Environment International 190 (2024) 108845

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Full length article

Common genetic variants associated with urinary phthalate levels in children: A genome-wide study

Mariona Bustamante [a,b,c,*,1], Laura Balagué-Dobón [a,1], Zsanett Buko [d], Amrit Kaur Sakhi [e], Maribel Casas [a,b,c], Lea Maitre [a,b,c], Sandra Andrusaityte [f], Regina Grazuleviciene [f], Kristine B. Gützkow [e], Anne-Lise Brantsæter [e], Barbara Heude [g], Claire Philippat [h], Leda Chatzi [i], Marina Vafeiadi [j], Tiffany C. Yang [k], John Wright [k], Amy Hough [k], Carlos Ruiz-Arenas [m], Ramil N. Nurtdinov [n], Geòrgia Escaramís [l,c], Juan R. González [a,b,c], Cathrine Thomsen [e], Martine Vrijheid [a,b,c]

[a] Environment and Health Over the Lifecourse, ISGlobal, Barcelona, Spain
[b] Universitat Pompeu Fabra (UPF), Barcelona, Spain
[c] CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain
[d] Department of Oncological Science, Huntsman Cancer Institute, Salt Lake City, United States
[e] Division of Climate and Environmental Health, Norwegian Institute of Public Health, Oslo, Norway
[f] Department of Environmental Science, Vytautas Magnus University, Kaunas, Lithuania
[g] Université Paris Cité and Université Sorbonne Paris Nord, Inserm, INRAE, Center for Research in Epidemiology and StatisticS (CRESS), F-75004, Paris, France
[h] University Grenoble Alpes, Inserm U-1209, CNRS-UMR-5309, Environmental Epidemiology Applied to Reproduction and Respiratory Health Team, Institute for Advanced Biosciences, 38000, Grenoble, France
[i] Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States
[j] Department of Social Medicine, Faculty of Medicine, University of Crete, Heraklion, Greece
[k] Bradford Institute for Health Research, Bradford Teaching Hospitals NHS Foundation Trust, Bradford, UK
[l] Departament de Biomedicina, Institut de Neurociències, Universitat de Barcelona (UB), Barcelona, Spain
[m] Computational Biology Program, CIMA University of Navarra, idiSNA, Pamplona 31008, Spain
[n] Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Barcelona 08003, Catalonia, Spain

ARTICLE INFO

Handling Editor: Adrian Covaci

Keywords:
Phthalate
Genome-wide association study (GWAS)
Genetic variant
Single nucleotide polymorphism (SNP)
Copy number variant (CNV)

ABSTRACT

Introduction: Phthalates, or dieters of phthalic acid, are a ubiquitous type of plasticizer used in a variety of common consumer and industrial products. They act as endocrine disruptors and are associated with increased risk for several diseases. Once in the body, phthalates are metabolized through partially known mechanisms, involving phase I and phase II enzymes.

Objective: In this study we aimed to identify common single nucleotide polymorphisms (SNPs) and copy number variants (CNVs) associated with the metabolism of phthalate compounds in children through genome-wide association studies (GWAS).

Abbreviations: **ABC**, ATP-binding cassette; **BAF**, B allele frequency; **BBP**, butylbenzyl phthalate; **chr**, chromosome; **CNV**, copy number variant; **CYP**, Cytochrome P450 monooxygenase; **DBzP**, dibenzyl phthalate; **DEP**, diethyl phthalate ; **DEHP**, bis(2-ethylhexyl) phthalate; **DiBP**, diisobutyl phthalate; **DINCH**, 2-cyclohexane dicarboxylic acid, diisononyl ester; **DnBP**, di-n-butyl-phthalate; **EDCs**, endocrine disrupting chemicals; **eQTL**, expression quantitative trait locus; **GSA**, Infinium Global Screening Array; **GWAS**, genome-wide association study; **HELIX**, Human Early Life Exposome project; **HMW**, high-molecular-weight; **HRC**, haplotype reference consortium; **LD**, linkage disequilibrium; **LMW**, low-molecular-weight; **LRR**, Log R ratio; **MBzP**, mono benzyl phthalate; **MECPP**, mono-2-ethyl-5-carboxypentyl phthalate; **MECPP_MEHP**, ratio between MECPP and MEHP; **MECPP_MEHHP**, ratio between MECPP and MEHHP; **MEHHP**, mono-2-ethyl-5-hydroxyhexyl phthalate; **MEHHP_MEHP**, ratio between MEHHP and MEHP; **MEHP**, mono-2-ethylhexyl phthalate; **MEOHP**, mono-2-ethyl-5-oxohexyl phthalate; **MEOHP_MEHP**, ratio between MEOHP and MEHP; **MEOHP_MEHHP**, ratio between MEOHP and MEHHP; **MEP**, monoethyl phthalate; **MiBP**, mono-*iso*-butyl phthalate; **MnBP**, mono-n-butyl phthalate; **oh-MiNP**, mono–4–methyl–7–hydroxyoctyl phthalate; **oxo-MiNP**, mono–4–methyl–7–oxooctyl phthalate; **oxo-MiNP_oh-MiNP**, ratio between oxo-MiNP and oh-MiNP; **PVC**, polyvinyl chloride; **PRS**, polygenic risk score; **QQ**, quantile–quantile; **SCL**, Solute carrier; **SNP**, single nucleotide polymorphism; **SULT**, Sulfotransferase; **UGT**, UDP-glucuronosyltransferase.

* Corresponding author at: ISGlobal, Dr. Aiguader 88 08003, Barcelona, Spain.
E-mail address: mariona.bustamante@isglobal.org (M. Bustamante).

[1] Equal contribution.

https://doi.org/10.1016/j.envint.2024.108845
Received 16 January 2024; Received in revised form 14 June 2024; Accepted 20 June 2024
Available online 23 June 2024

Metabolism
Toxicity
Phase I and II enzymes
Renal excretion

Methods: The study used data from 1,044 children with European ancestry from the Human Early Life Exposome (HELIX) cohort. Ten phthalate metabolites were assessed in a two-void pooled urine collected at the mean age of 8 years. Six ratios between secondary and primary phthalate metabolites were calculated. Genome-wide genotyping was done with the Infinium Global Screening Array (GSA) and imputation with the Haplotype Reference Consortium (HRC) panel. PennCNV was used to estimate copy number variants (CNVs) and CNVRanger to identify consensus regions. GWAS of SNPs and CNVs were conducted using PLINK and SNPassoc, respectively. Subsequently, functional annotation of suggestive SNPs (p-value < 1E-05) was done with the FUMA web-tool.
Results: We identified four genome-wide significant (p-value < 5E-08) loci at chromosome (chr) 3 (*FECHP1* for oxo-MiNP_oh-MiNP ratio), chr6 (*SLC17A1* for MECPP_MEHHP ratio), chr9 (*RAPGEF1* for MBzP), and chr10 (*CYP2C9* for MECPP_MEHHP ratio). Moreover, 115 additional loci were found at suggestive significance (p-value < 1E-05). Two CNVs located at chr11 (*MRGPRX1* for oh-MiNP and *SLC35F2* for MEP) were also identified. Functional annotation pointed to genes involved in phase I and phase II detoxification, molecular transfer across membranes, and renal excretion.
Conclusion: Through genome-wide screenings we identified known and novel loci implicated in phthalate metabolism in children. Genes annotated to these loci participate in detoxification, transmembrane transfer, and renal excretion.

1. Introduction

Phthalates, or dieters of phthalic acid, are a ubiquitous type of plasticizer used in a variety of common consumer and industrial products, whose exposure is widespread and ever-changing due to our constantly evolving environment and habits (Praveena et al., 2018; Wang et al., 2019). Phthalates are known to be endocrine disrupting chemicals (EDCs), and epidemiological research has suggested that exposure to phthalates is associated with increased risk for several diseases including infertility, allergy, obesity, diabetes, and behavioral problems (Praveena et al., 2018; Wang et al., 2019). Since children are especially vulnerable to contaminants as they are still developing, there is a great concern over the potential for phthalate exposure to disturb normal growth and development (Braun, 2017; Casale and Rice, 2023; Lee et al., 2022, 2023).

Phthalates can be separated into low-molecular-weight (LMW, 3–6 carbon atoms) and high-molecular-weight (HMW, 7–13 carbon atoms) compounds (Praveena et al., 2018; Wang et al., 2019). LMW phthalates are used as solvents and are usually found in medications and personal care items such as deodorants, lotions, and shampoos. HMW phthalates are used in the manufacturing of flexible plastics for purpose of vinyl flooring, adhesives, medical devices, and food packaging. The main intake of LMW phthalates is through skin and inhalation, while HMW phthalates are incorporated into the body through ingestion (Kim and Park, 2014). In the European Union, benzyl butyl phthalate (BBzP), bis (2-ethylhexyl) phthalate (DEHP), dibutyl phthalate (DBP) and bis(2-methoxyethyl) phthalate (DMEP) are now banned from use in cosmetics (Regulation (EC) No 1223/2009, 2009) and DEHP, DBP and BBzP are regulated in material intended to come into contact with food (Comission Directive EC/2007/19, 2007).

After exposure, phthalates are rapidly metabolized and excreted mainly in urine as a result of phase I and phase II enzymes (Domínguez-Romero and Scheringer, 2019; Frederiksen et al., 2007; Praveena et al., 2018). First, diester phthalates are hydrolyzed to monoester (primary metabolites) by phase I esterase and lipase enzymes (Bhattacharyya et al., 2022). Subsequently, LMW phthalates are primarily excreted in urine and feces as monoesters, without further metabolism. In contrast, HMW phthalates need further metabolism through hydroxylation and oxidation by phase I enzymes, such as *Cytochrome P450 (CYP) monooxygenases,* thus producing a number of oxidative metabolites (secondary metabolites). These oxidated metabolites can be directly excreted in the urine, or alternatively, can be additionally metabolized through phase II enzymes. Phase II detoxification consists of conjugation reactions where a compound is added to the parental congener to generate hydrophilic conjugates which can then be easily excreted in the urine through enzymes such as UDP-glucuronosyltransferases (UGTs), Sulfotransferases (SULTs), or Glutathione S-Transferases (GSTs). Phthalates and their metabolites can be measured in diverse biological specimens, such as blood, urine, breast milk, and feces. However, urinary phthalate metabolites are the most frequently used biomarkers to track exposure to phthalates (Sakhi et al., 2017; Wang et al., 2019). Phthalates' half-life is less than 24 h (Domínguez-Romero et al., 2023).

Phase I and phase II enzymes are highly polymorphic in humans, with slow and fast metabolic phenotypes determined by single nucleotide polymorphisms (SNPs) and copy number variants (CNVs) (Pinto and Eileen Dolan, 2011). This differential detoxification capacity might modify phthalate effects in the body and is rarely considered in epidemiological studies. In vitro studies demonstrated the importance for phthalate metabolism of genetic polymorphisms in *CYP* genes (Choi et al., 2012). In humans, urinary phthalate metabolite levels were reported to be associated with genetic variants in *GSTP1* and *SOD2* in children (Wang and Karmaus, 2017), with variants in *CYP2C9, CYP2C19* and *UGT1A7* in young adults (Stajnko et al., 2022), and with variants in *CYP2C9* and *Solute Carriers* (*SLC17A1*) in adults (Lu et al., 2024). The two first relied on a priori knowledge of candidate genes and SNPs, potentially overlooking additional loci important for phthalate metabolism, whereas only the latter study conducted a genome-wide association study (GWAS). GWAS, which allow the interrogation of millions of common genetic variants, both SNPs and CNVs, can help to overcome this limitation and reveal new phthalate detoxification pathways (Visscher et al., 2017).

Here, we aimed to identify genetic variants related to phthalate metabolism in children. For this, we analyzed the association of genome-wide SNPs and CNVs with levels of ten phthalate metabolites and six phthalate ratios measured in two-void pooled urine samples of 1,044 European ancestry children from the Human Early Life Exposome (HELIX) project, and performed functional annotation of the identified variants.

2. Materials and methods

2.1. Study population

HELIX project is an ongoing population-based birth cohort study in six birth cohorts from different European countries: (1) EDEN – Étude des Déterminants pré et postnatals du développement et de la santé de l'enfant, France, recruitment period: 2003–2006; (2) Rhea – the Rhea Mother-Child Study in Crete, Greece, 2007–2008; (3) KANC – Kaunus Cohort, Lithuania, 2007–2009; (4) MoBa – the Norwegian Mother, Father and Child Cohort Study, Norway, 1999–2009; (5) INMA – INfancia y Medio Ambiente, Spain, 2004–2006; (6) BiB – Born in Bradford, United Kingdom (UK), 2007–2010 (Maitre et al., 2018). The HELIX project aims to implement novel exposure assessment and biomarker methods to measure the early-life exposure to multiple environmental factors and associate these with omics biomarkers and child health outcomes, thus characterizing the "early-life exposome". The entire

study population is 31,472 mother–child pairs. The subcohort study includes 1,304 children with exposure, phenotypes, and molecular data measured at age of 6–12 years. The current study selected 1,044 children of European ancestry with genome-wide genetic and phthalate metabolite levels available from the HELIX subcohort (Appendix A – Supplementary Fig. 1). Child's ancestry was predicted from genome-wide genetic data. The following variables were used to describe the population: child's sex (male/female), child's age (in years), child's obesity in 3 categories (normal, overweight, obese) based on the World Health Organization (WHO) body mass index (BMI) classification, and self-reported maternal education (primary, secondary and university or higher).

