Cardiovascular Diabetology

RESEARCH

Open Access

Novel epigenetic marks of insulin resistance trajectories in a longitudinal study of childhood obesity

Augusto Anguita-Ruiz[1,2†], Álvaro Torres-Martos[2,3,4†], Mireia Bustos-Aibar[2,3,5], Adrià Setó-Llorens[1], Francisco Javier Ruiz-Ojeda[2,3,4,6], Luis A. Moreno[2,5], Ángel Gil[2,3,4], Mercedes Gil-Campos[2,7], Gloria Bueno[2,5,8], Rosaura Leis[2,9,10*], Jesús Alcalá-Fdez[4,11†] and Concepción M. Aguilera[2,3,4†]

Abstract

Background Childhood obesity is a major global public-health challenge. Insulin resistance (IR) is a critical driver of later cardiometabolic alterations. A comprehensive understanding of the molecular mechanisms underlying the initial development of childhood IR is essential for timely prevention and intervention. In this study, we aimed to assess the association between IR and blood DNA methylation in a longitudinal study from childhood into adolescence.

Methods The PUBMEP study included a longitudinal core of 90 children with paired blood samples collected at both pre-pubertal and pubertal stages. For cross-sectional analyses, this sample was expanded to 99 pre-pubertal and 129 pubertal participants. IR status was defined according to clinically relevant sex- and pubertal stage specific HOMA-IR cut-offs, as recommended by pediatric expert clinicians. Genotype data was obtained with the Infinium Global Screening Array, and blood DNA methylation sites with the Infinium MethylationEPIC BeadChip. Epigenome-wide associations with IR status and trajectories were tested using linear models in the longitudinal and cross-sectional sets. FDR-adjusted significant CpG sites were assessed with sex- and age-standardised cardiometabolic z-scores (adiposity, lipids, blood pressure, glycaemia and IR) at each stage. mQTL analyses were performed to identify genetic variants that drive IR-associated methylation signals.

Results We identified 120 CpG sites related to obesity-associated IR in the context of pubertal transition that remained significant after global FDR correction (FDR < 0.05). These CpG sites showed distinct methylation profiles that tracked IR trajectories from prepuberty to puberty, with consistent differences across children whose IR improved, worsened or remained stable, with several of them also related to cardiometabolic traits at pubertal stage, including adiposity measures, blood pressure and glycaemic indices. Among the FDR-significant CpG sites with biological relevance for IR, methylation at CpG sites annotated to *SLC2A9*, *PEPD*, *TSC2*, *EGLN3*, *EHD2* and *VASN* showed consistent associations with pubertal HOMA-IR z-score and, for several loci, with adiposity and blood pressure

†Augusto Anguita-Ruiz and Álvaro Torres-Martos contributed equally to this work. Jesús Alcalá-Fdez and Concepción M. Aguilera contributed equally to this work as senior authors.

*Correspondence:
Rosaura Leis
mariarosaura.leis@usc.es

Full list of author information is available at the end of the article

measures, with methylation changes paralleling IR worsening, improvement or stability across puberty. An mQTL look-up in GoDMC identified 25 cis SNP CpG associations corresponding to 20 of the 120 CpG sites, including CpG sites in *SLC2A9* and *TSC2*, indicating that only a fraction of these IR-associated CpG sites is likely to be partly influenced by nearby genetic variants.

Conclusion This longitudinal EWAS in children with obesity shows that specific blood DNA methylation signatures mirror IR status and track its evolution across the pubertal transition, with opposing methylation trajectories distinguishing improving from persistent IR. The identification of CpG sites at *VASN*, *SLC2A9*, *PEPD*, *EGLN3*, *EHD2* and *TSC2* links IR trajectories to pathways involved in vascular signalling, urate transport, extracellular matrix remodelling, and hypoxia sensing and nutrient signalling. Complementary mQTL analyses suggest that while some of this epigenetic variation is influenced by local genetic factors, a substantial component is likely acquired in response to metabolic and external exposures. If replicated and functionally characterised, these findings may help refine our understanding of the early molecular architecture of obesity-related IR and inform future strategies for cardiometabolic risk assessment and timing of preventive interventions.

Keywords DNA methylation, Epigenetics, Epigenome-wide association study, EWAS, Insulin resistance, Cardiometabolic risk factors, Longitudinal study, Pediatric obesity, Puberty, Whole-genome genotype, mQTL

Research insights

What is currently known about this topic?

- Insulin resistance is the gateway to youth-onset diabetes and later cardiovascular disease.
- Blood methylome trajectories of obesity and insulin resistance trajectories from childhood to adolescence remain unprofiled.

What is the key research question?

- Could epigenetic marks explain diverging insulin resistance trajectories during the pubertal transition in children with obesity?

What is new?

- This longitudinal study identified novel epigenetic marks that parallel insulin resistance trajectories in children with obesity during pubertal transition.
- mQTL analysis shows that some of these epigenetic marks might be partially genetically influenced.
- Our findings also suggest that methylation patterns linked to adult cardiometabolic risk may be established during childhood in the context of insulin resistance.

How might this study influence clinical practice?

- The dynamic nature of these marks suggests their reversibility upon metabolic improvement, identifying them as potential epigenetic therapeutic targets.
- Epigenetic stratification could refine monitoring of cardiometabolic risk.

Introduction

Insulin resistance (IR) is a prior condition to the development of glucose intolerance [1], which not only contributes to the development of type 2 diabetes mellitus (T2DM) but also plays a significant role in the pathogenesis of hypertension, dyslipidaemia, metabolic syndrome (MetS) and cardiovascular disease (CVD) in both children and adults [2, 3]. In particular, obesity is the most frequent driver of IR in children and adolescents [3]. A substantial number of prepubertal children with overweight or obesity continue to exhibit excess adiposity into early adulthood, a condition associated with increased morbidity and mortality [4–7]. Elevated mortality rates in individuals with obesity are predominantly attributable to the development of T2DM and an increased risk of CVD [8]. Consequently, early prevention and treatment of obesity and IR are essential [9–12].

Puberty is a process of maturation characterized by the activation of the hypothalamic pituitary gonadal axis, which drives gonadal growth, coordinates the physical and psychological transition to adulthood, and confers reproductive capacity [13]. This stage of development is associated with markedly different IR trajectories, highlighting its lasting impact on lifelong health [14, 15]. Nonetheless, the impact of puberty is not uniform across individuals [16, 17]. In healthy normal-weight youths, insulin sensitivity declines during mid-puberty, only to recover upon its completion. Conversely, in adolescents with obesity, there is evidence that IR persists [18, 19], potentially leading to an increased cardiometabolic risk later in life [20]. In fact, persistent pubertal IR is closely associated with youth-onset T2DM [21]. A comprehensive understanding of the molecular and biological processes underpinning these metabolic changes, and the additional influence of obesity, is crucial for the prevention of T2DM [22–24].

Although genome-wide association studies (GWAS) have identified numerous single nucleotide polymorphisms (SNPs) associated with T2DM and its related traits [25–31], these genetic variants explain only a modest fraction of the estimated heritability (15–18%), suggesting that additional genetic and regulatory factors still need to be elucidated [32]. DNA methylation (DNAm) is an epigenetic modification that involves the covalent addition of a methyl group to a cytosine followed by a guanine, forming a CpG dinucleotide. As a dynamic and potentially reversible mark, DNAm is susceptible to modulation by environmental factors, which can lead to alterations in gene expression and thus establish a crucial link between disease susceptibility and external exposome [33–36]. Crucially, a foundational epigenome-wide association study (EWAS) using data from nine cohorts accounted for up to 32% of BMI variance in adults, underscoring the role of complex genetic and epigenetic interactions in metabolic traits [37]. One plausible mechanism that may explain this observation is the interplay between SNPs and DNAm regulatory processes—referred to methylation quantitative trait loci (mQTL)—which has attracted increasing attention. Furthermore, previous research has demonstrated that such interactions can modulate mRNA expression and insulin secretion in adult human pancreatic islets [38].

Differential DNAm patterns at several loci have been associated with metabolic disorders, including obesity [39–41], systemic IR [42–46] and T2DM [47–51], in both blood and others metabolically active tissues. Given the central role of epigenetic mechanisms in metabolic regulation, elucidating DNAm dynamics during critical developmental windows is imperative. Puberty has been recognised as a key stage during which DNAm remains stable at certain CpG sites while undergoing significant modifications at others, potentially influencing long-term metabolic health outcomes [52–54]. Although progress has been made in characterising the epigenetic mechanisms underlying obesity-related IR [55–60], our understanding remains limited during the metabolically critical transition of puberty. Previous EWAS examining IR in children have relied predominantly on cross-sectional designs [61–63], which preclude the evaluation of dynamic methylation changes and therefore overlook the potential impact of pubertal development [64, 65]. Whereas previous studies have not examined whether mQTLs contribute to DNAm variation relevant to the IR phenotype, complementary mQTL analyses can address this gap by assessing the extent to which genetically driven methylation differences may reflect potential causal mechanisms rather than downstream metabolic effects [66]. This will further enrich our molecular understanding and help delineate clinically relevant molecular subtypes [67, 68].

This investigation is part of the prospective PUBMEP study, which evaluated the progression of cardiometabolic risk factors from pre-puberty to puberty in Spanish children [14, 15]. Given the above considerations, the present study analyzed epigenomic and genomic data specifically, DNAm at CpG sites and mQTLs interactions, associated with IR in children with obesity, assessed before and after the onset of puberty. The objectives are to clarify the molecular pathophysiology and identify epigenetic biomarkers and mQTLs that could support early risk stratification.

Methods

Study population and experimental design

In the PUBMEP study, all participants were initially recruited as prepubertal children and were subsequently recalled for follow-up medical consultations. Participants aged between 4 and 12.1 years recruited in three Spanish centres: Lozano Blesa University Clinical Hospital (Zaragoza), Santiago de Compostela University Clinical Hospital (Santiago de Compostela) and Reina Sofía University Clinical Hospital (Córdoba) between 2012 and 2015. At the follow-up visit, participants aged between 9.72 and 18.07 who exhibited clinical signs of the onset of puberty (Tanner > 1) were ultimately included in the longitudinal study between 2017 and 2020. The average time between baseline and follow-up visit was 5.84 years (median 6.19 years). Pubertal stage was assessed according to the Tanner classification [69]. Throughout the study period, participants were under regular medical monitoring by the same pediatric expert clinicians. The exclusion criteria were the use of any drug that could alter blood glucose, blood pressure or lipid metabolism, inability to comply with study procedures, and current or recent participation (within the last three months) in another research project.

The main inclusion criterion for this subcohort was the availability of high-quality DNA samples obtained during the prepubertal stage to perform whole-genome genotyping and array-based DNAm analysis. A total of 138 children were included in this subcohort with genomic and epigenomic assessments. Of these, 90 participants were evaluated at both the prepubertal and pubertal stages and therefore constitute the longitudinal population. The remaining participants correspond to independent cross-sectional samples: 9 children assessed only at the prepubertal stage (prepubertal cross-sectional population) and 39 assessed only at the pubertal stage (pubertal cross-sectional population). Accordingly, the prepubertal cross-sectional population comprised 99 participants and the pubertal cross-sectional population comprised 129 participants. Further details on the PUBMEP study can be found in [14, 15, 67].

A general overview of the study design, populations, omic analyses and findings performed is provided in Fig. 1. The main analysis undertaken in this work comprised: (1) an EWAS of IR; (2) the evaluation of the association between identified CpG sites and cardiometabolic risk outcomes; and (3) mQTL analysis.

Anthropometrics, blood pressure, biochemistry and inflammation biomarkers

Anthropometric measurements such as body weight (kg), height (cm) and waist circumference (WC) (cm) were measured at each time point using standardized procedures [70]. Fat mass was quantified at each time point using dual-energy X-ray absorptiometry (DXA) (General Electric, Lunar EnCore and QDR-Explorer TM 4500, Hologic Inc, Bedford, MA). Body mass index (BMI, $\text{weight (kg)}/\text{height (m)}^2$) and fat mass index (FMI, $\text{fat mass (kg)}/\text{height (m)}^2$) were calculated. BMI and WC z-scores was estimated based on the Spanish reference standards published [71, 72]. Blood pressure was measured three times for each individual by the same examiner using an electronic manometer (M6, HEM-7001-E, Omron, Tokio, Japan) and following international recommendations [73]. More detailed information about the blood pressure measures could be found in [74]. Measures of lipid and glucose metabolism, hormones, and classical biochemical parameters were performed at the laboratories of each participating hospital following internationally accepted quality control protocols [14, 15].

Blood samples from both time points were collected in overnight fasting conditions, centrifuged, and plasma and serum were stored at $-80°C$. Plasma adipokines, inflammation, and cardiovascular risk biomarkers (adiponectin, leptin, resistin, tumor necrosis factor alpha (TNF-α), high-sensitivity C-protein reactive (hs-CRP), interleukin-8 (IL-8), tissue plasminogen activator inhibitor-1 (tPAI-1), myeloperoxidase (MPO), monocyte chemoattractant protein 1 (MCP-1), and soluble intercellular cell adhesion molecule-1 (sICAM-1) were analyzed in all samples and time points using XMap technology (Luminex Corporation, Austin, TX) and human monoclonal antibodies (Milliplex Map Kit; Millipore, Billerica, MA) as previously reported [15, 75, 76].

