nature portfolio

Peer Review File

Multi-omics Architecture of Childhood Obesity and Metabolic Dysfunction Uncovers Biological Pathways and Prenatal Determinants

Corresponding Author: Dr Nikos Stratakis

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Paper Review: Multi-omics architecture of obesity and metabolic dysfunction in childhood: identifying biological pathways and prenatal determinants.

Manuscript number: NCOMMS-24-25298
Thank you for the opportunity to review this manuscript submission. In this study, the authors have assessed the molecular profiles of childhood obesity and associated metabolic dysfunction using extensive multi-omics profiling (methylome, miRNome, transcriptome, proteins and metabolites) across two parts of Europe within the population-based Human Early Life Exposome project. An untargeted approach uncovered three clusters of children defined by specific multi-omics profiles, one of which characterized by higher adiposity as well as a high degree of metabolic complications (as assessed with associations to the IDEFICS continuous MetS score) and inflammation-related cascades. Looking at early exposures from a DOHaD perspective, pre-pregnancy body mass index and environmental pollutants like perfluorooctanoate (PFOA) and mercury were found to be significant determinants of this high-risk cluster in the overarching aim to identify specific targets for mitigation of childhood obesity. This work is impressive, comprehensive, and the manuscript is organized and well-written.
The following remarks are to be considered/addressed:
Major comments:
1. While this is not to suggest for the authors to do any additional analyses, it would be useful to the reader if the authors would defend their choice of this bottom-up clustering approach versus using the IDEFICS MetS score that they've derived as their outcome and subsequently investigate the multiomics profiles associated with this score (e.g. using LASSO). Especially with the emphasis of metabolic health beyond obesity. One can clearly see that the score is most significantly and consistently associated with the derived clusters.
2. Line 57-end of introduction: this reads more like an abstract than the final paragraph of the introduction. Please refrain from having results in this paragraph (3 clusters, BMI, PFOA) and revise to have this section the summarize the objectives of the study.
3. Methods: Line 445: in the creation of the multiomics clusters and screening of published literature it would be useful to add (perhaps in supplemental Table 1) how many of the final included studies/molecues were from the same HELIX populations. As mentioned in Line 96 one objective was the replication of published literature and given HELIX is one of the largest birth cohorts, it is important to know how many of these molecules were from independent studies.
4. Aslo, when conducting the literature review, I think an important and comprehensive review on this topic in terms of metabolomic childhood obesity markers was not included, can the authors confirm? Would be useful to include it. [Handakas et al. Obesity Reviews. 2022;23(S1):e13384.]
5. Investigating the association between the high-risk multi-omics cluster in childhood and subsequent obesity traits in adolescence in a subsample of our population is a point of strength; as is the investigation of the prospective in-utero factors.
Minor comments:
1. Line 502: missing the word "to" handle….
2. Line 345: can you elaborate on the reproducible metabolomics measurements? Are these absolute concentrations or RPAs?
3. Discussion is very well written. Can you add insights to which omics layers were most predictive/valuable in the

clustering? If other cohorts had limited funding, which analyses should they prioritize?

(Remarks on code availability)
I have not reviewed the code but appreciate that it is publicly available.

Reviewer #2

(Remarks to the Author)
The manuscript investigates various omics within blood samples collected from children using several advanced techniques, including arrays for DNA methylation, miRNA, and gene expression, as well as multiplex assays for plasma protein and targeted LC-MS/MS for metabolites. The authors integrated omic targets reported in the published literature and categorized the cohorts into three distinct clusters associated with different prenatal exposures and metabolic outcomes. There are several commendable strengths in this study, particularly the large sample size and the extensive data available for childhood outcomes and pregnancy information. The analytical approach is robust and highlights key targets that may have significant biological relevance. Additionally, the manuscript is well-written, making it easy to read despite the complexity of the analyses. However, there are several limitations that decrease the enthusiasm of this reviewer:

1. The study designs and analyses do not elucidate the pathophysiology or mechanistic understanding of prenatal environments as suggested in the abstract and several parts of the manuscript. Instead, they demonstrate associations. While the authors acknowledge this in the discussion, the abstract and text within the manuscript present somewhat confounding information. For example, lines 15-17 state, "identify precise targets for prevention and intervention strategies early in the life-course...," yet the study identifies associated prenatal risk factors that are potentially modifiable without establishing them as precise targets for prevention (i.e. clinical studies) . Similarly, lines 71-73, "how prenatal environmental factors contribute to disease risks...," imply causation, which is not supported by the data.
2. The study's generalizability is limited by its demographic composition, with the majority of participants being White (85% in N/W and 100% in S/M) and the remainder categorized as Asian or Other. Given the study's European location, this is understandable. However, it raises questions about whether the same clusters and associations would be observed in more diverse populations. Addressing this factor would be important for reaching a wider, global audience.
3. While the study benefits from extensive prenatal data collection, several critical pieces of information are missing, which hampers the interpretation of prenatal and early life factors. Notably, maternal conditions such as diabetes (PMID: 30008062, 38820461) and preeclampsia (PMID: 36378310), which are known to impact offspring's metabolic and cardiovascular health, as well as cytokine levels, are not discussed. Including these data would strengthen the analyses and conclusions
4. Most results are presented using measures that demonstrate the magnitude of contribution (e.g., odds ratio, beta coefficient, iSHAP). However, presenting the mean levels of elevated cytokines and branched-chain amino acids in these clusters would be informative for both clinical and basic scientists. This data should be included, at least in a supplemental figure, to aid determining clinically relevant cutoffs in both clinical studies and basic science experiments.
5. In Figure 4a, maternal BMI, PFOA, and mercury levels appear to have different impacts in the N/W and S/M regions. For instance, maternal BMI differences may not be clinically actionable given the small differences within clusters, and the BMI in S/M region is actually normal in all clusters. The blood PFOA and mercury levels are somewhat modest in N/W (5-10% difference), raising some questions about the biological impact in general. Additionally, the mercury levels of both N/W and S/M appears to be extremely high. (1.65-4.1 ug/dl as mean) The 95% mean levels of US populations is ~0.4ug/L, with mean levels much lower than that. (https://www.epa.gov/americaschildrenenvironment/biomonitoring-mercury). A study in German also reported ranges comparable to the US population. (PMID: 34973943) This is the same for blood PFOA levels – the range reported by CDC is between 1.5-2 ug/L (https://www.cdc.gov/exposurereport/data_tables.html?NER_SectionItem=NHANES) and another study (PMID: 36574487), the unit listed in supp table is mg/dL and that would make the PFOA levels extremely high.
6. One key message that is lacking within the manuscript is the information regarding transitioning from metabolically healthy obesity to metabolically unhealthy obesity. (PMID: 38941611, PMID: 33125374). This is crucial, as it is unclear whether children in cluster B are genuinely healthy or if they will become metabolically unhealthy. This aspect should be emphasized to avoid giving false reassurance about the health of metabolically healthy obese children. Although briefly mentioned in lines 146-152 and supplemental figure 3, more comprehensive data on metabolic syndrome risk scores is needed.

Minor comment
- For data reproducibility, it is important to state if blood was collected after fasting and how the blood was collected, as these factors can significantly influence metabolite, miRNA, and transcriptome profiles.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
Understanding the molecular mechanism of childhood obesity, including biomarker signatures and environmental determinants, is crucial for developing preventive and therapeutic measures. This manuscript employs a systems biology approach with deep, multi-omics profiling to identify and validate molecular signatures of childhood obesity/metabolic

dysfunction and to characterize different subgroups based on metabolic traits and molecular pathways. Overall, it is a solid study and highly relevant to the field. Addressing the following questions will help enhance the study's impact and clarity.

