國立臺灣大學公共衛生學院職業醫學與工業衛生研究所
博士論文

Institute of Occupational Medicine and Environmental Hygiene

College of Public Health

National Taiwan University

Doctoral Dissertation

六輕石化工業區附近居民多重污染暴露與代謝體關係
之暴露體學研究

Exposomic study on the association between multiple
pollutants exposure and metabolome in residents living
near No. 6 Naphtha Cracking Complex

陳其欣

Chi-Hsin Chen

指導教授：詹長權 博士

Advisor: Chang-Chuan Chan, Sc.D.

中華民國 108 年 7 月

July 2019

國立臺灣大學博士學位論文
口試委員會審定書

六輕石化工業區附近居民多重污染暴露與代謝體關係之暴露體學研究

Exposomic study on the association between multiple pollutants exposure and metabolome in residents living near No. 6 Naphtha Cracking Complex

本論文係陳其欣君（D03841004）在國立臺灣大學職業醫學與工業衛生研究所完成之博士學位論文，於民國 108 年 5 月 22 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

<u>　　　　　　　　　　　　　　</u>（簽名）

<u>　　　　　　</u>（指導教授）　　<u>　　　　　　　　</u>

<u>　　　　　　　　</u>　　　　　　<u>　　　　　　　　</u>

<u>　　　　　　　　</u>　　　　　　<u>　　　　　　　　</u>

<u>　　　　　　　　</u>　　　　　　<u>　　　　　　　　</u>

中文摘要

研究背景：暴露體學已成為環境衛生學界的重要方法論，近年來更是發展出 "Public Health Exposome Approach"，探討特定地區的暴露特徵及健康影響。本論文針對台灣最大的石化工業區第六套輕油裂解廠（簡稱六輕）附近居民進行暴露體學研究，找出暴露程度、代謝體及早期健康效應生物指標物之間的相關性。

研究方法：本研究依住家與六輕距離、尿中暴露生物指標物濃度（釩與多環芳香烴暴露生物指標物 1－羥基芘）將 273 位研究對象分為高暴露組（9－15 歲小孩 43 人、＞55 歲老年人 77 人）與低暴露組（小孩 75 人、老年人 78 人），分析其（一）外在暴露：對六輕主要排放源的距離、住家附近道路面積、住家空氣中釩及多環芳香烴濃度；(二) 內在暴露：尿中石化工業污染暴露生物指標物砷、鎘、鉻、鎳、汞、鉛、釩、錳、銅、鍶、鉈與 1－羥基芘濃度；(三) 代謝體：利用二維氣相層析飛行時間質譜儀建立尿液代謝體，以超高壓液相層析－四極柱飛行時間質譜儀分析血液代謝體及血液脂質體；(四) 早期健康效應：尿中氧化壓力指標物與血中醯基肉鹼類濃度。本研究以「中途相遇法」找出潛在可作為連結暴露與早期健康效應的中間生物指標物，並以生物途徑分析找出多重石化工業污染物暴露可能影響的生理途徑。

研究成果：本研究結果顯示在小孩及老年人受試者中，高暴露組比低暴露組居住離六輕主要排放源較近、有較高的住家空氣中釩及多環芳香烴濃度，且高暴露組比低暴露組有較高的尿中暴露生物指標物濃度與氧化壓力生物指標物濃度。尿液代謝體在高低暴露組之間有年齡依賴性的改變可連結多重暴露與氧化壓力，在小孩中是色氨酸代謝等途徑的異常，在老年人中則是甘氨酸、絲氨酸與蘇胺酸代謝等途徑的異常，且在小孩與老年人的尿液代謝體中發現潛在暴露生物指標物癸烷、十二烷、十三烷。小孩的血液代謝體在高低暴露組之間有顯著差異，並找到十個潛在可做為中間生物指標物的代謝物質，連結多重工業致癌物暴露（國際癌症研究機構定義一

級致癌物：砷、鎘、鉻、鎳；二級致癌物：汞、鉛、釩、多環芳香烴）與早期健康效應氧化壓力增加、血中醯基肉鹼類濃度異常。生物途徑分析結果顯示小孩暴露於多重工業致癌物質可能造成嘌呤代謝途徑異常。小孩的血液脂質體在高低暴露組之間有顯著差異，並發現有 21 個脂質與多重工業污染物暴露相關，包括溶血卵磷脂類、卵磷脂類、神經鞘磷脂類及磷脂酸肌醇類，這四種脂質皆可連結到尿中氧化壓力生物指標物或血中醯基肉鹼類。

結論：Public health exposome approach 可用於探討石化工業影響地區內的易感族群，並釐清多重工業污染暴露如何影響重要生理途徑，導致與慢性和急性疾病相關的早期健康效應。氣相層析方法分析尿液代謝體可用於辨識石化工業附近的易感族群如小孩與老年人，並發現與年齡相關的生理途徑連結多重暴露與氧化壓力。液相層析方法分析血液代謝體可用於尋找多重工業致癌汙染物暴露在小孩與青少年體內影響的生理途徑，並連結癌症相關的早期健康效應。液相層析方法分析血液脂質體可用於辨識多重工業污染暴露在小孩及青少年體內造成與肝功能異常相關的脂質變化。基於本研究的發現，我們建議顯著降低石化工業污染排放量以減少暴露程度、改善代謝異常，並持續追蹤六輕附近居民的健康狀態。本研究也證實，暴露體學可作為公共衛生研究工具，探討工業污染對附近居民既有及潛在的健康效應，未來可作為尋找新的個人化健康效應指標及暴露生物指標物質、建立個人化風險評估指標的參考。
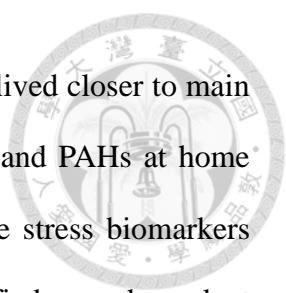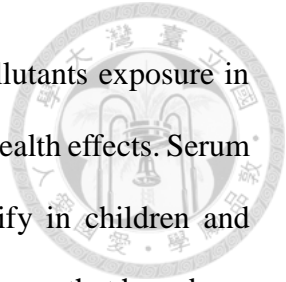
關鍵字：石化工業、暴露體學、代謝體學、脂質體學、重金屬、多環芳香烴

Abstract

Background: Exposomics is an important methodology in environmental health research. Recently, a branching paradigm, the Public Health Exposome Approach, focuses on the impact of exposures on the overall health of a population within a particular region. This dissertation focuses on the exposomics study of residents living near No. 6 Naphtha Cracking Complex, the largest petrochemical complex in Taiwan, and aim to clarify the association between exposure levels, metabolome, and early health effect biomarkers.

Material and Methods: We classified 273 study subjects as high exposure group (children aged 9-15 N=43; elderly aged > 55 N=77) and low exposure group (children N=75; elderly N=78) by the distance from their homes to the complex, and urinary levels of exposure biomarkers vanadium (V) and polycyclic aromatic hydrocarbon (PAHs) metabolite 1-hydroxypyrene (1-OHP). We analyzed (1) external exposures: distance from their homes to main emission points of the complex, road area surrounding homes, and ambient levels of V and PAHs at homes using previously established models; (2) internal exposures: urinary levels of exposure biomarkers, arsenic (As), cadmium (Cd), chromium (Cr), nickel (Ni), mercury (Hg), lead (Pb), vanadium (V), manganese (Mn), copper (Cu), strontium (Sr), thallium (Tl), and 1-OHP; (3) metabolome: urine metabolomics was analyzed using two dimensional gas chromatography coupled with time-of-flight mass spectrometry (GCxGC-TOFMS), and serum metabolomics and lipidomics were analyzed using ultra-high performance liquid chromatography-quadrupole time-of-flight mass spectrometry (UHPLC-qTOFMS); (4) early health effects: urinary levels of oxidative stress biomarkers, and serum acylcarnitines. We applied "meet-in-the-middle" approach to identify potential intermediate biomarkers connecting exposures with early health effects, and pathway analysis to find biological mechanisms affected by exposure to multiple pollutants.

Results: In both children and elderly subjects, high exposure group lived closer to main emission points of the complex, had elevated ambient levels of V and PAHs at home locations, and increased urinary exposure biomarkers and oxidative stress biomarkers compared to low exposure group. Urine metabolomics identified age-dependent biological pathways that associated multiple pollutants exposure with increased oxidative stress, including tryptophan metabolism in children, and serine, glycine, and threonine metabolism in elderly subjects. In addition, potential exposure biomarkers decane, dodecane, and tridecane were identified in both children and elderly subjects. Serum metabolomics found 10 potential metabolites possibly linking increased exposure to IARC group 1 carcinogens (As, Cd, Cr, Ni) and group 2 carcinogens (V, Hg, PAHs) with elevated oxidative stress and deregulated serum acylcarnitines. Purine metabolism was identified as the possible mechanism affected by children's exposure to carcinogens. Serum lipidomics results in children also showed significant difference between high and low exposure groups. We found 21 lipids associated with multiple industrial pollutants exposure, including lysophosphatidylcholines, phosphatidylcholines, sphingomyelins, and phosphatidylinositols. All four types of lipids were associated with urinary oxidative stress biomarkers and/or serum acylcarnitines.

Conclusion: Public health exposome approach could be used in a large petrochemical industry influenced region to identify vulnerable populations, and understand how multiple industrial pollutants exposure are affecting critical biological mechanisms, leading to early health effects that may be precursors to chronic and acute diseases. Urine metabolomics analyzed via GC-based method could be used to identify children and elderly as vulnerable populations in regions influenced by a large petrochemical industry, and found age-dependent pathways linking multiple exposures to increased oxidative stress. Serum metabolomics analyzed via LC-based method could be used to find

biological pathways affected by multiple industrial carcinogenic pollutants exposure in children and adolescents, that could be linked to cancer-related early health effects. Serum lipidomics analyzed via LC-based method could be used to identify in children and adolescents exposed to multiple industrial pollutants, lipid profile changes that have been implicated in liver dysfunctions. Based on our findings, we suggest significant reduction of petrochemical industrial emissions from the complex to decrease multiple pollutants exposure and metabolic abnormalities, and continued follow up on of residents' health. This dissertation also attests the application of exposomics as a public health research tool, in the investigation of current and potential health impacts of industrial pollution on nearby residents, providing information for future identification of novel personalized health indicators and exposure biomarkers, and establishment of individual risk index.
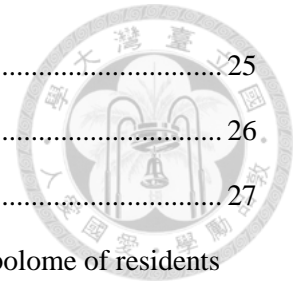
Keywords: petrochemical industry, exposomics, metabolomics, lipidomics, heavy metals, polycyclic aromatic hydrocarbons.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Background

Petrochemical industrial complex is a consortium of high-pollution facilities such as oil refineries and coal-fired power plants. These facilities emit multiple pollutants including sulfur oxides (SO_x), nitrogen oxides (NO_x), carbon dioxide (CO_2), carbon monoxide (CO), volatile organic compounds (VOCs), polycyclic aromatic hydrocarbons (PAHs), and heavy metals arsenic (As), cadmium (Cd), chromium (Cr), nickel (Ni), vanadium (V), mercury (Hg), lead (Pb), manganese (Mn), copper (Cu), strontium (Sr), and thallium (Tl) (Chan et al. 2006; Driscoll et al. 2015; George et al. 2015; Hu et al. 2011; Nadal et al. 2004; Nadal et al. 2009). Cumulative exposure to such complex chemical mixtures may have synergistic effects on health, and warrant the use of novel analytical approaches for a comprehensive evaluation (Carpenter et al. 2002).

In Taiwan, Chan et al. have conducted for the past ten years, extensive environmental and epidemiological studies near No. 6 Naphtha Cracking Complex, the largest petrochemical complex in Taiwan. To date, Chan et al. have published 15 research articles in SCI journals, 12 master theses and doctoral dissertations, and annual reports documenting the environmental and health impacts of No. 6 Naphtha Cracking Complex on surrounding areas and residents (Table 1).

Environmental studies found significant increase of ambient pollutants within 10 km radius of the complex, including NO_x, SO_x, VOCs such as ethylene, propylene, propane, butane, and benzene, PAHs such as anthracene, chrysene, fluoranthene, phenanthrene, and pyrene, vinyl chloride monomers (VCM), and metals (詹長權 2010, 2011, 2012, 2013). For his doctoral dissertation, Shie did a comprehensive study of air toxics pollution in areas surrounding No. 6 Naphtha Cracking Complex from accidental and routine

1
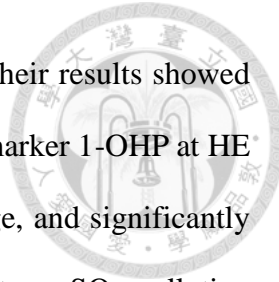
emissions (謝瑞豪 2014). In 2011, he deployed a variety of air-monitoring instruments to evaluate the air toxin levels inside and downwind of No. 6 Naphtha Cracking Complex before and during a fire at the complex caused by a liquefied petroleum gas fuel leak. They found high levels of combustion-related gaseous and particulate pollutants inside the complex and 10 km downwind for at least two days after the fire, demonstrating that a timely and comprehensive air monitoring is essential for tracing air pollution from industrial accidents (Shie and Chan 2013). Shie also used pollution rose to assess the level of SO_2 in townships downwind to the complex in preoperational period (1995–1999) and two postoperational periods (2000–2004 and 2005–2009), and showed that in the postoperational periods, hourly SO_2 levels exceeded the U.S. Environmental Protection Agency (EPA) health-based standard of 75 ppb (Shie et al. 2013). Pien established the protocols for analyzing heavy metals in particulate matters collected near No. 6 Naphtha Cracking Complex using Harvard Impactor, and in urine samples of residents living near the complex for her master thesis (邊瑋緒 2011). This methodology was later applied by Chio et al. to construct a two-stage dispersion model to assess the ambient concentrations of V and As in the vicinity of the complex, and by Yuan et al. to confirm association between model-estimated ambient V at home locations and individual urine concentrations of V in residents living near the complex (Chio et al. 2014; Yuan et al. 2015a). Yuan et al. established a kriging model to assess the ambient concentration of 16 PAHs surrounding the complex in 2015, and found significant association between estimated ambient levels of five PAHs including pyrene, benzo[a]anthracene, benzo[k]fluoranthene, fluoranthene, and dibenzo[a,h]anthracene, at home addresses and individual urinary concentrations of 1-hydroxypyrene (1-OHP) in residents living near the complex (Yuan et al. 2015b). These studies established urinary V and 1-OHP as exposure biomarkers for petrochemical industrial pollution within this area.
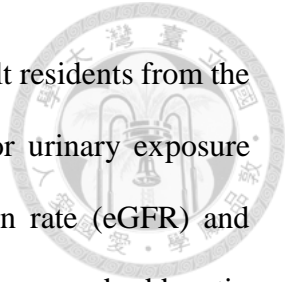
In order to examine the health impact of petrochemical industrialization, Chen et al. compared life expectancies and personal income between Yunlin County where No. 6 Naphtha Cracking Complex is located, and one reference county (Yilan County) which had no significant industrial activities, using data spanning 11 years before and after the complex began operations in 1999. Their findings showed Yunlin residents had lesser increases in life expectancy over time than Yilan residents, with male residents more vulnerable to the effects of industrialization, and no significant differences in individual income between the two counties (Chen et al. 2014).

Epidemiology studies were also conducted to further investigate the health impact of petrochemical industrial pollution on nearby residents, including chronic diseases such as cancer, chronic kidney disease, and hyperlipidemia, acute disease such as allergic diseases, asthma, and bronchitis, and subclinical abnormalities such as liver fibrosis. For his master thesis, Shen used primary data of demographic information, risk factors, biomarkers, and biochemical indices to investigate the adverse health effects, and secondary data of Taiwan Health Insurance Database (Registry for Catastrophic Illness Database) to retrospectively investigate the incidence of all cancers (ICD-9: 140-165, 170-176, 179-208) in 2,388 adults aged > 35 years at the time of recruitment (2009-2012), and aged > 20 years when the complex began operations in mid-1999, who have lived in Yunlin County for more than five years (沈育正 2014). Yuan et al. applied his methodology, and geographically classified the 2,388 participants into high exposure group (HE, lived in Mailiao and Taisi Townships, < 10 km from the complex), and low exposure group (LE, lived in Baojhong, Shihhu, Dongshih, Lunbei, Erlun, Citong, Yuanchang, and Huwei Townships, > 10 km from the complex). Temporally, Yuan et al. divided the 12 years participants lived near the complex since the operation of complex began with reported emissions of VOCs into the first period 1999-2007 (0-9 years after operation began) and

3
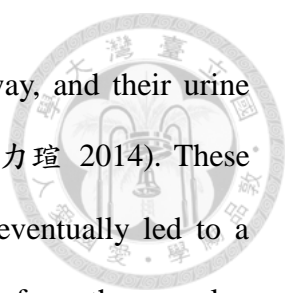
the second period 2008-2010 (10-12 years after operation began). Their results showed higher urine levels of carcinogens As, Cd, Hg, Pb, V, and PAHs biomarker 1-OHP at HE compared to LE, with Pb and V urine levels exceeding normal range, and significantly higher body mass index (BMI) and hepatitis C prevalence. Long-term SO_2 pollution levels were also significantly higher in HE than LE areas. Significant exposure area effect on elevating the relative risks (RRs) of the all cancer crude cumulative incidence rates (CIRs) were found for elder subjects (1.52; 1.04-2.22), female subjects (1.41; 1.00-1.97), and elder female subjects (1.91; 1.15-3.19) after the complex had operated for 10-12 years (Yuan et al. 2018). Chen et al. conducted a similar study in Changhua County which is north of the complex, with 1,934 adult participants (aged > 20) recruited in 2014-2016 who have lived in this area for more than five years, geographically divided into three study zones: Taisi Village (average 5.5 km from complex), Dacheng Township (average 9.2 km from complex), and Zhutang Township (average 19.9 km from complex), comparing all cause cancer incidence rate (ICD-9: 140-208), and urine exposure biomarkers. Results showed urine levels of carcinogenic pollutants As, Cd, Cr, Ni, and V, as well as other pollutants Mn, Cu, and Tl were significantly higher for participants in Taisi Village compared to the other two study zones. Temporal increase for all cause cancer incidence rates (IRs) were found in all three study zones when comparing 1999-2007 period (0-9 years after operation began) to 2008-2014 period (10-16 years after operation began), with the highest crude incidence rate ratios (IRRs) in Taisi Village compared to the other two study zones. All cause cancer IRRs were higher for Taisi Village compared with the other two study zones for all subjects and male subjects, and higher for Taisi Villange than Dacheng Township for female subjects, after the complex had operated for 10-16 years, with hepatitis C and age significantly associated with higher all cause cancer IRRs (Chen et al. 2018).

4
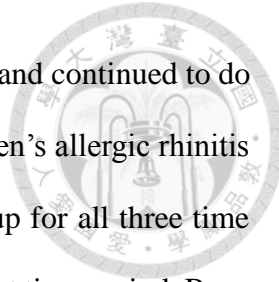
In addition to cancer, for his master thesis Ke analyzed 2,069 adult residents from the same epidemiology cohort that Shen used from Yunlin County, for urinary exposure biomarkers and the association with estimated glomerular filtration rate (eGFR) and chronic kidney disease (CKD). He found that decreased eGFR and increased odds ratio of CKD were associated with decreased distance from home address to complex, and increased levels of urine As (柯登元 2016). Jhuang conducted a similar study in Dacheng Township of Changhua County with 1,374 adult participants recruited from 2014-2016 for her master thesis. Her findings confirmed the association between decreased eGFR and increased risk of CKD with decreased distance from home location to complex, with increased urinary levels of Ni and Cr associated with decreased eGFR and increased risk of CKDs (莊明潔 2018). Shun's master thesis discussed the association between serum heavy metals levels and hyperlipidemia and CKD in 1,000 Yunlin adult residents aged > 35 years from the same cohort as Shen and Ke. Her findings showed significant and positive association between serum Cr, As, and Hg with total cholesterol levels, serum Hg with low-density lipoprotein cholesterol levels (LDL-C), and serum As and Hg with risk of hyperlipidemia. She also found association between increased serum As, Cr, and Tl with decreased eGFR, and increased serum As and Cr with increased risk of CKD (孫稚翔 2017).

Epidemiology studies were also conducted in children and adolescents who lived near the complex during critical periods of biological development. Liu established for her master thesis, an analytical method for exposure biomarker urinary thiodiglycolic acid (TDGA), a major metabolite of VCM, and used this method to analyze urine samples from 268 schoolchildren recruited from four elementary schools in Mailiao Township of Yunlin County. She found children attending an elementary school less than 1 km from the VCM/polyvinylchloride (PVC) plants within the complex had higher urine

5

concentration of TDGA than children attending schools further away, and their urine levels of TDGA significantly reduced during summer vacation (劉力瑄 2014). These findings were later published and gained media attention, which eventually led to a temporary relocation of the children to another school further away from the complex (Huang et al. 2016). In 2018, Wang et al. found association between urine TDGA levels and subclinical abnormal levels of hepatic fibrosis indicators serum aspartate aminotransferase (AST) and fibrosis-4 score (FIB-4) in the same group of schoolchildren (Wang et al. 2019). Chen applied a similar study design on 447 adult residents in Dacheng and Zhutang Townships of Changhua County for his master thesis, and found residents living closer to the complex had increased urine levels of TDGA, and significant association between urinary TDGA concentrations and liver fibrosis level indicator FIB-4 (陳俊霖 2018). For her master thesis, Chiang recruited 587 11-14 year old school children from junior high schools in Yunlin County from 2009 to 2011, who have lived at the same addresses for more than five years, and classified them as high exposure group (HE, lived in Mailiao, Taisi, Donshih Townships) and low exposure group (LE, lived in Erlun, Lunbei, Huwei, Baojhong, Sihhu, and Yuanchang Townships). Her study covered the time from 1999 to 2010, which was further divided into three periods: four years (1999-2002), eight years (1999-2006), and 12 years (1999-2010) after the complex began operations. Health data were obtained from Taiwan Health Insurance Database, choosing outpatient data for allergic rhinitis (ICD-9-CM: 477), bronchitis (ICD-9-CM: 490-491), and asthma (ICD-9-CM: 493). SO_2 was used as an indicator of exposure from the complex, using hourly data measured at two air quality monitoring stations set up by the Taiwan Environmental Protection Administration (TEPA) at HE area Taisi Township and LE area Lunbei Township from 1995 to 2010. From 2001, SO_2 concentration increased significantly in HE areas, and the three-year average of the 99th percentile of SO_2

6

concentration have exceeded U.S. EPA 75 ppb standards since 2003, and continued to do so with increasing concentrations up to 2010. Hazard ratios of children's allergic rhinitis and bronchitis were significantly higher in HE compared to LE group for all three time periods, while for asthma the difference was only significant in the first time period. Boys had higher risk of developing allergic rhinitis and asthma, and children living near roads had higher risk of developing allergic rhinitis. Her results showed that association between SO_2 exposure and acute respiratory effects occurred as early as < 2 years after the complex began operations, and lasted 8 to 12 years (Chiang et al. 2016; 江姿穎 2015). Killian recruited 168 preschool children aged 4-8 from four kindergartens within 13.7 km of the complex for her master thesis, and analyzed their urine concentrations of heavy metals and oxidative stress biomarkers, and at the same time used a food frequency questionnaire to assess individual's intake of antioxidants. Her findings showed preschool children living closer to the complex had increased urinary levels of As, Cd, Cr, Ni, Pb, Mn, Cu, and Sr which were associated with elevated levels of urinary oxidative stress biomarker 8-hydroxy-2'-deoxyguanosine (8-OHDG). Increased intake of total oxidants resulted in a decrease of urine 8-OHDG that did not reach statistical significance (柯昀君 2017).

