Molecular gatekeeper discovery: Workflow for linking multiple environmental biomarkers to metabolomics

Miao Yu, Susan L.Teitelbaum, Georgia Dolios, Lam-Ha T. Dang, Peijun Tu, Mary S. Wolff, Lauren M. Petrick*

Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, 10029, USA

* Corresponding author: lauren.petrick@mssm.edu. Icahn School of Medicine at Mount Sinai, 1428 Madison Ave, Atran Building 3rd floor, Box 1057, New York, NY, 20019

The authors declare they have no actual or potential competing financial interests

Abstract

The exposome reflects the many exposures to various factors across the life-course that can affect health. Sensitive techniques like metabolomics can reveal the underlying molecular basis linking exposures to disease and generate hypotheses for future quantitative toxicological studies. Current applications of metabolomics are primarily to identify metabolic changes linking a single exposure and a health outcome(s); there is no general framework for multiple exposures. Here, we explore the concept of 'molecular gatekeepers'—key metabolites that link single or multiple exposure biomarkers with correlated clusters of endogenous metabolites—to inform health-relevant biological targets. We performed untargeted metabolomics on plasma from 152 adolescent girls participating in the Growing Up Healthy Study in New York City, using liquid chromatography-high resolution mass spectrometry (LC-HRMS). We then performed network analysis to link metabolites to environmental biomarkers including five trace elements (Cd, Mn, Pb, Se, and Hg) and five perfluorinated chemicals (PFCs; n-PFOS, Sm-PFOS, n-PFOA, PFHxS, PFNA) previously measured in the same samples. We defined any metabolite associated with at least one environmental biomarker and correlated with at least one other metabolite (Spearman rho > 0.9) as a 'molecular gatekeeper'. Associations of gatekeepers with health outcomes (e.g., body mass index, age at menarche) were tested with linear models. After removing redundant peaks, 964 (positive mode) and 1784 (negative mode) metabolite features were used for network analysis. Of 95 and 138 metabolites, respectively, associated with at least one exposure, 28 and 43 were molecular gatekeepers. Further, lysophosphatidylcholine(16:0) and taurodeoxycholate were correlated with both n-PFOA and n-PFOS, suggesting a shared dysregulation from multiple xenobiotic exposures. One annotated gatekeeper, sphingomyelin(d18:2/14:0), was significantly associated with age at menarche; yet, no direct association was detected between any exposure biomarkers and age at menarche. Thus, molecular gatekeepers may provide a general data analysis framework to discover molecular linkages between exposure biomarkers and health outcomes that may otherwise be obscured by complex interactions in direct measurements. This framework may aid in identifying vulnerable biological pathways for future exposome research.

Keywords: exposome, metabolomics, network analysis, perfluorinated chemicals, trace metals, mixtures

Introduction

Exposomics centers on characterizing how various exposures, (e.g., trace metals, persistent and non-persistent organics, and psycho-social factors, across the human life-course can affect health[1]. For example, exposure to lead is associated with a broad range of adverse health outcomes in both adults and children[2], but the biological pathways linking lead exposure to such outcomes are not fully understood. Indeed, toxicology studies of lead exposure in animals implicate liver toxicity[3], but there is limited epidemiologic evidence to support such associations[4]. Such inconsistencies may reflect that interactions between exposures can act antagonistically or synergistically to impart influences on health. This complexity introduces challenges in uncovering relationships between the exposome and health.

Sensitive technologies may provide an avenue to resolve this complexity. In particular, metabolomics approaches enable unbiased measurement of thousands of metabolites to identify changes in the metabolome profile as a result of exposures or disease processes[5]. Metabolites are also connected by biological pathways[6] and biochemical reactions[7] that themselves can be associated with specific health conditions or diseases[8]. Further, exogenous exposures influence health outcomes via interaction with endogenous metabolites[9]. As such, the metabolome may mediate the health outcomes resulting from exposures. Yet, analytical triangulation among exposures, metabolites, and health outcomes is complex. For example, metabolite profiles have been reduced to latent variables using principal component analysis[10,11] for statistical testing. However, biological interpretation of the metabolites summarized within latent variables is challenging. Other approaches to reduce dimensionality include using biological pathway information[12]. These provide an easily interpreted biological link between exposures and health outcomes, but are limited to established pathways from databases. While the network-based approach, xMWAS, focuses on pairwise correlation among different omics datasets[13], this analysis does not consider the inner correlation network within the single omics dataset.

To overcome some of these limitations, here we explore the concept of ‘molecular gatekeepers’—key metabolites that link single or multiple exposure biomarkers with other endogenous metabolites. We posit that metabolites that are highly correlated with other metabolites contain more biological information than those metabolites that are isolated. As such, molecular gatekeepers that link highly correlated metabolites and exposure biomarkers may be particularly important targets for future toxicological studies and for uncovering the role of the exposome on health. We sought to validate this concept using network analysis between untargeted metabolomics and ten environmental biomarkers to identify molecular gatekeepers. Our findings suggest this method as a potential new approach to inform future exposomic and toxicological studies.

Methods

Study participants

Girls ages 6–8 years were enrolled at the Icahn School of Medicine at Mount Sinai in the Growing Up Healthy Study from 2004-2007 as described in previous studies. Participants provided assent and parents/guardians provided written consent. The study was approved by the Mount Sinai IRB. In addition to age, inclusion criteria required that girls have no underlying endocrine medical conditions and be of Black or Hispanic race/ethnicity. Blood samples were

collected from enrolled participants during subsequent annual visits at ages 7–16 years. During the examination visits, trained and certified staff members obtained standardized anthropometric measurements, including height and weight. BMI was expressed as age- sex-specific percentile based on the CDC algorithm, as described[14]. Age at menarche was ascertained through an algorithm combining parental information and self-report[15]. The current analysis includes 152 girls with data on exposure biomarkers, outcome, and potential confounding variables (Table 1).

Table 1. Descriptive characteristics of the Growing Up Healthy Study girls in the current study (N=152).

Characteristic	Category	N (%)	Mean (+-SD)	Range
Race/ethnicity	Black	50 (33)	-	-
	Hispanic	100 (66)	-	-
	White	2 (1)	-	-
Age at blood collection (yr)	-	-	5.9(3.2)	[1,12]
BMI percentile[a]	-	-	71.5(29.4)	[1.6,99.9]
BMI group[a]	high	76 (50)	-	-
	low	76 (50)	-	-
Age at menarche	-	-	11.7(1.2)	[9.1,14.9]

[a]:BMI sex- and age-specific. BMI group is dichotomized at the median (85.1%)

Exposure biomarker measurements

Samples were previously analyzed at the National Center for Environmental Health at the CDC using on-line solid phase extraction-HPLC-isotope dilution-tandem mass spectrometry (LC-MS)[16] for plasma perfluorinated chemicals (PFCs; n-PFOS, n-perfluorooctane sulfonate; n-PFOA, n-perfluorooctanoate; Sm-PFOS, monomethyl branched isomers of PFOS; PFHxS, perfluorohexane sulfonate; PFNA, perfluorononanoate) or using inductively coupled plasma mass spectrometry for whole blood trace elements (BCD, Cadmium; BMN, Manganese; BPB, Lead; BSE, Selenium; THG, Mercury). The CDC laboratory is certified by the Health Care Financing Administration to comply with the requirements set forth in the Clinical Laboratory Improvement Act of 1988 and is recertified biannually[17]. Spearman rhos among blood levels of trace metals and PFCs were all less than 0.9. Limit of detection for all PFCs was 0.1 μg/L while the limits of detection for the trace elements analytes were 0.1μg/L(BCD), 0.99 μg/L (BMN), 0.07 μg/dL(BPB), 0.28 μg/L(THG), and 24.48 μg/L(BSE).

