Current Environmental Health Reports (2021) 8:7–22
https://doi.org/10.1007/s40572-020-00298-x

SUSCEPTIBILITY FACTORS IN ENVIRONMENTAL HEALTH (Z LIEW AND K POLLITT, SECTION EDITORS)

Use of Untargeted Metabolomics to Explore the Air Pollution-Related Disease Continuum

Lan Jin[1] · Krystal J. Godri Pollitt[2] · Zeyan Liew[2] · Ana K. Rosen Vollmar[2] · Vasilis Vasiliou[2] · Caroline H. Johnson[2] · Yawei Zhang[1,2]

Accepted: 18 November 2020 / Published online: 9 January 2021

Abstract

Purpose of Review The purpose of this review is to summarize the application of untargeted metabolomics to identify the perturbation of metabolites or metabolic pathways associated with air pollutant exposures.

Recent Findings Twenty-three studies were included in this review, in adults, children, or pregnant women. The most commonly measured air pollutant is particulate matter smaller than 2.5 μm. Size-fractioned particles, particle chemical species, gas pollutants, or organic compounds were also investigated. The reviewed studies used a wide range of air pollution measurement techniques and metabolomics analyses. Identified metabolites were primarily related to oxidative stress and inflammatory responses, and a few were related to the alterations of steroid metabolic pathways. The observed metabolic perturbations can differ by disease status, sex, and age. Air pollution-related metabolic changes were also associated with health outcomes in some studies.

Summary Our review shows that air pollutant exposures are associated with metabolic pathways primarily related to oxidative stress, inflammation, as assessed through untargeted metabolomics in 23 studies. More metabolomic studies with larger sample sizes are needed to identify air pollution components most responsible for adverse health effects, elaborate on mechanisms for subpopulation susceptibility, and link air pollution exposure to specific adverse health effects.

Keywords Metabolomics · Air pollution · Particulate matter · Oxidative stress · Inflammation

Abbreviations

COPD	chronic obstructive pulmonary disease
IHD	ischemic heart disease
TRAP	traffic-related air pollution
O_3	ozone
PAHs	polycyclic aromatic hydrocarbons
PM	particulate matter
CO	carbon monoxide
NOx	nitrogen oxides
$PM_{2.5}$	particulate matter with diameter ≤ 2.5 μm
PM_{10}	particulate matter with diameter ≤ 10 μm
UFP	ultrafine particles (particulate matter of nano-scale size, < 0.1 μm)
$PM_{2.5–10}$	particulate matter with a diameter between 2.5 and 10 μm
TSP	total suspended particles
Pb-PAH	particle-bound polycyclic aromatic hydrocarbons
PNC	particle number concentration
BC	black carbon
EC	elemental carbon
OC	organic carbon
LDSA	lung deposited surface area
V	vanadium
1-OHP	1-hydroxypyrene
LC-MS	liquid chromatography coupled to mass spectrometry
HILIC	hydrophilic interaction liquid chromatography

Caroline H. Johnson and Yawei Zhang contributed equally to this work.

This article is part of the Topical Collection on *Susceptibility Factors in Environmental Health*

✉ Lan Jin
lan.jin@yale.edu

[1] Department of Surgery, School of Medicine, Yale University, New Haven, USA

[2] Department of Environmental Health Sciences, Yale School of Public Health, Yale University, New Haven, USA

GC-MS	gas chromatography coupled to mass spectrometry
C18	C18 hydrophobic reversed-phase chromatography
ESI	electrospray ionization
TOF	time-of-flight
HMBD	Human Metabolome Database
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC-MS/MS	liquid chromatography coupled to tandem mass spectrometry
PLS-DA	partial least squares discriminant analysis
OPLS-DA	orthogonal PLS-DA
ANOVA	analysis of variance
ANCOVA	analysis of covariance
ROS	reactive oxygen species
PUFA	polyunsaturated fatty acids
13-HODE	13-hydroperoxyoctadecadenoic acid
4-HNE	4-hydroxynonenal
FEV1	forced expiratory volume in 1 s
Pb-PAHs	particle-bound polycyclic aromatic hydrocarbons
AOA	adult-onset asthma
CCVD	cardio-cerebrovascular diseases

Introduction

Metabolomics is the agnostic or comprehensive assessment of low molecular-weight endogenous and exogenous metabolites in a biological sample. Knowledge of these compounds can be used to elucidate the molecular-level responses to a genetic alteration, or the effect of disease. They can also be used to monitor responses to various stimuli such as exposure to environmental and lifestyle factors or administration of a drug [1–3]. Metabolite abundance is regulated by various host processes such as protein-enzymatic reactions; thus, metabolomics sits at the end of the "omics cascade" within the central dogma of biology [3]. In addition to host processing, metabolites can be co-metabolized by the microbiome; a pool of these microbial metabolites is secondary metabolites, of which most are currently uncharacterized. As metabolites are the biochemical intermediates that carry out biological functions within the body, the metabolome closely reflects the phenotype or physiological state, compared to analysis of the genome, transcriptome, and proteome [3]. Recent developments in mass spectrometry have allowed for high-resolution measurement of tens of thousands of features in small volumes of biological samples, which enables an untargeted strategy to simultaneously analyze exogenous chemicals, their metabolites, and associated perturbations to the endogenous metabolome, allowing for an exposome-level evaluation [4].

Air pollutants have been associated with various health outcomes, such as cardiovascular diseases, respiratory diseases, cancer, central nervous system disorders, and adverse pregnancy outcomes [5–9]. However, the mechanisms that underlie these observed associations are not conclusive [10, 11]. Numerous studies have extensively investigated selected biomarkers of oxidative stress and inflammation, with a traditional, hypothesis-driven approach [12]. This approach, often focusing on a relatively small number of metabolites, cannot capture the full picture of biological responses to air pollution [11, 13]. To address this challenge, untargeted metabolomics has been incorporated into many recent mechanistic studies of air pollution. A recent book chapter reviewed the application of metabolic profiling in human, animal, and in vitro studies of air pollution, and discussed metabolites or pathways related to oxidative stress responses [13]; however, this review was based on a limited number of human studies ($n = 8$), and focused on oxidative responses only.

The objective of this review is to provide a comprehensive review of the emerging literature on air pollution and untargeted metabolomics (agnostic/global detection and relative quantitation of all small molecules in a sample), including assessments of study design, metabolites and pathways identified, discussion of methodological challenges, and to provide recommendations for future studies.

Study Populations and Study Designs

We conducted a search on human studies in PubMed using a combination of MeSH terms and keywords: (metabolomics[MeSH Terms] OR metabolome[MeSH Terms] OR metabonomics) AND (air pollution[MeSH Terms] OR particulate matter[MeSH Terms] OR nitrogen oxides[MeSH Terms] OR air pollution OR air pollutants OR polycyclic aromatic hydrocarbons). The last search was conducted on June 9, 2020. After screening the abstract and full text of 315 articles, we included a total of 23 studies on air pollution and untargeted metabolomics published between 2012 and 2019 (Fig. 1, Table 1) [10, 11, 14–34]. More details on these studies are in Table S1. Among these, 19 studies focused on adults aged 18 or older [10, 11, 14–22, 25–28, 31–34], one study on pregnant women [23], two studies on both elderly adults and children [29, 30], and one study on children and adolescents [24].

These 23 studies include 11 crossover designs [10, 11, 14–22], 10 cross-sectional studies [23–32], one case-control study, and one study using a combination of crossover and cross-sectional designs [34].