In order to clarify the biological mechanism between industrial pollutants exposure and oxidative stress, Yuan et al. used nuclear magnetic resonance spectroscopy (NMR) to analyze serum metabolites of 160 residents from a prospective cohort in Yunlin County. They found that exposure to V and PAHs may cause a reduction in amino acids and carbohydrates levels by elevating peroxisome proliferator-activated receptor (PPAR) signaling pathway, insulin signaling, and oxidative/nitrosative stress (Yuan et al. 2016). In vitro study was also conducted, and results showed that exposure to $PM_{2.5}$ from No. 6 Naphtha Cracking Complex emissions significantly correlated with reduced cell viability

7

and increased cytotoxicity-related lactate dehydrogenase, oxidative stress-related 8-isoprostane, and inflammation-related interleukin (IL)-6 (Chuang et al. 2018).

The findings of Chan et al. showed that adult, elderly, and children residents living near No. 6 Naphtha Cracking Complex are exposed to multiple hazardous industrial pollutants from routine and accidental emissions, and have increased risk of chronic and acute adverse health effects. Children and elderly residents may be more susceptible to these industrial pollutants exposure since children have immature physical development, and higher inhalation of air per unit time, and elderly residents may have compromised immune responses and underlying health conditions (Adler 2003; Makri and Stilianakis 2008). The complexity of industrial pollution and health effects on different age groups, with temporal and spatial differences in this industrial community, indicated that traditional models accessing single toxic exposure and disease are not sufficient in evaluating the health status of people living in this area. Comprehensive evaluation of multiple industrial pollutants exposure and the impact on biological mechanisms and pathways that underlie a range of common complex diseases are needed in order to provide information for future risk assessment and development of personal and community interventions. To achieve this, application of novel approaches were required (Juarez et al. 2014).

Table 1. Studies on the Environmental and Health Impacts of No. 6 Naphtha Cracking Complex

	Thesis / Dissertation	SCI Journals
External exposures		
Routine emissions	(謝瑞豪 2014)	
Heavy metals	(邊瑋緒 2011)	(Chio et al. 2014)
PAHs		(Yuan et al. 2015b)
SOx	(江姿穎 2015)	(Shie et al. 2013) (Chiang et al. 2016)
Accidental emissions	(謝瑞豪 2014)	(Shie and Chan 2013)
Internal exposures		
Heavy metals		
Urine	(邊瑋緒 2011) (柯登元 2016) (莊明潔 2018) (柯昀君 2017) (謝億廷 2019)	(Yuan et al. 2015a)
Serum	(孫稚翔 2017)	
PAHs		(Yuan et al. 2015b)
TDGA	(劉力瑄 2014) (陳俊霖 2018)	(Huang et al. 2016) (Wang et al. 2019)
Biological mechanisms	(陳其欣 2019) (謝億廷 2019)	(Yuan et al. 2016) (Chen et al. 2017) (Chen et al. 2019)
Health effects		
Cancer incidence	(沈育正 2014)	(Yuan et al. 2018) (Chen et al. 2018)
Respiratory disease	(江姿穎 2015)	(Chiang et al. 2016)
Chronic kidney disease	(柯登元 2016) (孫稚翔 2017) (莊明潔 2018)	
Liver fibrosis	(陳俊霖 2018)	(Wang et al. 2019)
Hyperlipidemia	(孫稚翔 2017)	
Oxidative stress	(柯昀君 2017) (陳其欣 2019)	(Yuan et al. 2016) (Chen et al. 2017) (Chen et al. 2019)
In vitro study		(Chuang et al. 2018)

9
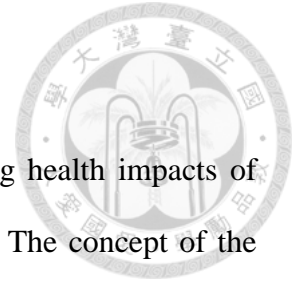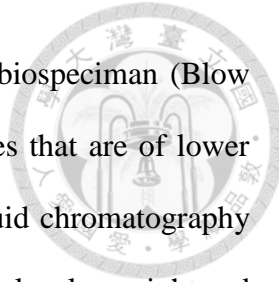
1.2 Exposomics

Exposomics has become the leading methodology for assessing health impacts of multiple environmental exposures in environmental health studies. The concept of the exposome was first proposed by Wild in 2005, and characterized as the comprehensive evaluation of all exposures and their contribution to disease causation or progression by Rappaport and Smith (Rappaport and Smith 2010; Wild 2005). A branching paradigm, the public health exposome, focuses on the impact of exposures on the overall health of a population within a particular region, with the intention of identifying vulnerable populations with higher risks of chronic illnesses (Juarez et al. 2014; Smith et al. 2015). The use of omics methods such as transcriptomics, proteomics, and metabolomics have been recommended in exposomics studies to identify the links between exposures and health outcomes, understand the mechanisms of disease development and progression, and potentially developing new biomarkers for exposure and early health effects (Vineis et al. 2013; Wild 2009).

1.3 Metabolomics

Recently, metabolomics was proposed to quantitatively measure exogenous chemicals and biological responses in order to provide "a snapshot measure of an individual's exposome" (Pennell 2016). Chadeau-Hyam et al. also proposed using metabolomics to identify "intermediate biomarkers" that could connect exposure with early health effects using the "meet-in-the-middle" approach (Chadeau-Hyam et al. 2011). Metabolites are the endpoint of biochemical activities and the metabolome could best reflect the effects of exposures and correlate with phenotype, since it is more sensitive to perturbations than transcriptome and proteome (Kell et al. 2005; Patti et al. 2012). Due to the complex chemical properties of metabolites, it is not possible to use one single

analytical platform to assess the complete metabolome within one biospeciman (Blow 2008). Gas chromatography (GC) -based methods detect metabolites that are of lower molecular weight and relatively polar metabolite classes, while liquid chromatography (LC) -based methods can be used to detect metabolites with higher molecular weight and medium-to-high lipophilicity (Dunn et al. 2011). Bouatra et al. and Psychogios et al. applied multiple metabolomics platforms including NMR, GC-MS, and LC-MS to assess urine and serum samples, respectively, and found relatively small overlaps in the metabolite profiles identified using different platforms and biospecimen (Bouatra et al. 2013; Psychogios et al. 2011). Urine and blood samples are most often used in metabolomics studies since the collection is relatively less invasive compared to other biological samples, and are integrative biofluids that reflect functions and phenotypes from different parts of the body (Dunn et al. 2011). So far metabolomics have been applied in pharmacology, clinical disease diagnosis, nutritional, and environmental studies (Robertson et al. 2011). However, most environmental metabolomics studies have focused on the assessment of single exposure (Ellis et al. 2012).

1.4 Lipidomics

Lipidomics is considered a sub-discipline of metabolomics that focuses on systemic analysis of lipids and their interacting partners (Wenk 2005). Deregulation of lipid profiles have been associated with disease onset and progression, and is often applied in clinical studies in search for novel biomarkers or understanding pathological mechanisms in diseases such as cancer, liver diseases, cardiovascular disease, diabetes, kidney diseases, and Alzheimer's disease (Stegemann et al. 2014; Yang et al. 2016; Zhao et al. 2015). When lipidomics approach was applied in exposure studies, most were based on animal models and focused on single toxic exposure (Chi et al. 2019; Hu et al. 2018; Li

11

et al. 2019).

1.5 Oxidative stress

Oxidative stress is an imbalance between production of free radicals, reactive oxygen species (ROS), and nitrogen reactive species (RNS), and their reduction by protective antioxidants. Accumulation of this imbalance can invoke early health effects such as inflammation, lipid peroxidation, and DNA damage. Oxidative stress is prevalent in CKD patients, elevated in hyperlipidemia patients, contributes to liver fibrosis, and mediates allergic respiratory diseases (Bowler and Crapo 2002; Hopps et al. 2010; Poli 2000; Xu et al. 2015). Oxidative stress also interacts with all three stages of cancer process: cancer initiation, cancer promotion, and cancer progression through ROS and RNS induced DNA damage, lipid peroxidation, and protein damage (Reuter et al. 2010; Valavanidis et al. 2009).

1.6 Serum acylcarnitines

Serum acylcarnitines are involved in transporting fatty acids into the mitochondria for β-oxidation and production of energy (Semba et al. 2017). Deregulations in serum acylcarnitines can activate inflammatory signaling pathways, and have been associated with chronic diseases including cancer, cardiovascular diseases, CKD, and Alzheimer's disease (Fouque et al. 2006; Ruiz-Canela et al. 2017; Rutkowsky et al. 2014; Toledo et al. 2017; Zhou et al. 2012b).

12

2. Objectives

Figure 1 shows the study framework for this thesis, and our objectives are to:

1. Establish comprehensive urine and serum metabolite profiles of children and elderly residents living near No. 6 Naphtha Cracking Complex.

2. Identify deregulations in biological mechanism associated with exposure to multiple industrial pollutants.

3. Use metabolomics to find the link connecting multiple industrial pollutants exposures with early health effect.
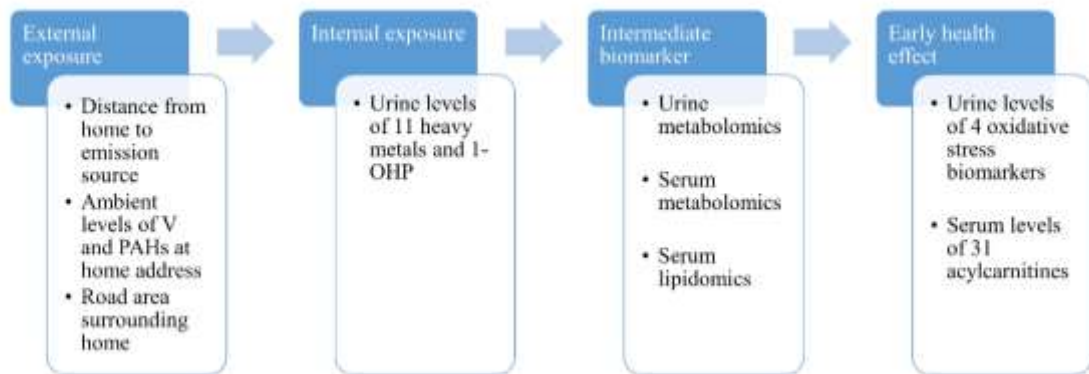
Figure 1. Study framework

3. Material and Methods

3.1 Study area and subjects

Our study area surrounded the largest petrochemical complex in Taiwan, No. 6 Naphtha Cracking Complex, located in Yunlin County on the west coast of central Taiwan. The complex began major operations in 1999. To date, the complex covers a total area of 2603 ha, housing 64 plants including one coal-fired power plant that generates 1800 MW of power, three oil refiner plants that processes 450,000 barrels of crude oil every day, two naphtha cracking plants that produces 160 million tons of ethylene per year, and three cogeneration plants that generates 2820 MW of power. Our subjects were selected from a prospective cohort of 3,230 residents who have lived in three townships in close vicinity to the complex (pink) and seven other townships further away (yellow) as shown in Figure 2 for more than five years. All 3,230 subjects (aged 5–88) have completed interview-administered questionnaire surveys including key factors related to exposure, provided one morning spot urine sample, and one fasting blood sample. Urine samples were stored at -20 $^\circ$C, serum samples were extracted from coagulated blood samples using centrifuge and stored at -80 $^\circ$C. All 3,230 participants' urine samples have been analyzed for As, Cd, Cr, Ni, Pb, Hg, V, Mn, Cu, Sr, Tl, and 1-OHP. We used urine concentrations of V and 1-OHP as well as residential address to identify 257 cohort members who lived in the three townships closest to the complex, with urine concentrations previously established petrochemical industrial exposure biomarkers V and 1-OHP in the top 60 % of the 3,230 residents as high exposure group, and another 337 cohort members who lived in townships further away, with urine concentrations of V and 1-OHP in bottom 40 % of the cohort as low exposure group. We then randomly select 43 children (aged 9–15) and 77 elderly (aged > 55) subjects from the 257 cohort members as our high exposure subjects,

14

and 75 children and 78 elderly participants from the 337 cohort members as our low exposure subjects. This study was approved by the Research Ethics Committee of National Taiwan University Hospital, and informed consent was obtained from each participant.
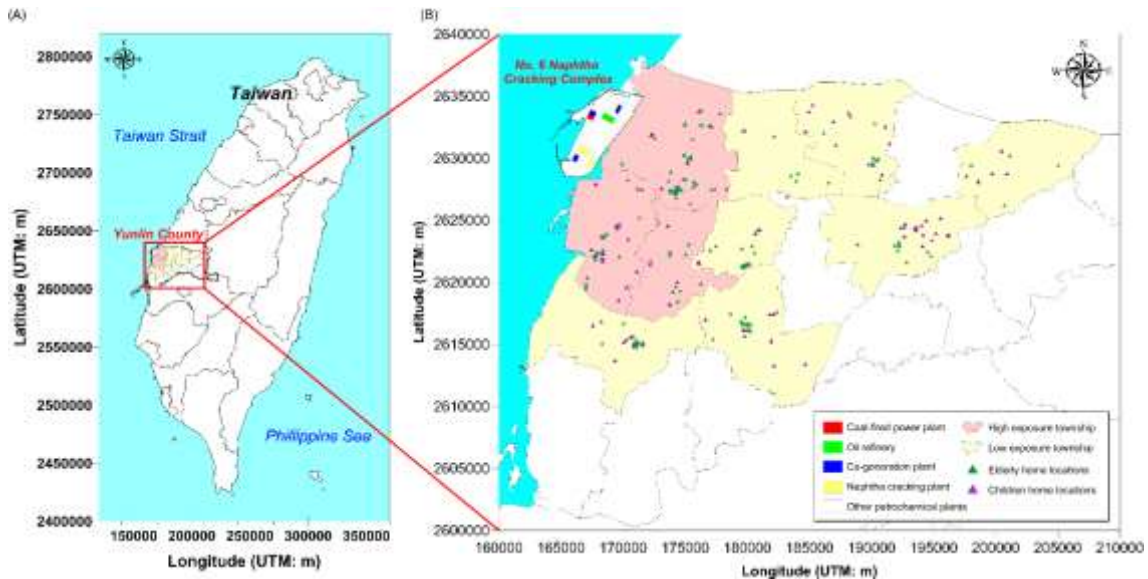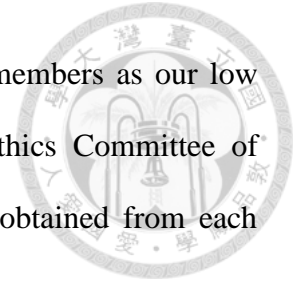
Figure 2. GIS map of (A) Yunlin County in central Taiwan and (B) location of study area, petrochemical plants, and 273 study subjects' homes

3.2 External exposure

Geographic coordinates for all 273 participant's home address were determined, and geological information system (GIS) software (ArcGIS version 10.1) was used to calculate the distances from each home address to previously identified main emission points of coal-fired power plant and oil refineries, respectively. GIS software was also used to measure road area surrounding homes, in order to estimate traffic contribution on air toxics levels. Ambient concentrations of V and five PAHs including pyrene, fluoranthene, dibenzo[a,h]anthracene, benzo[k]fluoranthene, and benzo[a]anthracene were calculated using previously established two-stage dispersion model and kriging method model, respectively (Chio et al. 2014; Yuan et al. 2015b).

15
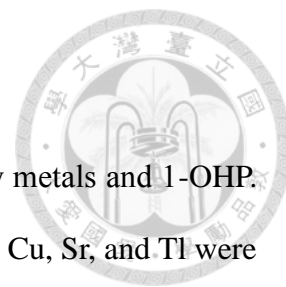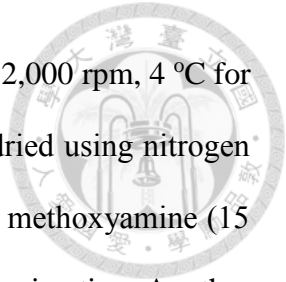
3.3 Internal exposure

All 273 participants' urine samples had been analyzed for heavy metals and 1-OHP. Urine concentrations of heavy metals As, Cd, Cr, Ni, Pb, Hg, V, Mn, Cu, Sr, and Tl were analyzed using previously reported inductively coupled plasma mass spectrometry (ICP-MS) method, and 1-OHP was analyzed using previously reported high performance liquid chromatography (HPLC) method. For heavy metals, spikes were examined to confirm measurement stability, and standard reference materials (SRM) for each metal were analyzed to assess accuracy. For 1-OHP analysis, detection limit was 0.01 ng/mL with an 89.6 % recovery rate and a 4.0 % coefficient of variation for repeated measurements. Urine concentration of exposure biomarkers below the method detection limit (MDL) was replaced by half of the MDL. Urinary creatinine analysis was conducted using enzyme-linked immunosorbent assay at National Taiwan University Hospital medical diagnosis laboratory and used for adjustment of urinary exposure biomarker levels.

3.4 Metabolomics

3.4.1 Urine metabolomics

252 participants had urine samples available for urine metabolomics analysis. Samples were prepared and analyzed following protocols derived from previous publications, using a Pegasus 4D GC $\times$ GC $-$ TOFMS (LecoCorp., St. Joseph, MI, USA) for analysis (Figure 3) (Chan et al. 2011; Pasikanti et al. 2013). Briefly, 200 μL urine sample was centrifuged at 12,000 rpm for 10 min at 4 °C. 50 μL supernatant was then transferred to a new Eppendorf tube. 10 μL urease (30 unit/10 μL) was added, vortexed thoroughly, and incubated at 37 °C for 1 hr. Internal standards 10 μL heptadecanoic acid (1 mg/mL) and 10 μL 2-chlorophenylalanine (0.3 mg/mL) were added to the mixture. 620

μL methanol was then added, vortexed for 1 min, then centrifuge at 12,000 rpm, 4 °C for 5 min. 550 μL supernatant was then transferred to a glass vial and dried using nitrogen gas. Derivatization steps involved adding 80 μL derivatization agent methoxyamine (15 mg/mL) to the dried sample and incubate at 60 °C for 2 hr for methoximation. Another derivatization agent was added for trimethylsililation, 80 μL N,O-Bis(trimethylsilyl)trifluoroacetamide (BSTFA), and the mixture incubated at 70 °C for 1hr. Derivatized sample was then centrifuged at 6,000 rpm for 3 min, then 100 μL supernatant was transferred to a glass insert, ready for GC analysis. Chromatographic separations were performed using a 30 m×250 μm (i.d.)×0.25 μm RXI-5 column fused together with a 2 m×180 μm (i.d.)×0.2 μm RTX-200 column as primary and secondary columns, respectively (RestekCorp., Bellefonte, PA, USA). Carrier gas was helium, at 1.5 mL/min constant flow rate. Primary oven temperature program began at 70 °C for 0.2 min, and increased at 5 °C/min to 270 °C where it was held for 7.5 min. Secondary oven temperature was maintained at 10 °C higher than the primary oven throughout the program. Thermal modulator was set at 20 °C higher than the primary oven. Modulation time was set at 4 sec with hot pulse of 0.8 sec and 1.2 sec cool time between stages. The inlet, transfer line, and ion source temperatures were set at 220, 200, and 250 °C, respectively. The mass spectrometer source was operated in EI mode with an electron energy of 70 eV at the detector voltage of 1450 V. Data was acquired over the range of m/z 40-600 at an acquisition rate of 100 Hz. For quality control (QC), blanks and pooled QC samples were analyzed at the beginning of each batch, and every five samples within batch. After data cleaning and preprocessing, acquired data were outputted as peak area for NIST library (version 08, National Institute of Standards and Technology, Gaithersburg, MD, USA) identified potential metabolite peaks, and we selected those with mass spectrum matched library spectrum for > 60 % (similarity score > 600 out of

17

1000). Obtained potential metabolite features were preprocessed by removing those with > 50 % missing values, and replacing the missing values of the remaining features with half of the minimum positive value in the original data. Preprocessed data were normalized by sum of total peak area, log transformed, and autoscaled (mean-centered and divided by the standard deviation of each variable) prior to further statistical analysis. NIST library match showed the derivatized form of potential metabolite peaks, which we converted to underivatized forms by replacing methoxyamine and trimethylsilyl groups with carbonyl and hydroxyl functional groups, respectively. ChemSpider was used for the identification of the underivatized forms of potential metabolite peaks (Royal Society of Chemistry, London, UK). All potential metabolites peaks were then put through an online repository (http://cts.fiehnlab.ucdavis.edu/) and searched under three online databases: the Human Metabolome Database (HMDB) (Wishart et al. 2018), Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al. 2014), and Chemical Entities of Biological Interest (ChEBI) (Hastings et al. 2013), for identification of known metabolites, chemical class, and involved biological pathways.
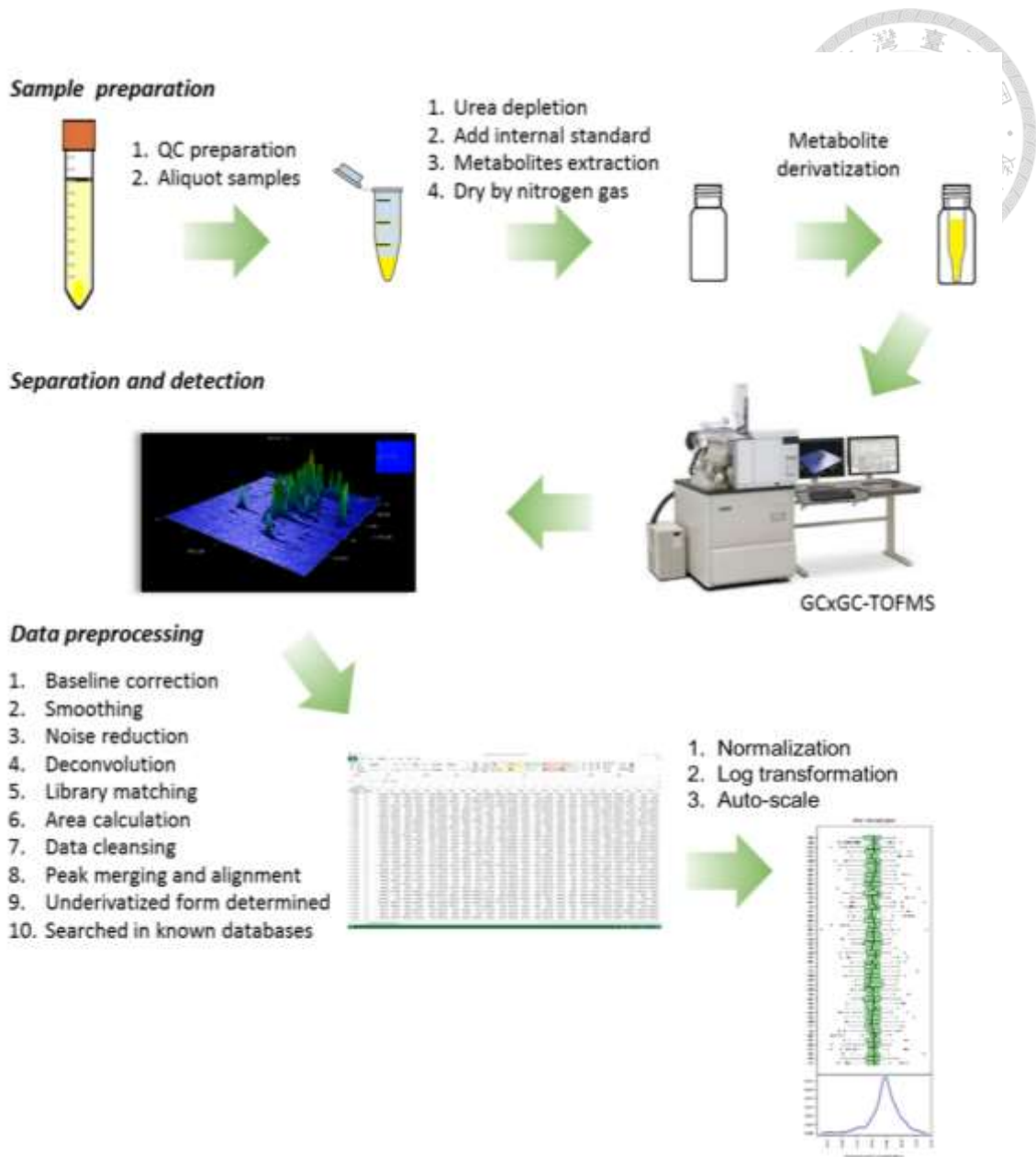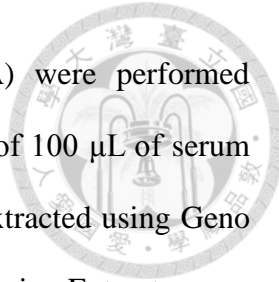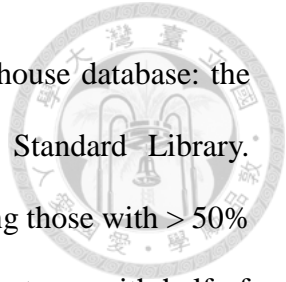
Figure 3. Flowchart of urine metabolomics analysis