(1) In the "Molecular drivers of the multi-omics clusters" section, the authors have identified several drivers that significantly contribute to defining high-risk Cluster C, including important genes. Please integrate with published bulk (GTEx, etc.) or scRNA-seq data to determine which tissues/cell types highly express these risk genes within Cluster C. This will reveal additional mechanisms of childhood obesity.

(2) Please clarify why cluster A was chosen over cluster B as the reference category in Fig 2? What criteria were used for this selection.

(3) Please describe how specific environmental exposures interact with genetic factors to influence obesity risk.

(4) Are there any untargeted approaches available to identify new biomarkers and risk factors?

(5) Line 100, "GpG sites" should be "CpG sites." Supplementary Table 1, there is a discrepancy: the number of DNA methylation sites (N=976) does not match the content on line 100, which refers to 977 CpG sites. Additionally, please include information regarding the protein section in Supplementary Table 1 and Supplementary Data 2A-D.

(Remarks on code availability)
I don't have the expertise to fully evaluate the code.

Reviewer #4

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors have answered all my questions and I am satisfied with the now strengthened manuscript.
A minor comment: line 326 has a word missing.

Furthermore, I leave it up to the authors if they can add as supplemental, a stratified analysis by sex. There have been few studies on sex differences in the metabolic profiles/obesity profiles in childhood and it can be interesting for the reader to have these results.

Reviewer #2

(Remarks to the Author)
Thank for addressing the comments. Most of the comments have been addressed adequately. The revised manuscript was not able to address causality given the nature of the study, and did not have sufficient data collected to inform role of other maternal contributions (e.g maternal diabetes, pre-eclampsia) or socioeconomic status. Authors have discussed these limitations within discussion section.

Reviewer #3

(Remarks to the Author)
The revised manuscript is significantly improved. For clarity and completeness, we recommend the following minor changes:

1. Please highlight genes with high expression in adipose tissue, liver, skeletal muscle, and pancreas in bold or a different color in supplementary Fig. 9 for readability.

2. Important discussions on transcriptome, proteome and metabolome are included in the Discussion sections, but little on methylome and miRNome. Please consider including discussion on how methylome and miRNome contribute to childhood obesity, especially in light of results in Fig. 3a.

Reviewer #4

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Dear Reviewers,

Thank you for the opportunity to respond to your evaluations of our manuscript # NCOMMS-24-25298 entitled "Multi-omics architecture of obesity and metabolic dysfunction in childhood: identifying biological pathways and prenatal determinants". We greatly appreciate your careful consideration of our manuscript and constructive comments, which were used to guide our revisions. Overall, we feel that your recommendations have greatly improved our manuscript. Below, you will find an itemized, point-by-point response to your comments (indicated in bold).

RESPONSE TO REVIEWER COMMENTS

Reviewer #1

Thank you for the opportunity to review this manuscript submission. In this study, the authors have assessed the molecular profiles of childhood obesity and associated metabolic dysfunction using extensive multi-omics profiling (methylome, miRNome, transcriptome, proteins and metabolites) across two parts of Europe within the population-based Human Early Life Exposome project. An untargeted approach uncovered three clusters of children defined by specific multi-omics profiles, one of which characterized by higher adiposity as well as a high degree of metabolic complications (as assessed with associations to the IDEFICS continuous MetS score) and inflammation-related cascades. Looking at early exposures from a DOHaD perspective, pre-pregnancy body mass index and environmental pollutants like perfluorooctanoate (PFOA) and mercury were found to be significant determinants of this high-risk cluster in the overarching aim to identify specific targets for mitigation of childhood obesity. This work is impressive, comprehensive, and the manuscript is organized and well-written.

We thank the reviewer for their thoughtful and positive feedback on our manuscript. We appreciate the recognition of the study's scope, methodological approach, and clarity in writing. We are pleased that our work was well-received.

Comment 1.1. While this is not to suggest for the authors to do any additional analyses, it would be useful to the reader if the authors would defend their choice of this bottom-up clustering approach versus using the IDEFICS MetS score that they've derived as their outcome and subsequently investigate the multiomics profiles associated with this score (e.g. using LASSO). Especially with the emphasis of metabolic health beyond obesity. One can clearly see that the score is most significantly and consistently associated with the derived clusters.

Our response: We thank the reviewer for this insightful comment and the opportunity to clarify our methodological choice. The primary aim of our study was to derive mechanistic insights into the molecular pathways involved in childhood obesity and metabolic dysfunction. We did not aim to create a predictive risk score, nor to substitute the MetS score which consists of biomarkers that are readily accessible and relatively affordable to measure in clinical practice. While the MetS score is indeed a valuable clinical tool, it does not necessarily capture the complex and heterogeneous molecular features implicated in the obesity phenotype. Our decision to use an unsupervised clustering approach allowed for a data-driven exploration of metabolic subtypes that reflect underlying molecular mechanisms beyond traditional MetS criteria. This bottom-up approach has the advantage of revealing novel molecular signatures and pathways potentially relevant to metabolic health, which may otherwise remain undetected when analyses are constrained to a predefined outcome. By focusing on multi-omics clusters, our approach provides a broader perspective on the intricate molecular complexity of obesity and associated metabolic dysfunction. Our goal was to offer a more comprehensive view of the molecular architecture of obesity and metabolic dysfunction, which could inform intervention strategies in the future. We hope this rationale clarifies our choice of the bottom-up clustering approach. We have clarified this point in the discussion.

Added text:

[Lines 280-289] The primary aim of this study was to uncover the molecular mechanisms underpinning childhood obesity and associated metabolic dysfunction, rather than to develop a predictive risk score for metabolic outcomes. The bottom-up multi-omics clustering approach was selected to allow for the emergence of distinct molecular subtypes, enabling a data-driven exploration of metabolic health. While the metabolic syndrome score remains a valuable clinical tool for risk classification, it does not reflect the intricate molecular heterogeneity underlying obesity-associated metabolic dysfunction. By focusing on clusters derived from multi-omics profiles, our approach provides an understanding of biological pathways involved in metabolic health, beyond established clinical scores.

Comment 1.2. Line 57-end of introduction: this reads more like an abstract than the final paragraph of the introduction. Please refrain from having results in this paragraph (3 clusters, BMI, PFOA) and revise to have this section the summarize the objectives of the study.

Our response: We have revised the relevant section accordingly.

Added text:

[Lines 83-105] We therefore sought to interrogate multiple blood omics layers and a number of metabolic health outcomes in children to advance our understanding of the pathobiology of obesity, early in disease genesis. Our objective was to offer insights into the diversity of molecular profiles in metabolically healthy and unhealthy individuals. We hypothesized that children at risk for metabolic disease would show broad and aggregated alterations in blood omics analytes, whether as causes or effects of biological changes leading up to the disease. To explore this hypothesis, we leveraged the unique database of the population-based Human Early Life Exposome (HELIX) project (PMID: 30206078), one of the largest datasets of the general pediatric population with comprehensive multi-omics profiling (five molecular layers) and a rich phenotypic characterization of metabolic health from two regions in Europe.

Motivated by the emerging field of precision environmental health (PMID: 37117186), we also sought to examine the role of the prenatal environment to provide an additional level of understanding of potential contributors to disease etiology. Prenatal life is a particularly important period to study the environmental triggers of disease; exposures during this developmentally vulnerable period may have pronounced effects at the molecular level and disease risk later in life (PMID: 18596274). Using a comprehensive approach that incorporates key environmental, social, and lifestyle factors, we aimed to understand how early-life exposures are associated with molecular profiles linked to childhood obesity and metabolic dysfunction.

Overall, this study seeks to expand our understanding of the molecular mechanisms underpinning obesity and associated metabolic dysfunction and to identify modifiable environmental factors that could guide future prevention strategies.

