Perturbation of metabolic pathways mediates the association of air pollutants with asthma and cardiovascular diseases

Ayoung Jeong^{a,b,1}, Giovanni Fiorito^{c,d,1}, Pekka Keski-Rahkonen^{e}, Medea Imboden^{a,b}, Agneta Kiss^{e}, Nivonirina Robino^{e}, Hans Gmuender^{f}, Jelle Vlaanderen^{g}, Roel Vermeulen^{g}, Soterios Kyrtopoulos^{h}, Zdenko Herceg^{e}, Akram Ghantous^{e}, Gianfranco Lovisoni^{i}, Claudia Galassi^{j}, Andrea Ranzi^{k}, Vittorio Krogli^{l}, Sara Grioni^{l}, Claudia Agnoli^{l}, Carlotta Sacerdoti^{m}, Nahid Mostafavi^{l}, Alessio Naccarati^{l}, Augustin Scalbert^{l}, Paolo Vineis^{c,n,2}, Nicole Probst-Hensch^{a,b,+,2}, for the EXPOsOMICS Consortium^{3}

^{a} Swiss Tropical and Public Health Institute, Basel, Switzerland
^{b} University of Basel, Basel, Switzerland
^{c} Italian Institute for Genomic Medicine (IIGM), Turin, Italy
^{d} Department of Medical Sciences - University of Turin, Italy
^{e} International Agency for Research on Cancer, Lyon, France
^{f} Genedata AG, Basel, Switzerland
^{g} Utrecht University, Institute for Risk Assessment Sciences, Environmental Epidemiology Division, Utrecht, Netherlands
^{h} National Hellenic Research Foundation, Athens, Greece
^{i} University of Palermo, Palermo, Italy
^{j} Unit of Cancer Epidemiology, Città della salute e della Scienza University-Hospital and Center for Cancer Prevention (CPO), Turin, Italy
^{k} Environmental Health Reference Center, Regional Agency for Prevention, Environment and Energy of Emilia-Romagna, Modena, Italy
^{l} Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy
^{m} Piedmont Reference Center for Epidemiology and Cancer Prevention (CPO Piemonte), Via Santena 7, 10126 Turin, Italy
^{n} MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, UK

A R T I C L E  I N F O

Handling Editor: Yong Guan Zhu

Keywords:
Air pollution
Untargeted metabolomics
Metabolic pathways
Cardio-cerebrovascular diseases

A B S T R A C T

Background: Epidemiologic evidence indicates common risk factors, including air pollution exposure, for respiratory and cardiovascular diseases, suggesting the involvement of common altered molecular pathways.

Objectives: The goal was to find intermediate metabolites or metabolic pathways that could be associated with both air pollutants and health outcomes (“meeting-in-the-middle”), thus shedding light on mechanisms and reinforcing causality.

Methods: We applied a statistical approach named ‘meet-in-the-middle’ to untargeted metabolomics in two independent case-control studies nested in cohorts on adult-onset asthma (AOA) and cardio-cerebrovascular diseases (CCVD). We compared the results to identify both common and disease-specific altered metabolic pathways.

Results: A novel finding was a strong association of AOA with ultrafine particles (UFP; odds ratio 1.80 [1.26, 2.55] per increase by 5000 particles/cm³). Further, we have identified several metabolic pathways that potentially mediate the effect of air pollution on health outcomes. Among those, perturbation of Linoleate metabolism pathway was associated with air pollution exposure, AOA and CCVD.
Refractory asthma is more likely to manifest with cardiovascular co-
molecular pathways altered in the etiology of diseases.

obesity, aging and air pollution exposure, consistent with common
cardiovascular disease share common risk factors such as smoking,
have long been established in adults and in children (Peel et al., 2005;
cline (Downs et al., 2007), with a decreased prevalence of respiratory
In the Swiss SAPALDIA cohort, long term improvement in air pollution
symptoms including wheezing and breathlessness (Schindler et al.,
2009), and with a decreased onset of asthma in adults (Kunzli et al.,
2010). However, few studies conducted an untargeted search for blood
biomarkers of air pollution exposure (Vlaarderen et al., 2017) or
asthma in adults, and none investigated the link between CCVD, asthma
and air pollution.

This study was conducted in the framework of EXPOsOMICS, an EU-
funded project to investigate the air- and water-borne exposome (Vineis
et al., 2017). One of the research questions EXPOsOMICS addresses is
the applicability of the ‘meet-in-the-middle (MITM)’ concept, i.e. inter-
mediate biomarkers as evidence of causality (Vineis et al., 2013). We
have applied the MITM approach within two independent case-control
studies nested in cohorts: one on adult-onset asthma (AOA) within the
SAPALDIA cohort, the other on CCVD within EPIC Italy cohort, and we
compared the results to identify both common and disease-specific al-
terred metabolic pathways.

1. Introduction

Asthmatics often suffer from comorbidities including cardiovascular
diseases. Comorbidity influences the disease prognosis and control.
Refractory asthma is more likely to manifest with cardiovascular co-
morbidities than controlled asthma (Hekking et al., 2018). Asthma and
cardiovascular disease share common risk factors such as smoking,
overweight, aging and air pollution exposure, consistent with common
molecular pathways altered in the etiology of diseases.

Short-term effects of air pollution exposure on asthma exacerbation
have long been established in adults and in children (Peel et al., 2005;
Schwartz et al., 1993; Sunyer et al., 1997). The role of air pollution in
asthma onset is less conclusive, particularly in adults (Anderson et al.,
2013; Jacquemin et al., 2012). Only a few studies used individually
assigned exposure estimates to study the effects of ambient air pollution
on adult-onset asthma. The largest study sample was based on over
600,000 subjects, including 27,000 asthmatics, and demonstrated an
association of PM10 exposure – derived from a pan-European land use
regression model – with asthma prevalence (Cai et al., 2017). The
‘European Study of Cohorts for Air Pollution Effects’ (ESCAPE) reported
a positive but not statistically significant association with asthma inci-
cidence in adults for all air pollution metrics (NO2, NO, PM10, PM2.5,
traffic load; traffic intensity) except PMcourse (Jacquemin et al., 2015).
In the Swiss SAPALDIA cohort, long term improvement in air pollution
levels was associated with an attenuated age-related lung function de-
cline (Downs et al., 2007), with a decreased prevalence of respiratory
symptoms including wheezing and breathlessness (Schindler et al.,
2009), and with a decreased onset of asthma in adults (Kunzli et al.,
2009).

