Nature Communications

https://doi.org/10.1038/s41467-026-70222-8

Article in Press

Genetic variations interact with polybrominated diphenyl ether exposure to alter lipid homeostasis

Received: 16 May 2025

Accepted: 23 February 2026

Cite this article as: Hu, N., Li, B., Lu, Y. *et al.* Genetic variations interact with polybrominated diphenyl ether exposure to alter lipid homeostasis. *Nat Commun* (2026). https://doi.org/10.1038/s41467-026-70222-8

Naifan Hu, Bin Li, Yifu Lu, Qi Jiang, Ying Zhu, Zheng Li, Yingli Qu, Tian Qiu, Donghui Zhang, Zhuo Wang, Yunfei Ma, Huibin Jin, Peijie Sun, Haocan Song, Yunhao Zhao, Yifan Zhao, Ming Zhang, Feng Zhao, Saisai Ji, Bifeng Yuan, Ying Zhu, Yuebin Lv, Jianbo Tian, Xiaoping Miao & Xiaoming Shi

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

Naifan Hu[1,2,3#], Bin Li[1#], Yifu Lu[2#], Qi Jiang[1#]; Ying Zhu[1#], Zheng Li[2#], Yingli Qu[2], Tian Qiu[2], Donghui Zhang[1], Zhuo Wang[1], Yunfei Ma[1], Huibin Jin[1], Peijie Sun[2], Haocan Song[2], Yunhao Zhao[1], Yifan Zhao[1], Ming Zhang[1], Feng Zhao[2], Saisai Ji[2], Bifeng Yuan[4], Ying Zhu[2*], Yuebin Lv[2,5*], Jianbo Tian[1,3*], Xiaoping Miao[1,3,6,7*], Xiaoming Shi[2,5*]

[1] Department of Epidemiology and Biostatistics, School of Public Health, State Key Laboratory of Metabolism and Regulation in Complex Organisms, TaiKang Center for Life and Medical Sciences, Wuhan University, Wuhan, 430071, China.
[2] China CDC Key Laboratory of Environment and Population Health, National Institute of Environmental Health, Chinese Center for Disease Control and Prevention, Beijing, 100021, China.
[3] Hubei Provincial Center for Disease Control and Prevention & NHC Specialty Laboratory of Food Safety Risk Assessment and Standard Development, Wuhan, 430079, China.
[4] Department of Occupational and Environmental Health, School of Public Health, Wuhan University, Wuhan, 430071, China.
[5] National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases (NITFID), National Institute of Environmental Health, Chinese Center for Disease Control and Prevention, Beijing, 100021, China.
[6] Department of Gastrointestinal Oncology, Zhongnan Hospital of Wuhan University, Wuhan University, Wuhan, 430071, China.
[7] Jiangsu Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing, 210000, China.

[#] These authors Naifan Hu, Bin Li, Yifu Lu, Qi Jiang, Ying Zhu[1], and Zheng Li contributed equally to this work.

[*]**Correspondence to**
Dr. Xiaoming Shi, China CDC Key Laboratory of Environment and Population Health, National Institute of Environmental Health, Chinese Center for Disease Control and Prevention, #7 Panjiayuan Nanli, Chaoyang, Beijing 100021, China; Email: shixm@chinacdc.cn
Dr. Xiaoping Miao, Department of Epidemiology and Biostatistics, School of Public Health, State Key Laboratory of Metabolism and Regulation in Complex Organisms,

TaiKang Center for Life and Medical Sciences, Wuhan University, Wuhan, 430071, China; Email: xpmiao@whu.edu.cn
Dr. Jianbo Tian, Department of Epidemiology and Biostatistics, School of Public Health, State Key Laboratory of Metabolism and Regulation in Complex Organisms, TaiKang Center for Life and Medical Sciences, Wuhan University, Wuhan, 430071, China; E-mail: tianjb@whu.edu.cn
Dr. Yuebin Lv, China CDC Key Laboratory of Environment and Population Health, National Institute of Environmental Health, Chinese Center for Disease Control and Prevention, #7 Panjiayuan Nanli, Chaoyang, Beijing 100021, China; E-mail: lvyuebin@nieh.chinacdc.cn
Dr. Ying Zhu, China CDC Key Laboratory of Environment and Population Health, National Institute of Environmental Health, Chinese Center for Disease Control and Prevention, #7 Panjiayuan Nanli, Chaoyang, Beijing 100021, China; Email: zhuying@nieh.chinacdc.cn

Abstract

Polybrominated diphenyl ethers (PBDEs) are implicated in dyslipidemia, but the molecular basis of individual susceptibility remains elusive. Here we report an analysis based on the China National Human Biomonitoring cohort, where we integrate exposome, genomic, and metabolomic data to identify 3,571 genetic variants that interact with PBDE exposure to influence dyslipidemia risk. Metabolomic analysis highlights glycine and glycerophosphate as key mediators. A polygenic risk score derived from these PBDE-interactive variants significantly enhances dyslipidemia prediction in highly exposed individuals. Among these, rs9869609 emerges as a candidate causal variant, showing the strongest association with hypercholesterolemia risk (β=1.18, FDR=0.0078). Further functional validation using single-base CRISPR/Cas9 editing reveals that the rs9869609-G allele downregulates *SLC6A20* expression by strengthening BHLHE40 binding, which further impairs glycine transport and promotes cholesterol accumulation, particularly under 2,2′,4,4′-Tetrabromodiphenyl ether exposure. Collectively, our study elucidates a gene-environment interaction mechanism through which genetic variants modulate lipid metabolism in response to PBDE exposure.

Introduction

Metabolic disorders such as obesity, type 2 diabetes (T2D), and dyslipidemia, account for approximately 60% of global mortality, posing a significant public health challenge[1]. Although genetic susceptibility plays a fundamental role, environmental exposures, particularly widespread pollutants like polybrominated diphenyl ethers (PBDEs), are also considered important contributors to metabolic dysregulation[2]. PBDEs are widely used as brominated flame retardants and have been implicated in metabolic toxicity due to their potential for bioaccumulation and capacity to disrupt lipid metabolism[3]. However, individual responses to PBDEs vary considerably, which cannot be explained solely by exposure levels[4]. This heterogeneity, along with inconsistent epidemiological findings linking PBDEs to metabolic diseases[5-8], underscores the pressing need to identify the determinants of individual susceptibility.

While genome-wide association studies (GWAS) have identified numerous genetic variants associated with metabolic traits, these findings explain less than 30% of disease heritability[9,10]. This points to the potential role of gene-environment (G × E) interactions, where genetic background modulates responses to environmental exposures[11,12]. Initial G × E interaction research has focused on established disease-related loci, such as the obesity-related gene *FTO*, and their interactions with lifestyle factors like diet and physical activity[13,14]. Emerging evidence also indicates that environmental pollutants, including dioxins and other endocrine disruptors, may mimic endogenous ligands and directly alter proteins encoded by risk variants, thereby disrupting metabolic pathways[15,16]. Recent genome-wide G × E interaction studies in neurological phenotypes have further highlighted how environmental factors influence epigenetic regulation of gene expression[17]. However, systematic investigations of G × E interactions involving metabolic traits, particularly in relation to PBDEs, remain limited, impeding a holistic grasp of the dynamics and underlying mechanisms of G × E interactions in metabolic health.

To address this gap, high-throughput metabolomics offers a powerful approach in mechanism insight, as perturbed metabolite profiles can reveal pathway-specific disruptions and serve as early dysfunction biomarkers[18]. Notably, alterations in amino acid metabolism, such as elevated branched-chain amino acids (BCAAs) and reduced glycine have been implicated in pollutant-induced metabolic dysregulation[19,20]. BCAAs may exacerbate insulin resistance through mTOR pathway activation, while glycine deficiency can impair glutathione synthesis and cholesterol metabolism, exacerbating oxidative stress and dyslipidemia[21-24]. Nevertheless, the specific metabolic pathways

mediating the interaction between PBDEs and specific genetic variants leading to lipid dysregulation remain unclear.

Against this backdrop, our study is driven by the hypothesis that PBDE exposure interacts with genetic variants to disrupt lipid homeostasis via metabolic pathway alterations. Using the China National Human Biomonitoring (CNHBM) cohort, which integrates exposome, multi-omics, and phenomics profiles in China[25], we aimed to systematically identify genetic variants that modulate PBDE-induced dyslipidemia, delineate the metabolic pathways mediating these gene-PBDE interactions, and provide mechanistic insights of individual susceptibility using single-base-edited cellular models (Fig. 1). By bridging epidemiological findings with experimental validation, this study seeks to clarify G × E mechanisms in metabolic disorders, uncover biomarkers for risk stratification, and inform precision intervention strategies to mitigate pollutant-related metabolic risks.

Results

Participant enrollment and basic characteristics

The 2017-2018 CNHBM baseline survey included 21,888 participants aged 3-79 years, with 10% randomly selected for multi-omics profiling (Supplementary Fig. 1a). From this sub-cohort, we restricted our analysis to adults with available lipid profiles and excluded individuals with missing data on PBDE exposure, genomic, metabolomic, or key covariate measurements, resulting in a final analytic dataset comprised 871 individuals (Supplementary Fig. 1b). Concentrations of 14 PBDE congeners were measured, with only 2,2′,4,4′-Tetrabromodiphenyl ether (BDE-47) and 2,2',4,4',5,6'-Hexabromodiphenyl ether (BDE-154) detected in > 50% of samples (Supplementary Fig. 1c). We calculated the sum of all 14 PBDEs (ΣPBDE) to represent overall exposure and stratified participants into high- and low-exposure groups based on the median ΣPBDE concentration (0.816 μg/L). The two exposure groups showed comparable distributions of age, sex, body mass index (BMI), and lifestyle factors (smoking, alcohol consumption). However, the high-exposure group demonstrated significantly increased levels of total cholesterol (TC; P = 0.015) and triglycerides (TG; P = 0.002) and decreased levels of high-density lipoprotein cholesterol (HDL-C; P = 0.004), along with marginally increased low-density lipoprotein cholesterol (LDL-C; P = 0.139), as depicted in Table 1. Association analyses revealed statistically significant relationships between ΣPBDE and the measured lipids except for LDL-C, with similar patterns observed for BDE-47 (Supplementary Fig. 1d).

Interactions of genetic variants and PBDE exposure on lipid profiles

To investigate genetic variants modulating lipid profiles in response to PBDE exposure, we performed genome-wide G × E interaction analysis using adjusted general linear regression models. Manhattan plots depicted the genomic distribution of significant PBDE-interacting single nucleotide polymorphisms (SNPs) at a false discovery rate (FDR) < 0.05 (Fig. 2a). We identified 1,108 SNPs across 178 loci exhibiting significant interactions on TC levels, with rs1392296, annotated to *SLC6A20* (Solute Carrier Family 6 Member 20), showing the strongest signal (FDR = 0.0078). *SLC6A20* encodes a membrane transporter in the SLC6 family, facilitating molecular transport across cellular boundaries. Compared to individuals with the AA genotype at rs1392296 under low PBDE exposure, those with the GG genotype displayed TC increases of 0.45 unit at low exposure and 1.18 unit at high exposure (Fig. 2b), corresponding to a standardized effect size of 1.13-fold standardized deviation (SD) of TC. 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[26]; *CPNE4* (FDR = 0.0104), a calcium-dependent phospholipid-binding protein potentially involved in lipid metabolism[27]; and *CAMK1D* (FDR = 0.0092), a calcium -dependent protein kinase implicated in hormonal regulation of metabolic processes[28]. These results suggest that PBDE exposure alters lipid profiles through specific genetic interactions, highlighting the critical role of G × E interactions in lipid metabolism regulation.

Building on these findings, we constructed lipid-specific polygenic risk scores (PRS) incorporating independent PBDE-interacting SNPs to evaluate their predictive utility for dyslipidemia risk. In populations with high PBDE exposure, the PRS exhibited significant positive associations with elevated TC, LDL-C, and TG (P_{trend} = 0.001, 0.026, and 0.003, respectively) and a negative association with HDL-C (P_{trend} = 0.001); whereas these associations were absent under low exposure (Fig. 2c). These data suggest that higher PRS values amplify dyslipidemia susceptibility in the context of elevated PBDE exposure, affirming the role of genetic predisposition in modulating pollutant-induced metabolic dysregulation.

Functional characterization of PBDE-interacting SNPs

To elucidate the gene-level mechanisms by which genetic variants affects lipids under PBDE exposure, we characterized significantly interactive SNPs basing on their genomic locations and molecular functions. For comparative analysis, a control set of non-

interactive SNPs was generated, matched to the interactive SNPs for linkage disequilibrium, minor allele frequency (MAF), and variant type. Genomic location annotation revealed that approximately 90% of the interactive SNPs resided in non-coding regions, predominantly intronic and intergenic regions (Fig. 3ab). These SNPs exhibited significant enrichment in histone marks linked to epigenetic regulatory elements, including enhancers (H3K27ac, H3K4me1), and other active marks (DNase I, H3K9ac, H3K4me2), while showing depletion in repressive marks (H3K27me3) (Fig. 3c). Analysis of transcription factor (TF) binding sites further indicated significant enrichment of motifs for key TFs, such as BHLHE40, CTCF, FOXA1, and CEBPB (Fig. 3d). These results suggest that interactive SNPs predominantly modulate lipid phenotypes by regulating epigenetic elements and gene transcription.