Comment 1.3. Methods: Line 445: in the creation of the multiomics clusters and screening of published literature it would be useful to add (perhaps in supplemental Table 1) how many of the final included studies/molecules were from the same HELIX populations. As mentioned in Line 96 one objective was the replication of published literature and given HELIX is one of the largest birth cohorts, it is important to know how many of these molecules were from independent studies.

Our response: Following the reviewer's suggestion, we have added a column to Supplemental Table 1 indicating how many of the included omics molecules were previously identified in studies involving the HELIX population. Specifically, out of the 976 identified CpG sites, 10 were previously shown to

associate with childhood BMI in a PACE consortium analysis which included HELIX (PMID: 33239103). Additionally, out of the 98 identified metabolites associated with adiposity measures, 41 were also previously shown to associate with childhood BMI in HELIX (PMID: 30404627). For transcriptomics and miRNAs, no previous HELIX study examined associations with adiposity.

Comment 1.4. Also, when conducting the literature review, I think an important and comprehensive review on this topic in terms of metabolomic childhood obesity markers was not included, can the authors confirm? Would be useful to include it. [Handakas et al. Obesity Reviews. 2022;23(S1):e13384.]

Our response: We appreciate the reviewer's suggestion regarding this key systematic review. We confirm that we included the review by Handakas et al. (2022) in our literature review process. This systematic review is referenced in Supplementary Data 2, Table S2D (PMID: 34797026).

Comment 1.5. Investigating the association between the high-risk multi-omics cluster in childhood and subsequent obesity traits in adolescence in a subsample of our population is a point of strength; as is the investigation of the prospective in-utero factors.

Our response: We thank the reviewer for acknowledging these aspects of our study. We agree that investigating the association between the high-risk multi-omics cluster in childhood and subsequent obesity traits in adolescence is a point of strength. Additionally, the exploration of in-utero factors aligns with the goal of understanding the early-life origins of obesity and metabolic dysfunction.

Comment 1.6. Line 502: missing the word "to" handle….

Our response: We thank the reviewer for pointing this out.

Comment 1.7. Line 345: can you elaborate on the reproducible metabolomics measurements? Are these absolute concentrations or RPAs?

Our response: We acquired the serum metabolomics data using a standardized, targeted LC-MS/MS assay, which provides absolute concentrations of metabolites. This analytical method exhibits high sensitivity and specificity of the quantification, has high interlaboratory reproducibility and has been widely used in large-scale epidemiology studies (e.g., PMID: 36627359, 31495913, 27959516, 23043162, 20037589). We have updated the manuscript to clarify this point.

Added text:

[Lines 414-416] However, this analytical method provides absolute concentrations of metabolites, with unambiguous annotation, has high interlaboratory reproducibility, and has been widely used in large-scale epidemiology studies (PMID: 27959516, 23043162).

Comment 1.8. Discussion is very well written. Can you add insights to which omics layers were most predictive/valuable in the clustering? If other cohorts had limited funding, which analyses should they prioritize?

Our response: We thank the reviewer for their positive feedback on the Discussion section. In our analysis, certain omics layers proved particularly valuable for clustering. Specifically, miRNA expression and transcriptomics emerged as key contributors, and they were highly informative in distinguishing clustering profiles (as shown in Fig.3a). The proteomics layer was also significant, contributing to our understanding of inflammation-related pathways associated with the high-risk cluster.

For cohorts with limited funding, we recommend prioritizing transcriptomics and proteomics analyses. Transcriptomics offers a broad view of gene expression changes associated with obesity and metabolic dysfunction, while proteomics offers direct insight into functional protein alterations, particularly those linked to inflammatory and metabolic pathways. Together, these two layers are well-suited to identify key biological processes and are likely to yield clinically relevant insights into metabolic health profiles. We have included this point in the Discussion.

Added text:

[Lines 323-331] Although our study utilized several omics layers, we recognize that resource constraints may limit the scope of omics layers feasible for other studies. Of the layers we analyzed, miRNA expression, transcriptomics, and proteomics proved particularly valuable for clustering profiles related to metabolic dysfunction and identifying key molecular pathways. For resource-limited settings, focusing on transcriptomics and proteomics could yield meaningful insights, as these layers capture broad gene expression changes and protein-level functional alterations, respectively. Together, they can provide a comprehensive view of the biological pathways central to obesity and metabolic dysfunction, thereby enhancing the utility of these analyses.

Reviewer #2

The manuscript investigates various omics within blood samples collected from children using several advanced techniques, including arrays for DNA methylation, miRNA, and gene expression, as well as multiplex assays for plasma protein and targeted LC-MS/MS for metabolites. The authors integrated omic targets reported in the published literature and categorized the cohorts into three distinct clusters associated with different prenatal exposures and metabolic outcomes. There are several commendable strengths in this study, particularly the large sample size and the extensive data available for childhood outcomes and pregnancy information. The analytical approach is robust and highlights key targets that may have significant biological relevance. Additionally, the manuscript is well-written, making it easy to read despite the complexity of the analyses. However, there are several limitations that decrease the enthusiasm of this reviewer:

We thank the reviewer for recognizing the strengths of our study, including our large sample size, the extensive data on childhood outcomes and pregnancy, and the robust analytical approach. We appreciate the positive feedback on the clarity and readability of the manuscript. We also acknowledge the limitations pointed out by the reviewer, and we welcome the opportunity to address them to improve the manuscript. We have carefully considered each limitation mentioned and have provided responses to each specific point below.

Comment 2.1. The study designs and analyses do not elucidate the pathophysiology or mechanistic understanding of prenatal environments as suggested in the abstract and several parts of the manuscript. Instead, they demonstrate associations. While the authors acknowledge this in the discussion, the abstract and text within the manuscript present somewhat confounding information. For example, lines 15-17 state, "identify precise targets for prevention and intervention strategies early in the life-course...," yet the study identifies associated prenatal risk factors that are potentially modifiable without establishing them as precise targets for prevention (i.e. clinical studies) . Similarly, lines 71-73, "how prenatal environmental factors contribute to disease risks...," imply causation, which is not supported by the data.

Our response: We thank the reviewer for this important observation. We agree that our study identifies associations rather than causal relationships and have clarified this in the manuscript. We have revised the abstract and main text to ensure that the language reflects the associative nature of our findings. Specifically:

- In the abstract, we have rephrased statements to clarify that the study identifies potential risk factors and highlights associative findings rather than definitive targets for prevention or causal mechanisms.
- Similarly, throughout the main manuscript, we have modified wording that may imply causation. For example, we now describe prenatal environmental factors as linked to or associated with disease risk, rather than contributing directly to it.

While causality cannot be determined within the scope of our study, the large-scale nature and depth of data provide a wealth of information not commonly available in the literature. We believe this unique dataset offers valuable insights into early-life risk factors for childhood obesity and metabolic dysfunction, laying the groundwork for further research into causality.

Added text:

[Lines 52-54] Overall, our work helps identify potential risk factors for prevention and intervention strategies early in the life-course, aimed at mitigating obesity and its long-term health consequences.

[Lines 99-102] Using a comprehensive approach that incorporates key environmental, social, and lifestyle factors, we aimed to understand how early-life exposures are associated with molecular profiles linked to childhood obesity and metabolic dysfunction.

[Lines 246-247] Finally, we sought to examine how the prenatal environment is associated with cluster membership.

[Line 250-254] Building on previous research showing that exposome components and their impact exhibit considerable variability across regions owning to different population characteristics and exposure patterns (PMID: 30530161, 30024382), we examined associations with the multi-omics clusters separately in the N/W and S/M cohorts.

