

Genetic Variations Interact with Polybrominated Diphenyl Ether Exposure to Alter Lipid Homeostasis

Corresponding Author: Professor Xiaoping Miao

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)

Hu N. et al. conducted an interesting study on individual heterogeneity in response to environmental exposure, with a specific focus on PBDE-induced dyslipidemia. By integrating exposome, genomics, and metabolomic data from the China National Human Biomonitoring cohort, the authors identified 3,571 genetic variants that interact with PBDE exposure to influence dyslipidemia risk with key metabolites serving as mediators. Among these, rs9869609 was pinpointed as the top causal variant, exhibiting the strongest association with hypercholesterolemia risk. Further functional validation through single-base CRISPR/Cas9 editing revealed that the rs9869609-G allele downregulates SLC6A20 expression by enhancing BHLHE40 binding. This, in turn, impaired glycine transport and promoted cholesterol accumulation, particularly under BDE-47 exposure.

This study provides a comprehensive genome-wide analysis of gene-environment interactions in metabolic diseases. It makes a significant contribution by identifying vulnerable populations and clarifying the mechanistic pathways underlying environmentally induced metabolic disorders. Additionally, it opens new avenues for research on the health effects of environmental pollutants. However, several key limitations need to be addressed to improve the quality and impact of the manuscript.

Major comments:

1. The authors constructed a polygenic risk score (PRS) using independent variants to predict lipid phenotypes, finding increased risk only in populations with high PBDE exposure. Why does the PRS not confer risk in low-exposure groups? This observation warrants further discussion

2. While the study identifies loci interacting with PBDE exposure, it does not clearly differentiate this approach from traditional GWAS for disease-associated loci. Please elaborate on why gene-environment interaction (G×E) loci may offer greater biological or clinical relevance than main-effect GWAS hits, particularly for environmental metabolic disorders.

3. In metabolomic analysis, the authors focus on overlapping metabolites for pathway enrichment and mediation analysis, potentially missing broader mechanistic insights. A supplementary correlation analysis of all significant metabolites could reveal additional candidate molecules and pathways.

4. The choice of rs9869609 (over the most statistically significant SNPs) lacks clear justification. Summing functional scores across different ranges may introduce bias; a weighted or standardized approach would be more rigorous. Please clarify the prioritization criteria.

5. The functional validation uses the maximum non-toxic dose of BDE-47. How does this dose compare to real-world exposure levels in the cohort? Are the observed effects plausible at environmentally relevant concentrations? Address potential limitations in extrapolating experimental findings to population exposures.

6. The enrichment of histone marks suggests epigenetic regulation, yet the study focuses solely on TF-mediated gene expression. Could other mechanisms (e.g., chromatin remodeling, splicing modulation) contribute? Please expand the

discussion on alternative mechanisms.

Minor comments:

1. SLC6A20 is implicated in glycine transport and cholesterol degradation via bile acid pathways. Were bile acid-related metabolites altered by PBDE exposure? Include relevant data or discuss their potential role.
2. HepG2 cells were used for validation—are these findings generalizable to primary hepatocytes or in vivo systems? Briefly justify the model choice and discuss translational limitations.
3. The term "dyslipidemia" is not defined, and the analysis uses quantitative lipid markers. How do these findings translate to clinical dyslipidemia risk? Clarify the connection.
4. Why total cholesterol was chosen over other lipid traits for experimental validation? Were the loci for different lipids compared? Discuss their interrelationships.
5. Provide basic genomic details of rs9869609 (e.g., distance to SLC6A20, intronic/intergenic status, MAF) to contextualize the variant's potential functional impact.
6. The PBDE-specific PRS lacks internal/external validation. Discuss its utility for future research (e.g., risk prediction, personalized interventions).
7. The experiments used BDE-47 as exposure, while analyses modeled total PBDEs. Explain this discrepancy and its implications for interpretation.
8. Other typographical and formatting issues:
 - In Fig. 2c, specify units for "Beta of blood lipids" in PRS models.
 - In Fig. 6, correct "gpro" to "g pro." in metabolite/lipid labels and describe normalization methods.
 - In Fig. 4b, indicate the significance threshold for pathway enrichment.

Reviewer #2

(Remarks to the Author)

This study investigates the interaction between genetic variants and polybrominated diphenyl ether (PBDE) exposure in modulating lipid homeostasis, combining epidemiological data from the China National Human Biomonitoring (CNHBM) cohort with multi-omics profiling and functional validation. The authors performed gene-environment (G×E) interaction analysis, identifying SNPs associated with PBDE-induced dyslipidemia, with rs9869609 (near SLC6A20) emerging as a top candidate (among others). Metabolomic analysis highlighted glycine and glycerophosphate as key mediators, while CRISPR/Cas9-edited HepG2 cells demonstrated that the rs9869609-G allele suppresses SLC6A20 expression via enhanced BHLHE40 binding, impairing glycine transport and exacerbating cholesterol accumulation under BDE-47 exposure.

The article is highly compelling due to its translational approach, bridging population-level findings with mechanistic insights. The identification of SLC6A20 as a susceptibility marker for PBDE-induced dyslipidemia and its validation through knockout experiments in HepG2 cells strengthens the biological plausibility of the results. The study is well-structured, clearly written, and addresses an important gap in understanding how environmental pollutants interact with genetic background to influence metabolic health.

While the findings are robust, some limitations should be noted. The functional validation was conducted only in vitro, and future studies using in vivo models (e.g., genetically modified mice or humanized animal systems) would further solidify the conclusions. Additionally, the generalizability of the findings need to confirm in other cohorts, and the clinical relevance of the polygenic risk score (PRS) warrants further investigation.

Despite these limitations, the study represents a significant advance in environmental health research, providing novel insights into gene-environment interactions in dyslipidemia. The work is of broad interest to the fields of toxicology, metabolism, and precision medicine. With minor revisions to address, this article is suitable for publication.

General remarks

1. Clarification of focus on SLC6A20

1.1) The authors focus on SLC6A20 due to its strong statistical association with TC (rs1392296, FDR=0.0078) and compelling functional evidence linking its suppression to glycine-mediated cholesterol metabolism. However, other candidate genes (e.g., CTCF, enriched across all lipid traits) are not explored experimentally. While prioritization is inevitable, a brief discussion justifying why SLC6A20 was selected over equally plausible candidates—such as CTCF, a global epigenetic regulator including of lipid-related genes, or CETP, directly involved in lipoprotein remodeling—would strengthen the manuscript. Future studies could address whether these genes act independently or synergistically with SLC6A20 in PBDE-induced dyslipidemia.

1.2) The polygenic risk score incorporates multiple loci, yet only SLC6A20 is functionally tested. A brief explanation of why other top genes were excluded (e.g., lack of eQTL support, lower effect sizes) would bolster transparency.

1.3) While the role of SLC6A20 in glycine transport is compelling, dyslipidemia is multifactorial. The study does not rule out additive effects of other candidate genes. Please discuss

2. Association between PBDE, SLC6A20 and dyslipidemia

2.1) The study lacks RNA-seq or single-cell transcriptomics to validate SLC6A20 expression changes in human tissues or additional PBDE-responsive genes.

2.2) The other top genes identified in genetic variants could had been investigated in HeparG cells.

2.3) The claim that rs9869609-G is the "top causal variant" is premature without: In vivo validation (e.g., knock-in mice). Please temper this in the article.

Evidence that this variant explains a significant proportion of dyslipidemia risk in the population. The PRS improves prediction but has modest effect sizes ($\beta=1.18$ for TC).

Here is some piece of advice that the authors could take into consideration and discuss : Clarify Exposure Relevance: It could be interesting to compare in vitro BDE-47 doses to human urinary levels and assess dose-response relationships.

Address Population Bias: Replicate findings in diverse cohorts or perform trans-ethnic meta-analyses.

Strengthen Mechanistic Claims:

Provide direct evidence of bile acid conjugation defects in SLC6A20-deficient models.

Integrate transcriptomics to identify downstream effects of PBDE-exposed genotypes.

Minor points

Figure 2 Color Consistency Issue:

In panel (a), colors distinguish chromosomes, but in panels (b) (stratified risk) and (c) (PRS trends), the same/similar colors represent PBDE exposure groups (low/high). This is confusing.

Use distinct color schemes (e.g., grayscale for chromosomes, blue/orange for exposure groups) or add explicit labels to legends.

Figure 4 (Metabolome Profiling)

There is a legend Mismatch in Panel (b3) (HDL-C):

The legend describes two circle sizes (likely representing effect size or significance), but three sizes appear in the plot (small/medium/large).

Please align the legend with the visual output

On panel c line widths effectively represent mediation strength, but key metabolites (e.g., glycine, citrate) could be highlighted with bold labels or arrows to emphasize central pathways.