In addition, a growing number of epidemiological studies showed
that air pollution is associated with coronary artery disease (McGuin
et al., 2016; Wolf et al., 2015), cardiovascular diseases (Brook et al.,
2010; Franklin et al., 2015), and cerebrovascular diseases (Stafoggia
et al., 2014) including ischemic stroke (Chung et al., 2017; Cox Jr,
2017). A recent meta-analysis within ESCAPE showed that increases in
PM2.5 and PM10 were associated with risks of fatal and total coronary
events, respectively (Cesaroni et al., 2014), and increased risk for cer-
ebrovascular diseases was reported for higher exposure to PM2.5 and
NO2 (Stafoggia et al., 2014).

Ultrafine particles (UFP) exposure has been less studied than ex-
posure to larger particles, and no regulatory agencies have established
guidelines for UFP so far. Compared to larger particulate matter, UFP
have distinctive characteristics that may lead to higher toxicity: their
extremely small size allows them to reach deeper into the tissues and
evade clearance, and higher surface-to-mass ratio facilitates adhesion of
larger amounts of hazardous materials. Whether this indeed translates
into a higher risk of respiratory or cardiovascular diseases in humans
remains to be ascertained (Herbert and Kumar, 2017).

The biological mechanisms explaining the effects of air pollution on
asthma and its phenotypes and cardio- and cerebrovascular disease
(CCVD) are still poorly understood. The best studied putative biological
mechanism is oxidative stress caused by air pollutants, followed by
pulmonary and systemic inflammation (Guarnieri and Balmes, 2014;
Herbert and Kumar, 2017; Newby et al., 2015; Uzoigwe et al., 2013).
Previous studies investigating the association between long-term ex-
posure to air pollution and various inflammatory blood biomarkers
reported inconsistent results, concerning specific cytokines and pro-
or anti-inflammatory effects (Chuang et al., 2011; Fiorito et al., 2018;
Mostafavi et al., 2015).

Large-scale profiling of small molecules in biological samples has
become available recently, opening the door to the agnostic inter-
rogation of disease processes at the molecular level in epidemiological
settings. The metabolome reflects endogenous processes as well as the
influences from environment and behaviors, and therefore metabo-
lomics provides a unique opportunity to link genome, exposome, and
disease. Metabolomics has been increasingly applied to investigate
asthma and major adverse cardiovascular events (Kelly et al., 2017;
Kordalewska and Markuszewski, 2015; Shah et al., 2012; Wurtz et al.,
2015). However, few studies conducted an untargeted search for blood
biomarkers of air pollution exposure (Vlaarderen et al., 2017) or
asthma in adults, and none investigated the link between CCVD, asthma
and air pollution.

This study was conducted in the framework of EXPOsOMICS, an EU-
funded project to investigate the air- and water-borne exposome (Vineis
et al., 2017). One of the research questions EXPOsOMICS addresses is
the applicability of the ‘meet-in-the-middle (MITM)’ concept, i.e. in-
termediate biomarkers as evidence of causality (Vineis et al., 2013). We
have applied the MITM approach within two independent case-control
studies nested in cohorts: one on adult-onset asthma (AOA) within the
SAPALDIA cohort, the other on CCVD within EPIC Italy cohort, and we
compared the results to identify both common and disease-specific al-
terred metabolic pathways.

2. Methods

2.1. Study population

2.1.1. Asthma in SAPALDIA

Adult-onset asthma (AOA) metabolomics was studied in a nested
case-control study from the Swiss Cohort Study on Air Pollution and
Lung and Heart Diseases in Adults (SAPALDIA). A total of 9651 adults
were recruited in eight cities representing different geographical and
meteorological environments in Switzerland in 1991 (SAPALDIA1);
8047 and 6088 of them participated in the first follow-up in 2001–3
(SAPALDIA2) and in the second follow-up in 2010–11 (SAPALDIA3),
respectively. The study protocol was described in detail previously
(Ackermann-Liebrich et al., 2005; Martin et al., 1997). The present
study examined blood samples from SAPALDIA3. A detailed description
of the population cohort and of the study protocol was described in
detail previously (Ackermann-Liebrich et al., 2005; Martin et al., 1997).
Briefly, asthma cases were selected among the self-reported diagnosis of
asthma occurred later than 16 years of age (n = 141) (Siroux et al.,
2014) and with archived blood sample available. Controls were ran-
domly sampled among the participants who never reported the fol-
lowing since SAPALDIA1: self-reported asthma; physician-diagnosed
asthma; asthma attack in the last 12 months; current asthma medica-
tion; wheezing without cold in the last 12 months; three or more
asthma-related symptoms in the last 12 months (symptoms considered:
breathless while wheezing; woken up with a feeling of chest tightness;
attack of shortness of breath during sleep; attack of shortness of breath
while at rest; woken by attack of shortness of breath) (Jacquemin et al.,
2015). All cases and controls had not smoked for at least 10 years before
blood was drawn. Study participants were non-fasted at the time of
blood collection and bench time was < 2 h for all but ten cases and five
controls. Subjects’ characteristics are summarized in Table 1.