2.2. Ethics approval and consent to participate

All studies were approved by the national research ethics committees and informed consent to participate was obtained for all participants.

2.3. Phthalate biomarkers measurement

At the 8-years visit, children were examined for clinical and neurological measurements and provided biological samples. Each HELIX participant collected two urine voids at home: one at bed time the day before the examination, and the other in the morning of the day of the examination. These urine voids were stored at –20C in the freezers of the volunteers and transported to the visit in a box with ice packs to maintain temperature. In the laboratory, with the aim of improving exposure assessment, a pool of seven ml of each of the two spot urine samples was prepared, and seven aliquots of 1.75 ml of the urine pool were stored at −80C.

Metabolites for 6 different phthalates were measured in an aliquot of pooled urine (Fig. 1). In particular, the following determinations were done using high performance liquid chromatography coupled to mass spectrometry (Haug et al., 2018): four LMW primary phthalate metabolites [monoethyl phthalate (MEP), mono-n-butyl phthalate (MnBP), mono-isobutyl phthalate (MiBP), mono benzyl phthalate (MBzP)], one HMW primary phthalate metabolite [mono-2-ethylhexyl phthalate (MEHP)], and five HMW secondary metabolites [mono-2-ethyl-5-hydroxyphenyl phthalate (MEHHP), mono-2-ethyl-5-oxohexyl phthalate (MEOHP), mono-2-ethyl-5-oxyhexyl phthalate (MECPP), mono–4–methyl–7–hydroxyoctyl phthalate (oh-MiNP), and mono–4–methyl–7–oxooctyl phthalate (oxo-MiNP)]. Both external control samples, in-house control samples and blank samples were analyzed in each batch of about fifty samples. The limits of detection (LOD) of the method ranged from 0.067 to 0.67 ng/mL. All the samples had values > LOD. MEHP, MEHHP and MEOHP had 34, 3 and 1 missing values, due to underperformance of sample quantification (either there was an interference in the chromatogram in the compound peak position, or the ratio between the quantifier and qualifier ions in the mass spectrometry were outside the acceptable range, and thus values were not reliable).

To account for urine dilution, a second aliquot of the pooled urine was analyzed for creatinine concentration, and phthalate metabolite concentrations were divided by urinary creatinine levels (Haug et al., 2018). All creatinine-adjusted concentrations were log2 transformed to obtain normal distributions, as the original distributions were right-skewed. Finally, as a proxy of enzymatic activity, we calculated six ratios between product and substrate, in our case, between secondary and primary phthalate metabolites or between two secondary phthalate metabolites, as follows: MECPP_MEHP, MEHHP_MEHP, MEOHP_MEHP, MECPP_MEHHP, MEOHP_MEHHP and oxo-MiNP_oh-MiNP. The ratios were calculated using untransformed phthalate values unadjusted for creatinine, and then they were log2 transformed. Urine dilution was not taken into account as the ratios were computed using two compounds within the same urine sample.

2.4. Genetic data

DNA was obtained from buffy coats collected in EDTA tubes at the mean age of 8 years. DNA was extracted by cohort using a Chemagen kit in batched of 12 samples. Two techniques were used to determine the DNA concentration and quality: NanoDrop 1000 UV–Vis Spectrophotometer (ThermoScientific), and Quant-iT™ PicoGreen® dsDNA Assay Kit (Life Technologies).

Infinium Global Screening Array (GSA) MD version 1 (Illumina) was used for genome-wide genotyping at the Human Genomics Facility (HuGe-F), Erasmus MC (https://www.glimdna.org). Genotype calling and annotation were done using the GenTrain2.8 algorithm based on a custom cluster file implemented in the GenomeStudio software and the GSAMD-24v1-0_20011747-A4 manifest. SNP coordinates were reported on human reference GRCh37 and on the source strand.

PLINK program was used for the quality control of the genetic data (Purcell et al., 2007). Briefly, from the 1397 samples initially genotyped, we filtered out those that had a call rate < 97 % (n exclusions = 43), had sex inconsistencies (n = 8), presented an heterozygosity > 3 standard deviations (n = 0), and were related (sharing > 18.5 % of alleles) (n = 10) or duplicated (n = 19) (Appendix A – Supplementary Fig. 1). For ancestry prediction from GWAS data, we used the Peddy program (Pedersen and Quinlan, 2017). Then genetic ancestry was contrasted with self-reported ethnicity and discordant samples were excluded (n = 12). Genetic variants were filtered out if they had a call rate < 95 %, if they were in the non-canonical pseudo-autosomal region (PAR), if they had a minor allele frequency (MAF) < 1 %, and if they were not in Hardy-Weinberg equilibrium (HWE) at a p-value < 1E −06. PLINK was also used to compute the first 20 principal components (PCs) from the GWAS data of European ancestry children using the linkage disequilibrium (LD) clumping option. Ten PCs explained up to 19.2 % of the genetic variance and the five first ones correlated with cohort.

Genome-wide imputation was performed with the Imputation Michigan server using the Haplotype Reference Consortium (HRC) panel, Version r1.1 2016. Before the imputation PLINK data was converted into VCF format and the variants were aligned with the reference genome. Eagle v2.4 was used for the phasing of the haplotypes and minimac4 for the imputation. In the end, we retrieved 40,405,505 variants after imputation. Then we filtered out the genetic variants according to the following parameters: imputation accuracy (R2) < 0.9, MAF < 5 %, and HWE p-value < 1E-06. After this the post-imputation dataset had 4,614,947 variants.

2.5. GWAS of SNPs

GWAS of SNPs were conducted with the PLINK program by applying a linear regression model for each phthalate metabolite and each SNP from chromosome 1 to 22, adjusting for sex, age and ten GWAS PCs as a proxy of ancestry (Purcell et al., 2007). We run a total of 16 GWAS (five primary metabolites, five secondary metabolites, and six ratios). Genome-wide statistical significance was established at a p-value < 5E-08 as proposed elsewhere (Xu et al., 2014), and suggestive statistical significance at an arbitrary p-value of < 1E-5 (The Severe Covid-19 GWAS Group, 2020). Multiple-testing correction did not account for the number of metabolites and ratios tested. The lambda genomic inflation factor was estimated as the median of the resulting chi-squared test statistics divided by the expected median of the chi-squared distribution. Quantile-quantile (QQ) plots were conducted with the qqman package and the Miami plot was conducted with the ggplot2 package in R (D. Turner, 2018; Wickham, 2016). The lead SNP is defined as the SNP with the lowest p-value in a locus, and a locus is defined as a region containing SNPs associated with phthalates at p-value<1E-05, and whose consecutive distance is <1 Mb.

In order to filter out associations between the SNPs and creatinine levels, we run sensitivity models adjusting for log2 transformed creatinine levels, instead of controlling (dividing) the phthalate

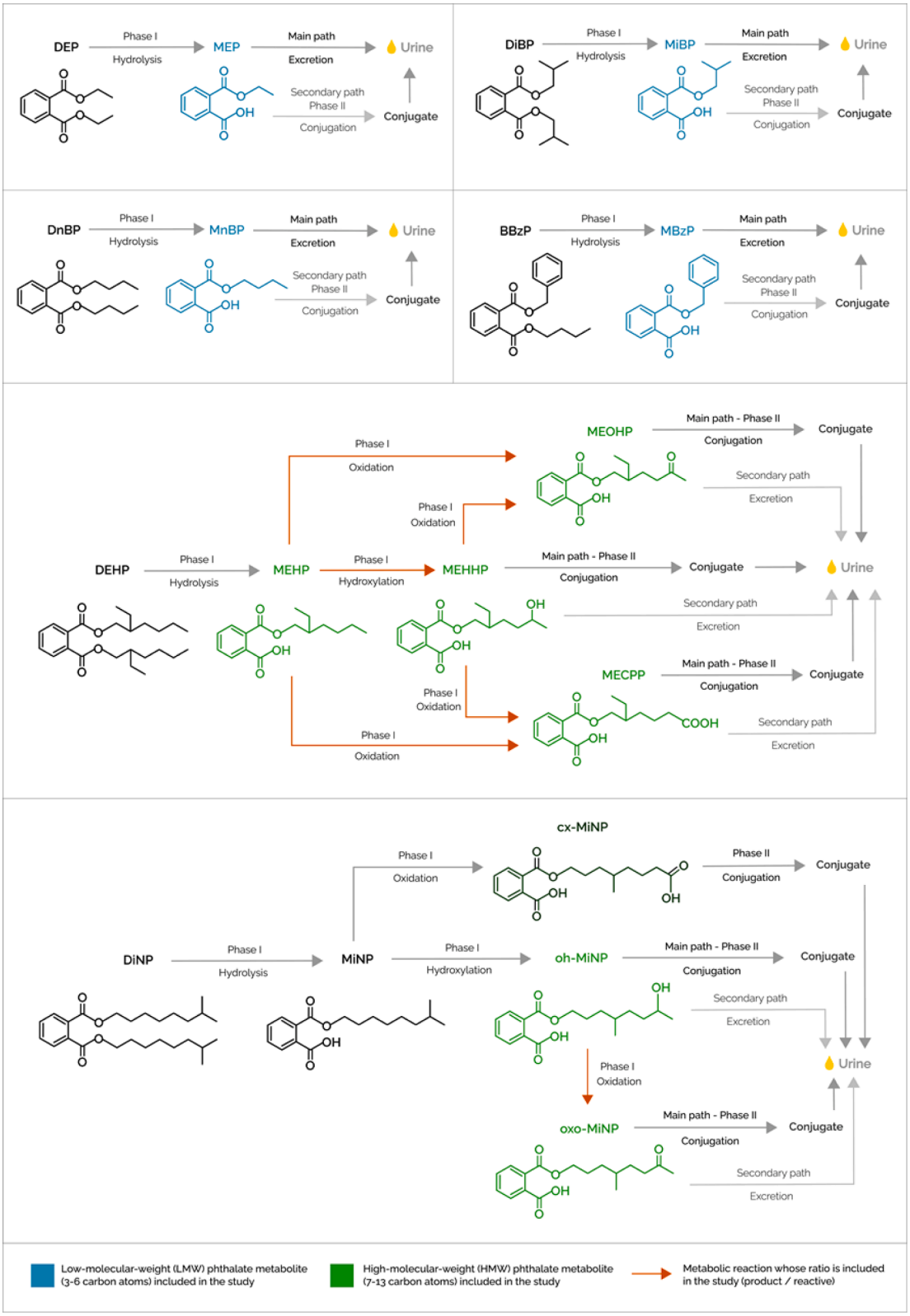

Fig. 1. Summary of the phthalate metabolites and ratios included in the study. Diagram showing the phthalate metabolites included in the study, along with their precursor phthalates and intermediate metabolites. Low-molecular-weight metabolites (LWM) in the study are coloured in blue and high-molecular weight metabolites (HMW) in the study are coloured in green. Orange arrows represent the metabolic reactions whose ratio is included in the study (product/reactant). **BBzP**: Benzyl butyl phthalate; **MBzP**: Mono benzyl phthalate; **DEP**: Diethyl phthalate; **MEP**: Monoethyl phthalate; **DiBP**: Di-*iso*-butyl phthalate; **MiBP**: Mono-*iso*-butyl phthalate; **DnBP**: Di-n-butyl phthalate; **MnBP**: Mono-n-butyl phthalate; **DEHP**: Di-2-ethylhexyl phthalate; **MEHP**: Mono-2-ethylhexyl phthalate; **MECPP**: Mono-2-ethyl-5-carboxypentyl phthalate; **MEHHP**: Mono-2-ethyl-5-hydroxyhexyl phthalate; **MEOHP**: Mono-2-ethyl-5-oxohexyl phthalate; **DINP**: Di-*iso*-nonyl phthalate; **MiNP**: Mono-*iso*-nonyl phthalate; **cx-MiNP**: Mono-carboxy isononyl phthalate; **oh-MiNP**: Mono-4-methyl-7-hydroxyoctyl phthalate; **oxo-MiNP**: Mono-4-methyl-7-oxooctyl phthalate. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

concentrations by creatinine. Finally, we also run a GWAS for log2 transformed creatinine levels (g per L of urine).