All participants in the crossover studies went through two to five exposure scenarios with varying expected air pollution levels. In these scenarios, participants were requested to conduct activities (e.g., walking, driving), or to use real or sham air purifiers. The order of purposefully designed exposure sessions was typically randomized, and any two exposure sessions were at least 7 days apart to avoid overlapping impacts

Table 1 Study design, air pollution exposures, and metabolomics analysis in the 23 reviewed studies

First author, year	Location	Number of participants	Study population	Air pollution metrics	Sample collection timing with respect to exposure	Short- or Long-term effects	Analytical platform	Significant Features (listing metabolites with level 1 and 2 confidence, if available)[2]
Crossover design Blood samples								
van Veldhoven 2019 [10]	London, UK; Barcelona, Spain	London: 60; Barcelona: 30	London: healthy adults, adults with COPD but no IHD, or adults with IHD but no COPD; walk on Oxford st (polluted by diesel vehicles) or walk in Hyde Park (traffic-free); Barcelona: healthy, non-smoking, non-medication, without high occupational exposures adults; physical activity or at rest at locations with low or high level of pollution	London: $PM_{2.5}$, PM_{10}, UFP, BC, NO_2; Barcelona: PM_{10}, $PM_{2.5}$, NOx, $PM_{2.5-10}$	London: 3 samples collected: 2 h before, 2 h after, and 24 h after each exposure session Barcelona: 1 sample collected ~ 7 h after each exposure session	Short-term (2 h)	UHPLC-QTOF MS	London: Level 1: phenylalanine (−), caffeine (+), both related to NO_2 Barcelona: Level 1: L-Tyrosine (+), related to $PM_{2.5-10}$, PM_{10}
Mu 2019 [14]	Beijing, China	26	Non-smoking adults without prior medical history of cancer or serious immunological or chronic respiratory diseases; Exposure scenarios: before, during, or after the Beijing Olympics	PM_1, $PM_{2.5}$, PM_7, PM_{10}, TSP	1 sample collected after each exposure scenario (i.e., before, during, or after the Olympics)	Short-term (29 days)	GC/MS; UHPLC-MS/MS	Level 1 and 2: 67 metabolites; 1st module ($n = 33$) are lipids (e.g., icosapentaenoic acid, stearic acid); 2nd module are primarily dipeptides (e.g., isoleucylglycine), and 4 other classes (e.g., hypoxanthine, 12-hydroxyeicosatetraenoic acid); 3rd and 4th modules are unknown
Liang 2019 [11]	Atlanta, USA	60	Adults, with or without asthma (Atlanta Commuters Exposure [ACE] study); Exposure sessions: commuting on highway, or surface street or visiting clinic	pb-PAH, PNC, $PM_{2.5}$ (3 factors identified using source apportionment based on 58 measured pollutant species); 19 PM metal species; 4 carbonaceous species: BC, EC, OC, WSOC	2 samples: before and 8-h after each exposure session	Short-term (2 h)	LC-HRMS	Level 1: 45 metabolites (92% endogenous metabolites related to oxidative stress, inflammatory responses, or nucleic acid damage; xenobiotics were identified); Generally, anti-inflammatory molecules (−); oxidative or pro-inflammatory metabolites (+)
Ladva 2018 [15]	Atlanta, USA	60	Adults, with or without asthma (ACE study); Exposure sessions: commuting on highway, side street, or visiting clinic	$PM_{2.5}$, BC, pb-PAHs, PNC, PM metal (Al, Pb, Fe), OC, WSOC, elemental concentrations, carbon fraction concentrations	2 samples: before and after each exposure session	Short-term (2 h)	Dual-chromatography - MS	Level 3/4: 110 features; 20-OH-leukotriene B4 (putative) is a top predicted metabolite in pathway analysis (−)
Shen 2018 [16]	Welding school in MA, USA	52	Boilermakers; Exposure scenarios: before and after a welding workshop	Measured in a previous study (Wei et al., 2013 [17])	1 sample after each exposure scenario	Short-term (5 h)	UPLC-MS/MS	Level 1: 113 metabolites (78 upregulated; 35 downregulated); Mainly involved in the lipid pathway (glucocorticoid class (cortisol, corticosterone, and cortisone), acylcarnitine class, and DiHOME species (9,10DiHOME and 12,13-DiHOME)), amino acid utilization (isoleucine, proline, and phenylalanine), and S-(3-hydroxypropyl) mercapturic acid (3-HPMA).
	Netherlands	31				Short-term (5 h)	UHPLC-QTOF MS	

Table 1 (continued)

First author, year	Location	Number of participants	Study population	Air pollution metrics	Sample collection timing with respect to exposure	Short- or Long-term effects	Analytical platform	Significant Features (listing metabolites with level 1 and 2 confidence, if available)[2]
Vlaanderen 2017 [18]			Healthy, non-smoking college students; Exposure sessions: cycling 20 min every hour at underground train station, a continuous traffic location, a stop-and-go traffic location, a farm, urban background site	$PM_{2.5}$, PM_{10}, $PM_{2.5-10}$, PNC, EC, trace metals, secondary organic and inorganic components, endotoxin content, NO_2, NOx, O_3, PM oxidative potential	3 samples collected 2 h before, 2 h and 18 h after each exposure session			Level 1: tyrosine, guanosine, hypoxanthine (+)
Li 2017 [19]	Shanghai, China	55	Healthy, non-smoking college students; Exposure sessions: real purification or sham purification	$PM_{2.5}$	1 sample collected after each exposure session	Short-term (9 days)	GC-MS; UHPLC-MS	Level 3: 97 features; Stress hormones: glucocorticoids (+), catecholamine (+), melatonin (+); amino acids: arginine (+), leucine (+), histamine (+), threonine (+), tryptophan (+); free fatty acids and metabolites involved in fatty acid oxidation: carnitine (+), acylcarnitine (+), hydroxy fatty acids (+), acetone (+); membrane phospholipids and their hydrolyzed products: lysolipids (+), 2 eicosanoids (+)
Miller 2016 [20]	NC, USA	24	Healthy young adults; Exposure sessions: 0.3 ppm O3 or filtered air	Predetermined O_3 exposure level	1 sample collected after each exposure session	Short-term (2 h)	LC-MS; GC-MS	Level 1: 121 features (85 increased, 36 decreased) primarily in lipid metabolism (including stress hormones)
Wei 2013 [17]	Welding school in MA, USA	11	Non-smoking, non-diabetic boilermakers at an apprentice welding school; Exposure scenarios: before and after a welding workshop	$PM_{2.5}$	1 sample at the end of each exposure scenario	Short-term (5 h)	LC-MS; GC-MS	Level 3: eicosapentaenoate (−), docosapentaenoate (−), docosapentaenoate (−)
Urine samples								
Chen 2019 [21]	Shanghai, China	45	Healthy, non-smoking college students; Exposure sessions: real purification or sham purification	$PM_{2.5}$	1 sample collected after each exposure session	Short-term (9 days)	UPLC-QTOF MS	Level 3: 16 lipids, 5 purine metabolites, 2 neurotransmitters, 3 coenzymes; others were metabolites of bile acids, corticosteroids, amino acids
Zhang 2019 [22]	Beijing, China	39	College students (healthy young adults, non-smokers, and did not take any medicine); Traveling in subway system, with either a respirator intervention phase (RIP) with face mask or no intervention phase (NIP)	$PM_{0.5}$, PM_1, $PM_{2.5}$, PM_5, PM_{10}	1 sample collected at the end of each exposure session	Short-term (4 h)	UPLC/QTOF MS/MS	Level 2: 11 metabolites Male: molybdopterin precursor Z (+), 8-hydroxydeoxyguanosine (8-OHdG) (+); Female: N-acetylglutamine (−), homovanillic acid (−)
Cross-sectional design								
Blood samples								
Yan 2019 [23]	The Central Valley, CA, US	160	Pregnant women	TRAP: CO, NOx, $PM_{2.5}$; classified into "highly exposed" (these 3 correlated pollutants > 75%) and "low exposed"	1 sample collected in mid-pregnancy (~ 16th week), which is after exposure during the 1st trimester	Sub-chronic (3 months)	LC-HRMS	Level 1: serine (−), creatinine (+), L-histidine (−), myo-inositol (+), heptadecanoic acid (−), linoleic acid (−)

Table 1 (continued)