We subsequently employed an integrative framework, combining nearest-gene, expression quantitative trait loci (eQTLs), and MAGMA annotations, to identify potential target genes regulated by these SNPs (Fig. 3e). For the four lipid phenotypes, we identified 1,241, 1,261, 957, and 921 interacting genes, respectively (Supplementary Data 1-4). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis delineated distinct metabolic roles: TC-related genes were enriched in cholesterol biosynthesis, LDL-C genes in lipid metabolism regulation via APOB-mediated lipoprotein assembly, HDL-C genes in reverse cholesterol transport, and TG genes in fatty acid *β*-oxidation (Fig. 3f). Gene Ontology (GO) term enrichment further clarified the functional dimensions of these pathways (Fig. 3g). In biological processes, steroid metabolism and lipid metabolic regulation are significantly enriched. In molecular functions, prominent LDL particle binding and cholesterol transport activity provide a functional context for *CPNE4* and *CAMK1D* in intracellular lipid signaling and metabolism. In cellular components, significant enrichment of lipoprotein particles and cholesterol synthesis complexes supports the structural basis for these processes. Together, these results indicate that PBDE-interacting SNPs exert regulatory influence over gene expression within key metabolic pathways, providing critical mechanistic insights into their contributions to lipid metabolism under PBDE exposure.

Metabolic pathways in dyslipidemia by Gene × PBDE interactions

Metabolomics offers a powerful approach to investigate how metabolic processes in organisms respond to varying environmental and genetic conditions by thoroughly examining shifts in small-molecule metabolites[29]. In order to identify the metabolic pathways and key metabolites involved in the interplay of PBDE exposure and

susceptible variants, we first screened the differential metabolites associated with an interaction index. The index was derived by multiplying PBDE exposure levels by the PRS constructed above for each lipid phenotype. Our analysis revealed 96, 89, 66, and 39 metabolites associated with the G × E interactions for the four lipid traits, respectively, with approximately one- to two- thirds of these metabolites also directly linked to the corresponding lipid phenotypes (Fig. 4a). Interestingly, the number of metabolites implicated to these interactions exceeded those associated solely with PBDE exposure, indicating a broader capture of responding molecules of PBDE exposure (Supplementary Fig. 2ab).

Among the 63 overlapping metabolites identified for TC, citrate, glycine, and bile acid were particularly prominent, pointing to disruptions in energy metabolism and the bile acid cycle that may impair cholesterol synthesis and catabolism. These metabolites also exerted significant effects across the other three lipid types, underscoring their central involvement in G × E-driven lipid dysregulation. For LDL-C, metabolites like arginine/proline, behenoyl dihydrosphingomyelin, and methionine sulfoxide stood out, likely contributing to LDL-C perturbations by influencing oxidative stress and vascular function. In the case of HDL-C, arginine and glycine levels were notably diminished in response to the G × E interaction index, yet they exhibited a positive correlation with HDL-C concentrations. This suggests that G × E interactions may compromise vascular endothelial function, exacerbate oxidative damage, impair reverse cholesterol transport, and reduce HDL functionality. For TG, key metabolites including glycerol 3-phosphate, glycocholate, and purine were identified, potentially worsening hypertriglyceridemia by promoting glycerophospholipid synthesis and purine-related lipid accumulation under PBDE exposure. These findings collectively underscore the complex metabolic alterations driven by the interplay of environmental exposure and genetic predisposition across lipid profiles.

Subsequently, we conducted KEGG enrichment analysis on the overlapping metabolites to pinpoint key metabolic pathways (Fig. 4b). For both TC and LDL-C, the enriched pathways included glycine metabolism, alanine metabolism, and glutathione metabolism, indicating that glycine's involvement in antioxidant defense and amino acid metabolism may exacerbate cholesterol dysregulation. In the case of TG, enriched pathways such as glycerophospholipid metabolism and the citrate cycle emerged, suggesting that heightened lipid synthesis and energy dysregulation contribute to TG accumulation. We also performed pathway enrichment analysis on all differential metabolites as a

supplement. Beyond validating the primary pathways from the overlapping metabolite analysis, the results further identified key pathways enriched across all differential metabolites, notably arachidonic acid metabolism and unsaturated fatty acids biosynthesis (Supplementary Fig. 2c). This further broadens our understanding of the G × E interaction metabolic network.

To further elucidate how specific genetic variants disrupt lipid metabolism through these metabolites, we performed mediation analyses to assess their roles in the associations between PBDE-interacting variants and elevated dyslipidemia risk. This analysis identified 235 significant genetic variants influencing lipid phenotypes via 13 critical metabolites (Fig. 4c). Notably, glycine showed a substantial mediation effect across all four lipid traits, mediating the impact of 11, 12, 8, and 9 SNPs, respectively, including the most significant SNP for TC (rs1392296). Behenoyl dihydrosphingomyelin demonstrated the strongest mediation effect (18.62%) on LDL-C disruption linked to rs12487683, possibly by altering sphingolipid metabolism and membrane dynamics. Additionally, arginine metabolism, encompassing arginine, proline, leucylhydroxyproline, and related compounds, mediated impairments in TC and HDL-C, potentially through alterations in nitric oxide production and protein metabolism. These key metabolites and pathways also displayed significance in the direct associations between PBDE exposure and dyslipidemia (Supplementary Fig. 2de). Consistent with leading candidates from differential analyses and pathway enrichment, these results collectively emphasized the critical role of metabolic pathways, particularly those involving glycine, citrate, and lipid-related intermediates, in driving lipid disturbances driven by G × E interactions. 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 (Supplementary Fig. 2f). 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.

Fine mapping of the top variant in hypercholesterolemia

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) (Supplementary Data 5). 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. Within this locus, seven another significant SNPs displayed strong linkage disequilibrium (LD) ($r^2 \geq 0.6$) 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.

The eQTL analysis based on the GTEx database (v8) indicated that rs9869609 could significantly downregulate gene expression of *SLC6A20* in lipid-metabolic tissues, such as skeletal muscle ($\beta = -0.09$, $P = 0.0098$) (Fig. 5d). 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. 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 (Supplementary Fig. 3ab).

To validate this regulatory effect, particularly under PBDE exposure, we conducted dual-luciferase reporter assays in HepG2 cells, comparing the A and G alleles of rs9869609. We used BDE-47, a prevalent PBDE homolog with established epidemiological ties to dyslipidemia, as the exposure agent. Constructs with the rs9869609-G allele exhibited lower transcriptional activity than those with the A allele. Exposure to BDE-47 further reduced activity for both alleles, with a more pronounced decrease observed for the G allele (Fig. 5e), suggesting that BDE-47 amplifies the suppressive effect of the G allele on *SLC6A20* transcription.

To delve deeper into the regulatory mechanisms, we considered that allele-specific activity at rs9869609 might arise from differential TF binding affinities[30]. TF motif analysis using JASPAR identified BHLHE40 as a candidate TF with a preference for binding the

G allele (Fig. 5f). BHLHE40, a member of the Basic Helix-Loop-Helix family, is a recognized transcriptional repressor involved in metabolic dysfunction and alcohol-related liver disease[31].To test whether BHLHE40 mediates the allele-specific effects of rs9869609 (A > G), we performed reporter gene assays in HepG2 cells following siRNA-mediated knockdown of BHLHE40. The knockdown resulted in increased luciferase activity for constructs carrying either the A or G allele, with minimal differences between them. However, when exposed to BDE-47, the activity of both alleles was suppressed again, with the G allele showing a more significant reduction compared to the A allele (Fig. 5g). Collectively, these results indicate that rs9869609-G functions as an allele-specific silencer, enhancing the suppression of *SLC6A20* transcription in the presence of BDE-47 through stronger binding of BHLHE40 to the G allele.

Mechanism validation in BDE-47-exposed rs9869609-edited cells

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, we identified 1,221 significantly differentially expressed genes (DEG), including *SLC6A20*, *CETP* and *CAMK1D* (Supplementary Fig. 4a). 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 (Supplementary Fig. 4b). Notably, *SLC6A20* was among the downregulated genes (log_2(FC) = -2.21, FDR = 0.0032), pointing to its role in the early cellular response to PBDE exposure.

We then utilized CRISPR/Cas9 genome editing to create HepG2 cell lines with AA, AG, and GG genotypes at rs9869609 (Fig. 6a) and evaluated the interaction effects on *SLC6A20* expression. Cells with the GG genotype showed reduced *SLC6A20* mRNA and protein levels compared to those with the AA genotype. Upon BDE-47 exposure, *SLC6A20* expression reduced in both AA and GG cells, with a more pronounced reduction observed in the GG genotype relative to AA (Fig. 6b). These results suggest that BDE-47 amplifies the suppressive impact of the rs9869609-G allele on *SLC6A20* expression.

Drawing from our earlier analyses, rs9869609, along with the tag SNP, interacts with PBDE exposure to disrupt cholesterol levels, probably through metabolites including glycine, which is transported by SLC6A20. To test this effect in gene-edited cells, we measured cholesterol and glycine levels in cells with different rs9869609 genotypes under BDE-47 exposure. GG Cells showed elevated cholesterol levels compared to those with

the AA genotype, with a more significant increase following BDE-47 exposure (Fig. 6c). Similarly, intracellular glycine levels were lower in GG cells than in AA cells, with a more substantial decrease observed upon BDE-47 exposure (Fig. 6d). These results confirm the G × E interaction that impairs glycine availability and exacerbates cholesterol dysregulation.

SLC6A20, a key membrane protein for glycine transport, may hinder cholesterol degradation via the bile acid pathway by reducing intracellular glycine, an essential co-factor in forming glycine-conjugated bile acids for cholesterol excretion[23,32,33]. To clarify SLC6A20's role in this G × E interaction, we transfected HepG2 cells with siRNAs targeting *SLC6A20* (Fig. 6e) and examined the downstream consequences. As expected, *SLC6A20* knockdown markedly elevated cholesterol levels compared to control cells. Upon BDE-47 exposure, cholesterol levels increased in both knockdown and control groups, with a more pronounced rise in the knockdown cells relative to their non-exposed counterparts (Fig. 6f). Likewise, intracellular glycine levels dropped significantly in the *SLC6A20* knockdown group, and BDE-47 exposure further reduced glycine levels in both knockdown and control cells, with the knockdown group exhibiting the lowest levels compared to non-exposed knockdown cells (Fig. 6g). To evaluate the impact of *SLC6A20* on glycine-conjugated bile acids, which are essential for cholesterol excretion, we measured glycocholic acid (GCA) and glycodeoxycholic acid (GDCA) levels in *SLC6A20* knockdown HepG2 cells. Ultra-High-Performance Liquid Chromatography-High Resolution Mass Spectrometry (UPLC-HRMS) 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). Collectively, these findings suggest that BDE-47 enhances the rs9869609-G allele's suppression of *SLC6A20*, impairing glycine transport and driving cholesterol accumulation. This provides a mechanistic framework for how G × E interactions disrupt lipid homeostasis.

Discussion

G × E interactions contribute to individual susceptibility to environmental toxicants and complex metabolic disorders such as dyslipidemia, a risk factor for cardiovascular diseases. In this study, we conducted a genome-wide interaction analysis within a nationwide Chinese cohort, identifying 3,571 genetic variants across 622 loci that modulate the lipid-disrupting effects of PBDEs, a group of ubiquitous flame retardants

linked to metabolic disruptions. Among these, rs9869609 emerged as a key causal variant. We further used single-base CRISPR/Cas9 editing cell models to revealed its pathogenic mechanism, where the A > G allele of rs9869609 suppresses *SLC6A20* expression through BHLHE40-mediated transcriptional repression, impairing glycine-dependent cholesterol catabolism and promoting cholesterol accumulation, an effect significantly amplified by BDE-47 exposure (Fig. 6j). As a genome-wide G × E interaction study integrating population-level omics profiles with experimental validation, this work establishes a mechanistic link between genetic predisposition and metabolic susceptibility to widespread environmental toxins. These findings offer valuable insights for precision prevention strategies aimed at mitigating pollutant-induced dyslipidemia and its significant public health implications.

Traditional GWAS have identified genetic variants associated with metabolic disorders, such as diabetes and dyslipidemia[9,34]. However, they only explain a fraction of disease heritability, neglecting the modifying influence of environmental factors like pollutants or diet[35]. G × E interaction analysis overcomes these shortcomings by revealing conditional genetic effects tied to specific environments and capturing more phenotypic variance[36], which our study also confirms, of 1,037 PBDE-interactive genes, only 210 overlapped with known GWAS genes, showing traditional GWAS overlook exposure-specific pathways. For example, Luo et al. demonstrated that the *LCT-13910* C > T variant's association with T2D risk emerged only in lactase-persistent individuals with high milk intake, a nuance masked in traditional GWAS without dietary stratification[20]. Similarly, our Chinese cohort data revealed a key pattern tied to PBDE exposure. 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. 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 ready for clinical application, this PRS may serve as a research tool to help identify population subgroups with higher genetic susceptibility. In the future, such a tool could inform the design of targeted exposure mitigation strategies for high-risk individuals[37].