3.4.2 Serum metabolomics

252 participants had serum samples available for serum metabolomics analysis. Serum metabolomics analysis was conducted by The Metabolomics Core Laboratory, Center of Genomic Medicine, National Taiwan University. Sample preparation and analytical method using Agilent 1290 UHPLC system coupled with 6540-QTOF

19
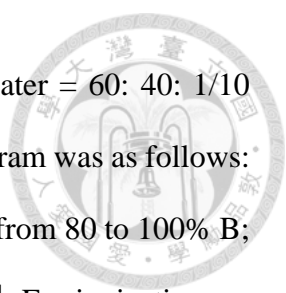
(UHPLC-QTOF) (Agilent Technologies, Santa Clara, CA, USA) were performed following previously reported protocols (Lai et al. 2015). Aliquots of 100 μL of serum samples were added to extraction solvents by the ratio of 1:4 and extracted using Geno Grinder 2010 (SPEX, Pittsburg, CA, USA) at 1000 rpm for 2 min. Extracts were centrifuged at 15000 g for 5 min at 4 °C. Supernatant was then collected and evaporated using centrifugal vaporizer (EYELA, Tokyo, Japan) for 2 h. Residue was reconstituted in 200 μL of 50% methanol and centrifuged at 15000 g for 5 min, and filtered with 0.2 μm Minisart RC 4 filter (Sartorius, Goettingen, Germany) before analysis. A total of 2 μL of sample was injected into an ACQUITY UPLC HSS T3 column (2.1 × 100 mm; 1.8 μm) (Waters, Milford, MA, USA). The mobile phase was composed of solvent A (water/0.1% formic acid) and solvent B (acetonitrile/0.1% formic acid). Gradient profile used was 0−1.5 min, 2% B; 1.5−9 min, linear gradient from 2 to 50% B; 9−14 min, linear gradient from 50% to 95% B; 14−15 min, 95% B; the column was then re-equilibrated. Flow rate was maintained at 0.3 mL/min. Column oven and auto-injection system temperatures were set at 40 °C and 10 °C, respectively. Jet Stream electrospray ion source with capillary voltage of 4 kV in positive mode and 3.5 kV in negative mode was used for sample ionization. For MS parameters, dry gas temperature, 325 °C; dry gas flow, 5 L/min; nebulizer, 40 psi; sheath gas temperature, 325 °C; sheath gas flow, 10 L/min; and fragmentor, 120 V were used. Scan range was set at m/z 50−1700. For QC, blanks and pooled QC samples were analyzed at the beginning of each batch, and every five samples within batch. Synthetic samples containing 40 chemical standards (QC standard) were analyzed at the beginning of each batch to check instrument performance. Three repeated analysis was performed for each sample and total ion chromatogram was manually checked for technical replicates. Acquired data was preprocessed using True Ion Pick (TIPick) algorithm for background subtraction and peak picking (Ho et al. 2013). Peak

identification was conducted by matching m/z to an established in-house database: the National Taiwan University MetaCore Metabolomics Chemical Standard Library. Obtained potential metabolite features were preprocessed by removing those with > 50% missing values, and replacing the missing values of the remaining features with half of the minimum positive value in the original data. Preprocessed data were normalized by sum of total peak area, log transformed, and autoscaled (mean-centered and divided by the standard deviation of each variable) prior to further statistical analysis.

3.4.3 Serum lipidomics

252 participants had serum samples available for serum lipidomics analysis. Serum lipidomics analysis was conducted by The Metabolomics Core Laboratory, Center of Genomic Medicine, National Taiwan University. Lipidomic profiling for seurm samples was performed using an Agilent 1290 UHPLC system coupled with a Bruker maXis impact QTOF (Bruker Daltonik, Bremen, Germany). 100 uL of serum sample was mixed with 100 uL water and 1000 uL methanol/chloroform (1:2) for extraction. After shaking at 1000 rpm for 5 minutes by using Geno/Grinder 2010 (SPEX SamplePrep., Metuchen, NJ, US), the extract was centrifuged by using Eppendorf Centrifuge 5810R at 15000 rcf for 5 minutes at 4 °C. 600 ul of the lower organic layer was taken to another eppendorf tube for dryness by nitrogen. Dried residues were reconstituted in 120 uL methanol, sonicated 10 minutes and centrifuged at 15000 rcf for 5 minutes at 4 °C. The supernatant were filtered through a 0.2 μm Minisart RC 4 filter (Sartorius Stedim Biotech GmbH, Goettingen, Germany) and subjected to LC-MS/MS analysis. 2 μL of sample from serum extract was injected into an Agilent ZORBAX Eclipse Plus C18 column (2.1 x 100 mm, 1.8μm, Agilent Technologies, Santa Clara, CA); the analytical column was maintained at 55 °C. The mobile phase was composed of solvent A (water/10 mM ammonium

21
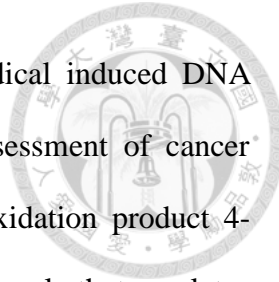
acetate/0.1% formic acid) and solvent B (methanol: isopropanol: water = 60: 40: 1/10 mM ammonium acetate/0.1% formic acid). The gradient elution program was as follows: 0–2 min: linear gradient from 35 to 80% B; 2–7 min: linear gradient from 80 to 100% B; and 10 min maintenance in 100% B. The flow rate was 350 μL min^{-1}. For ionization, an electrospray ionization source was used for sample ionization. The following parameters were used throughout the study: 180 °C dry gas temperature, 8 L min^{-1} dry gas flow, 2.0 bar nebulizer, 500 V end plate offset, 4500 V in positive mode for capillary voltage. The mass scan range and acquisition rate were m/z 150–1600 and 2 Hz, respectively. PITracer algorithm was applied for data format conversion, relative mass difference estimation, chromatogram generation, and peak detection (Wang et al. 2015). Peak identification was performed by matching m/z to an established in-house database: the National Taiwan University MetaCore In-House Lipidomics Library. Preprocessed data were normalized by sum of total peak area, log transformed, and autoscaled (mean-centered and divided by the standard deviation of each variable) prior to further statistical analysis.

3.5 Early health effects

3.5.1 Oxidative stress
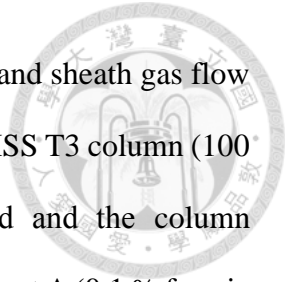
252 participants had urine samples available for urine oxidative stress biomarkers analysis. Urine concentrations of four oxidative stress biomarkers 8-hydroxy-2'-deoxyguanosine (8-OHDG), 4-hydroxy-2-nonenal-mercapturic acid (HNE-MA), 8-isoprostaglandin $F_{2\alpha}$ (8-isoPF$_{2\alpha}$), and 8-nitroguanine (8-NO$_2$Gua) were analyzed using previously published methods and adjusted with urinary creatinine concentrations (Wu et al. 2016). QC was conducted following European Medicines Agency guidelines (EMA 2011). The four oxidative stress biomarkers applied in this study represent the different effects of oxidative stress, and all four biomarkers participate in the process of

22

carcinogenesis. 8-OHDG is the most used biomarker for free radical induced DNA damage and has been reported as a good biomarker for risk assessment of cancer (Valavanidis et al. 2009). HNE-MA is a metabolite of lipid peroxidation product 4-hydroxy-2-nonenal (HNE), a cytotoxic and mutagenic signaling molecule that regulates cell cycle and forms DNA adducts leading to DNA damage (Ayala et al. 2014; Bartsch and Nair 2005; Valavanidis et al. 2009). Urine levels of 8-isoPF$_{2\alpha}$ is a biomarker for arachidonic acid peroxidation, and has been associated with increased risk of potential malignant oral disorders and breast cancer progression (Ferroni et al. 2017; Senghore et al. 2018). 8-NO$_2$Gua is formed from DNA damage induced by RNS generated under inflammatory conditions, and reported to participate in carcinogenesis (Hiraku 2010).

3.5.2 Serum acylcarnitines

252 participants had serum samples available for serum acylcarnitines analysis. Serum acylcarnitines were analyzed by The Metabolomics Core Laboratory, Center of Genomic Medicine, National Taiwan University. Serum levels of 31 acylcarnitines were analyzed using Agilent 1290 UHPLC coupled with an Agilent 6460 triple quadrupole mass spectrometer (Agilent Technologies). A total of 400 μL of methanol (Scharlau, Sentmenat, Spain) was added into 100 μL of human serum to extract metabolites. The extraction was performed on Geno/Grinder2010 (SPEX, Metuchen, NJ, US) at 1,000 rpm for 2 min followed by centrifugation at 15,000 rcf for 5 min at 4 ℃. Supernatant was collected and evaporated using EYELA CVE-200D Centrifugal Evaporator (TOKYO RIKAKIKAI CO., Tokyo, Japan) until dry. Dried extracts were reconstituted with 200 μL of 10 % methanol and centrifuged at 15,000 rcf for 5 min. The supernatant was then filtered with 0.2-μm Ministart RC 4 filter (Sartorius, Goettingen, Germany). All aliquots were transferred to glass inserts prior to UHPLC-MS analysis. The MS parameters were 325 ℃, 325 ℃, 500 V, and 3500 V for the drying gas temperature, sheath gas temperature,

nozzle voltage, and capillary voltage, respectively. The dry gas flow and sheath gas flow were 7 and 11 L min^{-1}, respectively. The nebulizer was set at 45 psi. HSS T3 column (100 x 2.1 mm, 1.8 μm, Waters, Milford, MA, USA) was employed and the column temperature was set at 40 ℃. The mobile phase was composed of solvent A (0.1 % formic acid in DI water) and solvent B (0.1 % formic acid in ACN) (J.T. Baker, Phillipsburg, NJ, USA). The gradient elution program was as followed: 0–1.5 min: 2% B; 1.5–9 min: linear gradient from 2 to 50% B; 9–14 min: linear gradient from 50 to 95% B; and 3 min maintenance in 95% B with the flow rate of 0.3 ml min^{-1}. A 3 min equilibrium was used before next injection. All analytes were monitored in positive MRM mode. All the peaks were integrated with MassHunter Quantitative Analysis software (Agilent Technologies). Pooled QC sample was analyzed every 20 samples and calculated for relative standard deviation (RSD). Out of the 31 analyzed acylcarnitines, 29 had RSD < 20%, and was used for statistical analysis of samples. Preprocessed data were normalized by sum of total peak area, log transformed, and autoscaled (mean-centered and divided by the standard deviation of each variable) prior to further statistical analysis.

3.6 Pathway analysis

Pathway analysis was performed using Metaboanalyst 4.0, which currently supports 80 pathways in the Homo sapiens pathway library (Chong et al. 2018). HMDB ID number and normalized peak area values were used as input. The method "Globaltest" was used for pathway enrichment analysis, and "betweenness centrality" for pathway topology analysis.

3.7 Meet-in-the-middle

Partial least squares discrimination analysis (PLS-DA) was performed using

24

Metaboanalyst 4.0 (The Metabolomics Innovation Center, Edmonton, Alberta, Canada) to identify exposures-related potential metabolites. PLS-DA models were validated using permutation test and cross-validation test. We further used Student's *t* test adjusted for false discovery rate (FDR) q value to compare the peak area of each potential metabolite between high and low exposure groups for urine metabolomics results, and analysis of covariance (ANCOVA) adjusted for age, sex, and BMI for serum metabolomics and lipidomics results. Pearson's correlation analysis was conducted to identify early health effect-related potential metabolites in urine metabolomics, and linear regression analysis was conducted for serum metabolomics and lipidomics, adjusting for age, sex, and BMI (Figure 4).
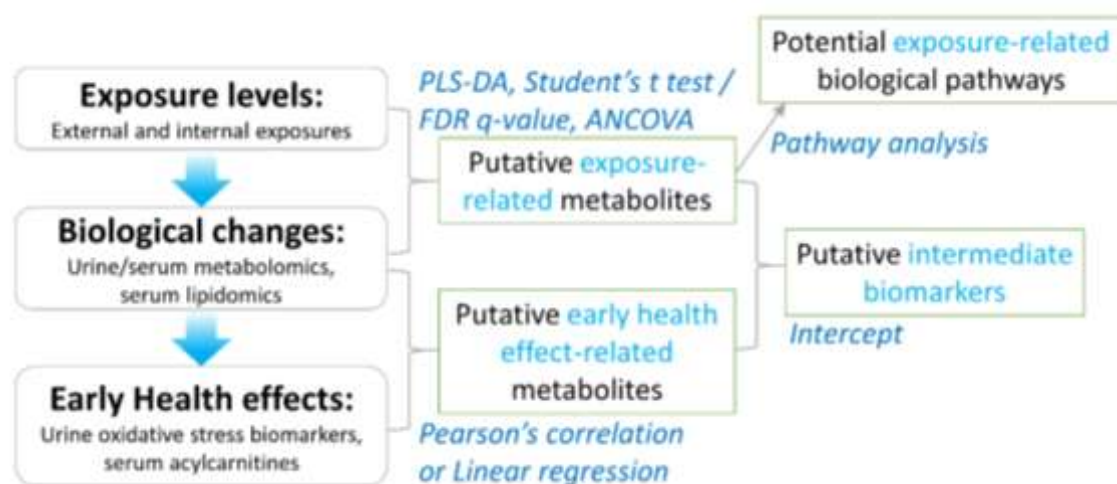
Figure 4. Meet-in-the-middle approach.

3.8 Association between exposure and early health effects

Individual association between eight carcinogens and four oxidative stress biomarkers were analyzed using linear regression analysis, while association of eight combined carcinogen exposures with four oxidative stress biomarkers were analyzed using weighted quantile sum (WQS) regression, both adjusted for age, sex, and BMI. The weighted contribution of quantile-scored exposures were derived based on bootstrap

25

sampling (N=100).

3.9 Statistical analysis

For comparison of basic characteristics and external exposure levels between high and low exposure groups, we used Student's t test to analyze continuous variables, and Chi-squared test or Fisher's exact test for discrete variables. Urine concentrations of internal exposure biomarkers were adjusted using urine creatinine concentrations and log transformed before comparing between high and low exposure groups using ANCOVA adjusting for age, sex, smoking, alcohol consumption, betel nut chewing, fish consumption, and BMI with a post comparison by Scheffe test. Oxidative stress biomarkers were adjusted using urine creatinine concentrations and log transformed before comparing between high and low exposure groups using Student's t test. Peak area of urine metabolite features were normalized before comparing between high and low exposure groups using Student's t test and FDR q value. Peak area of serum metabolite features, lipid features, and acylcarnitines were normalized before comparing between high and low exposure groups using ANCOVA, adjusting for age, sex, and BMI with a post comparison by Scheffe test. Student's t-test, Chi-squared test, Fisher's exact test, ANCOVA test, and linear regression analysis were performed using SAS 9.4 for Windows. PCA and PLS-DA were performed using Metaboanalyst 4.0. FDR q value was measured using fdrtool package in R 3.1.3 for Windows. WQS regression analysis was conducted using the gWQS package in R 3.5.1 for Windows.

4. Results and Discussion

4.1 Part 1. Linking sources to early effects by profiling urine metabolome of residents living near oil refineries and coal-fired power plants

4.1.1 Results

For 252 subjects used in urine metabolomics study, the comparison between high and low exposure groups in basic characteristics, external exposure at each subject's home locations, and internal exposures as urine biomarkers concentrations is shown in Table 2. Overall, high exposure subjects lived 10.07 ± 2.43 km away from the main emission point of the coal-fired power plant and 9.35 ± 2.65 km away from the main emission point of oil refineries, while low exposure subjects lived 21.64 ± 5.19 and 20.69 ± 5.00 km away from the two main emission points, respectively. High and low exposure subjects in children (13.76 ± 0.93 years old) and elderly (65.88 ± 6.92 years old) age groups showed no significant difference in gender distribution, smoking, drinking, and betel nut chewing history, body mass index, and systolic blood pressure. Ambient concentrations of V, and three PAHs pyrene, fluoranthene, and dibenzo[a,h]anthracene were significantly higher at the home locations of high exposure subjects when compared to low exposure subjects, for both children and elderly participants. Another PAH, benzo[k]fluoranthene, was significantly increased in high exposure group in elderly subjects, but showed no difference between high and low exposure groups in children subjects. Benzo[a]anthracene was decreased in high exposure groups for both children and elderly subjects. Road areas surrounding participants' homes, which we used to represent traffic contribution of air toxics levels, showed no difference between high and low exposure groups for both children and elderly residents at 25 m buffer. When we increased the buffer to 500 m, elderly subjects in the low exposure group had larger road areas

27

surrounding their homes than those in the high exposure group. Urine concentrations of 1-OHP, V, Ni, Cu, As, Sr, Cd, Hg, and Tl were significantly increased in high exposure groups compared to low exposure groups for both children and elderly subjects. The difference between high and low exposure groups was most profound for V, 1-OHP, and Tl, followed by Sr. Pearson's correlation analysis results showed significant correlation between ambient and urinary V levels; and between ambient pyrene, fluoranthene, and dibenzo[a,h]anthracene and urine 1-OHP concentrations for both children and elderly subjects (Table 3).

28

Table 2. Comparison of basic characteristics and exposure levels in 252 study subjects

	Children			Elderly		
	High (n=40)	Low (n=70)	*p*	High (n=71)	Low (n=71)	*p*
Basic Characteristics						
Age (years), mean±SD	13.78 ±0.93	13.83 ±0.89	0.88	66.23 ±6.54	66.36 ±7.47	0.76
Male, n (%)	22 (55.0)	38 (54.3)	0.94	28 (39.4)	35 (49.3)	0.24
Smoking history, n (%)	3 (7.5)	3 (4.3)	0.67	3 (4.2)	7 (9.9)	0.33
Drinking history, n (%)	2 (5.0)	1 (1.4)	0.30	8 (11.3)	7 (9.9)	0.79
Betel nut chewing history, n (%)	1 (2.5)	0 (0)	0.60	4 (5.6)	3 (4.2)	1.00
BMI (kg/m^2), mean±SD	21.13 ±3.20	20.05 ±3.43	0.10	26.30 ±3.89	26.36 ±3.35	0.93
SBP (mmHg), mean±SD	117.63 ±13.02	115.79 ±14.26	0.50	140.76 ±20.74	141.76 ±18.95	0.77
External Exposures at study subjects' homes[b], mean±SD						
Distance to coal-fired power plant	10.31 ±2.50	22.66 ±10.31	<0.05	9.94 ±2.39	20.62 ±4.72	<0.05
Distance to oil refinery	9.80 ±2.63	21.73 ±5.20	<0.05	9.09 ±2.65	19.66 ±4.59	<0.05
Road area surrounding homes						
25 m buffer	241.08 ±279.23	279.23 ±226.64	0.38	269.41 ±222.79	304.52 ±237.50	0.37
500 m buffer	68088.07 ±61248.67	61248.67 ±20502.19	0.18	71269.64 ±22808.14	82235.11 ±27217.15	<0.05
Ambient concentrations						
Vanadium	8.60 ±1.39	5.75 ±1.08	<0.05	8.97 ±1.63	6.22 ±0.96	<0.05
Polycyclic Aromatic Hydrocarbons[c]						
Pyrene	0.026 ±0.005	0.023 ±0.003	<0.05	0.030 ±0.005	0.022 ±0.003	<0.05
Fluoranthene	0.028 ±0.001	0.026 ±0.003	<0.05	0.027 ±0.001	0.024 ±0.003	<0.05
Dibenzo[a,h]anthracene	0.013 ±0.002	0.011 ±0.001	<0.05	0.014 ±0.002	0.011 ±0.001	<0.05

29

	Children			Elderly		
	High (n=40)	Low (n=70)	*p*	High (n=71)	Low (n=71)	*p*
Benzo[k]fluoranthene	0.017 ±0.003	0.017 ±0.002	0.98	0.019 ±0.003	0.018 ±0.002	<0.05
Benzo[a]anthracene	0.017 ±0.001	0.019 ±0.001	<0.05	0.017 ±0.001	0.020 ±0.002	<0.05
Internal Exposures[d], mean±SD						
1-hydroxypyrene	0.25 ±0.31	0.03 ±0.01	<0.05	0.42 ±0.70	0.03 ±0.01	<0.05
Vanadium	2.34 ±1.53	0.23 ±0.10	<0.05	4.02 ±2.23	0.17 ±0.08	<0.05
Chromium	3.89 ±4.56	2.06 ±1.65	0.11	5.32 ±7.33	2.98 ±2.62	0.09
Nickel	10.41 ±16.62	3.70 ±2.89	<0.05	11.28 ±15.34	8.33 ±29.64	<0.05
Copper	16.38 ±14.94	11.22 ±7.50	<0.05	22.87 ±24.33	17.36 ±30.23	<0.05
Arsenic	62.28 ±42.02	39.47 ±29.46	<0.05	119.60 ±205.04	64.92 ±51.73	<0.05
Strontium	170.77 ±249.93	70.47 ±64.55	<0.05	211.26 ±176.33	86.53 ±55.23	<0.05
Cadmium	0.34 ±0.34	0.19 ±0.15	<0.05	1.30 ±1.11	0.87 ±0.73	<0.05
Mercury	3.30 ±3.15	1.92 ±1.81	<0.05	2.59 ±2.35	1.49 ±1.27	<0.05
Thallium	2.13 ±3.92	0.21 ±0.11	<0.05	1.60 ±3.67	0.12 ±0.08	<0.05
Lead	0.77 ±1.08	0.65 ±0.59	0.71	1.79 ±2.30	1.09 ±1.21	0.17

[a] Comparison of basic characteristics between the high and low exposure groups for continuous variables was performed using Student's t-test, and for discrete variables, Chi-squared test or Fisher's exact test. Urinary exposure biomarker concentrations are log-transformed, high and low exposure groups compared by ANCOVA test adjusting age, gender, smoking, alcohol consumption, betel nut chewing, fish consumption, and source of drinking water with a post comparison by Scheffe test. [b] Distance to source: Average of home-to-coal-fired power plant and home-to-oil refinery distance, unit: km; Road area surrounding homes unit: m^2; Ambient V unit: ng/m^3; Polycyclic Aromatic Hydrocarbons unit: ng/m^3. [c] Children low exposure group n=35; Elderly low exposure group n=50. [d] For 1-hydroxypyrene, unit: µmol/mol-creatinine; for heavy metals, unit: µg/g-creatinine BMI: Body Mass Index; SBP: Systolic Blood Pressure

Table 3. Association between model-estimated ambient levels and analyzed urine concentrations of V and PAHs in children and elderly subjects.

