (left) and 7 (right) years.

id	MR/RT	Coefficient	q-value	p-value
1	X278.0873.5.621556	0.5814	1.25E-09	5.89E-06
2	X220.0162.0.898141	-0.0541	0.0000	0.0528
3	X446.3423.8.116376	0.1084	0.0002	0.2384
4	X505.4132.8.112494	0.0803	0.0004	0.3952
5	X193.1111.5.8661065	-0.1369	0.0004	0.3952
6	X188.9869.1.4949781	-0.0276	0.0007	0.4069
7	X355.2719.5.6389165	-0.0519	0.0007	0.4069
8	X357.2874.5.78511	-0.0377	0.0007	0.4069
9	X150.0284.1.0047277	0.2742	0.0008	0.4069
10	X73.0253.0.9259496	0.2212	0.0012	0.4581
11	X230.0111.0.89772815	-0.0432	0.0012	0.4581
12	X363.1896.4.809886	0.0769	0.0013	0.4581
13	X468.323.8.123368	0.0900	0.0013	0.4581
14	X166.006.1.0264771	0.2331	0.0014	0.4636
15	X183.0205.0.8958215	-0.0355	0.0015	0.4636
16	X246.0422.0.8965959	-0.0208	0.0017	0.4935
17	X341.2568.5.422301	-0.0519	0.0018	0.4942
18	X246.0415.1.4951686	-0.0178	0.0028	0.7370
19	X129.0003.0.9498402	0.2879	0.0035	0.8668
20	X232.0526.1.4952639	-0.0148	0.0040	0.9344

id	MR/RT	Coefficient	p-value	p-value
1	X278.0873.5.621556	0.4673	1.48E-08	6.99E-05
2	X325.2977.7.1644764	-0.0432	0.0003	0.4567
3	X446.3423.8.116376	0.0855	0.0004	0.4567
4	X150.0284.1.0047277	0.2499	0.0005	0.4567
5	X220.0162.0.898141	-0.0384	0.0005	0.4567
6	X73.0253.0.9259496	0.2039	0.0006	0.4567
7	X437.7798.6.9388313	0.0702	0.0011	0.4705
8	X375.7688.7.2666116	0.0711	0.0013	0.4705
9	X131.0948.0.8659426	-0.1019	0.0016	0.4705
10	X328.0247.0.83774406	0.0996	0.0017	0.4705
11	X132.0251.0.8538687	-0.0567	0.0018	0.4705
12	X765.5214.6.9398766	0.0568	0.0019	0.4705
13	X85.089.0.86572987	-0.0781	0.0020	0.4705
14	X290.0127.3.630484	-0.0269	0.0021	0.4705
15	X184.1087.0.52134717	0.0652	0.0023	0.4705
16	X468.323.8.123368	0.0717	0.0024	0.4705
17	X374.2418.7.134108	0.0515	0.0026	0.4705
18	X527.8746.7.1863704	0.0495	0.0027	0.4705
19	X589.4193.6.9426627	0.0611	0.0027	0.4705
20	X363.1896.4.809886	0.0601	0.0027	0.4705

Table s8. Top 20 strongest associations in metabolome wide analysis with rapid growth at 12 months

id	Metabolite	Coefficient	q-value	p-value
1	X481.2319.3.6582649	-0.1765	6.00E-09	2.83E-05
2	X384.3414.9.076708	0.5575	2.25E-08	5.31E-05
3	X406.3226.9.073516	0.4759	3.95E-07	0.0006
4	X443.4132.9.076516	0.5110	5.03E-07	0.0006
5	X446.1942.4.453114	-0.1882	3.44E-06	0.0032
6	X268.1821.5.3084226	-0.2219	5.97E-06	0.0047
7	X481.2323.4.453742	-0.1484	1.05E-05	0.0071
8	X257.2721.5.9162836	0.1180	3.28E-05	0.0193
9	X195.0527.2.9204202	0.2927	4.00E-05	0.0210
10	X471.1157.4.8988914	-0.1953	7.35E-05	0.0346
11	X486.1867.4.4549704	-0.1508	1.12E-04	0.0478
12	X368.3448.9.060744	0.7085	1.22E-04	0.0478
13	X385.7551.8.959619	0.3063	1.40E-04	0.0509
14	X550.3092.6.9024997	0.4307	0.0002	0.0550
15	X733.5627.8.960004	0.2610	0.0002	0.0550
16	X229.241.5.5392804	0.0872	0.0002	0.0550
17	X288.2084.4.8316393	-0.2059	0.0003	0.0800
18	X282.2184.6.8522468	-0.1265	0.0004	0.0971
19	X765.5742.8.858829	0.3486	0.0004	0.1010
20	X286.1939.4.8652577	-0.1818	0.0006	0.1497