2.6. GWAS of CNVs

Copy number variant (CNV) calls were obtained with PennCNV using B allele frequency (BAF) and Log R ratio (LRR) signal files from Illumina BeadStudio (Wang et al., 2007). A total of 22 samples from the initial 1,044 had to be removed due to poor quality in BAF or LRR values. After that, CNV calls overlapping with the centromeres were filtered out. After the quality control, consensus CNV regions were obtained with the CNVRanger R package following the reciprocal overlap procedure, that merges calls with sufficient mutual overlap, in our case 50 % (Conrad et al., 2010; Da Silva et al., 2020). From the resulting consensus regions, only common CNVs with an alternative minor allele frequency (either gain or loss) > 3 % were kept for the analysis. Consensus CNV regions were classified into: (i) only loss, when we only detected deletions (CN = 0 and CN = 1); (ii) only gain, when we only detected duplications (CN = 3; note that more than 3 copies were never detected); and (iii) loss and gain, when we detected both deletions and duplications (CN = 0, CN = 1 and CN = 3).

Association analyses between each of the resulting regions and each of the 16 metabolites and ratios were performed with SNPassoc R package (González et al., 2007). Distinct association models were applied to each region based on the prior CNV classification. For only-loss regions we tested loss-codominant (0 vs 1 vs 2), loss-dominant (0 + 1 vs 2), loss-recessive (0 vs 1 + 2), loss-overdominant (0 + 2 vs 1) and loss-log-additive models; and for only-gain regions we tested gain-codominant (2 vs 3) model. Finally, for loss-and-gain regions we tested all loss-specific and gain-specific models above, as well as an additive lineal model taking into account loss and gain. All analyses were adjusted for the same covariates as the GWAS of SNPs (sex, age and ten GWAS PCs). Multiple-testing correction was conducted by dividing the nominal p-value of 0.05 by the number of detected consensus CNVs (0.05/36 = 0.0014). As was done with SNPs, sensitivity models adjusting for log2 transformed creatinine levels were investigated in order to filter out associations between the CNVs and creatinine levels. Finally, we also run a model for log2 transformed creatinine levels (g/mL of urine). Annotation of genes in each CNV region was done with the biomaRt R package (Durinck et al., 2009).

2.7. Functional annotation of genetic variants, candidate genes, and enrichment for tissues and pathways

For the annotation, prioritization, and biological interpretation of the genetic variants identified in the GWAS, we used the Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA) tool, v1.6.1 (Watanabe et al., 2017). The SNP2GENE module computes LD structure, annotates SNPs using information from different databases, and prioritizes candidate genes. FUMA was run using default parameters except that the p-value threshold was set to 1E-05 instead of 1E-08, an r^2 of 0.6, and using European population from 1 KG/Phase3. For the SNP annotation we used: ANNOVAR, CADD, RegulomeDB, expression quantitative trait locus (eQTL) mapping based on Genotype-Tissue Expression (GTEx) v8 and other databases, comparison with SNPs reported in the GWAS catalog and gene mapping through chromatin interaction maps. Then, with the GENE2FUNC module from FUMA prioritized genes from SNP2GENE module were used to identify shared molecular functions using different databases: Gene Ontology (GO) terms, Kyoto Encyclopedia of Genes and Genomes (KEGG), and Reactome, among others. GENE2FUNC was also used to assess enrichment for 54 tissues included in the GTEXv8 dataset.

For gene mapping through chromatin interaction we also used the EPIraction tool, (https://epiraction.crg.es, accessed 29 June 2023) (Nurtdinov and Guigó, n.d.). Briefly, the tool contains a catalogue of tissue specific gene-enhancer interactions identified by applying linear models between gene expression and activities of nearby enhancers across 1,529 samples from 77 different tissues. In this study, we annotated SNPs to enhancers and then linked enhancers to genes in two candidate tissues for phthalate detoxification: liver and kidney.

3. Results

3.1. Description of the population

The description of the 1,044 HELIX participants is shown in Table 1. All participants were of European ancestry and from six different countries: Greece, Norway, Spain, Lithuania, France, and the UK. Forty-six percent of the participants were female, and the mean age of phthalate measurements was 8.08 years (inter-quartile range (IQR) = 6.49 – 8.89). Around half of the children were born from highly educated mothers. Urinary phthalate levels (μg per g of creatinine) were in the range or higher than levels reported in children and adolescents from 12 European countries for the period 2014–2021 (Vogel et al., 2023). After log2 transformation, levels tended to be normally distributed and the correlation among them ranged from 0.18 to 0.98 (Appendix A – Supplementary Fig. 2). Highest correlation coefficients (r > 0.79) were observed among MECPP, MEHHP, MEOHP and MEHP, which are all primary or secondary metabolites of the same parental

Table 1
Descriptive of the study population.

Variable	N (%), Median (IQR)	N missing
Cohort		
BiB (UK)	90 (8.62 %)	0
EDEN (France)	135 (12.93 %)	0
KAUNAS (Lithuania)	196 (18.77 %)	0
MoBa (Norway)	239 (22.89 %)	0
RHEA (Greece)	186 (17.82 %)	0
INMA (Spain)	198 (18.97 %)	0
Child's sex		
Females	473 (45.31 %)	0
Males	571 (54.69 %)	0
Child's age (years)	8.08 (6.49 – 8.89)	0
Child's obesity		
Normal	830 (79.50 %)	0
Overweight	155 (14.85 %)	0
Obese	59 (5.66 %)	0
Maternal education		
Primary	115 (11.02 %)	0
Secondary	353 (33.81 %)	0
University or higher	550 (52.68 %)	0
Urinary creatinine levels (g/mL of urine)	0.99 (0.79–1.23)	0
Urinary LMW phthalate levels (μg per g of creatinine)		
Mono benzyl phthalate (MBzP)	5.11 (3.18–8.53)	0
Monoethyl phthalate (MEP)	29.83 (15.54–70.03)	0
Mono-*iso*-butyl phthalate (MiBP)	39.82 (25.18–70.35)	0
Mono-n-butyl phthalate (MnBP)	23.24 (14.97–38.14)	0
Urinary HMW phthalate levels (μg per g of creatinine)		
Mono-2-ethylhexyl phthalate (MEHP)	2.81 (1.65–4.99)	34
Mono-2-ethyl-5-carboxypentyl phthalate (MECPP) *	35.37 (21.77–58.91)	0
Mono-2-ethyl-5-hydroxyhexyl phthalate (MEHHP) *	20.09 (12.11–33.02)	3
Mono-2-ethyl-5-oxohexyl phthalate (MEOHP) *	12.69 (7.47–20.86)	1
Mono-4-methyl-7-hydroxyoctyl phthalate (oh-MiNP)	5.20 (3.29–8.79)	0
Mono-4-methyl-7-oxooctyl phthalate (oxo-MiNP) **	2.84 (1.86–4.91)	0

For continuous variables median and IQR are shown, while for categorical variables the sample size and its percentage % are reported. LMW: low molecular weight. HMW: high molecular weight. N = 1,044.

* Metabolites of MEHP.

** Metabolite of oh-MiNP.

compound, DEHP. The correlation of oh-MiNP and oxo-MiNP was also high (r = 0.73).

3.2. GWAS of SNPs

To identify SNPs associated with the detoxification metabolism of phthalates in children, we run linear regressions for each of the ten phthalate metabolites and the six metabolic ratios adjusting for sex, age, and 10 GWAS PCs.

Genomic inflation factors (lambdas) ranged from 0.989 to 1.021 (Appendix A – Supplementary Fig. 3). Four loci reached genome-wide significance (p-value < 5E-08) as shown in the Miami plot (Fig. 2). They included a locus at chr3 associated with the oxo-MiNP_oh-MiNP ratio (lead SNP rs80064213_C, effect = 0.331, p-value = 4.15E-08, near *FECHP1* gene), one at chr6 associated with the MECPP_MEHHP ratio (rs1177442_A, effect = -0.134, p-value = 4.58E-17, *SLC17A1*), one at chr9 associated with MBzP (rs138702233_T, effect = 0.394, p-value = 4.29E-08, *RAPGEF1*), and another at chr10 associated with the MECPP_MEHHP, MEOHP_MEHHP and MECPP_MEHP ratios (rs74494115_T, effect = -0.309, p-value = 1.45E-63, *CYP2C9*) (Table 2). The locus zoom plots of these four loci are shown in Fig. 3.

Moreover, 2,140 SNPs representing 115 unique loci were identified at suggestive significance (p-value < 1E-05) for at least one phthalate metabolite or ratio (Appendix B – Supplementary Table B0 and B1). The number of loci with suggestive statistical significance per phthalate metabolite or ratio was between five and 16. Thirty-two of the 119 suggestive or genome-wide significant unique loci were shared between more than one phthalate compound or ratio (Appendix A – Supplementary Fig. 4). The genome-wide significant loci at chr3 (*FECHP1)*, chr6 (*SLC17A1)* and chr10 (*CYP2C9)*, also had suggestive associations with other phthalate compounds or ratios (Appendix B – Supplementary Table B1).

Sensitivity analyses adjusting for creatinine instead of using it to divide phthalate metabolite levels gave similar results with minor differences (Appendix B – Supplementary Table B2). Three of the four genome-wide significant loci had still a p-value < 5E-08, and the other hit (chr9 loci associated with MBzP) was marginally significant (p-value = 9.95E-07). A plot comparing p-values and effect sizes between the two models for all the suggestive SNPs can be seen at Appendix A – Supplementary Figures 5A and B. Moreover, the GWAS of creatinine levels did not retrieve any genome-wide significant SNP. Only two of the creatinine suggestive SNPs were also suggestively associated with MEOHP and oh-MiNP (Appendix B Supplementary Table B19).

3.3. GWAS of CNVs

Using the PennCNV tool, we identified a median of seven (interquartile range (IQR): 6–9) CNV regions per child with a median length of 43 (IQR: 29–64) kb. Only two children did not present any CNV event. After applying a reciprocal overlap of 50 %, 2,916 distinct CNV regions were detected, including 1,785 that were unique, meaning that they were detected in only one child. Thirty-six of the 2,916 CNVs were identified in > 3 % of the children and thus labeled as consensus CNV regions and tested in relation to phthalate levels and ratios (Appendix C – Supplementary Table C0).

In 12 of the 36 consensus regions, we detected single or double deletions (loss CNVs, CN = 0 and CN = 1), in seven of them we detected single duplications (gain CNVs, CN = 3) and in 17 of them we detected single and double deletions as well as single duplications (loss and gain CNVs, CN = 0, CN = 1 and CN = 3). We did not detect double duplications (CN = 4) in any CNV consensus region. Different models were tested in each type of region (see 2.6.) and the model with the highest significance was reported for each region. The results of the association of these common CNVs with phthalate levels and ratios can be found in Appendix C – Supplementary Tables C3-18.

Fifty-one CNVs presented nominal significance and two of them passed multiple-testing correction (Fig. 2, Appendix C – Supplementary Table C1). First, a single-copy gain in CNV region 11A (chr11:18.95 Mb-18.96 Mb) was associated with higher oh-MiNP levels (11A_gain, effect = 1.231, p-value = 1.77E-06, *MRGPRX1* gene), and second, a single-copy gain in CNV region 11B (chr11:107.65 Mb-107.67 Mb) was associated with higher MEP levels (11B_gain, effect = 0.792, p-value = 1.36E-03, *SLC35F2*). The locus zoom plots of these two CNVs are shown in Fig. 4.