Untargeted analysis

Plasma samples stored at -80°C were thawed on ice and vortexed, and 50-uL aliquots were transferred to a microcentrifuge tube. 150 uL of methanol containing internal standards were added, and the sample was vortexed and incubated at -80°C for 30 min. Samples were

centrifuged, and the supernatant dried using a Savant SC250EXP SpeedVac concentrator at 35°C for 90 minutes, and stored at -80°C until analysis. Before LC-HRMS analysis, dried extracts were reconstituted either in 100% methanol or in acetonitrile:water (8:2, v/v). An additional 10-uL aliquot from each sample was combined for use as a pooled quality control sample ('pooled QC') and processed similarly. Following the same protocol the matrix blank (replacing the plasma with water) and multiple pooled QCs were extracted. Samples were analyzed using reverse-phase (RP) and hydrophilic interaction liquid (ZH) chromatography connected to HRMS in negative (RPN) and positive mode (ZHP), respectively, as described elsewhere[18]. Samples were analyzed in a randomized order with pooled QCs injected routinely throughout the run.

Data pre-processing

For untargeted data in positive and negative mode, raw instrument data was converted into mzxml format[19] and analyzed using R programming platform (version 4.0.3). The xcms package[20] was used to generate a feature table with optimized parameters by the IPO package[21]. Features with relative standard deviation (RSD%) across the pooled QC samples smaller than 30% and fold change greater than 3 in blank samples were retained for further analysis. The GlobalStd algorithm[22] was used to reduce the redundant features such as isotopologue or adducts. Remaining peaks were further filtered by considering the base peak in the cluster with Pearson's correlation coefficients larger than 0.9 within GlobalStd retention time bins. Then, the peak lists (2058 and 989 peaks in RPN and ZHP, respectively) were refined to merge the peaks within 5s and mass accuracy within 5 ppm, resulting in a final detection of 1784 and 964 independent peaks in RPN and ZHP, respectively, for downstream analysis. Metabolite annotations were performed by MS/MS spectrum matching to Metlin[23], GNPS[24], MS-DIAL[25], and local databases with default settings.

Statistical and analysis

Analysis was performed on the ZHP and RPN data separately. Correlation between independent peaks was determined using a Spearman's rho > 0.9 threshold to distinguish potential pathway networks. For hypothesis testing, linear models using the empirical Bayes procedures[26] were firstly built between the 10 exposure biomarker concentrations and the log2-transformed intensity of independent peaks to identify significant exposure-metabolite relationships (p-value < 0.05 after FDR control Benjamini–Hochberg [BH] correction). Data visualization to show the intersections between exposure biomarkers and independent peaks was performed by UpSet plot[27]. Associations between log2-transformed intensity of "molecular gatekeeper" or exposure biomarkers and girls' BMI percentile (continuous), BMI groups (high/low based on median value of BMI percentage), and age at menarche (years) were determined using linear models or logistic regression (for BMI groups) with or without adjustment for age at blood collection, race/ethnicity and/or BMI percentage. Data processing R script is shared as supporting information, and a R implementation to discover gatekeepers between exposome and metabolome is available as the enet package at https://github.com/yufree/enet.

Results

"Molecular gatekeepers" are defined as metabolites that are 1) significantly associated with at least one exposure biomarker and 2) are correlated with at least one other metabolite. Such metabolites represent a potential role in bridging an exposure to other metabolites and possible downstream biological dysregulation. Therefore, as a workflow, we firstly determined the

metabolites that were significantly associated with exposure biomarkers. Then, we determined the metabolites that were correlated with other metabolites. Finally, we selected the metabolites that were found in both sets as gatekeepers.

Metabolite—exposure associations

For ZHP, there were 964 independent peaks following filtering. Of these(m/z range 71-1092, retention time range 47s-1139s), 95 metabolites were significantly associated with one exposure biomarker and 20 were significantly associated with multiple exposure biomarkers (2–3 exposures, see Figure 1). For RPN, there were 1784 independent peaks. Of these(m/z range 87-1197, retention time range 20s-791s),138 metabolites were significantly associated with one exposure biomarker and 42 were significantly associated with multiple exposure biomarkers (2–4 exposures, see Figure 1). Overall, a greater number of significant associations with metabolites were found with PFCs than with trace elements (total of 345 and 26, respectively). The greatest number of exposure biomarker–metabolite associations could be found for n-PFOS and n-PFOA for both modes. Only 41 significant associations were found for Sm-PFOS, PFHxS and PFNA. For trace elements, predominant associations were between metabolites and BMN (23) and metabolites and THG (3); no metabolites were significantly associated with BPB, BSE, and BCD. Interestingly, in contrast to the PFCs, the number of trace elements associated with metabolites was higher in ZHP than RPN (16 and 10, respectively).

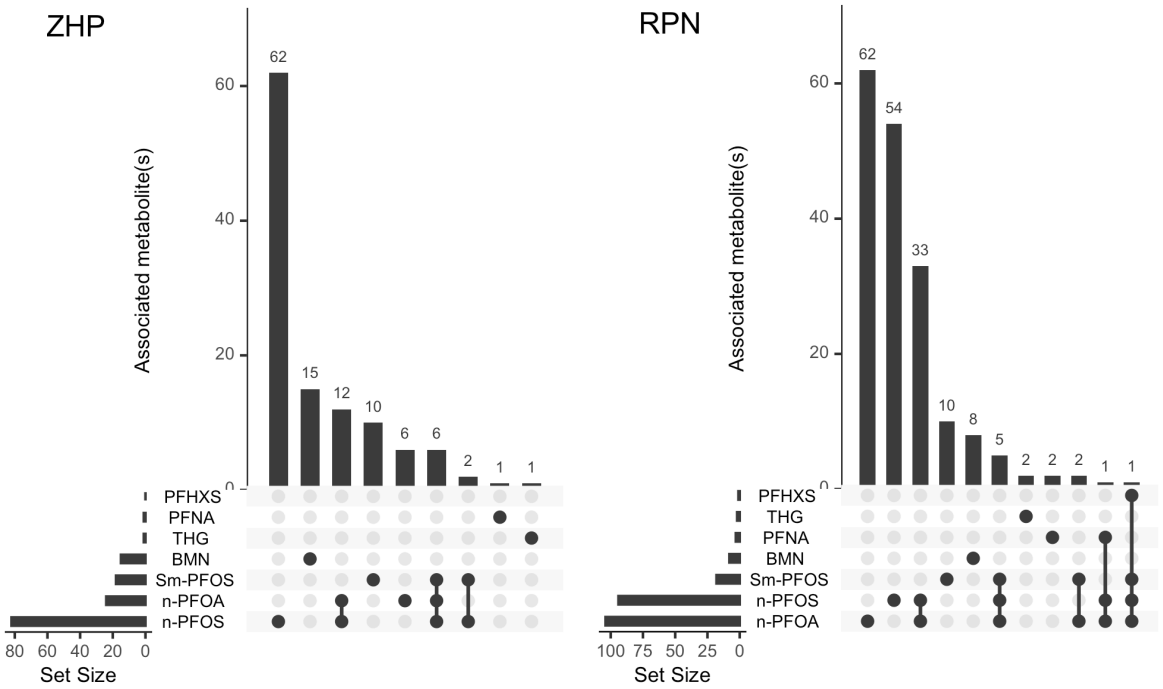


Figure 1. Upset plots of pairwise associations between metabolites and exposure biomarkers for ZHP and RPN modes. Associations were detected by linear models using the empirical Bayes procedures with p-values < 0.05 after FDR control using BH correction. The Set Size is the total number of metabolites associated with each exposure biomarker, while Associated Metabolites (vertical axis) describe the number of metabolites distributed across each intersection of multiple exposure biomarkers.

Molecular Gatekeeper Discovery

For ZHP, 178 out of 964 metabolites were correlated with at least one other metabolite (Spearman Rho >0.9). Of these, 28 peaks were significantly associated with at least one exposure (PFC or trace element analyte). For RPN, 368 out of 1784 metabolites were correlated with at least one other metabolite. Of these, 43 peaks were significantly associated with at least one PFC or trace element analyte. Thus, the 28 (ZHP) and 43 (RPN) peaks were considered gatekeepers, and those gatekeepers were highly correlated with a total of 58 (ZHP) and 101 (RPN) unique metabolites. The full list of 71 gatekeepers and their details can be found in Table S1, and the corresponding gatekeeper networks are shown in Figure 2. This figure depicts relationships between exposure biomarkers (blue points) and correlated metabolites (red points) and gatekeepers (red triangles). Compared with larger numbers of gatekeepers of PFC (27 in ZHP and 40 in RPN), only four gatekeepers were negatively associated with BMN (one in ZHP and three in RPN). Three gatekeepers in ZHP and twelve gatekeepers in RPN were significantly associated with multiple exposures (Table S1).