Multivariate analysis with prenatal risk factors and metabolites

Introduction

In order to explore the degree to which metabolites and pre-natal exposure factors can predict overweight, obesity and BMI at different ages using Artificial Intelligence machine learning approaches, we defined predictive models consisting of a combination of metabolites, prenatal factors and overweight, obesity and BMI. Random Forest regression and classification models have many advantages as suitable biomarker selection tools for metabolomic data analysis (Chen et al. 2013a), but the robustness of methods relies on the data quantity and quality. We aimed to assess whether metabolomics may improve prediction of overweight/obesity over using traditional risk factors.

Methods

We used scikit-learn (Pedregosa et al. 2011) for applying Random Forest (RF) regression and classification models in the predictive modelling analysis. RF models have the advantage that they are non-linear models, require relatively quick training for high dimensional datasets, can reduce the time to build patterns and increase the detection rate, provide a robust way to detect outliers (Pal 2005), provide relatively simple and easy way to model hyperparameter tuning and feature selection, and are well-studied approaches in clinical metabolomics (Chen et al. 2013b)

Random Forest models were initially setup using scikit-learn default parameters: e.g. criterion (the metric for selecting a split) set to “gini”, max_features (the number of features considered at each split) set to the square root of the number of features, max_depth (the maximum depth of each tree) set to “None” to permit unlimited depth. The optimal parameters were assessed using a five-fold cross-validation. The cross-validation routine was employed within the training set and the model performances evaluated using the area under the ROC curve for binary outcomes and average mean square error for continuous outcomes.

Finally, to demonstrate the robustness of the performance statistics computed for these models, and the degree to which they depend upon the random selection of test compounds, the training-test splits were repeated an additional 20 times for those splitting methods where randomness played a role (i.e. the random method and the single source method).

Results

A predictive model based on traditional prenatal risk factors only performed well in predicting overweight at age 7 (fig s2), with a mean area under the ROC curve (AUROC) of 0.79. Predictors in this model are shown in table s9. The most important predictors in the model were paternal BMI, parity and maternal weight gain. We compared the performance of the models to predict overweight, based on those using traditional prenatal risk factors only, those using metabolites only and those using risk factors and metabolites together (table s10). While metabolites performed reasonably well in predicting growth outcomes up to two years, generally models performed worse than those using risk factors alone.