In the sensitivity analysis adjusting for creatinine, 88 % of the nominally significant associations were reproduced (45 of 51). The associations between CNV 11A and oh-MiNP and between CNV 11B and MEP were very similar (Appendix C – Supplementary Table C2),

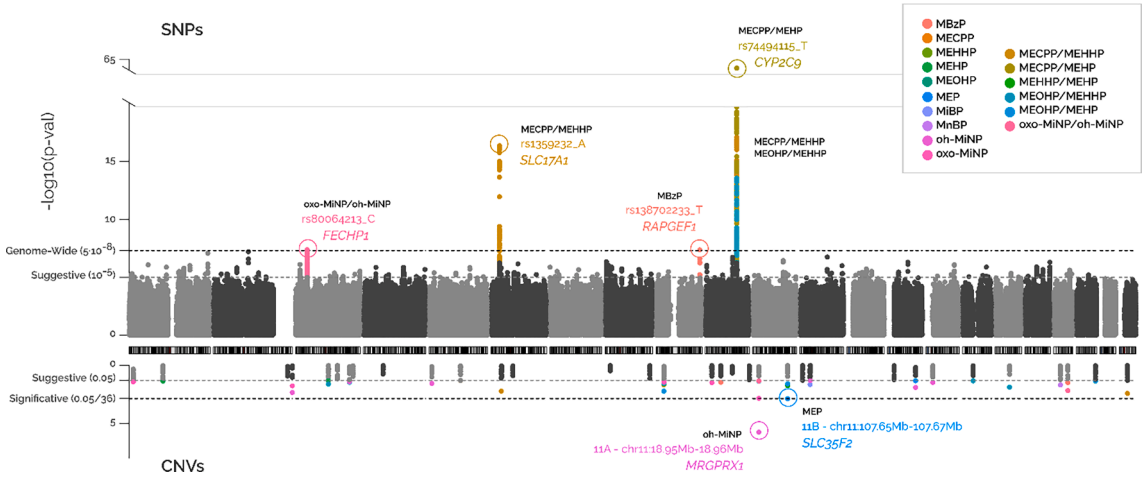

Fig. 2. Miami plot of the association between common SNPs (top panel) and CNVs (bottom panel) vs phthalate levels or ratios. Each dot represents the association of a SNP or CNV. The x-axis indicates the position of the SNP or CNV in the genome. The y-axis shows the statistical significance, the –log10(p-value). Colours indicate the phthalate compound or ratio the SNPs or CNVs are associated with. The SNPs/CNVs passing multiple-testing correction (5E-08 for SNPs and 1.39E-03 for CNVs) are annotated to the closest gene. **MBzP**: Mono benzyl phthalate; **MEP**: Monoethyl phthalate; **MiBP**: Mono-*iso*-butyl phthalate; **MnBP**: Mono-n-butyl phthalate; **MEHP**: Mono-2-ethylhexyl phthalate; **MECPP**: Mono-2-ethyl-5-carboxypentyl phthalate; **MEHHP**: Mono-2-ethyl-5-hydroxyhexyl phthalate; **MEOHP**: Mono-2-ethyl-5-oxohexyl phthalate; **oh-MiNP**: Mono-4-methyl-7-hydroxyoctyl phthalate; **oxo-MiNP**: Mono-4-methyl-7-oxooctyl phthalate; **MECPP_MEHP**: ratio between MECPP and MEHP; **MEHHP_MEHP**: ratio between MEHHP and MEHP; **MEOHP_MEHP**: ratio between MEOHP and MEHP; **MECPP_MEHHP**: ratio between MECPP and MEHHP; **MEOHP_MEHHP**: ratio between MEOHP and MEHHP; **oxo-MiNP_oh-MiNP**: ratio between oxo-MiNP and oh-MiNP.

Table 2
Genome-wide loci associated with phthalate levels and/or ratios, ordered by p-value.

Locus	Phthalate	Chr	Pos (hg19)	Lead SNP	Non-effect allele	Effect allele	Effect allele frequency	Effect	Standard error	P-value	Nearest gene to lead SNP	Relative gene position	Exonic variants (gene, CADD, amino-acid change)[a]	eQTLs (liver, kidney)	Gene-enhancer interactions (liver, kidney)	Other traits (GWAS catalog)	Candidate gene function	Evidences for phthalates
1[b]	MECPP_MEHHP	10	96,719,845	rs74494115	G	T	0.21	−0.3087	0.0171	1.45E-63	*CYP2C9*	Intronic	rs1799853_T (*CYP2C9*, 22.9, Arg > Cys); rs1057910_C (*CYP2C9*, 16.9, Ile > Leu); rs11572080_A (*CYP2C8*, 19.06, Arg > Lys)	Not in liver or kidney	Several genes in the region including *CYP2C* genes	Serum/plasma metabolite levels and drug metabolization, etc.	*Cytochrome P450 monooxygenases:* Detoxification, *CYP2C8* (xenobiotics); *CYP2C9* (xenobiotics); *CYP2C18* (unknown); and *CYP2C19* (xenobiotics)	SNPs at *CY2PC9* vs. MEHP metabolites (Lu et al. 2024; Stajnko et al. 2022)
	MECPP_MEHP	10	996,658,495	rs7910609	G	A	0.07	−0.6140	0.0649	2.10E-20	*RP11-400G3.3*	Intergenic						
	MEOHP_MEHHP	10	96,719,845	rs74494115	G	T	0.21	−0.0780	0.0101	2.81E-14	*CYP2C9*	Intronic						
2[c]	MECPP_MEHHP	6	25,809,069	rs1177442 (6:25809069:A:G)	G	A	0.46	−0.1335	0.0156	4.58E-17	*SCL17A1*	Intronic	rs11754288_A (*SLC17A4*, 13.27, Ala > Thr); 6:25813150:G:A = rs1165196_C (*SLC17A1*, 11.45, Ile > Thr)	*SLC* family genes	*SLC* family genes	Urinary metabolite levels (urate), estimated glomerular filtration rate (creatinine), etc.	*Solute Carrier Family 17 Members:* Transport of molecules across membranes, *SLC17A4* (phospate excretion); *SLC17A1* and *SLC17A3* (urate excretion); *SLC17A2* (sialic acid transport)	SNPs at *SLC17A1* vs. MECPP (Lu et al. 2024)
3[d]	oxo-MiNP_oh-MiNP	3	34,882,408	rs80064213	T	C	0.11	0.3307	0.0598	4.15E-08	*FECHP1*	Intergenic	–	–	–	Body mass index, breast cancer	*FECHP1 (Ferrochelatase Pseudogene 1):* Pseudogene // *STAC (SH3 And Cysteine Rich Domain):* Predicted to enable transmembrane transporter binding activity	–
4	MBzP	9	134,523,638	rs138702233	C	T	0.07	0.3938	0.0713	4.29E-08	*RAPGEF1*	Intronic	–	Not in liver or kidney	Several genes in the region including *RAPGEP1*	–	*RAPGEF1 (Rap Guanine Nucleotide Exchange Factor 1):* Human guanine nucleotide exchange factor.	–

A locus is defined as a region containing SNPs associated with phthalates at p-value < 1E-05, and whose consecutive distance is < 1 Mb.
The lead SNPs is the SNP with the lowest p-value in the locus.
MBzP: Mono benzyl phthalate; MEP: Monoethyl phthalate; MiBP: Mono-*iso*-butyl phthalate; MnBP: Mono-n-butyl phthalate; MEHP: Mono-2-ethylhexyl phthalate; MECPP: Mono-2-ethyl-5-carboxypentyl phthalate; MEHHP: Mono-2-ethyl-5-hydroxyhexyl phthalate; MEOHP: Mono-2-ethyl-5-oxohexyl phthalate; oh-MiNP: Mono-4-methyl-7-hydroxyoctyl phthalate; oxo-MiNP: Mono-4-methyl-7-oxooctyl phthalate; MECPP_MEHP: ratio between MECPP and MEHP; MEHHP_MEHP: ratio between MEHHP and MEHP; MEOHP_MEHP: ratio between MEOHP and MEHP; MECPP_MEHHP: ratio between MECPP and MEHHP; MEOHP_MEHHP: ratio between MEOHP and MEHHP; oxo-MiNP_oh-MiNP: ratio between oxo-MiNP and oh-MiNP.

[a] CADD: Combined Annotation Dependent Depletion; Alleles are ordered according to the linkage disequilibrium with the lead SNP shown in the table.

[b] Suggestive associations (p-value < 1E-05) with MECPP (−); MEOHP_MEHP (−).

[c] Suggestive associations (p-value < 1E-05) with MECPP (−); MEOHP_MEHP (+); MEHHP_MEHP (+).

[d] Suggestive associations (p-value < 1E-05) with MEOHP_MEHHP (+).

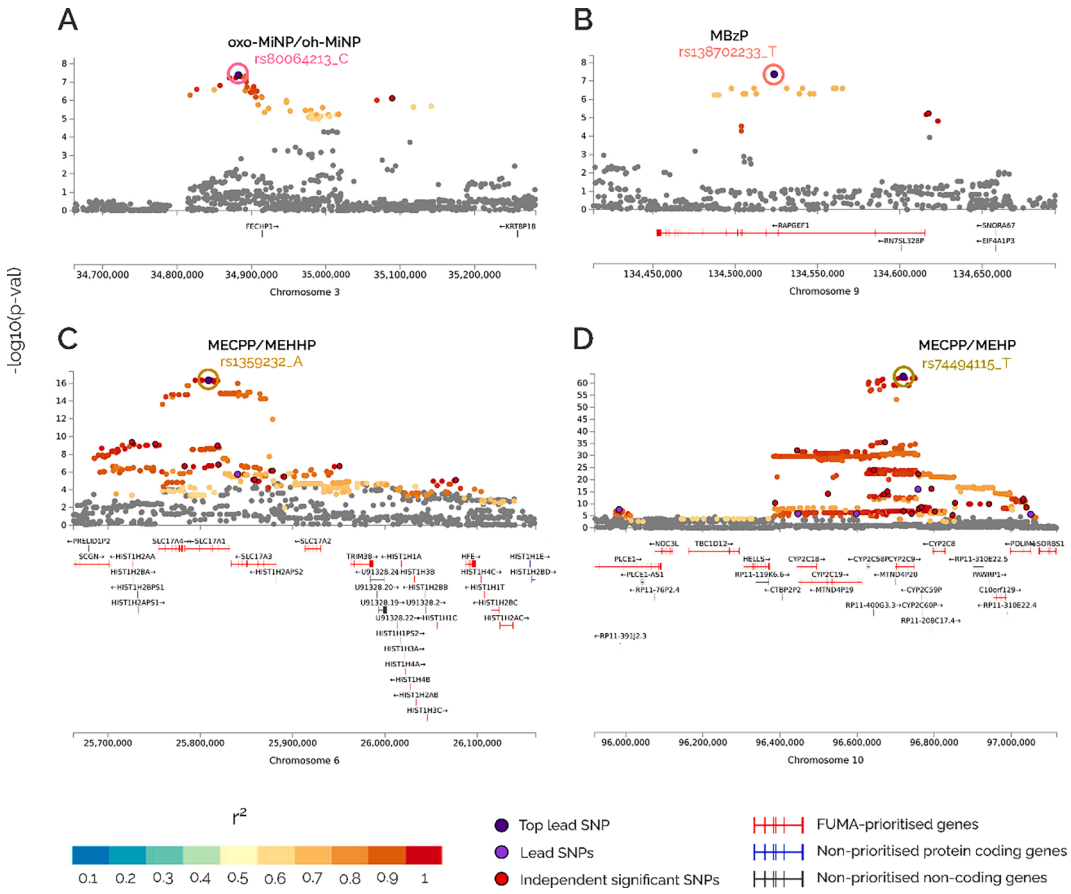

Fig. 3. Locus zoom of the SNP-associated regions identified after multiple-testing correction Each dot represents the association of a SNP. The lead SNP is coloured in dark purple, other lead SNPs (genotyped SNPs with $r^2 < 0.1$ with the lead SNP and p-value <1E-05) in light purple, and independent significant SNPs (genotyped SNPs with r^2 <0.6 with the lead SNP and p-value <1E-05) in red. Other non-genotyped SNPs are coloured according to their linkage disequilibrium (r^2) with the lead SNP. The y-axis shows the statistical significance, the –log10(p-value). The x-axis indicates the position of the SNP in the genome. FUMA prioritized coding genes and FUMA non-prioritized coding and non-coding genes are shown. **MBzP**: Mono benzyl phthalate; **MEHP**: Mono-2-ethylhexyl phthalate; **MECPP**: Mono-2-ethyl-5-carboxypentyl phthalate; **MEHHP**: Mono-2-ethyl-5-hydroxyhexyl phthalate; **oh-MiNP**: Mono-4-methyl-7-hydroxyoctyl phthalate; **oxo-MiNP**: Mono-4-methyl-7-oxooctyl phthalate; **MECPP_MEHP**: ratio between MECPP and MEHP; **MECPP_MEHHP**: ratio between MECPP and MEHHP; **oxo-MiNP_oh-MiNP**: ratio between oxo-MiNP and oh-MiNP. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

however the latter did not reach the multiple-testing threshold (chr11:18.95 Mb-18.96Mb_gain, effect = 1.120, p-value = 2.51E-06; chr11:107.65 Mb-107.67Mb_gain, effect = 0.695, p-value = 1.48E-03). A plot comparing p-values and effect sizes between the two models for all the suggestive CNVs can be seen at Appendix A – Supplementary Figures 5C and D. Results for creatinine GWAS did not show any genome-wide significant association, with only four CNV regions, not including 11A and 11B, showing a p-value < 0.05 (Appendix C – Supplementary Table C19).

3.4. *Functional annotation of genetic variants, candidate genes, and enrichment for tissues and pathways*

Next, SNPs with genome-wide or suggestive significance were explored with the FUMA tool providing information on SNP annotation (Appendix D), identification of eQTLs (Appendix E), and comparison with SNPs in the GWAS catalog (Appendix F). The list of candidate genes in each locus based on FUMA results and gene-enhancer interactions from the EPIraction tool can be found in Appendix G.