First author, year	Location	Number of participants	Study population	Air pollution metrics	Sample collection timing with respect to exposure	Short- or Long-term effects	Analytical platform	Significant Features (listing metabolites with level 1 and 2 confidence, if available)[2]
				(pollutants levels < 35%)				
Chen 2019 [24]	Taiwan	107	Children and adolescents (age 9–15) living near or far from a large petrochemical complex	Internal exposures to V and PAHs (1-OHP); Urine concentrations of As, Cd, Cr, Ni, Pb, V, Hg;	1 sample collected at the same time when internal exposures were measured in urine samples	Long-term (unspecified, exposures in participants who have lived in the area for more than 5 years)	UHPLC-QTOF	Level 3: 10 features; ketoleucine (+), carnitine (+), isovalerylcarnitine (+), aspartic acid (+), octenoyl-L-carnitine (+); pyroglutamic acid (−), adenosine monophosphate (AMP) (−); inosinic acid (IMP) (−), oxoglutaric acid (−), malic acid (−)
Walker 2018 [25]	Boston, USA	59	Non-smokers, non-Hispanic white adults (≥ 40), recruited from residents living in neighborhoods near (≤ 500 m) or far (≥ 1000 m) from highway (Community Assessment of Freeway Exposure and Health (CAFEH) study)	UFP (PNC); classified into low (mean 16,000 particles/cm^3) and high (mean 24,000 particles/cm^3) exposures based on annual averages	1 sample collected from 2009 to 2010, while exposure was estimated from 2009 to 2010	Long-term (annual average)	LC-HRMS	Level 1: L-aspartic acid (−); alpha/gamma-linolenic acid (−); methionine sulfoxide (+); L-cystine (+); Alanine (+); L-lysine (+); L-glutamine (+); cotinine (−); L-aspartic acid (−); L-glutamic acid (−); L-threonine (+)
Yuan 2016 [26]	Taiwan	160	Residents living near or far from a petrochemical complex (average 50 years old)	Internal exposure to V and PAHs	1 sample collected at the same time when internal exposures were measured in urine samples	Long-term (unspecified, exposures in participants who have lived in the area for more than 5 years)	NMR	Level 3: 18 features; lipid (-CH=CH-) (−); isoleucine (+), alanine (−), glutamine (−), phenylalanine (−), alpha- & beta-glucose (−), alpha-glucose (−), N-acetyl glycoprotein (−)
Menni 2015 [27]	London, UK	523	Participants in TwinsUK cohort who had PM exposure estimates	PM_{10}, $PM_{2.5}$	2 samples collected from 2004 to 2010, while exposures were estimated for 2003–2010	Long-term (annual average)	UPLC-MS/MS; GC-MS	Level 1: 8 metabolites; amino acids: asparagine (−), glycine (−), N-acetylglycine (−), serine (−); carbohydrates: glycerate (−); cofactor and vitamin: threonate (−), alpha-Tocopherol (−); Xenobiotics: benzoate (−)
Urine samples								
Huang 2018 [28]	Beijing, China	41	Elderly COPD patients and their healthy spouses, without any metabolic diseases	Indoor $PM_{2.5}$; 24 element components	1 sample collected after the exposure assessment	Short-term (22 h)	UPLC/MS	Level 2: Uric acid (−), glyceric acid 1,3-biphosphate (−), methyluric acid (−), indolelactic acid (+), 5-phosphoribosylamine (+), dopamine 4-sulfate (−), 4-pyridoxic acid (−)
Chen 2017 [29]	Taiwan	252	Residents near or far from a petrochemical complex (children or elderly)	Ambient V, 5 PAHs (pyrene, fluoranthene, dibenzo[a,-h]anthracene, benzo[k]fluoranthen-e, benzo[a]anthracene) concentrations; Internal exposures to V and PAHs	1 sample collected upon recruitment from 2009 to 2012, while air samples were collected from 2009 to 2012	Long-term (unspecified, exposures in participants who have lived in the area for more than 5 years)	GC*GC-TOFMS;	Level 2: 76 significant metabolites primarily in alanine, aspartate, glutamate metabolism, phenylalanine metabolism, tryptophan metabolism, glycine, serine, threonine metabolism, glutamate metabolism, aminoacyl-tRNA biosynthesis
Wang 2015 [30]	Shanxi, China	566	Healthy volunteers: 238 elderly non-smokers, 114 elderly		1 sample collected during the study period (the	Long-term (not specified, exposures	UHPLC-MS	Level 1: 5 metabolites; level 2: 13 metabolites;

Table 1 (continued)

First author, year	Location	Number of participants	Study population	Air pollution metrics	Sample collection timing with respect to exposure	Short- or Long-term effects	Analytical platform	Significant Features (listing metabolites with level 1 and 2 confidence, if available)[2]
			smokers, 214 children, living in an area polluted by a large coking plant, and the in a clean area that is 70 km away from the polluted area	Internal exposure to PAHs (9 urine PAH metabolites)	details of air sampling were not specified)	in living environment)		Uric acid (−); pyroglutamic acid (+); 3-Methylhistidine (+); azelaic acid (+); decenedioic acid (+); hydroxytetradecanedioic acid (+); decenedioylglucuronide (+); 11 medium-chain acylcarnitines (+)
Kuo 2012 [31]	A shipyard in Taiwan	51	Full time employees: 35 male welders and 16 make office workers (45–64 years old)	Fine PM; PM metals (Cr, Ni, Mg)	1 sample during the study period (the length of exposure assessment was not specified)	Long-term (not specified, long-term occupational exposures)	NMR	Level 2: creatine (−), taurine (+), trimethylamine-N-oxide/betaine (+), acetone (+), creatinine (+), glycine (+), gluconate (+), hippurate (+), S-sulfocystein (+), serine (+)
Blood and saliva samples								
Liang 2018 [32]	Atlanta, USA	54	Healthy, non-smoking college students, recruited from dorms near (~20 m) or far from highway (~ 1.4 km) (Dorm Room Inhalation to Vehicle Emission [DRIVE] study)	Indoor or outdoor BC, CO, NO, NO_2, NOx, $PM_{2.5}$	4 plasma (monthly) and 12 saliva (weekly) from Sep 8–Dec 15, 2014, while sampling was conducted at the same time	Sub-chronic (3 months)	LC-HRMS	Level 1: arginine (−), histidine (−), gamma-linolenic acid (−), hypoxanthine (+); adenosine 5′-monophosphate (+); bis(2-ethylhexyl) phthalate (+); cytosine (+); proline (−), s-Lactate glyceraldehyde (−), 3-Hydroxykynurenine (−)
Case-control study								
Blood samples								
Jeong 2018 [33]	Swiss; Italy	Swiss: 335; Italy: 386	Swiss: self-reported asthma cases, and controls with no asthmatic symptoms, no smoking for over 10 years; Italy: cardio-cerebrovascular disease (CCVD) cases and controls, who were never smokers or former smokers for over 1 year	Swiss: $PM_{2.5}$ (back-extrapolated up to 7 years), UFP, NO_2; Italy: $PM_{2.5}$, UFP, NO_2 (back-extrapolated up to 23 years)	Swiss: 1 sample collected from 2010 to 2011 when annual/biennial averages were estimated; Italy: 1 sample collected from 1993 to 1998 when annual averages were estimated	Long-term (annual or biennial averages at the biological sample collection)	UHPLC-QTOF-MS	Level 1: linoleate (+ for PM2.5 and PNC); octanoic acid (− for PM2.5); sphingosine (+ for PNC); L-carnitine (+ for NO2)
Combined crossover and cross-sectional design								
Exhaled breath condensate								
Maniscalco 2018 [34]	A factory that produces friction systems in Italy	30	20 blue-collar workers ("exposed" not wearing a mask, "unexposed" wearing mask), and 10 age-matched white-collar workers; healthy, non-smoking	Inhalable fractions of dust, phenol, formaldehyde, VOCs	Blue-collar workers: 2 samples collected before and 30 min after the work shift; White-collar workers: 30 min after work shift	Short-term (a work shift within a day)	NMR	Level 3/4: 8 features; Fatty acids: propionate (+), isopropanol (+), methanol (+), 1,2-propanediol (+), 3-hydroxyisobutyrate (+), phenylalanine (+); Alcohols: isocaproate (−), 3-hydroxybutyrate (−)

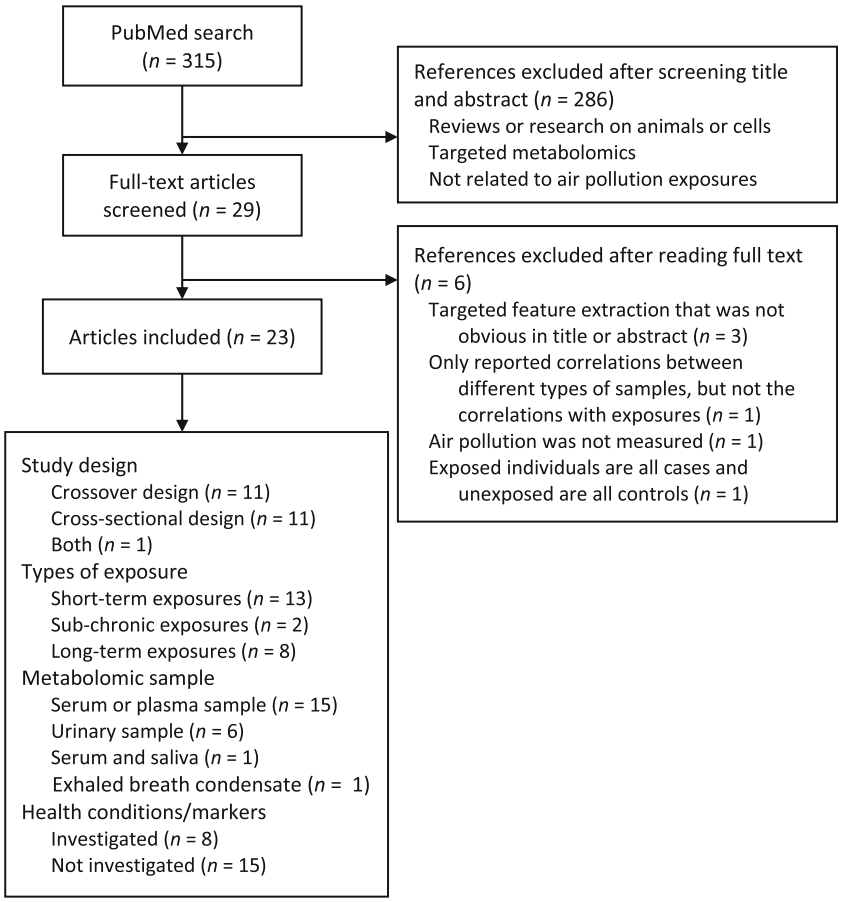

Fig. 1 Flow diagram of literature search and selection

of the exposures. In the cross-sectional or case-control studies, the selected participants were considered to capture the expected variation in air pollutant exposures of the study population. For example, residents living in neighborhoods near ($\leq$ 500 m) or far ($\geq$ 1000 m) from a highway were recruited in the Community Assessment of Freeway Exposure and Health (CAFEH) study in Boston, USA [25].