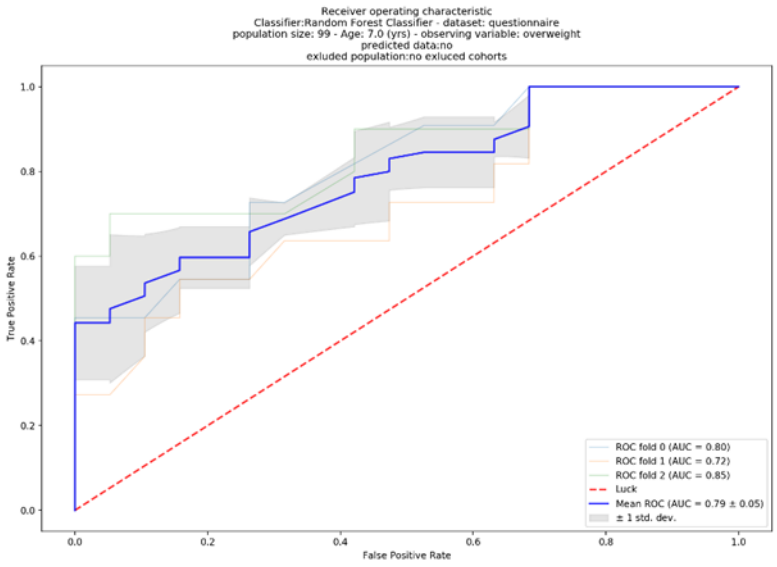


Figure s2. Receiver operating characteristic curve, of prediction model of overweight at 7 years, based on tradition prenatal risk factors

Table s9. Variable importance of risk factors in random forest prediction model of overweight at 7 years

num	Covariate	Variable improtance
1	Paternal BMI	1
2	Parity	0.99
3	Maternal weigth gain at the end of the pregnancy (kg)	0.95
4	Maternal pre-pregnancy BMI (kg/m2)	0.69
5	Maternal passive smoking at some time during pregnancy	0.58
6	Pregnancy-related pathologies(current pregnancy): Hypertension	0.53
7	Paternal BMI	0.51
8	Maternal pre-pregnancy BMI (kg/m2)	0.31
9	paternal education	0.27
10	maternal education	0.23
11	maternal nationality	0.18
12	mean average number of cigarettes per week during the pregnancy	0.17
13	cohort study centre	0.11
14	maternal active smoking during pregancy	0.11
15	maternal active smoking during trimester 1	0.04
16	birth weight (g)	0.04
17	maternal active smoking during trimester 2	0.01
18	period of conception of the pregnancy (based on the estimated fertilization date)	0
	gestational age based on best available records (weeks)	0

Table s10. Comparison of AUROC values for random forest models based on different sets of predictors at different ages.

Age/outcome	Prenatal Risk Factor	Birthweight Metabolites	Birthweight Metabolites + Prenatal Risk Factor
0.5			

obese			
overweight	0.69	0.26	0.34
rapid_growth_6	0.81	0.67	0.65
1			
obese			
overweight	0.69	0.53	0.56
rapid_growth_12	0.72	0.70	0.60
1.5			
obese			
overweight	0.63	0.71	0.66
2			
obese	0.00	0.00	0.00
overweight	0.59	0.40	0.32
3			
obese	0.63	0.55	0.65
overweight	0.59	0.48	0.46
4			
obese	0.68	0.39	0.54
overweight	0.62	0.47	0.51
5			
obese	0.87	0.61	0.65
overweight	0.69	0.35	0.44
6			
obese	0.67	0.56	0.52
overweight	0.72	0.53	0.59
7			
obese	0.53		0.55
overweight	0.74	0.49	0.63

Conclusion

Preliminary analyses do not suggest that metabolites may improve prediction of obesity over traditional risk factors. We will explore if the addition of further variable selection methods improves model performance with molecular markers.

References

Chen, Tianlu, Yu Cao, Yinan Zhang, Jiajian Liu, Yuqian Bao, Congrong Wang, Weiping Jia, and Aihua Zhao. 2013a. "Random forest in clinical metabolomics for phenotypic discrimination and biomarker selection." *Evidence-based complementary and alternative medicine : eCAM* 2013:298183-298183. doi: 10.1155/2013/298183.

Chen, Tianlu, Yu Cao, Yinan Zhang, Jiajian Liu, Yuqian Bao, Congrong Wang, Weiping Jia, and Aihua Zhao. 2013b. "Random forest in clinical metabolomics for phenotypic discrimination and biomarker selection." *Evidence-Based Complementary and Alternative Medicine* 2013.

Pal, Mahesh. 2005. "Random forest classifier for remote sensing classification." *International Journal of Remote Sensing* 26 (1):217-222.

Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, and Vincent Dubourg. 2011. "Scikit-learn: Machine learning in Python." *Journal of machine learning research* 12 (Oct):2825-2830.