The annotation of the four genome-wide significant loci is shown in Table 2. The lead SNP at chr10 (rs74494115_T), associated with lower levels of MECPP_MEHP, MECPP_MEHHP, and MEOHP_MEHHP ratios, was situated within an intron of the *CYP2C9* gene, part of a cluster containing several *CYP* genes. This SNP was in LD with two non-synonymous variants at *CYP2C9* (rs1799853_T, Arg > Cys, $r^2 = 0.82$; and rs1057910_C, Ile > Leu, $r^2 = 0.32$), one at *CYP2C8* (rs11572080_T, Arg > Lys, $r^2 = 0.99$), and several eQTLs for *CYP2C* genes, among others ($r^2 > 0.6$). The lead SNP at chr6 (rs1177442_A), associated with lower MECPP_MEHHP levels, was located within an intron of the *SCL17A1* gene. It exhibited LD with several eQTLs for *SLC17A* genes in liver and/or kidney ($r^2 > 0.6$), and with two non-synonymous variants at *SLC17A4* (rs11754288_A, Ala > Thr, $r^2 = 0.93$) and at *SLC17A1* (rs1165196_G, Ile > Thr, $r^2 = 1$). The lead SNP at chr3 (rs80064213_C), associated with higher oxo-MiNP_oh-MiNP levels, was located near *FECHP1*, a *Ferrochelatase* pseudogene. Finally, the lead SNP at chr9 (rs138702233_T), associated with higher MBzP levels, was located in an intron of the *RAPGEF1* gene, which encodes a guanine nucleotide exchange factor.

Appendix H provides further information about these loci, as well as eight additional loci selected from those exhibiting suggestive statistical significance. These additional loci are annotated to phase I (*CYP2A, B, C* clusters) and phase II (*SULT1B1*, *SULT1E1*, *SULT1C3*, *UGT2A1*, *UGT2A2* and *UGT2B11*) detoxification genes, glutathione related genes (*GSR*, *GPX5* and *GPX6*), and membrane transport genes (*SLC8A2*, *SLC35G1*, *ABCC3*, and *ATP9B*).

After the FUMA annotation, we took genome-wide and suggestive prioritized genes to run gene-set enrichment analyses (Appendix I).

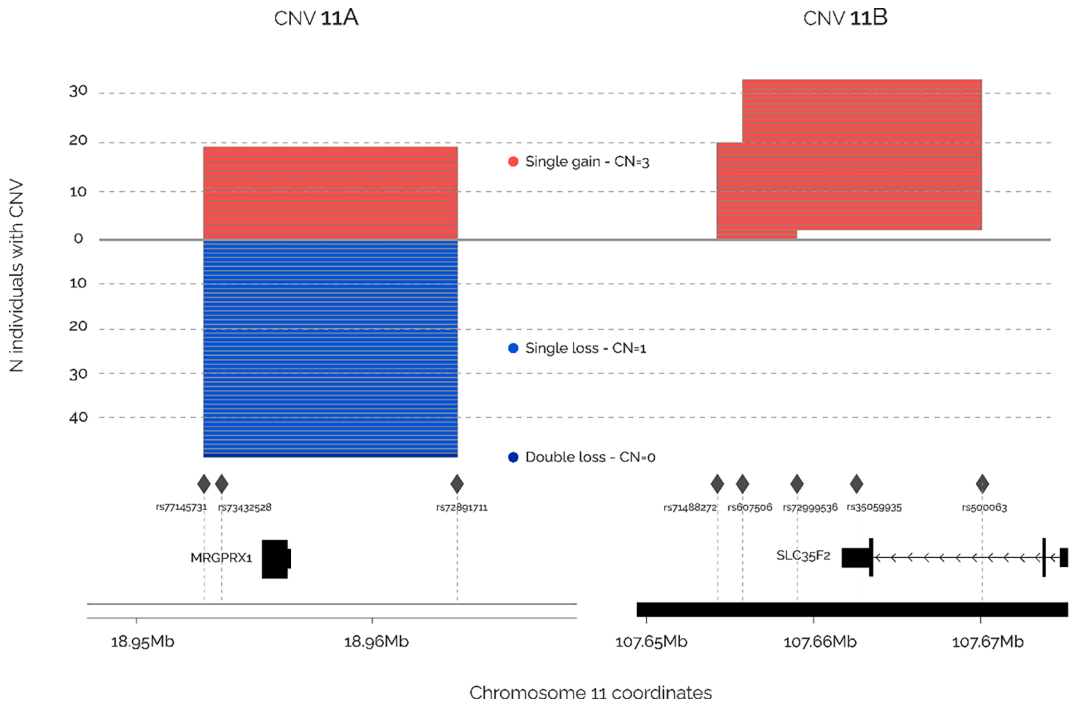

Fig. 4. Locus zoom of the consensus CNV-associated regions identified after multiple-testing correction Each line represents the CNV status and width in a single sample, for those samples with a copy number ≠ 2. Single gain CNVs are coloured in red, single loss CNVs in blue and double loss CNVs in dark blue. The y-axis shows the number of individuals presenting a CNV for the specific region. The x-axis shows the genomic coordinates. Rhombus represent the SNPs used to detect the CNVs, and genes in the regions are also shown. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Among others, we identified gene-sets related to xenobiotic detoxification, drug detoxification, CYP, oxidoreductase reactions, liver, phase I and phase II enzymes. These gene-sets were found for MEP, MECPP and its ratios, and MEOHP ratios. Furthermore, we searched for enriched GTEXv8 tissues (Appendix J). At adjusted and/or nominal statistical significance, we found enrichment for candidate tissues involved in phthalate metabolism: liver (MEOHP_MEHP and MEOHP_MEHHP), kidney cortex (MBzP, MECPP_MEHP and MEOHP_MEHHP), and bladder (MECPP_MEHP). Other significant enrichments were found for skin and vagina (oxo-MiNP_oh-MiNP), artery aorta (MiBP), and nerve tibial (MEHP).

3.5. Comparison with previous GWAS of phthalate levels in adults

Finally, we compared our findings with those reported in the GWAS conducted by Lu et al. in adults (discovery phase) (Lu et al., 2024). Consistent with our results, the loci at chr10 (*CYP2C9*) and chr6 (*SLC17A1)* were also associated with MEHP metabolites at genome-wide significance (chr10: rs74494115_T, effect = 0.086, p-value = 1.52E-29 for MECPP; and chr6: rs1177442_A, effect = 0.0322, p-value = 5.99E-10 for MECPP). In contrast, the locus at chr9 (*RAPGEF1* – MBzP) was not replicated (rs138702233_T, effect = 0.047, p-value = 2.05E-01), and the locus at chr3 (*FECHP1* – oxo-MiNP_oh-MiNP) could not be validated due to these metabolites not being measured in the study by Lu et al.

Regarding the eight selected suggestive loci of our study, four of them were nominally associated and exhibited consistent directionality in the study by Lu et al. However, two of these loci showed associations with metabolites of a different parental phthalate. More details of the comparison can be seen in Appendix K.

4. Discussion

In this study we conducted genome-wide screenings of SNPs and CNVs in order to identify genetic variants associated with phthalate metabolism in children. We identified four genome-wide significant loci. Two of them had previously been reported in a GWAS of phthalate levels in adults (Lu et al., 2024): chr6 (*SLC17A1* – MECPP_MEHHP), and chr10 (*CYP2C9* – MECPP_MEHHP; MECPP_MEHP; and MEOHP_MEHHP). The other two were novel: chr3 (*FECHP1* – oxo-MiNP_oh-MiNP) and chr9 (*RAPGEF1* – MBzP). In addition, we also identified two significant CNVs located at chr11 (*MRGPRX1* – oh-MiNP and *SLC35F2* – MEP). Functional annotation of genome-wide and suggestive SNPs highlighted pathways related to xenobiotic and drug detoxification and showed significant enrichment for liver and kidney, an indication of high gene expression level in these tissues. Furthermore, previous GWAS publications have reported significant associations of SNPs located in these loci with drug biotransformation and urinary metabolite levels of several compounds, such as urate and creatinine. Below we describe in detail the most relevant loci and genes identified, which are related to phase I and II metabolism, transport of molecules across membranes and kidney function.

4.1. Phase I metabolism

We identified three loci annotated to *CYP* genes, one of them achieving the smallest p-value of the study. This top locus at chr10 was associated with four DEHP metabolite ratios (MECPP_MEHHP, MECPP_MEHP, and MEOHP_MEHHP at genome-wide significance; and MEOHP_MEHP at suggestive significance) and contains *CYP2C8*, *CYP2C9*, *CYP2C18*, and *CYP2C19* genes. According to the EPIraction

tool, the locus can act as an enhancer of these genes in liver and/or kidney and several eQTLs were identified. *CYP* genes code for phase I monooxygenases that are key in multitude of metabolic processes of endogenous and exogenous compounds, and, indeed, SNPs in this locus have previously been reported to be associated with serum or plasma levels of several metabolites or proteins (Feofanova et al., 2020; Pazoki et al., 2021; Schlosser et al., 2020) and drug response (i.e., warfarin) (Takeuchi et al., 2009). Moreover, CYP2C enzymes are known to participate in phthalate metabolism, and *CYP2C9* variants have been associated with the biotransformation of MEHP into secondary metabolites in adults (Stajnko et al., 2022)(Lu et al., 2024). Our study confirms that similar detoxification mechanisms take place in children. Among all the SNPs in the locus, three non-synonymous SNPs, two at *CYP2C9* (rs1799853_T, Arg > Cys; and rs1057910_C, Ile > Leu) and one at *CYP2C8* (rs11572080_T, Arg > Lys) are the potential causal variants. The Arg, Leu and Lys alleles are known to confer lower enzymatic activity, and in our study they were associated with decreased MEHP metabolite ratios, indicative of decreased biotransformation of primary metabolites to more oxidized secondary metabolites (Agúndez et al., 2009).

The two additional *CYP* loci, one at chr7 harboring *CYP2W1* (p-value = 1.34E-06) and the other at chr19 including *CYP2A6*, *CYP2B6*, *CYP2A13* and *CYP2S1* (p-value = 5.61E-06), showed suggestive associations with urinary MEP levels (metabolite of diethyl phthalate (DEP)). Therefore, according to our results, DEP (LMW phthalate present in many personal care products, particularly those containing fragrances) and DEHP (HMW phthalate mainly used for food packaging) could be metabolized through different CYP genes, but further investigation is needed.

4.2. Phase II metabolism

We also identified several suggestive loci annotated to phase II metabolism genes, including *UDP-glucuronosyltransferases* (*UGTs*), *Sulfotransferases* (*SULTs*) and *Glutathione S-transferases* (*GSTs*). First, we found a locus at chr4, annotated to *UGT2B11*, *UGT2A1* and *UGT2A2,* to be suggestively associated with levels of MnBP (p-value = 7.64E-06), a metabolite of di-n-butyl-phthalate (DnBP), which is found in paints, adhesives, personal care products and deodorants. SNPs in another family member, *UGT1A7*, were previously related to urinary levels of 2-cyclohexane dicarboxylic acid, diisononyl ester (DINCH) metabolite in men and of diisobutyl phthalate (DiBP) and dibenzyl phthalate (DBzP) metabolites in women (Stajnko et al., 2022). Notably, DiBP and DnBP are structural isomers. Second, we found three loci annotated to *Sulfotransferases*: the same locus at chr4 associated with MnBP also contained the *SULT1B1* and *SULT1E1* genes, and two additional suggestive loci at chr2 (p-value = 6.17E-08) and chr19 (p-value = 2.42E-06) harboring SNPs near *SULT1C3* and *SULT2B1* genes, respectively, that showed associations with DEHP metabolites. Previous studies have described that multiple phthalates have strong inhibition potential towards *SULT* enzymes, potentially causing reduced body detoxification and liver injury (Ceauranu et al., 2023; Huang et al., 2022). Finally, we identified two loci that contain genes that participate in the glutathione metabolism: a genome-wide significant locus at chr6 annotated to *GPX5*, *GPX6* and *SLC* genes associated with MEHP and oh-MiNP levels; and a suggestive locus at chr8 near *GSR*, a central enzyme of cellular antioxidant defence, associated with MEHP ratios (p-value = 7.74E-06). This is consistent with previous research reporting that in zebrafish, MEHP exposure alters the expression of *gsr* (Kwan et al., 2021). Among all Phase II, only this last locus was nominally replicated in Lu et al., in relation to MBzP and MiBP.