The interplay between genetic susceptibility and environmental exposures shapes the

regulation of metabolic pathways[11]. Our investigation revealed that the majority of interactive variants are located in non-coding regions enriched with regulatory elements such as enhancers. This indicates that these variants likely function by interfering with the binding of TF to enhancers, thereby modulating transcriptional activity [38,39]. Through *in silico* analysis of TF motif sequences surrounding these interacting loci, we detected a pronounced presence of variants within binding sites of lipid-regulating TFs, including BHLHE40, FOXA1, and CTCF[40], with BHLHE40, a transcriptional repressor that modulates metabolic genes like *PPARGC1A* and *SREBP1* being key [31,41]. In luciferase assays and CRISPR-edited HepG2 cells, the rs9869609 A > G allele, under BDE-47 exposure, amplifies BHLHE40's repressive activity, an effect reversed by *BHLHE40* knockdown, emphasizing the pivotal role of this TF in mediating the metabolic impact of PBDE exposure. We therefore propose that BDE-47 triggers the activation of BHLHE40, enhancing its allele-specific, binding-mediated transcriptional repression. This observation is consistent with previous research demonstrating pollutant-induced upregulation of *BHLHE40* in hepatic and fibrotic models[42-44]. Taken together, our results outline a coherent mechanism in which the rs9869609 risk allele collaborates with PBDE exposure to impair lipid metabolism through BHLHE40-mediated suppression of *SLC6A20*. 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 to changes in histone modification patterns[11,17,45]. 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 events[46]. These potential mechanisms provide clear directions for our future research.

To elucidate the pathways affected by gene-PBDE interactions, we annotated genes linked to interactive SNPs and conducted enrichment analyses, uncovering distinct yet interconnected lipid metabolism pathways: cholesterol biosynthesis for TC, lipoprotein assembly for LDL-C, reverse cholesterol transport for HDL-C, and fatty acid *β*-oxidation for TG. These findings partially align with DEGs from transcriptomic studies of PBDE toxicity[4], reinforcing their biological relevance. For cholesterol-related phenotypes, interactive genes were significantly enriched in steroid biosynthesis pathways, including *HMGCR*, a rate-limiting enzyme in cholesterol production, which is sensitive to environmental stressors[47]. PBDEs, as endocrine-disrupting chemicals, may amplify these

genetic effects by altering transcriptional regulation of these genes via nuclear receptors like PPARs or LXRs, which are known targets of lipophilic toxins[48]. Additionally, GO analysis highlighted lipoprotein metabolism processes, with variants likely influencing *APOB* expression, crucial for LDL particle formation[49]. These regulatory mechanisms, driven by non-coding elements, position G × E interaction analysis as a powerful tool to identify susceptibility loci in regulatory hubs, which could be complemented by transcriptomics to capture downstream responses[37].

The mechanistic insights into how interactive loci regulate transcription in PBDE-induced lipid disruption could enhance our understanding of G × E interactions and identify potential targets for therapeutic or policy interventions[49]. Our CRISPR-edited cell models demonstrated that the rs9869609 variant functions as an enhancer regulating *SLC6A20,* a glycine transporter linked to energy metabolism, with significant eQTL effects despite not being prominent in lipid GWAS metabolism[50,51]. Suppression of *SLC6A20* diminishes intracellular glycine levels, impairing bile acid conjugation, which is a vital process in cholesterol catabolism[52]. Glycine serves as an essential substrate for the covalent binding to primary bile acids, leading to the formation of hydrophilic bile salts that promote cholesterol degradation[33,53,54]. Metabolomic analyses confirmed decreased glycine levels following high PBDE exposure. Notably, *SLC6A20* overexpression restored glycine bioavailability and reduced serum cholesterol, suggesting a potential compensatory mechanism against PBDE-induced metabolic toxicity[55]. These findings delineate a coherent pathway through which the interactive variant influences lipid homeostasis under environmental exposure[56]. 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. Beyond *SLC6A20*, other genes contribute to PBDE-induced dyslipidemia, underscoring the polygenic nature of this process *CETP* mediates cholesterol ester transfer between lipoproteins[57], *CTCF* acts as an epigenetic regulator of lipid-related genes[58], 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.

Other key metabolites implicated in G × E interactions also show promise for predicting pollutant-induced metabolic disorders. Metabolomic profiling demonstrates that PBDEs suppress glutathione biosynthesis, a vital process for sustaining cellular antioxidant defenses[59]. This depletion amplifies oxidative stress, triggering lipid peroxidation and

activating NF-κB-driven proinflammatory cascades, which in turn promote lipid accumulation and metabolic inflammation[60]. Additionally, PBDEs disrupt energy metabolism by inhibiting enzymes in the tricarboxylic acid (TCA) cycle, leading to decreased production of NADH and $FADH_2$[61]. This impairment redirects acetyl-CoA flux from energy production toward *de novo* lipogenesis, driven by the upregulation of fatty acid synthase (FASN) and acetyl-CoA carboxylase 1 (ACC1)[62]. These disruptions create a self-reinforcing cycle of lipid accumulation, oxidative stress, and inflammation, providing a mechanistic basis for PBDE-induced dyslipidemia and its associated increase in cardiovascular risk. 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. This metabolite-centric perspective underscores the intricate relationship between environmental toxicants and metabolic pathways, offering valuable insights into broader implications for disease prediction and control.

Several limitations warrant careful consideration when interpreting our findings. 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. 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. 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*. 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.

In summary, this nationwide study of G × E interactions in a Chinese population identified 3,571 genetic variants that interact with PBDE exposure to influence dyslipidemia susceptibility. A polygenic risk score derived from these variants improved risk prediction, which may aid in identifying higher-risk individuals. Mechanistically, we demonstrate that under BDE-47 exposure, the rs9869609 risk allele promotes BHLHE40-mediated suppression of *SLC6A20*, thereby disrupting glycine-dependent cholesterol metabolism, a pathway validated using CRISPR/Cas9 editing. These findings clarify links between

environmental exposure, genetic variation, and metabolic outcomes, providing a basis for risk assessment and environmental health policy.

Methods

Ethics

The study complies with all relevant regulations and the Declaration of Helsinki and was approved by the Ethical Review Committee of the National Institute of Environmental Health (NIEH), Chinese Center for Disease Control and Prevention (China CDC) as part of the CNHBM cohort (no. 201701). Moreover, written informed consent was obtained from each participant before the study began. For all participants who were minors, written informed consent was obtained from their parent or legal guardian, in accordance with CNHBM ethical protocols and regulations.

Data sources and study participants

This study leverages data from a subset of the China National Human Biomonitoring (CNHBM) baseline cohort, a cross-regional survey that enrolled 21,888 participants aged 3-79 years, representing multiple ethnicities including Han (86.4%), Tibetan, Zhuang, Mongolian, and others. For the present multi-omics analysis, a secondary sample of 1,824 participants was selected from the baseline cohort using stratified sampling across all age groups to maintain demographic representativeness. Peripheral venous blood samples were collected in a standardized manner from all enrolled individuals, immediately processed via centrifugation to separate serum and blood cells. Serum samples were utilized for metabolomics analysis as well as exposure metabolomics, while blood cells were reserved for genomics analysis. Detailed protocols for participant recruitment, sample collection and processing have been described previously[25]. Following the exclusion of participants under 18 years and those with incomplete data, the final analytical sample consisted of 871 individuals (Fig 1, Supplementary Fig. 1ab).

Multi-omics Data Generation

Exposure Data: Fourteen serum PBDE congeners (BDE-17, 28, 47, 66, 71, 85, 99, 100, 138, 153, 154, 183, 190, 209) were analyzed via the CDC-established method on an Agilent 7890B gas chromatograph coupled to a Waters Autospec Premier dual-focus magnetic mass spectrometer (VSIR mode, resolution 10,000). Serum samples were pretreated with acid-methanol, extracted with organic solvents, and extracts were concentrated, SPE-purified and reconstituted prior to analysis. Instrumental separation was performed on a capillary column (helium carrier gas) with pulsed splitless injection

and positive EI (35 eV) ionization. The limit of detection (LOD) of these pollutants ranged from 0.01 to 1.20 μg/L, with non-detectable measurements imputed as LOD/√2. Detailed LODs and median concentrations are presented in Supplementary Data 7. Participants were stratified into high- and low-exposure groups by the median of total PBDE levels (ΣPBDE; 0.8157 μg/L).

Genomics Data: A 1.5-mL peripheral blood sample was collected from each subject at baseline for DNA extraction. Samples were genotyped using the Asian Screening Array-China Health Industry Alliance (ASA-CHIA) BeadChip (Illumina, Inc., San Diego, California, USA), following the Illumina Infinium HTS chip standard operating procedures[63]. After applying quality control measures, raw genotyping data for ~700,000 SNPs were retained for each sample, with alignment to the human genome Hg19 assembly. Genotype imputation was performed using the Michigan Imputation Server (https://imputationserver.sph.umich.edu). SNPs were excluded based on the following criteria: (1) imputation quality score < 0.2; (2) MAF < 0.05; (3) SNP call rate < 95%; (4) Hardy-Weinberg equilibrium *P*-value < 1×10^{-6}.

Untargeted Metabolomics Data: Metabolomic profiling of serum was conducted on the CalOmics platform (Calibra Scientific Inc.). A total of 1824 human serum samples were analyzed as single injections without technical replicates. Sample preparation involved protein precipitation and metabolite extraction using methanol, and quality control (QC) measures included pooled QC samples, process blanks, and deuterated internal standards. The analysis was conducted using an ACQUITY 2D UPLC system (Waters) coupled to a Q Exactive Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific). Four complementary liquid chromatography- mass spectrometry (LC-MS/MS) methods were employed: two using C18 columns under different mobile phase conditions for electrospray ionization positive mode (ESI^+, denoted as POSa and POSb) and negative mode (ESI^-, denoted as NEGa), and one using a HILIC column (BEH Amide) for ESI^- mode (denoted as NEGb). Full scans were acquired at 35,000 resolutions (m/z 70–1000). Raw data were processed using the in-house software MetaPro[64] for peak picking, alignment, and integration. Metabolites were identified with Metabolomics Standards Initiative (MSI) Level 1 confidence[65,66] by matching retention time (±0.1 min), exact mass (error <10 ppm), and MS/MS spectra (match score >75%) against an in-house library of authentic standards.

Phenotypes and Covariates

The primary outcome variables in this study were four blood lipid components: TC, LDL-C, HDL-C, and TG. Blood samples were obtained through venipuncture, immediately centrifuged, and the resulting serum was analyzed using a standardized biochemical analyzer to determine the concentrations of these four lipid markers. To address right-skewed distributions and meet the assumptions of parametric statistical tests, the four lipid phenotypes underwent logarithmic transformation (log_{10}).

To adjust for potential confounding factors, the following covariates were included in all statistical models: sex (male and female), age group (19- 39, 40- 59, and ≥ 60 years), BMI (< 18.5, 18.5- 25, 25- 30, ≥ 30 kg/m^2), smoking status (smoked, defined as at least once a week, and never smoked), and alcohol consumption (drinking, defined at least once a week, and never drinking). BMI was derived from measured height and weight (kg/m²). These covariates were included as adjustment factors in the statistical models to minimize potential biases in assessing the relationship between PBDE exposure and lipid outcomes.

Genome-wide G × E interaction analysis

To examine the associations between PBDE exposure, SNPs, and the four lipid profiles (TC, LDL-C, HDL-C, and TG), we utilized a generalized linear regression model, implemented in PLINK (v.1.90), to evaluate the interaction effects of PBDE exposure and SNPs on lipid levels. As demonstrated by recent studies[67,68], such G × E interactions allow for the identification of interaction loci that respond to environmental exposure and modulate blood lipid levels, thereby facilitating the characterization of individuals at high risk of dyslipidemia associated with high environmental exposure. We then generated Manhattan plots to illustrate the genomic distribution of significant loci (FDR < 0.05), with the top five SNPs annotated to their corresponding genes; for the most significant SNP in each plot, a stratified analysis of dyslipidemia risk was performed to highlight the interaction effect.

PRS construction for lipid phenotype

To assess the cumulative interaction between lipid-specific genetic predisposition and PBDE exposure on elevated levels of each lipid phenotype, we constructed a PRS using significant SNPs identified from the prior interaction analysis. Independent SNPs were selected for the models through a linkage disequilibrium-driven clumping approach in PLINK (v.1.90), with parameters set as follows: --clump-p 0.0001, --clump-r^2 0.20 --clump-kb 250. A total of 172, 168, 130, and 124 independent SNPs were included in the PRS

model for the four lipid phenotypes (TC, LDL-C, HDL-C, and TG), respectively. Participants were then scored additively in a weighted manner based on the number of risk alleles they carried, utilizing the clumping-plus-thresholding (C+T) method according to the formula:

$$\mathrm{PRS} = \sum_{i=1}^{n} \beta_i SNP_i \quad (1)$$

where n represents the total number of variants, SNP_i (0, 1, or 2) denotes the count of risk alleles (corresponding to homozygous non-risk allele, heterozygous allele, or homozygous risk allele, respectively), and β_i is the effect size of the risk allele[69]. Using the lowest quintile (bottom 20%) of the PRS as the reference group, increased levels of corresponding lipid were estimated across the remaining four PRS quintiles in linear regression models, stratified by PBDE exposure levels. The dose-response relationship was assessed via trend analysis, with statistical significance defined as $P_{trend} < 0.05$.