Ambient / Urine	Children				Elderly			
	1-OHP		V		1-OHP		V	
	coefficient	p	coefficient	p	coefficient	p	coefficient	p
Vanadium[a]	-	-	0.634	<0.05	-	-	0.697	<0.05
PAHs[b]								
Pyrene	0.317	<0.05	-	-	0.555	<0.05	-	-
Fluoranthene	0.483	<0.05	-	-	0.521	<0.05	-	-
Dibenzo[a,h]anthracene	0.306	<0.05	-	-	0.574	<0.05	-	-
Benzo[k]fluoranthene	-0.097	0.41	-	-	0.237	<0.05	-	-
Benzo[a]anthracene	-0.594	<0.05	-	-	-0.343	<0.05	-	-

Association were assessed using Pearson's correlation analysis.

[a] Children n=110; Elderly n=142. [b] Children n=75; Elderly n=121.

Fig. 5 compared log-transformed urine concentrations of four oxidative stress biomarkers between high and low exposure groups in children and elderly subjects. In children, urine concentrations of 8-OHDG were 3.1± 2.52 and 2.59 ± 2.78 µg/g-creatinine in high and low exposure groups, respectively, with no significant statistical difference. For elderly subjects, 8-OHDG urine levels were 6.61 ± 20.34 and 3.16 ± 4.07 µg/g-creatinine for high and low exposure groups, respectively, with p value of 0.006 (Fig. 5A). Urine levels of HNE-MA were significantly increased in high exposure groups compared to low exposure groups for both children and elderly subjects (Fig. 5B), with urine concentrations 2.16±2.7 and 1.4±2.3 µg/g-creatinine for high and low exposure groups, respectively, in children, and 2.59± 3.16 and 1.82 ±3.66 µg/g-creatinine for high and low exposure groups, respectively, in elderly subjects. 8-isoPGF$_{2\alpha}$ was increased when comparing high (3.22 ± 3.4 µg/g-creatinine) to low exposure group (2.06 ± 2.14 µg/g-creatinine) in children, but the difference was not as significant in elderly subjects, with urine concentrations of high exposure group at 2.88 ± 2.94 µg/g-creatinine,

31

and low exposure group at 2.47±4.32 μg/g-creatinine (Fig. 5C). Fig. 5D showed 8-NO₂Gua was

significantly increased when comparing high and low exposure groups for both children and

elderly subjects. Urine concentrations were 6.88 ± 11.93 and 2.43 ± 2.97 μg/g-creatinine for the

high and low exposure groups, respectively, in children, while elderly subjects had 7.44±15.21

and 3.19±3.34 μg/g-creatinine for high and low exposure groups, respectively.
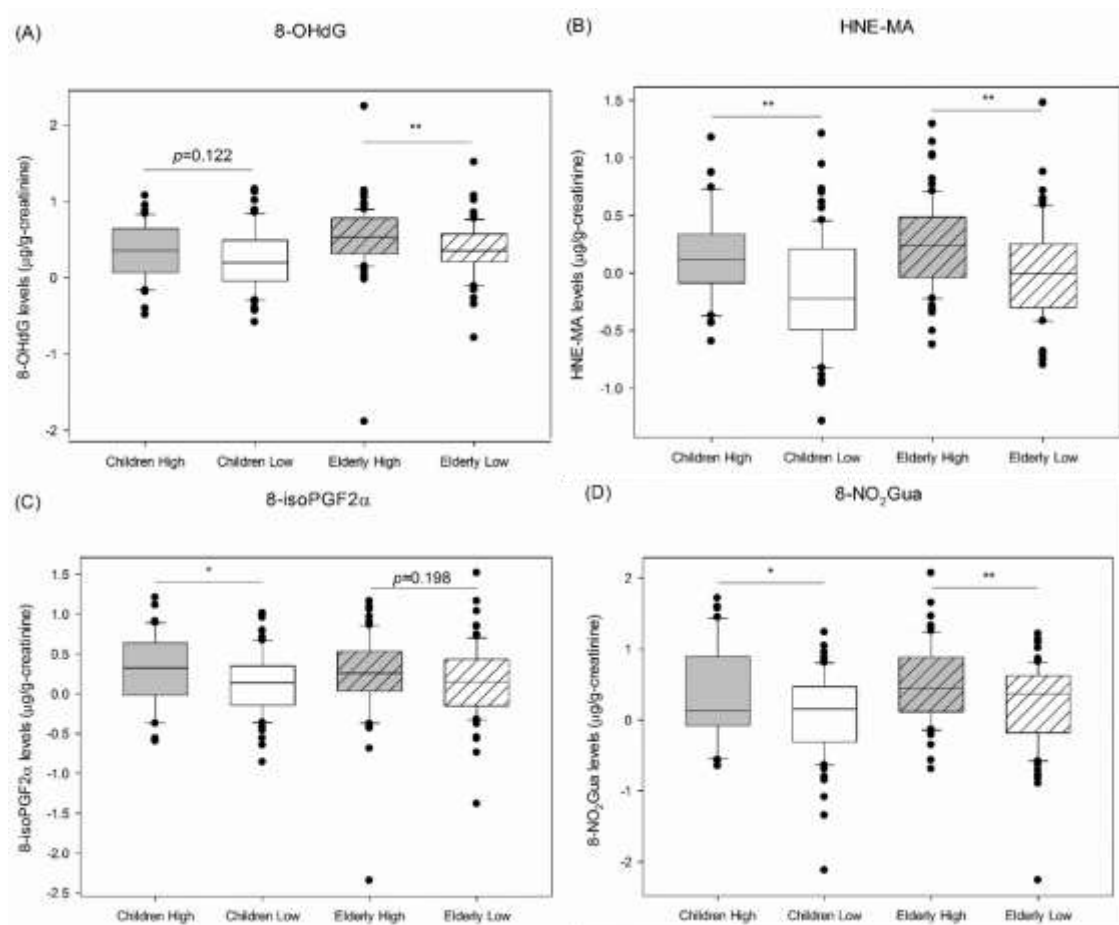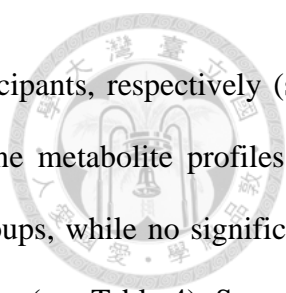
Figure 5. Urine concentrations of (A) oxidative DNA damage biomarker 8-OHDG (B)(C) lipid peroxidation biomarkers HNE-MA and 8-isoPGF₂α and (D) nitrative DNA damage biomarker 8-NO₂Gua levels between high and low exposure groups in children and elderly study subjects.
$*p<0.05$, $**p<0.01$

Urine metabolomics identified 405 potential metabolite peaks in each individual's urine

sample in children, and 391 in elderly participants, 216 and 209 of which were identified as

known human metabolites by HMDB in children and elderly participants, respectively (see Appendix 1). PLS-DA results showed separation between the urine metabolite profiles of children and elderly subjects, as well as high and low exposure groups, while no significant difference was found between metabolite profiles of different genders (see Table 4). Separate analysis of urine metabolite profiles in children and elderly participants showed clear separation of urine metabolite profiles between high and low exposure groups in both children (Fig. 6A) and elderly subjects (Figure 6B). Permutation test confirmed the validity of the PLS-DA models (see Figure 6C and 6D), and cross-validation test results showed that for both children and elderly subjects, best performance of the PLS-DA models were acquired after applying three components (see Figure 6E and 6F). The variable importance in the projection (VIP) value of each potential metabolite peak in first three components was listed in Table 4. Potential metabolite peaks with Component 1 VIP score > 1 were considered responsible for the separation between high and low exposure groups, 45 of these exposure-related potential metabolites were identified in children (FDR q < 0.1, p ≤ 0.033), and 42 in elderly subjects (FDR q < 0.15, p ≤ 0.040), of which only 11 were found in both children and elderly subjects. These 76 potential metabolites were classed by HMDB as 9 benzenoids, 3 hydrocarbons, 10 lipids and lipid-likemolecules, 19 organic acids and derivatives, 14 organoheterocyclic compounds, 2 organonitrogen compounds, 15 organooxygen compounds, 1 organosulfur compound, 2 phenylpropanoids and polyketides, and 1 homogeneous non-metal compound (Table 5).
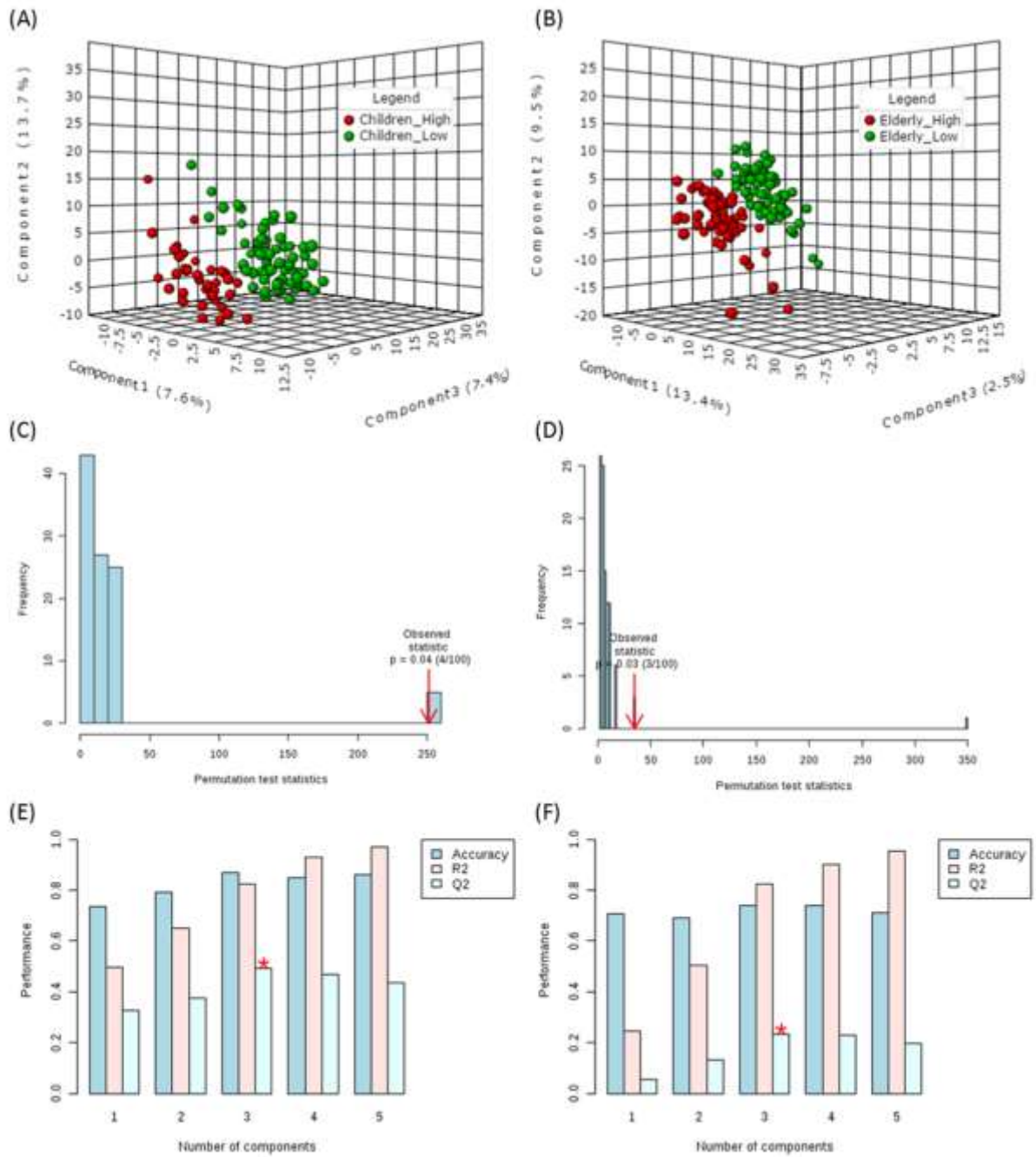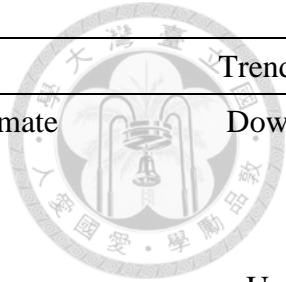
Figure 6. PLS-DA results for children and elderly urine metabolite profile analysis shown as (A)(B) PLS-DA score plots, (C)(D) permutation test results, and (E)(F) cross validation results.

34

Table 4. PLS-DA model validation results between different gender, age and exposure groups.
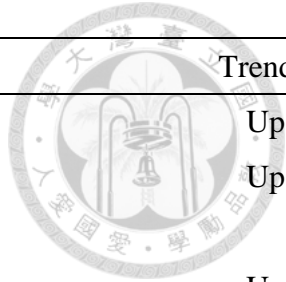
	n	Cross-validation			Permutation test
		accuracy	R^2	Q^2	(p value)
Male vs. Female					
Total	252	0.642	0.494	0.031	0.20
Children	110	0.602	0.223	-0.080	0.84
Elderly	142	0.618	0.897	0.056	0.48
Children vs. Elderly					
Total	252	0.907	0.882	0.674	<0.01
High exposure	111	0.984	0.957	0.736	0.66
Low exposure	141	0.922	0.923	0.640	0.01
High vs. Low exposure					
Total	252	0.708	0.714	0.207	<0.01
Children	110	0.826	0.826	0.494	0.04
Elderly	142	0.728	0.823	0.273	0.02

Table 5. Urine metabolic profiling of multiple exposures from refineries and coal-fired power plants in children and elderly subjects using GCxGC-TOFMS analysis

Metabolite Identification [a]	Involved pathway [b]	Trend [c]
***Children*[d]**		
Benzenoids		
2-phenylpropanal	−	Up
3-Hydroxyhippuric acid	−	Down
4-Hydroxybenzoic acid	Phenylalanine metabolism	Down
Hippuric acid	Phenylalanine metabolism	Down
o-Cymene	−	Up
Hydrocarbons		
Decane	−	Up
Dodecane	None	Up
Tridecane	None	Up
Lipids and lipid-like molecules		
Azelaic acid	None	Down
Glycerol 3-phosphate	Glycerolipid metabolism	Up
Myristic acid	Fatty acid biosynthesis	Down
Stearic acid	Fatty acid biosynthesis	Up
Tiglic acid	None	Up
Organic acids and derivatives		
3-Hydroxyoctanoic acid	cAMP signaling pathway	Down
Aminomalonic acid	None	Up
Fumaric acid	Citrate cycle (TCA cycle)	Up
Hydroxypyruvic acid	Glycine, serine, and threonine metabolism	Up
L-Alpha-aminobutyric acid	Cysteine and methionine metabolism	Up
L-Aspartic acid	Alanine, aspartate, and glutamate metabolism	Down
L-Histidine	Histidine metabolism	Down
N-acetylglutamic acid	Arginine biosynthesis	Down
Succinic acid	Citrate cycle (TCA cycle)	Up
Tiglylglycine	−	Down

Metabolite Identification [a]	Involved pathway [b]	Trend [c]
γ-Aminobutyric acid	Alanine, aspartate, and glutamate metabolism	Down
Organoheterocyclic compounds		
1H-Indole-3-acetamide	Tryptophan metabolism	Up
2-Deoxy-L-ribono-1,4-lactone	−	Down
4-Pyridoxic Acid	Vitamin B6 metabolism	Down
Cyanuric acid	None	Up
DL-Tryptophan	Tryptophan metabolism	Down
L-Gulonolactone	Tryptophan metabolism	Down
Quinolinic acid	Ascorbate and aldarate metabolism	Down
Sumiki's acid	Tryptophan metabolism	Down
Organonitrogen compounds		
Dimethylamine	Methane metabolism	Up
Organooxygen compounds		
4-Deoxyerythronic acid	−	Up
Cyanic acid	Nitrogen metabolism	Down
Cyclohexanone	Caprolactam degradation	Up
Diacetone alcohol	−	Up
D-Threitol	None	Up
Gluconic acid	Pentose phosphate pathway	Down
Hex-2-ulosonic acid	None	Down
L-Sorbose	None	Up
Rhamnose	Fructose and mannose metabolism	Down
Threonic acid	Ascorbate and aldarate metabolism	Down
Phenylpropanoids and polyketides		
Hydroxyphenyllactic acid	Tyrosine metabolism	Up
L-Phenylalanine	Phenylalanine metabolism	Up
Elderly[e]		
Benzenoids		
Catechol	Chlorocyclohexane and chlorobenzene degradation	Up

Metabolite Identification [a]	Involved pathway [b]	Trend [c]
m-Cresol	Toluene degradation	Up
Mono(2-ethylhexyl)phthalate [MEHP]	None	Up
o-Cymene	−	Up
Phenol	Tyrosine metabolism	Up
Homogeneous non-metal compounds		
Phosphoric acid	Oxidative phosphorylation	Up
Hydrocarbons		
Decane	−	Up
Dodecane	None	Up
Tridecane	None	Up
Lipids and lipid-like molecules		
2,4-Dihydroxybutanoic acid	−	Up
2-Hydroxyglutaric acid	None	Up
Borneol	None	Down
L-Threonine	Glycine, serine and threonine metabolism	Up
Myristic acid	Fatty acid biosynthesis	Up
Palmitic acid	Fatty acid biosynthesis	Up
Stearic acid	Fatty acid biosynthesis	Up
Organic acids and derivatives		
(S)-3-Hydroxyisobutyric acid	Valine, leucine and isoleucine degradation	Up
2-Ethylhydracrylic acid	−	Up
Alanine	Alanine, aspartate and glutamate metabolism	Up
Glutaric acid	Fatty acid degradation	Up
L-Aspartic acid	Alanine, aspartate and glutamate metabolism	Down
Leucine	Valine, leucine and isoleucine degradation	Up
L-Valine	Valine, leucine and isoleucine degradation	Up
Serine	Glycine, serine and threonine metabolism	Up
Thiodiacetic acid	Metabolism of xenobiotics by cytochrome	Up

38

Metabolite Identification [a]	Involved pathway [b]	Trend [c]
	P450	
Organoheterocyclic compounds		
5-Hydroxyindoleacetic acid	Tryptophan metabolism	Down
Cytosine	Pyrimidine metabolism	Up
D-Xylono-1,5-lactone	Pentose and glucuronate interconversions	Up
Hypoxanthin	Purine metabolism	Down
Picolinic acid	Tryptophan metabolism	Up
Uracil	Pyrimidine metabolism	Up
Organonitrogen compounds		
Diethanolamine	Glycerophospholipid metabolism	Up
Organooxygen compounds		
1,3-butanediol	−	Up
4-Deoxyerythronic acid	−	Up
Acetoin	None	Up
Arabinose	−	Up
Cyclohexanone	Caprolactam degradation	Up
Glyceric acid	Pentose phosphate pathway	Up
Inositol	−	Up
L-Sorbose	None	Up
Threonic acid	Ascorbate and aldarate metabolism	Up
Organosulfur compounds		
Dimethyl sulfone	Sulfur metabolism	Up

[a] Urine metabolites detected by GCxGC-TOFMS, identified using NIST library, and found in and classified via Human Metabolome Database.

[b] Metabolites were searched in KEGG database for involved biological pathways. If metabolite is involved in multiple pathways, only one is shown. None: Found in KEGG database but with no known involved pathways. −: Not found in KEGG database.

[c] The up- or downregulation of metabolites in high exposure group compared to low exposure group.

[d] VIP >1, FDR q <0.1 [e] VIP >1, FDR q <0.15

Pathway enrichment and pathway topology analysis on the 45 exposure-related potential metabolites in children identified three main biological pathways potentially affected by multiple exposures: alanine, aspartate, and glutamate metabolism, phenylalanine metabolism, and tryptophan metabolism. For the 42 potential metabolites found in elderly participants, pathway analysis showed glycine, serine, and threonine metabolism, alanine, aspartate, and glutamate metabolism, as well as aminoacyl-tRNA biosynthesis were the most important pathways disrupted by multiple exposures (Impact value > 0.1, FDR adjusted $p < 0.05$) (see Table 6). Through the Comparative Toxicogenomics Database (CTD), we found that pyrene, fluoranthene, As, Cu, Cd, and Ni were significantly associated with tryptophan metabolism, As, Cd, and Ni with phenylalanine metabolism, dibenzo[a,h]anthracene, As, Cu, Cd, Ni, and Hg with alanine, aspartate, and glutamate metabolism, and As, Cu, Cd, and Ni with glycine, serine, and threonine metabolism (Bonferroni adjusted $p < 0.01$) (Fig. 7) (Davis et al. 2019).

Table 6. Potential biological pathways affected by multiple exposures in children and elderly subjects.