2.1.2. Cardio-cerebrovascular diseases in EPIC Italy

Study participants were part of the Italian component (Turin and
Varese centers) of the EPICOR study (Bendinelli et al., 2011), which is
the cardiovascular section of the European Prospective Investigation

Conclusions: Our results suggest common pathway perturbations may occur as a consequence of chronic ex-
posure to air pollution leading to increased risk for both AOA and CCVD.
Table 1
SAPALDIA sample characteristics – adult-onset asthma.

| N | AOA cases | Controls | AOA cases<sup>a</sup> | Controls<sup>a</sup> |
|---|-----------|----------|-----------------------|---------------------|
| 139 | 196 | 73 | 115 |
| Age [year] | 59.4 (19.4) | 57.1 (15.8) | 60.3 (19.1) | 54.8 (15.5) |
| Female | 87 (63%) | 101 (52%) | 47 (64%) | 62 (54%) |
| BMI [kg/m²] | 25.7 (6.4) | 24.4 (4.8) | 27.0 (6.8) | 24.7 (4.8) |
| Smoking<sup>b</sup> | | | | |
| Former | 54 (39%) | 62 (32%) | 34 (47%) | 37 (32%) |
| Never | 85 (61%) | 134 (68%) | 39 (53%) | 78 (68%) |
| Education level<sup>c</sup> | | | | |
| Low | 3 (2%) | 2 (1%) | 1 (1%) | 2 (2%) |
| Middle | 86 (62%) | 121 (62%) | 46 (63%) | 72 (63%) |
| High | 50 (36%) | 73 (37%) | 26 (36%) | 41 (36%) |
| Fasting time [h] | 2.7 (1.2) | 2.8 (1.7) | 2.9 (1.8) | 2.7 (1.8) |
| Bench time [min] | 80.0 (34.5) | 80.0 (28.2) | 80.0 (30.0) | 80.0 (28.0) |
| PM<sub>2.5</sub> [μg/m³] | | | | |
| t = 1 | 14.6 (1.9) | 14.3 (1.7) | 15.4 (1.5) | 14.7 (2.0) |
| t = 2 | 14.7 (2.2) | 14.4 (1.8) | 15.7 (2.3) | 14.8 (2.4) |
| t = 3 | 14.6 (2.8) | 14.3 (2.2) | 16.0 (2.3) | 14.7 (2.4) |
| t = 4 | 16.0 (2.8) | 15.6 (2.2) | 16.7 (1.7) | 16.2 (1.9) |
| t = 5 | 17.3 (2.3) | 17.1 (2.1) | 17.8 (1.8) | 17.4 (1.9) |
| t = 6 | 16.5 (2.4) | 16.0 (2.2) | 17.2 (1.8) | 16.4 (2.3) |
| t = 7 | 16.8 (3.4) | 16.3 (2.0) | 17.6 (3.2) | 16.8 (3.2) |
| PNC [particles/cm³] | – | – | 13,418 (6376) | 9660 (7970) |
| LDSA [μm²/cm³] | – | – | 33.9 (16.1) | 27.1 (16.3) |
| NO<sub>2</sub> [μg/m³] | 25.0 (14.3) | 21.6 (10.9) | 29.3 (11.9) | 23.7 (15.0) |

Data are presented as count (%) or median (interquartile range). PM<sub>2.5</sub> annual mean estimates derived from the PolluMap in 2010; PNC and LDSA: biennial mean estimates derived from a SAPALDIA multi-area LUR in 2011/2012; NO<sub>2</sub>: annual mean estimates derived from a European LUR in 2010.

<sup>a</sup> Data set used for UFP MWASs, number of observation smaller due to limited availability of UFP estimates.

<sup>b</sup> Former smokers had not smoked for at least 1 year before blood was drawn.

<sup>c</sup> Education level: low (primary school or none), middle (vocational or apprenticeship); high: college or university.

<sup>d</sup> 365 days average t – 1 years before the examination.

Table 2
EPIC Italy sample characteristics – cardio-cerebrovascular diseases.

| N | CCVD cases | Controls | CCVD cases<sup>a</sup> | Controls<sup>a</sup> |
|---|-----------|----------|-----------------------|---------------------|
| 166 | 155 | 71 | 73 |
| Center | | | | |
| Turin | 71 (43%) | 73 (47%) | 71 (100%) | 73 (100%) |
| Varese | 95 (57%) | 82 (53%) | – | – |
| Age [years] | 56.16 (9.56) | 55.55 (9.44) | 58.19 (8.94) | 56.95 (9.90) |
| Female | 107 (64%) | 95 (61%) | 12 (17%) | 13 (18%) |
| BMI [kg/m²] | 26.34 (4.91) | 26.09 (4.91) | 26.34 (3.98) | 26.03 (4.01) |
| Smoking<sup>b</sup> | | | | |
| Former | 52 (31%) | 54 (35%) | 38 (53%) | 38 (52%) |
| Never | 114 (69%) | 101 (65%) | 33 (47%) | 35 (48%) |
| Education level<sup>c</sup> | | | | |
| Low | 103 (69%) | 84 (56%) | 32 (45%) | 22 (30%) |
| Middle | 48 (32%) | 44 (29%) | 29 (41%) | 31 (43%) |
| High | 12 (8%) | 22 (15%) | 10 (14%) | 20 (27%) |
| PM<sub>2.5</sub> [μg/m³] | 21.27 (2.19) | 21.27 (2.16) | | |
| PNC [particles/cm³] | – | – | 14,483 (2335) | 14,227 (2497) |
| NO<sub>2</sub> [μg/m³] | 50.30 (14.95) | 49.62 (16.48) | | |

Data are presented as count (%) or median (interquartile range). PM<sub>2.5</sub>: annual mean estimates derived a European LUR in 2010; PNC: annual mean estimates derived from a local LUR in 2014/2015; NO<sub>2</sub>: annual mean estimates derived from a European LUR in 2010.

<sup>a</sup> Data set used for UFP MWASs, number of observation smaller due to limited availability of UFP estimates.

<sup>b</sup> Former smokers had not smoked for at least 1 year before blood was drawn.

<sup>c</sup> Education level: low (primary school or none), middle (vocational or another secondary school), and high (university or vocational postsecondary school).
2.4. Statistical analyses

2.4.1. Association of air pollution exposure with AOA

We assessed the effect of air pollution exposure on AOA by fitting logistic regression models. AOA was regressed, with non-asthmatics as the reference, on air pollution exposure after adjustment for age, sex, education level, body mass index (BMI), and study area as random effect. For PM$_{2.5}$, the main predictors were two polynomial lag terms defined as $u_0 = \sum_{t=1}^{7} \text{PM}_{2.5}(t)$ and $u_1 = \sum_{t=1}^{7} \text{PM}_{2.5}(t)$. For PM$_{2.5}$, the main predictors were two polynomial lag terms defined as $u_0 = \sum_{t=1}^{7} \text{PM}_{2.5}(t)$ and $u_1 = \sum_{t=1}^{7} \text{PM}_{2.5}(t)$, where PM$_{2.5}(t)$ is average exposure to PM$_{2.5}$ of 365 days $t-1$ years before SAPALDIA3 examination. For UFP and NO$_2$, the main predictors were biennial and annual mean estimates respectively. The association was also assessed in the entire SAPALDIA subjects ($N = 3011; 272$ AOA cases).