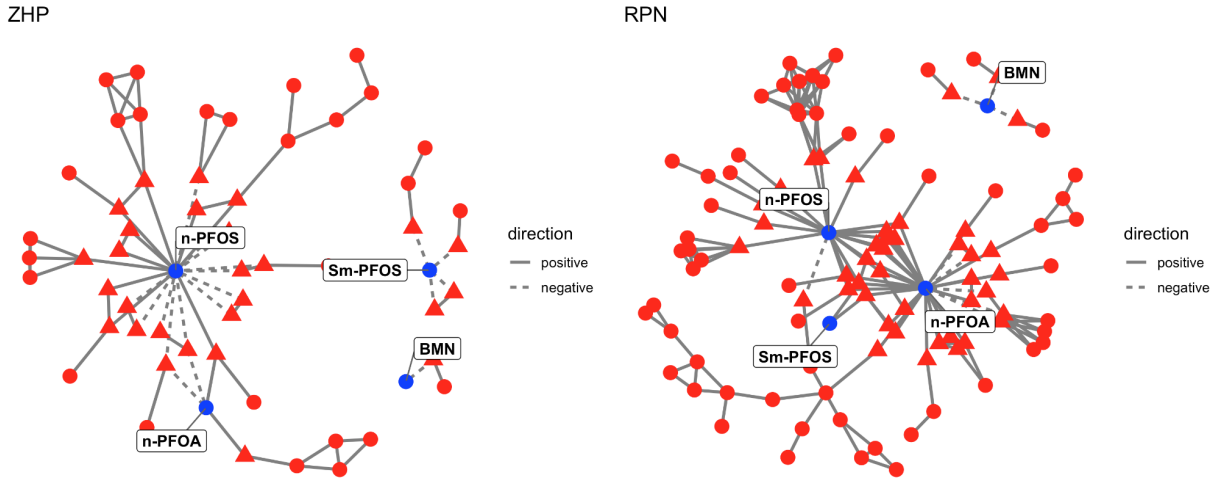
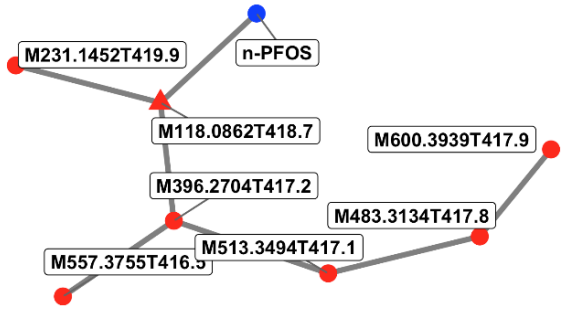


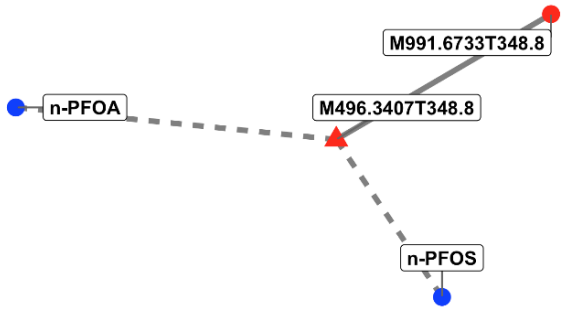
Figure 2. Molecular gatekeeper discovery network for metabolites measured in ZHP and RPN with PFCs and trace elements as exposure biomarkers. Red nodes represent independent metabolites, triangles represent gatekeeper metabolites, and blue nodes with labels represent exposure biomarkers. The edges among triangles and blue nodes represent significant associations (p-value < 0.05, empirical Bayes procedures after FDR control with BH correction). The edges among triangles and other nodes represent correlations (Spearman correlation coefficient > 0.9). Solid lines indicate a positive association or correlation while dashed lines indicate negative association or correlation. Gatekeeper molecules represent potentially important links between environmental exposures and metabolite sets.

Of the 71 gatekeepers, we ascertained high confidence annotations for ten. We then extracted the network for each annotated gatekeeper; we depict the linkages between the exposure biomarker and any correlated metabolites in Figure 3. From ZHP mode, we identified betaine, LPC(16:0), LPC(18:0), SM(d18:2/14:0), and PE(20:4/P-18:0) as gatekeepers. From RPN mode, we identified gatekeepers hippuric acid, dehydroepiandrosterone sulfate, androsterone sulfate, taurodeoxycholate and GPC(P-18:0/20:4). Two gatekeepers were associated with multiple exposures. LPC(16:0) was negatively associated with n-PFOS and n-PFOA and positively correlated with an unannotated metabolite (M991.6733T348.8). Taurodeoxycholate was positively associated with both n-PFOS and n-PFOA and one unannotated metabolite (M514.2835T307.5). There were two additional gatekeepers without annotation that were associated with three PFCs (Table S1).

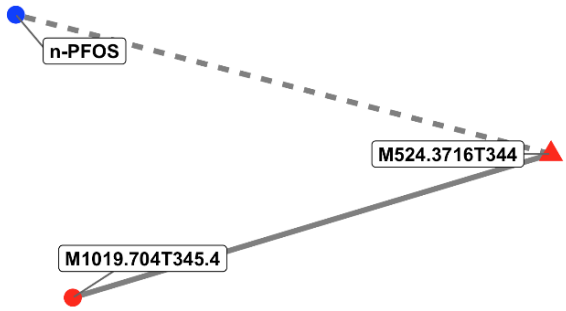
A: betaine (M118.0862T418.7)



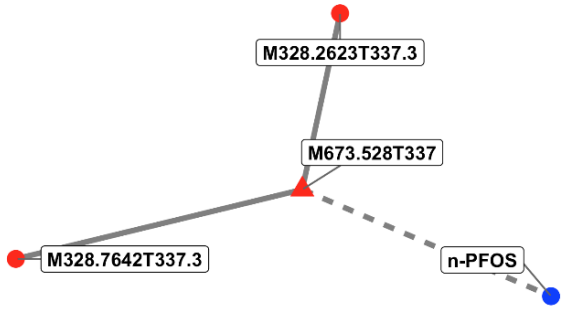
B: LPC(16:0) (M496.3407T348.8)



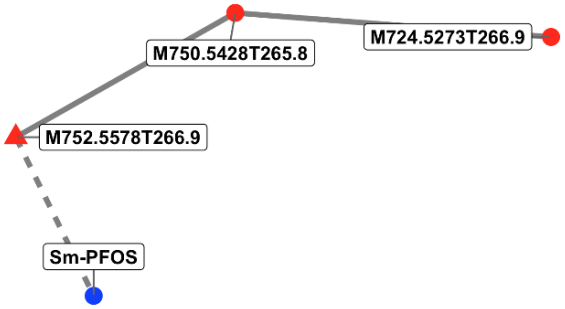
C: LPC(18:0) (M524.3716T344)



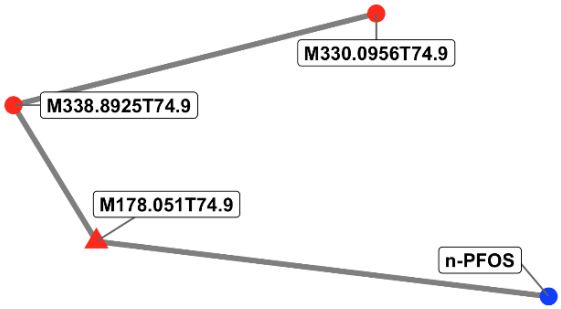
D: SM(d18:2/14:0) (M673.528T337)