Typos and suggestions

Line 164: unit instead of units

Line 168: Suggest rather than demonstrate

Line 1026 Characterized

Line 1085: H3K4me1 repeated twice

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have adequately addressed my concerns, and I have no further comments.

Reviewer #2

(Remarks to the Author)

The authors did an exceptionally good job in revising the manuscript. They have not only addressed all the points that I (and R1) raised, but have done it in a manner that has significantly strengthened the manuscript that is now more robust and impactful. I have no more reservations and recommend the manuscript for acceptance in its present form.

The new data and analyses the authors incorporated are impressive. Regarding the prioritization of SLC6A20, they provided a detailed clarification of the multi-tiered prioritization strategy. The explanation, based on the strongest G×E interaction signal (rs1392296 locus), followed by fine-mapping to identify rs9869609 as the variant with the highest integrated functional potential, is now completely clear and convincing. The inclusion of functional annotations and biological discussions for other top candidate genes (CETP, CPNE4, CAMK1D) in the results and discussion sections provides a more balanced view. This effectively addresses my initial concern.

I appreciated the careful revision of the language throughout the manuscript to refer to rs9869609 as a "candidate causal variant".

Regarding the relevance of in vitro doses, the addition of the benchmark dose (BMD) analysis for both the cellular model and the human cohort is a major strength of the revised manuscript. The authors now show that experimental concentration is multiple of the cellular BMD, and that a comparable multiple of the human BMD falls within a relevant exposure range (the third quartile) of the population. This bridges the gap between the experimental system and human environmental exposure. This greatly bolsters the translational significance of findings made by the authors.

The new LC-MS data quantifying bile acid levels in the knockdown model is an insightful piece of evidence. The results elegantly close the mechanistic loop, linking SLC6A20 knockdown to impaired glycine transport to reduced glycine-conjugated bile acids to compromised cholesterol excretion. This is a major improvement to the study's mechanistic narrative.

The new RNA-seq data from BDE-47 exposed HepG2 cells is highly valuable. It provides a systems-level view, identifying a network of dysregulated genes (CETP, CAMK1D, etc.) and enriched pathways. This successfully addresses my comment and highlights the polygenic nature of the response to PBDE exposure.

Authors completely meet the expectations for a revision and I congratulate them for this excellent work and have no further reservation.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Responses to Reviewers

Reviewer #1 (Comments to the Author):

Hu N. et al. conducted an interesting study on individual heterogeneity in response to environmental exposure, with a specific focus on PBDE-induced dyslipidemia. By integrating exposome, genomics, and metabolomic data from the China National Human Biomonitoring cohort, the authors identified 3,571 genetic variants that interact with PBDE exposure to influence dyslipidemia risk with key metabolites serving as mediators. Among these, rs9869609 was pinpointed as the top causal variant, exhibiting the strongest association with hypercholesterolemia risk. Further functional validation through single-base CRISPR/Cas9 editing revealed that the rs9869609-G allele downregulates SLC6A20 expression by enhancing BHLHE40 binding. This, in turn, impaired glycine transport and promoted cholesterol accumulation, particularly under BDE-47 exposure.

This study provides a comprehensive genome-wide analysis of gene-environment interactions in metabolic diseases. It makes a significant contribution by identifying vulnerable populations and clarifying the mechanistic pathways underlying environmentally induced metabolic disorders. Additionally, it opens new avenues for research on the health effects of environmental pollutants. However, several key limitations need to be addressed to improve the quality and impact of the manuscript.

Response: Thanks for the positive comments. We have addressed the concerns in the following responses.

Major comments:

1. The authors constructed a polygenic risk score (PRS) using independent variants to predict lipid phenotypes, finding increased risk only in populations with high PBDE exposure. Why does the PRS not confer risk in low-exposure groups? This observation warrants further discussion

Response: Thanks for your comments. The PRS was associated with increased dyslipidemia risk exclusively in the high-PBDE exposure group, with no comparable effect under low exposure. This demonstrates a clear G×E interaction, where genetic susceptibility becomes phenotypically significant only above a certain exposure threshold. Elevated PBDE levels appear to amplify the functional impact of risk variants, whereas low-level exposure remains insufficient to disrupt lipid metabolism. These findings underscore those genetic effects are not uniform but depend on specific environmental exposure. We have elaborated the discussions on these mechanisms in the revised manuscript.

In the Discussion section, page 14, line 453-460, we now state:

“Notably, this risk was negligible in low-exposure groups, suggesting a threshold effect wherein low PBDE levels are insufficient to disrupt metabolic pathways regulated by the interactive variants in the PRS models. In contrast, high PBDE exposure amplifies these disruptions through a synergistic interaction with these variants, significantly increasing the risk of lipid metabolic disorders. These findings demonstrate that genetic susceptibility to dyslipidemia is dynamically modulated by PBDE exposure.”

2. While the study identifies loci interacting with PBDE exposure, it does not clearly differentiate this approach from traditional GWAS for disease-associated loci. Please elaborate on why gene-

environment interaction (G×E) loci may offer greater biological or clinical relevance than main-effect GWAS hits, particularly for environmental metabolic disorders.

Response: Thanks for your insightful comment regarding the distinction between our G×E interaction approach and traditional GWAS. G×E interaction loci offer greater biological and clinical relevance than main-effect GWAS loci. Specifically, they reveal how PBDEs and other environmental exposures modulate genetic effects, elucidate specific pathogenic pathways that are elusive to GWAS, identify genetically susceptible subgroups for targeted exposure-reduction interventions, and link genetic susceptibility to modifiable environmental factors to provide actionable evidence for precision prevention.

To further demonstrate this advantage, we performed an overlap analysis between G×E interaction genes and known GWAS loci in the GWAS Catalog (<https://www.ebi.ac.uk/gwas/>; EFO_0004530; EFO_0004612; EFO_0004611; EFO_0004574). Among 1,037 genes associated with PBDE-lipid interactions phenotypes, 210 overlapped with known GWAS loci, while the remainder were novel. These findings highlight exposure-specific biological pathways that traditional GWAS may overlook. We have added this discussion in the revised manuscript.

In the Discussion section, page 13, line 444-449, we now state:

“G×E interaction analysis overcomes these shortcomings by revealing conditional genetic effects tied to specific environments and capturing more phenotypic variance, which our study also confirms, of the 1037 PBDE-interactive genes, only 210 overlapped with known GWAS genes, showing traditional GWAS overlook exposure-specific pathways.”

3. In metabolomic analysis, the authors focus on overlapping metabolites for pathway enrichment and mediation analysis, potentially missing broader mechanistic insights. A supplementary correlation analysis of all significant metabolites could reveal additional candidate molecules and pathways.

Response: Thanks for your constructive suggestions. Recognizing the limitation of focusing solely on overlapping metabolites, we expanded our analysis to include all significant metabolites through KEGG pathway and mediation analyses. This approach revealed additional lipid homeostasis pathways, including arachidonic acid metabolism and unsaturated fatty acid biosynthesis, and identified 14 novel mediating metabolites. These results have been added to the revised manuscript.

In the Result section, page 9, line 270-276, we added:

“We also performed pathway enrichment analysis on all differential metabolites. The results not only validated the primary pathways identified through overlapping metabolite analysis but also revealed novel pathways significantly enriched across all differential metabolites, including arachidonic acid metabolism and biosynthesis of unsaturated fatty acids (**Fig. S2c**). This further broadens our understanding of the G×E interaction metabolic network.”

In the Result section, page 10, line 297-302, we added:

“Furthermore, we expanded the scope of our mediator analysis to include all G×E-associated differential metabolites as potential mediators. This supplementary analysis identified 14 mediating metabolites (**Fig. S2f**). These metabolites primarily involve previously understudied pathways of

inflammation and amino acid metabolism disorders, providing more comprehensive mechanistic insights into G×E interactions affecting lipid metabolism.”

In the Discussion section, page 16, line 559-563, we added:

“Additionally, we highlighted the involvement of supplementary pathways such as arachidonic acid metabolism and biosynthesis of unsaturated fatty acid, thereby providing a more comprehensive perspective on the metabolic dysregulation induced by PBDEs.”

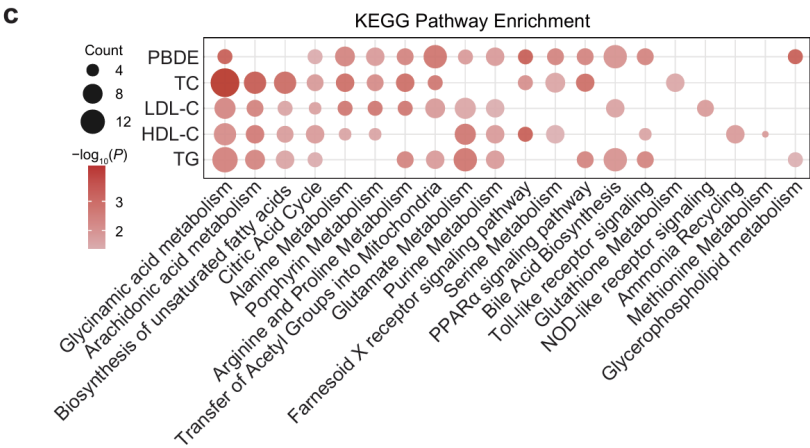


Fig. S2c, KEGG pathway enrichment analysis for PBDE, TC, LDL-C, HDL-C, and TG; bubble size corresponds to the number of metabolites in the pathway. Dot color and size represent P-values and gene counts, respectively.