Annotation and functional enrichment of SNPs

To investigate the enriched characteristics of PBDE-interacting SNPs, a control variant set was generated with a 1:1 match to the interactive SNPs for each lipid phenotype, ensuring equivalence in MAF, number of LD SNPs, and variant type. This was achieved using the web-based tool vSampler (version 8.0, http://mulinlab.org/vsampler/). Both interactive SNPs and their control variants were annotated using SnpEff software (version 8) with the GRCh37 reference genome from ENSEMBL. Variants were categorized into 5′ untranslated region (5'UTR), 3′ untranslated region (3′UTR), upstream, downstream, intron, and intergenic region (IGR). Functional annotation data, including histone modification ChIP-seq peaks (H3K27ac, H3K27me3, H3K4me1, H3K4me2 and H3K9ac), DNase I hypersensitive sites, and binding site of 59 transcription factors, were retrived from the ENCODE portal (https://www.encodeproject.org/data/annotations/)[70].

To identify the putative interactive genes, we employed an integrated framework for assigning target genes to interactive SNPs (Fig. 3e): (1) Physical distance strategy -- the nearest genes within a 10kb region surrounding the interactive SNPs were included as candidates; (2) eQTL strategy -- eQTL data from the GTEx database were leveraged to pinpoint target genes whose expression levels were affected by interactive SNPs; (3) MAGMA strategy -- target genes were determined by aggregating variant signals to the gene level using Multi-marker Analysis of GenoMic Annotation (MAGMA)[71]. Genes identified through this approach were designated as PBDE-interacting genes. To explore their biological significance, KEGG and GO term enrichment analyses[72] were performed, with the top enriched pathways visualized.

Fine mapping analysis

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 databases[73]. CADD first filters out potentially harmless variants, focusing on functionally significant sites[74]. RegulomeDB then determines whether these sites reside within regulatory regions and assesses their regulatory potential[75]. Finally, 3DSNP connects variants within regulatory regions to their actual distal target genes, completing the full biological narrative from variant to target gene[76]. Specifically, we apply minimum-maximum normalization to the raw outputs of these databases, uniformly converting their scores to the 0–1 range.

Cell culture and cell viability detection

Biological experiments were conducted using the HepG2 cell line, a model widely employed in drug metabolism and hepatotoxicity studies due to its functional resemblance to human liver tissue[77]. The HepG2 cells were obtained from the China Center for Type Culture Collection (Wuhan, China), and cultured in Dulbecco's Modified Eagle's Medium (DMEM; Gibco, USA) supplemented with 10% fetal bovine serum (FBS; Gibco, USA) and 1% antibiotics (100 U/mL penicillin and 0.1 mg/mL streptomycin). Cells were maintained in a humidified incubator at 37 °C with 5% CO_2. Cell line authenticity was verified through short tandem repeat profiling (Applied Biosystems), and the absence of Mycoplasma contamination was confirmed using the MycoAlert kit. The most recent testing data was in May 10, 2024.

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 pollutants[78].

Cell viability was assessed using the CCK-8 assay following the manufacturer's protocol. HepG2 cells were cultured in 96-well plates with exposure to gradient concentrations of BDE-47 (0, 6.25, 12.5, 25, 50, 100 μM) for durations of 24, 48, and 72 hours. Following exposure, the culture medium was replaced with medium containing 10% CCK-8 solution, and cells were incubated for an additional 50 minutes. Optical density (OD) at 450 nm was measured using a microplate reader (BioTek Epoch, USA), and cell viability (%) was calculated relative to the control group. Growth curves were plotted to assess the dose- and time-dependent effects of BDE-47 on HepG2 cell proliferation. Results indicated significant proliferation inhibition at 50 μM after 48-hour treatment (Supplementary Fig. 4c). Therefore, the 48-hour exposure to 25 μM BDE-47 was selected for subsequent toxicological assessments to balance observable effects with maintained cell viability.

Detection of lipids and metabolites

TC was quantified in HepG2 cells from BDE‑47 exposure and control groups (3 biological replicates per group, each with 3 technical replicates) using a commercial total cholesterol assay kit (Solarbio, Beijing, China). Intracellular glycine levels were determined in parallel with a glycine assay kit (Jiancheng Bioengineering Institute, Nanjing, China). Cell processing and analysis were performed according to the respective manufacturer's protocols. All measurements were normalized to the total protein concentration of each sample to ensure accurate quantification.

Targeted UPLC-HRMS quantified intracellular GCA and GDCA. Methanol-extracted samples (n = 18; BDE-47-exposed vs. control HepG2 cells, 3 biologicals × 3 technical replicates) were separated on a Hypersil GOLD C18 column (0.1% formic acid/acetonitrile, 15-min gradient) and analyzed in negative-ion full-scan mode (m/z 70–1050, resolution 70,000) using a Dionex UPLC/Q Exactive Orbitrap system. External calibration curves ($R^2 \geq 0.99$) were used for quantification, with pooled QC samples injected every 10 runs. Data were acquired with Xcalibur 4.0 and processed using TraceFinder 5.1 for peak integration and concentration calculation.

Dual-luciferase reporter assays

DNA fragments spanning 1,000 base pairs around the SNP rs9869609 (A or G allele)

were subcloned into the pGL3-promoter vector (Promega, USA) and synthesized by Tsingke Biological Technology (Wuhan, China). For transient transfection assays, HepG2 cells were cultured in 96-well plates and co-transfected with constructed luciferase vector (containing either the rs9869609-A or G allele) and the pRL-SV40 Renilla luciferase plasmid (Promega, USA) using Lipofectamine 3000 (Invitrogen, USA). Twenty-four hours post-transfection, the control group was maintained in standard medium, while the experimental group was treated with medium containing 12.5 μM BDE-47 for an additional 24 hours. Luciferase activity was measured using the Dual-Luciferase Reporter Assay Kit (Promega, USA) according to the manufacturer's protocol, with Renilla luciferase luminescence normalized to that of firefly luciferase for each sample.

RNA interference

For RNA interference, small interfering RNA (siRNA) oligonucleotides targeting *BHLHE40* and *SLC6A20*, along with a non-targeting siRNA control, were obtained from Tsingke Biological Technology (Wuhan, China), with the sequences provided in Supplementary Data 8. Transfection was performed using Lipofectamine RNAiMAX (Invitrogen, USA), and the knockdown efficiency was validated by quantitative reverse transcription PCR (qRT-PCR) (Supplementary Fig. 4de).

RNA-seq

HepG2 cells were treated with 12.5 μM BDE‑47 or vehicle control for 24 hours (n = 3 biological replicates per group). Total RNA was extracted using TRIzol reagent (Invitrogen), and RNA integrity was verified (RIN > 8.0) on an Agilent Bioanalyzer. Ribosomal RNA was depleted, and stranded cDNA libraries were constructed following the manufacturer's protocol (Vazyme). Libraries were sequenced on a DNBSEQ‑T7 platform (MGI Tech) in 150‑bp paired‑end mode. Raw reads were quality‑trimmed with fastp (v0.23.2) and aligned to the human reference genome (GRCh38) using HISAT2 (v2.2.1). Gene‑level counts were obtained with featureCounts (v2.0.1). Differential expression analysis was performed with edgeR (v3.40.2), applying a threshold of $|\log_2 FC| > 1$ and FDR < 0.05. The raw sequencing data have been deposited in the Genome Sequence Archive for Human (GSA-Human) under accession code HRA014233.

CRISPR/Cas9 mediated single-base mutation

CRISPR/Cas9-mediated editing of the rs9869609 variant was conducted in the HepG2 cell line. The single-guide RNA (sgRNA) was designed using the established CRISPR design tool (https://cctop.cos.uni-heidelberg.de/).The sgRNA was subcloned downstream

of the human U6 promoter into the plasmid pSpCas9(BB)-2A-Puro (PX459) V2.0 (Cat# 48138, GeneAdv, Suzhou) via BbsI restriction sites. Additionally, a double-stranded donor (dsDonor) sequence (1.6 kb) containing the rs9869609 variant (A > G) was designed based on the highly efficient sgRNA and inserted into the pUC57 vector through the EcoRV restriction site. All oligonucleotides used are listed in Supplementary Data 8.

Transfection was performed in HepG2 cells with 80% confluency. A total of 2.5 μg of the Cas9 plasmid (PX459-sgRNA) and 1 μg of dsDonor were co-transfected into cells using Lipofectamine 3000. After 48 hours, 0.8 mg/mL puromycin was added to select for transfected cells. Surviving cells were then trypsinized and seeded into 96-well plates via fluorescence-activated cell sorting (FACS) to isolate single-cell clones. These clones were subsequently subcultured and verified by Sanger sequencing over a period of 14-21 days.

QRT-PCR and western blotting (WB)

Total RNA was extracted from cells with Trizol Reagent (Thermo Fisher Scientific, USA). Reverse transcription was carried out with PrimeScript™ RT-PCR Kit (Takara, Japan), followed by qPCR on a QuantStudio 5 qPCR system (Applied Biosystems, Thermo Fisher) using SYBR™ Green PCR Master Mix (Takara, Japan). Gene expression levels were normalized to *GAPDH*, which served as an internal reference. The corresponding primers were listed in Supplementary Data 8.

Total protein was harvested from HepG2 cells of BDE47 exposure and control groups (3 biological replicates per group) using RIPA lysis buffer (Beyotime, China) supplemented with protease inhibitors PMSF (Beyotime, China), cOmplete cocktail (Sigma, USA) and PhosSTOP (Sigma, USA). Protein was quantified using BCA reaction (Beyotime, China) and denaturated at 99 °C for 5 min. Equal amount of protein was separated by electrophoresis in 10% SDS-PAGE gels and transferred onto 0.45 mm PVDF membrane. Protein was incubated with antibodies against SLC6A20 (1:1000, Proteintech, Cat# 30843-1-AP) and β-actin (1:1000, Proteintech, Cat# 20536-1-AP) at 4 °C overnight. HRP-conjugated anti-rabbit IgG (1:5,000, Proteintech, Cat# SA00001-2) was used as secondary antibody. Chemiluminescence signal was developed with enhanced chemiluminescence substrate (Beyotime, China) by Image Lab software. Unprocessed scans of the key blots (for Fig. 6B) are supplied in the Source Data file, and the detail information of antibodies was provided in the Reporting summary.

Statistics & Reproducibility

This study employed R (version 4.4.1) to perform the statistical analysis of epidemiological and experimental data. The generalized linear regression model (R package '*stats*'), adjusted by covariates, was employed to analyze the metabolomic data, identifying metabolites associated with the G × E interaction index and corresponding lipid levels. The interaction index, calculated as the product of PBDE exposure and the PRS for each lipid phenotype, represented the overall interaction effect. Overlapping metabolite sets were further enriched using the KEGG database within MetaboAnalyst 3.0[79].To examine the mediating role of serum differential metabolites in G × E interactions with lipid phenotypes, causal mediation analyses were conducted using the R package '*mediation*'. After covariates adjustment, we calculated the total effect (TE), natural direct effect (NDE), and natural indirect effect (NIE), with the mediation proportion of significant metabolites determined as the ratio of NIE to TE, reflecting the mediation extent. For experimental data, two-group comparisons were analyzed using unpaired *t*-tests in Prism software (GraphPad). For multiple-group comparisons involving normally distributed data, one-way analysis of variance (ANOVA) was applied, followed by Fisher's least significant difference (LSD) tests for post-hoc analysis. A P-value < 0.05 was considered statistically significant, with significance levels denoted in the corresponding figures.

Sample sizes were determined based on the predefined sample screening protocol for epidemiological and experimental cohorts. No statistical method was used to predetermine sample size. No data were excluded from the analyses. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment. All key experimental assays were independently replicated at least three times, with consistent results confirming the reproducibility of findings.

Data availability

Source data are provided with this paper. The raw RNA-seq data are publicly available in the GSA-Human repository under accession code HRA014233 [https://ngdc.cncb.ac.cn/gsa-human/s/b6tzKUmT] (RNA-seq of HepG2 cells exposed to BDE47 and control). The genomic, metabolomic, and exposomic data of the study population involve sensitive biomonitoring data. Due to restrictions specified in agreements with collaborating institutions (cohort initiators) and the informed consent forms signed by study participants, these data cannot be made publicly available and are subject to controlled access. The request of these individual data is suggested by sending an email to the corresponding author Professor Xiaoming Shi (shixm@chinacdc.cn) and

Professor Xiaoping Miao (xpmiao@whu.edu.cn). Requests should include name, affiliation and contact details of the person requesting the data, which data are requested and the purpose of requesting the data. Requests will be subject to consideration by the management committee of the corresponding institutes and the sample collection institutes. If approved, the corresponding author will send the request data by email. Time frame for a response will be within 3 months. Data requests under agreement will be considered for purposes of reproducing the data and subject to appropriate confidentiality obligations and restrictions. Applicants must promise that these individual data applied for will only be used for scientific research and cannot be publicly released.