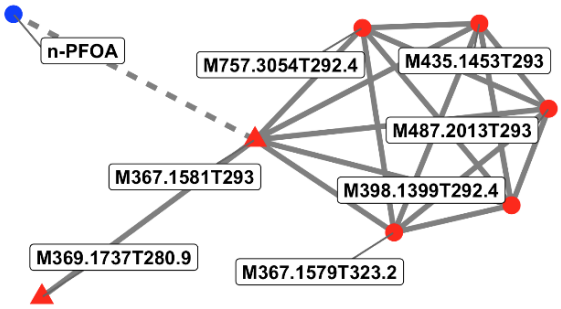
E: PE(20:4/P-18:0) (M752.5578T266.9)



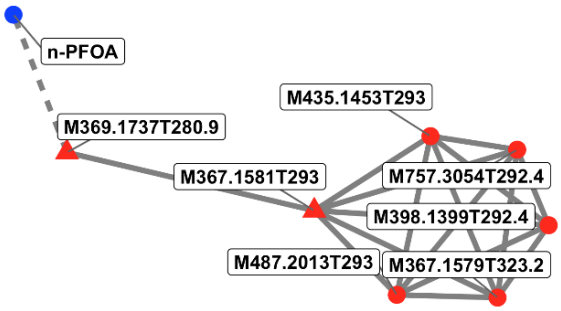
F: hippuric acid (M178.051T74.9)



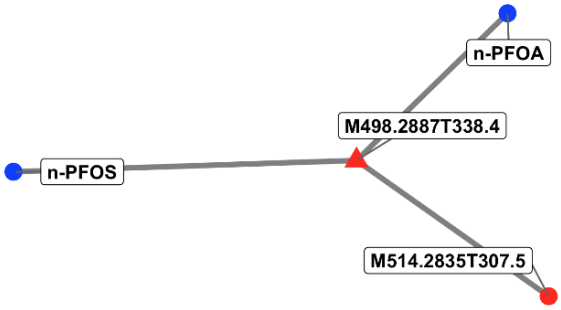
G: dehydroepiandrosterone sulfate (M367.1581T293)



H: androsterone sulfate (M369.1737T280.9)



I: taurodeoxycholate (M498.2887T338.4)



J: GPC(P-18:0/20:4) (M838.5957T725.4)

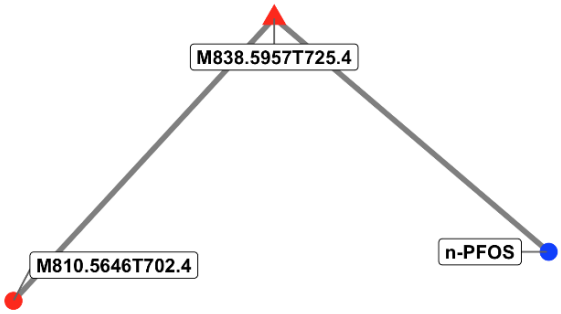


Figure 3. Networks for annotated gatekeepers. Each network for the ten annotated gatekeepers in both ZHP and RPN mode are displayed, depicting the linkages between the exposure biomarker and any correlated metabolites (A–J). Red nodes represent independent metabolites, triangles represent gatekeeper metabolites, and blue nodes with labels represent exposures. The edges among triangles and blue nodes represent significant associations (p-value < 0.05, empirical Bayes procedures after FDR control with BH correction). The edges among triangles and other nodes represent correlation (Spearman correlation coefficient > 0.9). Solid lines indicate positive association or correlation while dashed lines indicate negative association or correlation.

Gatekeepers linked with health outcomes

We next sought to test the hypothesis that gatekeepers represent conduits to downstream biological effects of exposures, and therefore hold relevance for exposome research. We first determined if there were direct associations between the exposure biomarkers and the health outcomes of interest. We estimated associations between the exposure biomarkers that are linked with gatekeepers (n-PFOS, n-PFOA, Sm-PFOS, and BMN) and age at menarche, BMI percentile, and BMI group (Table 2). Without adjustment for covariates, exposure biomarkers n-PFOS and n-PFOA were significantly associated with BMI percentile and exposure biomarker n-PFOS was associated with BMI group. After adjustment for race/ethnicity and age at blood collection n-PFOS remained associated with BMI percentile. No associations were detected between exposure biomarkers and age at menarche without adjustment or after adjusting for covariates.

Table 2. Associations between selected exposure biomarkers and health outcomes. Significant nominal p-values (< 0.05) are in bold. For continuous variables, linear regression was performed with or without adjustment for covariates. For the *BMI percentile group* (high versus low), logistic regression was performed with or without adjustment for covariates. N=152

Exposure biomarker (ug/L)	BMI percentile (unadjusted, β±SE)	BMI percentile (adjusted[a], β±SE)	BMI percentile group (unadjusted, β±SE)	BMI percentile group (adjusted[a], β±SE)	Age at menarche in years (unadjusted, β±SE)	Age at menarche in years (adjusted[b], β±SE)
n-PFOA	**-5.7±2.8**	-4.5±3.1	-0.4±0.2	-0.3±0.2	1.5±1.3	1.2±1.5
n-PFOS	**-5±1.8**	**-4.2±2.1**	**-0.3±0.1**	-0.3±0.2	0.5±0.9	-0.2±1
Sm-PFOS	-3.6±5.8	-2.5±5.9	-0.1±0.4	-0.1±0.4	1.9±2.8	2.4±2.8
BMN	0.5±0.4	0.3±0.4	0.02±0.03	0.002±0.03	-0.3±0.2	-0.2±0.2

[a]: adjusted for race/ethnicity, age at blood collection
[b]: adjusted for race/ethnicity, age at blood collection, and BMI percentile

As a proof-of-concept, we investigated associations between the annotated gatekeepers and *BMI percentile*, *BMI group,* and *age at menarche* (Table 3). Annotated gatekeepers SM(d18:2/14:0), dehydroepiandrosterone, and androsterone sulfate were positively associated with *BMI percentile* and *BMI group* both without adjustment and after adjusting for covariates. Taurodeoxycholate and GPC(P-18:0/20:4) were negatively associated with *BMI percentile* and *BMI group* both without adjustment and after adjusting for covariates. In addition, SM(d18:2/14:0) was positively associated with *age at menarche* after adjusting for covariates.

As the units of exposures (μg/L) are different from the log-transformed intensity data of metabolomics datasets, the estimates in Table 3 represented an effect several times larger than for the single exposures in Table 2.

Table 3. Associations among annotated gatekeepers, selected health outcomes and exposure biomarkers. Significant nominal p-values (< 0.05) are in bold. For continuous variables (log2), linear regression was performed with or without adjustment for relevant covariates. For the *BMI percentile group* (high versus low), logistic regression was performed with or without adjustment for relevant covariates. N=152.

Gatekeepers (log2 intensity)	mz	rt	mode	BMI percentile (unadjusted, β±SE)	BMI percentile (adjusted[e], β±SE)	BMI percentile group (unadjusted, β±SE)	BMI percentile group (adjusted[f], β±SE)	Age at menarche in years (unadjusted, β±SE)	Age at menarche in years (adjusted[b], β±SE)	Associated exposure(s)
Betaine	118.0862	418.7	ZHP	1.4 ±3.4	5.3 ±3.8	0.1 ±0.2	0.3 ±0.3	0 ±0.1	0 ±0.2	n-PFOS
LPC(16:0)[a]	496.3407	348.8	ZHP	**37 ±15.7**	26.1 ±16.6	1.2 ±1.1	0.4 ±1.2	-0.7 ±0.6	-0.3 ±0.7	n-PFOA, n-PFOS
LPC(18:0)	524.3716	344	ZHP	13 ±9	7.4 ±9.6	0.1 ±0.6	-0.3 ±0.7	0.1 ±0.4	0.3 ±0.4	n-PFOS
SM(d18:2/14:0)[b]	673.528	337	ZHP	**32.2 ±5**	**30.5 ±5.2**	**2.2 ±0.5**	**2.2 ±0.5**	0.2 ±0.2	**0.8 ±0.2**	n-PFOS
PE(20:4/P-18:0)[c]	752.5578	266.9	ZHP	-5.1 ±4.8	-2.5 ±5.1	**-0.8 ±0.3**	-0.7 ±0.4	0.2 ±0.2	0 ±0.2	Sm-PFOS
Hippuric acid	178.051	74.9	RPN	-4.1 ±2.2	-3.8 ±2.3	-0.2 ±0.2	-0.2 ±0.2	0 ±0.1	0 ±0.1	n-PFOS
Dehydroepiandrosterone sulfate	367.1581	293	RPN	**6.1 ±2.4**	**6.0 ±2.5**	**0.5 ±0.2**	**0.5 ±0.2**	-0.2 ±0.1	-0.1 ±0.1	n-PFOA
Androsterone sulfate	369.1737	280.9	RPN	**10.9 ±2.3**	**10.8 ±2.4**	**0.9 ±0.2**	**1.0 ±0.2**	-0.2 ±0.1	-0.1 ±0.1	n-PFOA
taurodeoxycholate	498.2887	338.4	RPN	**-4.3 ±1.6**	**-3.9 ±1.6**	**-0.3 ±0.1**	**-0.3 ±0.1**	0.1 ±0.1	0.1 ±0.1	n-PFOA, n-PFOS
GPC(P-18:0/20:4)[d]	838.5957	725.4	RPN	**-16 ±4.6**	**-13.4 ±5.1**	**-1.3 ±0.4**	**-1.2 ±0.4**	0.2 ±0.2	-0.1 ±0.2	n-PFOS