2.4.2. Association of air pollution exposure with CCVD

The association of exposure to air pollution with CCVD was assessed in the nested case-control study by logistic regression models adjusting for age at recruitment, center of recruitment, sex, BMI, smoking status, and education level (see Supplementary material for details). In addition, we conducted Cox proportional hazard regression to assess the association between air pollution exposure and the risk of future CCVD among all EPIC subjects (Turin and Varese centers; $N = 18,982; 948$ cases).

---

**Fig. 1.** Search for the MITM pathways.

*Adjusted for the corresponding air pollutant; ** by excluding the pathways not enriched in the other cohort.
CCVD events).

In both studies, odds ratios (OR), hazard ratios (HR), and 95% confidence intervals (CI) refer to an increase of 5 μg/m³ PM$_{2.5}$, 5000 particles/cm$^3$ PNC, 10 μm$^2$/cm$^3$ LDSA, and 10 μg/m³ NO$_2$.

2.4.3. Metabolome-wide association study (MWAS) on AOA

We conducted logistic regression analyses of AOA on each of the 7089 features after adjustment for age, sex, study area, bench time, fasting time, sines and cosines of venipuncture time with periods of 24 and 12 h, and their multiplicative interaction terms with fasting time. We did not adjust for smoking because all subjects were non-smokers since 10 years. Feature intensity, age, bench time, and fasting time were scaled to have mean equal 0 and standard deviation equal 1. We applied the Firth’s bias-reduction method (Firth, 1993; Perry, 2017) to obtain less biased estimates and the Benjamini-Hochberg method to correct for multiple testing (Benjamini and Hochberg, 1995).

2.4.4. MWAS on CCVD

For each of the 2790 features, we tested for their association with incident CCVD by logistic regression models adjusting for age at recruitment, center of recruitment, sex, BMI, smoking status, and education level.

2.4.5. MWAS on air pollution

In SAPALDIA and EPIC Italy separately, each feature was regressed on PM$_{2.5}$, UFP, or NO$_2$ after adjustment for the same covariates as in AOA MWAS and in CCVD MWAS, respectively. In SAPALDIA, a binary indicator for perfect geocoding quality was additionally included as a potential modifier of the effect of air pollution exposure on the metabolite level. Geocoding was declared perfect if the matching was possible at the level of residential address. As in the association of air pollution with AOA, first and second order polynomial lag terms were used for PM$_{2.5}$ while biennial and annual mean exposures were used for UFP and NO$_2$, respectively. In EPIC Italy, annual average exposure was used as the proxy for long-term exposure for each pollutant.

2.4.6. Link and variance functions

In EPIC Italy, feature intensities were Box-Cox transformed before regression (Han and Kronmal, 2004). In SAPALDIA, the best link and variance were sought for each feature and semi-partial pseudo-R$^2$ was computed as a measure of effect size (see Supplementary material for details).

2.5. Meet-in-the-middle (MITM) approach

2.5.1. Search for MITM features

We examined if any of the features associated with air pollution overlapped with the features associated with AOA or CCVD as an attempt to search for MITM features. As no single feature showed metabolome-wide significant association with AOA or CCVD, we found no single MITM features. Instead, we searched for MITM pathways as described below. The history of our analyses in this study is summarized as flowcharts in Supplementary materials (Fig. S1: MITM features; Fig. 1: MITM pathways).

2.5.2. Functional annotation and pathway enrichment tests using Mummichog

Mummichog is an algorithm developed to predict functional activities of metabolites (Li et al., 2013). Taking untargeted MWAS results as input, Mummichog searches for chemical identities by matching the measured mass (m/z) of the features to a reference metabolic model, integrated from KEGG (Kanehisa et al., 2006), UCSD BiGG (Duarte et al., 2007), and Edinburgh human metabolic network (Ma et al., 2007). Based on this putative annotation, it conducts pathway enrichment tests using Fisher's exact test. The statistical significance of pathway enrichment is estimated by permutation, where the features are randomly selected and mapped to each of the possible annotations to produce null distribution. We customized the types of ions that Mummichog searches for chemical identities, to match with the UHPLC-QTOF-MS method used. Cut-off p-value was chosen to have a reasonable number of significant features to ensure for the algorithm to conduct pathway enrichment analysis. We first used the 10th percentile of the p-values from each MWAS result as the cut-off and then the 5th percentile as a sensitivity analysis (Table S1).

2.5.3. Search for MITM pathways

Pathways found enriched (empirical p-value < 0.05) from Mummichog were listed. The pathways with overlap size – the number of features that contributed to the enrichment – smaller than 4 were ignored. This is an attempt to reduce the false positive findings as Mummichog annotates features only by matching m/z and hence matches are subject to error. The pathways that were not enriched for the same air pollution metric in both SAPALDIA and EPIC Italy were excluded. If the pathway enriched for air pollution metric was also enriched for AOA or CCVD after adjustment for the same metric, they were declared as “MITM” pathways (Figs. S3–S5). The MITM pathways were evaluated by confirmation of the putative annotation which Mummichog used to compute pathway enrichment (see Supplementary material for details).

3. Results

3.1. Exposure to UFP is associated with AOA

From logistic regression of AOA (n = 73) with non-asthmatics as the reference group (n = 115), we found a strong association of UFP exposure with AOA (Table 3). The odds ratios were 1.80 [95% CI 1.26, 2.55] for an increase in particle number concentration (PNC) by 5000 particles/cm$^3$, and 1.73 [95% CI 1.27, 2.36] for an increase in lung deposited surface area (LDSA) by 10 μm$^2$/cm$^3$. On the contrary, PM$_{2.5}$ and NO$_2$ did not show a significant association with AOA. The estimated risk for AOA due to UFP exposure is still significant after the inclusion of the other pollutants in the regression model, supporting the independence of the effect, although ORs were lower when estimated in the whole cohort (Table S2).