Reference

1. Chew, N.W.S., et al. The global burden of metabolic disease: Data from 2000 to 2019. Cell metabolism 35, 414-428.e413 (2023).
2. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet (London, England) 392, 1923-1994 (2018).
3. Lan, Y., Gao, X., Xu, H. & Li, M. 20 years of polybrominated diphenyl ethers on toxicity assessments. Water research 249, 121007 (2024).
4. Wei, J., et al. Metabolomics and lipidomics study unveils the impact of polybrominated diphenyl ether-47 on breast cancer mice. Journal of hazardous materials 390, 121451 (2020).
5. Yin, W., Xu, R., Zou, J., Wang, Y. & Zhang, Y. Single and combined association between brominated flame retardants and cardiovascular disease: a large-scale cross-sectional study. Frontiers in public health 12, 1357052 (2024).
6. Zhang, Q., et al. Associations between polybrominated diphenyl ethers (PBDEs) levels in adipose tissues and blood lipids in women of Shantou, China. Environmental research 214, 114096 (2022).
7. Boutot, M.E., et al. In Utero Exposure to Persistent Organic Pollutants and Childhood Lipid Levels. Metabolites 11(2021).
8. Zhao, X., et al. Polybrominated diphenyl ethers in serum from residents living in a brominated flame retardant production area: Occurrence, influencing factors, and relationships with thyroid and liver function. Environmental pollution (Barking, Essex : 1987) 270, 116046 (2021).
9. Klarin, D., et al. Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. Nature genetics 50, 1514-1523 (2018).
10. Noordam, R., et al. Multi-ancestry sleep-by-SNP interaction analysis in 126,926 individuals reveals lipid loci stratified by sleep duration. Nature communications 10, 5121 (2019).
11. Wu, H., Eckhardt, C.M. & Baccarelli, A.A. Molecular mechanisms of environmental exposures and human disease. Nature reviews. Genetics 24, 332-344 (2023).
12. Messerlian, C., Martinez, R.M., Hauser, R. & Baccarelli, A.A. 'Omics' and endocrine-disrupting chemicals - new paths forward. Nature reviews. Endocrinology 13, 740-748 (2017).
13. Ortega-Azorín, C., et al. Associations of the FTO rs9939609 and the MC4R rs17782313 polymorphisms with type 2 diabetes are modulated by diet, being higher when adherence to the Mediterranean diet pattern is low. Cardiovascular diabetology 11, 137 (2012).
14. Dietrich, S., et al. Gene-lifestyle interaction on risk of type 2 diabetes: A systematic review. Obesity reviews : an official journal of the International Association for the Study of Obesity 20, 1557-1571 (2019).

15. Brokken, L.J.S. & Giwercman, Y.L. Gene-environment interactions in male reproductive health: Special reference to the aryl hydrocarbon receptor signaling pathway. 16, 89-96 (2014).
16. Zhou, Z., et al. Relationships among N,N-dimethylformamide exposure, CYP2E1 and TM6SF2 genes, and non-alcoholic fatty liver disease. Ecotoxicology and environmental safety 228, 112986 (2021).
17. Migliore, L. & Coppedè, F. Gene-environment interactions in Alzheimer disease: the emerging role of epigenetics. Nature reviews. Neurology 18, 643-660 (2022).
18. Surendran, P., et al. Rare and common genetic determinants of metabolic individuality and their effects on human health. Nature medicine 28, 2321-2332 (2022).
19. Luukkonen, P.K., et al. The PNPLA3 I148M variant increases ketogenesis and decreases hepatic de novo lipogenesis and mitochondrial function in humans. Cell metabolism 35, 1887-1896.e1885 (2023).
20. Luo, K., et al. Variant of the lactase LCT gene explains association between milk intake and incident type 2 diabetes. Nature Metabolism 6, 169-186 (2024).
21. Solon-Biet, S.M., et al. Branched chain amino acids impact health and lifespan indirectly via amino acid balance and appetite control. Nat Metab 1, 532-545 (2019).
22. Rom, O., et al. Glycine-based treatment ameliorates NAFLD by modulating fatty acid oxidation, glutathione synthesis, and the gut microbiome. Science translational medicine 12(2020).
23. Ghrayeb, A., et al. Serine synthesis via reversed SHMT2 activity drives glycine depletion and acetaminophen hepatotoxicity in MASLD. Cell metabolism 36, 116-129.e117 (2024).
24. Fu, A., et al. Glucose metabolism and pyruvate carboxylase enhance glutathione synthesis and restrict oxidative stress in pancreatic islets. Cell reports 37, 110037 (2021).
25. Cao, Z., et al. Cohort profile: China National Human Biomonitoring (CNHBM)-A nationally representative, prospective cohort in Chinese population. Environment international 146, 106252 (2021).
26. Talmud, P.J., et al. Linkage of the cholesteryl ester transfer protein (CETP) gene to LDL particle size: use of a novel tetranucleotide repeat within the CETP promoter. Circulation 101, 2461-2466 (2000).
27. Reshetnikov, V.V., et al. Genes associated with cognitive performance in the Morris water maze: an RNA-seq study. Scientific reports 10, 22078 (2020).
28. Xue, A., et al. Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. Nature communications 9, 2941 (2018).
29. Tokarz, J., et al. Endocrinology Meets Metabolomics: Achievements, Pitfalls, and Challenges. Trends in endocrinology and metabolism: TEM 28, 705-721 (2017).

30. Shuai, Y., et al. TEAD4 modulated LncRNA MNX1-AS1 contributes to gastric cancer progression partly through suppressing BTG2 and activating BCL2. Molecular cancer 19, 6 (2020).
31. Podleśny-Drabiniok, A., et al. BHLHE40/41 regulate microglia and peripheral macrophage responses associated with Alzheimer's disease and other disorders of lipid-rich tissues. Nature communications 15, 2058 (2024).
32. Bae, M., et al. SLC6A20 transporter: a novel regulator of brain glycine homeostasis and NMDAR function. EMBO molecular medicine 13, e12632 (2021).
33. Chretien, Y., et al. Bile acid glycine and taurine conjugates in serum of patients with primary biliary cirrhosis: effect of ursodeoxycholic treatment. Gut 30, 1110-1115 (1989).
34. Mahajan, A., et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nature genetics 50, 1505-1513 (2018).
35. Manolio, T.A., et al. Finding the missing heritability of complex diseases. Nature 461, 747-753 (2009).
36. Thomas, D. Gene--environment-wide association studies: emerging approaches. Nature reviews. Genetics 11, 259-272 (2010).
37. Ritchie, M.D., Holzinger, E.R., Li, R., Pendergrass, S.A. & Kim, D. Methods of integrating data to uncover genotype-phenotype interactions. Nature reviews. Genetics 16, 85-97 (2015).
38. Turner, A.W., et al. Single-nucleus chromatin accessibility profiling highlights regulatory mechanisms of coronary artery disease risk. Nature genetics 54, 804-816 (2022).
39. Todeschini, A.L., Georges, A. & Veitia, R.A. Transcription factors: specific DNA binding and specific gene regulation. Trends in genetics : TIG 30, 211-219 (2014).
40. Deplancke, B., Alpern, D. & Gardeux, V. The Genetics of Transcription Factor DNA Binding Variation. Cell 166, 538-554 (2016).
41. Honma, S., et al. Dec1 and Dec2 are regulators of the mammalian molecular clock. Nature 419, 841-844 (2002).
42. Yao, W., et al. Transcriptome analysis reveals a protective role of liver X receptor alpha against silica particle-induced experimental silicosis. The Science of the total environment 747, 141531 (2020).
43. Andersen, M.E., et al. Assessing molecular initiating events (MIEs), key events (KEs) and modulating factors (MFs) for styrene responses in mouse lungs using whole genome gene expression profiling following 1-day and multi-week exposures. Toxicology and applied pharmacology 335, 28-40 (2017).
44. Andersen, M.E., et al. Combining transcriptomics and PBPK modeling indicates a primary role of hypoxia and altered circadian signaling in dichloromethane carcinogenicity in mouse lung and liver. Toxicology and applied pharmacology 332, 149-158 (2017).

45. Hua, W., et al. Transgenerational Effects of Arsenic Exposure on Learning and Memory in Rats: Crosstalk between Arsenic Methylation, Hippocampal Metabolism, and Histone Modifications. Environmental science & technology 58, 6475-6486 (2024).
46. Pajoro, A., Severing, E., Angenent, G.C. & Immink, R.G.H. Histone H3 lysine 36 methylation affects temperature-induced alternative splicing and flowering in plants. Genome biology 18, 102 (2017).
47. Brown, M.S. & Goldstein, J.L. The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. Cell 89, 331-340 (1997).
48. Martens, N., et al. Activation of Liver X Receptors and Peroxisome Proliferator-Activated Receptors by Lipid Extracts of Brown Seaweeds: A Potential Application in Alzheimer's Disease? Nutrients 15(2023).
49. Darnerud, P.O., Eriksen, G.S., Jóhannesson, T., Larsen, P.B. & Viluksela, M. Polybrominated diphenyl ethers: occurrence, dietary exposure, and toxicology. Environmental health perspectives 109 Suppl 1, 49-68 (2001).
50. Bröer, S. Amino acid transport across mammalian intestinal and renal epithelia. Physiological reviews 88, 249-286 (2008).
51. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science (New York, N.Y.) 369, 1318-1330 (2020).
52. Yue, J.T., Mighiu, P.I., Naples, M., Adeli, K. & Lam, T.K. Glycine normalizes hepatic triglyceride-rich VLDL secretion by triggering the CNS in high-fat fed rats. Circulation research 110, 1345-1354 (2012).
53. Guzior, D.V. & Quinn, R.A. Review: microbial transformations of human bile acids. Microbiome 9, 140 (2021).
54. Venkatesh, R., Srinivasan, K. & Singh, S.A. Effect of arginine:lysine and glycine:methionine intake ratios on dyslipidemia and selected biomarkers implicated in cardiovascular disease: A study with hypercholesterolemic rats. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie 91, 408-414 (2017).
55. Gao, H., et al. Perigestational low-dose BDE-47 exposure alters maternal serum metabolome and results in sex-specific weight gain in adult offspring. Chemosphere 233, 174-182 (2019).
56. Smith, G.D. & Ebrahim, S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? International journal of epidemiology 32, 1-22 (2003).
57. Borén, J., Taskinen, M.R., Björnson, E. & Packard, C.J. Metabolism of triglyceride-rich lipoproteins in health and dyslipidaemia. Nature reviews. Cardiology 19, 577-592 (2022).
58. Jin, L., et al. Dynamic chromatin architecture of the porcine adipose tissues with weight gain and loss. Nature communications 14, 3457 (2023).

59. Gianì, F., et al. Antioxidant Defense Capacity Is Reduced in Thyroid Stem/Precursor Cells Compared to Differentiated Thyrocytes. International journal of molecular sciences 24(2023).
60. Karunakaran, D., et al. RIPK1 Expression Associates With Inflammation in Early Atherosclerosis in Humans and Can Be Therapeutically Silenced to Reduce NF-κB Activation and Atherogenesis in Mice. Circulation 143, 163-177 (2021).
61. Jiang, Y., et al. Hemocytes in blue mussel Mytilus edulis adopt different energy supply modes to cope with different BDE-47 exposures. The Science of the total environment 885, 163766 (2023).
62. Xia, B., et al. BDE-47 induces metabolic dysfunction-associated steatotic liver disease (MASLD) through CD36-mediated increased fatty acid uptake and PPARα-induced abnormal fatty acid oxidation in BALB/c mice. Toxicology letters 391, 100-110 (2024).
63. Van Asselt, A.J. & Ehli, E.A. Whole-Genome Genotyping Using DNA Microarrays for Population Genetics. Methods in molecular biology (Clifton, N.J.) 2418, 269-287 (2022).
64. An, S., et al. MetaPro: a web-based metabolomics application for LC-MS data batch inspection and library curation. Metabolomics : Official journal of the Metabolomic Society 19, 57 (2023).
65. Sumner, L.W., et al. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics : Official journal of the Metabolomic Society 3, 211-221 (2007).
66. Phapale, P., Rai, V., Mohanty, A.K. & Srivastava, S. Untargeted Metabolomics Workshop Report: Quality Control Considerations from Sample Preparation to Data Analysis. Journal of the American Society for Mass Spectrometry 31, 2006-2010 (2020).
67. Ma, Y., Zhao, Y., Zhang, J.F. & Bi, W. Efficient and accurate framework for genome-wide gene-environment interaction analysis in large-scale biobanks. Nature communications 16, 3064 (2025).
68. Dong, Z., et al. Incorporating additive genetic effects and linkage disequilibrium information to discover gene-environment interactions using BV-LDER-GE. Genome biology 26, 332 (2025).
69. Cai, Y., et al. An atlas of genetic effects on cellular composition of the tumor microenvironment. Nature immunology 25, 1959-1975 (2024).
70. Luo, Y., et al. New developments on the Encyclopedia of DNA Elements (ENCODE) data portal. Nucleic acids research 48, D882-d889 (2020).
71. de Leeuw, C.A., Mooij, J.M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. PLoS computational biology 11, e1004219 (2015).
72. Gene Ontology, C. Gene Ontology Consortium: going forward. Nucleic acids research 43, D1049-1056 (2015).

73. Wang, H., et al. Identification of specific susceptibility loci for the early-onset colorectal cancer. Genome medicine 15, 13 (2023).
74. Rentzsch, P., Witten, D., Cooper, G.M., Shendure, J. & Kircher, M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic acids research 47, D886-d894 (2019).
75. Boyle, A.P., et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome research 22, 1790-1797 (2012).
76. Quan, C., Ping, J., Lu, H., Zhou, G. & Lu, Y. 3DSNP 2.0: update and expansion of the noncoding genomic variant annotation database. Nucleic acids research 50, D950-d955 (2022).
77. Li, J., et al. Association of Di(2-ethylhexyl) Terephthalate and Its Metabolites with Nonalcoholic Fatty Liver Disease: An Epidemiology and Toxicology Study. Environmental science & technology 58, 8182-8193 (2024).
78. 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).
79. Chong, J., et al. MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. Nucleic acids research 46, W486-w494 (2018).