[Lines 275-279] We further incorporated information from the environmental "riskscape" of the critical pregnancy period, capturing widespread chemical exposures, outdoor and built environmental exposures, and demographic and lifestyle factors, to explore potential environmental contributors to disease risk early in the life-course.

[Lines 370-372] In our study, we built upon this framework by introducing the prenatal environment as an additional dimension and employing a robust data science method to explore associations with health risk.

[Lines 393-395] Overall, our results underscore the importance of the modifiable prenatal environment linked to subsequent disease risk and highlight the need to tailor prevention guidelines to accommodate diverse country contexts.

[Lines 426-428] Further, given its observational design, our study identifies associations between prenatal factors and obesity-related outcomes but does not establish causation.

[Lines 444-449] In summary, our study delves into the intricate molecular biology of childhood obesity and metabolic dysfunction. Through the integration of multi-omics profiling and prenatal environmental and lifestyle data, we unveil distinct metabolic clusters of children and identify potential modifiable risk factors. Our results may help to inform early-life prevention and intervention strategies aimed at combating obesity and its long-term health consequences.

Comment 2.2. The study's generalizability is limited by its demographic composition, with the majority of participants being White (85% in N/W and 100% in S/M) and the remainder categorized as Asian or Other. Given the study's European location, this is understandable. However, it raises questions about whether the same clusters and associations would be observed in more diverse populations. Addressing this factor would be important for reaching a wider, global audience.

Our response: We appreciate the reviewer's observation on the demographic composition of our study population and the impact this may have on the generalizability of our findings. Given that our data were collected from European cohorts, the majority of participants are White/Caucasian, reflecting the demographics of the regions involved in the study. We recognize that the metabolic clusters and associations observed may vary across diverse populations due to genetic, environmental, and lifestyle differences. Expanding future research to include more diverse populations would be important for determining the generalizability of clusters across different demographic groups. We now acknowledge this limitation in the Discussion section to highlight the need for further studies involving more ethnically and geographically diverse populations.

Added text:

[Lines 439-443] Our study provides insights into childhood obesity and metabolic dysfunction in a European cohort that mostly consists of White/Caucasian participants. Expanding future research to include more ethnically and geographically diverse populations would be important for determining the generalizability of multi-omics clusters and associations across different demographic groups.

Comment 2.3. While the study benefits from extensive prenatal data collection, several critical pieces of information are missing, which hampers the interpretation of prenatal and early life factors. Notably, maternal conditions such as diabetes (PMID: 30008062, 38820461) and preeclampsia (PMID: 36378310), which are known to impact offspring's metabolic and cardiovascular health, as well as cytokine levels, are not discussed. Including these data would strengthen the analyses and conclusions.

Our response: We appreciate the reviewer's observation and agree that maternal conditions such as diabetes and preeclampsia are important factors influencing offspring metabolic and cardiovascular health. Unfortunately, in our study, a substantial amount of data for these conditions was missing, which limited our ability to include them in our multi-exposure models. Specifically, in the N/W cohort, we had 57% missing data for diabetes (n/N = 316/557) and 49% for preeclampsia (n/N = 273/557). Similarly, the S/M cohort had 21% missing data for diabetes (64/306) and 55.6% for preeclampsia (170/306). This high level of missing information would have prevented us from accurately and meaningfully assessing the impact of these maternal health conditions. Nonetheless, we considered pre-pregnancy BMI and gestational weight gain which have been strongly linked, across their full spectrum, to maternal metabolic health (e.g., PMID: 39366732, 21976280, 16029839, 39201703, 17416786, 26141788). We now acknowledge this in the Discussion section.

Added text:

[Lines 428-432] Although our study included extensive prenatal data, we were unable to assess the effects of maternal diabetes and preeclampsia- two conditions closely associated with offspring's metabolic health (PMID: 38820461, 36378310)- due to substantial missing data. However, we included pre-pregnancy BMI and gestational weight gain, which have been strongly linked to maternal metabolic health (PMID: 16029839, 26141788, 39366732).

Comment 2.4. Most results are presented using measures that demonstrate the magnitude of contribution (e.g., odds ratio, beta coefficient, iSHAP). However, presenting the mean levels of elevated cytokines and branched-chain amino acids in these clusters would be informative for both clinical and basic scientists. This data should be included, at least in a supplemental figure, to aid determining clinically relevant cutoffs in both clinical studies and basic science experiments.

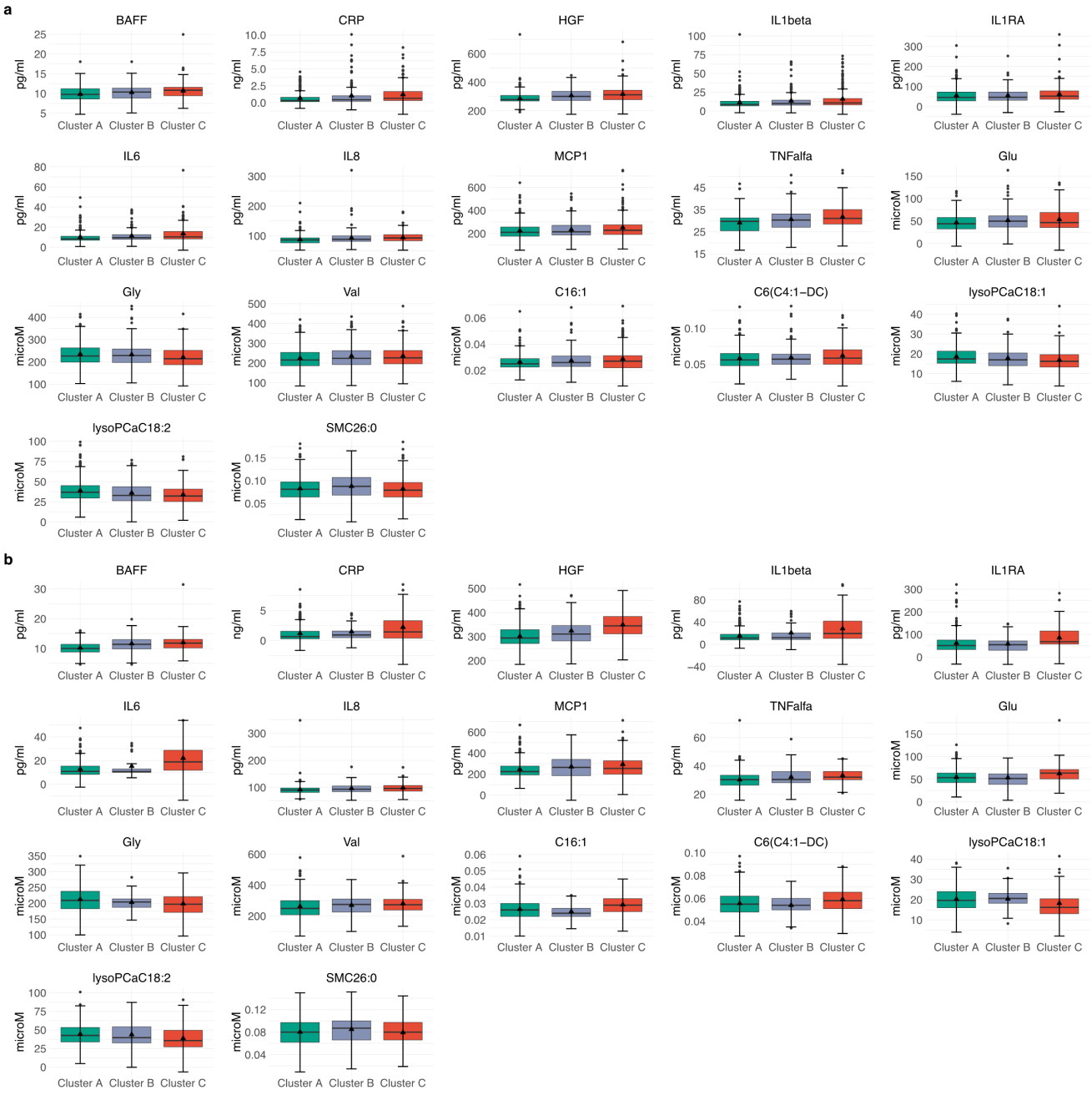
Our response: We agree with the reviewer and have now added a supplemental figure that displays the mean levels of cytokines and metabolites within each cluster across the participating cohorts.