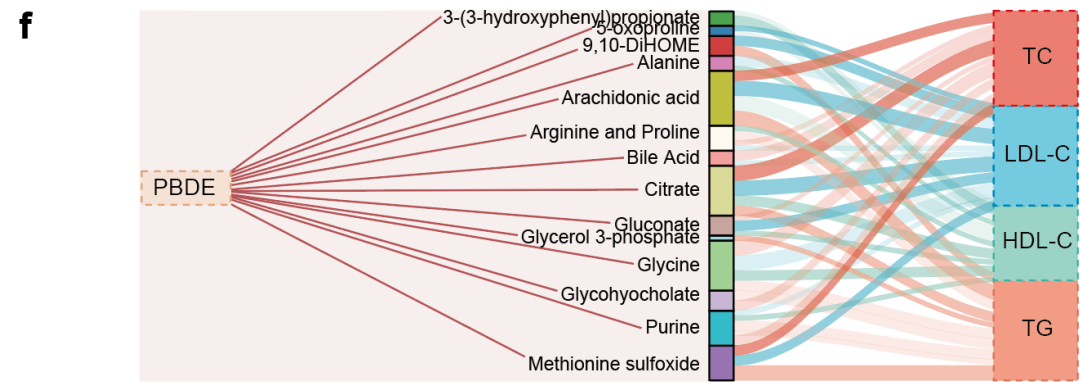


Fig. S2f, Sankey diagram of the mediating effect of PBDE on lipid characteristics, with lines representing mediating effects.

4. The choice of rs9869609 (over the most statistically significant SNPs) lacks clear justification. Summing functional scores across different ranges may introduce bias; a weighted or standardized approach would be more rigorous. Please clarify the prioritization criteria.

Response: Thank you for your insightful comments. The prioritization of rs9869609 was based on three criteria: (1) its significant association with hypercholesterolemia in G×E analyses, showing optimal FDR values; (2) functional annotations indicating the highest regulatory potential across three databases; and (3) its location in a regulatory region of *SLC6A20* with significant eQTL effects. We have added corresponding descriptions in the **Methods** and **Results** sections of the manuscript.

In the Method section, page 21, line 746-766, we now state:

“We next performed a functional annotation for the genetic variants using multiple bioinformatic tools and databases, including the HaploReg database (<https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php>), Roadmap Epigenomics Consortium (<https://www.ncbi.nlm.nih.gov/geo/roadmap/epigenomics/>), RegulomeDB database (<https://www.regulomedb.org/regulome-search>), CADD database (<https://cadd.gs.washington.edu/score>), 3DSNP (<https://omic.tech/3dsnpv2/>), which integrated multiple histone modification chromatin immunoprecipitation (ChIP)-seq peaks, transcription factor (TF) ChIP-seq peaks. To be specific, (1) regulatory chromatin histone ChIP-Seq were characterized by using the data from the Roadmap Epigenomics Consortium; (2) prioritizing these variants with the scores of regulatory elements by using RegulomeDB, CADD, and 3DSNP prediction databases⁶⁵. CADD first filters out potentially harmless variants, focusing on functionally significant sites⁶⁶. RegulomeDB then determines whether these sites reside within regulatory regions and assesses their regulatory potential⁶⁷. Finally, 3DSNP connects variants within regulatory regions to their actual distal target genes, completing the full biological narrative from variant to target gene⁶⁸. Specifically, we apply minimum-maximum normalization to the raw outputs of these databases, uniformly converting their scores to the 0–1 range.”

In the Result section, page 10, line 313-317, we now state:

“We calculated an integrated functional prediction score for these SNPs by combining outputs from CADD, RegulomeDB, and 3DSNP, which identified rs9869609 ($r^2 = 0.991$ with the tag SNP) as the candidate causal variant due to its greatest potential for functional regulation (**Fig. 5b, Table S6**).”

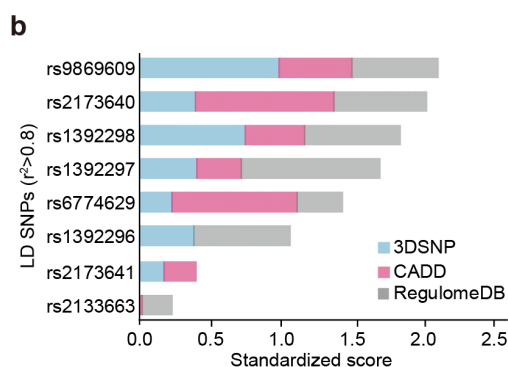


Fig. 5b, Bar Plot of SNP Annotation Scores based on different genomic databases (Min-max standardization of RegulomeDB, 3DSNP, and CADD scores).

5. The functional validation uses the maximum non-toxic dose of BDE-47. How does this dose compare to real-world exposure levels in the cohort? Are the observed effects plausible at environmentally relevant concentrations? Address potential limitations in extrapolating experimental findings to population exposures.

Response: Thank you for the excellent suggestion. To address this, we established dose-response relationships of BDE-47-induced lipid accumulation in both cellular models and human populations (**Fig. R2A**). The benchmark dose (BMD), a widely-used indicator to quantify the minimum concentration for chemical toxicity^{R1}, was 4.12 μM in HepG2 cells. Our *in vitro* concentration of

12.5 μM (approximately three times the cellular BMD) based on its consistently induced phenotypic alterations and alignment with established toxicological paradigms for acute BDE-47 exposure^{R2}.

In the human population (CNHBM cohort), the BMD for the association between BDE-47 exposure and dyslipidemia was 0.0711 $\mu\text{g/L}$ (**Fig. R2B**). Notably, three times the population BMD (0.21 $\mu\text{g/L}$) corresponds to the third quartile (Q3) of human BDE-47 exposure (**Fig. R2C**). This consistent dose-response pattern across experimental and human systems supports the translational relevance of our findings.

However, we acknowledge the limitations in extrapolating from acute *in vitro* exposure to chronic low-dose human exposure, particularly regarding toxicokinetic differences, and have addressed this in the **Discussion** section.

In the Discussion section, page 17, line 575–577, we now state:

“Third, the extrapolation from acute *in vitro* exposure to chronic low-dose human exposure remains challenging due to differences in toxicokinetic and the complex interplay of multiple organ systems *in vivo*.”

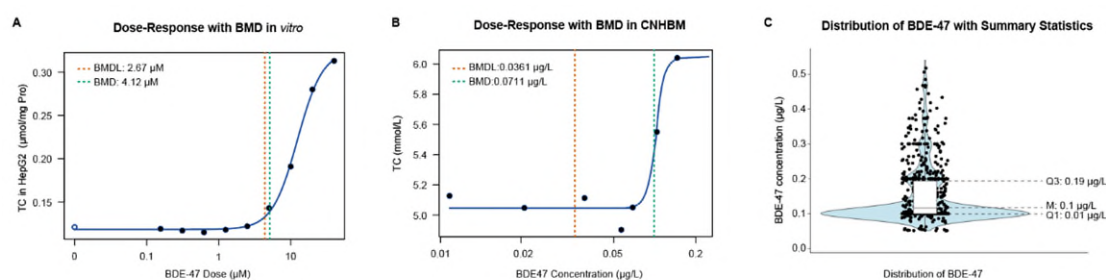


Fig. R1 Dose-response curves for BDE-47 concentration versus the total cholesterol (TC) *in vitro* (**A**) and in CNHBM (**B**), with BDE-47 distribution in CNHBM (**C**). Note: BMDL: Lower benchmark dose limit; BMD: Benchmark dose; Q1: First quartile; Q3: Third Quartile; M: Median.

6. The enrichment of histone marks suggests epigenetic regulation, yet the study focuses solely on TF-mediated gene expression. Could other mechanisms (e.g., chromatin remodeling, splicing modulation) contribute? Please expand the discussion on alternative mechanisms.

Response: Thank you for your insightful suggestions regarding alternative regulatory mechanisms beyond TF-mediated gene expression. We agree that the observed molecular features indicate a broader landscape of gene regulation. In response to your comment, we have expanded the discussions to include additional regulatory mechanisms.