3.2. Weak but consistent association of air pollution with CCVD

We have observed a positive association of exposure to PM$_{2.5}$, PNC, and NO$_2$ with the risk of CCVD (OR = 1.34 [95% CI 0.72, 2.52] for 10 μg/m³ increase in PM$_{2.5}$; OR = 1.09 [95% CI 0.60, 2.00] for 5000 particles/cm$^3$ increase in PNC; OR = 1.03 [95% CI 0.89, 1.18] for μg/cm$^3$ increase in NO$_2$), though the associations did not reach statistical significance (Table 3). However, when we expanded the analyses to the whole EPIC Turin-Varese subjects (N = 18,982; 948 CCVD events), the associations became stronger and significant (HR = 1.29 [95% CI 1.10, 1.55] for 10 μg/m³ increase in PM$_{2.5}$; HR = 1.16 [95% CI 0.97, 1.39] for 5000 particles/cm$^3$ increase in PNC (Turin subjects; N = 8753); HR = 1.12 [95% CI 0.99, 1.27] for μg/cm$^3$ increase in NO$_2$).

3.3. MWAS: no single metabolites are associated with both air pollution and AOA or CCVD

None of the 7089 features in SAPALDIA or 2790 features in EPIC Italy showed a significant association with AOA or CCVD after multiple testing corrections, respectively (Fig. S1). The air pollution MWAS in SAPALDIA showed 237, three, six and one features significantly associated with PM$_{2.5}$, PNC, LDSA, and NO$_2$, respectively (Fig. 2). One of the three PNC associated features coincided with the LDSA associated features. Five out of the eight UFP associated features were not
3.4. Several metabolic pathways are commonly associated with air pollution in both cohorts

Various pathways were associated with air pollution varying with the air pollutant and the cohort examined (Fig. 1, Tables S3–S9). The pathways that were enriched for the same air pollutant in both cohorts are summarized in Table 4 and Figs. S3–S5: Linoleate metabolism and Fatty acid activation were enriched for PM$_{2.5}$; Linoleate metabolism, glycophospholipid metabolism, and glycophospholipid metabolism for PM$_{2.5}$ and UFP; carnitine shuttle and pyrimidine metabolism for NO$_2$. No overlap was found looking at the list of features that contributed to the enrichment in the two studies (Table S10). We then repeated the same sensitivity analysis using the 5th percentile p-value as the cut-off, as a sensitivity analysis. Linoleate metabolism and glycophospholipid metabolism, associated with UFP, were confirmed in both cohorts. All the pathways associated to NO$_2$, carnitine shuttle and pyrimidine metabolism, were also confirmed.

3.5. Pathways enrichment and MITM analysis for AOA and CCVD

We found various altered metabolic pathways associated with AOA and CCVD (Fig. 1, Tables 5 and 6). The majority of the enriched pathways did not overlap between AOA and CCVD. Pathways associated with AOA and CCVD, respectively, after adjustment for single air pollution metrics to identify MITM pathways are presented in Tables S11–17.

3.6. Linoleate metabolism is a common MITM pathway linking air pollution to AOA and CCVD

Linoleate metabolism was enriched for PM$_{2.5}$ and UFP in both cohorts and for AOA after adjustment for PM$_{2.5}$ or UFP (Tables S11–S13) as well as for CCVD after adjustment for PM$_{2.5}$ (Table S15). Therefore, we considered Linoleate metabolism as MITM linking PM$_{2.5}$ and UFP to AOA and PM$_{2.5}$ to CCVD. Similarly, we considered glycophospholipid metabolism as MITM linking UFP to AOA (Table S13); fatty acid activation, glycophospholipid metabolism, and carnitine shuttle as MITM linking PM$_{2.5}$, UFP, or NO$_2$ to CCVD, respectively (Tables S15–S17).

Linoleate metabolism and glycophospholipid metabolism were confirmed as MITM pathways linking UFP to AOA after the sensitivity analysis (5th percentile of p-values as the cut-off), as well as glycopshingolipid metabolism linking UFP to CCVD, and carnitine shuttle linking NO$_2$ to CCVD.

3.7. Confirmed annotation of metabolites in MITM pathways

A total of 108 features mapping to the aforementioned MITM pathways were selected for confirmation of the putative annotation. Table 7 summarizes all the features whose annotation was confirmed using chemical standards and fragmentation spectra. Linoleate was confirmed in both cohorts with confidence level 1 according to the classification of the Chemical Analysis Working Group (CAWG) (Sumner et al., 2007). In SAPALDIA, linoleate was considered as a signal for the AOA MWAS further adjusted for UFP and contributed to the enrichment of linoleate metabolism and glycophospholipid metabolism. In EPIC Italy, linoleate was considered as a signal for the PM$_{2.5}$ MWAS and contributed to the enrichment of linoleate metabolism. Also confirmed were octanoic acid, sphingosine, and L-carnitine, contributing in EPIC Italy to the enrichment of fatty acid activation for PM$_{2.5}$, glycophospholipid metabolism for UFP, and carnitine shuttle for CCVD adjusted for NO$_2$, respectively. Five additional features were confirmed for their chemical classes with confidence level 3 for the CAWG (Sumner et al., 2007).

3.8. Additional sensitivity analyses

For consistency between the two studies, we performed further sensitivity analyses on AOA. Additional adjustment for education level resulted in a non-relevant change of the results, while adjustment for BMI slightly changed the results (Table S18). In the pathway enrichment analyses, glycophospholipid metabolism remained as MITM linking UFP to AOA after adjustment for BMI or for education level. Linoleate metabolism remained as MITM linking UFP to AOA after adjustment for education level but not after adjustment for BMI.