Homeostasis model assessment of IR (HOMA-IR) and quantitative insulin sensitivity test index (QUICKI) were computed as surrogate markers of IR [3]. Since HOMA-IR strongly varies between ages, sexes and diseases and no reference values have been yet established in either children or adult populations [77], own cut-off points were extracted from a previous well-described Spanish cohort composed of 1669 children and adolescents [75, 78]. For the prepubertal stage, a single cut-off value of HOMA-IR=2.5 was considered for IR [75, 78]. For the pubertal stage instead, sex information was considered and different cut-off points were adopted for IR

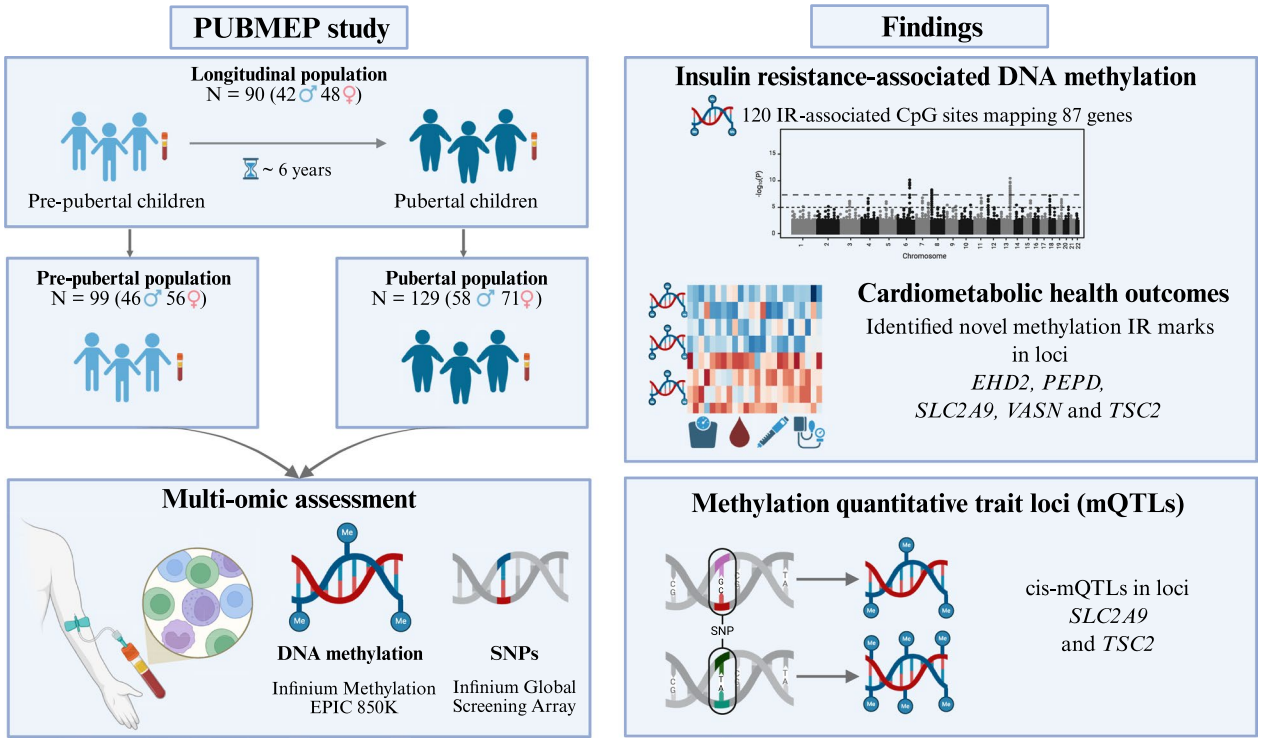

Fig. 1 Study overview

according to the 95th HOMA-IR percentile. Extracted from a subset of 778 pubertal Spanish children, cut-off values corresponded to HOMA-IR = 3.38 in boys and HOMA-IR = 3.905 in girls. These cut-off points have already been tested and validated as good metabolic risk classifiers in our population according to the results from a previous paper [79]. The obesity status was defined using the cut-offs proposed by International Obesity Task Force (IOTF) [80]. Also, we constructed age- and sex-specific z-scores for the classical cardiometabolic components, including systolic and diastolic blood pressure (SBP, DBP), triacylglycerols (TAG), HDL-cholesterol (HDL-c), LDL-cholesterol (LDL-c), glucose, insulin, and HOMA-IR. All calculations used to define obesity/IR status categorizations and z-scores were performed through ObMetrics [81].

Whole-genome genotyping

Whole-genome genotyping and EWAS were performed in all participants (Fig. 1). Buffy coat fractions from blood samples were selected for genomic and epigenomic analysis. DNA was extracted from peripheral white blood cells (PWBC) using two automated kits, the Qiamp DNA Investigator Kit for coagulated samples and the Qiamp DNA Mini & Blood Mini Kit for non-coagulated samples (QIAgen Systems, Inc., Valencia, CA, USA). All extractions were purified using the DNA Clean and Concentrator kit from Zymo Research (Zymo Research, Irvine, CA, USA). Whole-genome genotyping was performed on the i-SCan platform using the Infinium HTS Assay (Illumina, San Diego, CA, USA). The BeadChip selected for the project was the Infinium Global Screening Array (GSA) 24v3.0 Kit, which includes 654,000 genetic variants. After the quantification of DNA samples by fluorimetry, they were normalized to 200-400 ng of DNA per sample in deep well plates, as established in the Infinium HTS Assay Protocol.

The first step of the bioinformatic pipeline consisted of the extraction of genotype calls from fluorescence data and the construction of work data files for data manipulation and analysis. Using the Illumina GenomeStudio software, we obtained genotype calls for all individuals and generated the standard format files (.ped and .map). Data were then imported into PLINK 1.9 software [82], and converted into binary format files using the –make-bed flag. These binary formats (.bed, .bim and .fam) are a more compact representation of the data that saves space and speeds up subsequent analyses. We implemented a quality control (QC) process in PLINK 1.9 software, in accordance with the literature, by applying standard QC filtering criteria as follows: 1) Exclusion of SNPs and individuals with a missing data rate $\geq$ 10%, and 2) Exclusion of SNPs with a minor allele frequency (MAF) < 1% or a Hardy-Weinberg Equilibrium (HWE) p-value $< 1 \times 10^{-5}$ in controls. As a result, 471,192 SNPs remained in the dataset. These filters were selected following procedures described in [83] to minimize the influence of genotype-calling artefacts. Further details on the preprocessing of whole-genome genotyping data can be found in [67, 84].

Epigenome-wide association study of IR

Quality control and preprocessing

DNAm was measured with the Infinium Methylation EPIC 850K array using BeadChip technology (Illumina, San Diego, CA, USA) in high-quality DNA samples ($\geq$ 500 ng) from PWBC prior treatment with bisulfite using the EZ-96 DNAm Kit (Zymo Research Corporation, Irvine, CA). Raw intensity signals from .IDAT files were loaded into the R environment [85] using the minfi R package [86]. As a result, we obtained an RGChannelSet object containing all the raw intensity data, from both the red and green color channels, for each of the samples and time points. We generated a detection p-value for every CpG site in each sample by comparing the total signal for each probe to the background signal, which was estimated from the negative control probes. To minimize the unwanted variation within and between samples, we applied Beta-Mixture Quantile (BMIQ) intra-array normalization [87], including all individuals and time points. Poor performing probes were filtered out according to different criteria: probes with a detection p-value above 0.01 in more than 10 % of the samples to identify probes that have failed to hybridize (number of probes = 230) using the detectionP function, probes with SNPs (number of probes = 30,432) using the dropLociWithSnps function, cross-reactive probes aligning to multiple locations (number of probes = 25,570), as reported by [88], and probes located on the Y chromosome (number of probes = 246). After applying all these filters, 809,381 probes remained in the dataset.

Intensity values were used to determine the proportion of methylation at each CpG site. Methylation levels were reported as either β-Values $\left(\beta = \frac{M}{M+U}\right)$ or M-values $\left(M = \log_2\left(\frac{M}{U}\right)\right)$, where M and U correspond to the methylated and unmethylated signals, respectively. β-values and M-values are related through a logit transformation $\left(M = \log_2\left(\frac{\beta}{1-\beta}\right)\right)$ [89]. Because percentage methylation is easily interpretable, β-values in the present paper were employed for describing the level of methylation at each locus and for graphical presentation of results. On the other hand, due to their distributional properties, M-values were selected for statistical testing [90, 91].

DNAm is inherently cell-type specific, and measurements obtained from mixed cell populations such as PWBC can be confounded by variations in cell

proportions. To mitigate this, surrogate estimates of cell proportions were obtained using the Houseman method and incorporated as covariates in our models [92]. We also performed a correlation analysis [93] between the principal components and potential confounders, detecting a batch effect in DNAm data (Supplementary Figs. 1 and 2). To address this, we removed the variance attributable to batch by extracting residuals from a linear mixed-effects model in which batch was included as a random effect [94–96]. This residuals's method based correction effectively eliminated the batch effect [97], as confirmed by post-correction assessments (Supplementary Figs. 1 and 2). All subsequent EWAS models were then fitted using these batch-corrected methylation values and were adjusted for age, sex, recruitment centre and estimated white blood cell proportions.

Throughout the bioinformatics pipeline, each stage—encompassing QC, preprocessing, normalisation, and the estimation of blood cell proportions—was conducted in accordance with the recommended protocols [65, 67, 84]. All described pipeline were performed in R environment [85].

Bioinformatic analysis

After pre-processing, 809,381 CpG site probes that passed quality filters were selected for analysis. The experimental design adopted for DNAm bioinformatics analysis of IR is presented in Fig. 2. Participants were classified into five longitudinal groups (G1–G5) according to their obesity and IR status at the prepubertal baseline and pubertal follow-up visits. For clarity, we use the following abbreviations throughout the manuscript: Nw (normal-weight) and Ov/Ob (overweight/obesity). The five longitudinal groups were defined as follows: G1: Nw and non-IR at both time points, G2: Ov/Ob and non-IR at baseline remaining non-IR at follow-up, G3: Ov/Ob

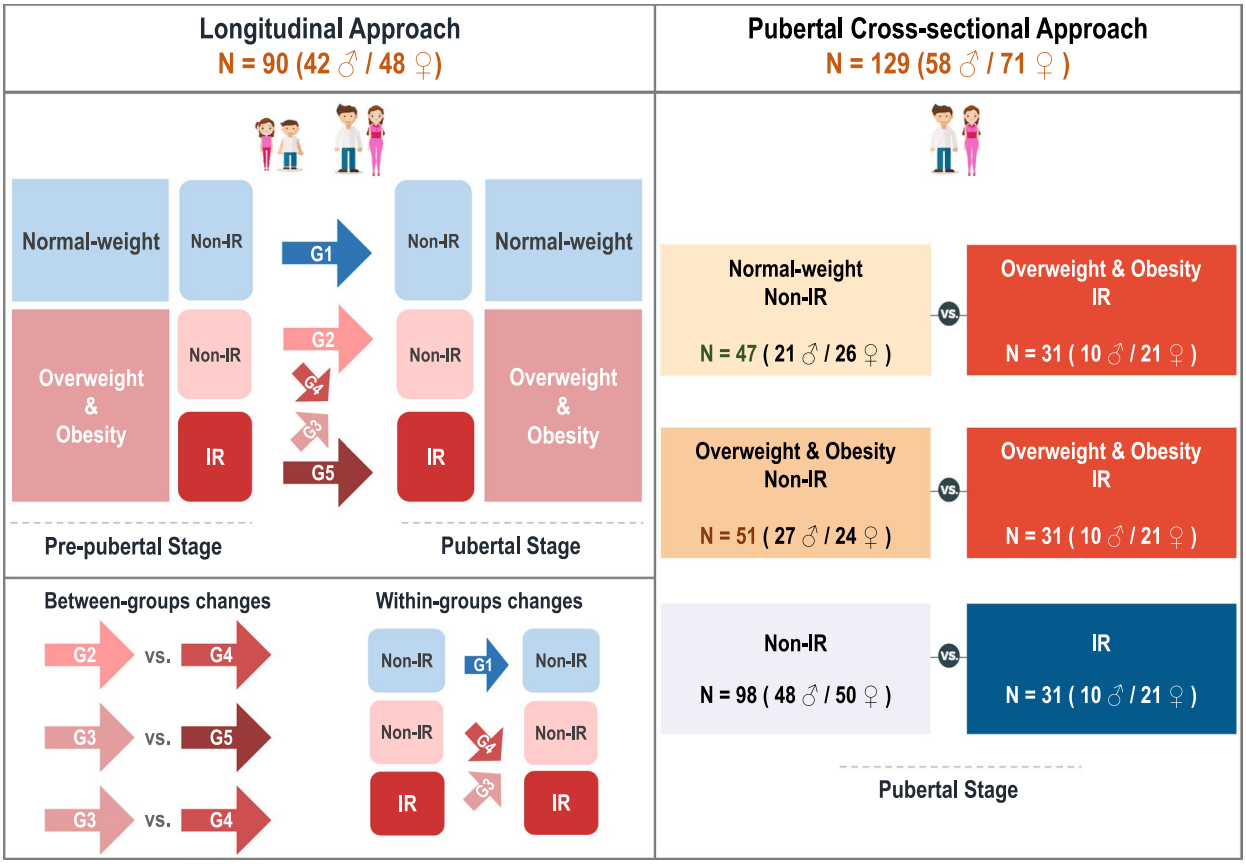

Fig. 2 Longitudinal and pubertal cross-sectional analytical designs in EWAS of IR. The longitudinal approach included 90 Spanish children (42 males and 48 females), classified into five experimental groups (G1 G5) based on their obesity and IR trajectories from prepuberty (baseline) to puberty (follow-up). G1 comprises normal-weight, non-IR children at both time points, serving as a control group. G2 and G4 consist of children with overweight/obesity who were non-IR at baseline and either remained non-IR (G2) or developed IR (G4) by follow-up. In contrast, G3 and G5 include children with overweight/obesity who were IR at baseline and either transitioned to non-IR (G3) or remained IR (G5) by follow-up. The pubertal cross-sectional approach included 129 pubertal children (58 males and 71 females) classified into three groups based on their obesity and IR status: normal-weight non-IR, overweight/obesity non-IR, and overweight/obesity IR. Additionally, a comparison was conducted between all non-IR children and those with IR. Longitudinal and pubertal cross-sectional comparisons were performed to assess epigenetic differences associated with IR. All analyses were adjusted for age, sex, recruitment centre, and white blood cells proportions

and IR at baseline; improving to non-IR at follow-up, G4: Ov/Ob and non-IR at baseline; progressing to IR at follow-up and G5: Ov/Ob and IR at both time points. Additional body weight-trajectory patterns (e.g., Nw to Ov/Ob or the reverse), as well as a few cases of Nw children with IR, were observed only in a small number of participants and, due to their insufficient representation, these individuals were excluded from the analyses. The primary objective of this analysis was to identify the DNAm patterns associated with improvement, worsening or stability of the IR status in children with Ov/Ob during the onset of puberty.

In the longitudinal approach (N = 90), methylation changes between the prepubertal baseline and the pubertal follow-up were evaluated through both within-group and between-group comparisons. The within-group analysis performed in G1 (Nw non-IR to Nw non-IR) was used to identify CpG sites whose methylation changes reflect normal growth and pubertal development rather than metabolic alterations. CpG sites identified in this comparison were therefore excluded from all subsequent analyses. All remaining longitudinal comparisons were designed to investigate DNAm patterns associated with IR during pubertal development, each one addressing a distinct biological question. The within-group analyzes in G3 (Ov/Ob IR to non-IR) and G4 (Ov/Ob non-IR to IR) evaluate intra-individual methylation changes linked to IR improvement and IR worsening across puberty, respectively. These comparisons capture epigenetic shifts that accompany opposite IR transitions within the same individuals. Complementary between-group comparisons were performed to assess whether children starting from the same baseline IR status diverge epigenetically according to their IR trajectory. The comparison between G2 (Ov/Ob non-IR remaining non-IR) and G4 (Ov/Ob non-IR to IR) identifies methylation differences associated with the worsening of IR among children initially classified as non-IR. Conversely, the comparison between G3 (Ov/Ob IR to non-IR) and G5 (Ov/Ob IR remaining IR) reveals CpG sites associated with IR improvement among children who were IR at baseline. Additionally, the G3 (Ov/Ob IR to non-IR) versus G4 (Ov/Ob non-IR to IR) comparison contrasts children with opposite IR transitions, providing insight into methylation patterns specifically related to the directionality of IR change over puberty. In addition to the longitudinal analyses, we also performed cross-sectional comparisons in the pubertal population (N = 129). Three comparisons were conducted to capture complementary aspects of IR-associated methylation: Nw non-IR versus Ov/Ob IR to identify CpG sites reflecting the combined signature of adiposity and IR; Ov/Ob non-IR versus Ov/Ob IR to isolate CpG differences specifically linked to IR under similar degrees of adiposity; and all non-IR versus all IR participants to evaluate methylation changes associated with IR irrespective of adiposity.

To evaluate whether the study was adequately powered to detect DNAm differences, we performed post-hoc power simulations using the pwrEWAS R package [98]. These simulations were used to estimate the minimum sample size required for detecting significant differential CpG sites under varying effect-size assumptions, and they indicated that statistical power is strongly dependent on the magnitude of the expected methylation changes (Supplementary Fig. 3). Although the longitudinal cohort comprised 90 participants, each individual contributed repeated measurements at two time points (180 observations), which provides additional power intrinsic to the within-subject design and enhances our ability to detect biologically meaningful DNAm differences.