In the Discussion section, page 14, line 466–497, we now state:

“The intricate interplay between genetic susceptibility and environmental exposures collectively constructs a sophisticated regulatory network that governs metabolic pathways¹¹. Our investigation revealed that the majority of interactive variants are located in non-coding regions enriched with regulatory elements such as promoters and enhancers. This indicates that these variants likely function by interfering with TF binding at promoter or enhancer sites, thereby modulating transcriptional activity. ... Additionally, we observed that these interactive variants are significantly

enriched in epigenetic features. Emerging evidence suggests that environmental pollutants or their metabolites may act as substrate mimics or inhibitors that directly bind to histone-modifying enzymes, thereby altering catalytic activity and leading changes in histone modification patterns^{11,17,42}. Such epigenetic alterations can subsequently influence downstream RNA processing events. For instance, histone modifications such as H3K36me3 can recruit splicing factors and directly influence the elongation rate of RNA polymerase II, thereby modulating alternative splicing events⁴³. These potential mechanisms provide clear directions for our future research.”

Minor comments:

1. SLC6A20 is implicated in glycine transport and cholesterol degradation via bile acid pathways. Were bile acid-related metabolites altered by PBDE exposure? Include relevant data or discuss their potential role.

Response: Thank you for this suggestion. To strengthen the mechanistic evidence for bile acid conjugation defects in *SLC6A20* deficiency, we performed LC-MS analysis of bile acid metabolites in *SLC6A20* knockdown HepG2 cells. The results demonstrated significantly reduced levels of glycocholic acid (GCA) and glycochenodeoxycholic acid (GDCA) in *SLC6A20* knockdown HepG2 cells, with a more pronounced decrease under BDE-47 exposure. These findings indicate that both *SLC6A20* deficiency and BDE-47 exposure impair glycine transport, thereby reducing the synthesis of glycine-conjugated bile acids that are critical for cholesterol excretion. We have added a corresponding description in the **Result** section of the manuscript.

In the Result section, page 12, line 408-414, we now state:

“To evaluate the impact of *SLC6A20* on glycine-conjugated bile acids, which are essential for cholesterol excretion, we measured GCA and GDCA levels in *SLC6A20* knockdown HepG2 cells. LC-MS analysis revealed significantly reduced levels of GCA and GDCA concentrations in siSLC6A20 cells compared to siControl cells. Following BDE-47 exposure, GCA and GDCA levels exhibited a further decrease in siSLC6A20 cells, while siControl cells showed a more moderate reduction (**Fig. 6hi**).”

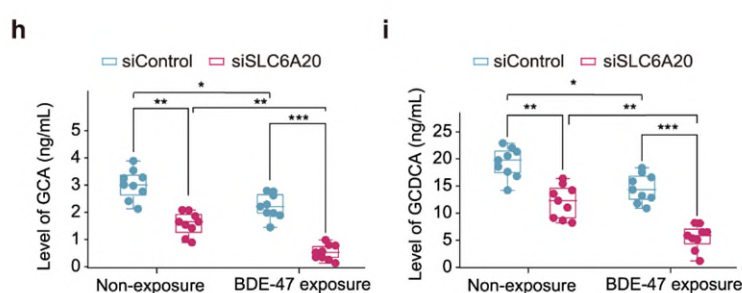


Fig.6hi, GCA levels and GDCA levels in HepG2 cells with *SLC6A20* knockdown, evaluated under control conditions and BDE-47 exposure.

2. HepG2 cells were used for validation—are these findings generalizable to primary hepatocytes or *in vivo* systems? Briefly justify the model choice and discuss translational limitations.

Response: We appreciate this important note. HepG2 cells were selected for initial validation due to their retained hepatocyte functions, though we acknowledge their limitations in metabolic

competency compared to primary hepatocytes. Our findings warrant further validation in more physiologically relevant models.

In the Discussion section, page 17, line 571-575, we now state:

“Second, though used as a typical model for lipid metabolism, HepG2 cells exhibit reduced metabolic competency, altered signaling pathways, and lack the full phenotypic complexity of primary human hepatocytes or in vivo systems, which alert for future validations in other cell lines and in vivo models.”

3. The term "dyslipidemia" is not defined, and the analysis uses quantitative lipid markers. How do these findings translate to clinical dyslipidemia risk? Clarify the connection.

Response: Thank you for this insightful comment. To more precisely analyze the quantitative association between G×E interactions and lipid phenotypes, our models utilize quantitative lipid indicators. The term "dyslipidemia" is used in this study as a broad clinical concept referring to abnormal lipid levels, including elevated total cholesterol, elevated LDL-C, low HDL-C, or high triglycerides.

To illustrate the association between these quantitative indices and disease dyslipidemia, we analyzed the effect of variants rs98696609 on disease risk and found significant G×E interactions with PBDE exposure on TC (FDR = 0.0082) and LDL-C (FDR = 0.0104). Under high exposure, GG genotype carriers showed substantially increased hypercholesterolemia prevalence (26.1%) compared to AA carriers (16.8%). We have added a corresponding description in the **Result** section of the manuscript.

In the Result section, page 10, line 328-334, we added:

We further confirmed the population-level effect of this variant, in which significant interaction effects were observed for TC (FDR = 0.0082) and LDL-C (FDR = 0.0104) levels. Especially, under high PBDE exposure, individuals carrying the GG genotype exhibited substantially altered lipid profiles compared to those with AA genotype, with hypercholesterolemia prevalence increasing from 16.8% to 26.1%, while the change was neglectable under low exposure (**Fig.S3ab**).

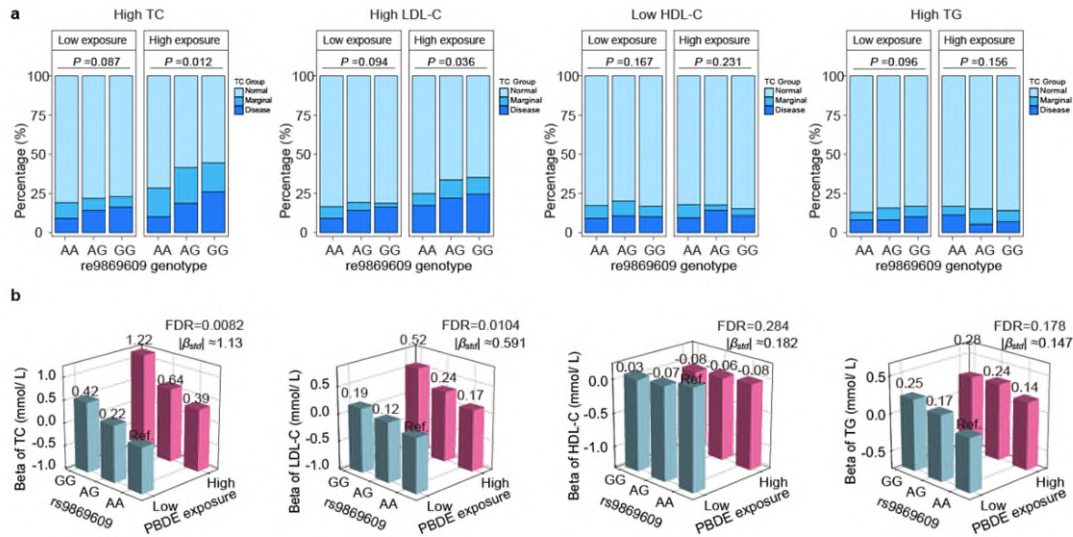


Fig. S3. PBDE exposure, genetic variation, and their impacts on lipid metabolism

a, Cumulative distributions of lipid phenotypes (TC, LDL-C, HDL-C, TG) by rs9869609 genotype and exposure level. **b**, Stratified risk analysis of the rs9869609 in lipid phenotypes (TC, LDL-C, HDL-C, TG). Abbreviations: $|\beta_{std}|$, absolute standardized β .

4. Why total cholesterol was chosen over other lipid traits for experimental validation? Were the loci for different lipids compared? Discuss their interrelationships.

Response: Thank you for your thoughtful comments. Our priority of TC for experimental validation is supported by phenotypic associations, G×E interaction analyses and biological relevance of TC. Phenotypically, TC showed the strongest positive associations with both Σ PBDEs and BDE-47 exposure (**Fig. S1d**). Genetically, TC exhibited the broadest spectrum of G×E interaction variants among all lipid traits (**Fig.2a**). Through systematic screening of the top 20 loci prioritized by statistical significance and interaction effect size (**Table S5**), the rs1392296 locus emerged as the most compelling candidate, exhibiting the lowest FDR (0.0078) and highest $|\beta_{std}|$ (0.6517). This robust statistical evidence, combined with the biological relevance of its annotated gene *SLC6A20* in cholesterol metabolism, supported our focus on TC for subsequent mechanistic investigation. These results have been incorporated into the revised manuscript.