Added figure:

[Supplementary Material]

Supplementary Fig. 5: Distribution of selected cytokines and metabolites across Clusters.

a *Northern/Western cohort.* ***b*** *Southern/Mediterranean cohort. The box-plots represent descriptive statistics for each selected omics biomarker across Clusters. The median value is shown by the black horizontal line within the box, the mean value by the black triangle within the box, the first and third quartiles as the lower and upper border of the box, respectively, and the 1.5 interquartile range by the vertical black lines. Outliers are shown as black circles.*



Comment 2.5. In Figure 4a, maternal BMI, PFOA, and mercury levels appear to have different impacts in the N/W and S/M regions. For instance, maternal BMI differences may not be clinically actionable given the small differences within clusters, and the BMI in S/M region is actually normal in all clusters. The blood PFOA and mercury levels are somewhat modest in N/W (5-10% difference), raising some questions about the biological impact in general. Additionally, the mercury levels of both N/W and S/M appears to be extremely high. (1.65-4.1 ug/dl as mean) The 95% mean levels of US populations is ~0.4ug/L, with mean levels much lower than that. (https://www.epa.gov/americaschildrenenvironment/biomonitoring-mercury). A study in German also reported ranges comparable to the US population. (PMID: 34973943) This is the same for blood PFOA levels – the range reported by CDC is between 1.5-2 ug/L (https://www.cdc.gov/exposurereport/data_tables.html?NER_SectionItem=NHANES) and another study (PMID: 36574487), the unit listed in supp table is mg/dL and that would make the PFOA levels extremely high.

Our response: We thank the reviewer for these insightful comments. As described in the Results and Methods Section, we examined the associations of prenatal environmental factors with the multi-omics clusters separately in the N/W and S/M cohorts, given prior research indicating that

environmental components and their impacts exhibit considerable variability across regions due to population-specific characteristics and exposure patterns. In the N/W cohort, maternal pre-pregnancy BMI and exposure to PFOA emerged as the main factors in LASSO, while in the S/M cohort, maternal mercury exposure was highlighted.

There was an error in Supplementary Table 9, where the original units for blood PFOA were incorrectly listed; the correct unit is μg/L. We have corrected this and also updated the table to include additional univariate metrics (geometric mean, 25th percentile, median, 75th percentile, 90th percentile, and 95th percentile) to facilitate comparison with NHANES and other datasets.

While the differences in maternal BMI within clusters may appear modest, even small shifts at the population level can have significant public health implications. An extensive body of research suggests that increasing maternal BMI associates, across its full spectrum, with higher offspring obesity and metabolic risk (PMID: 31185012, 30232419, 26525168, 22344037). Assuming causality in this association, even slight increases in maternal BMI within a population could shift the overall distribution of metabolic health outcomes toward the upper end (i.e., higher-risk categories), where complications can become more prevalent.

The observed differences in blood PFOA and mercury levels between clusters in the N/W cohort are somewhat modest, particularly at lower percentiles. However, the between-cluster differences at the upper level of the chemical distribution, and especially for PFOA, become more pronounced (reaching up to 16%). Given the cumulative, long-term, and widespread exposure to PFOA, even small variations can have substantial biological relevance at the population level. Our study was able to capture an association between increasing PFOA concentration and higher risk for membership to Cluster C. This is in agreement with previous studies including participants with similar or even lower PFOA concentrations that report an increasing risk of childhood obesity or metabolic complications with increasing PFOA levels in a dose-response manner (PMID: 33418195, 27857130, 22935244, 27352404, 29347948, 26554535).

Recruitment of pregnant women in our study cohorts, and consequently measurement of prenatal chemical concentrations, occurred between 1999-2010. PFAS have been in production for over 55 years, resulting in ubiquitous exposure levels. Maternal blood PFAS concentrations in our study are lower than those reported in the U.S. NHANES female population during the same period (collection cycles: 1999-2010, geometric mean ranging from 2.69 to 4.80 μg/L) and higher than those reported in the most recent collection cycles of 2011–2018 (geometric mean of 1.26-1.84 μg/L). Similarly, PFOA levels in our study are lower than those reported in the U.S. HOME cohort (median: 5.3 μg/L, PMID: 26554535) and U.S. Project Viva cohort (median: 5.6 μg/L, PMID: 27352404), in which PFOA measurement occurred between 1999-2006, and slightly higher than the New Hampshire Birth Cohort (median: 1.44 μg/L) which measured blood maternal PFAS levels between 2009-2018 (PMID: 36162478). This difference in PFOA levels assessed over time probably reflects global efforts to reduce PFAS usage through regulatory and voluntary actions (PMID: 39155039, 34682663).

Regarding mercury, our study population has levels that are broadly relevant to many populations around the world. The maternal levels of mercury in the N/W cohort (median: 1.36 μg/L) and S/M cohorts (median: 3.16 μg/L) are between those reported in Faroes (geometric mean: 22.2 μg/L, PMID: 24681285), Seychelles (median: 15.97 μg/L, PMID: 30660840), China (median: 4.93 μg/L, PMID: 24847687) and South Korea (median: 2.94 μg/L, PMID: 24847687), and those reported in Germany (geometric mean between 2001-2010: 0.77-1.76 μg/L, PMID: 34973943) and the U.S. NHANES (2005–2006-cycle geometric mean: 0.86 μg/L, 2017–2018-cycle geometric mean: 0.63 μg/L). Variations in blood mercury levels may reflect differences in dietary exposure patterns and particularly in fish consumption, which is a major source of human exposure to mercury. Coastal regions tend to have higher fish intake.

Added table:

[Supplementary Material]

Supplementary Table 9: Distribution of selected prenatal factors by multi-omics cluster membership