4.3. Molecular transport across membranes and kidney function

Besides detoxification, we identified several loci annotated to transmembrane transporter genes which could potentially play a role in the excretion of phthalates in the intestine or kidney. These encompass several *Solute carrier* (*SLC*) genes that code for transporters of a wide array of substrates across membranes (Liu, 2019). A locus at chr6 harboring *SLC17A1*, *SLC17A2*, *SLC17A3* and *SLC17A4* was associated with MEHP ratios at genome-wide significance and showed suggestive associations with oh-MiNP. This locus was also identified by Lu et al. at genome-wide significance. Both gene-enhancer and eQTL information, suggested that these SNPs regulate *SLC* transcription in liver and/or kidney. In particular, the effect alleles were associated with lower expression of these genes. Moreover, it is known that *SLC17A1* and *SLC17A3* participate in urate regulation by facilitating the excretion of intracellular urate from the bloodstream into renal tubule cells. Consistently, previous GWAS reported associations between this locus and serum uric acid levels (Hollis-Moffatt et al., 2012), urinary metabolite levels (Schlosser et al., 2020), estimated glomerular filtration rate (creatinine) (Stanzick et al., 2021), and chronic kidney disease (Torres et al., 2021), among others. Among all SNPs in the locus, two of them were non-synonymous, one at *SCL217A4* (rs11754288_A, Ala > Thr), and the other at *SLC17A1* (rs1165196_G, Ile > Thr). *SLC17A1* gene is located in the apical membrane of proximal tubular cells in kidney and mediates urate transport into urine, with the Thr allele more efficient in this process and therefore being associated with higher urate levels in urine, and lower in serum (Middelberg et al., 2011)(Yang et al., 2010) (Chiba et al., 2015). In our study, the Thr allele was related to lower levels of urinary MECPP_MEHHP ratio and MECPP, and higher levels of MEOHP_MEHP and MEHHP_MEHP, suggesting higher excretion of MEOHP and MEHHP over MECPP and MEHP metabolites. Besides, other associations involving the *SLC* family were identified, including the association between MEP urinary levels and a single-copy gain in a CNV that overlaps the final exon and 3′UTR of the *SLC35F2* gene, predicted to enable transmembrane transporter activity.

ABCC3 encodes a member of the ATP-binding cassette (ABC) transporters that play a role in the transport of biliary and intestinal excretion of organic anions. Moreover, it is reported to be involved in multi-drug resistance (Deeley et al., 2006; Ramírez-Cosmes et al., 2021) and variants in the gene are related to serum/plasma metabolite levels (Bruhn and Cascorbi, 2014). In our study, SNPs in the *ABCC3* intronic region (chr17) were found to be suggestively associated with the MEOHP_MEHP ratio (p-value = 4.54E-06). Notably, another study showed that flies exposed to dibutyl-phthalate (DBP) have increased expression of several genes, including an homologous of the *ABCC3* gene (Williams et al., 2016). Two more transporters were identified. On the one hand, SNPs in a liver and kidney enhancer for *ATP9B* gene (chr18), which is predicted to enable ATPase-coupled intramembrane lipid transport, were suggestively associated with DEHP metabolites and ratios (p-value = 5E-07). These SNPs were also in moderate LD with a non-synonymous variant at *ATP9B* with unknown functional consequences (rs4078115_G, Ser > Gly, $r^2 = 0.63$). On the other hand, SNPs near *STAC* (chr3) were associated with the oxo-MiNP_oh-MiNP ratio at genome-wide significance and with the MEOHP_MEHHP ratio at suggestive significance. *STAC* is predicted to enable transmembrane transporter binding activity and participate in positive regulation of voltage-gated calcium channels and skeletal muscle contraction. The associations at *ABCC3* and *ATP9B* were nominally replicated by Lu et al. for MEP and MEHP, respectively. However, the correlation between MEOHP_MEHP and MEP in our data was null (r = -0.05). The association with *STAC* could not be replicated as Lu et al. had not measured MiNP metabolites.

The last group of results is comprised by two genes potentially implicated in kidney function: *RAPGEF1* and *MRGPRX1*. First, SNPs located in a liver and kidney enhancer for *RAPGEF1* at chr9 were genome-wide associated with MBzP, a metabolite of butylbenzyl phthalate (BBP) used as plasticizer for polyvinyl chloride (PVC). *RAPGEF1* encodes a guanine nucleotide exchange factor, and mutations in genes of the same family have been found in patients with nephrotic syndrome type 7 and membranoproliferative glomerulonephritis (Maywald et al., 2022; Zhu et al., 2019). This association, however, was

not replicated by Lu et al. Second, a loss-gain CNV of 10.8 kb overlapping the whole gene *MRGPRX1* was associated with oh-MiNP, a metabolite of DINP which is used primarily in the manufacture of PVC. *MRGPRX1* encodes a Mas-related G-protein-coupled receptor reported to be responsible for itch sensation, pain transmission and inflammatory reactions, including those to the antimalarial drug chloroquine (Gan et al., 2023). Additionally, mutations in other G-protein-coupled receptors have been implicated in nephrogenic diabetes insipidus, a disease characterized by polyuria and polydipsia (Wang et al., 2018). We found that individuals with three copies of the gene had increased levels of urinary levels of oh-MiNP, whereas individuals with zero or one copy had no difference versus the reference group with two copies. Given the low numbers presenting the alternative allele, further investigation in a larger sample would be required.

Our study found several associations between phthalates and SNPs known to regulate urinary metabolite levels, glomerular filtration rate, and kidney function. These associations might indicate an important role of these SNPs in phthalate excretion. Nevertheless, as our phthalate measurements were normalized by creatinine to account for urine dilution, some of our results could also just reflect renal excretion capacity. Although both explanations are possible, there is some evidence supporting the former. First, in vitro studies suggest that phthalates can be a substrate for the *SLC* transporters (Klaassen and Aleksunes, 2010); second, the associations found with the phthalate ratios were unaffected by the creatinine normalization; third the sensitivity model adjusting for creatinine instead of normalizing by it provided similar results; and four the GWAS of creatinine did not find many suggestive loci overlapping with loci found for phthalates. Finally, children of the study were, in general, healthy, with disease prevalence similar to those found in the general population, and thus we do not anticipate many children having kidney function problems that could explain associations with these loci.

4.4. Strengths of the study

The study has several strengths. First, we measured a large number of phthalate metabolites using a robust analytical method in a daily pooled urine sample (morning and night). For highly variable non-persistent chemicals such as phthalates, measuring the levels in pooled urine samples provides somewhat longer-term assessment of exposure. Second, in addition to assess the association between genetic variants and phthalate levels, we analyzed the ratio between secondary and primary metabolites as a proxy of the activity of the genes involved in detoxification. This has allowed us to increase the statistical power. Third, instead of analyzing individual SNPs in candidate genes, we performed genome-wide screenings of SNPs and several downstream annotation methods that allowed us to identify novel loci implicated in phthalate metabolism. Fourth, we did not only focus on SNPs, but also incorporated the analysis of CNVs. Fifth, the study was conducted in a large sample of children from the HELIX project, similar to the sample size in the GWAS by Lu et al. conducted in adults (Lu et al., 2024), but still limited in comparison with GWAS of other traits. Finally, this study, besides producing novel knowledge about the genes participating in phthalate detoxification and renal excretion, also provides a preliminary list of SNPs, that could be used to stratify individuals in epidemiological research according to their detoxification capacity through the creation of polygenic risk scores (PRSs).

4.5. Limitations of the study

However, the study also has some limitations. First, individual phthalate metabolite levels reflect both the exposure levels and the detoxification process, thus, in comparison to the analysis of the metabolite ratios, the statistical power might be reduced. Second, associations reported for the first time in this study have to be interpreted with caution and warrant further replication, particularly those observed in suggestive loci that have not been replicated in epidemiological studies and lack supporting evidence from other sources. Third, we have restricted our discussion on a few genes in genome-wide loci or with interesting functions for phthalate metabolism. However, since SNPs can regulate long distance genes, other genes not highlighted here could also be relevant. Functional analyses in in vitro studies or animal models should be conducted in order to identify the causal variants and genes. Fourth, we analyzed phthalate metabolites in two-void pooled urine sample and then standardized for creatinine. This has some limitations, especially for pooled samples as discussed elsewhere (O'Brien et al., 2017; Philippat and Calafat, 2021), but currently there is no consensus on how to account for urine dilution in equal volume pools. As discussed above, this could have resulted in the identification of genetic variants that regulate glomerular filtration in general, rather than phthalate excretion in particular. Finally, the analyses were restricted to European ancestry children, thus the translation to other ancestries is unknown.

5. Conclusions

In summary, through genome-wide screenings we identified known and novel loci potentially implicated in phthalate metabolism in children. Genes annotated to these loci participate in phase I and phase II metabolism, transport of molecules across membranes, and renal excretion process. Replication of the findings, especially the suggestive signals, in independent studies and validation in in vitro or in vivo systems is required.

Credit authorship contribution statement

Mariona Bustamante: Writing – original draft, Supervision, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Laura Balagué-Dobón:** Writing – original draft, Visualization, Investigation, Formal analysis, Data curation. **Zsanett Buko:** Writing – original draft, Investigation, Formal analysis. **Amrit Kaur Sakhi:** Writing – review & editing, Methodology, Investigation, Data curation. **Maribel Casas:** Writing – review & editing, Investigation, Data curation. **Lea Maitre:** Writing – review & editing, Investigation. **Sandra Andrusaityte:** Writing – review & editing, Investigation. **Regina Grazuleviciene:** Writing – review & editing, Investigation, Funding acquisition. **Kristine B. Gützkow:** Writing – review & editing, Investigation, Funding acquisition. **Anne-Lise Brantsæter:** Writing – review & editing, Investigation. **Barbara Heude:** Writing – review & editing, Investigation, Funding acquisition. **Claire Philippat:** Writing – review & editing, Investigation. **Leda Chatzi:** Writing – review & editing, Investigation, Funding acquisition. **Marina Vafeiadi:** Writing – review & editing, Investigation. **Tiffany C. Yang:** Writing – review & editing, Investigation. **John Wright:** Writing – review & editing, Investigation, Funding acquisition. **Amy Hough:** Writing – review & editing, Investigation. **Carlos Ruiz-Arenas:** Writing – review & editing, Investigation, Data curation. **Ramil N. Nurtdinov:** Writing – review & editing, Software, Investigation. **Geòrgia Escaramís:** Writing – review & editing, Investigation, Funding acquisition. **Juan R. González:** Writing – review & editing, Supervision, Software, Methodology, Investigation, Conceptualization. **Cathrine Thomsen:** Writing – review & editing, Methodology, Investigation, Data curation. **Martine Vrijheid:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We would like to thank all the families for their generous contribution.

Funding

The study has received funding from the European Community's Seventh Framework Programme (FP7/2007-206) under grant agreement no 308333 (HELIX project) and the H2020-EU.3.1.2. - Preventing Disease Programme under grant agreement no 874583 (ATHLETE project). The genotyping was supported by the project PI17/01225 and PI17/01935, funded by the Instituto de Salud Carlos III and co-funded by European Union (ERDF, "A way to make Europe") and the Centro Nacional de Genotipado-CEGEN (PRB2-ISCIII).

BiB received core infrastructure funding from the Wellcome Trust (WT101597MA) and a joint grant from the UK Medical Research Council (MRC) and Economic and Social Science Research Council (ESRC) (MR/N024397/1), and National Institute for Health Research Applied Research Collaboration Yorkshire and Humber (NIHR200166). The views expressed are those of the author(s), and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. INMA data collections were supported by grants from the Instituto de Salud Carlos III, CIBERESP, and the Generalitat de Catalunya-CIRIT. KANC was funded by the grant of the Lithuanian Agency for Science Innovation and Technology (6-04-2014_31V-66). The Norwegian Mother, Father and Child Cohort Study is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research. The Rhea project was financially supported by European projects (EU FP6-2003-Food-3-NewGeneris, EU FP6. STREP Hiwate, EU FP7 ENV.2007.1.2.2.2. Project No 211250 Escape, EU FP7-2008-ENV-1.2.1.4 Envirogenomarkers, EU FP7-HEALTH-2009- single stage CHICOS, EU FP7 ENV.2008.1.2.1.6. Proposal No 226285 ENRIECO, EU-FP7- HEALTH-2012 Proposal No 308333 HELIX), and the Greek Ministry of Health (Program of Prevention of obesity and neurodevelopmental disorders in preschool children, in Heraklion district, Crete, Greece: 2011-2014; "Rhea Plus": Primary Prevention Program of Environmental Risk Factors for Reproductive Health, and Child Health: 2012-15). ISGlobal acknowledges support from the Spanish Ministry of Science and Innovation through the "Centro de Excelencia Severo Ochoa 2019-2023" Program (CEX2018-000806-S), and support from the Generalitat de Catalunya through the CERCA Program. CRG acknowledge the support of the Spanish Ministry of Science, Innovation, and Universities to the EMBL partnership, the Centro de Excelencia Severo Ochoa, and the CERCA Programme/Generalitat de Catalunya.

LM is funded by a Juan de la Cierva-Incorporación fellowship (IJC2018-035394-I) awarded by the Spanish Ministerio de Economía, Industria y Competitividad. MC holds a Miguel Servet fellowship (MS16/00128) funded by Instituto de Salud Carlos III and co-funded by European Social Fund "Investing in your future".

Appendixes A–K. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2024.108845.