In the Result section, page 10, line 305-313, we now state:

“To investigate the molecular mechanisms by which PBDE-interacting variants disrupt lipid metabolism, we first screened 20 top loci (the top five for each lipid phenotype) from **Fig.2a** and prioritized those with the most significant effects based on two core indicators: FDR and absolute standardized β (interaction β per SD of outcome) (**Table S5**). As a result, the locus with tag SNP rs1392296 outperformed others, exhibiting the lowest FDR (0.0078) and highest $|\beta_{std}|$ (0.6517). The latter indicator suggests a relatively high impact of PBDE-SNP interaction on TC.”

Regarding whether genetic variants for different lipids should be compared, we analyzed the overlap of G×E interaction variants across the four lipid phenotypes. Results showed TC and LDL-C shared 193 interactive variants (**Fig. R2A**) and 80 overlapping annotated genes, mostly concentrated in

cholesterol metabolism pathways (**Fig. R2B**), while other lipid phenotypes had few common variants. This confirms that validating TC's mechanism not only clarifies its own response to PBDEs but also provides a shared mechanistic reference for understanding LDL-C and other lipid dysregulation, balancing the study's focus and efficiency. We have added a corresponding discussion in the revised manuscript.

In the Discussion section, page 16, line 534-537, we now state:

“Notably, the interactive effect between rs9869609 and PBDEs on LDL-C, together with other shared variants influencing both TC and LDL-C, providing a mechanistic basis for the coordinated dysregulation of LDL-C and TC.”

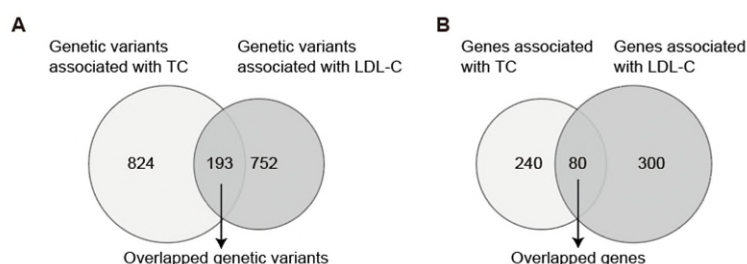


Fig. R2 Venn diagram of genetic variants (**A**) and genes (**B**) associated with TC and LDL-C

5. Provide basic genomic details of rs9869609 (e.g., distance to *SLC6A20*, intronic/intergenic status, MAF) to contextualize the variant's potential functional impact.

Response: Thank you for this suggestion. As detailed in **Table S6**, rs9869609 is an intergenic variant located approximately 347 kb downstream of *SLC6A20*. With a MAF of 6% and high integrated functional scores from CADD, RegulomeDB, and 3DSNP, this variant was prioritized as a candidate causal variant potentially functioning through distal cis-regulatory mechanisms.

In the Result section, page 10, line 313-328, we now state:

“We calculated an integrated functional prediction score for these SNPs by combining outputs from CADD, RegulomeDB, and 3DSNP, which identified rs9869609 ($r^2 = 0.991$ with the tag SNP, MAF = 6%) as the candidate causal variant due to its greatest potential for functional regulation (**Fig. 5b**, **Table S6**). ... This variant is located in an intergenic region, approximately 347 kb downstream of the *SLC6A20* gene, suggesting its distal cis-enhancer role to regulate the gene expression.”

6. The PBDE-specific PRS lacks internal/external validation. Discuss its utility for future research (e.g., risk prediction, personalized interventions).

Response: Thank you for this suggestion. We acknowledge the challenges in accessing databases that combine genome-wide genotype data with PBDE exposure measurements, which limited our ability to validate findings in an independent cohort. This limitation has been explicitly noted in the Discussion section. As suggested, we have also discussed the potential utility of the PRS for future research directions.

In the Discussion section, page 14, line 460-464, we now state:

“The PRS developed in this study validates polygenic G×E interactions, showing significant association with total cholesterol and confirming a concerted genetic mechanism. Although not yet clinically applicable, this PRS serves as a foundational tool for identifying high-risk individuals who may benefit from targeted exposure mitigation.”

In the Discussion section, page 17, line 577-581, we now state:

“Lastly, the findings’ generalizability may be constrained by the specific demographic and environmental context of the Chinese cohort studied. Validation across more diverse populations is essential to establish the broader applicability of these G×E interactions.”

7. The experiments used BDE-47 as exposure, while analyses modeled total PBDEs. Explain this discrepancy and its implications for interpretation.

Response: Thank you for this comment. BDE-47 was selected for experimental validation based on two considerations: it showed the highest detection rate in our cohort and correlated strongly with total PBDEs levels, representing the primary exposure effect; and as a highly bioactive congener, it enables clearer mechanistic investigation in experimental systems^{R3}. The use of BDE-47 in experiments provides a clear and representative entry point for mechanism studies, while employing total PBDEs in population analyses better reflects the cumulative effects of actual exposure.

In the Method section, page 22, line 780-787, we now state:

“In the population effects analysis, to better reflect real-world exposure conditions, we used total PBDEs as a quantitative indicator of mixture exposure to assess its impact on blood lipids. Analysis results indicate that the primary effects of BDE-47 in the population align with those of the total PBDEs mixture. In the experimental section, to streamline procedures and ensure representativeness of exposure levels, effect intensity, and mechanisms, we selected BDE-47 as the representative compound for this class of pollutants to infer potential mechanisms of action for similar pollutants⁷⁵.”

8. Other typographical and formatting Issues:

In Fig. 2c, specify units for “Beta of blood lipids” in PRS models.

In Fig. 6, correct "gpro" to "g pro." in metabolite/lipid labels and describe normalization methods.

In Fig. 4b, indicate the significance threshold for pathway enrichment.

Response: We appreciate this suggestion. We have carefully addressed all the typographical and formatting issues raised, including the addition of units in **Fig. 2c** and correction of metabolite labels in **Fig. 6**. Metabolite levels were normalized to total protein (mmol/g pro) using BCA assay. The significance threshold in **Fig. 4b** is specified as $P_{adj} < 0.05$. All modifications have been incorporated into the revised manuscript.

Reviewer #2 (Remarks to the Author):

This study investigates the interaction between genetic variants and polybrominated diphenyl ether (PBDE) exposure in modulating lipid homeostasis, combining epidemiological data from the China National Human Biomonitoring (CNHBM) cohort with multi-omics profiling and functional validation. The authors performed gene-environment (G×E) interaction analysis, identifying SNPs associated with PBDE-induced dyslipidemia, with rs9869609 (near SLC6A20) emerging as a top candidate (among others). Metabolomic analysis highlighted glycine and glycerophosphate as key mediators, while CRISPR/Cas9-edited HepG2 cells demonstrated that the rs9869609-G allele suppresses SLC6A20 expression via enhanced BHLHE40 binding, impairing glycine transport and exacerbating cholesterol accumulation under BDE-47 exposure.

The article is highly compelling due to its translational approach, bridging population-level findings with mechanistic insights. The identification of SLC6A20 as a susceptibility marker for PBDE-induced dyslipidemia and its validation through knockout experiments in HepG2 cells strengthens the biological plausibility of the results. The study is well-structured, clearly written, and addresses an important gap in understanding how environmental pollutants interact with genetic background to influence metabolic health.

While the findings are robust, some limitations should be noted. The functional validation was conducted only in vitro, and future studies using in vivo models (e.g., genetically modified mice or humanized animal systems) would further solidify the conclusions. Additionally, the generalizability of the findings needs to be confirmed in other cohorts, and the clinical relevance of the polygenic risk score (PRS) warrants further investigation.

Despite these limitations, the study represents a significant advance in environmental health research, providing novel insights into gene-environment interactions in dyslipidemia. The work is of broad interest to the fields of toxicology, metabolism, and precision medicine. With minor revisions to address, this article is suitable for publication.

Response: Thanks for the positive comments. We have addressed the concerns in the following responses.

General remarks

1. Clarification of focus on SLC6A20

1.1) The authors focus on SLC6A20 due to its strong statistical association with TC (rs1392296, FDR=0.0078) and compelling functional evidence linking its suppression to glycine-mediated cholesterol metabolism. However, other candidate genes (e.g., CTCF, enriched across all lipid traits) are not explored experimentally. While prioritization is inevitable, a brief discussion justifying why SLC6A20 was selected over equally plausible candidates—such as CTCF, a global epigenetic regulator including of lipid-related genes, or CETP, directly involved in lipoprotein remodeling—would strengthen the manuscript. Future studies could address whether these genes act independently or synergistically with SLC6A20 in PBDE-induced dyslipidemia.