All 11 crossover studies and the one study combining crossover and cross-sectional designs investigated short-term effects of air pollution exposures (i.e., hours or days) on metabolomic perturbations. As participants served as their own comparisons in crossover designs, the changes in metabolite levels between exposure scenarios were investigated. All studies measured metabolites at the end of exposure scenarios. Five studies also measured metabolites before each scenario [10, 11, 14, 15, 18]; that is, these five studies investigated the changes in metabolites between exposure scenarios, as well as the changes within each scenario.

Among the 10 cross-sectional studies and the one case-control study, 10 investigated sub-chronic (i.e., multiple months) or long-term effects (i.e., $\geq$ 1 year). Nine collected one biological sample for each participant, and thus investigated the between-participant variations and related variations in metabolic profiles. The other two studies on sub-chronic or long-term effects had repeated measures of metabolites to also account for within-participant changes in metabolites over time [27, 32]. In the 10 sub-chronic or long-term studies, the timing of biological sample collection with respect to the exposure windows varied: five studies collected one or multiple samples during the exposure windows [25, 27, 31–33], one study collected one sample after the exposure window [23], and four studies involving exposure biomarkers collected the samples for measuring metabolites and internal exposures at the same time [21, 26, 29].

Among the 23 studies reviewed in this paper, eight also investigated metabolites related to health conditions such as adult-onset asthma, ischemic heart disease (IHD) [10, 28, 33], or had measured health markers of early health effects such as lung function, blood pressure, and biomarkers of oxidative stress or inflammation [18, 21, 22, 27, 29] (Table 2). These studies used case-control or cross-sectional study designs. All but one of the eight studies collected biological samples after or at the same time of the assessment of health outcomes. The other one study selected cases and controls from two cohorts: (1) in Italy, biological samples were collected before the identification of incident cardiocerebrovascular disease (CCVD); and (2) in Switzerland, biological samples were collected 7–10 years after identifying cases of adult-onset asthma (AOA) [33].

Table 2 Summary of studies investigating the metabolites or metabolic pathways related to both air pollution exposures and health outcomes

First author, year	Health outcomes	Sample collection timing with respect to health outcomes	Findings
van Veldhoven 2019 [10]	COPD, IHD	Samples collected after identifying health outcomes (diseases presented upon recruitment)	Out of the 15 compounds associated with NO_2, two were also associated with COPD (data not shown).
Chen 2019 [24]	4 oxidative stress biomarkers: 8-OHdG, HNE−MA, 8-isoPF2α, and 8-NO2Gua; 31 acylcarnitines	Samples collected at the same time when biomarkers were evaluated (in urine samples)	10 potential metabolites associated with higher exposures to industrial emissions were also related to elevated oxidative stress and deregulated serum acylcarnitiens: inosine monophosphate and adenosine monophosphate (purine metabolism), malic acid and oxoglutaric acid (citrate cycle), carnitine (fatty acid metabolism), and pyroglutamic acid (glutathione metabolism).
Zhang 2019 [22]	Electrocardiogram (ECG) parameters and ambulatory blood pressure (BP) monitored during the whole riding period	Samples collected shortly after the measurements of health markers (i.e., health markers were measured during the exposure session, while samples were collected at the end of each session)	8-OHdG was significantly associated with both heart rate variability indices, heart rate, and size-fractioned PM ($PM_{0.5}$, PM_1, $PM_{2.5}$, PM_5, PM_{10}, TPM), but only among males. Decreased 8-OHdG and increased prolyl-arginine were observed while wearing masks, compared to not wearing masks. And they both were associated with ECG parameters. These findings were only observed in males.
Jeong 2018 [33]	Switzerland: adult-onset asthma (AOA); Italy: cardio-cerebrovascular diseases (CCVD)	Switzerland: Samples collected after case identification; Italy: samples collected before case identification	Overlapping exists among significant pathways: linoleate metabolism and glycerophospholipid metabolism linking UFP to AOA; glycosphingolipid metabolism linking UFP to CCVD; and carnitine shuttle linking NO_2 to CCVD; No metabolites were associated with both air pollution exposure and health outcomes (i.e., AOA or CCVD)
Huang 2018 [28]	COPD	Samples collected after case identification	$PM_{2.5}$- and COPD-related metabolic biomarkers were correlated. These biomarkers include uric acid, altered glucose and dopamine metabolism.
Vlaanderen 2017 [18]	3 parameters of lung function: FENO, FVC, FEV1; 6 markers of inflammation and coagulation measured in peripheral blood: IL-6, fibrinogen, tPA/PAI−1, VWF, platelets, CRP	Samples collected at the same time when health markers were evaluated	At 2 h after exposure session, 21 features were associated with > 1 air pollutant and to > 1 health marker. The reported associations were primarily for FEV1 and fibrinogen. At 18 h after exposure session, 62 features were associated with >1 air pollutant and to >1 health marker. The observed associations were primarily for FEV1, fibrinogen, or $PM_{2.5}$–10-SO_4. These metabolites are primarily involved in tyrosine metabolism, urea cycle/amino group metabolism, and N-Glycan degradation.
Chen 2017 [29]	Oxidative stress biomarkers: 8-OHdG, HNE-MA, 8-isoPF2α, and 8NO2Gua	Samples were used for measuring both metabolome and health markers (at the same time)	In children: tryptophan metabolism (e.g., tryptophan, indole-3-acetamide), phenylalanine metabolism (e.g., phenylalanine, hippuric acid, 4-hydroxy benzoic acid), and alanine, aspartate, and glutamate metabolism (aspartic acid) linking industrial emissions to oxidative stress;

Table 2 (continued)

First author, year	Health outcomes	Sample collection timing with respect to health outcomes	Findings
			In elderly: glycine, serien, and threonine metabolism (e.g., threonine, serine, glyceric), and alanine, aspartate, and glutamate metabolism (aspartic acid) linking industrial emissions to oxidative stress
Menni 2015 [27]	FEV1 and FVC	Samples were collected at the same time when lung functions were measured	Eight metabolites were significantly associated with both lung function and PM (i.e., $PM_{2.5}$ and PM_{10}): asparagine, glycine, N-acetylglycine, serine, glycerate, threonate, alpha-Tocopherol, benzoate. Vitamin E showed the strongest associations with both $PM_{2.5}$ exposure and FEV1.

To find the metabolic pathways linking air pollution to health outcomes, these studies first investigated the metabolites or pathways associated with air pollution, and those associated with health outcomes, separately. Then, they investigated whether there was overlap or correlation in the metabolites or pathways related to air pollution and those related to health outcomes. Three of the eight studies explicitly referred to this method of finding overlapping metabolites as a "meet-in-the-middle" approach [21, 29, 33].

Air Pollution Exposure Assessment

In 13 studies on short-term effects, 12 studies used portable or stationary monitors [10, 11, 14–19, 21, 22, 28, 34], and one study used human exposure chamber to challenge study participants with ozone (O_3) or filtered air [20]. Among the 10 studies on sub-chronic or long-term effects, five studies used dispersion, land use regression (LUR), or kriging models to estimate air pollutant exposure levels [23, 25, 27, 29, 33]; four studies measured exposure biomarkers of ambient polycyclic aromatic hydrocarbons (PAHs) or vanadium (V) [24, 26, 29, 30]; and one study used portable or stationary monitors [31].

Air pollution mixtures from different sources differ in the chemical composition and size distribution of particulate matter (PM), which could influence the decision on which pollutants or components to measure in these studies. All but one of the 14 studies on ambient air pollution or traffic-related air pollution (TRAP) measured PM with diameter ≤ 2.5 μm ($PM_{2.5}$) mass concentration [10, 11, 14, 15, 18, 19, 21–23, 25, 27, 28, 32, 33]. Five of the eight TRAP studies also measured PM components related to combustion, such as elemental carbon (EC), organic carbon (OC), and particle-bound polycyclic aromatic hydrocarbons (pb-PAH) [10, 11, 15, 18, 32]. Among the four studies on industrial emissions from petrochemical or coking plants, external or internal exposures to PAHs and/or heavy metals were investigated among residents [21, 26, 29, 30]. Pollutants measured in occupational settings varied widely depending on the anticipated exposures [16, 17, 31, 34].