Acknowledgments

This research was supported by the National Key R&D Program of China (2022YFA0806600, 2024YFC3405804), the National Natural Science Foundation of China (82388102, 82230111, 82025030, 82222063, 82404355 U22A20404), the National Science Fund for Excellent Young Scholars (NSFC-82322058), the Young Elite Scientists Sponsorship Program by CAST (2022QNRC001), Open Fund of the Key Laboratory of Environmental and Population Health, Chinese Center for Disease Control and Prevention (2026-CKL-01), the National Science Fund for Distinguished Young Scholars of Hubei Province of China (2023AFA046), the Fundamental Research Funds for the Central Universities (2042025kf0027), the National Key R&D Program of China (2024YFC3405803) and National Natural Science Foundation of China (NSFC-82373663). We also thank the staff and participants of CNHBM for their important contributions in the cohort establishment and follow-up.

Author Contributions Statement

N.H., B.L. and Y.L drafted the manuscript, with Q.J. revising and finalizing it. N.H., Q.J., and Y.Z.[1] conducted the statistical analyses and interpreted the results. Y.L. and Z.L. coordinated cohort data collection, sample testing, and omics profiling. N.H., B.L., D.Z., Y.M., H.J., and Y.F.Z. conducted experimental work. D.Z., Z.W., Y.H.Z., and M.Z. processed data and prepared the visualizations. Y.Q. T.Q., Y.L., P.S., H.S., F.Z., and S.J. participated in population investigation and specimen collection. B.Y. provided technical expertise in pollutant and metabolite quantification. Y.Z.[2], Y.L., X.M., J.T., and X.S. conceived the study design and supervised the project. All authors critically reviewed and approved the final manuscript for publication.

Competing Interests Statement

The authors declare no conflicts of interest.

Tables

Table 1. Population characteristics by PBDE exposure in 871 adults from the CNBHM cohort during 2017-2018

Variables	Total (*N* = 871)	High PBDE-exposed (*N* = 435)	Low PBDE-exposed (*N* = 436)
Age group (years)			
18-39	302 (34.67)	164 (37.70)	138 (31.65)
40-59	287 (32.95)	141 (32.41)	146 (33.49)
60-79	282 (32.38)	130 (29.89)	152 (34.86)
Sex			
Male	433 (49.71)	223 (51.26)	210 (48.17)
Female	438 (50.29)	212 (48.74)	226 (51.83)
Smoking			
Yes	637 (73.13)	322 (74.02)	315 (72.25)
No	234 (26.87)	113 (25.98)	121 (27.75)
Drinking			
Yes	515 (59.13)	263 (60.46)	252 (57.80)
No	356 (40.87)	172 (39.54)	184 (42.20)
BMI group			
Underweight	43 (4.94)	21 (4.83)	22 (5.05)
Normal weight	418 (47.99)	207 (47.59)	211 (48.39)
Overweight	274 (31.46)	139 (31.95)	135 (30.96)
Obesity	136 (15.61)	68 (15.63)	68 (15.60)
BMI (kg/m^2)	23.74 (21.7, 26.37)	23.76 (21.8, 26.37)	23.74 (21.6, 26.38)
TC (mmol/L) *	5.01 (4.4, 5.76)	5.11 (4.50, 5.83)	4.92 (4.30, 5.67)
LDL-C (mmol/L)	2.73 (2.21, 3.45)	2.74 (2.23, 3.46)	2.72 (2.20, 3.42)
HDL-C (mmol/L) *	1.39 (1.15, 1.7)	1.39 (1.15, 1.66)	1.41 (1.17, 1.73)
TG (mmol/L) *	1.36 (0.97, 1.98)	1.39 (0.99, 2.08)	1.33 (0.91, 1.90)

Note: Category variables (age, sex, smoking, drinking, BMI group) were presented as sample size (N) and percentage (%). Continuous variables (BMI, TC, LDL-C, HDL-C, TG) were presented as median (25th, 75th percentiles) (non-normal via Shapiro-Wilk test). Two-sided Pearson's *Chi-squared* tests (categorical) and Mann-Whitney U test (continuous, for non-normal data) compared high/low PBDE groups. An asterisk (*) denotes significant differences ($P < 0.05$), with exact P: TC = 0.015, HDL-C = 0.004, TG = 0.002. Abbreviations: PBDE, polybrominated diphenyl ethers; BMI, body mass index; TC, total cholesterol; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol; TG, triglyceride.

Figure legends

Fig. 1. An overview of the study design.

The CNHBM is an ongoing prospective cohort study involving 21,888 Chinese participants at the 2017-2018 baseline. Using a cluster-stratified random sampling approach, we selected 871 subpopulations for multi-omics analyses that integrated exposome, genome, metabolome, and dyslipidemia phenotypes, characterized by the total cholesterol (TC), low-density lipoprotein cholesterol (LDL-C), and high-density lipoprotein cholesterol (HDL-C), and triglycerides (TG). We began by conducting a genome-wide association study (GWAS) targeting gene-environment interactions to identify genetic variants that modulate lipid metabolism in response to environmental exposures. These variants were subsequently incorporated into polygenic risk scores (PRS) and subjected to functional annotation. We also performed a detailed analysis of metabolites linked to lipid profile changes driven by gene-environment interactions, identifying enriched metabolic pathways and critical mediating metabolites. For mechanistic understanding, we employed bioinformatic tools to predict the regulatory roles of prioritized genetic variants and generated single-base CRISPR-edited HepG2 cell models to investigate genotype-specific effects of environmental exposures on metabolite profiles and lipid homeostasis. Dual-luciferase reporter assays further validated allele-specific effects on transcriptional activity. This integrated approach offers a holistic view of how gene-environment interactions contribute to dyslipidemia pathogenesis, connecting molecular mechanisms to population-level health implications. Created in BioRender. Hu, N. (2026) https://BioRender.com/c75ezjr. By figdraw.com.

Fig. 2. Gene × PBDE interactions in lipid profiles in CNHBM. a, Manhattan plots of genome-wide gene-environment interaction analyses on the lipid phenotypes (TC, LDL-C**,** HDL-C, and TG). Each dot represents a SNP, and the top five significant loci (FDR < 0.05) are highlighted in red with annotations to their corresponding genes. b, Stratified risk analysis of the top SNPs in each phenotype. c, Trend analysis of the associations between PRS and corresponding lipid profiles. Red and blue points distinguish high and low PBDE exposure groups, error bars denote 95% confidence intervals (CI), and statistical analysis used a two-sided linear regression. Sample size: $N = 871$ for all groups. Abbreviations: FDR, false discovery rate; PRS, polygenic risk score; SNP, single nucleotide polymorphism; TC, total cholesterol; LDL-C, low-density-lipoprotein cholesterol; HDL-C, high-density-lipoprotein cholesterol; TG, triglycerides. Source data are provided as a Source Data file.

Fig. 3. Functional characterization of PBDE-interacting SNPs. a, Distribution of genomic locations of interaction loci affecting lipid phenotypes.
b, Enrichment analysis of genomic locations of interactive SNPs (vs. non-interactive SNPs) across lipid phenotypes. Bars represent OR of interactive SNP enrichment in each genomic region; error bars indicate 95% confidence intervals of OR. Non-interactive SNPs were 1:1 matched to interactive SNPs (by match number, MAF, and variant type). *P* value was calculated by two-tailed Fisher's exact test. The error bars represent the 95% confidence interval. Sample sizes per phenotype: PBDE × SNP→TC (interactive SNPs: *n* = 1108; matched non-interactive SNPs: *n* = 1108), PBDE × SNP→LDL-C (interactive SNPs: *n* = 955; matched non-interactive SNPs: *n* = 955), PBDE × SNP→HDL-C (interactive SNPs: *n* = 769; matched non-interactive SNPs: *n* = 769), PBDE × SNP→TG (interactive SNPs: *n* = 739; matched non-interactive SNPs: *n* = 739). c, Enrichment analysis of interactive SNPs in regulatory elements (vs. non-interactive SNPs; elements from ENCODE: H3K27ac, H3K4me1, DNase I, H3K9ac, H3K4me2, H3K27me3). *P* value was calculated by two-tailed Fisher's exact test. The error bars represent the 95% confidence interval. Sample sizes are consistent with panel b. d, Enrichment of interactive SNPs in transcription factor binding sites compared to non-interactive SNPs. *P* value was calculated by two-tailed Fisher's exact test. The error bars represent the 95% confidence interval. Sample sizes are consistent with panel b. e, The workflow of identification of interaction Genes. f, KEGG pathway enrichment analyses of genes identified through the framework above. g, GO pathway enrichment analyses of genes identified through the framework above. Source data are provided as a Source Data file.

Fig. 4. Metabolome profiling and mediating pathways. a, Volcano plot illustrating differential metabolites ($P_{adj} < 0.05$) linked to the lipid profiles and Gene × PBDE interaction index, calculated by multiplying PBDE levels with polygenic risk scores (PRS) for each lipid phenotype (constructed in *Fig. 2c)*. The Venn diagram below indicates the number of overlapping differential metabolites identified from paired volcano plots. Red and blue dots represent upregulated and downregulated metabolites, respectively, and Venn diagram overlaps indicate shared differential metabolites. Statistical test: Linear regression (adjusted for sex, age, BMI, smoking, and alcohol consumption); two-sided; multiple comparisons adjusted by FDR. b, Top 15 KEGG pathways enriched among the overlapping metabolites associated with both the Gene × PBDE interaction and lipid profiles. The red dashed line indicates the significance threshold $P_{adj} < 0.05$, corresponding to $-\log_{10}(P_{adj}) < 1.3$. Statistical test: Hypergeometric test; two-sided; multiple comparisons adjusted by FDR. c, Mediation network depicting the role of

metabolites in the associations between gene-environment interactions and lipid profiles, with line width representing the magnitude of the mediating effects. Source data are provided as a Source Data file.

Fig. 5. Fine mapping of top interactive SNP in hypercholesterolemia. a, Regional association plots for loci achieving genome-wide significance. Dot color corresponds to linkage disequilibrium (LD) with rs9869609 as indicated in the plot legend. b, Bar Plot of SNP Annotation Scores based on different genomic databases (Min-max standardization of RegulomeDB, 3DSNP, and CADD scores). c, Epigenetic annotation for the region surrounding SNP rs9869609 in HepG2 cell lines. Data including multiple histone moderation markers (H3K4me1, H3K4me3, H3K9ac, H3K27ac) modification peaks and TF (BHLEH40) peaks were obtained from the Cistrome database. d, eQTL analyses for rs9869609 genotypes and *SLC6A20* expression in the GTEx database. Violin plot definition: The violin shape represents the kernel density distribution of normalized *SLC6A20* expression; the box plot embedded within each violin follows: center line = median; box bounds = 25th and 75th percentiles. e, Relative reporter gene activity of the constructs containing the variant rs9869609-A and rs9869609-G alleles in HepG2 cell lines under non and BDE-47 exposure. Box plot definition: The center line represents the median; the bounds of the box represent the 25th and 75th percentiles (interquartile range, IQR); whiskers extend to the maximum and minimum values of the dataset. Replicates: n = 3 biological replicates per group, with 3 technical replicates per biological replicate. P value was calculated by two-sided Student's t-test. f, rs9869609-G allele resides within an BHLHE40-binding motif predicted by the JASPAR website. g, Effects of *BHLHE40* knockdown on luciferase activity of combined constructs containing rs9869609 under non and BDE-47 exposure. Box plot definition: The center line represents the median; the bounds of the box represent the 25th and 75th percentiles (interquartile range, IQR); whiskers extend to the maximum and minimum values of the dataset. Replicates: n = 3 biological replicates per group, with 3 technical replicates per biological replicate. P value was calculated by two-sided Student's t-test. Source data are provided as a Source Data file.

Fig. 6. Mechanism validation in BDE-47-exposed rs9869609-edited cells. a, Schematic overview of the CRISPR-Cas9 genome editing strategy used to generate rs9869609 genotypes (AA, AG, GG) in HepG2 cells. Created in BioRender. Hu, N. (2026) https://BioRender.com/c75ezjr. By figdraw.com. b, c, d, Relative *SLC6A20* mRNA expression levels and protein levels, cholesterol levels, and glycine levels

in CRISPR-edited HepG2 cells of different genotypes, assessed under control conditions and BDE-47 exposure. Box plot definition: The center line represents the median; the bounds of the box represent the 25th and 75th percentiles (interquartile range, IQR); whiskers extend to the maximum and minimum values of the data set. Replicates: *n* = 3 biological replicates per group, with 3 technical replicates per biological replicate. *P* value was calculated by two-sided Student's t-test. e, f, g, h, i, Relative *SLC6A20* mRNA expression levels, cholesterol levels, glycine levels, GCA levels and GDCA levels in HepG2 cells with *SLC6A20* knockdown, evaluated under control conditions and BDE-47 exposure. Box plot definition: The center line represents the median; the bounds of the box represent the 25th and 75th percentiles (interquartile range, IQR); whiskers extend to the maximum and minimum values of the dataset. Replicates: *n* = 3 biological replicates per group, with 3 technical replicates per biological replicate. *P* value was calculated by two-sided Student's t-test. j, Graphical summary of the interaction between BDE-47 and rs9869609 in cholesterol metabolism. Compared to the rs9869609-A allele, the risk rs9869609-G allele promotes binding of the suppressive transcription factor BHLHE40, resulting in decreased *SLC6A20* expression. This reduction impairs glycine transport, leading to cholesterol accumulation, an effect that is intensified by increased PBDE exposure. Source data are provided as a Source Data file. Created in BioRender. Hu, N. (2026) https://BioRender.com/c75ezjr.