[a]: LPC, lysophosphatidylcholine
[b]: SM, sphingomyelin
[c]: PE, phosphatidylethanolamine
[d]: GPC, glycerophosphocholine
[e]: adjusted for race/ethnicity, age at blood collection
[f]: adjusted for race/ethnicity, age at blood collection, and BMI percentile

Discussion

Gatekeepers are characterized by their connective roles between exposure biomarkers and other endogenous metabolites. As shown by the KEGG pathway database[28], metabolites within pathways tend to be correlated with each other instead of isolated from other metabolites. In this case, metabolites that are highly correlated with both other metabolites and exposure biomarkers should contain more biologically relevant information than metabolites associated with a single exposure biomarker in isolation. The purpose of the gatekeeper discovery process is to find those information-rich metabolites among the thousands of metabolites that are measured, as *a priori* targets for testing associations with health outcomes. Therefore, gatekeeper discovery can be considered as a dimension-reduction method to highlight biologically relevant metabolites based on network analysis.

Our results showed that seven out of ten exposure biomarkers—PFNA, THG, BMN, Sm-PFOS, n-PFOA, n-PFOS, and PFHxS— were significantly associated with a total of 233 metabolites in RPN and ZHP modes combined, highlighting the complex interactions between plasma metabolites and both PFCs and trace elements. Further network analysis identified 28 gatekeepers in ZHP and 43 gatekeepers in RPN associated with sm-PFOS, n-PFOS, n-PFOA, and MNE, indicating that these three PFCs and BMN may be particularly biologically important exposures. While studies of exposures to PFCs and metabolomics in human populations are emerging in the literature[29–31], studies on manganese exposure are sparse. We found only a single human study investigating associations between manganese exposure and metabolite profiles during pregnancy[32], although studies in rat models have been performed[33,34]. Since manganese exposure has been associated with both beneficial and harmful health effects[35,36], future metabolomics studies investigating this exposure in humans are encouraged.

We found gatekeepers that were associated with more than one exposure. While 13 out of the 15 gatekeepers that were linked to multiple exposures are unannotated, we found that LPC(16:0), a glycerophospholipid, was negatively associated with both PFOA and PFOS while taurodeoxycholate, an active bile acid derivative, was positively associated with both PFOA and PFOS. These results are consistent with observations in epidemiological studies. Dysregulated glycerophospholipid metabolism has been associated with PFC exposure in children and adults[37,38]. In addition, recent literature suggests associations between PFCs and cholesterol levels in human plasma which may be mediated by reabsorption of bile acids in the gut[39]. Bile acid metabolism is influenced by PFOA and PFOS exposures in human HepaRG hepatoma cells[40], and a recent pilot study found positive associations between several PFCs, including PFOA and PFOS, with bile acids[41]. Therefore, gatekeeper discovery facilitated the selection of metabolites involved in important biological response pathways following exposures. Further, two unannotated gatekeepers (Table S1, M418.0859T38.6 and M717.7553T25.9) were associated with three PFCs (PFOS/PFOA/Sm-PFOS) suggesting that PFCs may work synergistically to alter specific pathways. These metabolites, associated with multiple exposures, may be particularly important for understanding health impacts of exposure groups.

Several gatekeepers linked exposure biomarkers to health outcomes, even when direct associations were absent. Since n-PFOS is associated with *BMI* (Table 2), and SM(d18:2/14:0), taurodeoxycholate, and GPC(P-18:0/20:4) are associated with n-PFOS and with *BMI* (Table 3), these gatekeeper metabolites may play key roles in the n-PFOS–BMI interaction at a molecular level, providing hypotheses for future research. Similarly, while there were no associations between exposures and *age at menarche* (Table 2), SM(d18:2/14:0) was positively associated with *age at menarche* after adjustment and negatively associated with PFOS. In this case, direct associations may have been masked by antagonist relationships, and gatekeeper discovery revealed SM(d18:2/14:0) as a sensitive endogenous marker of this exposure-health interaction. SM(d18:2/14:0) was also associated with BMI, as has been observed in other studies[42]. Taurodeoxycholate has been observed at higher levels in prepubertal obese children with insulin resistance compared with their non-insulin resistant counterparts[43]. However, we found taurodeoxycholate was negatively correlated with BMI in the adolescent girls, most of whom were not obese. Together, gatekeeper discovery generated hypotheses linking biomarkers to health outcomes to guide future mechanistic research.

In summary, we demonstrated that the gatekeeper discovery workflow selects key metabolites from untargeted data that encompass biologically important information linking exposures to health outcomes. The associations between paired exposure–metabolite relationships were built using a simple linear regression. However, the gatekeeper discovery framework can be extended in future work to multivariate linear regression to consider covariates, or other machine learning algorithms such as random forest or support vector machine. Additionally, the correlation threshold among metabolites can be reduced by the user to reveal additional biological pathways or gatekeepers, or correlation can be replaced by other relationships such as reactomics or paired mass distances[7]. As a general data analysis framework, gatekeeper discovery is flexible for direct adoption to different environmental health studies and even different omics. Limitations of this study include a small sample size and limited health outcome data. Therefore, results reported here may not be generalizable to other populations.

Acknowledgement

We acknowledge Dr. Antonia Calafat and the laboratory at CDC for their biomarker exposure measures. We are grateful for support from the National Institutes of Health from grants P30ES023515, U2CES030859, U2CES026561, U2CES026555, R21ES030882, and R01ES031117.

References

(1) Polderman, T. J. C.; Benyamin, B.; de Leeuw, C. A.; Sullivan, P. F.; van Bochoven, A.; Visscher, P. M.; Posthuma, D. Meta-Analysis of the Heritability of Human Traits Based on Fifty Years of Twin Studies. *Nat. Genet.* **2015**, *47* (7), 702–709. https://doi.org/10.1038/ng.3285.

(2) Lanphear, B. P.; Rauch, S.; Auinger, P.; Allen, R. W.; Hornung, R. W. Low-Level Lead Exposure and Mortality in US Adults: A Population-Based Cohort Study. *Lancet Public Health* **2018**, *3* (4), e177–e184. https://doi.org/10.1016/S2468-2667(18)30025-2.

(3) Cordner, A.; De La Rosa, V. Y.; Schaider, L. A.; Rudel, R. A.; Richter, L.; Brown, P. Guideline Levels for PFOA and PFOS in Drinking Water: The Role of Scientific Uncertainty, Risk Assessment Decisions, and Social Factors. *J. Expo. Sci. Environ. Epidemiol.* **2019**, *29* (2), 157–171. https://doi.org/10.1038/s41370-018-0099-9.

(4) Steenland, K.; Fletcher, T.; Savitz, D. A. Epidemiologic Evidence on the Health Effects of Perfluorooctanoic Acid (PFOA). *Environ. Health Perspect.* **2010**, *118* (8), 1100–1108. https://doi.org/10.1289/ehp.0901827.