Response: We appreciate the insightful comments and have supplemented additional information in response. First, our locus prioritization was based on G×E interaction significance (FDR) and effect size (**Table S5**). The rs1392296 locus, annotated to *SLC6A20*, exhibited the strongest association with TC (FDR = 0.0078, $|\beta_{\text{std}}| = 0.6517$). Subsequent fine-mapping revealed rs9869609 as the top functional candidate based on CADD, RegulomeDB, and 3DSNP annotations among LD SNPs (**Table S6**). Additionally, we have supplemented functional annotations for other top loci in

Table S5. Taking the TG-associated rs2431621 locus as an example, this variant was annotated to *CAMK1D*, with GTEx eQTL data supporting its regulatory role (**Fig. R1**). *CAMK1D* encodes a kinase involved in metabolic regulation^{R4}. These additions have been incorporated into the results and discussion sections.

In the Result section, page 17, line 305-322, we now state:

“To investigate the molecular mechanisms by which PBDE-interacting variants disrupt lipid metabolism, we first screened 20 top loci (the top five for each lipid phenotype) from **Fig. 2a** and prioritized those with the most significant effects based on two core indicators: FDR and absolute standardized β (interaction β per SD of outcome) (**Table S5**). As a result, the locus with tag SNP rs1392296 outperformed others, exhibiting the lowest FDR (0.0078) and highest $|\beta_{\text{std}}|$ (0.6517). The latter indicator suggests a relatively high impact of PBDE-SNP interaction on TC. Within this locus, seven another significant SNPs displayed strong LD ($r^2 \geq 0.8$) with the tag SNP (**Fig. 5a**). We calculated an integrated functional prediction score for these SNPs by combining outputs from CADD, RegulomeDB, and 3DSNP, which identified rs9869609 ($r^2 = 0.991$ with the tag SNP, MAF = 6%) as the candidate causal SNP due to its greatest potential for functional regulation (**Fig. 5b**). Genomic visualization further confirmed that rs9869609 is significantly enriched in DNase hypersensitive sites and active histone modification peaks (H3K4me1, H3K4me3, H3K9ac, and H3K27ac), underscoring its robust regulatory potential (**Fig. 5c**). Therefore, we prioritized functional validation of this variant to explore how SNPs interact with PBDE to drive lipid dysregulation.”

In the Result section, page 6, line 165-173, we now state:

“Similarly, PBDE exposure interacted with 955 SNPs (177 loci), 769 SNPs (140 loci), and 739 SNPs (127 loci) to influence LDL-C, HDL-C, and TG levels, respectively. The top SNPs from these analyses were annotated to biologically relevant genes: *CETP* (FDR = 0.0089), which encodes cholesteryl ester transfer protein critical for lipoprotein remodeling; *CPNE4* (FDR = 0.0104), a calcium-dependent phospholipid-binding protein potentially involved in lipid metabolism; and *CAMK1D* (FDR = 0.0092), a calcium -dependent protein kinase implicated in hormonal regulation of metabolic processes.”

In the Discussion section, page 16, line 537-544, we now state:

“Beyond *SLC6A20*, other genes contribute to PBDE-induced dyslipidemia, underscoring the polygenic nature of this process *CETP* mediates cholesterol ester transfer between lipoproteins, *CTCF* acts as an epigenetic regulator of lipid-related genes⁵⁵, while *CPNE4* and *CAMK1D* may participate in intracellular lipid signaling and metabolic regulation. These genes represent promising targets for future investigation into the complex interplay between genetic susceptibility and environmental triggers in dyslipidemia.”

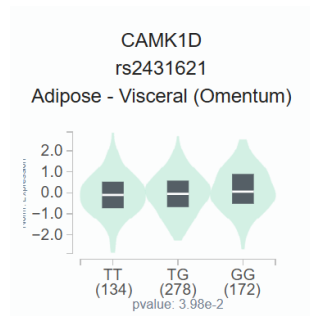


Fig. R1 eQTL analyses for rs2431621 genotypes and *CAMK1D* expression in the GTEx database

1.2) The polygenic risk score incorporates multiple loci, yet only *SLC6A20* is functionally tested. A brief explanation of why other top genes was excluded (e.g., lack of eQTL support, lower effect sizes) would bolster transparency.

Response: We appreciate your insightful comment. The PRS was constructed using established methods and incorporated all genome-wide significant variants from our G×E interaction analysis. Among these, *SLC6A20* was prioritized for experimental validation based on statistical and functional evidence: it exhibited both the lowest FDR (0.0078) and the largest interaction effect size ($|\beta_{\text{std}}| = 0.6517$) for TC, and the regulatory variant rs9869609 within this locus demonstrated the highest integrated functional score and strong eQTL.

Due to practical experimental constraints, we focused our validation efforts on the most promising candidate. To ensure transparency, we have provided complete functional annotations for all candidate genes in **Tables S5** and **S6**, including functional prediction scores, and literature-based biological functions for each gene. We have also expanded discussion of other key genes such as *CETP*, *CPNE4* and *CAMK1D* in the manuscript.

In the Discussion section, page 16, line 537-544, we now state:

“Beyond *SLC6A20*, other genes contribute to PBDE-induced dyslipidemia, underscoring the polygenic nature of this process *CETP* mediates cholesterol ester transfer between lipoproteins⁵⁴, *CTCF* acts as an epigenetic regulator of lipid-related genes⁵⁵, while *CPNE4* and *CAMK1D* may participate in intracellular lipid signaling and metabolic regulation. These genes represent promising targets for future investigation into the complex interplay between genetic susceptibility and environmental triggers in dyslipidemia.”

1.3) While the role of *SLC6A20* in glycine transport is compelling, dyslipidemia is multifactorial. The study does not rule out additive effects of other candidate genes. Please discuss

Response: Thanks for your insightful comment. We acknowledge the polygenic nature of dyslipidemia, which motivated the development of a PRS in this study. Our transcriptomic analysis in HepG2 cells exposed to BDE-47 identified 1,221 differentially expressed genes, including key lipid regulators such as *CETP* and *CAMK1D* alongside *SLC6A20*. These results indicate that PBDE-induced dyslipidemia likely results from coordinated disruption across multiple genetic pathways, rather than the effect of a single gene.

We agree that investigating additive or interactive effects among these genes would be valuable.

However, due to technical and resource limitations, we focused validation on the most statistically compelling candidate. We have acknowledged this limitation and plan to pursue functional studies of additional genes in future work.

In the Discussion section, page 17, line 568-571, we now state:

“First, our mechanistic experiments focused solely on the rs9869609 variant’s role in PBDE-induced hypercholesterolemia, future studies on other key genes, especially their independently or synergistically act, are needed for comprehensive validation.”

2. Association between PBDE, SLC6A20 and dyslipidemia

2.1) The study lacks RNA-seq or single-cell transcriptomics to validate SLC6A20 expression changes in human tissues or additional PBDE-responsive genes.

Response: Thank you for your insightful suggestion. In response, we performed RNA-seq in HepG2 cells exposed to BDE-47 (12.5 μ M, 24 h). The analysis identified 1221 significantly differentially expressed genes, including *SLC6A20*, *CETP* and *CAMK1D*. KEGG pathway analysis revealed enrichment in cholesterol metabolism and lipoprotein biosynthesis pathways, indicating PBDE exposure disrupts lipid homeostasis through these metabolic pathways. Specifically, *SLC6A20* showed significant downregulation ($\log_2FC = -2.21$, FDR = 0.0032), supporting its involvement in response to PBDE exposure. These results have been included in the revised manuscript.

In the Result section, page 11, line 364-372, we now state:

“To investigate the downstream responsive genes of BDE-47 exposure, we performed RNA-seq in HepG2 cells following BDE-47 exposure (12.5 μ M, 24 h). As a result (**Fig S4**), we identified 1221 significantly differentially expressed genes (DEG), including *SLC6A20*, *CETP* and *CAMK1D* ($|\log_2(FC)| > 1$, FDR < 0.05). KEGG analysis revealed that the DEG were primarily enriched in cholesterol metabolism and lipoprotein biosynthesis pathways, indicating that PBDE exposure disrupts lipid homeostasis through these key metabolic pathways. Notably, *SLC6A20* was among the downregulated genes ($\log_2(FC) = -2.21$, FDR = 0.0032), pointing to its role in response to PBDE exposure.”

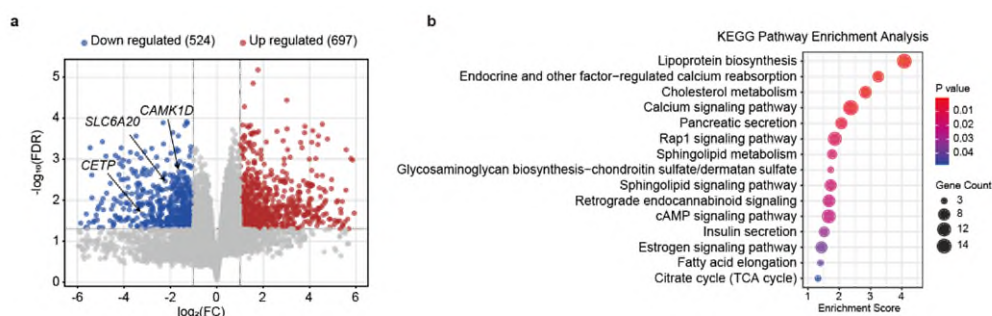


Fig. S4. Transcriptome Analysis following BDE-47 exposure in HepG2 cells.

a, Volcano plot of $\log_2(FC)$ for Differentially expressed genes (DEG) versus false discovery rate ($-\log_{10}(FDR)$). **b,** KEGG pathway enrichment analysis for DEG; Dot color and size represent P-values and gene counts, respectively.