	Northern/Western region				Southern/Mediterranean region			
	Cluster A (N=227)	Cluster B (N=150)	Cluster C (N=180)	Total (N=557)	Cluster A (N=238)	Cluster B (N=21)	Cluster C (N=47)	Total (N=306)
Pre-pregnancy BMI, kg/m²								
Mean (SD)	25.01 (4.78)	26.19 (5.64)	26.02 (5.45)	25.66 (5.26)	24.09 (4.69)	24.13 (5.20)	24.73 (4.01)	24.19 (4.62)
Geometric Mean (SD)	24.59 (1.2)	25.63 (1.23)	25.48 (1.23)	25.16 (1.22)	23.7 (1.19)	23.7 (1.21)	24.45 (1.16)	23.81 (1.19)
25th Percentile	21.22	22.06	21.88	21.61	21.16	21.53	22.43	21.34
Median	24.37	25.37	25.27	24.78	23.01	22.73	23.75	23.11
75th Percentile	27.56	29.29	28.95	28.62	25.71	23.95	25.71	25.6
90th Percentile	32.09	33.61	33.51	33.25	29.23	31.68	29.22	29.38
95th Percentile	34.27	36.7	35.92	35.3	35.02	36.43	31.39	35.25
Maternal PFOA, µg/L								
Mean (SD)	2.17 (1.31)	2.29 (1.55)	2.46 (1.69)	2.29 (1.51)	2.77 (1.72)	3.19 (1.33)	3.43 (4.52)	2.90 (2.35)
Geometric Mean (SD)	1.81 (1.85)	1.83 (2.02)	2 (1.9)	1.88 (1.91)	2.26 (2.09)	2.62 (2.46)	2.65 (1.92)	2.33 (2.09)
25th Percentile	1.16	1.14	1.22	1.18	1.83	2.33	2.27	1.91
Median	1.84	1.88	2.09	1.92	2.57	3.34	2.54	2.6
75th Percentile	2.94	3.05	3.2	3.04	3.30	3.88	3.30	3.39
90th Percentile	4.18	4.4	4.56	4.28	4.56	4.23	4.10	4.51
95th Percentile	4.75	4.88	5.61	5.13	5.24	4.43	4.70	5.23
Maternal mercury, µg/L								
Mean (SD)	1.66 (1.38)	1.56 (1.12)	1.71 (1.31)	1.65 (1.29)	3.94 (3.16)	3.41 (2.13)	5.19 (5.53)	4.10 (3.60)
Geometric Mean (SD)	1.23 (2.22)	1.19 (2.19)	1.29 (2.18)	1.24 (2.2)	2.99 (2.13)	2.86 (1.86)	3.69 (2.3)	3.08 (2.14)
25th Percentile	0.71	0.72	0.73	0.72	1.8	1.79	2.29	1.91
Median	1.27	1.36	1.45	1.36	3.06	3.07	3.74	3.16
75th Percentile	2.12	2.09	2.16	2.15	5.07	4.25	6.68	5.18
90th Percentile	3.52	2.98	3.29	3.39	7.65	6.68	8.65	7.71
95th Percentile	4.04	3.59	3.92	3.95	9.41	7.34	11.32	9.41

Comment 2.6. One key message that is lacking within the manuscript is the information regarding transitioning from metabolically healthy obesity to metabolically unhealthy obesity. (PMID: 38941611, PMID: 33125374). This is crucial, as it is unclear whether children in cluster B are genuinely healthy or if they will become metabolically unhealthy. This aspect should be emphasized to avoid giving false reassurance about the health of metabolically healthy obese children. Although briefly mentioned in lines 146-152 and supplemental figure 3, more comprehensive data on metabolic syndrome risk scores is needed.

Our response: We thank the reviewer for this important observation. We agree that understanding the transition from metabolically healthy to metabolically unhealthy obesity is essential. Although our findings indicate that children in Cluster B exhibit a relatively healthier metabolic profile compared to those in Cluster C, we recognize that these children may still be at risk of transitioning to a metabolically unhealthy state over time. Given the cross-sectional nature of our primary multi-omics and metabolic health data in childhood, we were not able to fully assess the dynamics of these transitions within the cohorts. However, we assessed prospective associations of multi-omics clusters with weight-related outcomes in adolescence for a subsample with available anthropometric measures (Supplementary Fig 3) and we observed relatively weak associations for Cluster B.

We acknowledge that further detailed longitudinal data would be necessary to evaluate the trajectory of multi-omics profiles and their association with metabolic health status over time. We have now added a section highlighting this point in the Discussion.

Added text:

[Lines 315-322] Children in Cluster B appeared to have a relatively healthier metabolic profile; however, the potential for transition to a metabolically unhealthy state remains an important consideration, as these children may still be at risk for future metabolic complications. Recent studies highlight that metabolically healthy individuals with obesity carry a health risk and may transition to an unfavorable metabolic state (PMID: 28919065, 29581366, 38941611, 33125374). Further research incorporating longitudinal follow-up data is needed to examine the trajectory of multi-omics profiles and provide insights into the risk of metabolic complications in the future.

Comment 2.7. For data reproducibility, it is important to state if blood was collected after fasting and how the blood was collected, as these factors can significantly influence metabolite, miRNA, and transcriptome profiles.

Our response: We appreciate the reviewer's attention to these details, as they are indeed important for data reproducibility. In our study, blood samples were collected at the end of clinical examination, following a median (IQR) fasting time of 3.4 (2.8, 4.0) hours, and processed according to standardized protocols. We have now included this information in the Methods section for clarity.

Added text:

[Lines 473-478] Blood samples were collected during the childhood follow-up visit at the end of the clinical examination, following a median (IQR) fasting time of 3.4 (2.8, 4.0) hours. Blood samples were collected using a 'butterfly' vacuum clip and processed into a variety of sample matrices: EDTA used for plasma proteomics, miRNAs, and DNA isolation; tempus tubes for RNA isolation; and plastic silica Vacutainers for serum metabolomics. All samples were processed and frozen at −80°C under standardized procedures.

Reviewer #3

Understanding the molecular mechanism of childhood obesity, including biomarker signatures and environmental determinants, is crucial for developing preventive and therapeutic measures. This manuscript employs a systems biology approach with deep, multi-omics profiling to identify and validate molecular signatures of childhood obesity/metabolic dysfunction and to characterize different subgroups based on metabolic traits and molecular pathways. Overall, it is a solid study and highly relevant to the field. Addressing the following questions will help enhance the study's impact and clarity.

We thank the reviewer for their positive feedback on the relevance and impact of our study. We appreciate the recognition of our systems biology approach and its potential contributions to advancing the understanding of childhood obesity and metabolic dysfunction. We have addressed each of the reviewer's questions and suggestions in detail to further enhance the clarity and impact of the manuscript. We believe that these additions have strengthened the overall presentation and relevance of our findings.

Comment 3.1. In the "Molecular drivers of the multi-omics clusters" section, the authors have identified several drivers that significantly contribute to defining high-risk Cluster C, including important genes. Please integrate with published bulk (GTEx, etc.) or scRNA-seq data to determine which tissues/cell types highly express these risk genes within Cluster C. This will reveal additional mechanisms of childhood obesity.