These analyses were implemented using linear models from the limma R package [99], with M-values as the outcome and experimental group as a categorical predictor. The limma models were appropriately adjusted for confounders, including age, sex, recruitment centre, and white blood cell proportions. For each CpG site, an iterative linear model was fitted using lmFit function, and contrasts corresponding to the specific comparisons of interest were applied through contrasts.fit function. In the longitudinal analyses, we accounted for within-subject variability using the duplicateCorrelation function. This method estimates the intra-subject correlation in repeated measures through a mixed linear model fitted for each CpG site, and provides a consensus correlation used as input for the lmFit function. The subsequent fitting of the model was performed using generalized least squares to appropriately account for the correlation between repeated observations of the same individual. Finally, statistical inference was performed using eBayes function, which computes moderated t-statistics by shrinking probe-wise standard errors toward a global empirical Bayes estimate. This moderation improves the stability of variance estimates, particularly in small to moderate sample sizes. Moderated t-tests were used to evaluate each contrast, while moderated F-statistics tested whether all contrasts for a CpG were simultaneously equal to zero, analogous to ANOVA but with empirically moderated residual variances [99]. In addition, genomic inflation factors (λ) were computed for each comparison, using bacon R package [100], to assess the presence of test-statistic inflation and to ensure that our results were not driven by systematic bias or deviations from the expected null distribution. These quality-control metrics substantiate that the observed associations reflect true biological differences rather than analytical artifacts, confirming the absence of inflation or systematic bias (Supplementary Table 1). Given the large number of statistical tests performed, false discovery

rate (FDR) was controlled using the Benjamini Hochberg procedure [101] applied once to the pooled set of raw p-values from all comparisons, including all longitudinal and cross-sectional analyses. This approach ensured control of the expected FDR across the entire set of CpG site comparison tests, while exploiting the information contained in the global p-value distribution and thereby minimising the risk of false-positive findings arising from multiple testing.

Beyond single CpG level associations, commonly referred to as differentially methylated positions (DMPs), we also explored regional methylation differences to capture coordinated DNAm changes across contiguous CpG sites. Differentially methylated regions (DMRs) are genomic segments composed of multiple adjacent DMPs that often share regulatory functions and can provide stronger biological insights than isolated CpG sites. To identify DMRs, we fitted linear models using a design matrix that included pubertal IR status as the main outcome and age, sex, recruitment centre and estimated white blood cell proportions as counfounders. Analogous to the modelling strategy applied in the limma framework, we contrasted M-values between non-IR and IR pubertal children. We defined candidate DMRs as genomic windows of 1,000 bp containing at least two adjacent DMPs, a commonly used criterion that captures most regulatory gene features [65]. Following current recommendations for DMR identification [65], regional analysis was conducted using the DMRcate R package [102]. P-values are computed by modelling the kernel-smoothed test statistic as a scaled chi-square variable (via the Satterthwaite approximation), comparing it to the corresponding chi-square distribution, and applying Benjamini Hochberg correction to obtain the smoothed FDR [102].

After, we applied iterative linear models to examine the association between DNAm levels of the statistically significant CpG sites and continuous measures of obesity (BMI, WC and FMI) and cardiometabolic health (glucose, insulin, HOMA-IR, QUICKI, DBP, SBP, TAG, HDL-c and LDL-c), analysing the prepubertal and pubertal cross-sectional groups separately. These additional analyses aimed to identify CpG sites that, beyond their association with IR, also exhibited statistical relationships with other obesity-associated cardiometabolic alterations. To reduce the risk of spurious associations driven by presence of outliers, cardiometabolic outcomes were winsorized. For this analysis, we used the lm and Winsorize functions. All analyses were adjusted for age, sex, recruitment centre, and white blood cell proportions. Due to the number of outcomes analysed, FDR correction was applied using the Benjamin Hochberg procedure [101] to adjust for multiple hypothesis testing.

Pathway enrichment analysis

Selected CpG sites were further annotated using the IlluminaHumanMethylationEPICanno.ilm10b4.hg19 R package [103]. Genes associated with each CpG site were obtained using the getMappedEntrezIDs function from missMethyl R package [104]. Genomic and epigenomic datasets were functionally annotated based on GO and KEGG ontologies, utilising Entrez gene identifiers and the database org.Hs.eg.db [105].

The gometh function was employed to determine the enrichment of CpG-annotated genes in KEGG (Kyoto Encyclopedia of Genes and Genomes) terms, biological pathways, and cellular and molecular functions [106]. This function accepts a character vector of significant CpG sites, maps these sites to Entrez Gene IDs, and tests for GO (Gene Ontology) term or KEGG pathway enrichment using a Wallenius' noncentral hypergeometric test [107, 108], considering the number of CpG sites per gene on the Infinium Methylation EPIC 850K array. All GO and KEGG terms were tested and FDR correction was calculated [101].

mQTL analysis

Although DNAm serves as an epigenetic mark that modulates gene expression in response to environmental stimuli, certain SNPs may influence the methylation of CpG sites, whether they are located proximally or distally to their genomic positions. This interplay between SNPs and DNAm is crucial for deciphering epigenetic signatures related to IR, particularly in instances where the external exposome may not serve as the primary causal mediator but rather interacts with genetic factors. Consequently, mQTL analysis was conducted to elucidate the genetic regulation of DNAm associated with IR. Importantly, no GWAS was performed in this work; genotyping data were used exclusively to evaluate whether the DNAm levels of the CpG sites identified in the EWAS analyses were influenced by genetic variants through mQTL analysis.

To investigate whether the methylation levels of the selected CpG sites were influenced by underlying genetic variation, we explored their association with SNPs through mQTL analyses. First, we queried the GoDMC (Genetics of DNA Methylation Consortium) catalogue [109], a large-scale reference resource that aggregates harmonized mQTL results from multiple population-based studies and provides robust evidence for cis- and trans-mQTL effects across the epigenome. This external lookup allowed us to evaluate whether the CpG sites associated with IR in our EWAS analyses had been previously reported as genetically regulated. In parallel, we performed an internal mQTL analysis within our own cohort to assess the presence of SNP CpG associations in our study population. These two complementary

approaches—external validation using the GoDMC atlas and internal testing using our genotyping data—enabled a comprehensive assessment of the genetic contribution to DNAm variation at the CpG sites identified in this study.

For internal mQTL analysis within our own cohort, we used a linear model implemented with the MatrixEQTL R package [110]. In the linear model, DNAm values were modelled as the outcome and SNPs were encoded as 0, 1, or 2 according to the number of minor alleles (following the additive genetic model), and age, sex, and recruitment centre was included as covariates. To distinguish between cis-mQTLs and trans-mQTLs associations, a boundary of 500 bp between SNPs and CpG sites was used to define cis-mQTLs; all other SNP-CpG pairs were considered as trans-mQTLs. P-values were adjusted using a correction factor for multiple testing that accounts for the dependency among SNPs due to linkage disequilibrium (LD) via LD-based pruning, reflecting the total number of independent tests. The correction value for the trans-analysis was calculated as the total number of analysed CpG sites multiplied by the number of SNPs in the whole dataset, after FDR correction for such number of tests correction was calculated [101].

Statistical analysis

Statistical analysis was performed using the R environment (version 4.3.1). Normality of all continuous non-omics variables was assessed using the Shapiro Wilk test [111]. Categorical variables were evaluated with Pearson's Chi-squared test and Fisher's exact test, while differences in continuous variables between experimental groups were examined using one-way analysis of means (not assuming equal variances) and the Kruskal Wallis rank sum test, as appropriate [112]. Differences were evaluated for both the longitudinal groups (G1 G5) and the cross-sectional groups (Nw Non-IR, Ov/Ob Non-IR, and Ov/Ob IR) at baseline and follow-up. Descriptive statistics are available in Supplementary Tables 2 (longitudinal approach), 3 (prepubertal cross-sectional approach) and 4 (pubertal cross-sectional approach).

Results

Clinical and biochemical profile of participants

We observed five experimental groups, each characterized by a distinct trajectory of obesity-related IR. Descriptive statistics for the longitudinal approach are presented in Supplementary Table 2. At both time points, there were no statistically significant differences in terms of age or sex distribution among the experimental groups. In line with the study design, the groups that developed or maintained IR during the onset of puberty (G4: Ov/Ob non-IR → Ov/Ob IR and G5: Ov/Ob IR → Ov/Ob IR) displayed the highest increases in FMI, fasting serum insulin levels and HOMA-IR compared to the non-IR groups (G1: Nw non-IR → Nw non-IR, G2: Ov/Ob non-IR → Ov/Ob non-IR and G3: Ov/Ob IR → Ov/Ob non-IR). Although SBP z-score did not differ significantly between groups at baseline, significant differences were evident at follow-up, with the highest values observed in groups (G4: Ov/Ob non-IR → Ov/Ob IR) and (G5: Ov/Ob IR → Ov/Ob IR). This finding highlights the importance of puberty as a critical period influencing metabolic health.

Additionally, G4 (Ov/Ob non-IR → Ov/Ob IR) exhibited the most pronounced increases in leptin, hs-CRP, and t-PAI-1 concentrations. In contrast, G3 (Ov/Ob IR → Ov/Ob non-IR) showed reductions in these biomarkers. Overall, the experimental groups clearly reflected the intended clinical trajectories in IR status. As would be expected with pubertal progression, LDL-c and adiponectin concentrations declined across all groups following pubertal maturation onset. Cross-sectional groups similarly demonstrated coherent and expected differences across anthropometric, metabolic, inflammatory, and cardiovascular biomarkers (Supplementary Tables 3 and 4).

White blood cells DNAm associated with IR

To comprehensively characterise epigenetic marks associated with IR, we evaluated multiple longitudinal and cross-sectional comparisons, each addressing a complementary aspect of the clinical evolution of IR across the pubertal transition (Fig. 2). Although all comparisons targeted the same overarching scientific question, DNAm associated with IR during puberty, they captured distinct IR trajectories (improvement, worsening, or stability) as well as cross-sectional differences in the pubertal population.

Given the large number of statistical tests performed, FDR was controlled using the Benjamini Hochberg procedure [101] applied once to the pooled set of raw p-values from all comparisons, including all longitudinal and cross-sectional analyses (see Section 3.4.2, Bioinformatic analysis, for details). In total, 120 CpG sites surpassed the global FDR threshold (FDR < 0.05) and were considered associated with IR. These CpG sites are listed in Supplementary Table 5, whereas Supplementary Table 6 presents the complete output for all CpG sites and all comparisons, including both significant and non-significant results.

For interpretative clarity, the FDR-significant CpG sites were grouped into three categories: (1) 22 CpG sites from longitudinal within-group comparisons, (2) 93 CpG sites from longitudinal between-group comparisons and (3) 9 CpG sites from pubertal cross-sectional contrasts. These results are summarised in Fig. 3. Notably, 85 out of the 120 significant CpG sites belonged to the comparison

between group G3 (Ov/Ob IR → non-IR) *vs.* G5 (Ov/Ob IR → IR). Within the longitudinal analyses, children whose IR worsened (G4: Ov/Ob non-IR → IR) tended to exhibit increasing hypermethylation at several CpG sites, whereas those whose IR improved (G3: Ov/Ob IR → non-IR) showed opposite hypomethylation trajectory patterns. These trends were consistent with cross-sectional findings, where all FDR-significant CpG sites were hypomethylated in non-IR pubertal children.

Among the 120 global FDR-significant CpG sites, 42 were located in CpG islands, of which only six were annotated to promoter regions. These promoter-associated CpG sites were cg16147221 annotated in *SLC2A9*, cg03994651 annotated in *LMTK3*, cg20196543 annotated in *SACM1L*, cg14166149 annotated in *ZNF341*, cg03962365 annotated in *NR4A1*, and cg15063951 annotated in *FAR1* (Supplementary Table 5). A region-level analysis of the CpG sites identified, comparing pubertal non-IR *vs.* IR children, yielded 195 FDR-significant DMRs, mapping 174 loci (Supplementary Table 7). GO and KEGG pathway enrichment analyses of the genes annotated to the 120 FDR-significant CpG sites highlighted several biologically relevant pathways, including cortisol synthesis and secretion, ATP-dependent chromatin remodelling, and aldosterone synthesis and secretion. These pathways met the nominal significance threshold ($p<0.05$) but did not remain significant after FDR correction (Supplementary Tables 8 and 9).

Associations between DNAm and overall metabolic health

Cross-sectional associations between DNAm levels at the 120 IR-associated CpG sites and continuous cardiometabolic outcomes (BMI z-score, WC z-score, FMI, glucose z-score, insulin z-score, HOMA-IR z-score, QUICKI, DBP z-score, SBP z-score, TAG z-score, HDL-c z-score, and LDL-c z-score) were cross-sectionally evaluated in