4. Discussion

In short-term studies, UFP exposure has been reported to have cardio-respiratory effects that were stronger than for larger particles. Peters et al. reported that UFP exposure had a stronger effect on peak expiratory flow than larger particles (Peters et al., 1997). Exposure to UFP but not to larger particles was associated with asthma exacerbations in children (Evans et al., 2014). However, a recent in vitro study showed that coarse particles might have stronger effects on airway epithelium, possibly due to the higher iron content in coarse particles (Kumar et al., 2015). Studies investigating the long-term cardio-respiratory effects of UFP exposure remain very limited. In the California Teachers Study cohort, UFP exposure derived from a chemical transport model was associated with all-cause and ischemic heart disease mortality (Ostro et al., 2015). In the SAPALDIA cohort, UFP exposure was associated with carotid-intima media thickness, a marker of subclinical...
atherosclerosis (Aguilera et al., 2016). UFP exposure derived from a city-specific LUR model in Toronto linked to health registry data of 1.1 million adult city residents found no positive association of UFP exposure with respiratory disease incidence including AOA (Weichenthal et al., 2017). This is in contrast to our findings, which are based on individual reports of asthma and which provide evidence of UFP effects being stronger than, and independent of, those of larger particles. Traffic-related pollutants contribute mainly to the fine or ultrafine particles, while specks of dust of geological origin including metals link to the coarse particles (Kelly and Fussell, 2012; Yamada et al., 2005). Particulates of various sizes may have different toxicity dependent on their composition (Kumar et al., 2015; Schwarze et al., 2007).

4.1. Meet-in-the-middle (MITM) approach

We applied the ‘meet-in-the-middle (MITM)’ approach, which helps in developing a causal hypothesis and improve biological understanding for air pollution-cardio-respiratory health associations, making use of high-resolution metabolomic data. In the MITM approach, one searches for intermediate biomarkers that are associated with both the exposure and the outcome (Vineis et al., 2013). Ideally, this applies to longitudinal studies where the exposure precedes the biomarker measurement, and the biomarker measurement precedes the outcome, e.g. incidence of cardiovascular events, as we did for CCVD in EPIC Italy. It is much less straightforward to define incident cases for asthma than for CCVD. Asthma is a complex chronic disease phenotype.

Fig. 2. Volcano plots of MWAS results in SAPALDIA.
Note the asymmetric distribution of points in air pollution MWASs due to the positive nature of semi-partial pseudo-R² used as a measure of effect size. Linoleate (m/z 281.2464; RT = 7.283) whose annotation was confirmed with confidence level 1 is highlighted in red; Metabolome-wide signals after Benjamini-Hochberg correction in black. Dotted line depicts Benjamini-Hochberg adjusted p = 0.05.
that develops over a long period of time, can go unnoticed for years if not for decades, and can also disappear as well as resurface. This difficulty inherent to asthma research is complicating the assessment of causality to identified risks such as air pollution. Realizing this difficulty, we pursued the MITM approach for asthma even though our study is by design cross-sectional. For all these reasons, we restricted the outcome to adult (after the 16 years of age) onset of asthma which is less susceptible to reverse causation bias and exposure misclassification.

4.2. MWAS analyses

At the level of single metabolites, we found no intermediate biomarkers among the 7089 and 2790 features investigated in SAPALDIA and EPIC Italy respectively, due to lacking metabolome-wide significant associations. Multiple testing corrections can be too stringent, given the highly inter-correlated nature of the metabolome. The effective number of tests (ENT) computed for the SAPALDIA metabolome was 2728, indicating a high degree of dependency in the data. Given this highly correlated, high dimensional data structure, our study likely suffers from low power to detect subtle differences related to chronic diseases, after Benjamini-Hochberg correction.

Table 4
Pathways associated to air pollution in both SAPALDIA and EPIC Italy.

| Air pollutant | Pathway                        | SAPALDIA | EPIC Italy |
|--------------|--------------------------------|----------|------------|
| PM<sub>2.5</sub> | Linoleate metabolism<sup>b,c</sup> | 17       | 6          | 0.0007     | 0.0249     |
|               | Fatty acid activation<sup>c</sup> | 10       | 5          | 0.0054     | 0.0180     |
| UFP<sup>a</sup> | Linoleate metabolism<sup>b</sup> | 12       | 7          | 0.0007     | 0.0084     |
|               | Glycerophospholipid metabolism<sup>b</sup> | 12       | 13         | 0.0023     | 0.0022     |
| NO<sub>2</sub> | Glycosphingolipid metabolism<sup>b</sup> | 12       | 6          | 0.0079     | 0.0367     |
|              | Carnitine shuttle<sup>c</sup> | 10       | 6          | 0.0063     | 0.0040     |
|              | Pyrimidine metabolism<sup>c</sup> | 12       | 8          | 0.0074     | 0.0035     |

<sup>a</sup> Either PNC or LDSA in SAPALDIA and PNC in EPIC Italy.

<sup>b</sup> Also enriched for AOA after further adjustment for the corresponding air pollutant.

<sup>c</sup> Also enriched for CCVD after further adjustment for the corresponding air pollutant.
and in particular to asthma, where distinguishing sub-phenotypes may be essential for understanding risk and etiology of the disease (Jeong et al., 2017; Siroux et al., 2014; Wenzel, 2012). Therefore, heterogeneity and misclassification might have attenuated the associations with biomarkers. Distinguishing further sub-phenotypes requires larger data in future metabolome studies. Given the above, we focused on pathway enrichment analyses.

### 4.3. Pathway enrichment analyses

Metabolomics, given the high dimensionality and high dependency, benefits much from multivariate systems approaches like pathway enrichment tests. Yet, the challenge unique to metabolomics in this context is annotation. Unlike other omics, annotation of the features obtained from untargeted metabolomics requires laborious manual work. The Mummichog software offers an opportunity to bypass this step and to conduct pathway enrichment tests directly from untargeted MWAS results. Using Mummichog, we found various pathways enriched for AOA, CCVD, and air pollution exposures. Air pollution MWASs and pathway enrichment tests conducted in two cohorts served as each other's validation. Although we found no single overlapping features between the two cohorts when comparing validated pathways, lack of such overlap does not exclude the possibility that the pathways truly reflect air pollution-induced metabolic changes, involving different molecules. The specific molecules affected in a pathway may, for example, depend on the particle composition which can vary across different areas (Kelly and Fussell, 2012).