Pathway	Total	Hits	Impact[c]	FDR p[d]
Children [a]				
Alanine, aspartate and glutamate metabolism	24	4	0.370	0.0004
Phenylalanine metabolism	45	5	0.151	0.0003
Tryptophan metabolism	79	2	0.123	0.0347
Elderly [b]				
Alanine, aspartate and glutamate metabolism	24	1	0.265	0.0001
Glycine, serine and threonine metabolism	48	4	0.233	0.0059
Aminoacyl-tRNA biosynthesis	75	5	0.113	0.0174

[a] Pathway analysis results of exposure-related metabolites identified via PLS-DA (q value <0.1)

[b] Pathway analysis results of exposure-related metabolites identified via PLS-DA (q value <0.15)

[c] Pathway impact value calculated from pathway topology analysis. (Impact value >1 are shown)

[d] FDR adjusted p value calculated from pathway enrichment analysis. (FDR p <0.05 are shown)
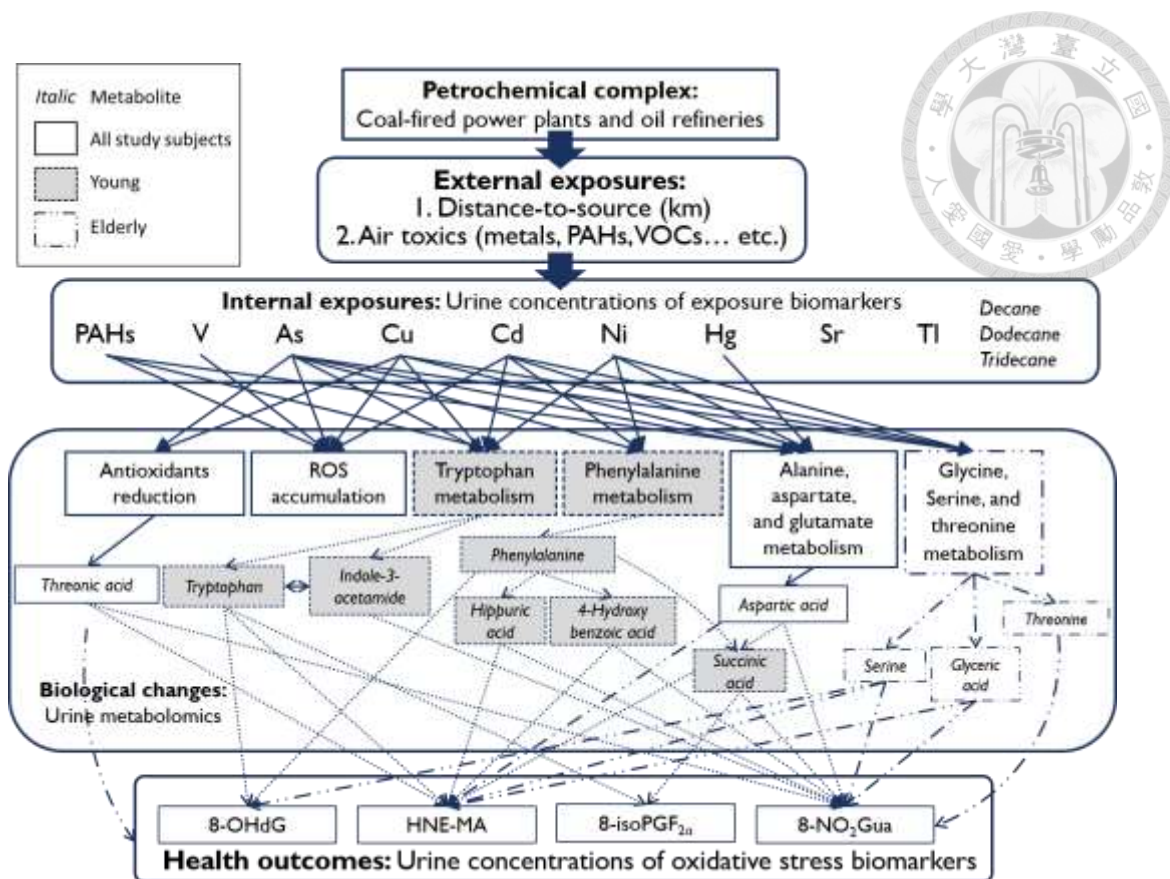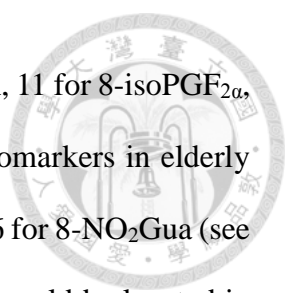
Figure 7. Exposure pathways of petrochemical air pollution and the effects on urine metabolic profile changes and increased oxidative stress

We also identified 163 oxidative stress-related potential metabolites in children subjects, and 144 in elderly subjects that were associated with at least one of the four oxidative stress biomarkers ($p < 0.05$) (Appendix 1). In children, we found 54 potential metabolites associated with 8-OHDG, 108 with HNE-MA, 69 with 8-isoPGF$_{2\alpha}$, and 56 with 8-NO$_2$Gua, while in elderly subjects, 33 were correlated with 8-OHDG, 113 with HNE-MA, 65 with 8-isoPGF$_{2\alpha}$, and 49 with 8-NO$_2$Gua. Identified oxidative stress-related potential metabolites showed more overlap (113 out of 194) than multiple exposure-related potential metabolites (11 out of 76) between children and elderly subjects.

Through the "Meet-in-the-middle" approach, we identified a profile of putative intermediate biomarkers that were associated with both exposures and oxidative stress

41

biomarkers in children participants: 10 for 8-OHDG, 23 for HNE-MA, 11 for 8-isoPGF$_{2\alpha}$, and 22 for 8-NO$_2$Gua. We also identified another profile of such biomarkers in elderly subjects: 17 for 8-OHDG, 32 for HNE-MA, 10 for 8-isoPGF$_{2\alpha}$, and 26 for 8-NO$_2$Gua (see Table 7). Fig. 7 showed that these putative intermediate biomarkers could be located in previously identified exposures-related pathways. In children, tryptophan and indole-3-acetamide could be located in tryptophan metabolism pathway, phenylalanine, hippuric acid, and 4-hydroxy benzoic acid in phenylalanine metabolism pathway, and succinic acid in both phenylalanine metabolism and alanine, aspartate, and glutamate metabolism pathways. In elderly participants, threonine, serine, and glyceric could be located in glycine, serine, and threonine metabolism. Aspartic acid, which was identified in both children and elderly subjects, is involved in alanine, aspartate, and glutamate metabolism.
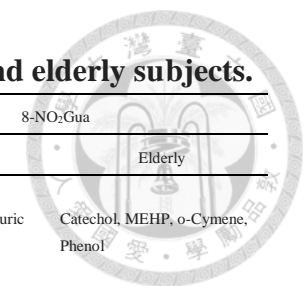
Table 7. Putative intermediate biomarkers associated with both exposure and oxidative stress identified in children and elderly subjects.

Biomarker / Metabolite	8-OHDG Children	8-OHDG Elderly	HNE-MA Children	HNE-MA Elderly	8-isoPGF$_{2\alpha}$ Children	8-isoPGF$_{2\alpha}$ Elderly	8-NO$_2$Gua Children	8-NO$_2$Gua Elderly
Benzenoids		Catechol, MEHP, o-Cymene, Phenol	2-phenylpropanal, 4-Hydroxybenzoic acid, Hippuric acid, o-Cymene	Catechol, o-Cymene, Phenol	3-Hydroxyhippuric acid		2-phenylpropanal, Hippuric acid, o-Cymene	Catechol, MEHP, o-Cymene, Phenol
Hydrocarbons	Tridecane	Dodecane	Decane, Dodecane, Tridecane	Decane, Dodecane, Tridecane		Decane	Decane, Dodecane, Tridecane	Dodecane, Tridecane
Lipids and lipid-like molecules		2-Hydroxyglutaric acid, Myristic acid, Stearic acid	Azelaic acid, Stearic acid	2,4-Dihydroxybutanoic acid, 2-Hydroxyglutaric acid, Myristic acid, Stearic acid	Azelaic acid	Myristic acid, Stearic acid	Myristic acid, Stearic acid, Tiglic acid	2-Hydroxyglutaric acid, L-Threonine, Myristic acid, Stearic acid
Organic acids and derivatives	N-acetylglutamic acid, Tiglylglycine	(S)-3-Hydroxyisobutyric acid, Serine	Aminomalonic acid, N-acetylglutamic acid, Succinic acid, Tiglylglycine	(S)-3-Hydroxyisobutyric acid, 2-Ethylhydracrylic acid, Glutaric acid, L-Aspartic acid, Leucine, L-Valine, Serine, Thiodiacetic acid	L-Alpha-aminobutyric acid, N-acetylglutamic acid, Succinic acid	Leucine	L-Aspartic acid, L-Histidine, Succinic acid	(S)-3-Hydroxyisobutyric acid, Glutaric acid, L-Valine, Serine, Thiodiacetic acid
Organoheterocyclic compounds	2-Deoxy-L-ribono-1,4-lactone, DL-Tryptophan, L-Gulonolactone, Quinolinic acid	Cytosine, D-Xylono-1,5-lactone	DL-Tryptophan	5-Hydroxyindoleacetic acid, Cytosine, D-Xylono-1,5-lactone, Picolinic acid, Uracil	Cyanuric acid, DL-Tryptophan	5-Hydroxyindoleacetic acid, Cytosine, D-Xylono-1,5-lactone	1H-Indole-3-acetamide, Cyanuric acid, Quinolinic acid, Sumiki's acid	D-Xylono-1,5-lactone, Hypoxanthin
Organonitrogen compounds				Diethanolamine			Dimethylamine	
Organooxygen compounds	4-Deoxyerythronic acid	Arabinose, Cyclohexanone, L-Sorbose, Threonic acid	4-Deoxyerythronic acid, Cyclohexanone, Diacetone alcohol, D-Threitol, L-Sorbose, Rhamnose, Threonic acid	4-Deoxyerythronic acid, Arabinose, Cyclohexanone, Glyceric acid, Inositol, L-Sorbose, Threonic acid	4-Deoxyerythronic acid, Diacetone alcohol, D-Threitol, Rhamnose	Arabinose, Threonic acid	4-Deoxyerythronic acid, Cyclohexanone, Diacetone alcohol, Gluconic acid, Threonic acid	1,3-butanediol, 4-Deoxyerythronic acid, Arabinose, Cyclohexanone, Glyceric acid, Inositol, L-Sorbose, Threonic acid
Phenylpropanoids and polyketides	Hydroxyphenyllactic acid, L-Phenylalanine		Hydroxyphenyllactic acid, L-Phenylalanine					
Organosulfur compounds		Dimethyl sulfone		Dimethyl sulfone		Dimethyl sulfone		Dimethyl sulfone
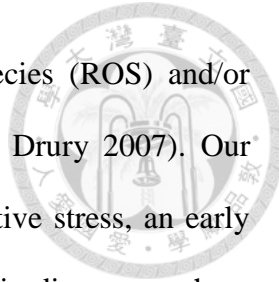
4.1.2 Discussion

Here we demonstrate a urine metabolomic approach to evaluate how a mixture of air toxics from an industrial emission source can increase the external and internal exposures of nearby residents in a distance-to-source-related manner, inducing age-dependent responses in children and elderly residents that led to elevated oxidative stress. The strength of this study was that we comprehensively evaluated multiple air toxics, and successfully used untargeted metabolomics to find the association between exposures and early health effect oxidative stress through the "Meet-in-the-middle" approach.

External and internal exposure results suggest petrochemical complex is the major source of both external and internal exposures to multiple air toxics for our study subjects. High exposure groups lived closer to the emission sources, had higher levels of ambient V and PAHs at the location of their home addresses, with increased urinary concentrations of heavy metals and PAHs metabolite 1-OHP, of which PAHs, V, Ni, As, and Cu exposures have been reported near both coal-fired power plants and oil refineries, and Cd, Hg, Sr, and Tl near coal-fired power plants (Dybing et al. 2013; George et al. 2015; IARC 1989a, b; O'Rourke and Connolly 2003; Peter and Viraraghavan 2005; Williams et al. 2006). Lack of difference in surrounding road area at home addresses of high and low exposure groups suggest limited traffic influence on exposures.
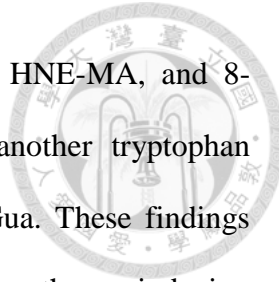
Urine oxidative stress biomarkers were elevated in high exposure groups, including 8-OHDG, a biomarker for oxidative DNA damage, HNE-MA, produced by lipid peroxidation, 8-isoPGF$_{2\alpha}$, a metabolite from arachidonic acid peroxidation, and 8-NO$_2$Gua, potential biomarker for nitrative DNA damage (Wu et al. 2016). Previous studies have reported exposure to PAHs and heavy metals V, Ni, Cu, As, Cd, and Hg can

44
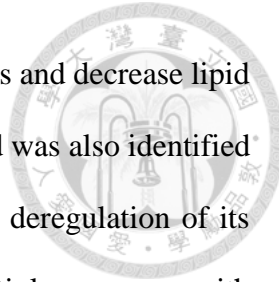
induce oxidative stress through increment of reactive oxygen species (ROS) and/or reduction of anti-oxidants (Jomova and Valko 2011; Penning and Drury 2007). Our findings suggest that exposure to multiple air toxics induces oxidative stress, an early health effect that contributes to numerous common complex chronic diseases such as cancer, cardiovascular disease, diabetes, and neurodegenerative diseases, as well as acute respiratory diseases such as allergic rhinitis and asthma (Bowler and Crapo 2002; Reuter et al. 2010).

Potential metabolites responsible for the separation between high and low exposure groups were different for children and elderly subjects, and pathway analysis results showed three different and one common biological pathway affected by exposures in the two age groups. "Meet-in-the-middle" approach found putative intermediate biomarkers that were involved in these four pathways, suggesting that exposures can disrupt diverse biological mechanisms in different age groups, inducing common early health effect oxidative stress.

Tryptophan metabolism was one of the exposures-related pathways identified in children subjects. Tryptophan can be metabolized through different pathways, the most important route, kynurenine pathway, is often upregulated by activated immune responses, leading to the depletion of tryptophan, and has shown involvement in increased oxidative stress as well as numerous diseases including cancer, neurodegenerative diseases, and allergic disorders such as rhinitis and asthma (Chen and Guillemin 2009; Ciprandi et al. 2010; Gostner et al. 2016; Stoy et al. 2005). Our previous study in the same area also showed increased incidence of allergic rhinitis, bronchitis, and asthma in children living near the petrochemical complex (Chiang et al. 2016). Tryptophan was downregulated in high exposure group compared to low exposure group in children subjects, identified as

a putative intermediate biomarker and correlated with 8-OHDG, HNE-MA, and 8-isoPGF$_{2\alpha}$. Downstream metabolite 1H-indole-3-acetamide from another tryptophan metabolism route was also identified and associated with 8-NO$_2$Gua. These findings suggest exposures could affect at least one tryptophan metabolism pathway, inducing multiple oxidative stress outcomes. Given the mediating roles of tryptophan metabolism and oxidative stress in allergic respiratory diseases (Bowler and Crapo 2002; Ciprandi et al. 2010; Gostner et al. 2016), our findings may provide further information on the biological mechanisms deregulated by petrochemical industry exposures that led to increased risks in children. Phenylalanine metabolism was also identified in children subjects. Phenylalanine, which have been used as an oxidative stress biomarker (Orhan et al. 2004), was significantly correlated with exposures and oxidative stress biomarkers 8-OHDG and HNE-MA, and with HNE-MA, 8-isoPGF$_{2\alpha}$, and 8-NO$_2$Gua through its downstream metabolites hippuric acid, 4-hydroxy benzoic acid, and succinic acid in children participants. Since the increase of 8-OHDG levels in high exposure group for children participants increased but did not reach statistical significance, our findings suggest that urinary tryptophan and phenylalanine could be a potential intermediate biomarkers of exposure-induced oxidative stress in children study subjects, before DNA damage became significant. In elderly subjects, glycine, serine, and threonine metabolism was identified, with related-compounds threonine associating with 8-NO$_2$Gua, serine with 8-OHDG, HNE-MA, and 8-NO$_2$Gua, and glyceric acid with HNE-MA and 8-NO$_2$Gua. The disruption of this pathway is closely related to oncogenic transformation, and the biosynthesis of antioxidant glutathione (Amelio et al. 2014). Alanine, aspartate, and glutamate metabolism was identified in both children and elderly participants, with aspartic acid downregulated in high exposure subjects of both age groups, associated with 8-NO$_2$Gua in children participants, and with HNE-MA in elderly participants. Previous

46

studies have shown that aspartic acid could increase glutathione levels and decrease lipid peroxidation in animal models (Sivakumar et al. 2011). Threonic acid was also identified in both age groups as a putative intermediate biomarker, indicating deregulation of its precursor, antioxidant ascorbic acid (Gao et al. 2012), associating multiple exposures with HNE-MA and 8-NO$_2$Gua in children and all four oxidative stress biomarkers in elderly participants. We can draw from these results a complicated web showing the relation between exposures and different oxidative stress induced health effects through age-dependent diverse biological pathways.

Accidentally but not unexpectedly, we also found some exposure biomarkers in our untargeted urine analysis of metabolomes. Decane, dodecane, and tridecane were elevated in high exposure groups for both children and elderly subjects. These compounds are intermediates in petrochemical industrial productions, and were previously reported as potential health risks (IARC 1989b; Williams et al. 2006). We did not locate decane, dodecane, or tridecane in human biological pathway analysis, suggesting these compounds are from external sources, and can be used as exposure biomarkers of petrochemical emissions.

There are limitations to the present study. First, biomarkers measured in one spot urine may have daily variability, and may not be as stable in reflecting exposure to heavy metals as those measured in other biospecimens such as hair, which unfortunately was not collected in this study. Nevertheless, we believe this would not systemically bias our exposure classification because distance-to-source and model-based ambient concentrations all showed the same pattern as urine biomarkers. In addition, the petrochemical complex has had continuous emissions since operation started in 1999, and our study subjects have lived in the area for at least five years. Secondly, the air toxics in

this study was limited to PAHs and heavy metals for which we performed biomonitoring. SO₂, NOx, black carbon, and other toxics whose biomarkers were not measured in this study may also contribute to the increase of oxidative stress. Thirdly, we could not rule out routes of exposure other than air that could induce oxidative stress, such as water and food. Lastly, we could not rule out the possibility of inaccurate metabolite identification by library match.

Urine exposure biomarkers of PAHs and heavy metals V, Ni, As, Cu, Sr, Cd, Hg, and Tl were elevated in children and elderly residents living near a petrochemical complex. These internal exposures were associated with model-estimated ambient concentrations at residential addresses, and could possibly be traced to air toxics emitted by oil refineries and coal-fired power plants within the complex. Both children and elderly residents living in the pollution-affected area with higher levels of urine exposure biomarkers showed changes in urine metabolite profiles which could be linked to increased oxidative stress, including oxidative and nitrative DNA damage, and lipid peroxidation. We conclude that urine metabolomics could possibly serve as the link to trace multiple air toxics exposure to oxidative stress through age-specific biological pathways including tryptophan metabolism and phenylalanine metabolism in children subjects, glycine, serine, and threonine metabolism in elderly subjects, and alanine, aspartate, and glutamate metabolism in both age groups. The identified exposures and metabolic pathways will improve risk assessments on developing common complex chronic diseases, such as cancers and cardiovascular diseases, as well as allergic respiratory diseases, such as allergic rhinitis and asthma, for the residents living near the petrochemical complex if air toxics exposure continues in the future. We recommend to significantly reduce air toxics emissions from the petrochemical complex to lower residents' health risks. Our findings

also warrant a follow-up study on residents who continue to be affected by petrochemical pollution.

4.2 Part 2. Metabolomics of Children and Adolescents Exposed to Industrial Carcinogenic Pollutants
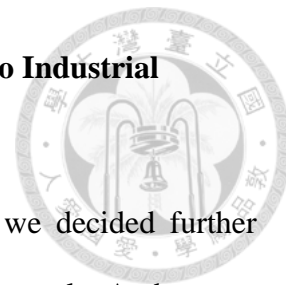
Based on the age-dependent findings from our Part 1 study, we decided further metabolomics studies for the two age groups should be conducted separately. At the same time, Yuan et al. and Chen et al. reported increased risk of all cancers for residents living near No. 6 Naphtha Cracking Complex (Chen et al. 2018; Yuan et al. 2018). Since children exposed to carcinogenic pollutants at critical periods of development have more time for chronic adverse health effects such as cancer to manifest, we decided to first focus on our children participants in order to identify the association between industrial carcinogenic pollutants exposure, serum metabolic changes, and cancer-related early health effects including oxidative stress and serum acylcarnitines. We used serum samples in our Part 2 study because blood circulates the body covering every tissue and organ, carrying all the molecules that are secreted, excreted, or discarded in response to physiological needs and stress, and alterations in blood metabolite profile could reflect pathological states and the body's attempt to maintain homeostasis (Psychogios et al. 2011).

4.2.1   Results

Table 8 showed the comparison of basic characteristics, external and internal exposure levels, and urine oxidative stress biomarker levels between high and low exposure groups of the 107 children participants with serum samples available for metabolomics analysis. High exposure group lived 10.57 ± 2.52 km and 10.02 ± 2.73 km away from the main emission points of coal-fired power plant and oil refineries, respectively, while low exposure group lived 21.81 ± 5.71 km and 20.91 ± 5.44 km away,

respectively. High and low exposure group showed no significant difference in age, sex distribution, systolic blood pressure (SBP), smoking history, alcohol history, and betelnut history. However, high exposure group had higher BMI compared to low exposure group. Road area surrounding homes showed no significant difference between the two exposure groups at neither 25 m or 500 m buffer. Urine concentrations of exposure biomarkers As, Cd, Cr, Ni, 1-OHP, V, and Hg were increased in high exposure group compared to low exposure group, with As, Cd, 1-OHP, V, and Hg reaching statistical significance. The difference was most profound in 1-OHP and V. Urine concentrations of oxidative stress biomarkers showed all four biomarkers were increased in high exposure group compared to low exposure group, but the differences were more statistically significant for lipid peroxidation biomarkers HNE-MA ($p=0.006$) and 8-isoPF$_{2\alpha}$ ($p=0.09$) than DNA damage biomarkers 8-OHDG ($p=0.21$) and 8-NO$_2$Gua ($p=0.11$). Only 99 out of the 107 subjects had available data for urine oxidative stress biomarkers concentrations.

51

Table 8. Comparison of basic characteristics, carcinogens exposure levels, and oxidative stress biomarker levels in 107 study subjects.

	High (n=37)	Low (n=70)	p^a
Basic characteristics			
Age, mean±SD	13.67±0.92	13.70±0.90	0.84
Male, n(%)	20(54.1)	38(54.3)	0.98
Systolic blood pressure (SBP), mean±SD	118.7±12.96	116.2±13.73	0.37
Body Mass Index (BMI), mean±SD	21.67±3.41	20.15±3.48	0.04
Smoke history, n(%)	5(13.5)	5(7.1)	0.28
Drink history, n(%)	5(13.5)	3(4.3)	0.12
Betelnut history, n(%)	1(2.7)	3(4.4)	1.00
External exposures[b], mean±SD			
Distance to coal-fired power plant	10.57±2.52	21.81±5.71	<0.0001
Distance to oil refinery	10.02±2.73	20.91±5.44	<0.0001
Road area surrounding homes			
25 m buffer	304.1±211.4	329.4±204.7	0.58
500 m buffer	70938.4±26594.8	64120.1±20016.4	0.18
Internal exposures[c], mean±SD			
Group 1 carcinogen			
Arsenic	60.27±42.16	39.62±30.18	0.01
Cadmium	0.34±0.34	0.19±0.15	0.02
Chromium	3.24±2.96	2.14±1.63	0.10
Nickel	6.69±8.72	3.89±2.85	0.31
Group 2A carcinogen			
Lead	0.64±0.64	0.66±0.65	0.80
1-OHP	0.19±0.14	0.03±0.01	<0.0001
Group 2B carcinogen			
Vanadium	2.46±1.64	0.24±0.10	<0.0001
Mercury	3.13±2.89	1.86±1.88	0.04
Oxidative stress[d], mean±SD			
8-OHDG	3.01±2.15	2.63±2.86	0.21
HNE-MA	2.00±2.58	1.30±2.20	0.006
8-isoPF$_{2\alpha}$	3.27±3.54	2.06±2.18	0.09
8-NO$_2$Gua	6.49±11.77	2.54±3.05	0.11

[a] Comparison of basic characteristics between the high and low exposure groups for continuous variables was performed using Student's t-test, and for discrete variables, Chi-squared test or Fisher's exact test. Urinary exposure biomarker concentrations are log-transformed, high and low exposure groups compared by ANCOVA test adjusting

52

age, sex, smoking, alcohol consumption, betel nut chewing, fish consumption, and BMI with a post comparison by Scheffe test. Urinary oxidative stress biomarker concentrations are log-transformed, high and low exposure groups compared by Student's t-test.

[b] Distance to source: Average of home-to-coal-fired power plant and home-to-oil refinery distance, unit: km; Road area surrounding homes unit: m^2.

[c] For urine 1-OHP, unit: μmol/mol-creatinine; for urine heavy metals, unit: μg/g-creatinine

[d] Urine oxidative stress biomarkers unit: μg/g-creatinine. High exposure group N=34, low exposure group N=65.