Editor's Summary

Polybrominated diphenyl ethers (PBDEs) are linked to dyslipidemia, but why individuals respond differently is unclear. Here, the authors use a national biomonitoring cohort to reveal how specific genetic variants interact with PBDE exposure to increase risk, identifying a key mechanism that disrupts cholesterol metabolism.

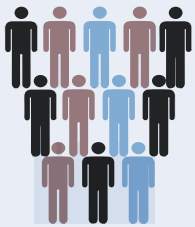
Peer Review Information: *Nature Communications* thanks Jean-Baptiste Fini and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Study population

Chinese population-based cohort (*N* = 21,888)

Omic subcohort (*N* = 1,824)

Dyslipidemia dataset (*N* = 871)

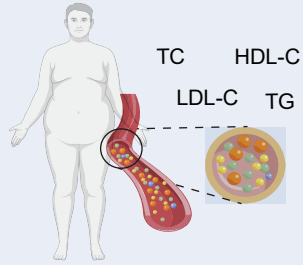


Blood sample → Exposome

DNA extract → Genomics

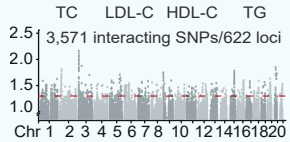
Blood sample → Metabolomics

China National Human Biomonitoring (CNHBM)



Gene×PBDE interaction

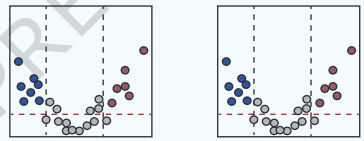
G×E interaction screening



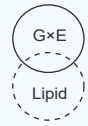
Genome-wide interacting variants

PRS construction

Metabolome profiling

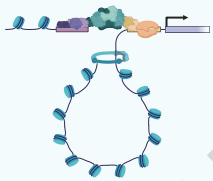


Differential metabolites



Overlapping

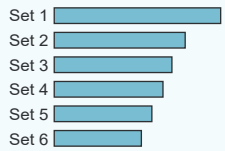
Functional annotation



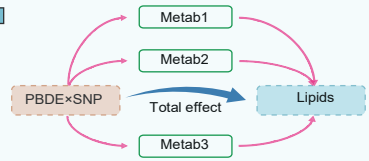
TF/Histone marker enrichment

Gene enrichment

Metabolic pathway



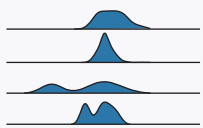
Metabolite enrichment



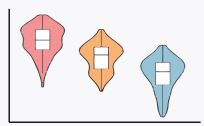
Mediation analysis

Mechanism interpretation

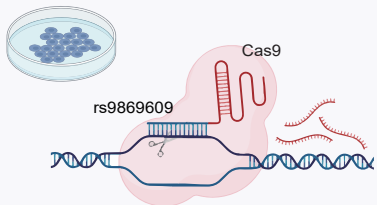
Bioinformatic prediction → Single-base editing → Experimental validation