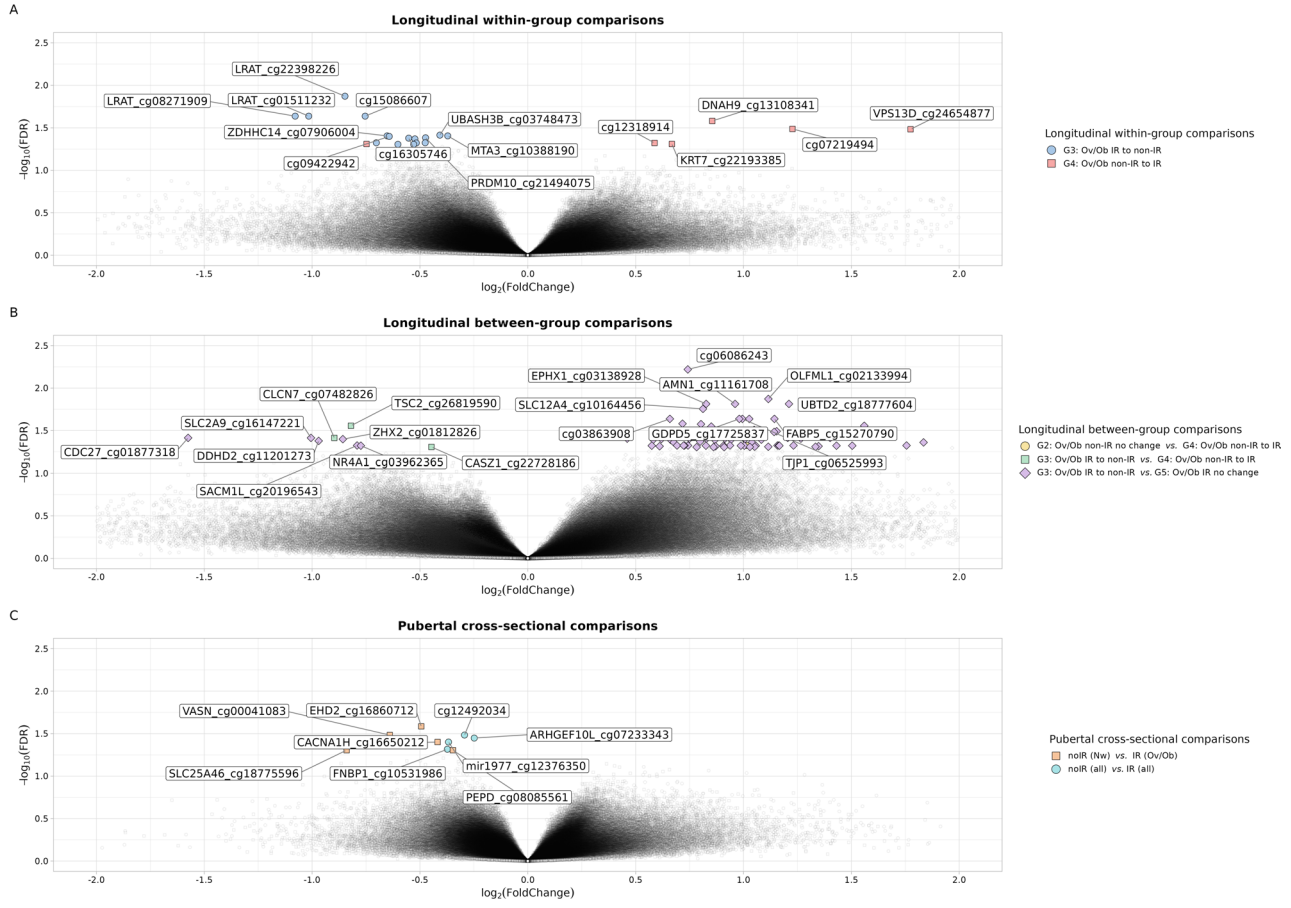

Fig. 3 Multipanel volcano plots illustrating the results of the EWAS analyses (longitudinal within-group, longitudinal between-group, and pubertal cross-sectional comparisons) fully described in Fig. 2. **A** Longitudinal within-group comparisons: G3 (Ov/Ob IR → Ov/Ob non-IR) and G4 (Ov/Ob non-IR → Ov/Ob IR). **B** Longitudinal between-group comparisons, showing each contrast explicitly: 1) G2 (Ov/Ob non-IR with no change) vs. G4 (Ov/Ob non-IR → IR), 2) G3 (Ov/Ob IR → non-IR) vs. G4 (Ov/Ob non-IR → IR), 3) G3 (Ov/Ob IR → non-IR) vs. G5 (Ov/Ob IR with no change)(C) Pubertal cross-sectional comparisons: non-IR (Nw) vs. IR (Ov/Ob); non-IR (Ov/Ob) vs. IR (Ov/Ob); and all non-IR vs. all IR participants. In each plot, the x-axis represents the log_2 fold change, and the y-axis shows the $-\log_{10}$ of the global FDR. CpG sites surpassing the significance threshold (global FDR < 0.05) are highlighted. For simplification, only the top 10 hypermethylated and top 10 hypomethylated CpG sites meeting the global FDR threshold are labelled with their CpG ID and mapped gene symbol, when available. All analyses were adjusted for age, sex, recruitment centre and white blood cells proportions

both the prepubertal and pubertal populations. These analyses were designed to further explore the relationship between the 120 global FDR-significant CpG sites, IR and other adiposity and cardiometabolic complications tipically observed in children with obesity. Global FDR-significant associations (FDR < 0.05) were observed for BMI z-score, WC z-score, glucose z-score, insulin z-score, HOMA-IR z-score, QUICKI, DBP z-score, and SBP z-score outcomes. Supplementary Table 10 report all findings, and Fig. 4 illustrates the heatmap for CpG sites that displayed at least a FDR-significant association with one continuous outcome.

Several CpG sites showed at least one FDR-significant association with pubertal IR-related outcomes (insulin z-score, HOMA-IR z-score or QUICKI). These included cg16650212 (*CACNA1H*), cg26810726 (*EGLN3*), cg16860712 (*EHD2*), cg10531986 (*FNBP1*), cg155864606 (*LINC00708*), cg03994651 (*LMTK3*), cg08086561 (*PEPD*), cg16147221 (*SLC2A9*), cg17389219 (*TJP3*), cg20681950 (*TSC2*), cg00041083 (*VASN*) and cg27416312 (*XPO6*). Examination of effect directions revealed a consistent pattern of hypermethylation associated with higher pubertal IR across nearly all CpG sites (i.e., positive associations with insulin and HOMA-IR and negative associations with QUICKI). The only exception was cg16147221 (*SLC2A9*), which showed the opposite trend, displaying lower methylation levels in adolescents with greater IR. Curiosly, cg08086561 (*PEPD*) and cg00041083 (*VASN*) were the only CpG sites that showed FDR-significant associations with both IR-related measures plus adiposity outcomes, highlighting them as key epigenetic biomarkers at the intersection of adiposity and IR during puberty. Among all genes that emerged consistently across longitudinal and cross-sectional analyses and showed associations with multiple cardiometabolic outcomes as *EHD2*, *PEPD*, *SLC2A9*, *TSC2* and *VASN* loci. For these five CpG sites, effect directions were consistently aligned with IR severity trajectories (Fig. 5).

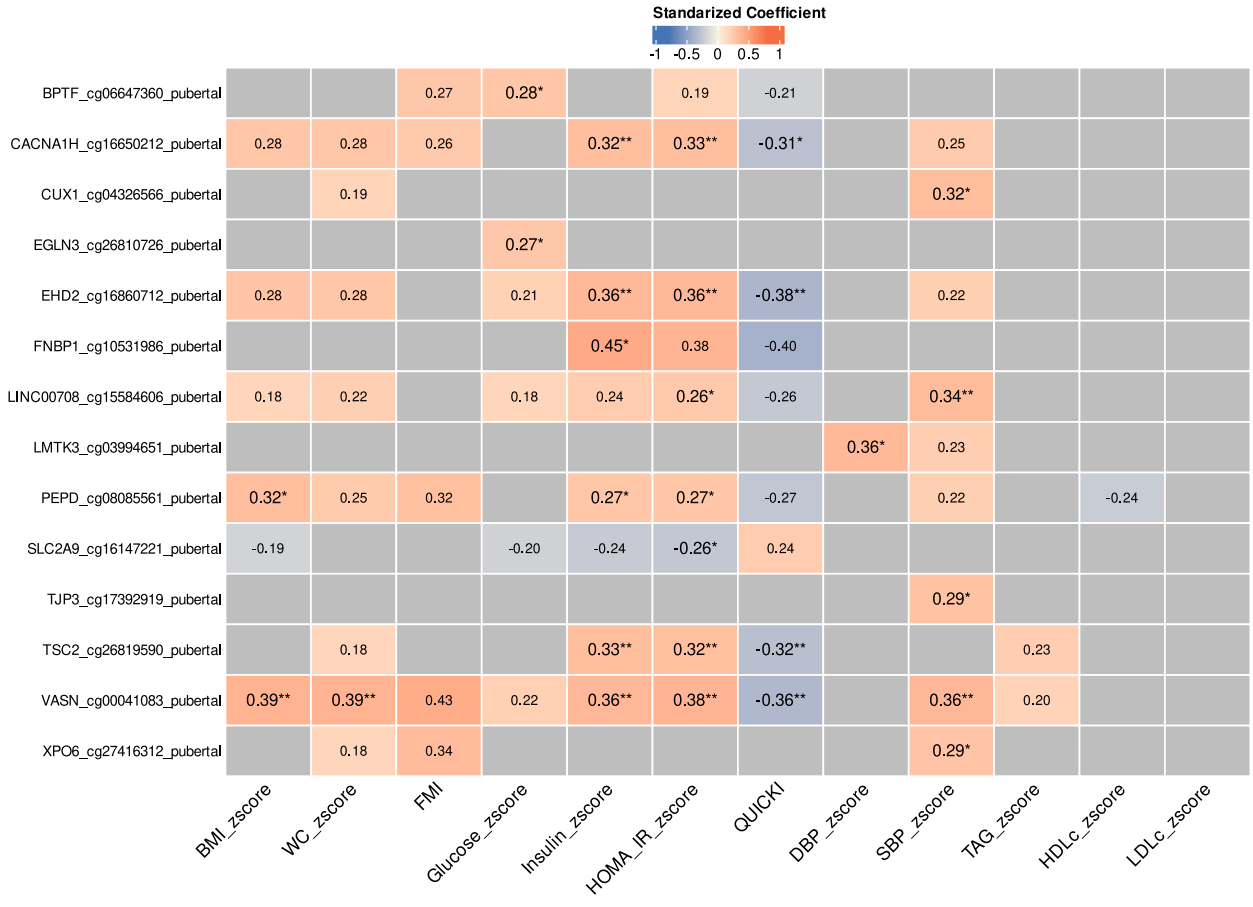

Fig. 4 Associations between DNAm and continuous cardiometabolic outcomes. The heatmap displays the standardized coefficients representing the associations between DNAm levels at selected CpG sites and various continuous metabolic traits, including BMI z-score, WC z-score, FMI, Glucose z-score, Insulin z-score, HOMA-IR z-score, QUICKI, DBP z-score, SBP z-score, TAG z-score, HDL-c z-score and LDL-c z-score. Only CpG sites exhibiting at least one significant association (FDR < 0.05) are presented. Statistical significance is indicated using asterisks: * (FDR < 0.05), ** (FDR < 0.01). If no asterisk is present but a coefficient is displayed, the association is significant at a nominal *p*-value < 0.05. Non-significant associations (raw *p*-value > 0.05) are shown in grey. All analyses were adjusted for age, sex, recruitment centre, and white blood cells proportions

Associations between genetic variants and DNAm (mQTL)

To investigate whether the IR-associated DNAm signals were influenced by underlying genetic variation, we queried the 120 FDR-significant CpG sites identified in our study against the external GoDMC catalogue, which compiles cis-mQTL information from multiple large population-based studies (Table 1). Twenty-five cis SNP CpG associations were identified in the GoDMC look-up, indicating that part of the IR-associated methylation variation may be explained by nearby genetic variants. These corresponded to 20 unique IR-associated CpG sites from our list. As suggested by our significant mQTL findings, the number of minor alleles of a given genetic variant was generally associated with lower methylation at the corresponding CpG site, although opposite effects were also observed for some loci. Most genetically regulated CpG sites (n = 11) had been originally detected in the comparison between G3 (Ov/Ob IR → Ov/Ob non-IR) *vs.* G5 (Ov/Ob IR → Ov/Ob IR). In particular, it is noteworthy to highlight cg16147221 (*SLC2A9*) and cg26819590 (*TSC2*), two CpG sites associated with IR and with longitudinal IR trajectories in our study, which showed associations with rs62409886 and rs34104130, and with rs79518513 respectively. For cg16147221 (*SLC2A9*), the minor alleles of rs62409886 and rs34104130 were associated with lower and higher methylation levels, respectively. Whereas for cg26819590 (*TSC2*), the minor allele of rs79518513 was associated with lower methylation. We also performed mQTL analyses within our own cohort, at both prepubertal and pubertal stages. However, none of the SNP CpG associations surpassed the FDR threshold (Supplementary Table 11).

Discussion

In this longitudinal EWAS, we investigated whether blood genome-wide DNAm patterns capture divergent trajectories of obesity-related IR from childhood into puberty. Leveraging a deeply phenotyped cohort of children with paired prepubertal and pubertal blood samples, complemented with cross-sectional analyses at each stage, we focused specifically on IR status and trajectories within the context of Ov/Ob. Across multiple longitudinal and cross-sectional contrasts, we identified 120 differentially methylated CpG sites associated with IR after a global FDR correction. These CpG sites delineated clear methylation patterns accompanying IR improvement,

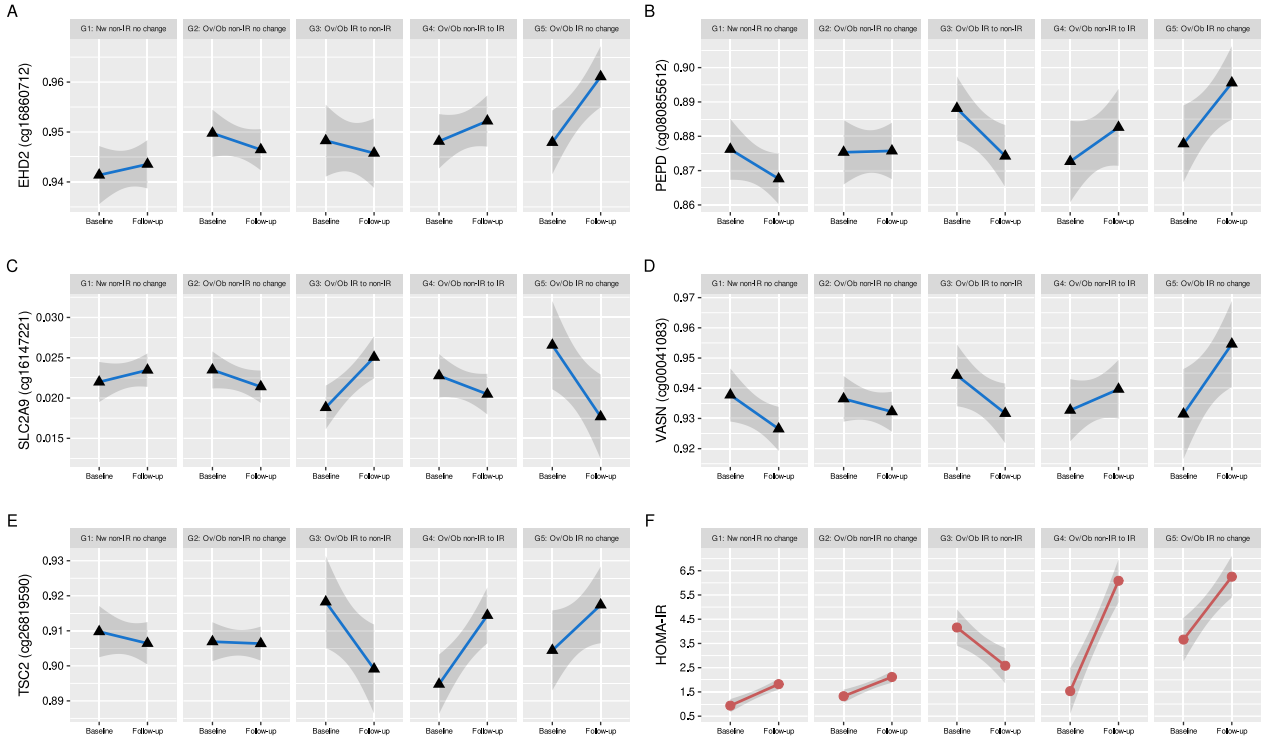

Fig. 5 Multipanel longitudinal trajectories of DNAm and HOMA-IR across longitudinal groups. This figure displays the temporal patterns of DNAm (β-values) or HOMA-IR at baseline and follow-up, with the x-axis indicating both time points and the y-axis representing either methylation levels for each IR-associated CpG site or HOMA-IR values. Panels **A**–**E** illustrate DNAm trajectories for five IR-associated CpG sites: **A***EHD2* (cg16860712), **B***PEPD* (cg08086561), **C***SLC2A9* (cg16147221), **D***VASN* (cg00041083), and **E***TSC2* (cg26819590). Panel **F** displays the corresponding longitudinal trajectories of HOMA-IR. Faceting by longitudinal groups (G1–G5) enables visualization of methylation and HOMA-IR changes according to obesity and IR trajectories: G1 (Nw non-IR → Nw non-IR), G2 (Ov/Ob non-IR → Ov/Ob non-IR), G3 (Ov/Ob IR → Ov/Ob non-IR), G4 (Ov/Ob non-IR → Ov/Ob IR) and G5 (Ov/Ob IR → Ov/Ob IR). Across CpG sites, these plots reveal consistent patterns of hypo- or hypermethylation that align with the direction of IR change—improvement, worsening, or stability—mirroring the metabolic transitions reflected by HOMA-IR during puberty

Table 1 Identification of mQTLs from FDR-significant 120 CpG sites using GoDMC catalogue