Figure 8 showed serum levels of six acylcarnitines that were significantly different in high exposure group compared to low exposure group in 107 study subjects. Samples are in columns and arranged according to high exposure (red) and low exposure (green) groups. Acylcarnitines are in rows and were arranged according to hierarchical clustering using Euclidean distance measure and Ward algorithm. The colors vary from deep blue to dark brown to indicate data values change from down-regulation (blue) to up-regulation (brown). We found long-chain acylcarnitines were clustered together and down-regulated in high exposure group compared to low exposure group (Dodecanoylcarnitine, C12; Tetradecanoylcarnitine, C14; Tetradecenoylcarnitine, C14:1; Hexadecenoylcarnitine, C16:1; Pristanoylcarnitine, C19), while short-chain acylcarnitine (Hexanoylcarnitine, C6) was up-regulated in high exposure group compared to low exposure group.

Figure 8. Heat map of serum acylcarnitine levels in 107 study subjects

PCA showed clear clustering of pooled QC on the score plots shown in Figure 9. Metabolomics identified 84 potential metabolite features in study subjects serum samples after removing features missing in more than 50% of the samples, 80 of which had available HMDB ID number as shown in Appendix 2. 84 potential metabolite features were put through PLS-DA analysis, and results showed metabolic profiles between high and low exposure groups could be significantly separated by two components that accounted for 5.8 % and 9.0 % of variability of metabolic profiles between high and low exposure groups, respectively (Accuracy=0.78, R^2=0.53, Q^2=0.23) (Figure 10A). Permutation test was performed to confirm the validity of PLS-DA model (p=0.01). PLS-DA and ANCOVA analysis adjusting for age, sex, and BMI found 11 exposure-related potential metabolite features (average variable importance in projection (VIP) score >1, ANCOVA p <0.05), which through in house library search was identified as 10 potential metabolites. Two potential metabolites, one detected under positive mode and one under negative mode of UHPLC-qTOFMS analysis, were both identified as pyroglutamic acid. Figure 10B showed the up- and down-regulation of exposure-related potential metabolites in high and low exposure groups. Samples are in columns arranged according to high exposure (red) and low exposure (green) groups. Potential metabolites are in rows and were arranged according to hierarchical clustering using Euclidean distance measure and

54

Ward algorithm. The colors vary from deep blue to dark brown to indicate data values change from down-regulation (blue) to up-regulation (brown). We found potential metabolites up-regulated in high exposure group compared to low exposure group were clustered together, including ketoleucine, carnitine, isovalerylcarnitine, aspartic acid, and octenoyl-L-carnitine, while down-regulated potential metabolites were also clustered together, including pyroglutamic acid, adenosine monophosphate (AMP), inosinic acid (inosine monophosphate, IMP), oxoglutaric acid, and malic acid (Figure 10B). Pathway analysis results showed purine metabolism was the main biological pathway affected by multiple exposures ($p < 0.05$, Impact > 0.1) (data not shown). We identified two exposure-related potential metabolites involved in purine metabolism, nucleotides AMP and IMP (Figure 10B) (Simoni et al. 2007). Through the Comparative Toxicogenomics Database (CTD), we found that group 1 carcinogens As, Cd, Cr, and Ni were significantly associated with purine metabolism pathway (Bonferroni adjusted $p < 0.01$).
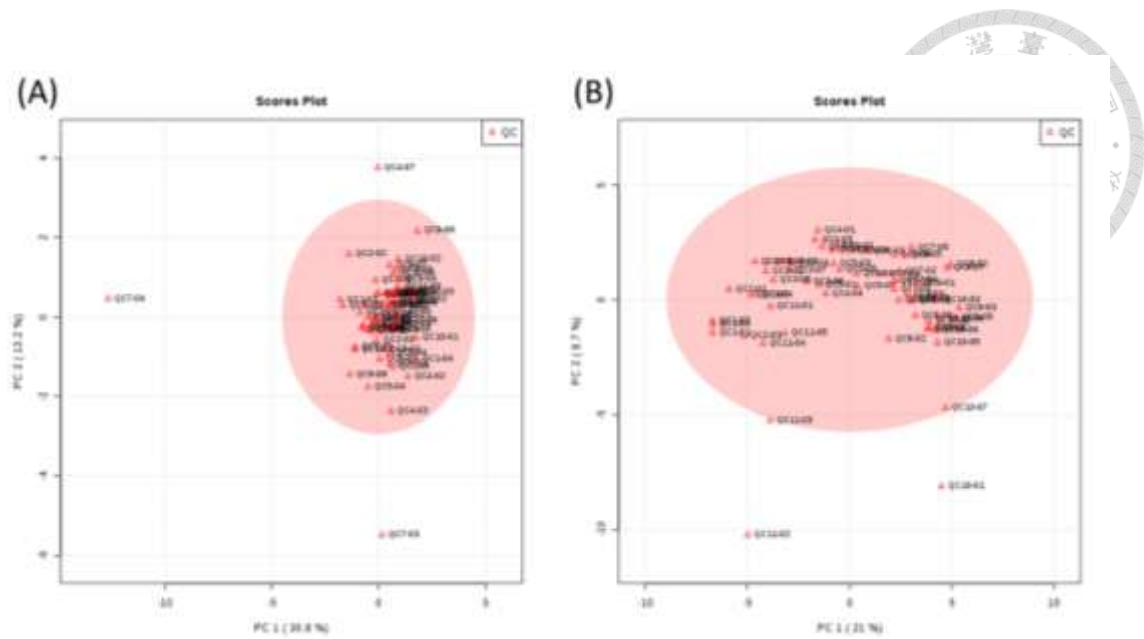
Figure 9. Principle component analysis score plot of 61 pooled quality control (QC) samples data from 11 batches detected under (A) negative mode and (B) positive mode of UHPLC-qTOFMS metabolomics analysis

Pyroglutamic acid was detected in both [1] negative and [2] positive mode.
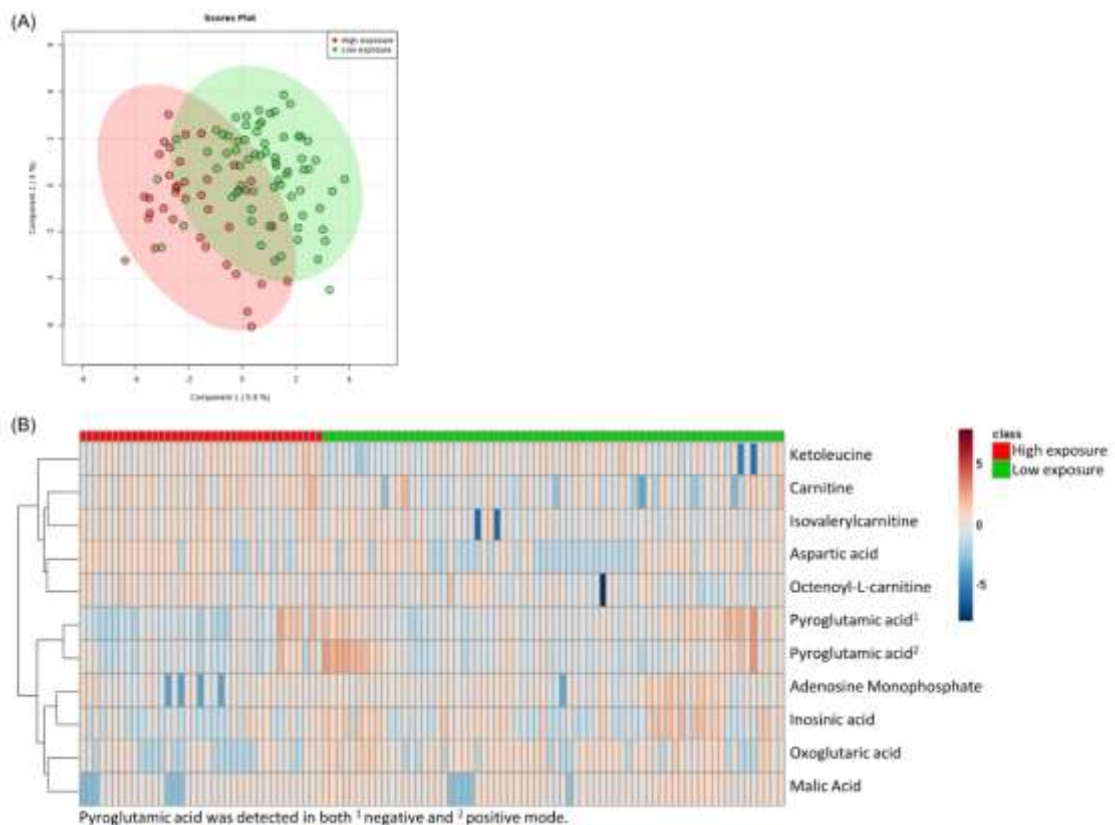
Figure 10. Comparison of serum metabolic profile in 107 study subjects using (A) PLS-DA score plot (Accuracy=0.78, R^2=0.53, Q^2=0.23, Permutation p=0.01) and (B) heat map of exposure-related potential metabolite levels (average VIP score >1, ANOVA p <0.05 are shown adjusted for sex, age, and BMI)

56
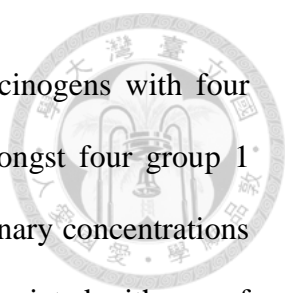
Table 9 showed the association between eight individual carcinogens with four oxidative stress biomarkers in 99 study subjects. Individually, amongst four group 1 carcinogens, urinary levels of Cd was positively associated with urinary concentrations of 8-NO$_2$Gua (p=0.029). The two group 2A carcinogens was not associated with any of the four oxidative stress biomarkers. For the two group 2B carcinogens, V was associated with HNE-MA (p=0.003), 8-isoPF$_{2\alpha}$ (p=0.023), and 8-NO$_2$Gua (p=0.004), and Hg was associated with 8-OHDG (p=0.008) and HNE-MA (p=0.012). Figure 11A to 11D showed the WQS regression analysis of the association of combined eight carcinogens exposure with four oxidative stress biomarkers, respectively. Association with all four oxidative stress biomarkers were positive and statistically significant with 8-OHDG, HNE-MA, and 8-NO$_2$Gua, while association with 8-isoPF$_{2\alpha}$ was borderline significant. For 8-OHDG, group 2B carcinogen Hg predominated in the mixture index (49.7%) and group 1 carcinogens Ni, As, and Cd also contributed to the association (p=0.002) (Figure 11A). Figure 11B showed group 2B carcinogens Hg (43.3%) and V (31.1%) contributed to over half of the mixture index positively associated with HNE-MA levels (p=0.0006), and group 1 carcinogens As, Cd, Ni, and Cr also showed contribution. Associations with 8-isoPF$_{2\alpha}$ was predominated by group 2B carcinogens Hg (36.9%) and V (31.6%), followed by group 1 carcinogen Ni (19.4%), with contribution from Cr and As (p=0.08) (Figure 11C). In Figure 11D, we can see in the mixture index positively associated with 8-NO$_2$Gua (p=0.0001), group 2B carcinogen V contributed to half of the association (50.1%), followed by group 1 carcinogen Ni (32.0%), with contributions from Cd, Cr, and As.
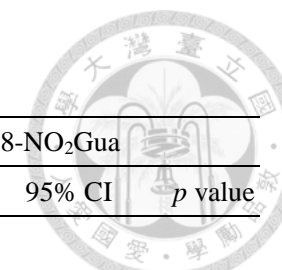
Table 9. Individual association between urine carcinogens and oxidative stress biomarkers in 99 study subjects.

	8-OHDG			HNE-MA			8-isoPF$_{2\alpha}$			8-NO$_2$Gua		
	Estimate	95% CI	*p* value	Estimate	95% CI	*p* value	Estimate	95% CI	*p* value	Estimate	95% CI	*p* value
Group 1 carcinogen												
Arsenic	0.049	(-0.165, 0.263)	0.653	0.092	(-0.163, 0.347)	0.476	-0.096	(-0.321, 0.130)	0.402	0.016	(-0.324, 0.356)	0.925
Cadmium	-0.025	(-0.252, 0.202)	0.831	0.097	(-0.173, 0.366)	0.480	-0.141	(-0.379, 0.097)	0.243	0.391	(0.040, 0.743)	0.029
Chromium	-0.186	(-0.380, 0.007)	0.059	-0.016	(-0.251, 0.219)	0.892	-0.061	(-0.269, 0.147)	0.562	0.120	(-0.192, 0.432)	0.448
Nickel	0.018	(-0.172, 0.209)	0.848	0.033	(-0.194, 0.260)	0.775	0.126	(-0.073, 0.326)	0.213	0.224	(-0.075, 0.523)	0.140
Group 2A carcinogen												
Lead	0.115	(-0.112, 0.341)	0.318	0.158	(-0.112, 0.429)	0.249	-0.135	(-0.374, 0.104)	0.265	-0.033	(-0.397, 0.331)	0.858
1-OHP	0.094	(-0.057, 0.244)	0.219	0.146	(-0.032, 0.325)	0.107	0.073	(-0.087, 0.232)	0.369	0.188	(-0.050, 0.425)	0.121
Group 2B carcinogen												
Vanadium	0.132	(-0.013, 0.276)	0.073	0.254	(0.087, 0.422)	0.003	0.176	(0.025, 0.327)	0.023	0.330	(0.107, 0.554)	0.004
Mercury	0.246	(0.064, 0.427)	0.008	0.279	(0.062, 0.496)	0.012	0.119	(-0.078, 0.316)	0.233	0.005	(-0.293, 0.304)	0.973

Linear regression analysis adjusted for sex, age, and BMI
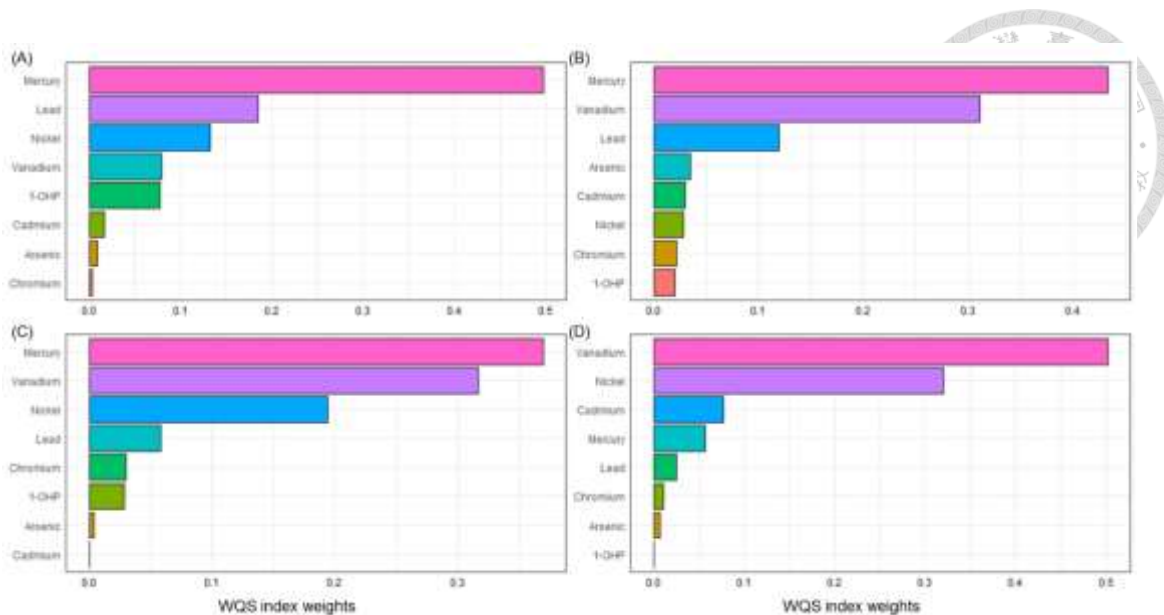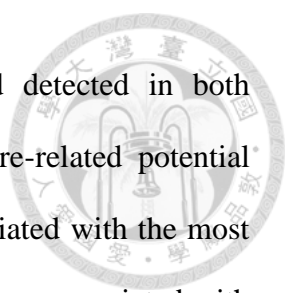
Figure 11. Combined associations between internal exposure levels and (A) 8-OHDG (*p*=0.002), (B) HNE-MA (*p*= 0.0006), (C) 8-isoPF$_{2\alpha}$ (*p*= 0.08), and (D) 8-NO$_2$Gua (*p*=0.0001) levels based on weighted quantile sum (WQS) regression analysis in 99 study subjects. (Adjusted for sex, age, and BMI)

"Meet-in-the middle" approach identified eight potential metabolites that were both carcinogens exposure-related and associated with biomarkers of early health effects. Table 10 and 11 showed the level of association between carcinogens exposure-related potential metabolites (in rows) and biomarkers of early health effects (in columns). For oxidative stress biomarkers, 8-OHDG was significantly associated with pyroglutamic acid and inosinic acid, HNE-MA was significantly associated with ketoleucine, octenoyl-L-carnitine, pyroglutamic acid, AMP, and IMP, 8-isoPF$_{2\alpha}$ was significantly associated with octenoyl-L-carnitine, and 8-NO$_2$Gua was not significantly associated with any exposure-related potential metabolites (Table 10). Long-chain acylcarnitines C14 and C19 were associated with the most number of exposure-related potential metabolites, five for C14 (carnitine, octenoyl-L-carnitine, pyroglutamic acid detected in both positive and negative modes, IMP) and five for C19 (carnitine, octenoyl-L-carnitine, pyroglutamic acid, AMP, IMP). C16:1 was associated with carnitine and octenoyl-L-carnitine. Short-chain acylcarnitine C6 was associated with four exposure-related potential metabolites

59

including ketoleucine, isovalerylcarnitine, and pyroglutamic acid detected in both positive and negative modes (Table 11). Overall, for the exposure-related potential metabolites, octenoyl-L-carnitine and pyroglutamic acid were associated with the most number of biomarkers of early health effects. Octenoyl-L-carnitine was associated with two oxidative stress biomarkers and three long-chain acylcarnitines, and pyroglutamic acid was associated with two oxidative stress biomarkers, short-chain acylcarnitine, and two long-chain acylcarnitines. Aspartic acid, oxoglutaric acid, and malic acid were not significantly associated with any of the biomarkers of early health effects.

60

Table 10. Association between exposure-related potential metabolites and oxidative stress biomarkers in 99 study subjects

(Estimates of linear regression analysis are shown. 95% CI are in brackets. *$p<0.05$ †$p<0.01$ ‡$p<0.001$).

	8-OHDG	HNE-MA	8-isoPF$_{2\alpha}$	8-NO$_2$Gua
Ketoleucine	0.03	0.10*	0.05	0.09
	(-0.05, 0.11)	(2.19E-03, 0.19)	(-0.04, 0.13)	(-0.03, 0.22)
Carnitine	2.79E-03	0.05	0.01	-0.02
	(-0.08, 0.09)	(-0.05, 0.14)	(-0.08, 0.10)	(-0.15, 0.11)
Isovalerylcarnitine	-0.05	-0.01	0.03	0.03
	(-0.13, 0.03)	(-0.11, 0.09)	(-0.05, 0.12)	(-0.10, 0.16)
Aspartic acid	-0.03	-0.03	2.84E-04	0.11
	(-0.11, 0.04)	(-0.12, 0.07)	(-0.08, 0.08)	(-0.01, 0.24)
Octenoyl-L-carnitine	0.05	0.09*	0.08*	0.11
	(-0.03, 0.13)	(2.00E-03, 0.18)	(4.28E-03, 0.17)	(-0.01, 0.23)
Pyroglutamic acid[a]	-0.04	-0.09	0.02	-0.07
	(-0.12, 0.04)	(-0.19, 1.89E-03)	(-0.07, 0.10)	(-0.20, 0.05)
Pyroglutamic acid[b]	-0.08*	-0.09*	0.01	-0.04
	(-0.16, -3.36E-03)	(-0.19, -3.48E-04)	(-0.07, 0.09)	(-0.16, 0.09)
Adenosine Monophosphate	-0.05	-0.11*	0.02	-0.11
	(-0.13, 0.03)	(-0.20, -0.02)	(-0.07, 0.10)	(-0.24, 0.01)
Inosinic acid	-0.08*	-0.13†	-0.04	-0.08
	(-0.16, -9.14E-04)	(-0.23, -0.04)	(-0.12, 0.05)	(-0.21, 0.05)
Oxoglutaric acid	-0.04	-0.01	1.48E-03	-0.09
	(-0.12, 0.04)	(-0.11, 0.08)	(-0.08, 0.09)	(-0.22, 0.03)
Malic Acid	-0.01	-0.03	0.05	0.01
	(-0.09, 0.07)	(-0.12, 0.07)	(-0.03, 0.14)	(-0.11, 0.14)

Linear regression model adjusted for age, sex, and BMI.

Pyroglutamic acid was detected in both [a] negative and [b] positive mode.

Table 11. Association between exposure-related potential metabolites and acylcarnitines in 107 study subjects

(Estimates of linear regression analysis are shown. 95% CI are in brackets. *$p<0.05$ †$p<0.01$ ‡$p<0.001$).

	C6[b]	C14:1[b]	C12[b]	C19[b]	C14[b]	C16:1[b]
Ketoleucine	0.31 (0.12, 0.49)†	-0.01 (-0.20, 0.17)	0.01 (-0.18, 0.19)	-0.04 (-0.23, 0.15)	-0.04 (-0.23, 0.15)	4.66E-03 (-0.19, 0.20)
Carnitine	0.11 (-0.09, 0.31)	-0.13 (-0.32, 0.05)	-0.16 (-0.34, 0.02)	-0.19 (-0.38, -3.39E-03)*	-0.25 (-0.43, -0.06)*	-0.23 (-0.42, -0.04)*
Isovalerylcarnitine	0.20 (1.39E-03, 0.40)*	0.08 (-0.11, 0.28)	0.02 (-0.17, 0.21)	-0.09 (-0.29, 0.10)	-0.14 (-0.34, 0.05)	-0.09 (-0.29, 0.11)
Aspartic acid	0.04 (-0.15, 0.24)	-0.09 (-0.27, 0.09)	-0.14 (-0.31, 0.04)	-0.02 (-0.21, 0.16)	-0.10 (-0.28, 0.11)	-0.08 (-0.27, 0.11)
Octenoyl-L-carnitine	0.13 (-0.06, 0.32)	-0.12 (-0.30, 0.06)	-0.12 (-0.30, 0.05)	-0.32 (-0.49, -0.14)‡	-0.21 (-0.39, -0.03)*	-0.19 (-0.38, -0.01)*
Pyroglutamic acid[a]	-0.39 (-0.57, -0.32)‡	0.17 (-0.01, 0.36)	0.13 (-0.05, 0.32)	0.27 (0.09, 0.46)†	0.21 (0.02, 0.39)*	0.19 (-0.01, 0.38)
Pyroglutamic acid[b]	-0.32 (-0.50, -0.13)†	0.13 (-0.05, 0.31)	0.07 (-0.11, 0.25)	0.17 (-0.01, 0.35)	0.20 (0.02, 0.39)*	0.14 (-0.05, 0.33)
Adenosine Monophosphate	0.03 (-0.17, 0.22)	0.08 (-0.10, 0.27)	0.15 (-0.02, 0.33)	0.34 (0.17, 0.52)‡	0.18 (-2.75E-03, 0.36)	0.09 (-0.10, 0.28)
Inosinic acid	-0.06 (-0.25, 0.13)	0.01 (-0.17, 0.19)	0.05 (-0.13, 0.23)	0.48 (0.31, 0.64)‡	0.20 (0.02, 0.38)*	0.05 (-0.15, 0.24)
Oxoglutaric acid	0.05 (-0.14, 0.25)	0.13 (-0.05, 0.32)	0.08 (-0.10, 0.26)	0.13 (-0.06, 0.31)	0.09 (-0.09, 0.28)	0.12 (-0.07, 0.31)
Malic Acid	0.06 (-0.14, 0.25)	0.14 (-0.04, 0.32)	0.14 (-0.04, 0.32)	0.17 (-0.01, 0.36)	0.15 (-0.03, 0.34)	0.08 (-0.11, 0.27)

Linear regression model adjusted for age, sex, and BMI.