(5) Pezzatti, J.; Boccard, J.; Codesido, S.; Gagnebin, Y.; Joshi, A.; Picard, D.; González-Ruiz, V.; Rudaz, S. Implementation of Liquid Chromatography–High Resolution Mass Spectrometry Methods for Untargeted Metabolomic Analyses of Biological Samples: A Tutorial. *Anal. Chim. Acta* **2020**, *1105*, 28–44. https://doi.org/10.1016/j.aca.2019.12.062.

(6) Li, S.; Park, Y.; Duraisingham, S.; Strobel, F. H.; Khan, N.; Soltow, Q. A.; Jones, D. P.; Pulendran, B. Predicting Network Activity from High Throughput Metabolomics. *PLOS Comput. Biol.* **2013**, *9* (7), e1003123. https://doi.org/10.1371/journal.pcbi.1003123.

(7) Yu, M.; Petrick, L. Untargeted High-Resolution Paired Mass Distance Data Mining for Retrieving General Chemical Relationships. *Commun. Chem.* **2020**, *3* (1), 1–6. https://doi.org/10.1038/s42004-020-00403-z.

(8) Mi, K.; Jiang, Y.; Chen, J.; Lv, D.; Qian, Z.; Sun, H.; Shang, D. Construction and Analysis of Human Diseases and Metabolites Network. *Front. Bioeng. Biotechnol.* **2020**, *8*. https://doi.org/10.3389/fbioe.2020.00398.

(9) Bessonneau, V.; Gerona, R. R.; Trowbridge, J.; Grashow, R.; Lin, T.; Buren, H.; Morello-Frosch, R.; Rudel, R. A. Gaussian Graphical Modeling of the Serum Exposome and Metabolome Reveals Interactions between Environmental Chemicals and Endogenous Metabolites. *Sci. Rep.* **2021**, *11* (1), 7607. https://doi.org/10.1038/s41598-021-87070-9.

(10) Nyamundanda, G.; Brennan, L.; Gormley, I. C. Probabilistic Principal Component Analysis for Metabolomic Data. *BMC Bioinformatics* **2010**, *11*, 571. https://doi.org/10.1186/1471-2105-11-571.

(11) Xu, Y.; Goodacre, R. Multiblock Principal Component Analysis: An Efficient Tool for Analyzing Metabolomics Data Which Contain Two Influential Factors. *Metabolomics* **2012**, *8* (1), 37–51. https://doi.org/10.1007/s11306-011-0361-9.

(12) Langfelder, P.; Horvath, S. WGCNA: An R Package for Weighted Correlation Network Analysis. *BMC Bioinformatics* **2008**, *9*, 559. https://doi.org/10.1186/1471-2105-9-559.

(13) Uppal, K.; Ma, C.; Go, Y.-M.; Jones, D. P. XMWAS: A Data-Driven Integration and Differential Network Analysis Tool. *Bioinformatics* **2018**, *34* (4), 701–702. https://doi.org/10.1093/bioinformatics/btx656.

(14) Biro, F. M.; Pajak, A.; Wolff, M. S.; Pinney, S. M.; Windham, G. C.; Galvez, M. P.; Greenspan, L. C.; Kushi, L. H.; Teitelbaum, S. L. Age of Menarche in a Longitudinal US Cohort. *J. Pediatr. Adolesc. Gynecol.* **2018**, *31* (4), 339–345. https://doi.org/10.1016/j.jpag.2018.05.002.

(15) Biro, F. M.; Greenspan, L. C.; Galvez, M. P.; Pinney, S. M.; Teitelbaum, S.; Windham, G. C.; Deardorff, J.; Herrick, R. L.; Succop, P. A.; Hiatt, R. A.; Kushi, L. H.; Wolff, M. S. Onset of Breast Development in a Longitudinal Cohort. *Pediatrics* **2013**. https://doi.org/10.1542/peds.2012-3773.

(16) Kato, K.; Basden, B. J.; Needham, L. L.; Calafat, A. M. Improved Selectivity for the Analysis of Maternal Serum and Cord Serum for Polyfluoroalkyl Chemicals. *J. Chromatogr. A* **2011**, *1218* (15), 2133–2137. https://doi.org/10.1016/j.chroma.2010.10.051.

(17) National Report on Human Exposure to Environmental Chemicals | CDC https://www.cdc.gov/exposurereport/index.html (accessed 2021 -04 -21).

(18) Yu, M.; Dolios, G.; Yong-Gonzalez, V.; Björkqvist, O.; Colicino, E.; Halfvarson, J.; Petrick, L. Untargeted Metabolomics Profiling and Hemoglobin Normalization for Archived Newborn Dried Blood Spots from a Refrigerated Biorepository. *J. Pharm. Biomed. Anal.* **2020**, *191*, 113574. https://doi.org/10.1016/j.jpba.2020.113574.

(19) Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M.-Y.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.; Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer, A.; Langridge, J.; Connolly, B.; Chadick, T.; Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.; Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P. A Cross-Platform Toolkit for Mass Spectrometry and Proteomics. *Nat. Biotechnol.* **2012**, *30*, 918–920. https://doi.org/10.1038/nbt.2377.

(20) Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Anal. Chem.* **2006**, *78* (3), 779–787. https://doi.org/10.1021/ac051437y.

(21) Libiseller, G.; Dvorzak, M.; Kleb, U.; Gander, E.; Eisenberg, T.; Madeo, F.; Neumann, S.; Trausinger, G.; Sinner, F.; Pieber, T.; Magnes, C. IPO: A Tool for Automated Optimization of XCMS Parameters. *BMC Bioinformatics* **2015**, *16*, 118. https://doi.org/10.1186/s12859-015-0562-8.

(22) Yu, M.; Olkowicz, M.; Pawliszyn, J. Structure/Reaction Directed Analysis for LC-MS Based Untargeted Analysis. *Anal. Chim. Acta* **2019**, *1050*, 16–24. https://doi.org/10.1016/j.aca.2018.10.062.

(23) Xue, J.; Guijas, C.; Benton, H. P.; Warth, B.; Siuzdak, G. METLIN MS 2 Molecular Standards Database: A Broad Chemical and Biological Resource. *Nat. Methods* **2020**, *17* (10), 953–954. https://doi.org/10.1038/s41592-020-0942-5.

(24) Aron, A. T.; Gentry, E. C.; McPhail, K. L.; Nothias, L.-F.; Nothias-Esposito, M.; Bouslimani, A.; Petras, D.; Gauglitz, J. M.; Sikora, N.; Vargas, F.; van der Hooft, J. J. J.; Ernst, M.; Kang, K. B.; Aceves, C. M.; Caraballo-Rodríguez, A. M.; Koester, I.; Weldon, K. C.; Bertrand, S.; Roullier, C.; Sun, K.; Tehan, R. M.; P, C. A. B.; Christian, M. H.; Gutiérrez, M.; Ulloa, A. M.; Tejeda Mora, J. A.; Mojica-Flores, R.; Lakey-Beitia, J.; Vásquez-Chaves, V.; Zhang, Y.; Calderón, A. I.; Tayler, N.; Keyzers, R. A.; Tugizimana, F.; Ndlovu, N.; Aksenov, A. A.; Jarmusch, A. K.; Schmid, R.; Truman, A. W.; Bandeira, N.; Wang, M.; Dorrestein, P. C. Reproducible Molecular Networking of Untargeted Mass Spectrometry Data Using GNPS. *Nat. Protoc.* **2020**, *15* (6), 1954–1991. https://doi.org/10.1038/s41596-020-0317-5.

(25) Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M. MS-DIAL: Data-Independent MS/MS Deconvolution for Comprehensive Metabolome Analysis. *Nat. Methods* **2015**, *12* (6), 523–526. https://doi.org/10.1038/nmeth.3393.

(26) Ritchie, M. E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C. W.; Shi, W.; Smyth, G. K. Limma

Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res.* **2015**, *43* (7), e47–e47. https://doi.org/10.1093/nar/gkv007.

(27) Lex, A.; Gehlenborg, N.; Strobelt, H.; Vuillemot, R.; Pfister, H. UpSet: Visualization of Intersecting Sets. *IEEE Trans. Vis. Comput. Graph.* **2014**, *20* (12), 1983–1992. https://doi.org/10.1109/TVCG.2014.2346248.