References

Agúndez, J.A.G., García-Martín, E., Martínez, C., 2009. Genetically based impairment in CYP2C8- and CYP2C9-dependent NSAID metabolism as a risk factor for gastrointestinal bleeding: Is a combination of pharmacogenomics and metabolomics required to improve personalized medicine? Expert Opin. Drug Metab. Toxicol. 5, 607–620. https://doi.org/10.1517/17425250902970998.

Bhattacharyya, M., Basu, S., Dhar, R., Dutta, T.K., 2022. Phthalate hydrolase: distribution, diversity and molecular evolution. Environ. Microbiol. Rep. 14, 333–346. https://doi.org/10.1111/1758-2229.13028.

Braun, J.M., 2017. Early-life exposure to EDCs: role in childhood obesity and neurodevelopment. Nat. Rev. Endocrinol. 13, 161–173. https://doi.org/10.1038/nrendo.2016.186.

Bruhn, O., Cascorbi, I., 2014. Polymorphisms of the drug transporters ABCB1, ABCG2, ABCC2 and ABCC3 and their impact on drug bioavailability and clinical relevance. Expert Opin. Drug Metab. Toxicol. 10, 1337–1354. https://doi.org/10.1517/17425255.2014.952630.

Casale, J., Rice, A.S., 2023. Phthalates Toxicity, StatPearls Treasure Island (FL). StatPearls Publishing.

Ceauranu, S., Ciorsac, A., Ostafe, V., Isvoran, A., 2023. Evaluation of the toxicity potential of the metabolites of di-isononyl phthalate and of their interactions with members of family 1 of sulfotransferases - A computational study. Molecules 28, 6748–6765. https://doi.org/10.3390/molecules28186748/S1.

Chiba, T., Matsuo, H., Kawamura, Y., Nagamori, S., Nishiyama, T., Wei, L., Nakayama, A., Nakamura, T., Sakiyama, M., Takada, T., Taketani, Y., Suma, S., Naito, M., Oda, T., Kumagai, H., Moriyama, Y., Ichida, K., Shimizu, T., Kanai, Y., Shinomiya, N., 2015. NPT1/SLC17A1 is a renal urate exporter in humans and its common gain-of-function variant decreases the risk of renal underexcretion gout. Arthritis Rheumatol. https://doi.org/10.1002/art.38884.

Choi, K., Joo, H., Campbell, J.L., Clewell, R.A., Andersen, M.E., Clewell, H.J., 2012. In vitro metabolism of di(2-ethylhexyl) phthalate (DEHP) by various tissues and cytochrome P450s of human and rat. Toxicol. Vitr. https://doi.org/10.1016/j.tiv.2011.12.002.

Comission Directive EC/2007/19, 2007. Commission Directive 2007/19/EC amending Directive 2002/72/EC relating to plastic materials and articles intended to come into contact with food.

Conrad, D.F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., Aerts, J., Andrews, T.D., Barnes, C., Campbell, P., Fitzgerald, T., Hu, M., Ihm, C.H., Kristiansson, K., MacArthur, D.G., MacDonald, J.R., Onyiah, I., Pang, A.W.C., Robson, S., Stirrups, K., Valsesia, A., Walter, K., Wei, J., Tyler-Smith, C., Carter, N.P., Lee, C., Scherer, S.W., Hurles, M.E., 2010. Origins and functional impact of copy number variation in the human genome. Nature 464, 704–712. https://doi.org/10.1038/nature08516.

Da Silva, V., Ramos, M., Groenen, M., Crooijmans, R., Johansson, A., Regitano, L., Coutinho, L., Zimmer, R., Waldron, L., Geistlinger, L., 2020. CNVRanger: association analysis of CNVs with gene expression and quantitative phenotypes. Bioinformatics 36, 972–973. https://doi.org/10.1093/bioinformatics/btz632.

Deeley, R.G., Westlake, C., Cole, S.P.C., 2006. Transmembrane transport of endo- and xenobiotics by mammalian ATP-binding cassette multidrug resistance proteins. Physiol. Rev. 86, 849–899. https://doi.org/10.1152/PHYSREV.00035.2005.

Domínguez-Romero, E., Komprdová, K., Kalina, J., Bessems, J., Karakitsios, S., Sarigiannis, D.A., Scheringer, M., 2023. Time-trends in human urinary concentrations of phthalates and substitutes DEHT and DINCH in Asian and North American countries (2009–2019). J. Expo. Sci. Environ. Epidemiol. 33, 244. https://doi.org/10.1038/S41370-022-00441-W.

Domínguez-Romero, E., Scheringer, M., 2019. A review of phthalate pharmacokinetics in human and rat: what factors drive phthalate distribution and partitioning? Drug Metab. Rev. https://doi.org/10.1080/03602532.2019.1620762.

Durinck, S., Spellman, P.T., Birney, E., Huber, W., 2009. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat. Protoc. 4, 1184–1191. https://doi.org/10.1038/NPROT.2009.97.

Feofanova, E.V., Chen, H., Dai, Y., Jia, P., Grove, M.L., Morrison, A.C., Qi, Q., Daviglus, M., Cai, J., North, K.E., Laurie, C.C., Kaplan, R.C., Boerwinkle, E., Yu, B., 2020. A genome-wide association study discovers 46 loci of the human metabolome in the Hispanic community health study/study of Latinos. Am. J. Hum. Genet. 107, 849–863. https://doi.org/10.1016/J.AJHG.2020.09.003.

Frederiksen, H., Skakkebæk, N.E., Andersson, A.M., 2007. Metabolism of phthalates in humans. Mol. Nutr. Food Res. https://doi.org/10.1002/mnfr.200600243.

Gan, B., Yu, L., Yang, H., Jiao, H., Pang, B., Chen, Y., Wang, C., Lv, R., Hu, H., Cao, Z., Ren, R., 2023. Mechanism of agonist-induced activation of the human itch receptor MRGPRX1. PLOS Biol. 21, e3001975. https://doi.org/10.1371/JOURNAL.PBIO.3001975.

González, J.R., Armengol, L., Solé, X., Guinó, E., Mercader, J.M., Estivill, X., Moreno, V., 2007. SNPassoc: An R package to perform whole genome association studies. Bioinformatics 23, 644–645. https://doi.org/10.1093/bioinformatics/btm025.

Haug, L.S., Sakhi, A.K., Cequier, E., Casas, M., Maitre, L., Basagana, X., Andrusaityte, S., Chalkiadaki, G., Chatzi, L., Coen, M., de Bont, J., Dedele, A., Ferrand, J., Grazuleviciene, R., Gonzalez, J.R., Gutzkow, K.B., Keun, H., McEachan, R., Meltzer, H.M., Petraviciene, I., Robinson, O., Saulnier, P.-J., Slama, R., Sunyer, J., Urquiza, J., Vafeiadi, M., Wright, J., Vrijheid, M., Thomsen, C., 2018. In-utero and childhood chemical exposome in six European mother-child cohorts. Environ. Int. 121, 751–763. https://doi.org/10.1016/J.ENVINT.2018.09.056.

Hollis-Moffatt, J.E., Phipps-Green, A.J., Chapman, B., Jones, G.T., van Rij, A., Gow, P.J., Harrison, A.A., Highton, J., Jones, P.B., Montgomery, G.W., Stamp, L.K., Dalbeth, N., Merriman, T.R., 2012. The renal urate transporter SLC17A1 locus: Confirmation of association with gout. Arthritis Res. Ther. https://doi.org/10.1186/ar3816.

Huang, H., Lan, B.D., Zhang, Y.J., Fan, X.J., Hu, M.C., Qin, G.Q., Wang, F.G., Wu, Y., Zheng, T., Liu, J.H., 2022. Inhibition of human sulfotransferases by phthalate monoesters. Front. Endocrinol. (Lausanne) 13, 1–9. https://doi.org/10.3389/fendo.2022.868105.

Kim, S.H., Park, M.J., 2014. Phthalate exposure and childhood obesity. Ann. Pediatr. Endocrinol. Metab. 19, 69–75. https://doi.org/10.6065/APEM.2014.19.2.69.