CpG site	SNP	Gene	Comparison	Num studies	Sample size	P-value	Beta
cg10531986	rs56224285	*FNBP1*	Pubertal non-IR vs. IR	35	27246	0	0.92
cg22099241	chr6:31067074:D	*MICA*	G3 vs. G5	26	19095	0	0.08
cg25066224	rs1004490	*SLC47A2*	G3 vs. G5	35	27232	0	−0.06
cg03138928	rs4653680	*EPHX1*	G3 vs. G5	31	23646	0	−0.06
cg16147221	rs62409886	*SLC2A9*	G3 vs. G5	18	13946	0	−0.08
cg22099241	rs28771422	*MICA*	G3 vs. G5	18	13381	0	-0.11
cg18193726	rs73454919	*STX2*	G3 vs. G5	33	25901	0	-0.12
cg22193385	rs7970078	*KRT7*	G4: Ow/Ob non-IR to IR	36	27742	0	−0.12
cg02678476	rs75033680	–	G3 vs. G5	27	21036	0	−0.15
cg16244711	rs10458345	*SIPA1L2*	G3 vs. G5	33	25906	0	−0.21
cg03863908	rs7734252	–	G3 vs. G5	32	23526	0	−0.70
cg03886085	rs2246614	*MUPCDH*	G3 vs. G5	32	25549	1.59e-223	−0.29
cg14395885	rs3739821	*DPM2*	G3 vs. G5	35	27257	3.99e-147	−0.26
cg03994651	rs1320401	*LMTK3*	G3: Ow/Ob IR to non-IR	25	16364	1.53e-130	−0.78
cg08271909	rs201822	*LRAT*	G3: Ow/Ob IR to non-IR	35	27223	6.339e-83	−0.17
cg16147221	rs34104130	*SLC2A9*	G3 vs. G5	34	26476	2.297e-82	0.19
cg07233343	rs12562076	*ARHGEF10L*	Pubertal non-IR vs. IR	36	27746	1.79e-41	0.15
cg01511232	rs12512714	*LRAT*	G3: Ow/Ob IR to non-IR	34	26339	1.027e-43	−0.12
cg10531986	rs113335566	*FNBP1*	Pubertal non-IR vs. IR	33	25825	5.767e-42	0.26
cg07219494	rs384227	–	G4: Ow/Ob non-IR to IR	29	23335	9.356e-61	0.17
cg22398226	rs1910849	*LRAT*	G3: Ow/Ob IR to non-IR	34	26362	3.246e-32	−0.10
cg02678476	rs6858295	–	G3 vs. G5	34	26471	8.034e-15	−0.07
cg26819590	rs79518513	*TSC2*	G3 vs. G4	19	13013	2.604e-12	−0.14
cg05624582	rs2899319	*CACNA1I*	G3 vs. G5	30	23791	2.438e-08	0.05
cg22099241	chr6:31535381:I	*MICA*	G3 vs. G5	18	13765	1.007e-08	0.07

deterioration, or stability, and several of them were additionally associated with cardiometabolic traits such as adiposity indices, blood pressure and glycaemic markers. Previous evidence from large-scale mQTL studies further suggests that only a subset of these IR-associated marks is likely to be partially driven by nearby genetic variants. Altogether, our findings indicate that DNAm signatures in childhood obesity not only mirror current IR status but also track its dynamic evolution across the pubertal transition, raising the possibility that some of these epigenetic alterations might be amenable to modification alongside metabolic improvement.

Our study design was explicitly conceived to isolate epigenetic marks related to obesity-associated IR in the context of pubertal transition. Both the longitudinal groupings (G1–G5) and the cross-sectional comparisons were defined around clinical IR as the main outcome, while BMI, FMI, WC, lipids, blood pressure, glucose and IR indices were evaluated as continuous traits to refine cardiometabolic interpretation. By pooling all nominal p-values from longitudinal and cross-sectional analyses into a single distribution and applying a global FDR correction, we obtained a conservative set of 120 IR-associated CpG sites (Fig. 3). These signals arose from three complementary analytical layers: (1) 22 CpG sites from longitudinal within-group comparisons, capturing intra-individual methylation changes from prepruberty to puberty in each trajectory group; (2) 89 CpG sites from longitudinal between-group comparisons, predominantly arising from contrasts between children with improving IR throughout puberty development (G3) and those with either persistent (G5) or incident IR (G4), thereby highlighting loci that distinguish divergent IR trajectories across puberty; and (3) 9 CpG sites from cross-sectional contrasts of pubertal non-IR versus IR participants. All three dimensions relate to IR in the metabolically active window of puberty, yet the longitudinal contrasts—especially the comparison between children who improved from IR to non-IR (G3) and those who remained persistently Ov/Ob and IR (G5)—were particularly informative, yielding the largest number of FDR-significant CpG sites. Interestingly, children whose IR worsened tended to show hypermethylation at IR-associated CpG sites, whereas those whose IR improved showed opposing hypomethylation patterns, consistent with the cross-sectional finding that non-IR adolescents were invariably more hypomethylated at these CpG sites. Furthermore, a subset of 14 from these 120 IR-associated CpG sites were also robustly linked to higher adiposity, blood pressure or adverse glycaemic profiles at puberty, suggesting that these marks lie at the intersection of multiple cardiometabolic disturbances rather than reflecting IR in isolation (Fig. 4). In line with this, a region-level analysis comparing pubertal non-IR and IR children identified 195

FDR-significant DMRs highlighting two regions mapping to *EHD2* and *VASN*. Consistent with the DMP-level findings, our finding indicated that IR is accompanied by coordinated methylation shifts across specific genomic regions rather than isolated single-site changes.

Our analysis also highlighted several well-established epigenetic loci previously linked to obesity and T2DM in adults, though these loci did not surpass our predefined global FDR threshold (i.e., only significant if we considered the nominal p-value) (Supplementary Table 5). In particular, we identified sites mapping to *ABCG1* and *HDAC4*, two of the most consistently replicated DNAm markers of adiposity, dyslipidaemia and incident T2DM in adult EWAS and candidate-gene studies [113–116]. We also detected *PTPRN2*, whose methylation has been associated with incident CVD and T2DM [117, 118], and *ADCY5*, a canonical T2DM locus repeatedly highlighted in adult epigenetic and genetic studies [119]. Complementarily, our DMR analysis reported FDR-significant differential methylation across regions annotated into the *ABCG1*, and *HIF3A* whose blood and adipose-tissue methylation has been associated with adipose dysfunction in adults [120–122]. Notably, several CpG sites previously associated with IR in adult populations [42, 44, 45, 56, 58] also showed consistent directions of association with IR in our paediatric cohort, though again none of these surpassed our global FDR correction (please, see Supplementary Table 12 for details). Together, the convergence of these established cardiometabolic loci with our paediatric signals suggests that key components of the molecular architecture of IR are already present by late childhood, and that the weaker statistical evidence observed here may reflect the earlier disease stage in our cohort, before overt cardiometabolic complications have fully developed.

Beyond confirming known biology, our results highlight a subset of five loci that, to our knowledge, have not been previously reported in the context of pediatric IR. These loci showed converging evidence across longitudinal and cross-sectional analyses, as well as across multiple cardiometabolic outcomes, and an in-depth literature review of their biological functions supported a plausible link with IR (Figs. 4, 5). These include CpG sites annotated to genes involved in hypoxia sensing (*EGLN3*), extracellular matrix remodelling (*PEPD*), nutrient and urate transport (*SLC2A9* (also known as *GLUT9*)), vascular signalling (*VASN*), mTOR regulation (*TSC2*), and caveolar dynamics (*EHD2*). Notably, several of these loci were discovered in the most biologically informative contrasts of our design. CpG sites annotated to *VASN*, *SLC2A9* and *EGLN3* emerged from the longitudinal comparison between G3 (Ov/Ob IR → Ov/Ob non-IR) and G5 (Ov/Ob IR → Ov/Ob IR), capturing epigenetic differences between children with obesity who improve IR versus those who remain persistently insulin-resistant during puberty. Furthermore, *TSC2* was detected in the longitudinal comparison between G3 (Ov/Ob IR → Ov/Ob non-IR) and G4 (Ov/Ob non-IR → Ov/Ob IR), directly contrasting opposite IR trajectories within the Ov/Ob group. Finally, *PEPD* and *EHD2,* and again *VASN*, were identified in the pubertal cross-sectional comparison between Nw non-IR and Ov/Ob IR youth, thus reflecting methylation differences associated with an established IR phenotype at this critical developmental stage. For all five loci, effect directions were coherent with IR severity (Fig. 5) and, in several cases, with adiposity, blood pressure and glycaemic traits (Fig. 4), indicating that these CpG sites do not simply tag excess body weight but instead mark dynamic biological processes underlying the transition between metabolically healthier and metabolically unhealthy obesity, and, importantly, the reverse transition as well. This tight coupling between methylation change and clinical course, together with the molecular functions of the implicated genes in nutrient signalling, urate transport, hypoxia sensing, extracellular matrix remodelling and vascular regulation, suggests that these epigenetic marks are plausibly embedded in the mechanisms driving obesity-related IR. While we cannot infer causality, the observation that DNAm at these loci shifts in parallel with IR improvement raises the possibility that at least part of this epigenetic signature may be modifiable alongside metabolic recovery, thereby pointing to potential epigenetic entry points for future therapeutic or lifestyle interventions.

Among these candidates, vasorin (*VASN*) stands out as a particularly attractive translational target. In our data, hypermethylation at cg00041083 (*VASN*) was associated with higher pubertal IR and adverse adiposity measures, and methylation levels tracked IR trajectories across the longitudinal groups (G3 (Ov/Ob IR → Ov/Ob non-IR) *vs.* G5 (Ov/Ob IR → Ov/Ob IR)) (Supplementary Table 5). *VASN* encodes a TGF-β binding type I transmembrane glycoprotein implicated in vascular remodelling, protection against oxidative stress and modulation of inflammatory responses, particularly under hypoxic conditions relevant for expanding adipose tissue. Although initially described in hepatic and vascular disease, its role in metabolic tissues is largely unexplored. Importantly, *VASN* is a secreted protein readily measurable in serum, which, together with its vascular and adipose-related functions, supports its candidacy as a biologically plausible modulator of metabolic health at the interface of vascular function, inflammation and adipose tissue architecture in pediatric obesity. If validated in independent cohorts and linked to prospective outcomes, such biomarker might help identify children at particularly high risk of progressing from obesity with compensated IR to overt dysglycaemia and cardiometabolic disease.

Our results also support a role for transport of glucose-metabolism relevant macromolecules (notably urate) and extracellular matrix (ECM) remodelling pathways as candidate mechanistic nodes. Hypomethylation at cg16147221 in the promoter of *SLC2A9* was associated with greater pubertal IR and, unlike most other CpG sites, lower methylation tracked worse metabolic profiles. *SLC2A9* encodes the urate and fructose transporter GLUT9, a major determinant of serum uric acid levels, as demonstrated in large genetic studies showing that *SLC2A9* variants explain a substantial proportion of interindividual variation in uricemia and gout risk [123–127]. Serum uric acid rises early in youth with obesity and has been associated with features of the MetS, IR and later hypertension and cardiometabolic risk [128–131]. In this context, the presence of cis-mQTL signals for cg16147221 in external resources suggests that both inherited and acquired factors may converge on this locus to shape the urate IR axis. Similarly, cg08086561 in *PEPD*, the gene encoding prolidase, a key enzyme in collagen turnover, was associated with both IR-related measures and adiposity in pubertal children, suggesting that ECM remodelling in expanding adipose tissue is tightly linked to systemic insulin sensitivity. Prolidase (PEPD) catalyses the rate-limiting step of collagen recycling and is essential for collagen metabolism and matrix remodelling [132–136]. Recent experimental work further shows that dysregulation of macrophage PEPD in obesity drives adipose tissue fibro-inflammation and IR, reinforcing the causal link between defective collagen remodelling and metabolic dysfunction [137, 138]. Given that PEPD-generated proline and collagen degradation products influence adipose architecture and inflammatory tone, methylation or activity readouts at this locus may therefore reflect the balance between healthy adipose expansion and fibro-inflammatory remodelling in childhood obesity.

Other loci highlighted by our findings further underscore the involvement of nutrient sensing, oxygen supply and cellular architecture in the early stages of metabolic dysfunction. Methylation at a CpG within *TSC2*, a central inhibitor of mTORC1 signalling, was associated with pubertal IR and IR trajectories, and this site harboured a cis-mQTL in GoDMC, consistent with genetic influences on the mTOR axis. *TSC2* forms, together with *TSC1*, the tuberous sclerosis complex that restrains mTORC1 activity in response to nutrient and growth factor cues; chronic mTORC1 overactivation is known to impair insulin signalling via feedback inhibition of IRS proteins and to promote metabolic dysfunction in obesity [139–143]. Epigenetic modulation at *TSC2* in blood may therefore serve as a surrogate marker of systemic "mTOR tone" and chronic nutrient surplus in metabolically stressed children. Likewise, CpG sites annotated to *EGLN3* (PHD3), a key oxygen sensor that regulates HIF stability, align with the concept that rapid adipose tissue expansion during childhood obesity may outstrip its vascular supply, creating hypoxic, pro-inflammatory conditions that favour IR; PHD/HIF signalling has been implicated in adipose tissue hypoxia, inflammation and altered insulin sensitivity [144–146]. In addition, loci in *EHD2* point to complementary mechanisms at the plasma membrane: EHD2 stabilises caveolae and regulates their dynamics, thereby influencing lipid uptake, membrane organisation and insulin receptor signalling at the cell surface; reduced EHD2 function has been linked to altered fatty acid uptake and obesity-related phenotypes in experimental models [147, 148].

The origin of the IR-associated methylation marks identified in this study is likely multifactorial, reflecting both genetic predisposition and environmental exposures. By querying our 120 FDR-significant CpG sites in the GoDMC catalogue, we identified 25 cis SNP CpG associations corresponding to 20 unique CpG sites, indicating that a subset of the observed methylation variation might be partly influenced by nearby genetic variants. Notably, several of these genetically influenced CpG sites were located in loci that we highlight as biologically compelling, including cg16147221 (*SLC2A9*) and cg26819590 (*TSC2*), supporting a model in which inherited variants in key metabolic genes shape baseline methylation states and, potentially, individual susceptibility to IR. However, our own mQTL analyses within PUBMEP did not yield FDR-significant associations, likely due to limited sample size, and most of the 120 CpG sites did not show robust cis-mQTL signals in external data. Moreover, the fact that methylation at several of these loci dynamically tracked IR trajectories—worsening, improvement or stability—across puberty argues against a purely genetically predetermined pattern and instead points to an important modifiable component. Taken together, this pattern suggests that a substantial fraction of the epigenetic signal is acquired, plausibly in response to energy imbalance, adiposity, inflammation, hormonal changes during puberty or other environmental factors. Disentangling cause and consequence will require larger, multi-ethnic cohorts, repeated methylation assessments at finer temporal resolution, and integrative methods such as Mendelian randomisation and mediation analyses linking exposures, methylation and metabolic outcomes.