2.2) The other top genes identified in genetic variants could had been investigated in HepaRG cells.

Response: Thank you for your suggestion. In response, we have supplemented **Tables S5** and **S6** with comprehensive functional annotations, including prediction scores, and literature-supported biological roles for all candidate genes. We have also expanded the discussion of *CETP*, *CPNE4*, and *CAMK1D* in the manuscript. However, due to technical and resource limitations, we focused experimental validation on the most statistically compelling candidate, *SLC6A20*. We have explicitly acknowledged this constraint in the revised manuscript.

In the Discussion section, page 16, line 537-544, we now state:

“Beyond *SLC6A20*, other genes contribute to PBDE-induced dyslipidemia, underscoring the polygenic nature of this process *CETP* mediates cholesterol ester transfer between lipoproteins⁵⁴, *CTCF* acts as an epigenetic regulator of lipid-related genes⁵⁵, while *CPNE4* and *CAMK1D* may participate in intracellular lipid signaling and metabolic regulation. These genes represent promising targets for future investigation into the complex interplay between genetic susceptibility and environmental triggers in dyslipidemia.”

In the Discussion section, page 17, line 568-571, we now state:

“First, our mechanistic experiments focused solely on the rs9869609 variant’s role in PBDE-induced hypercholesterolemia, future studies on other key genes, especially their independently or synergistically act, are needed for comprehensive validation.”

2.3) The claim that rs9869609-G is the "top causal variant" is premature without: In vivo validation (e.g., knock-in mice). Please temper this in the article.

Evidence that this variant explains a significant proportion of dyslipidemia risk in the population. The PRS improves prediction but has modest effect sizes ($\beta=1.18$ for TC).

Response: Thank you for your critical comment. We agree that claiming causality requires *in vivo* validation and have revised the terminology to "candidate causal variant" throughout the manuscript.

Regarding the PRS effect size, the observed effect ($\beta = 0.62$ mmol/L, standardized $\beta \approx 0.59$) meets Cohen's criterion^{R5} for a large effect (standardized $\beta > 0.5$), confirming the substantial impact of G×E interactions on lipid levels. It should be noted that the single-variant effect ($\beta = 1.18$ mmol/L) and PRS represent distinct aspects of genetic susceptibility. The former reflects a strong interaction under specific environmental conditions, while the latter captures cumulative risk across multiple loci. Both approaches provide complementary insights into the genetic architecture of PBDE-induced dyslipidemia.

3. Here is some piece of advice that the authors could take into consideration and discuss: Clarify Exposure Relevance: It could be interesting to compare *in vitro* BDE-47 doses to human urinary levels and assess dose-response relationships.

Response: Thank you for the excellent suggestion. To address this, we established dose-response relationships of BDE-47-induced lipid accumulation in both cellular models and human populations (**Fig. R2A**). The benchmark dose (BMD), a widely-used indicator to quantify the minimum concentration for chemical toxicity^{R1}, was 4.12 μ M in HepG2 cells. Our *in vitro* concentration of 12.5 μ M (approximately three times the cellular BMD) based on its consistently induced phenotypic alterations and alignment with established toxicological paradigms for acute BDE-47 exposure^{R2}.

In the human population (CNHBM cohort), the BMD for the association between BDE-47 exposure and dyslipidemia was 0.0711 $\mu\text{g/L}$ (**Fig. R2B**). Notably, three times the population BMD (0.21 $\mu\text{g/L}$) corresponds to the third quartile (Q3) of human BDE-47 exposure (**Fig. R2C**). This consistent dose-response pattern across experimental and human systems supports the translational relevance of our findings.

However, we acknowledge the limitations in extrapolating from acute *in vitro* exposure to chronic low-dose human exposure, particularly regarding toxicokinetic differences, and have addressed this in the **Discussion** section.

In the Discussion section, page 17, line 575-577, we now state:

“Third, the extrapolation from acute *in vitro* exposure to chronic low-dose human exposure remains challenging due to differences in toxicokinetic and the complex interplay of multiple organ systems *in vivo*.”

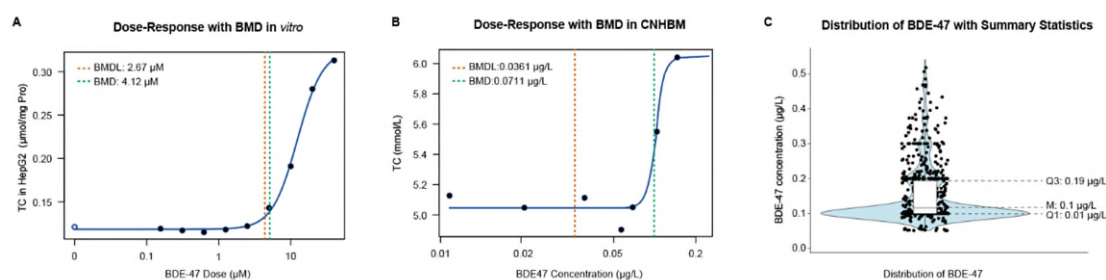


Fig. R2 Dose-response curves for BDE-47 concentration versus the total cholesterol (TC) *in vitro* (A) and in CNHBM (B), with BDE-47 distribution in CNHBM (C). Note: BMDL: Lower benchmark dose limit; BMD: Benchmark dose; Q1: First quartile; Q3: Third Quartile; M: Median.

4. Address Population Bias: Replicate findings in diverse cohorts or perform trans-ethnic meta-analyses.

Response: We appreciate the suggestion regarding trans-ethnic validation. To clarify, the CNHBM cohort provides substantial diversity, encompassing participants from 31 provinces across China and representing multiple ethnic groups, as detailed in the Methods section. This geographical and ethnic heterogeneity strengthens the reliability of our findings within the Chinese population.

We acknowledge that external validation in non-East Asian cohorts would further reinforce the trans-ethnic relevance of our results. However, the limited availability of large-scale cohorts with integrated genotype data and detailed PBDE exposure measurements has constrained our ability to perform trans-ethnic meta-analyses. We have explicitly recognized this as a limitation in the revised manuscript.

In the Discussion section, page 18, line 607-611, we now state:

“Notably, the CNHBM cohort comprises a multinational population, including Han (18,794, 86.4%), Tibet (551, 2.5%), Zhuang (392, 1.8%), Mongolia (340, 1.6%), Hui (279, 1.3%), Man (249, 1.1%), Li (180, 0.8%), Uyghur (138, 0.6%), Dong (138, 0.6%), and other ethnicities (685, 3.2%), reflecting

its broad ethnic and geographic representation.”

In the Discussion section, page 17, line 577-581, we now state:

“Lastly, the findings’ generalizability may be constrained by the specific demographic and environmental context of the Chinese cohort studied. Validation across more diverse populations is essential to establish the broader applicability of these G×E interactions.”

5. Strengthen Mechanistic Claims:

Provide direct evidence of bile acid conjugation defects in *SLC6A20*-deficient models.

Response: Thank you for this suggestion. To strengthen the mechanistic evidence for bile acid conjugation defects in *SLC6A20* deficiency, we performed LC-MS analysis of bile acid metabolites in *SLC6A20* knockdown HepG2 cells. The results demonstrated significantly reduced levels of glycocholic acid (GCA) and glycochenodeoxycholic acid (GDCA) in *SLC6A20* knockdown HepG2 cells, with a more pronounced decrease under BDE-47 exposure. These findings indicate that both *SLC6A20* deficiency and BDE-47 exposure impair glycine transport, thereby reducing the synthesis of glycine-conjugated bile acids that are critical for cholesterol excretion. We have added a corresponding description in the **Result** section of the manuscript.

In the Result section, page 12, line 408-414, we now state:

“To evaluate the impact of *SLC6A20* on glycine-conjugated bile acids, which are essential for cholesterol excretion, we measured GCA and GDCA levels in *SLC6A20* knockdown HepG2 cells. LC-MS analysis revealed significantly reduced levels of GCA and GDCA concentrations in siSLC6A20 cells compared to siControl cells. Following BDE-47 exposure, GCA and GDCA levels exhibited a further decrease in siSLC6A20 cells, while siControl cells showed a more moderate reduction (**Fig. 6hi**).”