Studies in China (Beijing and Shanghai) [14, 19, 21, 22, 28] tended to show higher $PM_{2.5}$ mass concentrations than the studies performed in Europe (UK, Switzerland, and Italy) [27, 33] or the USA [32]. For the four studies on industrial emissions, the urinary 1-hydroxypyrene (1-OHP) level in the elderly residing near a coking plant in Shanxi, China (1.42 μg/g creatinine for non-smokers and 3.13 μg/g creatinine for smokers), was higher than the levels in elderly adults residing near the largest petrochemical complex in Taiwan (0.42 μg/g creatinine) [21, 26, 29, 30]. In addition, the 1-OHP levels in the exposed elderly were higher than the levels in exposed children (0.25 μg/g creatinine) in Taiwan [29].

Untargeted Metabolomics Analysis

From the 23 studies we examined, 15 studies analyzed the serum or plasma metabolome [10, 11, 14–21, 23, 25–27, 33], six analyzed the urinary metabolome [21, 22, 28–31], one analyzed both serum and saliva [32], and one analyzed exhaled breath condensate [34].

For analytical platforms, 15 studies used liquid chromatography mass spectrometry (LC-MS) to measure metabolites [10, 11, 15, 16, 18, 21–25, 28, 30, 32, 33, 35], four studies used both gas chromatography (GC)-MS and LC-MS [14, 19, 20, 27], three studies used nuclear magnetic resonance (NMR) spectroscopy [26, 31, 34], and one study used two-dimensional GC × GC-time-of-flight (ToF)-MS [29]. The combination of LC and GC has been shown to achieve

complementary coverage [36]. LC-MS analyses can be heterogenous between research labs due to the many different configurations of columns (e.g., hydrophilic interaction liquid chromatography (HILIC), reversed-phase liquid chromatography (RPLC)), the type of ionization (e.g., electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI)), the mode of ionization (i.e., positive and/or negative), and the type of mass analyzer (e.g., time-of-flight (ToF), quadrupole, ion trap).

All studies preprocessed their metabolomics data to align, filter, and deconvolute the data using a wide range of software. After obtaining the deconvoluted peak area/heights for each metabolic feature, studies narrowed down the features for further analysis, by selecting those that were detected in at least 10–80% of samples. Such a wide range of cut-off points for feature selection could lead to variation between data reported from different studies; keeping features detected in $\geq$ 10% samples might be more likely to find more significant features than a study with more strict inclusion criteria.

As most of the studies (20 out of 23) used MS, in this review, we only discuss the common methods for annotating features in MS studies. All of the studies searched the mass-to-charge ratios (*m/z*) of each feature against metabolite, and metabolic pathway libraries (e.g., Human Metabolome Database (HMDB) [37], Kyoto Encyclopedia of Genes and Genomes (KEGG) [38], METLIN [39]). Five studies conducted additional validation using LC coupled to tandem mass spectrometry (LC-MS/MS), to fragment the metabolic features and confirm the compound identity by comparing MS/MS spectra to those available in the public libraries [18, 22, 28, 29, 31]. In 12 studies, the spectra were matched against an authentic chemical standard analyzed on the same analytical platform. To do this, the *m/z* of the parent ion and MS/MS fragments were compared, as well as the retention time to perform the highest level of confidence in metabolite identification [10, 11, 14, 16, 20, 23, 25, 27, 30, 32, 33, 35].

For statistical analysis, supervised partial least squares discriminant analysis (PLS-DA) or orthogonal partial least squares discriminant analysis (OPLS-DA) was used in 11 studies to determine which metabolites drive the separation of high- or low-exposure groups [19, 21, 23–26, 28–30, 34, 35]. The key features that drive the separation of the exposure groups were identified by variable importance in projection (VIP) scores (cut-offs, >1, >1.5, or >2). Network analysis and principal component analysis were also used to identify groups of metabolites that were correlated with each other [14, 17]. These methods are dimension reduction tools for high-dimensional data. To obtain effect direction, fold changes, *t* test, analysis of variance (ANOVA), analysis of covariance (ANCOVA), or regression models (e.g., mixed effect models) were used in these studies. As a large number of metabolites were investigated in these analyses, multiple comparison correction methods, such as Benjamini-Hochberg false discovery rate adjustment or Bonferroni correction, were used to account for multiple comparisons.

Metabolomics Methods Assessment

We developed an assessment tool to compare methods between studies, based on recommendations from the Metabolomics Standards Initiative, and other related studies and quality assessment tools [40–46]. We constructed five metrics for scoring the studies (Fig. 2). Higher scores indicate either more complete reporting on the experimental design and methods for metabolomics analysis, or more rigorous exposure assessment. The scores are not final judgments on the scientific validity of a study, but rather an attempt to understand where reporting could be more transparent, so that findings can be robustly compared across heterogeneous studies. This is a novel tool designed for evaluation of metabolomics analysis of air pollution, but it can be adapted to any metabolomics study with human subjects. Most of the studies we reviewed had moderate to high scores (Table 3). Fifteen of the 23 studies did not report how missing values were treated [11, 17, 19, 21, 22, 24–32, 34]. Ten studies conducted more rigorous exposure assessment by accounting for time-activity patterns or requested participants to perform standardized activities, and/or had taken into account temporality, by having the exposure windows preceding the metabolite measurements [10, 11, 14–16, 18–23, 25, 28, 31, 32, 34]. Most of the studies reviewed were found to have high confidence levels for chemical annotation (level 1 or 2).

Perturbations of Metabolites or Pathways Related to Air Pollutant Exposure

While untargeted metabolomic analysis of biological samples can detect thousands of metabolic features, only a small number of these features correlate with air pollution exposure (ranging from three to 121 across the reviewed studies), and could be positively identified (MS/MS spectra matched to authentic chemical standards or spectra in the libraries, i.e., level 1/2, Fig. 2). The detected metabolites were mostly endogenous metabolites (i.e., lipids, amino acids, carbohydrates, nucleotides, steroids, cofactors, and vitamins). Few studies identified xenobiotics, such as PAH metabolites (i.e., catechol, 3-(2-hydroxyphenyl)propanoate, naphthylamine, nicotine metabolites (i.e., cotinine), and benzoate [11, 25, 27]. These xenobiotics were identified through an untargeted process. Pathway analyses identified 1–50 enriched metabolic pathways.

All 23 studies detected perturbations in lipid levels or pathways related to lipid metabolism, primarily as responses to oxidative stress and/or inflammation. Air pollutant exposures were associated with higher levels of reactive oxygen species

Table 3 Metabolomics methods assessment for reviewed studies

Author year	1. Study participants described	2. Exposure assessment and study design	3. Sample collection and analysis	4. Data processing workflow	5. Certainty of metabolite identification	Total points
Crossover design						
Blood samples						
Miller 2016 [20]	0	1	0	0	0	6
van Veldhoven 2019 [10]	0	1	0	0	0	6
Vlaanderen 2017 [18]	0	1	0	0	0	6
Mu 2019 [14]	0	1	0	0	0	6
Liang 2019 [11]	0	1	0	− 1	0	5
Ladva 2018 [15]	0	1	0	0	−1	5
Shen 2018 [16]	0	− 1	0	0	0	4
Li 2017 [19]	0	1	0	− 1	− 1	4
Wei 2013 [17]	0	1	0	− 1	− 1	4
Urine samples						
Zhang 2019 [22]	0	1	0	− 1	0	5
Chen 2019 [24]	0	0	0	− 1	− 1	3
Cross-sectional design						
Blood samples						
Yan 2019 [23]	0	1	0	0	0	6
Walker 2018 [25]	0	1	0	− 1	0	5
Chen 2019 [21]	0	1	0	− 1	− 1	4
Menni 2015 [27]	0	0	0	− 1	0	4
Yuan 2016 [26]	0	0	− 1	− 1	− 1	2
Blood and saliva						
Liang 2018 [32]	0	1	0	− 1	0	5
Urine samples						
Huang 2018 [28]	0	1	0	− 1	0	5
Kuo 2012 [31]	0	1	0	− 1	0	5
Chen 2017 [29]	0	0	0	− 1	0	4
Wang 2015 [30]	0	0	0	− 1	0	4
Case-control study						
Blood samples						
Jeong 2018 [33]	0	0	0	0	0	5
Combined crossover and cross-sectional design						
Exhaled breath condensate						
Maniscalo 2018 [34]	0	1	0	− 1	− 1	4

(ROS) or free radicals that cause cellular membranes to break into free fatty acids. Polyunsaturated fatty acids (PUFA), such as linoleic acid and arachidonic acid, can be oxidized, leading to increased pro-inflammatory metabolites, such as leukotrienes and prostaglandins [11, 14, 15, 21, 23, 31, 32]. A lipid peroxidation biomarker, 4-hydroxynonenal (4-HNE), was positively associated with UFP in the Boston CAFEH study [25]. Linolenic acid, a downstream metabolite of PUFA with anti-inflammatory effects (inhibits the biosynthesis of leukotriene B4), showed negative association with TRAP in Boston, Atlanta, and California, in the USA [25, 32].