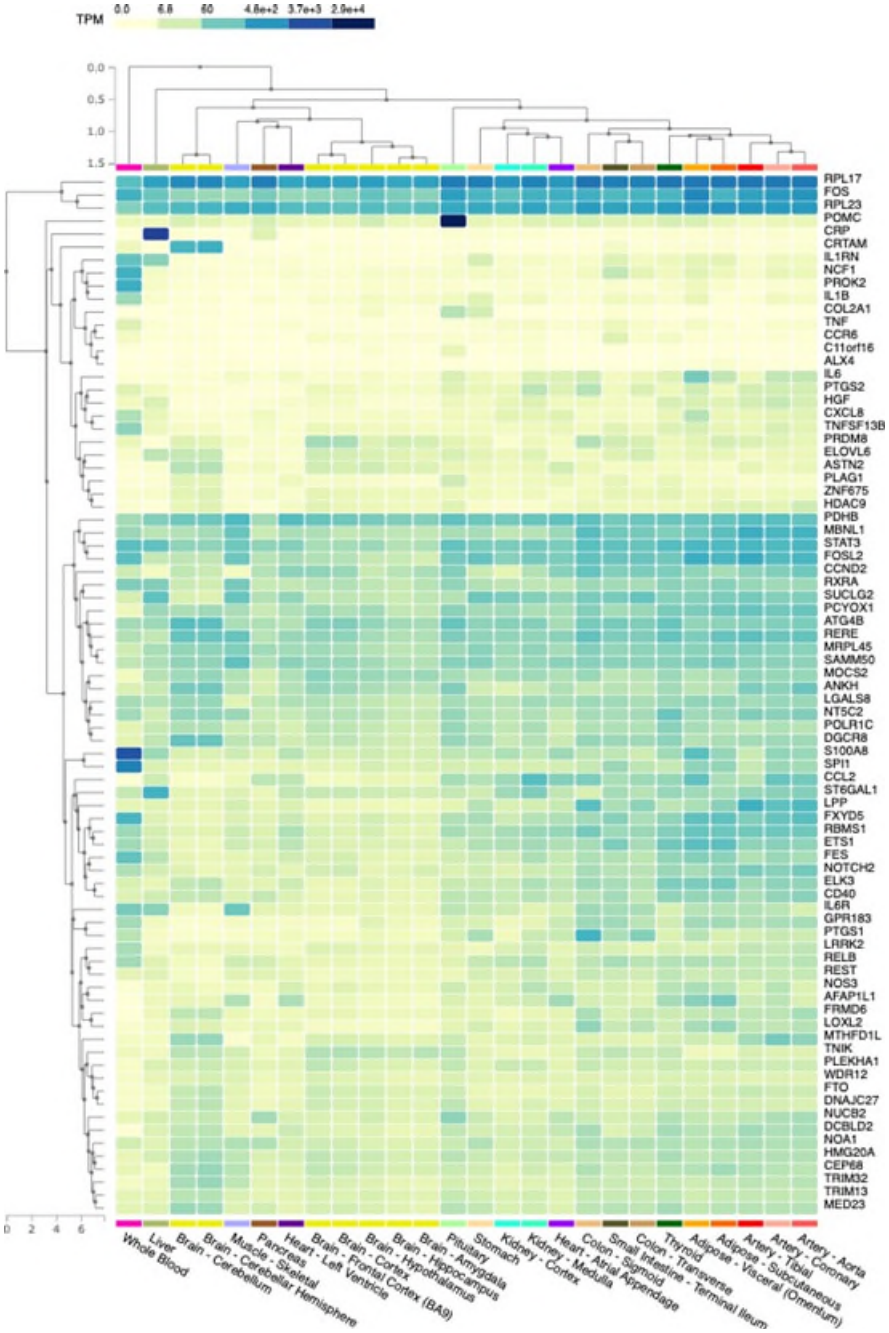
Our response: We thank the reviewer for this great suggestion. We agree that integrating our findings with published expression data would provide additional insights into the mechanisms underlying childhood obesity in high-risk Cluster C. In response, we analyzed the identified genes using publicly available expression data from GTEx for bulk tissue expression to determine which tissues exhibit high expression of these genes.

Added text and figure:

[Lines 237-242] To complement these findings, we further integrated the annotated genes with data from the Genotype-Tissue Expression (GTEx) public database (PMID: 32913098) to identify the human tissues in which these genes are most highly expressed (Supplementary Fig. 9). Several of these genes showed high expression in adipose tissue, liver, skeletal muscle, and pancreas, underscoring the role of these tissues in the metabolic and inflammatory characteristics of Cluster C.

[Lines 622-625] We further evaluated tissue-specific expression patterns of the mapped genes by using the publicly available dataset of the GTEx project (reference according to recommended cite usage of GTEx data*: The data used for the analysis described in this manuscript were obtained from the GTEx Portal (https://www.gtexportal.org/home/) on 10/12/24.). Expression levels were measured in Transcripts Per Million (TPM) to enable accurate comparisons across tissues.*

[Supplementary Material] ***Supplementary Fig. 9: Tissue-specific gene expression linked to the top features contributing to the definition of Cluster C****. Genes presented were mapped using the top molecular features identified to contribute to the definition of Cluster C. Expression levels are measured in Transcripts Per Million (TPM) and are based on data from the Genotype-Tissue Expression (GTEx) database. The data used were obtained from the GTEx Portal on 10/12/24.*



Comment 3.2. Please clarify why cluster A was chosen over cluster B as the reference category in Fig 2? What criteria were used for this selection.

Our response: Cluster A was selected as the reference category because it represents the group with the lowest levels of adiposity and metabolic complications. Using Cluster A as the reference allowed us to compare the more adverse metabolic profiles observed in Clusters B and C against a relatively healthier baseline. This approach provides a clearer contrast in terms of metabolic health, and hence, a more straightforward comparison, facilitating interpretation of the relative differences among the clusters. We have now clarified this in the Methods section.

Added text:

[Lines 576-578] Cluster A was chosen as the reference category for comparative analyses because it exhibited the lowest levels of adiposity and metabolic complications, thereby providing a clearer baseline for assessing metabolic health differences across clusters.

Comment 3.3. Please describe how specific environmental exposures interact with genetic factors to influence obesity risk.

Our response: Thank you for this insightful comment. We agree that understanding the interactions between environmental exposures and genetic factors is critical to elucidate disease risk. Examining gene-environment interactions is beyond the scope of this manuscript. Our study's primary focus was to characterize multi-omics clusters in childhood obesity and metabolic dysfunction and identify associated environmental exposures. Investigating gene-environment interactions would necessitate additional complex analytical frameworks to accurately assess these relationships, such as environmental-wide interaction models, and a focused study design including identification of specific target genes and construction of polygenic risk scores which capture cumulative effects of multiple variants, improve predictive power and can be more representative of population-level risk. We are currently conducting a separate, detailed analysis that involves also this cohort and is specifically focused on polygenic risk and gene-environment interactions. The findings from this ongoing analysis will be reported in a future manuscript, allowing for a more in-depth exploration of these complex relationships. Additionally, we have included relevant text in the Discussion section to acknowledge the potential impact of gene-environment interactions on disease risk.

Added text:

[Lines 395-402] We acknowledge that potential gene-environment interactions may influence disease risk. For instance, a small study in the Faroes cohort recently showed that PFAS health effects and in particular associations observed with insulin sensitivity could vary between individuals as a result of genetic predisposition involving gene variants related to lipid and glucose metabolism (PMID: 36868448). Large-scale studies incorporating polygenic risk scores and environment-wide interaction models will be needed to comprehensively assess how genes interact with the environment to shape health outcomes.

Comment 3.4 Are there any untargeted approaches available to identify new biomarkers and risk factors?

Our response: In our study, we employed both targeted and untargeted approaches to capture a broad range of molecular features across multiple omics layers. Blood DNA methylation, blood gene expression, and blood miRNA expression were assessed using untargeted methods. Plasma proteins and serum metabolites were assessed using targeted approaches. We now better clarify this in the Methods section. For the prenatal risk factors, we employed a targeted approach, assessing 37 key environmental characteristics previously reported to associate with maternal and/or offspring health, to provide relevant context on prenatal exposures. We recognize that other untargeted technologies—such as high-resolution mass spectrometry (allowing profiling of both internal metabolites and environmental exposure agents), shotgun proteomics (which can identify a broad array of proteins including low-abundance proteins), single-cell RNA sequencing, and metagenomics (which can give

insight into the diversity and functional potential of microbiota)—could provide even greater insights. We have noted this in the limitations section, emphasizing that future studies using these advanced techniques could enhance the understanding of molecular pathways involved in childhood obesity and help identify novel biomarkers and risk factors.

Added text:

[Lines 416-420] While our study employed both targeted and untargeted methods, additional untargeted technologies—such as high-resolution mass spectrometry, shotgun proteomics, single-cell RNA sequencing, and metagenomics—could enhance the understanding of molecular pathways involved and help identify novel biomarkers and risk factors (PMID: 38066102).

[Lines 479-484] We used both targeted and untargeted methods to assess molecular features across five omics layers. For untargeted profiling, we assessed blood DNA methylation with the Illumina 450 K array, blood gene expression using the Affymetrix HTA v2.0 array and blood miRNA expression using the Agilent SurePrint Human miRNA rel 21 array. For targeted profiling, we assessed plasma proteins using three Luminex multiplex assays and serum metabolites with the LC-MS/MS Biocrates AbsoluteIDQ p180 kit.