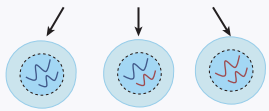


Functional annotation



eQTL analysis

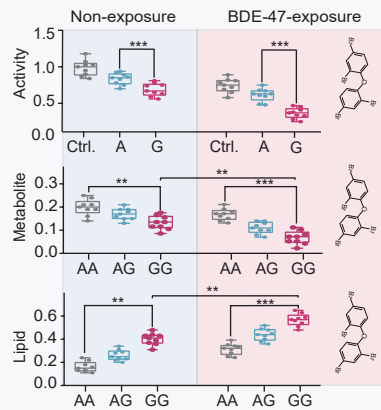


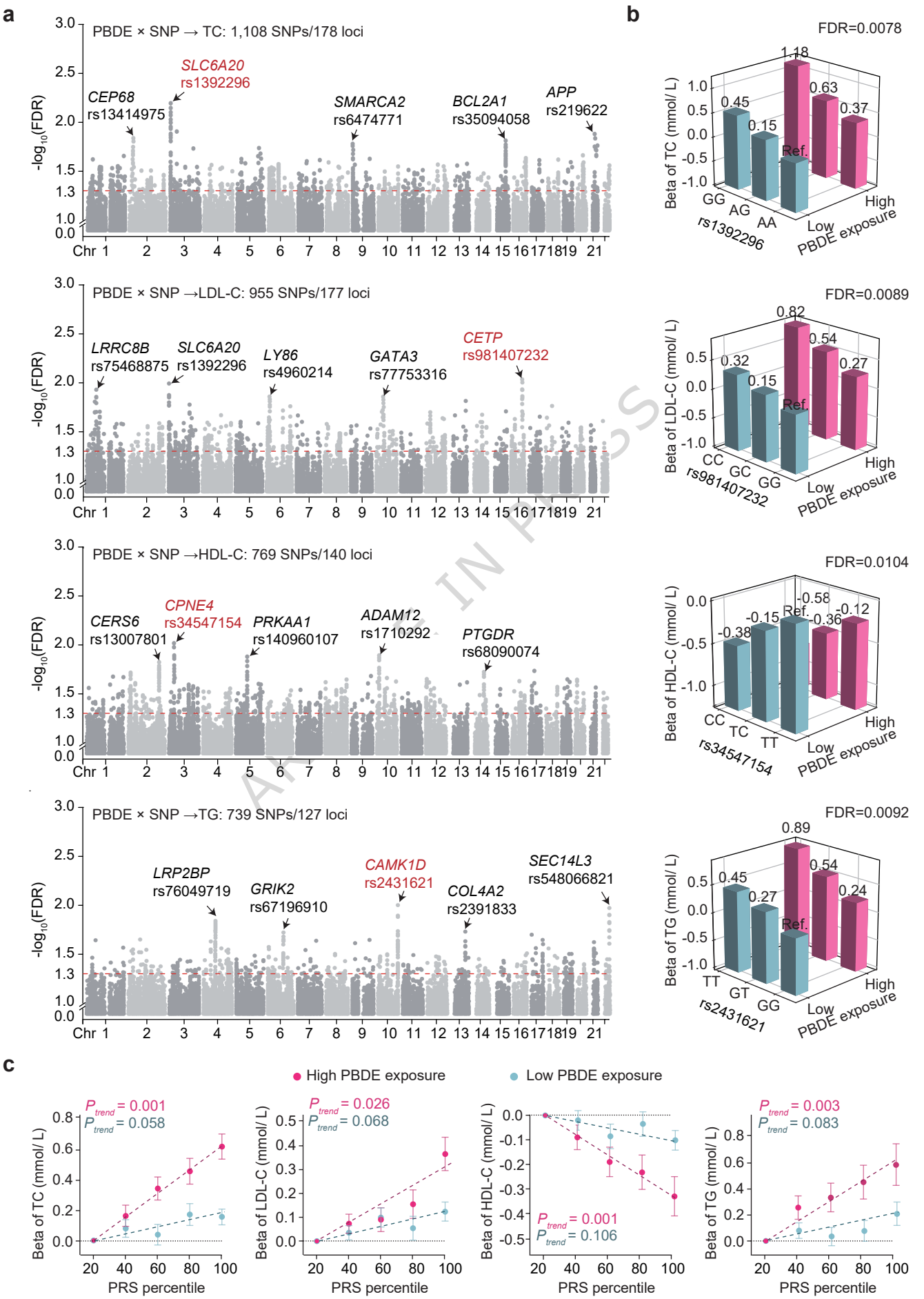


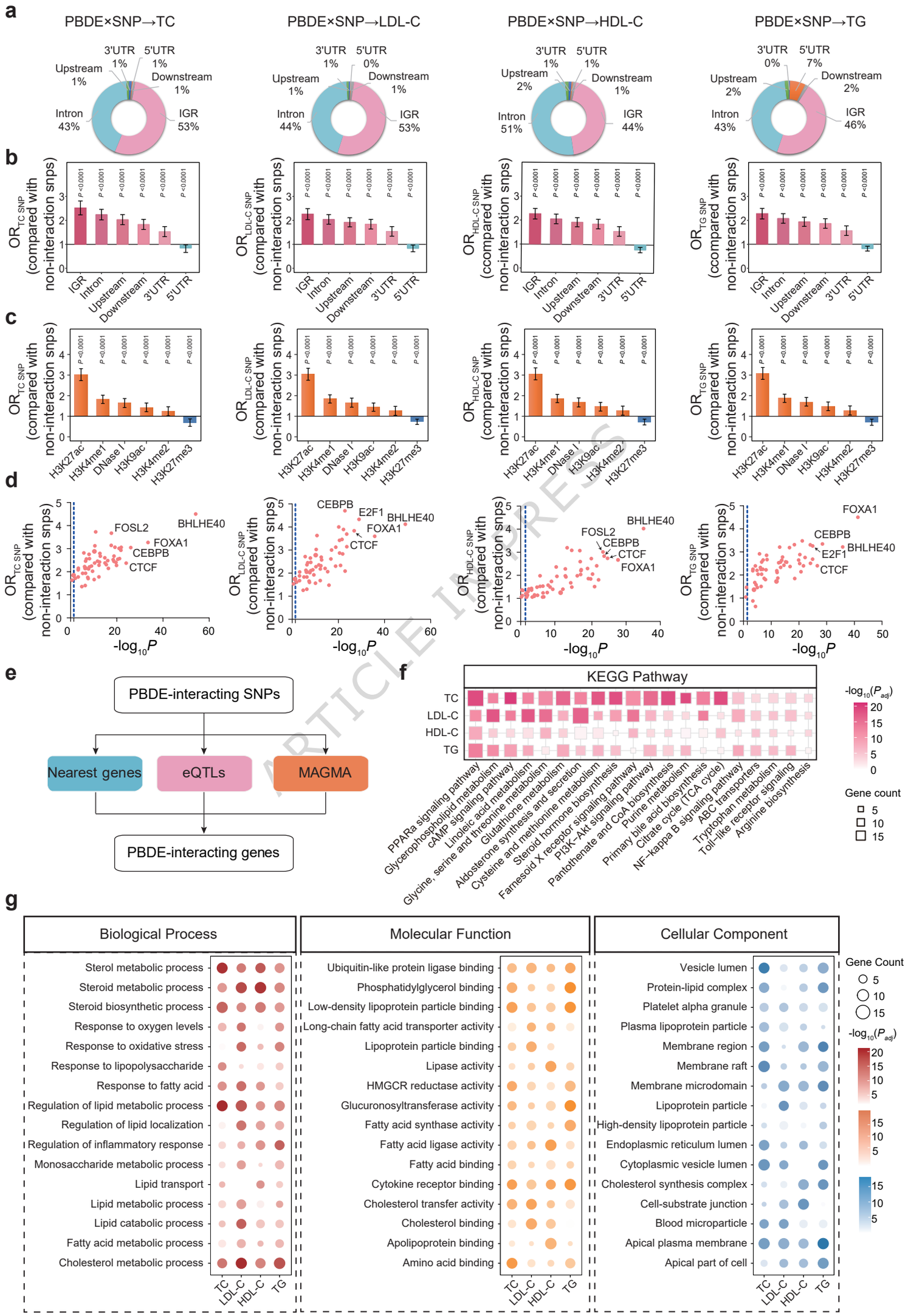
Dual luciferase reporter gene

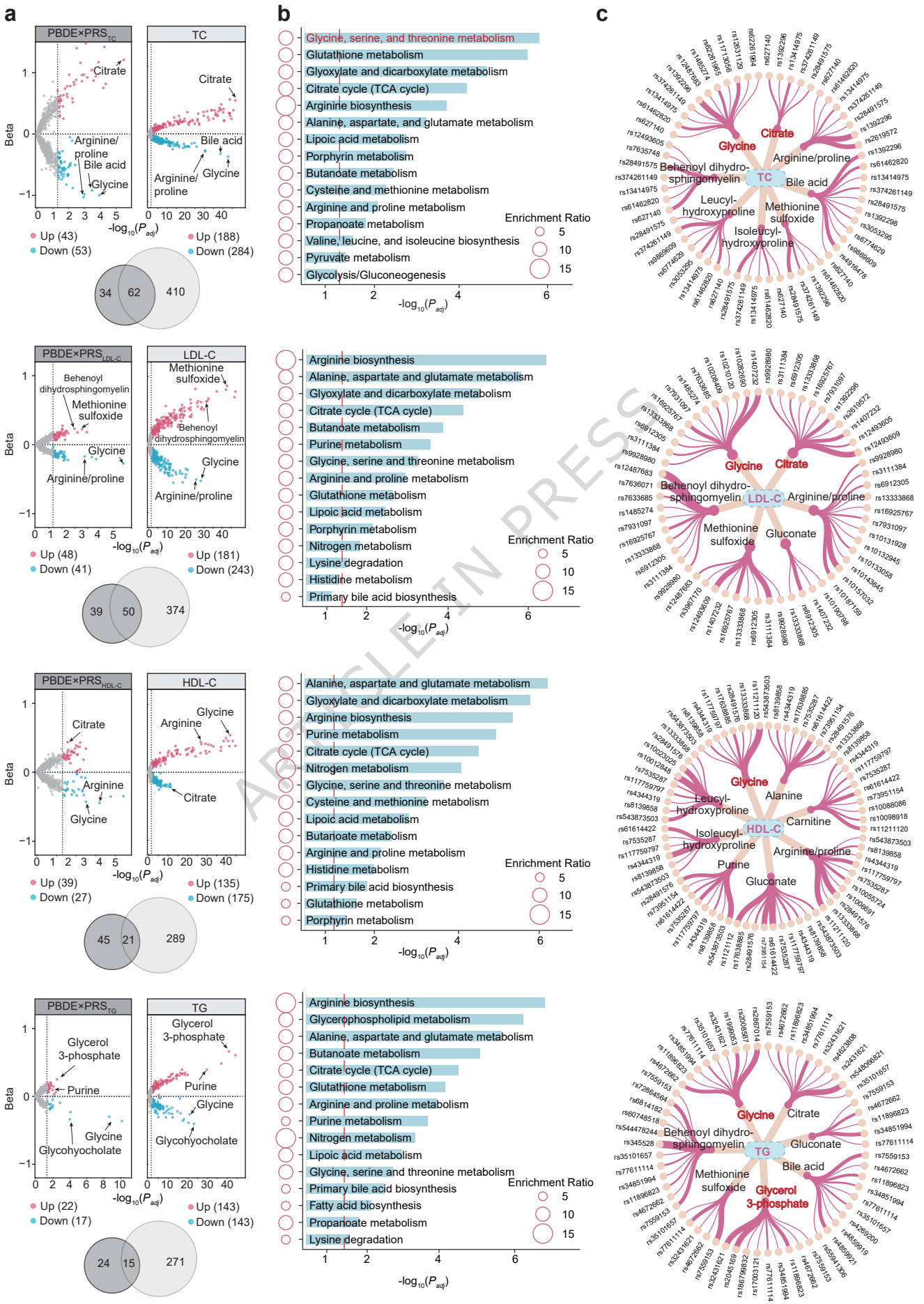
Metabolite level

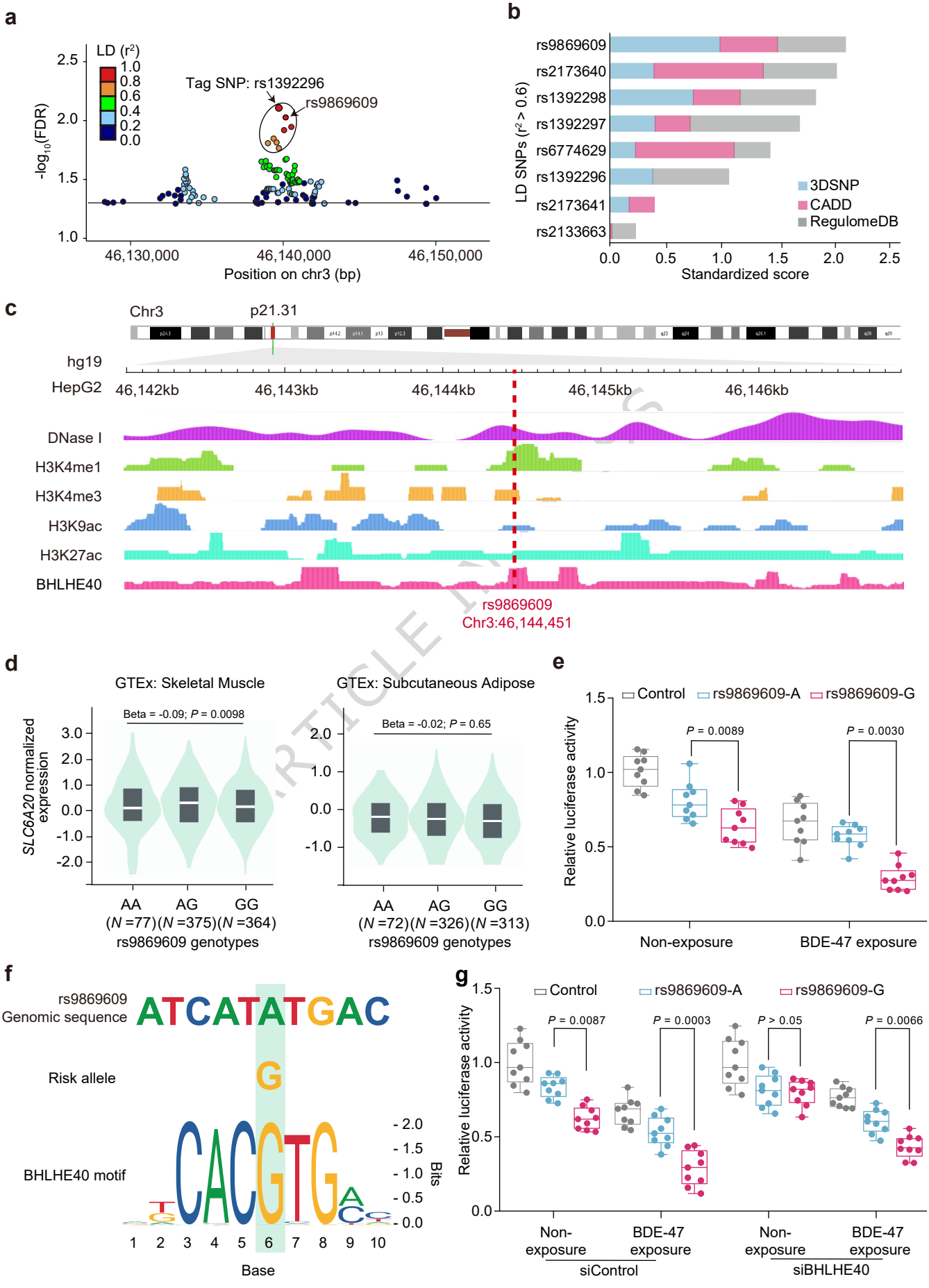
Lipid level




a
PBDE × SNP → TC: 1,108 SNPs/178 loci
SLC6A20
rs1392296
CEP68
rs13414975
SMARCA2
rs6474771
BCL2A1
rs35094058
APP
rs219622
PBDE × SNP →LDL-C: 955 SNPs/177 loci
LRRC8B
rs75468875
SLC6A20
rs1392296
LY86
rs4960214
GATA3
rs77753316
CETP
rs981407232
PBDE × SNP →HDL-C: 769 SNPs/140 loci
CERS6
rs13007801
CPNE4
rs34547154
PRKAA1
rs140960107
ADAM12
rs1710292
PTGDR
rs68090074
PBDE × SNP →TG: 739 SNPs/127 loci
LRP2BP
rs76049719
GRIK2
rs67196910
CAMK1D
rs2431621
COL4A2
rs2391833
SEC14L3
rs548066821
-log10(FDR)
Chr 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 21
b
FDR=0.0078
Beta of TC (mmol/ L)
GG AG AA
rs1392296
FDR=0.0089
Beta of LDL-C (mmol/ L)
CC GC GG
rs981407232
FDR=0.0104
Beta of HDL-C (mmol/ L)
CC TC TT
rs34547154
FDR=0.0092
Beta of TG (mmol/ L)
TT GT GG
rs2431621
Low High
PBDE exposure
c
High PBDE exposure
Low PBDE exposure
Ptrend = 0.001
Ptrend = 0.058
Beta of TC (mmol/ L)
Ptrend = 0.026
Ptrend = 0.068
Beta of LDL-C (mmol/ L)
Ptrend = 0.001
Ptrend = 0.106
Beta of HDL-C (mmol/ L)
Ptrend = 0.003
Ptrend = 0.083
Beta of TG (mmol/ L)
PRS percentile


a
PBDE×SNP→TC
PBDE×SNP→LDL-C
PBDE×SNP→HDL-C
PBDE×SNP→TG
3'UTR 1%
5'UTR 1%
Upstream 1%
Downstream 1%
Intron 43%
IGR 53%
3'UTR 1%
5'UTR 0%
Upstream 1%
Downstream 1%
Intron 44%
IGR 53%
3'UTR 1%
5'UTR 1%
Upstream 2%
Downstream 1%
Intron 51%
IGR 44%
3'UTR 0%
5'UTR 7%
Upstream 2%
Downstream 2%
Intron 43%
IGR 46%
b
OR TC SNP (compared with non-interaction snps)
OR LDL-C SNP (compared with non-interaction snps)
OR HDL-C SNP (ccompared with non-interaction snps)
OR TG SNP (compared with non-interaction snps)
P <0.0001
IGR
Intron
Upstream
Downstream
3'UTR
5'UTR
c
H3K27ac
H3K4me1
DNase I
H3K9ac
H3K4me2
H3K27me3
d
FOSL2
BHLHE40
FOXA1
CEBPB
CTCF
E2F1
-log10P
e
PBDE-interacting SNPs
Nearest genes
eQTLs
MAGMA
PBDE-interacting genes
f
KEGG Pathway
TC
LDL-C
HDL-C
TG
PPARa signaling pathway
Glycerophospholipid metabolism
cAMP signaling pathway
Linoleic acid metabolism
Glycine, serine and threonine metabolism
Glutathione metabolism
Cysteine and methionine metabolism
Aldosterone synthesis and secretion
Steroid hormone biosynthesis
Farnesoid X receptor signaling pathway
PI3K−Akt signaling pathway
Pantothenate and CoA biosynthesis
Purine metabolism
Primary bile acid biosynthesis
Citrate cycle (TCA cycle)
NF−kappa B signaling pathway
ABC transporters
Tryptophan metabolism
Toll−like receptor signaling
Arginine biosynthesis
-log10(Padj)
Gene count
g
Biological Process
Sterol metabolic process
Steroid metabolic process
Steroid biosynthetic process
Response to oxygen levels
Response to oxidative stress
Response to lipopolysaccharide
Response to fatty acid
Regulation of lipid metabolic process
Regulation of lipid localization
Regulation of inflammatory response
Monosaccharide metabolic process
Lipid transport
Lipid metabolic process
Lipid catabolic process
Fatty acid metabolic process
Cholesterol metabolic process
Molecular Function
Ubiquitin-like protein ligase binding
Phosphatidylglycerol binding
Low-density lipoprotein particle binding
Long-chain fatty acid transporter activity
Lipoprotein particle binding
Lipase activity
HMGCR reductase activity
Glucuronosyltransferase activity
Fatty acid synthase activity
Fatty acid ligase activity
Fatty acid binding
Cytokine receptor binding
Cholesterol transfer activity
Cholesterol binding
Apolipoprotein binding
Amino acid binding
Cellular Component
Vesicle lumen
Protein-lipid complex
Platelet alpha granule
Plasma lipoprotein particle
Membrane region
Membrane raft
Membrane microdomain
Lipoprotein particle
High-density lipoprotein particle
Endoplasmic reticulum lumen
Cytoplasmic vesicle lumen
Cholesterol synthesis complex
Cell-substrate junction
Blood microparticle
Apical plasma membrane
Apical part of cell
Gene Count
-log10(Padj)




a
LD (r²)
1.0
0.8
0.6
0.4
0.2
0.0
Tag SNP: rs1392296
rs9869609
-log10(FDR)
2.5
2.0
1.5
1.0
46,130,000
46,140,000
46,150,000
Position on chr3 (bp)
b
rs9869609
rs2173640
rs1392298
rs1392297
rs6774629
rs1392296
rs2173641
rs2133663
LD SNPs (r² > 0.6)
3DSNP
CADD
RegulomeDB
0.0
0.5
1.0
1.5
2.0
2.5
Standardized score
c
Chr3
p21.31
hg19
HepG2
46,142kb
46,143kb
46,144kb
46,145kb
46,146kb
DNase I
H3K4me1
H3K4me3
H3K9ac
H3K27ac
BHLHE40
rs9869609
Chr3:46,144,451
d
GTEx: Skeletal Muscle
Beta = -0.09; P = 0.0098
SLC6A20 normalized expression
3.0
2.0
1.0
0.0
-1.0
-2.0
AA
AG
GG
(N =77)(N =375)(N =364)
rs9869609 genotypes
GTEx: Subcutaneous Adipose
Beta = -0.02; P = 0.65
2.0
1.0
0.0
-1.0
AA
AG
GG
(N =72)(N =326)(N =313)
rs9869609 genotypes
e
Control
rs9869609-A
rs9869609-G
Relative luciferase activity
1.5
1.0
0.5
0.0
P = 0.0089
P = 0.0030
Non-exposure
BDE-47 exposure
f
rs9869609
Genomic sequence
ATCATATGAC
Risk allele
G
BHLHE40 motif
CACGTG
Bits
2.0
1.5
1.0
0.5
0.0
1
2
3
4
5
6
7
8
9
10
Base
g
Control
rs9869609-A
rs9869609-G
Relative luciferase activity
1.5
1.0
0.5
0.0
P = 0.0087
P = 0.0003
P > 0.05
P = 0.0066
Non-exposure
BDE-47 exposure
Non-exposure
BDE-47 exposure
siControl
siBHLHE40