(28) Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a Reference Resource for Gene and Protein Annotation. *Nucleic Acids Res.* **2016**, *44* (D1), D457–D462. https://doi.org/10.1093/nar/gkv1070.

(29) Li, R.; Guo, C.; Tse, W. K. F.; Su, M.; Zhang, X.; Lai, K. P. Metabolomic Analysis Reveals Metabolic Alterations of Human Peripheral Blood Lymphocytes by Perfluorooctanoic Acid. *Chemosphere* **2020**, *239*, 124810. https://doi.org/10.1016/j.chemosphere.2019.124810.

(30) Wang, X.; Liu, L.; Zhang, W.; Zhang, J.; Du, X.; Huang, Q.; Tian, M.; Shen, H. Serum Metabolome Biomarkers Associate Low-Level Environmental Perfluorinated Compound Exposure with Oxidative /Nitrosative Stress in Humans. *Environ. Pollut.* **2017**, *229*, 168–176. https://doi.org/10.1016/j.envpol.2017.04.086.

(31) Shao, X.; Ji, F.; Wang, Y.; Zhu, L.; Zhang, Z.; Du, X.; Chung, A. C. K.; Hong, Y.; Zhao, Q.; Cai, Z. Integrative Chemical Proteomics-Metabolomics Approach Reveals Acaca/Acacb as Direct Molecular Targets of PFOA. *Anal. Chem.* **2018**, *90* (18), 11092–11098. https://doi.org/10.1021/acs.analchem.8b02995.

(32) Wang, M.; Xia, W.; Liu, H.; Liu, F.; Li, H.; Chang, H.; Sun, J.; Liu, W.; Sun, X.; Jiang, Y.; Liu, H.; Wu, C.; Pan, X.; Li, Y.; Rang, W.; Lu, S.; Xu, S. Urinary Metabolomics Reveals Novel Interactions between Metal Exposure and Amino Acid Metabolic Stress during Pregnancy. *Toxicol. Res.* **2018**, *7* (6), 1164–1172. https://doi.org/10.1039/c8tx00042e.

(33) Wang, H.; Liu, Z.; Wang, S.; Cui, D.; Zhang, X.; Liu, Y.; Zhang, Y. UHPLC-Q-TOF/MS Based Plasma Metabolomics Reveals the Metabolic Perturbations by Manganese Exposure in Rat Models. *Metallomics* **2017**, *9* (2), 192–203. https://doi.org/10.1039/C7MT00007C.

(34) Fordahl, S.; Cooney, P.; Qiu, Y.; Xie, G.; Jia, W.; Erikson, K. M. Waterborne Manganese Exposure Alters Plasma, Brain, and Liver Metabolites Accompanied by Changes in Stereotypic Behaviors. *Neurotoxicol. Teratol.* **2012**, *34* (1), 27–36. https://doi.org/10.1016/j.ntt.2011.10.003.

(35) Livingstone, C. Manganese Provision in Parenteral Nutrition: An Update. *Nutr. Clin. Pract.* **2018**, *33* (3), 404–418. https://doi.org/10.1177/0884533617702837.

(36) Iyare, P. U. The Effects of Manganese Exposure from Drinking Water on School-Age Children: A Systematic Review. *NeuroToxicology* **2019**, *73*, 1–7. https://doi.org/10.1016/j.neuro.2019.02.013.

(37) Stratakis, N.; V Conti, D.; Jin, R.; Margetaki, K.; Valvi, D.; Siskos, A. P.; Maitre, L.; Garcia, E.; Varo, N.; Zhao, Y.; Roumeliotaki, T.; Vafeiadi, M.; Urquiza, J.; Fernández-Barrés, S.; Heude, B.; Basagana, X.; Casas, M.; Fossati, S.; Gražulevičienė, R.; Andrušaitytė, S.; Uppal, K.; McEachan, R. R. C.; Papadopoulou, E.; Robinson, O.; Haug, L. S.; Wright, J.; Vos, M. B.; Keun, H. C.; Vrijheid, M.; Berhane, K. T.; McConnell, R.; Chatzi, L. Prenatal Exposure to Perfluoroalkyl Substances Associated With Increased Susceptibility to Liver Injury in Children. *Hepatol. Baltim. Md* **2020**, *72* (5), 1758–1770. https://doi.org/10.1002/hep.31483.

(38) Vrijheid, M.; Fossati, S.; Maitre, L.; Márquez, S.; Roumeliotaki, T.; Agier, L.; Andrusaityte, S.; Cadiou, S.; Casas, M.; de Castro, M.; Dedele, A.; Donaire-Gonzalez, D.; Grazuleviciene, R.; Haug, L. S.; McEachan, R.; Meltzer, H. M.; Papadopouplou, E.; Robinson, O.; Sakhi, A. K.; Siroux, V.; Sunyer, J.; Schwarze, P. E.; Tamayo-Uria, I.; Urquiza, J.; Vafeiadi, M.; Valentin, A.; Warembourg, C.; Wright, J.; Nieuwenhuijsen, M. J.; Thomsen, C.; Basagaña, X.; Slama, R.; Chatzi, L. Early-Life Environmental Exposures and Childhood Obesity: An Exposome-Wide Approach. *Environ. Health Perspect.* **2020**,

128 (6), 67009. https://doi.org/10.1289/EHP5975.

(39) Minutes of the expert meeting on perfluooroctane sulfonic acid and perfluorooctanoic acid in food assessment /paper/Minutes-of-the-expert-meeting-on-perfluooroctane-in/a7f56cd59eb20355f2f33e6162af56790eb74570 (accessed 2021 -06 -01).

(40) Behr, A.-C.; Kwiatkowski, A.; Ståhlman, M.; Schmidt, F. F.; Luckert, C.; Braeuning, A.; Buhrke, T. Impairment of Bile Acid Metabolism by Perfluorooctanoic Acid (PFOA) and Perfluorooctanesulfonic Acid (PFOS) in Human HepaRG Hepatoma Cells. *Arch. Toxicol.* **2020**, *94* (5), 1673–1686. https://doi.org/10.1007/s00204-020-02732-3.

(41) Salihović, S.; Dickens, A. M.; Schoultz, I.; Fart, F.; Sinisalu, L.; Lindeman, T.; Halfvarson, J.; Orešič, M.; Hyötyläinen, T. Simultaneous Determination of Perfluoroalkyl Substances and Bile Acids in Human Serum Using Ultra-High-Performance Liquid Chromatography–Tandem Mass Spectrometry. *Anal. Bioanal. Chem.* **2020**, *412* (10), 2251–2259. https://doi.org/10.1007/s00216-019-02263-6.