Perturbations to amino acids, metabolites in purine metabolism, and acylcarnitines that are involved in energy metabolism were also commonly reported (Table 4). Histidine, with anti-inflammatory effects, showed consistent inverse associations with air pollution exposures, including TRAP, $PM_{2.5}$, and emissions from a petrochemical plant [11, 23, 25, 26, 32]. Uric acid, a powerful antioxidant and end product of purine metabolism, which is measured in urine samples, was negatively associated with air pollution exposures, including $PM_{2.5}$ and coking plant emissions [21, 28, 30]. The effect directions of other metabolites,

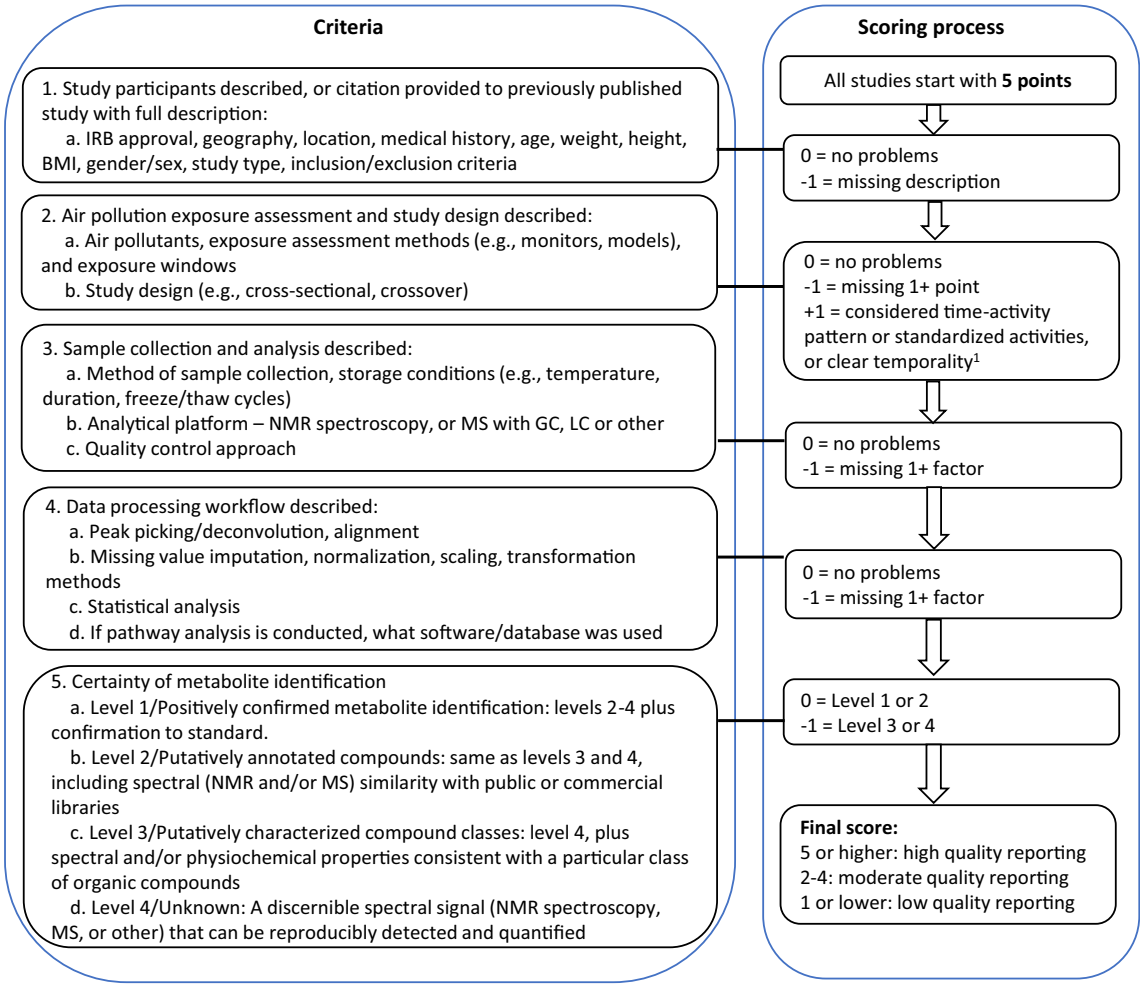

Fig. 2 Scoring process for metabolomics methods assessment ([1]clear temporality means that the air pollution exposures precede the metabolite measurements)

including oxidants or antioxidants, were variable between studies (Table 4).

Three studies reported perturbations in metabolites in steroid metabolic pathways (e.g., glucocorticoid metabolism) [16, 19, 21]. A randomized, double-blind crossover trial in Shanghai, China, implemented functional or sham air purifiers in dorms for 9 days with a 12-day washout period, and observed that stress hormones measured in blood serum (i.e., glucocorticoids (cortisone and cortisol), catecholamine (epinephrine and norepinephrine), and melatonin) were associated with higher $PM_{2.5}$ mass concentration exposures suggesting a response from the central nervous system by activation of the hypothalamus-pituitary-adrenal axis [19]. In a randomized crossover study, 24 participants were exposed to O_3 and filtered air in human chambers during two clinical visits that were at least 2 weeks apart. O_3 was significantly associated with elevated cortisol and corticosterone, consistent with studies in rodent models [20]. However, a study on boilermakers at an apprentice welding school in MA, USA, showed that glucocorticoids (cortisol, cortisone, corticosterone) were significantly lower after a 5-h shift in the welding workshop (higher exposure to $PM_{2.5}$ and PM enriched with heavy metals), compared to before their shift commenced [16].

Current evidence cannot conclude which air pollutant or chemical component of PM has stronger impacts on metabolic perturbations. Some PM components exhibited stronger associations with metabolites within studies, which might suggest exposure contributions from different emission sources, specific to each study area. In the Atlanta Commuters Exposure (ACE) study on TRAP, EC and V showed the largest number of associations with metabolite features (i.e., unconfirmed metabolites), among various $PM_{2.5}$ components [11]. Specifically, EC and V had 802 and 762 significantly associated features, respectively, while $PM_{2.5}$ was associated with just 215 significant features [11]. On the other hand, compared to other $PM_{2.5}$ components, S showed the strongest associations with 5-phosphoribosylamine, and S and V were related to the largest decline in 4-pyridoxic acid, among the elderly residents in

Table 4 Summary of commonly detected metabolites (detected in at least three studies) and effect directions