Several strengths of our work warrant mention. First, the longitudinal design, incorporating well-characterised IR trajectories across puberty and defined using sex- and pubertal stage- specific HOMA-IR cut-offs, offers a unique opportunity to capture dynamic cardiometabolic changes during a critical developmental window. Second, the use of a global FDR correction across all longitudinal and cross-sectional comparisons provides a conservative

and robust set of IR-associated CpG sites. Third, integrating DNAm with a broad panel of cardiometabolic traits, together with external mQTL information, enables us to move beyond traditional case control frameworks and begin delineating the underlying mechanistic landscape. From a translational perspective, the study addresses an urgent public-health need by linking epigenetic variation to clinically defined trajectories of paediatric IR—a recognised gateway to later T2DM and CVD. Importantly, the longitudinal groups were constructed using clinically validated HOMA-IR cut-offs already applied in paediatric practice, reinforcing the clinical relevance of the potential biomarkers identified.

Nevertheless, several limitations should be acknowledged. The sample size is moderate for EWAS standards, which may have reduced our ability to detect smaller methylation effects or weaker mQTLs, and our findings have not yet been replicated in independent cohorts. Although replication was explored in other longitudinal paediatric cohorts [68, 149], no suitable replication cohorts were found, because cohorts lacked repeated Tanner stage assessment and/or fasting insulin measurements. DNAm was measured in whole blood, which may not fully capture epigenetic variation in others metabolically active tissues such as liver, muscle, or adipose tissue, although systemic metabolic stress would be expected to leave at least partial signatures in circulating cells. In addition, our models were not adjusted for a number of potential confounders—such as smoking exposure, maternal BMI, circulating sex hormones, or environmental pollutants—that could also influence DNAm. We were likewise unable to assess associations with lifestyle-related environmental factors, including dietary patterns or physical activity, which might have helped clarify the potential causal pathways linking DNAm to obesity-related IR. Finally, a recognised challenge in EWAS research, including the present study, is the difficulty of replicating individual CpG sites rather than broader gene-level methylation patterns. Although multiple investigations have repeatedly implicated the same genes, replication at the exact CpG-site level remains inconsistent, highlighting the need for targeted validation and functional follow-up studies.

Conclusions

In conclusion, this longitudinal EWAS in children with obesity demonstrates that specific DNAm signatures in blood parallel the clinical evolution of IR across the pubertal transition. The contrasting methylation trajectories observed between children who resolve IR and those who remain persistently IR support the notion that at least part of the epigenetic landscape of IR is dynamic and potentially reversible. Within this landscape, we identify several loci—most notably *VASN*, *SLC2A9*, *PEPD*, *EGLN3*, *EHD2*, and *TSC2*—as promising biomarkers and mechanistic candidates that integrate adiposity, vascular function, nutrient signalling and tissue remodelling. If replicated and functionally validated, these marks could eventually contribute to epigenetic risk stratification approaches and help refine interventions that target metabolic improvement, potentially complementing traditional lifestyle and pharmacological strategies in children with obesity.

Abbreviations

BMI	Body mass index
BMIQ	Beta-mixture quantile
CVD	Cardiovascular disease
DBP	Diastolic blood pressure
DMP	Differentially methylated position
DNAm	DNA methylation
DMR	Differentially methylated region
DXA	Dual-energy X-ray absorptiometry
EWAS	Epigenome wide association study
FDR	False discovery rate
FMI	Fat mass index
GO	Gene ontology
GoDMC	Genetics of DNA methylation consortium
GSA	Global screening array
GWAS	Genome wide association study
HDL-c	HDL-cholesterol
HOMA-IR	Homeostasis model assessment of insulin resistance
hs-CRP	High-sensitivity C-reactive protein
HWE	Hardy-weinberg equilibrium
KEGG	Kyoto encyclopedia of genes and genomes
IOTF	International obesity task force
IR	Insulin resistance
LD	Linkage disequilibrium
LDL-c	LDL-cholesterol
MAF	Minor allele frequency
MetS	Metabolic syndrome
mQTL	Methylation quantitative trait loci
Nw	Normal-weight
Ov/Ob	Overweight/obesity
PWBC	Peripheral white blood cells
QC	Quality control
QUICKI	Quantitative insulin sensitivity check index
SBP	Systolic blood pressure
SNPs	Single nucleotide polymorphisms
TAG	Triacylglycerols
T2DM	Type 2 diabetes mellitus
WC	Waist circumference

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12933-026-03101-7.

Supplementary file 1.

Supplementary file 2.

Supplementary file 3.

Supplementary file 4.

Supplementary file 5.

Supplementary file 6.

Supplementary file 7.

Supplementary file 8.

Supplementary file 9.

Supplementary file 10.

Supplementary file 11.

Supplementary file 12.

Supplementary file 13.

Supplementary file 14.

Supplementary file 15.

Acknowledgements
The authors wish to thank the children and parents participating in the study and appreciate their effort.

Author contributions
A.A.R, A.T.M, M.B.A and A.S.L: Data curation; formal analysis; investigation; methodology; software; visualisation; roles/writing original draft; and writing review and editing. FJ.R.O, LA.M and A.G: Conceptualisation; methodology; validation and writing review and editing. M.G.C, G.B, R.L, J.A.F and CM. A: Conceptualisation; funding acquisition; methodology; project administration; resources; supervision; validation; and writing review and editing. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Funding
This study has been funded by Instituto de Salud Carlos III (ISCIII) through the projects PI23/00165, PI23/00028, PI23/00129, PI23/01032 and PI23/00191 and co-funded by the European Union.
This research was also supported by the Instituto de Salud Carlos III co-funded by the European Union and ERDF A way of making Europe (grant numbers PI11/01425, PI11/02042, PI11/02059, PI16/01301, PI16/01205, PI16/00871, PI20/00563, PI20/00711, PI20/00924 and P20/00988), and by the European Union through the Horizon Europe Framework Programme (eprObes project, grant number GA 101080219). The authors also acknowledge Instituto de Salud Carlos III for personal funding of Álvaro Torres-Martos and Mireia Bustos-Aibar: i-PFIS and PFIS contracts: IIS doctorates - company in health sciences and technologies of the Strategic Health Action (IFI22/00013 and FI23/00042). The project that gave rise to these results received the support of a fellowship from the "la Caixa" Foundation (ID 100010434)". The fellowship code is LCF/BQ/PR25/12110010.

Data availability
The data that support the findings of this study are not openly available due to reasons of sensitivity and are available from the corresponding authors (mariarosaura.leis@usc.es and jalcala@decsai.ugr.es) upon reasonable request. Data are located in controlled access data storage at the University of Granada.

Code availability
To promote open access and reproducibility, the code is publicly available on Github: https://github.com/AlvaroTorresMartos/EWAS_IR.

Declarations

Competing interests
The authors declare no competing interests.

Author details
[1]Barcelona Institute for Global Health, ISGlobal, 08003 Barcelona, Spain
[2]CIBER de Fisiopatología de la Obesidad y Nutrición (CIBEROBN), Instituto de Salud Carlos III, 28029 Madrid, Spain
[3]Department of Biochemistry and Molecular Biology II, School of Pharmacy, "José Mataix Verdú" Institute of Nutrition and Food Technology (INYTA) and Center of Biomedical Research (CIBM), University of Granada, 18071 Granada, Spain
[4]Instituto de investigación Biosanitaria, ibs.GRANADA, Granada, Spain
[5]Growth, Exercise, Nutrition and Development (GENUD) Research Group, Universidad de Zaragoza, Instituto Agroalimentario de Aragón (IA2), Instituto de Investigación Sanitaria de Aragón (IISA), 50009 Zaragoza, Spain
[6]RG Adipocytes and Metabolism, Institute for Diabetes and Obesity, Helmholtz Diabetes Center at the Helmholtz Zentrum München, 85764 Munich, Germany
[7]Metabolism and Investigation Unit, Reina Sofia University Hospital. Maimnides Institute of Biomedicine Research of Cordoba (IMIBIC), University of Córdoba, 14004 Córdoba, Spain
[8]Pediatric Endocrinology Unit, Facultad de Medicina, Clinic University Hospital Lozano Blesa, University of Zaragoza, 50009 Zaragoza, Spain
[9]Unidad de Gastroenterología, Hepatología y Nutrición Pediátrica del Hospital Clínico Universitario de Santiago. Research Group of Pediatric Nutrition, Instituto de Investigación Sanitaria de Santiago (IDIS), 15706 Santiago de Compostela, Spain
[10]Unit of Investigation in Nutrition, Growth and Human Development of Galicia, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain
[11]Department of Computer Science and Artificial Intelligence, Andalusian Research Institute in Data Science and Computational Intelligence (DaSCI), University of Granada, 18071 Granada, Spain

Received: 4 August 2025 / Accepted: 27 January 2026
Published online: 24 February 2026

References

1. Sinha R, Fisch G, Teague B, Tamborlane WV, Banyas B, Allen K, et al. Prevalence of impaired glucose tolerance among children and adolescents with marked obesity. N Engl J Med. 2002;346(11):802–10. https://doi.org/10.1056/nejmoa012578.
2. Drozdz D, Alvarez-Pitti J, Wójcik M, Borghi C, Gabbianelli R, Mazur A, et al. Obesity and cardiometabolic risk factors: from childhood to adulthood. Nutrients. 2021;13(11):4176. https://doi.org/10.3390/nu13114176.
3. Luo Y, Luo D, Li M, Tang B. Insulin resistance in pediatric obesity: from mechanisms to treatment strategies. Pediatr Diabetes. 2024;2298306(1):1–22. https://doi.org/10.1155/2024/2298306.
4. Gao M, Piernas C, Astbury NM, Hippisley-Cox J, O'Rahilly S, Aveyard P, et al. Associations between body-mass index and COVID-19 severity in 6 9 million people in England: a prospective, community-based, cohort study. Lancet Diabetes Endocrinol. 2021;9(6):350–9. https://doi.org/10.1016/s2213-8587(21)00089-9.
5. Jacobs DR, Woo JG, Sinaiko AR, Daniels SR, Ikonen J, Juonala M, et al. Childhood cardiovascular risk factors and adult cardiovascular events. N Engl J Med. 2022;386(20):1877–88. https://doi.org/10.1056/nejmoa2109191.
6. Magnussen C, Ojeda FM, Leong DP, Alegre-Diaz J, Amouyel P, Aviles-Santa L, et al. Global effect of modifiable risk factors on cardiovascular disease and mortality. N Engl J Med. 2023;389(14):1273–85. https://doi.org/10.1056/nejmoa2206916.
7. Altunkaya J, Piernas C, Pouwels KB, Jebb SA, Clarke P, Astbury NM, et al. Associations between BMI and hospital resource use in patients hospitalised for COVID-19 in England: a community-based cohort study. Lancet Diabetes Endocrinol. 2024;12(7):462–71. https://doi.org/10.1016/s2213-8587(24)00129-3.
8. Lim LL, Jones S, Cikomola JC, Hivert MF, Misra S. Understanding the drivers and consequences of early-onset type 2 diabetes. Lancet. 2025;405(10497):2327–40. https://doi.org/10.1016/s0140-6736(25)01012-8.
9. Jebeile H, Kelly AS, O'Malley G, Baur LA. Obesity in children and adolescents: epidemiology, causes, assessment, and management. Lancet Diabetes Endocrinol. 2022;10(5):351–65. https://doi.org/10.1016/s2213-8587(22)00047-x.
10. Lister NB, Baur LA, Felix JF, Hill AJ, Marcus C, Reinehr T, et al. Child and adolescent obesity. Nat Rev Dis Prim. 2023;9(24):1–19. https://doi.org/10.1038/s41572-023-00435-4.
11. Hannon TS, Arslanian SA. Obesity in adolescents. N Engl J Med. 2023;389(3):251–61. https://doi.org/10.1056/nejmcp2102062.
12. Hampl SE, Hassink SG, Skinner AC, Armstrong SC, Barlow SE, Bolling CF, et al. Clinical practice guideline for the evaluation and treatment of children and adolescents with obesity. Pediatrics. 2023;151(2):1–100. https://doi.org/10.1542/peds.2022-060640.
13. Abbassi V. Growth and normal puberty. Pediatrics. 1998;102(3):507–11. https://doi.org/10.1542/peds.102.S3.507.
14. Lamas C, Kalen A, Anguita-Ruiz A, Perez-Ferreiros A, Picans-Leis R, Flores K, et al. Progression of metabolic syndrome and associated cardiometabolic risk factors from prepuberty to puberty in children: the PUBMEP study. Front