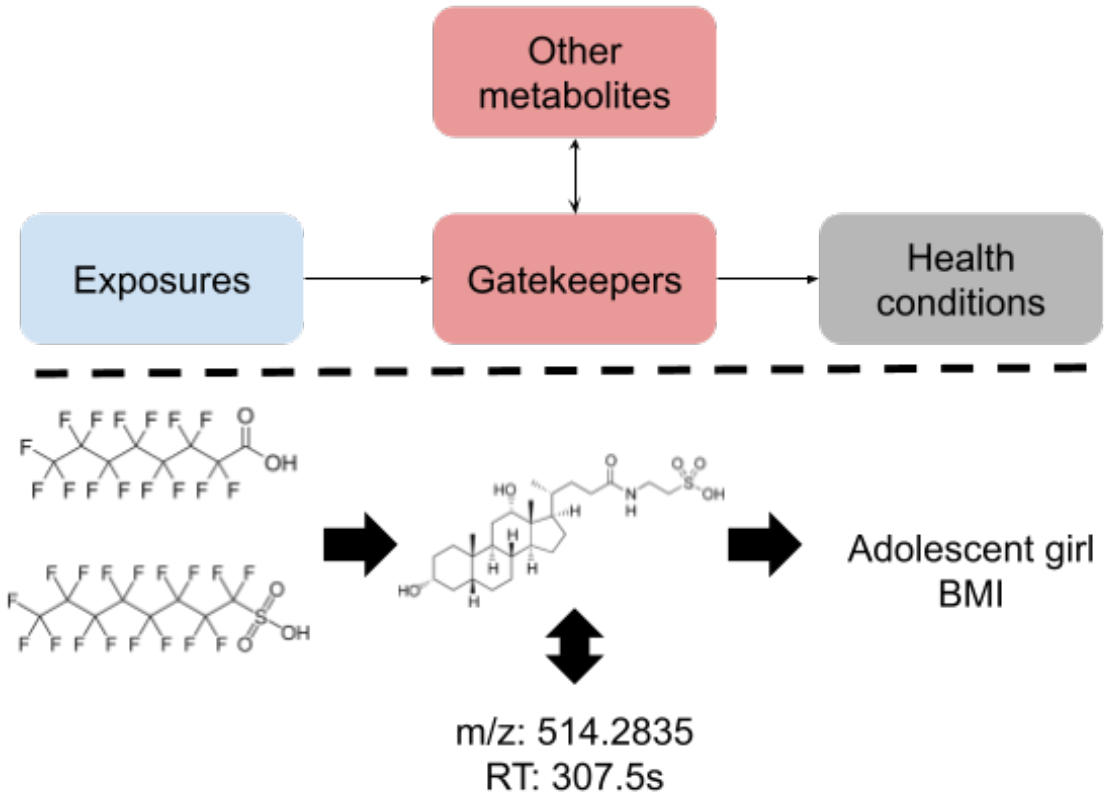
(42) Huynh, K.; Barlow, C. K.; Jayawardana, K. S.; Weir, J. M.; Mellett, N. A.; Cinel, M.; Magliano, D. J.; Shaw, J. E.; Drew, B. G.; Meikle, P. J. High-Throughput Plasma Lipidomics: Detailed Mapping of the Associations with Cardiometabolic Risk Factors. *Cell Chem. Biol.* **2019**, *26* (1), 71-84.e4. https://doi.org/10.1016/j.chembiol.2018.10.008.

(43) Mastrangelo, A.; Martos-Moreno, G. Á.; García, A.; Barrios, V.; Rupérez, F. J.; Chowen, J. A.; Barbas, C.; Argente, J. Insulin Resistance in Prepubertal Obese Children Correlates with Sex-Dependent Early Onset Metabolomic Alterations. *Int. J. Obes. 2005* **2016**, *40* (10), 1494–1502. https://doi.org/10.1038/ijo.2016.92.

534
535

536 TOC

537



538
539

Supporting Information

Table S1. Gatekeepers found in this study and their association with exposure(s).The associations were detected by linear models using the empirical Bayes procedures and the coefficients of the association show p-values < 0.05 after FDR control BH correction.

Gatekeeper	**m/z**	**Retention time (s)**	**mode**	**Associated environmental biomarkers**
M90.9767T553.8	90.9767	553.8	ZHP	n-PFOA, n-PFOS
M104.107T418.5	104.107	418.5	ZHP	n-PFOS
M118.0862T418.7	118.0862	418.7	ZHP	n-PFOS
M202.1549T409	202.1549	409	ZHP	n-PFOS
M202.1802T389.4	202.1802	389.4	ZHP	n-PFOS
M231.1452T419.9	231.1452	419.9	ZHP	n-PFOS
M243.1831T414.4	243.1831	414.4	ZHP	n-PFOS
M280.1543T151	280.1543	151	ZHP	n-PFOS
M280.2382T156.9	280.2382	156.9	ZHP	n-PFOS
M305.207T413.8	305.207	413.8	ZHP	n-PFOS
M328.2623T337.3	328.2623	337.3	ZHP	n-PFOS
M357.2117T233.8	357.2117	233.8	ZHP	n-PFOA
M439.3297T389.4	439.3297	389.4	ZHP	n-PFOS
M496.3407T348.8	496.3407	348.8	ZHP	n-PFOA, n-PFOS
M520.3404T346.8	520.3404	346.8	ZHP	n-PFOS
M523.0476T455.7	523.0476	455.7	ZHP	Sm-PFOS
M524.3716T344	524.3716	344	ZHP	n-PFOS
M586.3599T175.6	586.3599	175.6	ZHP	n-PFOS
M627.5345T260.8	627.5345	260.8	ZHP	n-PFOS
M673.528T337	673.528	337	ZHP	n-PFOS
M728.558T271	728.558	271	ZHP	Sm-PFOS

M752.5578T266.9	752.5578	266.9	ZHP	Sm-PFOS
M773.0767T455.7	773.0767	455.7	ZHP	Sm-PFOS
M907.5782T305.5	907.5782	305.5	ZHP	n-PFOA, n-PFOS
M991.6733T348.8	991.6733	348.8	ZHP	n-PFOS
M1019.704T345.4	1019.704	345.4	ZHP	n-PFOS
M1039.6725T345.4	1039.6725	345.4	ZHP	n-PFOS
M1091.7023T342.7	1091.7023	342.7	ZHP	BMN
M178.051T74.9	178.051	74.9	RPN	n-PFOS
M188.0105T105.8	188.0105	105.8	RPN	n-PFOS
M231.052T210	231.052	210	RPN	n-PFOA
M231.5534T209.4	231.5534	209.4	RPN	n-PFOA
M340.1413T37.4	340.1413	37.4	RPN	n-PFOS
M367.1581T293	367.1581	293	RPN	n-PFOA
M369.1737T280.9	369.1737	280.9	RPN	n-PFOA
M380.8143T29.6	380.8143	29.6	RPN	n-PFOS
M389.2469T532.2	389.2469	532.2	RPN	n-PFOA, n-PFOS
M390.0555T31.3	390.0555	31.3	RPN	n-PFOA
M391.2042T525.5	391.2042	525.5	RPN	n-PFOA, n-PFOS
M391.2621T559.4	391.2621	559.4	RPN	n-PFOA, n-PFOS
M401.2022T68.9	401.2022	68.9	RPN	n-PFOA, n-PFOS
M407.2199T472.8	407.2199	472.8	RPN	n-PFOS
M411.3473T603.7	411.3473	603.7	RPN	n-PFOA
M412.0889T38.6	412.0889	38.6	RPN	n-PFOA
M418.0859T38.6	418.0859	38.6	RPN	n-PFOA, n-PFOS, Sm-PFOS
M425.3992T655.7	425.3992	655.7	RPN	n-PFOA
M435.1453T293	435.1453	293	RPN	n-PFOA
M438.7727T30.1	438.7727	30.1	RPN	n-PFOS
M451.3777T619.4	451.3777	619.4	RPN	n-PFOA

M454.2654T517.1	454.2654	517.1	RPN	n-PFOA
M467.3735T597	467.3735	597	RPN	n-PFOA, n-PFOS
M481.2935T560.7	481.2935	560.7	RPN	n-PFOA
M487.2906T502.5	487.2906	502.5	RPN	n-PFOA
M491.3228T560.1	491.3228	560.1	RPN	n-PFOA
M493.205T214.2	493.205	214.2	RPN	n-PFOA
M498.2887T338.4	498.2887	338.4	RPN	n-PFOA, n-PFOS
M539.8857T26.5	539.8857	26.5	RPN	n-PFOA, n-PFOS
M540.8686T24.7	540.8686	24.7	RPN	n-PFOA, n-PFOS
M540.9461T26.5	540.9461	26.5	RPN	n-PFOA, n-PFOS
M561.4873T701.2	561.4873	701.2	RPN	n-PFOA
M578.3009T470.4	578.3009	470.4	RPN	BMN
M614.3461T481.9	614.3461	481.9	RPN	BMN
M641.3532T338.4	641.3532	338.4	RPN	n-PFOA
M641.3534T331.1	641.3534	331.1	RPN	n-PFOA
M657.3304T366.9	657.3304	366.9	RPN	n-PFOA
M717.7553T25.9	717.7553	25.9	RPN	n-PFOA, n-PFOS, Sm-PFOS
M718.7816T30.1	718.7816	30.1	RPN	n-PFOA, n-PFOS
M836.5798T706.6	836.5798	706.6	RPN	n-PFOS
M838.5957T725.4	838.5957	725.4	RPN	n-PFOS
M857.5969T370.5	857.5969	370.5	RPN	n-PFOS
M1131.6617T470.4	1131.6617	470.4	RPN	BMN

547