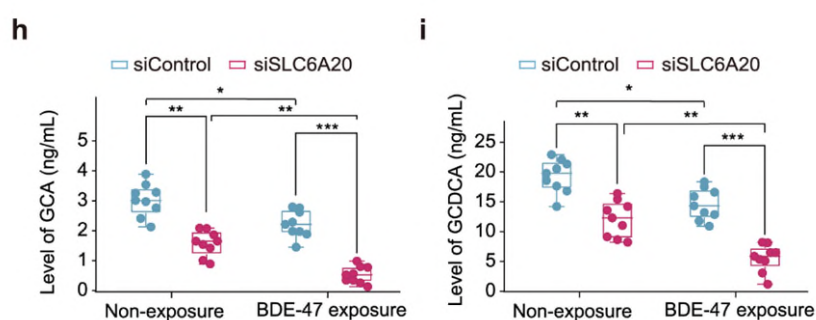


Fig.6. Boxplots showing GCA (**h**) and GDCA (**i**) levels in *SLC6A20* knockdown and control HepG2 cells under non-exposure or BDE-47 exposure.

6. Integrate transcriptomics to identify downstream effects of PBDE-exposed genotypes.

Response: Thank you for your insightful suggestion. In response, we performed RNA-seq in HepG2 cells exposed to BDE-47 (12.5 μ M, 24 h). The analysis identified 1221 significantly differentially expressed genes, including *SLC6A20*, *CETP* and *CAMK1D*. KEGG pathway analysis revealed enrichment in cholesterol metabolism and lipoprotein biosynthesis pathways, indicating PBDE exposure disrupts lipid homeostasis through these metabolic pathways. Specifically, *SLC6A20*

showed significant downregulation ($\log_2FC = -2.21$, $FDR = 0.0032$), supporting its involvement in response to PBDE exposure.

In the Result section, page 11, line 364-372, we now state:

“To investigate the downstream responsive genes of BDE-47 exposure, we performed RNA-seq in HepG2 cells following BDE-47 exposure (12.5 μM , 24 h). As a result (**Fig S4**), we identified 1221 significantly differentially expressed genes (DEG), including *SLC6A20*, *CETP* and *CAMK1D* ($|\log_2(FC)| > 1$, $FDR < 0.05$). KEGG analysis revealed that the DEG were primarily enriched in cholesterol metabolism and lipoprotein biosynthesis pathways, indicating that PBDE exposure disrupts lipid homeostasis through these key metabolic pathways. Notably, *SLC6A20* was among the downregulated genes ($\log_2(FC) = -2.21$, $FDR = 0.0032$), pointing to its role in response to PBDE exposure.”

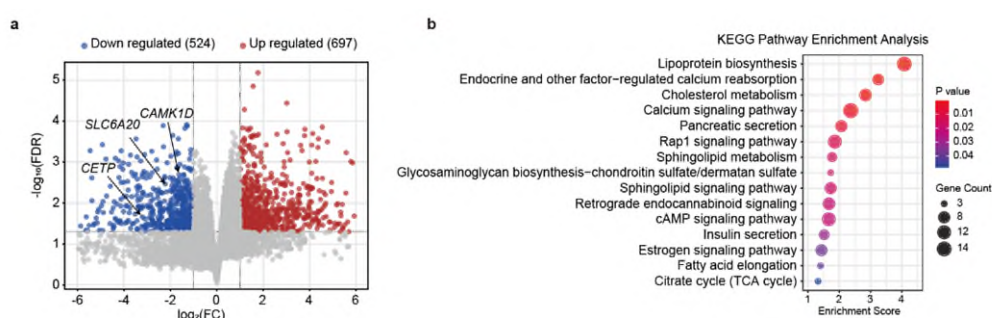


Fig. S4. Transcriptome Analysis following BDE-47 exposure in HepG2 cells.

a, Volcano plot of $\log_2(FC)$ for Differentially expressed genes (DEG) versus false discovery rate ($-\log_{10}(FDR)$). **b**, KEGG pathway enrichment analysis for DEG; Dot color and size represent P-values and gene counts, respectively.

Minor points

1. Figure 2 Color Consistency Issue:

In panel (a), colors distinguish chromosomes, but in panels (b) (stratified risk) and (c) (PRS trends), the same/similar colors represent PBDE exposure groups (low/high). This is confusing. Use distinct color schemes (e.g., grayscale for chromosomes, blue/orange for exposure groups) or add explicit labels to legends.

Response: Thanks for your suggestion. We have revised **Fig. 2a** by using grayscale to represent chromosomes.

2. Figure 4 (Metabolome Profiling)

There is a legend mismatch in Panel (b3) (HDL-C):

The legend describes two circle sizes (likely representing effect size or significance), but three sizes appear in the plot (small/medium/large). Please align the legend with the visual output.

On panel c line widths effectively represent mediation strength, but key metabolites (e.g., glycine, citrate) could be highlighted with bold labels or arrows to emphasize central pathways.

Response: Thanks for your reminding. We have standardized the legends by categorizing circle sizes into three distinct levels, corresponding to pathway enrichment ratios (low: 5; medium: 10;

high: 15). We have added bold labels to key metabolites, including glycine, citrate, and glycerol 3-phosphate, enhancing their visibility. These revisions have been incorporated into the updated **Fig. 4** in the revised manuscript.

3. Typos and suggestions

Line 164: unit instead of units

Line 168: Suggest rather than demonstrate

Line 1026 Characterized

Line 1085: H3K4me1 repeated twice

Response: Thanks for your reminding. We have corrected the inappropriate description in the revised manuscript.

References

- R1. Jin, Y., et al. High throughput data-based, toxicity pathway-oriented development of a quantitative adverse outcome pathway network linking AHR activation to lung damages. *Journal of hazardous materials* 425, 128041 (2022).
- R2. Liu, X., et al. The role of lysosomes in BDE 47-mediated activation of mitochondrial apoptotic pathway in HepG2 cells. *Chemosphere* 124, 10-21 (2015).
- R3. Zhu, L., Ma, B. & Hites, R.A. Brominated flame retardants in serum from the general population in northern China. *Environmental science & technology* 43, 6963-6968 (2009).
- R4. Vivot, K., et al. CaMK1D signalling in AgRP neurons promotes ghrelin-mediated food intake. *Nat Metab* 5, 1045-1058 (2023).
- R5. Zou, G.Y. Exact confidence interval for Cohen's effect size is readily available. *Statistics in medicine* 26, 3054-3056 (2007).

Point-by-point responses to reviewer comments

Reviewer #1 (Comments to the Author):

The authors have adequately addressed my concerns, and I have no further comments.

Response to Reviewer #1:

We thank the reviewer for the positive feedback.

Reviewer #2 (Remarks to the Author):

The authors did an exceptionally work in revising the manuscript. They have not only addressed all the points that I (and R1) raised, but have done it in a manner that has significantly strengthened the manuscript that is now more robust and impactful. I have no more reserves and recommend the manuscript for acceptance in its present form.

The new data and analyses the authors incorporated are impressive. Regarding the prioritization of SLC6A20, they provided a detailed clarification of the multi-tiered prioritization strategy. The explanation, based on the strongest G×E interaction signal (rs1392296 locus), followed by fine-mapping to identify rs9869609 as the variant with the highest integrated functional potential, is now completely clear and convincing. The inclusion of functional annotations and biological discussions for other top candidate genes (CETP, CPNE4, CAMK1D) in the results and discussion sections provides a more balanced view. This effectively addresses my initial concern.

I appreciated the careful revision of the language throughout the manuscript to refer to rs9869609 as a "candidate causal variant".

Regarding the relevance of in vitro doses, the addition of the benchmark dose (BMD) analysis for both the cellular model and the human cohort is a major strength of the revised manuscript. The authors now show that experimental concentration is multiple of the cellular BMD, and that a comparable multiple of the human BMD falls within a relevant exposure range (the third quartile) of the population. This bridges the gap between the experimental system and human environmental exposure. This greatly bolsters the translational significance of findings made by the authors.

The new LC-MS data quantifying bile acid levels in the knockdown model is an insightful piece of evidence. The results elegantly close the mechanistic loop, linking SLC6A20 knockdown to impaired glycine transport to reduced glycine-conjugated bile acids to compromised cholesterol excretion. This is a major improvement to the study's mechanistic narrative.

The new RNA-seq data from BDE-47 exposed HepG2 cells is highly valuable. It provides a systems-level view, identifying a network of dysregulated genes (CETP, CAMK1D, etc.) and enriched pathways. This successfully addresses my comment and highlights the polygenic nature of the response to PBDE exposure.

Authors completely meet the expectations for a revision and I congratulate them for this excellent work and have no further reservation.

Response to Reviewer #2:

We sincerely appreciate your detailed and constructive assessment. We are glad that the additional analyses (BMD, LC-MS, RNA-seq) and clarifications have strengthened the manuscript's mechanistic narrative and translational relevance to your satisfaction.