Comment 3.5. Line 100, "GpG sites" should be "CpG sites." Supplementary Table 1, there is a discrepancy: the number of DNA methylation sites (N=976) does not match the content on line 100, which refers to 977 CpG sites. Additionally, please include information regarding the protein section in Supplementary Table 1 and Supplementary Data 2A-D.

Our response: We thank the reviewer for identifying these points. We analyzed 976 CpG sites. We have now corrected the typographical errors on the discrepancy in the number and wording regarding the CpG sites.

As for the proteome data, no filtering was applied. Only 35 proteins were measured using our panels, and all were included in the analysis; this is now further clarified in the Methods section. Therefore, additional information regarding protein selection is not applicable for Supplementary Table 1 or Supplementary Data 2A-D.

Added text:

[Lines 540-542] For the proteome data, no filtering was performed; our panels assessed a total of 35 proteins, all of which were included in the analysis.

Reviewer #4

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

We thank the reviewer for their collaborative effort and feedback.

Dear Reviewers,

Thank you for the opportunity to respond to the evaluations of our manuscript #NCOMMS-24-25298A, titled "Multi-omics architecture of obesity and metabolic dysfunction in childhood uncovers biological pathways and prenatal determinants." [Please note the change in title to adhere to the journal's guidelines]. We are grateful for the reviewers' thoughtful feedback and constructive suggestions throughout the review process, which have significantly improved our manuscript. Below, we provide a detailed, point-by-point response to the remaining comments (indicated in bold).

RESPONSE TO REVIEWERS' COMMENTS

Reviewer #1

The authors have answered all my questions and I am satisfied with the now strengthened manuscript.
We thank the reviewer for their thoughtful feedback and the time devoted on our manuscript during the review process.

Comment 1.1. A minor comment: line 326 has a word missing.
Our response: We appreciate your attention to detail. The missing word has been added, and the sentence has been corrected in the revised manuscript.

Comment 1.2. Furthermore, I leave it up to the authors if they can add as supplemental, a stratified analysis by sex. There have been few studies on sex differences in the metabolic profiles/obesity profiles in childhood and it can be interesting for the reader to have these results.
Our response: Thank you for this suggestion. We acknowledge the importance of exploring sex differences in metabolic/obesity profiles. We have now included a post-hoc stratified analysis by sex (Supplementary Table 4).

Added text:

[Lines 170-174] In a post-hoc sensitivity analysis, we examined whether sex had any influence in the observed associations for fat mass and metabolic risk (Supplementary Table 4). We observed no substantial evidence to indicate effect modification by sex, as reflected in the largely overlapping 95% confidence intervals of the sex-specific effect estimates.

Supplementary Table 4: Associations of multi-omics clusters with metabolic health outcomes by cohort and sex

	Northern/Western cohort			Southern/Mediterranean cohort			Pooled population		
	Males	Females	P interaction[c]	Males	Females	P interaction[c]	Males	Females	P interaction[c]
Fat mass[a]									
Cluster A	Ref.			Ref.			Ref.		
Cluster B	0.29 (0.01, 0.57)	0.21 (-0.09, 0.51)	0.707	0.08 (-0.52, 0.68)	0.77 (0.14, 1.41)	0.121	0.3 (0.05, 0.55)	0.33 (0.06, 0.60)	0.775
Cluster C	0.30 (0.02, 0.58)	0.30 (0.04, 0.57)	0.962	0.76 (0.37, 1.15)	0.66 (0.16, 1.15)	0.745	0.43 (0.21, 0.66)	0.40 (0.17, 0.64)	0.935
Metabolic syndrome score[a,b]									
Cluster A	Ref.			Ref.			Ref.		
Cluster B	0.53 (-0.05, 1.11)	0.72 (0.01, 1.43)	0.703	1.03 (-0.28, 2.35)	1.17 (-0.33, 2.67)	0.891	0.66 (0.13, 1.18)	0.90 (0.26, 1.53)	0.383
Cluster C	1.09 (0.51, 1.67)	0.80 (0.17, 1.44)	0.497	1.65 (0.79, 2.51)	2.1 (0.91, 3.28)	0.545	1.27 (0.80, 1.75)	1.11 (0.55, 1.67)	0.834

[a] Effect estimates represent beta coefficients (expressed in SD) and their 95% CIs, derived from generalized linear regression models controlled for study site and age at examination.
[b] The metabolic syndrome score was derived using z scores for waist circumference, HDL cholesterol level, triglyceride level, insulin level, and systolic and diastolic blood pressure.
[c] Values represent P-values for sex*cluster (interaction) estimates.

Reviewer #2

Thank for addressing the comments. Most of the comments have been addressed adequately. The revised manuscript was not able to address causality given the nature of the study, and did not have sufficient data collected to inform role of other maternal contributions (e.g maternal diabetes, pre-eclampsia) or socioeconomic status. Authors have discussed these limitations within discussion section.

No further changes were suggested. We are grateful for your constructive feedback and time spent throughout this process.

Reviewer #3

The revised manuscript is significantly improved. For clarity and completeness, we recommend the following minor changes:

We are grateful for your constructive feedback and time spent throughout the review process.

Comment 3.1. Please highlight genes with high expression in adipose tissue, liver, skeletal muscle, and pancreas in bold or a different color in supplementary Fig. 9 for readability.

Our response: As suggested by the reviewer, we have updated Supplementary Fig.9 to emphasize genes with high expression (TPM ≥80th percentile) in adipose tissue, liver, skeletal muscle, and pancreas. These genes are now highlighted with bold labels to improve clarity and readability.

Added text:

Supplementary Fig. 9: Tissue-specific gene expression linked to the top features contributing to the definition of Cluster C. Genes presented were mapped using the top molecular features identified to contribute to the definition of Cluster C. Expression levels are measured in Transcripts Per Million (TPM) and are based on data from the Genotype-Tissue Expression (GTEx) database. The data used were obtained from the GTEx Portal on 10/12/24. Genes with high expression values (TPM ≥80th percentile) in adipose tissue, liver, pancreas, and skeletal muscle are highlighted in bold.

Comment 3.2. Important discussions on transcriptome, proteome and metabolome are included in the Discussion sections, but little on methylome and miRNome. Please consider

including discussion on how methylome and miRNome contribute to childhood obesity, especially in light of results in Fig. 3a.

Our response: We agree that these layers may provide important insights into the biological processes underlying obesity and metabolic dysfunction, and have now expanded the discussion to provide some perspective on the role of methylome and miRNome in light of our results.

Added text:

[Lines 364-379] We identified a signature of 14 miRNAs playing a significant role in the definition of the high-risk Cluster C. Among the miRNAs identified, four in particular (miR-23a, miR-24, miR-130a and miR-21) have gathered increased attention for their role in regulating key metabolic processes. Specifically, miR-23a and miR-24 are part of the miR-23~27~24 family which has been suggested to control effector immune cell responses and has been implicated in various physiological and pathological processes, including the atherosclerotic process.[56-62] Moreover, miR-130a has been reported to regulate the proliferation of vascular smooth muscle cells and angiogenesis,[63-65] potentially contributing to vascular remodeling and altered blood pressure.[66-68] miR-21 exerts pleiotropic effects in human metabolism, including roles in insulin action[69,70] and hepatic inflammation.[71] Additionally, several DNA methylation markers were identified as contributing to the definition of Cluster C, albeit to a lesser extent than miRNAs. The CpGs cg08462942 and cg09615786 emerged as the strongest contributors. Notably, cg08462942 is annotated to ATG4B, an autophagy-related gene that has been implicated in adipogenesis.[72-74] The CpG cg09615786 is annotated to DCBLD2 which has been linked to endothelial function.[75]

Reviewer #4

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

We appreciate your collaborative effort in reviewing the manuscript.