### 4.4. Linoleate metabolism is a common MITM pathway for AOA and CCVD

AOA and CCVD were mostly associated with different sets of pathways and hence MITM pathways linking air pollution exposure to both chronic diseases differed. The two chronic diseases may involve different biological mechanisms and the same environmental insults may act through different pathways. One exception was linoleate metabolism pathway, which was found not only as MITM pathway linking PM2.5 and UFP to AOA but also linking PM2.5 to CCVD. Laboratory analysis confirmed the annotation of linoleate in both cohorts. The feature confirmed as linoleate showed a positive association with AOA, while it did not show statistically significant association with UFPs exposure and did not contribute to the pathway enrichment for UFPs. Still, the linoleate MITM-pathway finding seems biologically interesting. Linoleate was reported in an in vitro experiment to regulate the pro-inflammatory cytokine IL8 (Maruyama et al., 2014) and induce smooth muscle contraction via the free fatty acid receptor 1 (FFAR1) (Mizuta et al., 2015). Another in vitro study demonstrated that α1-antitrypsin bound to linoleate reduced the expression and secretion of IL1β in LPS-stimulated neutrophils, while free α1-antitrypsin did not (Aggarwal et al., 2016). In observational studies in children, eczema was positively associated with linoleate intake (Miyake et al., 2011) and atopy with circulating linoleate (Yen et al., 2008). A recent targeted metabolomic study investigated 64 lipid metabolites and reported Linoleate metabolism and Arachidonic acid metabolism as the top pathways albeit not statistically significantly associated with asthma control (McGeachie et al., 2015). Few studies associated linoleate with CCVD, although in general ω-6 fatty acids have long been believed to have pro-inflammatory effects in the cardiovascular system. An early in vitro study suggested that linoleate may lead to atherogenesis by NFκB signaling mediated vascular adhesion molecule-1 (VCAM-1) expression (Dichtl et al., 2002).

### 4.5. CCVD specific MITM pathways

Glycosphingolipid metabolism was found as MITM pathway linking exposure to UFP and CCVD and annotation of sphingosine was confirmed as one of the modulated metabolites in this pathway. Sphingolipids are structural components of cell membrane but known to play a crucial role in apoptosis, cell growth, senescence, and cell cycle control (Yang et al., 2004). Sphingolipids in blood have been associated with cardiovascular diseases including acute coronary syndrome (Pan et al., 2014) and myocardial infarction (Park et al., 2015). A recent clinical trial reported a strong association between blood sphingolipids and incident cardiovascular diseases (Wang et al., 2017). Sphingolipids have also been associated with asthma (Petrache and Berdyshov, 2016) in contrast to our findings. Perturbation of sphingolipid metabolism may be more relevant for allergic or child-onset asthma (Ono et al., 2015).

Carnitine shuttle pathway was identified as a MITM pathway linking exposure to NO2 and CCVD. Carnitines facilitate the transport of long-chain fatty acids from the cytosol into the mitochondria and play an important role in fatty acid metabolism and carbohydrate utilization. The role of i-carnitine in CCVD has been extensively described, reporting protective effects of i-carnitine administration for various cardiovascular diseases including coronary artery disease, congestive heart failure, and hypertension (Ferrari et al., 2004). A recent meta-

---

**Table 5**

| Pathway                                    | Overlap size | Pathway size | p-Value |
|--------------------------------------------|--------------|--------------|---------|
| Tryptophan metabolism                      | 20           | 54           | 0.0009  |
| Vitamin B6 (pyridoxine) metabolism         | 4            | 6            | 0.0017  |
| Bioterin metabolism                        | 6            | 13           | 0.0021  |
| TCA cycle                                  | 4            | 8            | 0.0041  |
| Hexose phosphorylation                     | 5            | 12           | 0.0048  |
| Fatty acid metabolism                      | 5            | 14           | 0.0101  |
| De novo fatty acid biosynthesis            | 7            | 22           | 0.0102  |
| Drug metabolism - cytochrome P450           | 12           | 42           | 0.0102  |
| Valine, leucine and isoleucine degradation | 7            | 23           | 0.0137  |

Mummichog pathway enrichment test on the results from AOA MWAS adjusted for age, sex, study area, bench time, fasting time, sine and cosine functions of venipuncture time with periods of 24 and 12 h, and their multiplicative interaction terms with fasting time.

**Table 6**

| Pathway                                    | Overlap size | Pathway size | p-Value |
|--------------------------------------------|--------------|--------------|---------|
| De novo fatty acid biosynthesis            | 9            | 14           | 0.0011  |
| Hexose phosphorylation                     | 8            | 12           | 0.0012  |
| Phosphatidylinositol phosphate metabolism  | 6            | 10           | 0.0031  |
| Carnitine shuttle                          | 9            | 19           | 0.0047  |
| Starch and sucrose metabolism              | 6            | 11           | 0.0051  |
| Linoleate metabolism                       | 9            | 20           | 0.0070  |
| Glycosphingolipid metabolism               | 9            | 21           | 0.0105  |
| Glutamate metabolism                       | 5            | 10           | 0.0139  |
| Caffeine metabolism                        | 5            | 11           | 0.0249  |
| Fatty acid activation                      | 6            | 15           | 0.0398  |
| Glycolysis and gluconeogenesis             | 4            | 9            | 0.0479  |
| Fructose and mannose metabolism            | 4            | 9            | 0.0479  |