Klaassen, C.D., Aleksunes, L.M., 2010. Xenobiotic, bile acid, and cholesterol transporters: Function and regulation. Pharmacol. Rev. https://doi.org/10.1124/pr.109.002014.
Kwan, W.S., Roy, V.A.L., Yu, K.N., 2021. Review on toxic effects of Di(2-ethylhexyl) phthalate on zebrafish embryos. Toxics 9, 193–215. https://doi.org/10.3390/TOXICS9080193.
Lee, D.W., Lim, H.M., Lee, J.Y., Min, K.B., Shin, C.H., Lee, Y.A., Hong, Y.C., 2022. Prenatal exposure to phthalate and decreased body mass index of children: a systematic review and meta-analysis. Sci. Reports 2022 121 12, 1–18. Doi: 10.1038/s41598-022-13154-9.
Lee, S., Lee, H.A., Park, B., Han, H., Hong, Y.S., Ha, E.H., Park, H., 2023. Prospective association between phthalate exposure in childhood and liver function in adolescence: the Ewha Birth and Growth Cohort Study. Environ. Heal. A Glob. Access Sci. Source 22, 1–12. https://doi.org/10.1186/S12940-022-00953-W/FIGURES/2.
Liu, X., 2019. SLC family transporters. Adv. Exp. Med. Biol. 1141, 101–202. https://doi.org/10.1007/978-981-13-7647-4_3.
Lu, X., van der Meer, T.P., Kamali, Z., van Faassen, M., Kema, I.P., van Beek, A.P., Xu, X., Huo, X., Ani, A., Nolte, I.M., Wolffenbuttel, B.H.R., van Vliet-Ostaptchouk, J.V., Snieder, H., 2024. A genome-wide association study of 24-hour urinary excretion of endocrine disrupting chemicals. Environ. Int. 183, 108396 https://doi.org/10.1016/J.ENVINT.2023.108396.
Maitre, L., de Bont, J., Casas, M., Robinson, O., Aasvang, G.M., Agier, L., Andrušaitytė, S., Ballester, F., Basagaña, X., Borràs, E., Brochot, C., Bustamante, M., Carracedo, A., de Castro, M., Dedele, A., Donaire-Gonzalez, D., Estivill, X., Evandt, J., Fossati, S., Giorgis-Allemand, L., R Gonzalez, J., Granum, B., Grazuleviciene, R., Bjerve Gützkow, K., Småstuen Haug, L., Hernandez-Ferrer, C., Heude, B., Ibarluzea, J., Julvez, J., Karachaliou, M., Keun, H.C., Hjertager Krog, N., Lau, C.-H.E., Leventakou, V., Lyon-Caen, S., Manzano, C., Mason, D., McEachan, R., Meltzer, H.M., Petraviciene, I., Quentin, J., Roumeliotaki, T., Sabido, E., Saulnier, P.-J., Siskos, A.P., Siroux, V., Sunyer, J., Tamayo, I., Urquiza, J., Vafeiadi, M., van Gent, D., Vives-Usano, M., Waiblinger, D., Warembourg, C., Chatzi, L., Coen, M., van den Hazel, P., Nieuwenhuijsen, M.J., Slama, R., Thomsen, C., Wright, J., Vrijheid, M., 2018. Human Early Life Exposome (HELIX) study: a European population-based exposome cohort. BMJ Open 8, e021311. Doi: 10.1136/bmjopen-2017-021311.
Maywald, M.L., Picciotto, C., Lepa, C., Bertgen, L., Yousaf, F.S., Ricker, A., Klingauf, J., Krahn, M.P., Pavenstädt, H., George, B., 2022. Rap1 activity is essential for focal adhesion and slit diaphragm integrity. Front. Cell Dev. Biol. 10 (790365), 1–14. https://doi.org/10.3389/FCELL.2022.790365/BIBTEX.
Middelberg, R.P.S., Ferreira, M.A.R., Henders, A.K., Heath, A.C., Madden, P.A.F., Montgomery, G.W., Martin, N.G., Whitfield, J.B., 2011. Genetic variants in LPL, OASL and TOMM40/APOE-C1-C2-C4 genes are associated with multiple cardiovascular-related traits. BMC. Med. Genet. https://doi.org/10.1186/1471-2350-12-123.
Nurtdinov, R., Guigó, R., n.d. EPIraction - an atlas of candidate enhancer-gene interactions in human tissues and cell lines. [Manuscript in preparation].
O'Brien, K.M., Upson, K., Buckley, J.P., 2017. Lipid and creatinine adjustment to evaluate health effects of environmental exposures. Curr. Environ. Heal. Reports 4, 44–50. https://doi.org/10.1007/S40572-017-0122-7/FIGURES/2.
Pazoki, R., Vujkovic, M., Elliott, J., Evangelou, E., Gill, D., Ghanbari, M., van der Most, P. J., Pinto, R.C., Wielscher, M., Farlik, M., Zuber, V., de Knegt, R.J., Snieder, H., Uitterlinden, A.G., Boezen, H.M., Franke, L., van der Harst, P., Navis, G., Rots, M., Swertz, M., Wolffenbuttel, B.H.R., Wijmenga, C., Lynch, J.A., Jiang, X., Said, S., Kaplan, D.E., Lee, K.M., Serper, M., Carr, R.M., Tsao, P.S., Atkinson, S.R., Dehghan, A., Tzoulaki, I., Ikram, M.A., Herzig, K.H., Järvelin, M.R., Alizadeh, B.Z., O'Donnell, C.J., Saleheen, D., Voight, B.F., Chang, K.M., Thursz, M.R., Elliott, P., Ballas, Z.K., Bhushan, S., Boyko, E.J., Cohen, D.M., Concato, J., Aslan, M., Zhao, H., Constans, J.I., Dellitalia, L.J., Fayad, J.M., Fernando, R.S., Florez, H.J., Gaddy, M.A., Gappy, S.S., Gibson, G., Godschalk, M., Greco, J.A., Gupta, S., Gutierrez, S., Hammer, K.D., Hamner, M.B., Harley, J.B., Hung, A.M., Huq, M., Hurley, R.A., Iruvanti, P.R., Ivins, D.J., Jacono, F.J., Jhala, D.N., Kaminsky, L.S., Klein, J.B., Liangpunsakul, S., Lichy, J.H., Moser, J., Huang, G.D., Muralidhar, S., Mastorides, S. M., Mathew, R.O., Mattocks, K.M., McArdle, R., Meyer, P.N., Meyer, L.J., Moorman, J.P., Morgan, T.R., Murdoch, M., Okusaga, O.O., Oursler, K.A.K., Ratcliffe, N.R., Rauchman, M.I., Robey, R.B., Ross, G.W., Servatius, R.J., Sharma, S. C., Sherman, S.E., Sonel, E., Sriram, P., Stapley, T., Striker, R.T., Tandon, N., Villareal, G., Wallbom, A.S., Wells, J.M., Whittle, J.C., Whooley, M.A., Wilson, P.W., Sun, Y.V., Xu, J., Yeh, S.S., Connor, T., Argyres, D.P., Hauser, E.R., Beckham, J.C., Stephens, B., Aguayo, S.M., Ahuja, S.K., Pyarajan, S., Cho, K., Gaziano, J.M., Kinlay, S., Nguyen, X.M.T., Brewer, J.V., Brophy, M.T., Do, N.V., Humphries, D.E., Selva, L.E., Shayan, S., Whitbourne, S.B., Breeling, J.L., Romero, J.P.C., Ramoni, R. B., 2021. Genetic analysis in European ancestry individuals identifies 517 loci associated with liver enzymes. Nat. Commun. 12 (2579), 1–12. https://doi.org/10.1038/S41467-021-22338-2.
Pedersen, B.S., Quinlan, A.R., 2017. Who's who? Detecting and resolving sample anomalies in human DNA sequencing studies with peddy. Am. J. Hum. Genet. 100, 406–413. https://doi.org/10.1016/j.ajhg.2017.01.017.
Philippat, C., Calafat, A.M., 2021. Comparison of strategies to efficiently combine repeated urine samples in biomarker-based studies. Environ. Res. 192 (110275), 1–9. https://doi.org/10.1016/J.ENVRES.2020.110275.
Pinto, N., Eileen Dolan, M., 2011. Clinically relevant genetic variations in drug metabolizing enzymes. Curr. Drug Metab. https://doi.org/10.2174/138920011795495321.
Praveena, S.M., Teh, S.W., Rajendran, R.K., Kannan, N., Lin, C.C., Abdullah, R., Kumar, S., 2018. Recent updates on phthalate exposure and human health: a special focus on liver toxicity and stem cell regeneration. Environ. Sci. Pollut. Res. https://doi.org/10.1007/s11356-018-1652-8.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575.
Ramírez-Cosmes, A., Reyes-Jiménez, E., Zertuche-Martínez, C., Hernández-Hernández, C.A., García-Román, R., Romero-Díaz, R.I., Manuel-Martínez, A.E., Elizarrarás-Rivas, J., Vásquez-Garzón, V.R., 2021. The implications of ABCC3 in cancer drug resistance: Can we use it as a therapeutic target? Am. J. Cancer Res. 11, 4127–4140.
Regulation (EC) No 1223/2009, 2009. Regulation (EC) No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products.
Sakhi, A.K., Sabaredzovic, A., Cequier, E., Thomsen, C., 2017. Phthalate metabolites in Norwegian mothers and children: Levels, diurnal variation and use of personal care products. Sci. Total Environ. 599–600, 1984–1992. https://doi.org/10.1016/J.SCITOTENV.2017.05.109.
Schlosser, P., Li, Y., Sekula, P., Raffler, J., Grundner-Culemann, F., Pietzner, M., Cheng, Y., Wuttke, M., Steinbrenner, I., Schultheiss, U.T., Kotsis, F., Kacprowski, T., Forer, L., Hausknecht, B., Ekici, A.B., Nauck, M., Völker, U., Walz, G., Oefner, P.J., Kronenberg, F., Mohney, R.P., Köttgen, M., Suhre, K., Eckardt, K.U., Kastenmüller, G., Köttgen, A., 2020. Genetic studies of urinary metabolites illuminate mechanisms of detoxification and excretion in humans. Nat. Genet. 52, 167–176. https://doi.org/10.1038/S41588-019-0567-8.
Stajnko, A., Runkel, A.A., Kosjek, T., Snoj Tratnik, J., Mazej, D., Falnoga, I., Horvat, M., 2022. Assessment of susceptibility to phthalate and DINCH exposure through CYP and UGT single nucleotide polymorphisms. Environ. Int. https://doi.org/10.1016/j.envint.2021.107046.
Stanzick, K.J., Li, Y., Schlosser, P., Gorski, M., Wuttke, M., Thomas, L.F., Rasheed, H., Rowan, B.X., Graham, S.E., Vanderweff, B.R., Patil, S.B., Robinson-Cohen, C., Gaziano, J.M., O'Donnell, C.J., Willer, C.J., Hallan, S., Åsvold, B.O., Gessner, A., Hung, A.M., Pattaro, C., Köttgen, A., Stark, K.J., Heid, I.M., Winkler, T.W., 2021. Discovery and prioritization of variants and genes for kidney function in >1.2 million individuals. Nat. Commun. 12, 4350, 1–17. Doi: 10.1038/S41467-021-24491-0.
Takeuchi, F., McGinnis, R., Bourgeois, S., Barnes, C., Eriksson, N., Soranzo, N., Whittaker, P., Ranganath, V., Kumanduri, V., McLaren, W., Holm, L., Lindh, J., Rane, A., Wadelius, M., Deloukas, P., 2009. A genome-wide association study confirms VKORC1, CYP2C9, and CYP4F2 as principal genetic determinants of warfarin dose. PLoS Genet. 5 (e1000433), 1–9. https://doi.org/10.1371/JOURNAL.PGEN.1000433.
The Severe Covid-19 GWAS Group, 2020. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. N. Engl. J. Med. Doi: 10.1056/nejmoa2020283.
Torres, A.M., Dnyanmote, A.V., Granados, J.C., Nigam, S.K., 2021. Renal and non-renal response of ABC and SLC transporters in chronic kidney disease. Expert Opin. Drug Metab. Toxicol. https://doi.org/10.1080/17425255.2021.1899159.
D. Turner, S., 2018. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. J. Open Source Softw. Doi: 10.21105/joss.00731.
Visscher, P.M., Wray, N.R., Zhang, Q., Sklar, P., McCarthy, M.I., Brown, M.A., Yang, J., 2017. 10 years of GWAS discovery: Biology, function, and translation. Am. J. Hum. Genet. https://doi.org/10.1016/j.ajhg.2017.06.005.
Vogel, N., Schmidt, P., Lange, R., Gerofke, A., Sakhi, A.K., Haug, L.S., Jensen, T.K., Frederiksen, H., Szigeti, T., Csákó, Z., Murinova, L.P., Sidlovska, M., Janasik, B., Wasowicz, W., Tratnik, J.S., Mazej, D., Gabriel, C., Karakitsios, S., Barbone, F., Rosolen, V., Rambaud, L., Riou, M., Murawski, A., Leseman, D., Koppen, G., Covaci, A., Lignell, S., Lindroos, A.K., Zvonar, M., Andryskova, L., Fabelova, L., Richterova, D., Horvat, M., Kosjek, T., Sarigiannis, D., Maroulis, M., Pedraza-Diaz, S., Cañas, A., Verheyen, V.J., Bastiaensen, M., Gilles, L., Schoeters, G., Esteban-López, M., Castaño, A., Govarts, E., Koch, H.M., Kolossa-Gehring, M., 2023. Current exposure to phthalates and DINCH in European children and adolescents – Results from the HBM4EU Aligned Studies 2014 to 2021. Int. J. Hyg. Environ. Health 249 (114101), 1–11. https://doi.org/10.1016/J.IJHEH.2022.114101.
Wang, I.J., Karmaus, W.J.J., 2017. Oxidative stress-related genetic variants may modify associations of phthalate exposures with asthma. Int. J. Environ. Res. Public Health 14, 162–175. https://doi.org/10.3390/IJERPH14020162.
Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S.F.A., Hakonarson, H., Bucan, M., 2007. PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res. 17, 1665–1674. https://doi.org/10.1101/GR.6861907.
Wang, X.X., Wang, D., Luo, Y., Myakala, K., Dobrinskikh, E., Rosenberg, A.Z., Levi, J., Kopp, J.B., Field, A., Hill, A., Lucia, S., Qiu, L., Jiang, T., Peng, Y., Orlicky, D., Garcia, G., Herman-Edelstein, M., D'Agati, V., Henriksen, K., Adorini, L., Pruzanski, M., Xie, C., Krausz, K.W., Gonzalez, F.J., Ranjit, S., Dvornikov, A., Gratton, E., Levi, M., 2018. FXR/TGR5 dual agonist prevents progression of nephropathy in diabetes and obesity. J. Am. Soc. Nephrol. 29, 118–137. https://doi.org/10.1681/ASN.2017020222.
Wang, Y., Zhu, H., Kannan, K., 2019. A review of biomonitoring of phthalate exposures. Toxics. https://doi.org/10.3390/TOXICS7020021.
Watanabe, K., Taskesen, E., Van Bochoven, A., Posthuma, D., 2017. Functional mapping and annotation of genetic associations with FUMA. Nat. Commun. 8, 1826. https://doi.org/10.1038/s41467-017-01261-5.
Wickham, H., 2016. ggplot2: Elegant graphics for data analysis, use R! Springer Cham. https://doi.org/10.1007/978-3-319-24277-4.
Williams, M.J., Wiemerslage, L., Gohel, P., Kheder, S., Kothegala, L.V., Schiöth, H.B., 2016. Dibutyl phthalate exposure disrupts evolutionarily conserved insulin and glucagon-like signaling in drosophila males. Endocrinology 157, 2309–2321. https://doi.org/10.1210/EN.2015-2006.

Xu, C., Tachmazidou, I., Walter, K., Ciampi, A., Zeggini, E., Greenwood, C.M.T., 2014. Estimating genome-wide significance for whole-genome sequencing studies. Genet. Epidemiol. 38, 281. https://doi.org/10.1002/GEPI.21797.

Yang, Q., Köttgen, A., Dehghan, A., Smith, A. V., Glazer, N.L., Chen, M.H., Chasman, D.I., Aspelund, T., Eiriksdottir, G., Harris, T.B., Launer, L., Nalls, M., Hernandez, D., Arking, D.E., Boerwinkle, E., Grove, M.L., Li, M., Linda Kao, W.H., Chonchol, M., Haritunians, T., Li, G., Lumley, T., Psaty, B.M., Shlipak, M., Hwang, S.J., Larson, M. G., O'Donnell, C.J., Upadhyay, A., Van Duijn, C.M., Hofman, A., Rivadeneira, F., Stricker, B., Uitterlinden, A.G., Paré, G., Parker, A.N., Ridker, P.M., Siscovick, D.S., Gudnason, V., Witteman, J.C., Fox, C.S., Coresh, J., 2010. Multiple genetic loci influence serum urate levels and their relationship with gout and cardiovascular disease risk factors. Circ. Cardiovasc. Genet. Doi: 10.1161/CIRCGENETICS.109.934455.

Zhu, B., Cao, A., Li, J., Young, J., Wong, J., Ashraf, S., Bierzynska, A., Menon, M.C., Hou, S., Sawyers, C., Campbell, K.N., Saleem, M.A., He, J.C., Hildebrandt, F., D'Agati, V.D., Peng, W., Kaufman, L., 2019. Disruption of MAGI2-RapGEF2-Rap1 signaling contributes to podocyte dysfunction in congenital nephrotic syndrome caused by mutations in MAGI2. Kidney Int. 96, 642–655. https://doi.org/10.1016/J.KINT.2019.03.016.