Pyroglutamic acid was detected in both [a] negative and [b] positive mode.

62
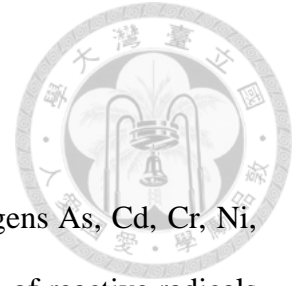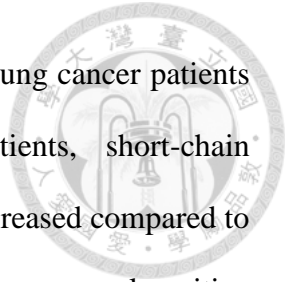
4.2.2 Discussion

Previous studies have reported exposure to individual carcinogens As, Cd, Cr, Ni, Pb, PAHs, V, and Hg can induce oxidative stress through production of reactive radicals and/or depletion of anti-oxidants (Fu et al. 2012; Jomova and Valko 2011; Valko et al. 2005). However, these studies mostly focused on the association between single carcinogen exposure and oxidative stress, and only occupational exposure studies in adults reported the association between multiple heavy metals exposure and oxidative stress (Ko et al. 2017). Our subjects were exposed to multiple carcinogens and therefore it was difficult to find one-to-one association between specific carcinogens and oxidative stress. The level of which each carcinogen induced oxidative stress may also vary, especially in a mixture. The strength of our study is that we applied WQS regression analysis and showed in children and adolescents, exposure to a mixture of eight environmental carcinogens was positively associated with four oxidative stress biomarkers, and both group 1 and group 2 carcinogens contributed to this association.

Our study is the first to report multiple carcinogens exposure could be associated with alterations in serum acylcarnitine levels in children and adolescents. Our findings support a previous study of adults occupationally exposed to metal-containing welding fumes who had significant decrease in short- and long-chain acylcarnitines (Shen et al. 2018). Previous studies have suggested acylcarnitines to be suitable candidates for cancer diagnosis (Ni et al. 2016). Interestingly, the up- and down-regulation of short- and long-chain acylcarnitines vary by cancer types and studies. Ni et al. reported both short-chain and long-chain acylcarnitines were significantly increased in lung cancer patients compared to healthy control subjects (Ni et al. 2016). Another study showed significant

63

decrease of short-chain acylcarnitines in early stage non-small-cell lung cancer patients (Klupczynska et al. 2017). In hepatocellular carcinoma patients, short-chain acylcarnitines were decreased and long-chain acylcarnitines were increased compared to control subjects (Chen et al. 2013; Zhou et al. 2012b). We identified serum acylcarnitine deregulations in children and adolescents exposed to multiple carcinogens that has been reported in cancer patients, which may imply the possibility of increased cancer risk.

Deregulation in purine metabolism has been associated with early stage cancer development and cancer progression (Bester et al. 2011). Purine metabolism is involved in energy production and signal transduction, and the enzymes and metabolites from this pathway can mediate oxidative stress through reactive species and anti-oxidant productions (Cantu-Medellin and Kelley 2013; Maiuolo et al. 2016; Pedley and Benkovic 2017). Our findings suggest multiple carcinogens exposure can induce perturbations in purine metabolism and link to increased oxidative stress and altered serum acylcarnitine levels.

Multiple carcinogens exposure were also associated with several potential metabolites in this study which could not be summarized in pathway analysis, but are involved in important biological mechanisms and have been reported in cancer studies. These potential metabolites included aspartic acid, an amino acid that has been reported to be involved in oxidative stress regulations (Sivakumar et al. 2011). Carnitine and citrate cycle-related metabolites malic acid and oxoglutaric acid were also identified, and carnitine was associated with acylcarnitines. Carnitine cooperates with acylcarnitines transporting fatty acids into mitochondria for β-oxidation, forming acetyl-CoA that enters the citrate cycle (Semba et al. 2017). These findings suggest multiple carcinogens exposure in children and adolescents may affect fatty acid oxidation and energy

64

production mechanisms leading to deregulation of acylcarnitines. Multiple carcinogens exposure in children and adolescents also affected pyroglutamic acid, an intermediate metabolite of anti-oxidant glutathione, and was linked to oxidative stress biomarkers and acylcarnitines (Kumar and Bachhawat 2012). Interestingly, four of the exposure-related potential metabolites we identified also showed similar patterns of alteration in early stage non-small cell carcinoma patients, including increased serum aspartic acid and carnitine, and decreased serum malic acid and pyroglutamic acid (Klupczynska et al. 2016a; Klupczynska et al. 2016b; Klupczynska et al. 2017). Our findings suggest multiple carcinogens exposure may have diverse effects on children and adolescents, causing disruptions in various biological mechanisms such as fatty acid oxidation, energy production, oxidative stress, and amino acid metabolism.

In this study, we found children and adolescents living near a petrochemical complex had increased exposure to multiple carcinogens which induced metabolic changes associated with early health effects including increased oxidative stress and altered serum acylcarnitines, both of which may lead to increased cancer risk. Our findings may provide an explanation for increased cancer incidence among adult residents living near the same petrochemical complex reported in previous studies (Chen et al. 2018; Yuan et al. 2018).

There are limitations to this study. Firstly, we analyzed metabolomics using single analytical platform which limited the number of potential metabolite features detected, and cannot provide a comprehensive view of the metabolome. Secondly, we applied in-house library match using m/z for metabolite identification and therefore could not rule out the possibility of inaccurate metabolite identification and could not provide exact quantification of potential metabolites. Thirdly, our sample size was limited, which could possibly explain why three of the four oxidative stress biomarkers were increased but did

65

not reach statistically significant difference between high and low exposure groups. Lastly, this is a cross-sectional study using single urine and serum samples, and therefore we could not confirm biomarker stability and could not be certain if the potential metabolites we identified can serve as life-long indicators of increased cancer risk.
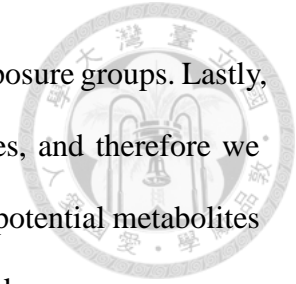
Our findings imply multiple carcinogens exposure during critical periods of childhood and adolescence development induce metabolic perturbations in children and adolescents linking to early health effects that may contribute to cancer risk later in life. This indicates significant reduction of toxic emissions from the complex could decrease carcinogens exposure and metabolic abnormities, which may potentially reduce cancer risks in children and adolescents living nearby. We recommend longitudinal epidemiological studies in this area to follow up on children and adolescents' health if carcinogens emission continues in the near future.
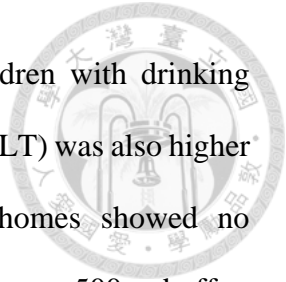
4.3 Part 3. Lipidomics of Children and Adolescents Exposed to Industrial Pollutants.

Part 2 of our study focused on carcinogenic pollutants and cancer-related biological mechanisms and pathways. Previous epidemiological studies have reported increased risks of other chronic and acute adverse health effects for residents in this industrial community, including CKD, hyperlipidemia, subclinical liver dysfunctions, and allergic respiratory diseases (Chiang et al. 2016; Wang et al. 2019; 柯登元 2016; 孫稚翔 2017; 莊明潔 2018). We aim in Part 3 of our study to use serum lipidomics to identify intermediate biomarkers linking multiple industrial pollutants exposure and biomarkers of early health effects, that have also been implicated in chronic and acute diseases.

4.3.1 Results

Serum samples from the same 107 children and adolescents that participated in the Part 2 study were analyzed for untargeted lipidomics. Figure 12 showed the PCA score plot of 61 QC samples from 13 batches of analysis, with all five QCs analyzed in batch 13 (displayed names s13-1 to s13-5) outside of the 95% confidence interval. We removed data from eight participants in batch 13, leaving 99 out of 107 participants for statistical analysis. Table 12 showed the comparison of basic characteristics, external and internal exposure levels, and urine oxidative stress biomarker levels between high and low exposure groups of the 99 children participants included in serum lipidomics analysis. High exposure group lived 10.71±2.33 km and 10.20±2.18 km away from the main emission points of coal-fired power plant and oil refineries, respectively, while low exposure group lived 21.66±5.81 km and 20.77±5.54 km away, respectively. High and low exposure group showed no significant difference in age, sex distribution, SBP, aspartate transaminase (AST), smoking history, and betelnut history. However, high

exposure group had significantly higher BMI and number of children with drinking history compared to low exposure group, and alanine transaminase (ALT) was also higher with borderline statistical significance. Road area surrounding homes showed no significant difference between the two exposure groups at neither 25 m or 500 m buffer. Urine concentrations of exposure biomarkers As, Cd, Cr, Ni, 1-OHP, V, Hg, Mn, Cu, Sr, and Tl were increased in high exposure group compared to low exposure group, with As, Cd, 1-OHP, V, Hg, Sr, and Tl reaching statistical significance. The difference was most profound in 1-OHP, V, and Tl. Urine concentrations of oxidative stress biomarkers showed all four biomarkers were increased in high exposure group compared to low exposure group, but only HNE-MA reached statistical significance (p=0.010), and 8-NO$_2$Gua reaching borderline significance (p=0.059). Only 97 out of the 99 subjects had available data for urine oxidative stress biomarkers concentrations.
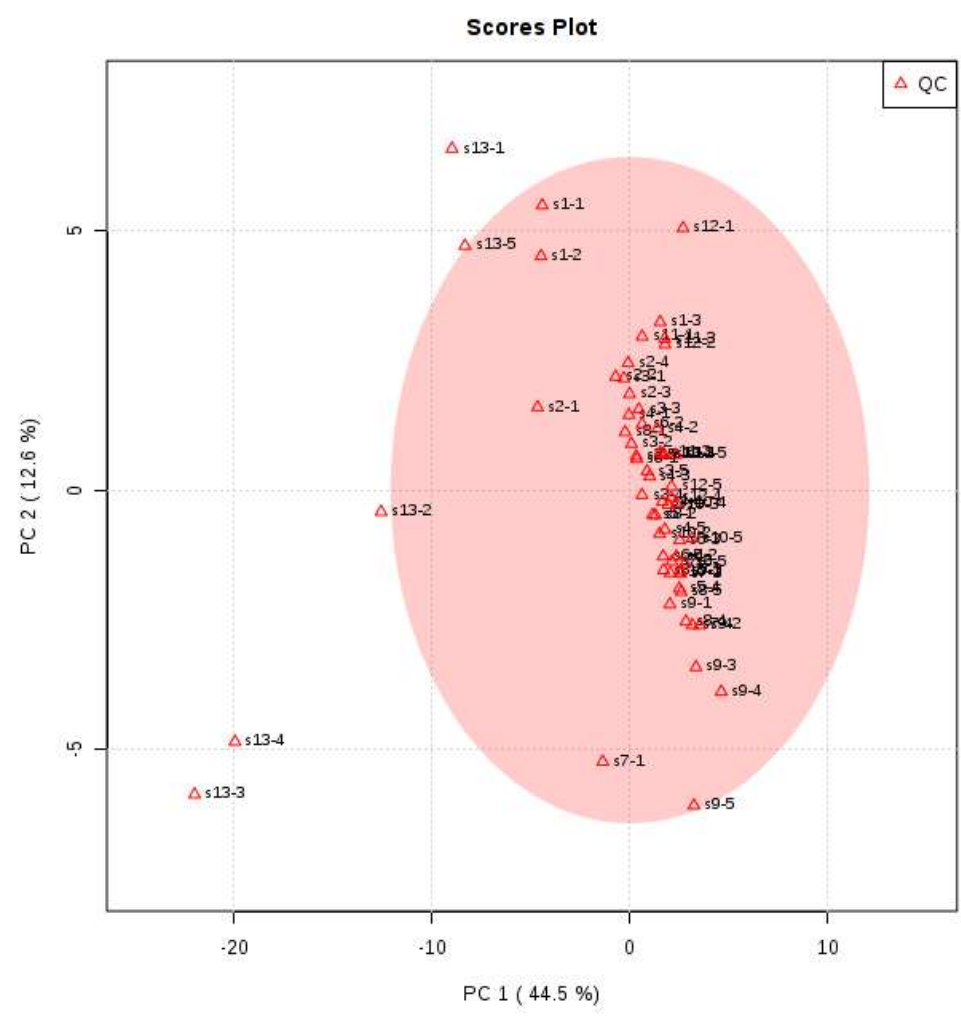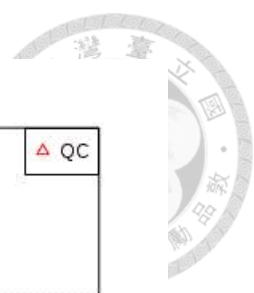
Figure 12. Principle component analysis score plot of 61 pooled quality control (QC) samples data from 13 batches detected using UHPLC-qTOFMS analysis

Table 12. Comparison of basic characteristics, exposure levels, and oxidative stress biomarker levels in 99 study subjects.

	High (n=34)	Low (n=65)	p^a
Basic characteristics			
Age, mean±SD	13.65±0.94	13.70±0.93	0.785
Male, n(%)	18(52.9)	35(53.9)	0.932
SBP, mean±SD	118.3±13.44	116.2±14.06	0.480
BMI, mean±SD	21.54±3.25	20.07±3.55	0.047
AST[b], mean±SD	20.24±3.03	20.54±4.25	0.683
ALT[b], mean±SD	14.68±5.98	12.57±5.14	0.070
Smoke history, n(%)	4(11.8)	4(6.2)	0.441
Drink history, n(%)	5(14.7)	2(3.1)	0.045
Betelnut history, n(%)	1(2.9)	2(3.1)	1.000
External exposures[c], mean±SD			
Distance to coal-fired power plant	10.71±2.33	21.66±5.81	<0.0001
Distance to oil refinery	10.20±2.18	20.77±5.54	<0.0001
Road area surrounding homes			
25 m buffer	303.2±213.5	324.4±210.5	0.665
500 m buffer	69973.4±26069.9	63749.4±20308.0	0.195
Internal exposures[d], mean±SD			
Arsenic	61.99±43.53	38.67±27.47	0.006
Cadmium	0.34±0.35	0.19±0.15	0.014
Chromium	3.20±3.08	2.14±1.66	0.289
Nickel	7.13±8.97	3.95±2.92	0.258
Lead	0.66±0.66	0.67±0.67	0.899
1-OHP	0.20±0.15	0.03±0.01	<0.0001
Vanadium	2.27±1.46	0.23±0.10	<0.0001
Mercury	3.24±2.96	1.86±1.86	0.033
Manganese	2.10±2.16	1.22±0.76	0.108
Copper	14.63±8.60	10.83±7.71	0.076
Strontium	144.68±155.63	72.95±65.90	0.004
Thallium	2.05±3.94	0.21±0.12	<0.0001
Oxidative stress[e], mean±SD			
8-OHDG	3.01±2.18	2.64±2.88	0.219
HNE-MA	1.60±1.19	1.31±2.21	0.010
8-isoPF$_{2\alpha}$	2.88±2.77	2.07±2.19	0.145
8-NO$_2$Gua	6.68±11.90	2.53±3.07	0.059

[a] Comparison of basic characteristics between the high and low exposure groups for

70

continuous variables was performed using Student's t-test, and for discrete variables, Chi-squared test or Fisher's exact test. Urinary exposure biomarker concentrations are log-transformed, high and low exposure groups compared by ANCOVA test adjusting age, sex, smoking, alcohol consumption, betel nut chewing, fish consumption, and BMI with a post comparison by Scheffe test. Urinary oxidative stress biomarker concentrations are log-transformed, high and low exposure groups compared by Student's t-test.

[b] Unit: IU/L

[c] Distance to source: Average of home-to-coal-fired power plant and home-to-oil refinery distance, unit: km; Road area surrounding homes unit: m^2.

[d] For urine 1-OHP, unit: μmol/mol-creatinine; for urine heavy metals, unit: μg/g-creatinine

[e] Urine oxidative stress biomarkers unit: μg/g-creatinine. High exposure group N=33, low exposure group N=64.

SBP: Systolic blood pressure; BMI: Body Mass Index; AST: aspartate transaminase; ALT: alanine transaminase; TC: total cholesterol; HDL-C: high-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol; TC: triacylglycerol;

Figure 13 showed serum levels of eight acylcarnitines that were significantly different in high exposure group compared to low exposure group in 99 study subjects. Samples are in columns and arranged according to high exposure (red) and low exposure (green) groups. Acylcarnitines are in rows and arranged according to hierarchical clustering using Euclidean distance measure and Ward algorithm. The colors vary from deep blue to dark brown to indicate data values change from down-regulation (blue) to up-regulation (brown). We found long-chain acylcarnitines were clustered together and down-regulated in high exposure group compared to low exposure group (Dodecanoylcarnitine, C12; Tetradecanoylcarnitine, C14; Tetradecenoylcarnitine, C14:1; Pentadecanoylcarnitine, C15; Hexadecenoylcarnitine, C16:1; Linoleylcarnitine, C18:2; Pristanoylcarnitine, C19), while short-chain acylcarnitine (Hexanoylcarnitine, C6) was up-regulated in high exposure group compared to low exposure group.

71
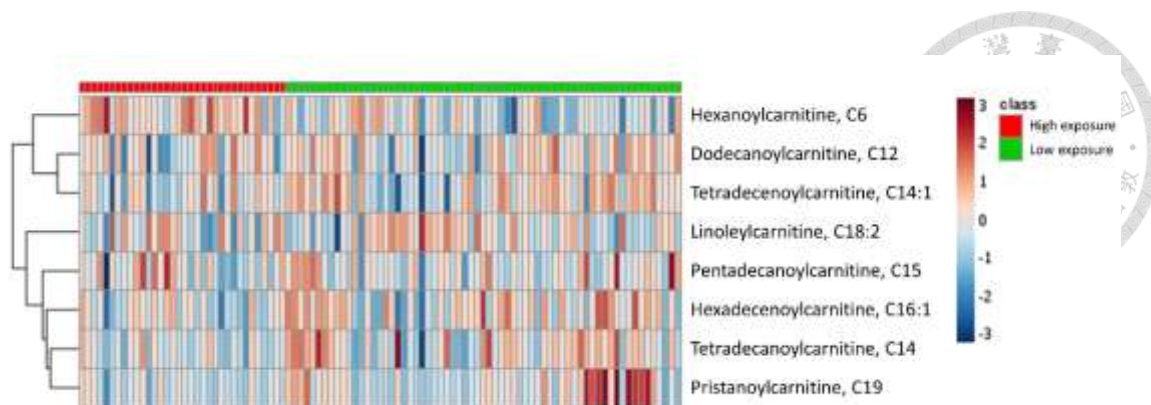
Figure 13. Heat map of serum acylcarnitine levels in 99 study subjects

Lipidomics identified 134 potential lipid features in study subjects serum samples after removing features missing in more than 50% of the samples, and with >20% RSD in QC samples. These 134 lipid features corresponded to 64 lipid species from six lipid subclasses that could be classified into two lipid categories, glycerophospholipids (including three lipid subclasses: lysophosphatidylcholine (LPC), phosphatidylcholine (PC), and phosphatidylinositol (PI)) and sphingolipids (including three lipid subclasses: ceramide (CER), cerebroside (CB), and sphingomyelin (SM)) (Appendix 3). All 134 potential lipid features were put through PLS-DA analysis, and results showed lipid profiles between high and low exposure groups could be significantly separated by three components that accounted for 13.6, 6.9, and 7.7 % of variability of lipid profiles between high and low exposure groups, respectively (Accuracy=0.85, R^2=0.66, Q^2=0.42) (Figure 14A). Permutation test was performed to confirm the validity of PLS-DA model ($p < 0.01$). PLS-DA and ANCOVA analysis adjusting for age, sex, and BMI found 18 exposure-related potential lipid features (average variable importance in projection (VIP) score >1, ANCOVA p <0.05). Figure 14B showed the up- and down-regulation of exposure-related potential lipids in high and low exposure groups. Samples are in columns arranged according to high exposure (red) and low exposure (green) groups. Potential lipids are in rows and arranged according to hierarchical clustering using

72

Euclidean distance measure and Ward algorithm. The colors vary from deep blue to dark brown to indicate data values change from down-regulation (blue) to up-regulation (brown). We found potential lipids up-regulated in high exposure group compared to low exposure group included one LPC, four PCs, and two SMs, while down-regulated potential lipids included three LPCs, five PCs, two PIs, and one SM (Figure 14B).

73
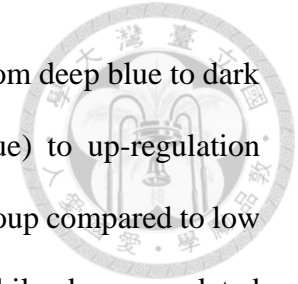
(A)
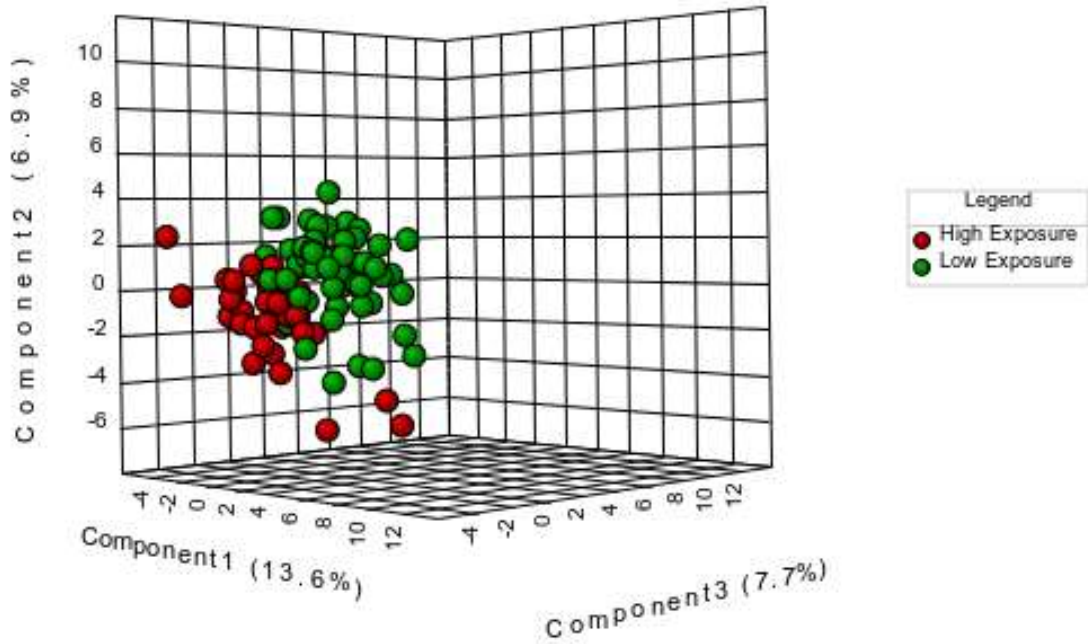
(B)
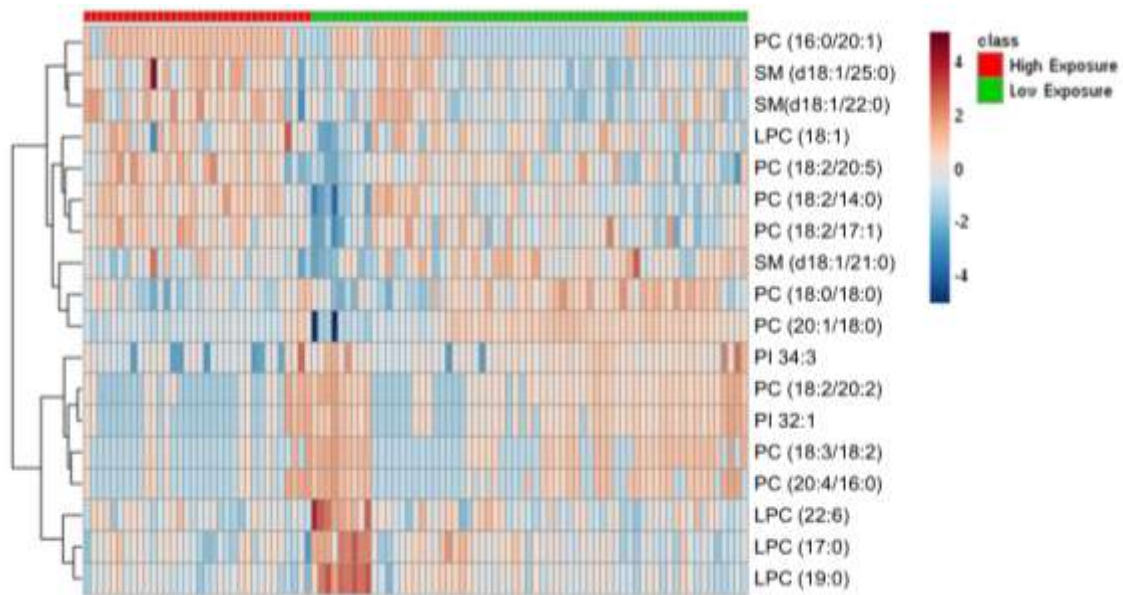
Figure 14. Comparison of serum lipid profile in 99 study subjects using (A) PLS-DA score plot (Accuracy=0.85, R^2=0.66, Q^2=0.42, Permutation test p <0.01) and (B) heat map of exposure-related potential metabolite levels (average VIP score >1, ANOVA p <0.05 are shown adjusted for sex, age, and BMI).