Endocrinol. 2022;13(1082684):1–14. https://doi.org/10.3389/fendo.2022.1082684.
15. González-Gil EM, Anguita-Ruiz A, Kalén A, Perez DLL, C, Rupérez AI, Vázquez-Cobela R, et al. Longitudinal associations between cardiovascular biomarkers and metabolic syndrome during puberty: the PUBMEP study. Eur J Pediatr. 2023;182(1):419–29. https://doi.org/10.1007/s00431-022-04702-6.
16. Reinehr T, Wolters B, Knop C, Lass N, Holl RW. Strong effect of pubertal status on metabolic health in obese children: a longitudinal study. J Clin Endocrinol Metab. 2015;100(1):301–8. https://doi.org/10.1210/jc.2014-2674.
17. Reinehr T, Roth CL. Is there a causal relationship between obesity and puberty? Lancet Child Adolesc Health. 2019;3(1):44–54. https://doi.org/10.1016/S2352-4642(18)30306-7.
18. Pilia S, Casini MR, Foschini ML, Minerba L, Musiu MC, Marras V, et al. The effect of puberty on insulin resistance in obese children. J Endocrinol Invest. 2009;32(5):401–5. https://doi.org/10.1007/BF03346475.
19. Kelly LA, Lane CJ, Weigensberg MJ, Toledo-Corral CM, Goran MI. Pubertal changes of insulin sensitivity, acute insulin response, and -cell function in overweight latino youth. J Pediatr. 2011;158(3):442–6. https://doi.org/10.1016/j.jpeds.2010.08.046.
20. Kelsey MM, Zeitler PS. Insulin resistance of puberty. Curr Diabetes Rep. 2016;16(7):64. https://doi.org/10.1007/s11892-016-0751-5.
21. Dabelea D, Bell RA, D'Agostino RB Jr, Imperatore G, Johansen JM, Linder B, et al. Incidence of diabetes in youth in the United States. JAMA. 2007;297(24):2716–24. https://doi.org/10.1001/jama.297.24.2716.
22. James DE, Stöckli J, Birnbaum MJ. The aetiology and molecular landscape of insulin resistance. Nat Rev Mol Cell Biol. 2021;22(11):751–71. https://doi.org/10.1038/s41580-021-00390-6.
23. Luk A, Wild SH, Jones S, Anjana RM, Hivert MF, McCaffrey J, et al. Early-onset type 2 diabetes: the next major diabetes transition. Lancet. 2025;405(10497):2313–26. https://doi.org/10.1016/s0140-6736(25)00830-x.
24. Misra S, Khunti K, Goyal A, Gable D, Armocida B, Tandon N, et al. Managing early-onset type 2 diabetes in the individual and at the population level. Lancet. 2025;405(10497):2341–54. https://doi.org/10.1016/s0140-6736(25)01067-0.
25. Goodarzi MO. Genetics of obesity: what genetic association studies have taught us about the biology of obesity and its complications. Lancet Diabetes Endocrinol. 2018;6(3):223–36. https://doi.org/10.1210/jc.2015-3336.
26. Scott RA, Scott LJ, Mägi R, Marullo L, Gaulton KJ, Kaakinen M, et al. An expanded genome-wide association study of type 2 diabetes in Europeans. Diabetes. 2017;66(11):2888–902. https://doi.org/10.2337/db16-1253.
27. Zhao W, Rasheed A, Tikkanen E, Lee JJ, Butterworth AS, Howson JMM, et al. Identification of new susceptibility loci for type 2 diabetes and shared etiological pathways with coronary heart disease. Nat Genet. 2017;49(10):1450–7. https://doi.org/10.1038/ng.3943.
28. Mahajan A, Wessel J, Willems SM, Zhao W, Robertson NR, Chu AY, et al. Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. Nat Genet. 2018;50(4):559–71. https://doi.org/10.1038/s41588-018-0084-1.
29. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PIW, Chen H, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science. 2007;316(5829):1331–6. https://doi.org/10.1126/science.1142358.
30. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature. 2007;445(7130):881–5. https://doi.org/10.1126/science.1142358.
31. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science. 2007;316(5829):1341–5. https://doi.org/10.1126/science.1142382.
32. Smit RAJ, Wade KH, Hui Q, Arias JD, Yin X, Christiansen MR, et al. Polygenic prediction of body mass index and obesity through the life course and across ancestries. Nat Med. 2025;31(9):3151–68. https://doi.org/10.1038/s41591-025-03827-z.
33. de Prado-Bert P, Ruiz-Arenas C, Vives-Usano M, Andrusaityte S, Cadiou S, Carracedo Á, et al. The early-life exposome and epigenetic age acceleration in children. Environ Int. 2021;155(106683):106683. https://doi.org/10.1016/j.envint.2021.106683.
34. Maitre L, Bustamante M, Hernández-Ferrer C, Thiel D, Lau CHE, Siskos AP, et al. Multi-omics signatures of the human early life exposome. Nat Commun. 2022;13(7024):1–24. https://doi.org/10.1038/s41467-022-34422-2.
35. Plaza-Florido A, Anguita-Ruiz A, Esteban FJ, Aguilera CM, Labayen I, Reitzner SM, et al. Integrated analysis of methylome and transcriptome responses to exercise training in children with overweight/obesity. Physiol Genom. 2025;57(2):91–102. https://doi.org/10.1152/physiolgenomics.00059.2024.
36. García-Calzón S, Maguolo A, Eichelmann F, Edsfeldt A, Perfilyev A, Maziarz M, et al. Epigenetic biomarkers predict macrovascular events in individuals with type 2 diabetes. Cell Rep Med. 2025;6(8):102290. https://doi.org/10.1016/j.xcrm.2025.102290.
37. Aslibekyan S, Demerath EW, Mendelson M, Zhi D, Guan W, Liang L, et al. Epigenome-wide study identifies novel methylation loci associated with body mass index and waist circumference: epigenetics of Obesity. Obesity. 2015;23(7):1493–501. https://doi.org/10.1002/oby.21111.
38. Olsson AH, Volkov P, Bacos K, Dayeh T, Hall E, Nilsson EA, et al. Genome-wide associations between genetic and epigenetic variation influence mRNA expression and insulin secretion in human pancreatic islets. PLoS Genet. 2014;10(11):e1004735. https://doi.org/10.1371/journal.pgen.1004735.
39. Wahl S, Drong A, Lehne B, Loh M, Scott WR, Kunze S, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. Nature. 2017;541(7635):81–6. https://doi.org/10.1038/nature20784.
40. Do WL, Sun D, Meeks K, Dugué PA, Demerath E, Guan W, et al. Epigenome-wide meta-analysis of BMI in nine cohorts: examining the utility of epigenetically predicted BMI. Am J Hum Genet. 2023;110(2):273–83. https://doi.org/10.1016/j.ajhg.2022.12.014.
41. Li W, Xia M, Zeng H, Lin H, Teschendorff AE, Gao X, et al. Longitudinal analysis of epigenome-wide DNA methylation reveals novel loci associated with BMI change in East Asians. Clin Epigenet. 2024;16(70):1–13. https://doi.org/10.1186/s13148-024-01679-x.
42. Hidalgo B, Irvin MR, Sha J, Zhi D, Aslibekyan S, Absher D, et al. Epigenome-wide association study of fasting measures of glucose, insulin, and HOMA-IR in the genetics of lipid lowering drugs and diet network study. Diabetes. 2014;63(2):801–7. https://doi.org/10.2337/db13-1100.
43. Arner P, Sahlqvist AS, Sinha I, Xu H, Yao X, Waterworth D, et al. The epigenetic signature of systemic insulin resistance in obese women. Diabetologia. 2016;59(11):2393–405. https://doi.org/10.1007/s00125-016-4074-5.
44. Kriebel J, Herder C, Rathmann W, Wahl S, Kunze S, Molnos S, et al. Association between DNA methylation in whole blood and measures of glucose metabolism: KORA F4 study. PLoS ONE. 2016;11(3):e0152314.
45. Arpón A, Milagro FI, Ramos-Lopez O, Mansego ML, Santos JL, Riezu-Boj JI, et al (2019) Epigenome-wide association study in peripheral white blood cells involving insulin resistance. Sci Rep, 9(2445):1–11. https://doi.org/10.1038/s41598-019-38980-2
46. Sharma NK, Comeau ME, Montoya D, Pellegrini M, Howard TD, Langefeld CD, et al. Integrative analysis of glucometabolic traits, adipose tissue DNA methylation, and gene expression identifies epigenetic regulatory mechanisms of insulin resistance and obesity in African Americans. Diabetes. 2020;69(12):2779–93. https://doi.org/10.2337/db20-0117.
47. Florath I, Butterbach K, Heiss J, Bewerunge-Hudler M, Zhang Y, Schöttker B, et al. Type 2 diabetes and leucocyte DNA methylation: an epigenome-wide association study in over 1,500 older adults. Diabetologia. 2016;59(1):130–8. https://doi.org/10.1007/s00125-015-3773-7.
48. Cardona A, Day FR, Perry JRB, Loh M, Chu AY, Lehne B, et al. Epigenome-wide association study of incident type 2 diabetes in a British population: EPIC-Norfolk study. Diabetes. 2019;68(12):2315–26. https://doi.org/10.2337/db18-0290.
49. Ling C, Rönn T. Epigenetics in human obesity and type 2 diabetes. Cell Metab. 2019;29(5):1028–44. https://doi.org/10.1016/j.cmet.2019.03.009.
50. Juvinao-Quintero DL, Marioni RE, Ochoa-Rosales C, Russ TC, Deary IJ, van Meurs JBJ, et al. DNA methylation of blood cells is associated with prevalent type 2 diabetes in a meta-analysis of four European cohorts. Clin Epigenet. 2021;13(14):2–14. https://doi.org/10.1186/s13148-021-01027-3.
51. Lai L, Juntilla DL, Del M, Gomez-Alonso DC, M, Grallert H, Thorand B, et al. Longitudinal association between DNA methylation and type 2 diabetes: findings from the KORA F4/FF4 study. Cardiovasc Diabetol. 2025;24(19):1–15. https://doi.org/10.1186/s12933-024-02558-8.
52. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet. 2008;9(6):465–76. https://doi.org/10.1038/nrg2341.
53. Han L, Zhang H, Kaushal A, Rezwan FI, Kadalayil L, Karmaus W, et al. Changes in DNA methylation from pre- to post-adolescence are associated with pubertal exposures. Clin Epigenet. 2019;11(176):1–14. https://doi.org/10.1186/s13148-019-0780-4.
54. Vehmeijer FOL, Küpers LK, Sharp GC, Salas LA, Lent S, Jima DD, et al. DNA methylation and body mass index from birth to adolescence: meta-analyses

of epigenome-wide association studies. Genome Med. 2020;12(105):1–15. https://doi.org/10.1186/s13073-020-00810-w.
55. Ott R, Stein R, Hauta-alus HH, Ronkainen J, Fernández-Barrés S, Spielau U, et al. Epigenome-wide meta-analysis reveals associations between dietary glycemic index and glycemic load and DNA methylation in children and adolescents of different body sizes. Diabetes Care. 2023;46(11):2067–75. https://doi.org/10.2337/dc23-0474.
56. Fragoso-Bargas N, Elliott HR, Lee-Ødegård S, Opsahl JO, Sletner L, Jenum AK, et al. Cross-ancestry DNA methylation marks of insulin resistance in Pregnancy: an integrative epigenome-wide association study. Diabetes. 2023;72(3):415–26. https://doi.org/10.2337/db22-0504.
57. Salama OE, Hizon N, Del Vecchio M, Kolsun K, Fonseca MA, Lin DTS, et al. DNA methylation signatures of youth-onset type 2 diabetes and exposure to maternal diabetes. Clin Epigenet. 2024. https://doi.org/10.1186/s13148-024-01675-1.
58. Shin J, Bressler J, Grove ML, Brown M, Selvin E, Pankow JS, et al. DNA methylation markers of insulin resistance surrogate measures in the atherosclerosis risk in communities (ARIC) study. Epigenetics. 2025;20(1):2498857. https://doi.org/10.1080/15592294.2025.2498857.
59. Linares-Pineda TM, Lendínez-Jurado A, Piserra-López A, Suárez-Arana M, Pozo M, Molina-Vega M, et al. Longitudinal DNA methylation profiles in saliva of offspring from mothers with gestational diabetes: associations with early childhood growth patterns. Cardiovasc Diabetol. 2025;24(1):15. https://doi.org/10.1186/s12933-024-02568-6.
60. Llauradó-Pont J, Stratakis N, Fiorito G, Handakas E, Neumann A, Barros H, et al. A meta-analysis of epigenome-wide association studies of ultra-processed food consumption with DNA methylation in European children. Clin Epigenet. 2025;17(1):3. https://doi.org/10.1186/s13148-024-01782-z.
61. Su S, Zhu H, Xu X, Wang X, Dong Y, Kapuku G, et al. DNA methylation of the ly86 gene is associated with obesity, insulin resistance, and inflammation. Twin Res Hum Genet. 2014;17(3):183–91. https://doi.org/10.1017/thg.2014.22.
62. van Dijk SJ, Peters TJ, Buckley M, Zhou J, Jones PA, Gibson RA, et al. DNA methylation in blood from neonatal screening cards and the association with BMI and insulin sensitivity in early childhood. Int J Obes. 2018;42(1):28–35. https://doi.org/10.1038/ijo.2017.228.
63. Barbosa P, Landes RD, Graw S, Byrum SD, Bennuri S, Delhey L, et al. Effect of excess weight and insulin resistance on DNA methylation in prepubertal children. Sci Rep. 2022;12(8430):1–10. https://doi.org/10.1038/s41598-022-12325-y.
64. Schüssler-Fiorenza Rose SM, Contrepois K, Moneghetti KJ, Zhou W, Mishra T, Mataraso S, et al. A longitudinal big data approach for precision health. Nat Med. 2019;25(5):792–804. https://doi.org/10.1038/s41591-019-0414-6.
65. Campagna MP, Xavier A, Lechner-Scott J, Maltby V, Scott RJ, Butzkueven H, et al. Epigenome-wide association studies: current knowledge, strategies and recommendations. Clin Epigenet. 2021;13(214):1–24. https://doi.org/10.1186/s13148-021-01200-8.
66. Ye Y, Zhang Z, Liu Y, Diao L, Han L. A multi-omics perspective of quantitative trait loci in precision medicine. Trends Genet. 2020;36(5):318–36. https://doi.org/10.1016/j.tig.2020.01.009.
67. Torres-Martos Á, Anguita-Ruiz A, Bustos-Aibar M, Ramírez-Mena A, Arteaga M, Bueno G, et al. Multiomics and eXplainable artificial intelligence for decision support in insulin resistance early diagnosis: a pediatric population-based longitudinal study. Artif Intell Med. 2024;156(102962):1–14. https://doi.org/10.1016/j.artmed.2024.102962.
68. Stratakis N, Anguita-Ruiz A, Fabbri L, Maitre L, González JR, Andrusaityte S, et al. Multi-omics architecture of childhood obesity and metabolic dysfunction uncovers biological pathways and prenatal determinants. Nat Commun. 2025;16(654):1–15. https://doi.org/10.1038/s41467-025-56013-7.
69. Tanner JM, Whitehouse RH. Clinical longitudinal standards for height, weight, height velocity, weight velocity, and stages of puberty. Arch Dis Childood. 1976;51(3):170–9. https://doi.org/10.1136/adc.51.3.170.
70. Stewart A, Marfell-Jones M, Olds T, De Ridder J. International standards for anthropometric assessment (3rd ed.). International Society for the Advancement of Kinanthropometry; 2012.
71. Sobradillo B, Aguirre A, Aresti U, Bilbao A, Fernández-Ramos C, Lizarraga A, et al. Curvas y Tablas de Crecimiento (Estudios Longitudinal y Transversal). 1st ed. Bilbao: Fundación Faustino Orbegozo Eizaguirre; 2004.
72. Ferrández A. Curvas y tablas de crecimiento: estudios longitudinal y transversal. Zaragoza, España: Fundación Andrea Prader; 2005.
73. McCrindle BW. Assessment and management of hypertension in children and adolescents. Nat Rev Cardiol. 2010;7(3):155–63. https://doi.org/10.1038/nrcardio.2009.231.
74. Pérez-Gimeno G, Ruperez AI, Gil-Campos M, Aguilera CM, Anguita A, Vázquez-Cobela R, et al. Height-based equations as screening tools for high blood pressure in pediatric practice, the GENOBOX study. J Clin Hypertens. 2022;24(6):713–22. https://doi.org/10.1111/jch.14489.
75. Anguita-Ruiz A, Plaza-Diaz J, Ruiz-Ojeda FJ, Rupérez AI, Leis R, Bueno G, et al. X chromosome genetic data in a Spanish children cohort, dataset description and analysis pipeline. Sci Data. 2019;6(130):1–10. https://doi.org/10.1038/s41597-019-0109-3.
76. Ruiz-Ojeda FJ, Anguita-Ruiz A, Rupérez AI, Gomez-Llorente C, Olza J, Vázquez-Cobela R, et al. Effects of X-chromosome tenomodulin genetic variants on obesity in a children's cohort and implications of the gene in adipocyte metabolism. Sci Rep. 2019;9(3979):1–16. https://doi.org/10.1038/s41598-019-40482-0.
77. Tang Q, Li X, Song P, Xu L. Optimal cut-off values for the homeostasis model assessment of insulin resistance (HOMA-IR) and pre-diabetes screening: developments in research and prospects for the future. Drug Discov Ther. 2015;9(6):380–5. https://doi.org/10.5582/ddt.2015.01207.
78. Rupérez AI, Olza J, Gil-Campos M, Leis R, Bueno G, Aguilera CM, et al. Cardiovascular risk biomarkers and metabolically unhealthy status in prepubertal children: comparison of definitions. Nutr Metab Cardiovasc Dis. 2018;28(5):524–30. https://doi.org/10.1016/j.numecd.2018.02.006.
79. Anguita-Ruiz A, Mendez-Gutierrez A, Ruperez AI, Leis R, Bueno G, Gil-Campos M, et al. The protein S100A4 as a novel marker of insulin resistance in prepubertal and pubertal children with obesity. Metabolism. 2020;105(154187):1–15. https://doi.org/10.1016/j.metabol.2020.154187.
80. Cole TJ, Lobstein T. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. Pediatr Obes. 2012;7(4):284–94. https://doi.org/10.1111/j.2047-6310.2012.00064.x.
81. Torres-Martos Á, Requena F, López-Rodríguez G, Hernández-Cabrera J, Galván M, Solís-Pérez E, et al. ObMetrics: a shiny app to assist in metabolic syndrome assessment in paediatric obesity. Pediatr Obes. 2025;20(18):e70016. https://doi.org/10.1111/ijpo.70016.
82. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–75. https://doi.org/10.1086/519795.
83. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. Nat Protoc. 2010;5(9):1564–73. https://doi.org/10.1038/nprot.2010.116.
84. Torres-Martos Á, Bustos-Aibar M, Ramírez-Mena A, Cámara-Sánchez S, Anguita-Ruiz A, Alcalá R, et al. Omics Data Preprocessing for Machine Learning: a case study in childhood obesity. Genes. 2023;14(2):1–16. https://doi.org/10.3390/genes14020248.
85. R Core Team.: R: A language and environment for statistical computing. Vienna, Austria. Available from: https://www.R-project.org/.
86. Fortin JP, Triche TJ Jr, Hansen KD. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. Bioinformatics. 2017;33(4):558–60. https://doi.org/10.1093/bioinformatics/btw691.
87. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics. 2013;29(2):189–96. https://doi.org/10.1093/bioinformatics/bts680.
88. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics. 2013;8(2):203–9. https://doi.org/10.4161/epi.23470.
89. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. Bioinformatics. 2008;24(13):1547–8.
90. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics. 2010;11(587):1–9. https://doi.org/10.1186/1471-2105-11-587.
91. Kruppa J, Sieg M, Richter G, Pohrt A. Estimands in epigenome-wide association studies. Clin Epigenet. 2021;13(98):1–16. https://doi.org/10.1186/s13148-021-01083-9.
92. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinf. 2012;13(86):1–16. https://doi.org/10.1186/1471-2105-13-86.