Metabolites	Air pollution exposures	Effect directions	First author year
Histidine (anti-inflammatory effects)	TRAP (high vs. low exposure)	Negative	Yan 2019 [23]
	EC, V	Negative only in asthma group	Liang 2019 [11]
	$PM_{2.5}$	Negative	Liang 2018 [32]
	Petrochemical plant emissions (high vs. low exposure)	Negative	Yuan 2016 [26]
Arginine (Related to endothelial function, inflammation, and airway hyperresponsiveness)	UFP	Positive	Walker 2018 [25]
	$PM_{2.5}$	Positive	Li 2017 [19]
	EC and V	Negative only in asthma group	Liang 2019 [11]
	BC and NOx	Negative	Liang 2018 [32]
Glutamate (a precursor to the antioxidant glutathione)	$PM_{2.5}$	Positive	Mu 2019 [14]
	EC, OC	Negative	Liang 2019 [11]
	UFP	Negative	Walker 2018 [25]
Phenylalanine (a biomarker of oxidative stress)	Occupational exposure to dust, phenol, formaldehyde, and VOCs (high vs. low exposure)	Positive	Maniscalco 2018 [34]
	Petrochemical plant emissions (high vs. low exposure)	Positive only in children	Chen 2017 [29]
	Petrochemical plant emissions (high vs. low exposure)	Negative	Yuan 2016 [26]
	NO_2	Negative	van Veldhoven 2019 [10]
Serine (glycine, serine, and threonine metabolism, a pathway closely related to oncogenic transformation and the biosynthesis of glutathione)	Petrochemical plant emissions (high vs. low exposure)	Positive only in elderly	Chen 2017 [29]
	Occupational exposure to welding fume (welders vs. controls)	Positive	Kuo 2012 [31]
	TRAP (high vs. low exposure)	Negative	Yan 2019 [23]
	EC and V	Negative only in asthma group	Liang 2019 [11]
	$PM_{2.5}$ and PM_{10}	Negative	Menni 2015 [27]
Aspartic acid (increase glutathione levels and decrease lipid peroxidation in animal models)	Petrochemical plant emissions (high vs. low exposure)	Positive (only investigated children and adolescents)	Chen 2019 [24]
	TRAP	Negative	Liang 2019 [11]
	UFP	Negative	Walker 2018 [25]
	Petrochemical plant emissions (high vs. low exposure)	Negative	Chen 2017 [29]
Hypoxanthine (a substrate for ROS formation)	$PM_{2.5}$ (before/after vs. during the Beijing Olympics)	Positive	Mu 2019 [14]
	$PM_{2.5}$-Zn	Positive	Vlaanderen 2017 [18]
	CO	Positive	Liang 2018 [32]
	$PM_{2.5}$ (sham vs. real purifications)	Positive	Chen 2019 [21]
	$PM_{2.5}$ and $PM_{2.5}$ components (i.e., Ba, Cd, Mn, P, Sb)	Negative only in asthma group	Liang 2019 [11]
Uric acid (a powerful antioxidant and end product of purine metabolism measured in urine samples)	$PM_{2.5}$ (sham vs. real purifications)	Negative	Chen 2019 [21]
	$PM_{2.5}$	Negative	Huang 2018 [28]
	Coking plant emissions (high vs. low exposure)	Negative	Wang 2015 [30]
Acylcarnitine (involving energy metabolism)	O_3	Positive	Miller 2016 [20]
	$PM_{2.5}$	Positive and negative	Chen 2019 [21]
	NO_2	Positive	Jeong 2018 [33]
	Coking plant emissions (high vs. low exposure)	Positive	Wang 2015 [30]
	$PM_{2.5}$-Zn	Negative	Liang 2019 [11]
	NO_2	Negative	van Veldhoven 2019 [10]

Beijing, where coal combustion for heating and industries is a major pollution source [28]. Little overlap in the metabolites that showed associations with PM (i.e., $PM_{2.5}$, PM_{10}, $PM_{2.5–10}$, UFP, and BC) and NOx, indicated differential effects of PM and gas pollutants [10].

The perturbations of metabolites or pathways related to air pollution exposure differed by disease status, age, and sex, which potentially provides mechanistic basis for susceptibility of subpopulations. The ACE study observed that most of the TRAP-related metabolites were differently expressed by asthmatic status of the participants [11]. Specifically, decreased arginine and histidine were significantly associated with exposures to EC or $PM_{2.5}$-V only among asthmatic participants. In addition, in these participants, increased methionine (an essential amino acid promoting ROS) was significantly associated with $PM_{2.5}$-Colbat. The pathways only significant in asthmatic participants tend to be those related to acute inflammatory responses. A randomized crossover study, where 39 healthy volunteers commuted in the Beijing subway for 4 h with or without masks (3 M respirator), found strong associations between 8-hydroxy-deoxyguanosine (8-OHdG), a biomarker of DNA damage, and size-fractioned PM and cardiovascular indicators in men only. This indicates that men might be more prone to air pollution-related cardiovascular effects than women [22]. In the randomized crossover trial on using real and a sham air purifier (each for 9 days) in university dormitories in Shanghai, the analysis of serum samples showed significant interactions of $PM_{2.5}$ with sex for five metabolites, among which the effect sizes for hydroxylamine, arginine, tryptophan, and phytosphingosine in men were larger than those in women [19]. The studies on residents near a large petrochemical complex in Taiwan found that in blood and urine samples, different metabolic pathways were significantly dysregulated and the observed changes were age-specific [24, 26, 29].

Eight studies investigated whether air pollution-related metabolites or pathways were also perturbed by health outcomes, including diseases, (e.g., chronic obstructive pulmonary disease, COPD; IHD), as well as health markers (e.g., measures of lung function, blood pressure, biomarkers of oxidative stress, or inflammation). Seven of the eight studies found overlapping metabolites or pathways related to both air pollution exposures and health outcomes [10], and one found a correlation between $PM_{2.5}$- and COPD-related metabolites among elderly COPD patients and their healthy spouses in Beijing [28]. Table 2 summarizes the associations of air pollution or health outcomes with the following metabolites or pathways: 8-OHdG, prolyl-arginine, vitamin E, metabolites related to purine metabolism, citrate cycle, fatty acid metabolism, glutathione metabolism, linoleate metabolism, glycosphingolipid metabolism, carnitine shuttle, tyrosine metabolism, urea cycle/amino group metabolism, N-Glycan degradation, tryptophan metabolism, phenylalanine metabolism, glycine, serine, and threonine metabolism, alanine, aspartate, and glutamate metabolism.

Conclusions

Over the past decade, the application of untargeted metabolomics to air pollution epidemiology has gained popularity. However, the methodologies used in these studies vary widely for both air pollution exposure assessment and untargeted metabolomics profiling. Most studies investigated ambient PM measured for various size fractions, and the number concentration and chemical composition. Gas pollutants (NO, NOx, CO, O_3) were also measured by these studies. A wide range of exposure assessment methods was used, including portable monitors, stationary monitors, and spatial models. For analysis of the metabolome, most studies used LC-MS as their primary analytical platform. All the studies report the detection of thousands of features, but only a few metabolites can be confirmed with high confidence level. The processes of annotation, statistical analysis, and pathway enrichment analysis also varied widely.

A wide range of metabolites was associated with air pollutant exposures, most of which were endogenous, and a few of which were xenobiotics. Most detected metabolites or pathways were related to oxidative stress or inflammation responses, and perturbations in stress hormones were reported in three studies [16, 19, 21]. Pro-inflammatory metabolites (e.g., leukotrienes) or related metabolism was upregulated [11, 14, 15, 21, 23, 31, 32] and metabolites with anti-inflammation effects (e.g., histidine, linolenic acid) tended to be downregulated under elevated air pollution exposures [11, 23, 25, 26, 32]. Although air pollutants were consistently reported to disrupt antioxidant-oxidant balance, mixed effect directions were reported for numerous oxidants or antioxidants detected in the metabolomics analyses.

Existing evidence cannot conclude which air pollutants or chemical components are most responsible for adverse health effects. The physiochemical characteristics of air pollutants differed across study locations and were likely attributable to heterogeneity in findings. In one study reporting two cross-sectional studies in London and Barcelona with similar metabolomics analysis, the authors found no overlap in either the confirmed metabolites or pathways related to air pollution between the two cities [10]. The metabolic perturbations by air pollution could differ by disease status, age, and sex, but did not seem to differ by racial and ethnic groups [23], which could provide evidence on mechanisms for susceptible subpopulations.

A limited number of studies involving health outcomes (including diseases or health markers) reported overlapping or correlated metabolites or pathways between air pollution exposures and health outcomes. However, all but one of these

studies collected biological samples after or at the same time of health outcome assessment. It is challenging to draw conclusions on whether these overlapping metabolites or pathways mediate the impact of air pollution on health outcomes, or whether air pollution exacerbates health outcomes through these shared metabolisms. Future studies prospectively evaluating exposures to airborne contaminants, metabolic changes, and health outcomes are needed.

Recommendations

The studies documented in this review conducted extensive exposure assessment, and most extensively documented the details of metabolomic analyses. However, additional studies or further improvement can be considered in the following aspects:

- Identifying air pollutants or PM chemical components that have the largest effect on the metabolome;
- Characterizing the joint effects of complex air pollutant mixtures;
- Investigating various potential effect modifiers, such as sex, age, or disease status, to provide a mechanistic basis for susceptibility to a wide range of health effects;
- Expanding sample sizes by leveraging new technologies in personal exposure assessment [47];
- Adopting a more standardized reporting methodology for sample collection and data preprocessing, facilitating comparisons across labs and between studies; and,
- Longitudinal design that prospectively evaluates air pollution exposures, metabolic changes, and health outcomes.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40572-020-00298-x.