Mummichog pathway enrichment test on the results from CCVD MWAS adjusted for age at recruitment, center of recruitment, sex, BMI, smoking status, and education level.
| Metabolite                        | Putative annotation from Mummichog         | Level of confidence (Summer et al., 2007) | Pathway                                      | MWAS                        | Regression model   | Coefficient | p-Value | Pseudo-R²a |
|----------------------------------|--------------------------------------------|------------------------------------------|---------------------------------------------|-----------------------------|-------------------|-------------|---------|------------|
| m/z = 281.2464, RT = 7.283       | Linoleate                                   | Level 1                                  | Linoleate metabolism; glycerophospholipid metabolism | AOA, PNC adjusted Logistic 0.29 0.10 – |                  |             |         |            |
|                                 |                                             |                                          |                                             | AOA, LDSA adjusted Logistic 0.32 0.071 – |                  |             |         |            |
|                                 |                                             |                                          |                                             | PNC in SAPALDIA Gamma with log link 7.9e−7 0.86 0.0030 |                  |             |         |            |
|                                 |                                             |                                          |                                             | LDSA in SAPALDIA Gamma with log link −0.00026 0.69 0.0059 |                  |             |         |            |
| m/z = 281.2481, RT = 7.306       | Linoleate                                   | Level 1                                  | Fatty acid activation; linoleate metabolism | CCVD, PM2.5 adjusted Logistic 1.05 0.40 – |                  |             |         |            |
| m/z = 127.1119, RT = 4.388       | Octanoic acid                               | Level 1                                  | Fatty acid activation                       | PM2.5 in EPIC Italy Logistic 0.06 0.001 – |                  |             |         |            |
|                                 |                                             |                                          |                                             | CCVD, PM2.5 adjusted Logistic 0.93 0.17 – |                  |             |         |            |
| m/z = 301.2903, RT = 6.019       | Sphinganine                                 | Level 1                                  | Glycosphingolipid metabolism                | CCVD, PNC adjusted Logistic 0.65 0.30 – |                  |             |         |            |
|                                 |                                             |                                          |                                             | CCVD, PM2.5 adjusted Logistic 0.0003 0.08 – |                  |             |         |            |
| m/z = 162.1128, RT = 0.601       | γ-Carnitine                                 | Level 1                                  | Carnitine shuttle                           | CCVD, NO2 adjusted Logistic 3.24 0.07 – |                  |             |         |            |
| m/z = 279.2221, RT = 7.166       | α-Linolenic acid; γ-linolenic acid          | Level 3                                  | Linoleate metabolism; fatty acid activation | CCVD, PM2.5 adjusted Logistic 0.51 0.01 – |                  |             |         |            |
|                                 |                                             |                                          |                                             | NO2 in EPIC Italy Linear 0.001 0.59 – |                  |             |         |            |
| m/z = 145.0495, RT = 0.646       | d-Glucose; galactose                        | Level 3                                  | Glycosphingolipid metabolism                | CCVD, PNC adjusted Logistic 1.59 0.05 – |                  |             |         |            |
|                                 |                                             |                                          |                                             | CCVD, NO2 adjusted Logistic 0.0001 0.97 – |                  |             |         |            |
| m/z = 424.3428, RT = 6.337       | Linoleaidyl carnitine; linoleyl carnitine    | Level 3                                  | Carnitine shuttle                           | CCVD, NO2 adjusted Logistic 0.79 0.08 – |                  |             |         |            |
| m/z = 426.3590, RT = 6.337       | Octadecenoyl carnitine; vaccenyl carnitine; | Level 3                                  | Carnitine shuttle                           | CCVD, NO2 adjusted Logistic 0.0007 0.20 – |                  |             |         |            |
|                                 | Stearoylcarnitine                           |                                          |                                             | NO2 in EPIC Italy Linear −0.005 0.22 – |                  |             |         |            |
|                                 |                                             |                                          |                                             | NO2 in EPIC Italy Linear −0.006 0.08 – |                  |             |         |            |

Level 1: identified compounds; Level 3: putatively characterized compound class.

a  Semi-partial pseudo R² as a measure of effect size for air pollution MWASs in SAPALDIA: for details see Supplementary information.
analysis of randomized controlled trials demonstrated the efficacy of L-carnitine against chronic heart failure (Song et al., 2017). In an experimental study in rats, inflammation accompanied with hypertension was attenuated by L-carnitine administration (Miguel-Carrasco et al., 2008). In this study, however, L-carnitine was associated with increased risk of CCVD.

4.6. Strengths and limitations

Strengths of our study include its prospective nature (nested in longitudinal cohorts), the individual assessment of exposure to air pollution, the accurate diagnoses for the outcomes, the agnostic nature of our metabolome-wide measurements, and the application of ‘meet-in-the-middle’ as a novel approach helping in the causal interpretation of the results. We focused on biological pathways that were associated with air pollution (mostly UFP) in both studies, supporting the robustness and replicability of our findings. Limitations include the small sample size for metabolome-wide analyses; we focused on pathways enrichment but we were not able to identify single features associated with both air pollution and at least one disease due to the lack of statistical power. Also, we used slightly different statistical methods (including the set of confounders) in the two studies, mainly due to the nature of the outcomes and the quite different estimation of exposure in the two studies. For example, unlike in EPIC Italy, we did not adjust for BMI in SAPALDIA. Air pollution exposure can increase the risk of obesity (Eze et al., 2015; Wei et al., 2016) and obesity may have a causal effect on asthma (Wenzel, 2012), therefore adjustment for BMI can lead to missing some signals. Given the smaller sample size and expected subtle effects, parsimony was more strongly sought in AOA MWAS. And a previous study observed less strong association between socioeconomic status and air pollution exposure in Switzerland than in Italy (Temam et al., 2017). Sensitivity analysis showed that the additional adjustment did not change the results.

5. Conclusions

In summary, we successfully applied a MTTM approach in untargeted metabolomics to produce evidence of common and disease-specific pathway perturbations in the etiological relationship between air pollution exposure, AOA, and CCVD. Our findings need to be confirmed in future targeted and untargeted studies.

Data availability

Raw metabolomic data that support the findings of this study are available from EXPOSOMICS but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of EXPOSOMICS.

Acknowledgments

This work was supported by the grant FP7 of the European Commission “Enhanced exposure 36 assessment and omic profiling for high priority environmental exposures in Europe” (EXPOSOMICS grant 308610 to PV). The SAPALDIA cohort and biobank is funded by the Swiss National Science Foundation grant no 33CS30-148470/1 to NP. EPIC-Italy was financially supported by the Italian Association for Cancer Research (AIRC). We thank Mr. Vincent Cahais for coordinating the data storage and transfer. The authors declare they have no actual or potential competing financial interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2018.06.025.