93. Fox, J. Polycor: polychoric and polyserial correlations. R package version 0.8-1. Available from: https://CRAN.R-project.org/package=polycor.
94. Frisch R, Waugh FV. Partial time regressions as compared with individual trends. Econometrica. 1933;1(4):387. https://doi.org/10.2307/1907330.
95. Ding P. The Frisch-Waugh-Lovell theorem for standard errors. Stat Probab Lett. 2021;168(108945):108945. https://doi.org/10.1016/j.spl.2020.108945.
96. Lussier AA, Zhu Y, Smith BJ, Simpkin AJ, Smith ADAC, Suderman MJ, et al. Updates to data versions and analytic methods influence the reproducibility of results from epigenome-wide association studies. Epigenetics. 2022;17(11):1373–88. https://doi.org/10.1080/15592294.2022.2028072.
97. Perrier F, Novoloaca A, Ambatipudi S, Baglietto L, Ghantous A, Perduca V, et al. Identifying and correcting epigenetics measurements for systematic sources of variation. Clin Epigenet. 2018;10(1):38. https://doi.org/10.1186/s13148-018-0471-6.
98. Graw S, Henn R, Thompson JA, Koestler DC. pwrEWAS: a user-friendly tool for comprehensive power estimation for epigenome wide association studies (EWAS). BMC Bioinf. 2019;20(1):218. https://doi.org/10.1186/s12859-019-2804-7.
99. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):1–13. https://doi.org/10.1093/nar/gkv007.
100. van Iterson M, van Zwet EW, BIOS Consortium, Heijmans BT. Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. Genome Biol. 2017;18(1):19. https://doi.org/10.1186/s13059-016-1131-9.
101. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc: Ser B (Methodol). 1995;57(1):289–300. https://doi.org/10.2307/2346101.
102. Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K, Lord RV, et al. De novo identification of differentially methylated regions in the human genome. Epigenet Chromatin. 2015;8(1):6. https://doi.org/10.1186/1756-8935-8-6.
103. Hansen KD. IlluminaHumanMethylationEPICmanifest: manifest for illumina's epic methylation arrays. R package version 0.3.0. Available from: https://bitbucket.com/kasperdanielhansen/Illumina_EPIC.
104. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from illumina's humanmethylation450 platform. Bioinformatics. 2016;32(2):286–8. https://doi.org/10.1093/bioinformatics/btv560.
105. Carlson M. org.Hs.eg.db: genome wide annotation for Human. R package version 3.18.0.
106. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol. 2010;11(2):1–12. https://doi.org/10.1186/gb-2010-11-2-r14.
107. Geeleher P, Hartnett L, Egan LJ, Golden A, Raja Ali RA, Seoighe C. Gene-set analysis is severely biased when applied to genome-wide methylation data. Bioinformatics. 2013;29(15):1851–7. https://doi.org/10.1093/bioinformatics/btt311.
108. Garcia-Moreno A, López-Domínguez R, Villatoro-García JA, Ramirez-Mena A, Aparicio-Puerta E, Hackenberg M, et al. Functional enrichment analysis of regulatory elements. Biomedicines. 2022;10(3):590. https://doi.org/10.3390/biomedicines10030590.
109. Min JL, Hemani G, Hannon E, Dekkers KF, Castillo-Fernandez J, Luijk R, et al. Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. Nat Genet. 2021;53(9):1311–21. https://doi.org/10.1038/s41588-021-00923-x.
110. Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. Bioinformatics. 2012;28(10):1353–8. https://doi.org/10.1093/bioinformatics/bts163.
111. Royston P, Remark AS. R94: a remark on algorithm AS 181: the W-test for normality. J R Stat Soc: Ser C: Appl Stat. 1995;44(4):547. https://doi.org/10.2307/2986146.
112. Sjoberg DD, Whiting K, Curry M, Lavery JA, Larmarange J. Reproducible summary tables with the gtsummary package. R J. 2021;13:570–80.
113. Braun KVE, Voortman T, Dhana K, Bramer WM, Troup J, Troup J, et al. The role of DNA methylation in dyslipidaemia: a systematic review. Prog Lipid Res. 2016;64:178–91. https://doi.org/10.1016/j.plipres.2016.10.002.
114. Walaszczyk E, Luijten M, Spijkerman AMW, Bonder MJ, Lutgers HL, Snieder H, et al. DNA methylation markers associated with type 2 diabetes, fasting glucose and HbA1c levels: a systematic review and replication in a case–control sample of the Lifelines study. Diabetologia. 2017;61(2):354–68. https://doi.org/10.1007/s00125-017-4497-7.
115. Mazaheri-Tehrani S, Khoshhali M, Heidari-Beni M, Poursafa P, Kelishadi R. A systematic review and meta-analysis of observational studies on the effects of epigenetic factors on serum triglycerides. Arch Endocrinol Metab. 2022;66(33):407–19.
116. Barouti Z, Heidari-Beni M, Shabanian-Boroujeni A, Mohammadzadeh M, Pahlevani V, Poursafa P, et al. Effects of DNA methylation on cardiometabolic risk factors: a systematic review and meta-analysis. Arch Public Health. 2022;80(150):1–16. https://doi.org/10.1186/s13690-022-00907-1.
117. Krolevets M, Vt C, Prochaska JH, Schulz A, Rapp S, Tenzer S, et al. DNA methylation and cardiovascular disease in humans: a systematic review and database of known CpG methylation sites. Clin Epigenet. 2023;15(56):1–16. https://doi.org/10.1186/s13148-023-01468-y.
118. Nadiger N, Veed JK, Chinya Nataraj P, Mukhopadhyay A. DNA methylation and type 2 diabetes: a systematic review. Clin Epigenet. 2024;16(67):1–25. https://doi.org/10.1186/s13148-024-01670-6.
119. Ling C. Epigenetic regulation of insulin action and secretion – role in the pathogenesis of type 2 diabetes. J Intern Med. 2020;288(2):158–67. https://doi.org/10.1111/joim.13049.
120. Dick KJ, Nelson CP, Tsaprouni L, Sandling JK, Aïssi D, Wahl S, et al. DNA methylation and body-mass index: a genome-wide analysis. Lancet. 2014;383(9933):1990–8. https://doi.org/10.1016/S0140-6736(13)62674-4.
121. Pfeiffer S, Krüger J, Maierhofer A, Böttcher Y, Klöting N, El Hajj N, et al. Hypoxia-inducible factor 3A gene expression and methylation in adipose tissue is related to adipose tissue dysfunction. Sci Rep. 2016;6(1):27969. https://doi.org/10.1038/srep27969.
122. Main AM, Gillberg L, Jacobsen AL, Nilsson E, Gjesing AP, Hansen T, et al. DNA methylation and gene expression of HIF3A: cross-tissue validation and associations with BMI and insulin resistance. Clin Epigenet. 2016;8(1):89. https://doi.org/10.1186/s13148-016-0258-6.
123. Li S, Sanna S, Maschio A, Busonero F, Usala G, Mulas A, et al. The GLUT9 gene is associated with serum uric acid levels in Sardinia and Chianti cohorts. PLoS Genet. 2007;3(11):e194. https://doi.org/10.1371/journal.pgen.0030194.
124. Le MT, Shafiu M, Mu W, Johnson RJ. SLC2A9-a fructose transporter identified as a novel uric acid transporter. Nephrol Dial Transpl. 2008;23(9):2746–9. https://doi.org/10.1093/ndt/gfn349.
125. Jing J, Ekici AB, Sitter T, Eckardt KU, Schaeffner E, Li Y, et al. Genetics of serum urate concentrations and gout in a high-risk population, patients with chronic kidney disease. Sci Rep. 2018;8(1):13184. https://doi.org/10.1038/s41598-018-31282-z.
126. Ruiz A, Gautschi I, Schild L, Bonny O. Human mutations in SLC2A9 (Glut9) affect transport capacity for urate. Front Physiol. 2018. https://doi.org/10.3389/fphys.2018.00476.
127. Pavelcova K, Bohata J, Pavlikova M, Bubenikova E, Pavelka K, Stiburkova B. Evaluation of the influence of genetic variants of SLC2A9 (GLUT9) and SLC22A12 (URAT1) on the development of hyperuricemia and gout. J Clin Med. 2020;9(8):2510. https://doi.org/10.3390/jcm9082510.
128. Niu Y, Tang Q, Zhao X, Zhao X, Mao X, Sheng J, et al. Obesity-induced insulin resistance is mediated by high uric acid in obese children and adolescents. Front Endocrinol (Lausanne). 2021;12:773820. https://doi.org/10.3389/fendo.2021.773820.
129. Dalili S, Hassanzadeh Rad A, Salkhori O, Dabbaghi S, Karambin MM, Badeli H, et al. Potentials of hyperuricemia and insulin levels in predicting hypertension in obese children: a cross-sectional study. J Compr Pediatr. 2023. https://doi.org/10.5812/jcp-139577.
130. Villasis-Keever MA, Zurita-Cruz JN, Alcaraz-Hurtado IA, Klünder-Klünder M, Vilchis-Gil J, Romero-Guerra AL, et al. Association of serum uric acid levels with cardiometabolic factors in adolescents with obesity: a cross-sectional study. Metabolites. 2025;15(4):237. https://doi.org/10.3390/metabo15040237.
131. Rao J, Yan Y, Cheng H, Hou D, Zhao X, Shan X, et al. Uric acid mediated the relationship between obesity and hypertension in children and adolescents: a population?based cohort study. Nutr Metab Cardiovasc Dis. 2024;34(1):214–22.
132. Palka JA, Phang JM. Prolidase activity in fibroblasts is regulated by interaction of extracellular matrix with cell surface integrin receptors. J Cell Biochem. 1997;67(2):166–75.
133. Besio R, Gioia R, Cossu F, Monzani E, Nicolis S, Cucca L, et al. Kinetic and structural evidences on human prolidase pathological mutants suggest strategies for enzyme functional rescue. PLoS ONE. 2013;8(3):e58792. https://doi.org/10.1371/journal.pone.0058792.
134. Eni-Aganga I, Lanaghan ZM, Balasubramaniam M, Dash C, Pandhare J. PROLIDASE: a review from discovery to its role in health and disease. Front Mol Biosci. 2021;8:723003. https://doi.org/10.3389/fmolb.2021.723003.

135. Misiura M, Miltyk W. Current understanding of the emerging role of prolidase in cellular metabolism. Int J Mol Sci. 2020;21(16):5906. https://doi.org/10.3390/ijms21165906.
136. Jaworowska A, Miltyk W. Where does prolidase come from and where is it heading? - A review of recent findings. Int J Biol Macromol. 2025;309(4):143044. https://doi.org/10.1016/j.ijbiomac.2025.143044.
137. Pellegrinelli V, Rodriguez-Cuenca S, Rouault C, Figueroa-Juarez E, Schilbert H, Virtue S, et al. Dysregulation of macrophage PEPD in obesity determines adipose tissue fibro-inflammation and insulin resistance. Nat Metab. 2022;4(4):476–94. https://doi.org/10.1038/s42255-022-00561-5.
138. Pellegrinelli V, Figueroa-Juárez E, Samuelson I, U-Din M, Rodriguez-Fdez S, Virtue S, et al. Defective extracellular matrix remodeling in brown adipose tissue is associated with fibro-inflammation and reduced diet-induced thermogenesis. Cell Rep. 2023;42(6):112640. https://doi.org/10.1016/j.celrep.2023.112640.
139. Inoki K, Zhu T, Guan KL. TSC2 mediates cellular energy response to control cell growth and survival. Cell. 2003;115(5):577–90. https://doi.org/10.1016/s0092-8674(03)00929-2.
140. Kwiatkowski DJ, Manning BD. Tuberous sclerosis: a GAP at the crossroads of multiple signaling pathways. Hum Mol Genet. 2005. https://doi.org/10.1093/hmg/ddi260.
141. Laplante M, Sabatini DM. mTOR signaling in growth control and disease. Cell. 2012;149(2):274–93. https://doi.org/10.1016/j.cell.2012.03.017.
142. Manning AK, Hivert MF, Scott RA, Grimsby JL, Bouatia-Naji N, Chen H, et al. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. Nat Genet. 2012;44(6):659–69. https://doi.org/10.1038/ng.2274.
143. Saxton RA, Sabatini DM. MTOR signaling in growth, metabolism, and disease. Cell. 2017;168(6):960–76. https://doi.org/10.1016/j.cell.2017.02.004.
144. Jiang C, Qu A, Matsubara T, Chanturiya T, Jou W, Gavrilova O, et al. Disruption of hypoxia-inducible factor 1 in adipocytes improves insulin sensitivity and decreases adiposity in high-fat diet-fed mice. Diabetes. 2011;60(10):2484–95. https://doi.org/10.2337/db11-0174.
145. Palmer BF, Clegg DJ. Oxygen sensing and metabolic homeostasis. Mol Cell Endocrinol. 2014;397(1–2):51–8. https://doi.org/10.1016/j.mce.2014.08.001.
146. Morin R, Mauger JF, Amaratunga R, Imbeault P. The effect of acute intermittent hypoxia on postprandial triglyceride levels in humans: a randomized crossover trial. J Transl Med. 2021;19(1):268. https://doi.org/10.1186/s12967-021-02933-z.
147. Matthaeus C, Lahmann I, Kunz S, Jonas W, Melo AA, Lehmann M, et al. EHD2-mediated restriction of caveolar dynamics regulates cellular fatty acid uptake. Proc Natl Acad Sci U S A. 2020;117(13):7471–81. https://doi.org/10.1073/pnas.1918415117.
148. Matthaeus C, Taraska JW. Energy and dynamics of caveolae trafficking. Front Cell Dev Biol. 2020, 8:614472. https://doi.org/10.3389/fcell.2020.614472.
149. Almstrup K, Lindhardt Johansen M, Busch AS, Hagen CP, Nielsen JE, Petersen JH, et al. Pubertal development in healthy children is mirrored by DNA methylation patterns in peripheral blood. Sci Rep. 2016, https://doi.org/10.1038/srep28657.

Publisher's Note