References

1. Walker, D.I., et al., *The metabolome: a key measure for exposome research in epidemiology.* Current Epidemiology Reports, 2019: p. 1–11.
2. Patti GJ, Yanes O, Siuzdak G. Metabolomics: the apogee of the omics trilogy. Nat Rev Mol Cell Biol. 2012;13(4):263–9.
3. Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. Mass Spectrom Rev. 2007;26(1):51–78.
4. Rattray NJ, et al. Beyond genomics: understanding exposotypes through metabolomics. Human genomics. 2018;12(1):4.
5. Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. Lancet. 2017;389(10082):1907–18.
6. Turner MC, Jerrett M, Pope CA III, Krewski D, Gapstur SM, Diver WR, et al. Long-term ozone exposure and mortality in a large prospective study. Am J Respir Crit Care Med. 2016;193(10):1134–42.
7. Malley CS, Kuylenstierna JCI, Vallack HW, Henze DK, Blencowe H, Ashmore MR. Preterm birth associated with maternal fine particulate matter exposure: a global, regional and national assessment. Environ Int. 2017;101:173–82.
8. Schraufnagel DE, Balmes JR, Cowl CT, de Matteis S, Jung SH, Mortimer K, et al. Air pollution and noncommunicable diseases: a review by the Forum of International Respiratory Societies' Environmental Committee, part 1: the damaging effects of air pollution. Chest. 2019;155(2):409–16.
9. Schraufnagel DE, Balmes JR, Cowl CT, de Matteis S, Jung SH, Mortimer K, et al. Air pollution and noncommunicable diseases: a review by the Forum of International Respiratory Societies' Environmental Committee, part 2: air pollution and organ systems. Chest. 2019;155(2):417–26.
10. van Veldhoven K, Kiss A, Keski-Rahkonen P, Robinot N, Scalbert A, Cullinan P, et al. Impact of short-term traffic-related air pollution on the metabolome - results from two metabolome-wide experimental studies. Environ Int. 2019;123:124–31.
11. Liang DH, et al. Perturbations of the arginine metabolome following exposures to traffic-related air pollution in a panel of commuters with and without asthma. Environ Int. 2019;127:503–13.
12. de Oliveira BFA, Chacra APM, Frauches TS, Vallochi A, Hacon S. A curated review of recent literature of biomarkers used for assessing air pollution exposures and effects in humans. Journal of Toxicology and Environmental Health, Part B. 2014;17(7–8):369–410.
13. Kelly, F.J. and J.C. Fussell, *Metabolomics as a tool to unravel the oxidative stress-induced toxicity of ambient air pollutants*, in *Oxidative Stress*. 2020, Elsevier. p. 463–488.
14. Mu, L.N., et al., *Metabolomics profiling before, during, and after the Beijing Olympics: a panel study of within-individual differences during periods of high and low air pollution.* Environ Health Perspect, 2019. **127**(5).
15. Ladva, C.N., et al., *Particulate metal exposures induce plasma metabolome changes in a commuter panel study.* PLoS One, 2018. **13**(9).
16. Shen S, Zhang R, Zhang J, Wei Y, Guo Y, Su L, et al. Welding fume exposure is associated with inflammation: a global metabolomics profiling study. Environ Health. 2018;17(1):68.
17. Wei, Y., et al., *Global metabolomic profiling reveals an association of metal fume exposure and plasma unsaturated fatty acids.* PLoS One, 2013. **8**(10).
18. Vlaanderen JJ, Janssen NA, Hoek G, Keski-Rahkonen P, Barupal DK, Cassee FR, et al. The impact of ambient air pollution on the human blood metabolome. Environ Res. 2017;156:341–8.
19. Li, H.C., et al., *Particulate matter exposure and stress hormone levels a randomized, double-blind, crossover trial of air purification.* Circulation, 2017. **136**(7): p. 618−+.
20. Miller DB, Ghio AJ, Karoly ED, Bell LN, Snow SJ, Madden MC, et al. Ozone exposure increases circulating stress hormones and lipid metabolites in humans. Am J Respir Crit Care Med. 2016;193(12):1382–91.
21. Chen C, Li H, Niu Y, Liu C, Lin Z, Cai J, et al. Impact of short-term exposure to fine particulate matter air pollution on urinary metabolome: a randomized, double-blind, crossover trial. Environ Int. 2019;130:104878.
22. Zhang Y, Chu M, Zhang J, Duan J, Hu D, Zhang W, et al. Urine metabolites associated with cardiovascular effects from exposure of size-fractioned particulate matter in a subway environment: a randomized crossover study. Environ Int. 2019;130:104920.
23. Yan Q, Liew Z, Uppal K, Cui X, Ling C, Heck JE, et al. Maternal serum metabolome and traffic-related air pollution exposure in pregnancy. Environ Int. 2019;130:**104872**.

24. Chen C-HS, Kuo TC, Kuo HC, Tseng YJ, Kuo CH, Yuan TH, et al. Metabolomics of children and adolescents exposed to industrial carcinogenic pollutants. Environmental science & technology. 2019;53(9):5454–65.
25. Walker, D.I., et al., *Metabolomic assessment of exposure to near-highway ultrafine particles.* Journal of exposure science & environmental epidemiology, 2018: p. 1.
26. Yuan T-H, et al. Metabolic profiling of residents in the vicinity of a petrochemical complex. Sci Total Environ. 2016;548:260–9.
27. Menni C, Metrustry SJ, Mohney RP, Beevers S, Barratt B, Spector TD, et al. Circulating levels of antioxidant vitamins correlate with better lung function and reduced exposure to ambient pollution. Am J Respir Crit Care Med. 2015;191(10):1203–7.
28. Huang, Q., Hu D., Wang X., Chen Y., Wu Y., Pan L., Li H., Zhang J., Deng F., Guo X., Shen H., *The modification of indoor PM2.5 exposure to chronic obstructive pulmonary disease in Chinese elderly people: a meet-in-metabolite analysis.* Environ Int, 2018. 121: p. 1243–1252.
29. Chen C-HS, Yuan TH, Shie RH, Wu KY, Chan CC. Linking sources to early effects by profiling urine metabolome of residents living near oil refineries and coal-fired power plants. Environ Int. 2017;102:87–96.
30. Wang Z, Zheng Y, Zhao B, Zhang Y, Liu Z, Xu J, et al. Human metabolic responses to chronic environmental polycyclic aromatic hydrocarbon exposure by a metabolomic approach. J Proteome Res. 2015;14(6):2583–93.
31. Kuo C-H, Wang KC, Tian TF, Tsai MH, Chiung YM, Hsiech CM, et al. Metabolomic characterization of laborers exposed to welding fumes. Chem Res Toxicol. 2012;25(3):676–86.
32. Liang DH, et al. Use of high-resolution metabolomics for the identification of metabolic signals associated with traffic-related air pollution. Environ Int. 2018;120:145–54.
33. Jeong A, Fiorito G, Keski-Rahkonen P, Imboden M, Kiss A, Robinot N, et al. Perturbation of metabolic pathways mediates the association of air pollutants with asthma and cardiovascular diseases. Environ Int. 2018;119:334–45.
34. Maniscalco M, Paris D, Melck D, Chiariello N, di Napoli F, Manno M, et al. Biomonitoring of workers using nuclear magnetic resonance-based metabolomics of exhaled breath condensate: a pilot study. Toxicol Lett. 2018;298:4–12.
35. Cheng W, Duncan KE, Ghio AJ, Ward-Caviness C, Karoly ED, Diaz-Sanchez D, et al. Changes in metabolites present in lung-lining fluid following exposure of humans to ozone. Toxicol Sci. 2018;163(2):430–9.
36. Haggarty J, Burgess KE. Recent advances in liquid and gas chromatography methodology for extending coverage of the metabolome. Curr Opin Biotechnol. 2017;43:77–85.
37. Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, Young N, et al. HMDB: the Human Metabolome Database. Nucleic Acids Res. 2007;35(Database issue):D521–6.
38. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 2000;28(1):27–30.
39. Smith CA, Maille GO??, Want EJ, Qin C, Trauger SA, Brandon TR, et al. METLIN: a metabolite mass spectral database. Ther Drug Monit. 2005;27(6):747–51.
40. Griffin JL, Nicholls AW, Daykin CA, Heald S, Keun HC, Schuppe-Koistinen I, et al. Standard reporting requirements for biological samples in metabolomics experiments: mammalian/in vivo experiments. Metabolomics. 2007;3(3):179–88.
41. Spicer RA, Salek R, Steinbeck C. Compliance with minimum information guidelines in public metabolomics repositories. Scientific data. 2017;4:170137.
42. Viant MR, Kurland IJ, Jones MR, Dunn WB. How close are we to complete annotation of metabolomes? Curr Opin Chem Biol. 2017;36:64–9.
43. Barnes S, Benton HP, Casazza K, Cooper SJ, Cui X, du X, et al. Training in metabolomics research. I. Designing the experiment, collecting and extracting samples and generating metabolomics data. J Mass Spectrom. 2016;51(7):461–75.
44. Barnes S, Benton HP, Casazza K, Cooper SJ, Cui X, du X, et al. Training in metabolomics research. II. Processing and statistical analysis of metabolomics data, metabolite identification, pathway analysis, applications of metabolomics and its future. J Mass Spectrom. 2016;51(8):535–48.
45. Di Guida R, et al. Non-targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling. Metabolomics. 2016;12(5):93.
46. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64(4):383–94.
47. Lin, E.Z., et al., The fresh air wristband*: a wearable air pollutant sampler.* Environmental Science & Technology Letters, 2020.