PC: phosphatidylcholine; LPC: lysophosphatidylcholine; SM: sphingomyelin; PI: phosphatidylinositol; chain length and number of cis-double bonds are shown in brackets

Table 13 showed the association between 12 individual carcinogens with four oxidative stress biomarkers in 97 study subjects. Individually, 8-OHDG was significantly associated with Hg (p=0.010), while association with Cr (p=0.060) and Sr (p=0.078) were borderline significant. HNE-MA was significantly associated with V (p=0.013), Hg (p=0.013), and Mn (p=0.039), while association with Tl was borderline significant (p=0.068). For 8-isoPF$_{2\alpha}$, there was only borderline significant association with V (p=0.053). For 8-NO$_2$Gua, the association with Cd (p=0.024), V (p=0.002), and Tl (p=0.001) all reached statistical significance. Figure 15A to 15D showed the WQS regression analysis of the association of combined 12 carcinogens exposure with four oxidative stress biomarkers, respectively. Association were positive and statistically significant with 8-OHDG, HNE-MA, and 8-NO$_2$Gua. Associations with 8-isoPF$_{2\alpha}$ was positive but did not reach statistical significance (p=0.102). For 8-OHDG, Hg predominated in the mixture index (38.7%) followed by Sr (15.9%) (p=0.009) (Figure 15A). Figure 15B showed Hg (25.6%), Mn (21.9%), and V (21.0%) contributed to over half of the mixture index positively associated with HNE-MA levels (p=0.0005), and Tl also contributed to the association (10.3%). Associations with 8-isoPF$_{2\alpha}$ was predominated by Hg (25.7%), V (22.9%), and Ni (20.9%), with contributions from Mn (14.4%) (Figure 15C). In Figure 15D, we can see in the mixture index positively associated with 8-NO$_2$Gua (p=0.00008), V (43.2%) contributed to nearly half of the association, followed by Tl (14.8%) and Ni (14.4%).

Table 13. Individual association between urine exposure biomarkers and oxidative stress biomarkers in 97 study subjects.

	8-OHDG			HNE-MA			8-isoPF$_{2\alpha}$			8-NO$_2$Gua		
	Estimate	95% CI	*p* value	Estimate	95% CI	*p* value	Estimate	95% CI	*p* value	Estimate	95% CI	*p* value
Arsenic	0.026	(-0.069, 0.121)	0.588	0.053	(-0.056, 0.161)	0.337	-0.043	(-0.141, 0.055)	0.388	0.002	(-0.149, 0.153)	0.978
Cadmium	-0.014	(-0.116, 0.086)	0.773	0.035	(-0.080, 0.151)	0.545	-0.060	(-0.164, 0.043)	0.251	0.179	(0.024, 0.335)	0.024
Chromium	-0.081	(-0.166, 0.003)	0.060	0.010	(-0.109, 0.089)	0.846	-0.031	(0.120, 0.058)	0.491	0.054	(-0.082, 0.190)	0.434
Nickel	0.009	(-0.075, 0.919)	0.840	0.016	(-0.079, 0.111)	0.740	0.055	(-0.030, 0.141)	0.200	0.097	(-0.034, 0.227)	0.145
Lead	0.049	(-0.050, -.148)	0.331	0.066	(-0.048, 0.180)	0.252	-0.059	(-0.161, 0.043)	0.256	-0.013	(-0.172, 0.146)	0.871
1-OHP	0.038	(-0.027, 0.105)	0.247	0.054	(-0.021, 0.130)	0.157	0.025	(-0.044, 0.093)	0.471	0.087	(-0.017, 0.191)	0.098
Vanadium	0.054	(-0.010, 0.119)	0.099	0.093	(0.020, 0.165)	0.013	0.065	(-0.001, 0.132)	0.053	0.161	(0.062, 0.259)	0.002
Mercury	0.106	(0.026, 0.185)	0.010	0.117	(0.025, 0.208)	0.013	0.051	(-0.033, 0.136)	0.231	0.005	(-0.126, 0.135)	0.943
Manganese	0.029	(-0.087, 0.144)	0.624	0.136	(0.007, 0.265)	0.039	0.049	(-0.070, 0.167)	0.418	0.098	(-0.084, 0.279)	0.288
Copper	0.014	(-0.124, 0.151)	0.842	0.029	(-0.128, 0.186)	0.716	-0.038	(-0.179, 0.104)	0.596	0.086	(-0.131, 0.303)	0.434
Strontium	0.075	(-0.009, 0.159)	0.078	0.054	(-0.043, 0.151)	0.270	-0.008	(-0.096, 0.080)	0.861	0.103	(-0.030, 0.236)	0.128
Thallium	-0.016	(-0.088, 0.055)	0.650	0.075	(-0.006, 0.156)	0.068	0.038	(-0.036, 0.116)	0.308	0.192	(0.086, 0.299)	0.001

Linear regression analysis adjusted for sex, age, and BMI
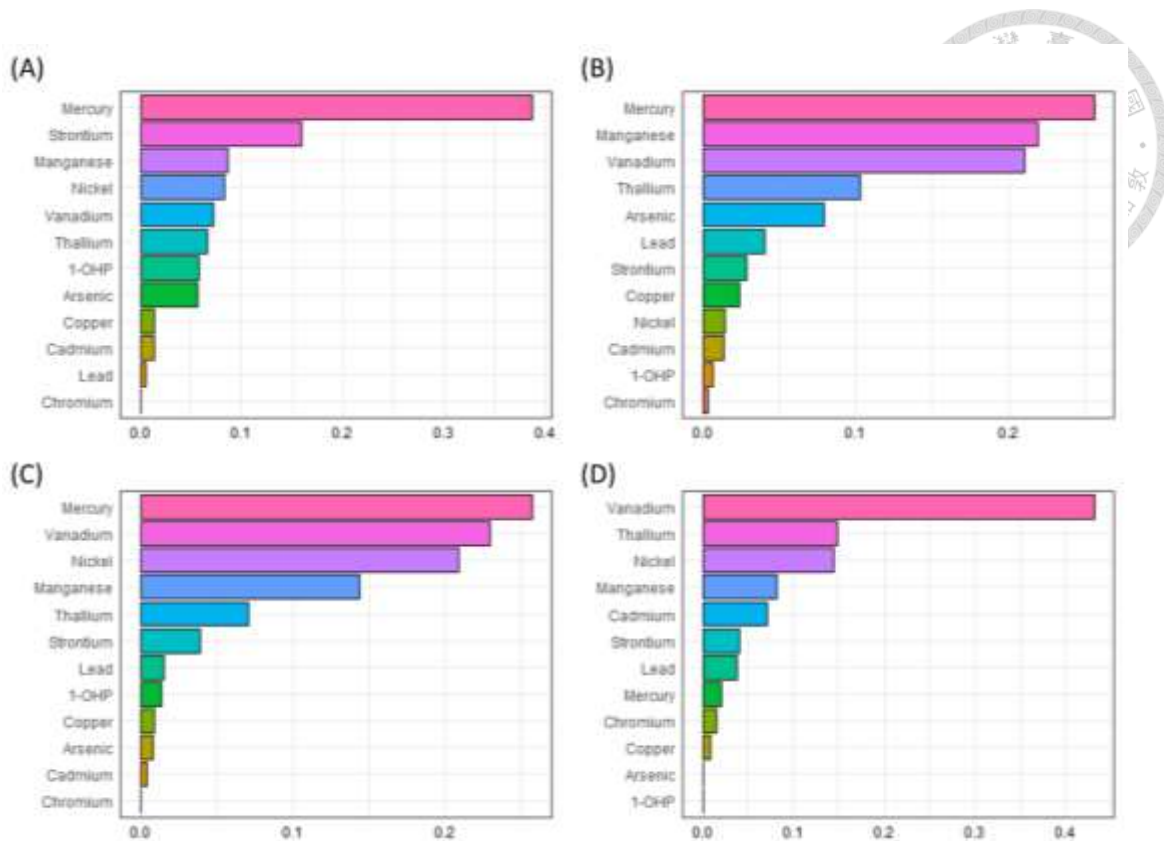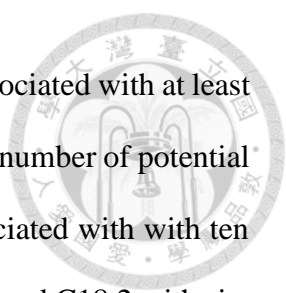
Figure 15. Combined associations between internal exposure levels and (A) 8-OHDG (p=0.009), (B) HNE-MA (p=0.0005), (C) 8-isoPF$_{2\alpha}$ (p= 0.1022), and (D) 8-NO₂Gua (p=0.00008) levels based on weighted quantile sum (WQS) regression analysis in 97 study subjects.
(Adjusted for sex, age, and BMI)

"Meet-in-the middle" approach identified ten potential lipids that were both exposure-related and associated with at least one oxidative stress biomarker (Table 14), and 16 that were both exposure-related and associated with at least one serum acylcarnitine (Table 15). Table 14 and 15 showed the level of association between exposure-related potential lipids (in rows) and biomarkers of early health effects (in columns). For oxidative stress biomarkers, 8-OHDG was significantly associated with three LPCs and two PCs. HNE-MA was significantly associated with five PCs, and one SM. 8-isoPF$_{2\alpha}$ was not significantly associated with any potential lipids, and 8-NO₂Gua was significantly associated with one PC. For serum acylcarnitines, short-chain acylcarnitine (C6) was significantly associated with one LPC, three PCs, two PIs, and

77

two SMs. While long-chain acylcarnitines were each significantly associated with at least eight potential lipids. C14 was significantly associated with the most number of potential lipids, 11 including one LPC, eight PCs, and two PIs. C19 was associated with with ten potential lipids, C12 and C16.1 with nine, C14.1 and C15 with eight, and C18.2 with six. Overall, only C18.2 was associated with LPC(18:1), and it is also the only long-chain acylcarnitine associated with SMs.

Table 14. Association between exposure-related potential lipids and oxidative stress biomarkers in 97 study subjects

(Estimates of linear regression analysis are shown. 95% CI are in brackets. *$p<0.05$ †$p<0.01$ ‡$p<0.001$).

	8-OHDG	HNEMA	8-isoPF$_{2\alpha}$	8-NO$_2$Gua
LPC (17:0)	-0.084*	-0.066	-0.056	-0.081
	(-0.166, -0.002)	(-0.161, 0.028)	(-0.142, 0.030)	(-0.213, 0.050)
LPC (18:1)	0.023	0.014	-0.010	0.071
	(-0.058, 0.103)	(-0.078, 0.107)	(-0.094, 0.073)	(-0.057, 0.198)
LPC (19:0)	-0.092*	-0.082	-0.026	-0.022
	(-0.174, -0.010)	(-0.177, 0.013)	(-0.113, 0.061)	(-0.156, 0.112)
LPC (22:6)	-0.080*	-0.069	-0.012	-0.073
	(-0.159, -0.002)	(-0.160, 0.022)	(-0.095, 0.070)	(-0.200, 0.053)
PC (16:0/20:1)	0.053	0.104*	0.069	0.145*
	(-0.030, 0.135)	(0.012, 0.197)	(-0.016, 0.153)	(0.017, 0.273)
PC (18:0/18:0)	-0.052	-0.120*	-0.008	-0.087
	(-0.137, 0.033)	-(0.214, -0.025)	(-0.096, 0.080)	(-0.221, 0.047)
PC (18:2/14:0)	0.097*	0.115*	0.015	0.047
	(0.017, 0.176)	(0.025, 0.206)	(-0.070, 0.099)	(-0.082, 0.176)
PC (18:2/17:1)	0.078	0.117*	0.032	0.031
	(-0.003, 0.159)	(0.026, 0.208)	(-0.053, 0.117)	(-0.099, 0.161)
PC (18:2/20:2)	-0.079	-0.087	-0.024	-0.114
	(-0.161, 0.003)	-(0.181, 0.007)	(-0.110, 0.062)	(-0.244, 0.016)
PC (18:2/20:5)	0.061	0.128†	0.065	0.067
	(-0.025, 0.146)	(0.033, 0.223)	(-0.023, 0.153)	(-0.069, 0.203)
PC (18:3/18:2)	-0.090*	-0.029	0.022	-0.048
	(-0.170, -0.010)	(-0.122, 0.065)	(-0.062, 0.106)	(-0.177, 0.082)
PC (20:1/18:0)	0.047	0.046	-0.020	-0.036
	(-0.036, 0.130)	(-0.049, 0.141)	(-0.106, 0.066)	(-0.168, 0.096)
PC (20:4/16:0)	-0.037	-0.036	0.058	-0.032
	(-0.119, 0.045)	(-0.130, 0.058)	(-0.026, 0.142)	(-0.162, 0.099)
PI (32:1)	-0.055	-0.063	0.011	-0.095
	(-0.138, 0.027)	(-0.157, 0.032)	(-0.075, 0.097)	(-0.225, 0.035)
PI (34:3)	0.003	-0.039	0.018	-0.058
	(-0.082, 0.088)	(-0.136, 0.057)	(-0.070, 0.105)	(-0.192, 0.076)

79

SM (d18:1/21:0)	0.004	0.093*	0.020	0.017
	(-0.078, 0.086)	(0.001, 0.185)	(-0.065, 0.105)	(-0.112, 0.147)
SM (d18:1/22:0)	-0.018	-0.045	-0.020	0.083
	(-0.100, 0.063)	(-0.138, 0.048)	(-0.104, 0.064)	(-0.045, 0.211)
SM (d18:1/25:0)	0.002	0.064	0.032	0.064
	(-0.080, 0.084)	(-0.029, 0.157)	(-0.053, 0.116)	(-0.065, 0.194)

Linear regression model adjusted for age, sex, and BMI. PC: phosphatidylcholine; LPC: lysophosphatidylcholine; SM: sphingomyelin; PI: phosphatidylinositol; tail lengths and number of cis-double bonds are shown in brackets.

80
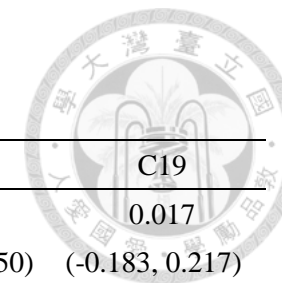
Table 15. Association between exposure-related potential lipids and acylcarnitines in 99 study subjects

(Estimates of linear regression analysis are shown. 95% CI are in brackets. *$p<0.05$ †$p<0.01$ ‡$p<0.001$).

	C6	C12	C14	C14.1	C15	C16.1	C18.2	C19
LPC (17:0)	-0.164	-0.006	0.085	0.045	0.062	0.064	0.040	0.017
	(-0.353, 0.026)	(-0.198, 0.185)	(-0.115, 0.284)	(-0.151, 0.242)	(-0.139, 0.263)	(-0.139, 0.267)	(-0.169, 0.250)	(-0.183, 0.217)
LPC (18:1)	-0.169	-0.045	-0.079	-0.137	-0.102	-0.137	0.357‡	-0.150
	(-0.353, 0.015)	(-0.231, 0.141)	(-0.273, 0.115)	(-0.327, 0.052)	(-0.297, 0.092)	(-0.333, 0.058)	(0.167, 0.547)	(-0.342, 0.043)
LPC (19:0)	-0.209*	0.081	0.219*	0.105	0.104	0.089	-0.027	0.147
	(-0.398, -0.020)	(-0.111, 0.274)	(0.022, 0.415)	(-0.092, 0.302)	(-0.098, 0.305)	(-0.115, 0.293)	(-0.238, 0.184)	(-0.052, 0.347)
LPC (22:6)	-0.001	0.096	0.185	0.185	0.291†	0.225*	-0.155	0.188
	(-0.187, 0.185)	(-0.088, 0.280)	(-0.004, 0.375)	(-0.001, 0.372)	(0.106, 0.476)	(0.035, 0.416)	(-0.355, 0.045)	(-0.002, 0.377)
PC (16:0/20:1)	0.146	-0.305‡	-0.298†	-0.349†	-0.306†	-0.377‡	-0.272‡	-0.397‡
	(-0.041, 0.333)	(-0.483, -0.127)	(-0.485, -0.110)	(-0.529, -0.169)	(-0.494, -0.118)	(-0.561, -0.192)	(-0.471, -0.074)	(-0.576, -0.217)
PC (18:0/18:0)	-0.114	0.355‡	0.340†	0.270†	0.068	0.217*	0.149	0.443‡
	(-0.309, 0.080)	(0.175, 0.536)	(0.148, 0.531)	(0.077, 0.462)	(-0.137, 0.272)	(0.016, 0.419)	(-0.062, 0.360)	(0.261, 0.626)
PC (18:2/14:0)	0.261‡	-0.197*	-0.356‡	-0.269†	-0.276†	-0.339‡	-0.045	-0.452‡
	(0.082, 0.440)	(-0.379, -0.015)	(-0.536, -0.175)	(-0.453, -0.086)	(-0.463, -0.088)	(-0.524, -0.155)	(-0.249, 0.158)	(-0.623, -0.281)
PC (18:2/17:1)	0.037	-0.236*	-0.347‡	-0.186	-0.146	-0.192	0.215*	-0.423‡
	(-0.151, 0.225)	(-0.417, -0.056)	(-0.529, -0.166)	(-0.375, 0.002)	(-0.341, 0.048)	(-0.387, 0.002)	(0.016, 0.415)	(-0.598, -0.248)
PC (18:2/20:2)	-0.243*	0.223*	0.336†	0.306†	0.348‡	0.333‡	0.086	0.357‡
	(-0.428, -0.057)	(0.037, 0.408)	(0.149, 0.522)	(0.120, 0.492)	(0.160, 0.535)	(0.143, 0.523)	(-0.122, 0.294)	(0.171, 0.542)

PC (18:2/20:5)	0.269‡	-0.178	-0.310†	-0.156	-0.186	-0.101	0.014	-0.295‡
	(0.077, 0.460)	(-0.372, 0.016)	(-0.506, -0.113)	(-0.357, 0.044)	(-0.391, 0.018)	(-0.309, 0.108)	(-0.203, 0.230)	(-0.493, -0.097)
PC (18:3/18:2)	-0.056	0.339‡	0.371‡	0.374‡	0.306‡	0.320‡	0.005	0.427‡
	(-0.243, 0.132)	(0.166, 0.513)	(0.191, 0.551)	(0.197, 0.550)	(0.120, 0.493)	(0.133, 0.507)	(-0.199, 0.210)	(0.252, 0.602)
PC (20:1/18:0)	-0.068	0.220*	0.118	0.134	0.057	0.141	0.371‡	0.123
	(-0.259, 0.123)	(0.035, 0.404)	(-0.079, 0.315)	(-0.060, 0.327)	(-0.143, 0.257)	(-0.058, 0.341)	(0.177, 0.565)	(-0.074, 0.321)
PC (20:4/16:0)	-0.135	0.280‡	0.364‡	0.361‡	0.315‡	0.365‡	0.115	0.352‡
	(-0.322, 0.052)	(0.101, 0.459)	(0.182, 0.546)	(0.182, 0.539)	(0.128, 0.502)	(0.180, 0.550)	(-0.090, 0.319)	(0.169, 0.535)
PI (32:1)	-0.192*	0.238*	0.318†	0.320†	0.335‡	0.345‡	0.103	0.330‡
	(-0.379, -0.005)	(0.054, 0.421)	(0.130, 0.506)	(0.136, 0.505)	(0.147, 0.523)	(0.156, 0.534)	(-0.104, 0.310)	(0.143, 0.517)
PI (34:3)	-0.204*	0.177	0.318†	0.212*	0.311‡	0.280‡	0.200	0.285‡
	(-0.390, -0.017)	(-0.010, 0.363)	(0.130, 0.506)	(0.021, 0.402)	(0.121, 0.501)	(0.086, 0.473)	(-0.004, 0.404)	(0.095, 0.475)
SM (d18:1/21:0)	-0.048	0.025	-0.048	0.047	0.130	0.006	-0.008	-0.095
	(-0.237, 0.140)	(-0.162, 0.213)	(-0.243, 0.148)	(-0.145, 0.240)	(-0.066, 0.326)	(-0.193, 0.205)	(-0.213, 0.198)	(-0.291, 0.100)
SM (d18:1/22:0)	0.186*	-0.100	-0.123	-0.066	-0.114	-0.180	-0.327‡	0.063
	(0.002, 0.369)	(-0.286, 0.086)	(-0.316, 0.071)	(-0.258, 0.126)	(-0.309, 0.081)	(-0.375, 0.014)	(-0.520, -0.134)	(-0.132, 0.258)
SM (d18:1/25:0)	0.216*	-0.136	-0.143	-0.118	-0.037	-0.144	-0.255†	-0.153
	(0.032, 0.400)	(-0.322, 0.050)	(-0.337, 0.052)	(-0.310, 0.074)	(-0.235, 0.161)	(-0.341, 0.053)	(-0.454, -0.056)	(-0.347, 0.041)
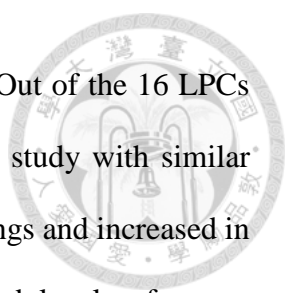
Linear regression model adjusted for age, sex, and BMI. PC: phosphatidylcholine; LPC: lysophosphatidylcholine; SM: sphingomyelin; PI: phosphatidylinositol; tail lengths and number of cis-double bonds are shown in brackets.
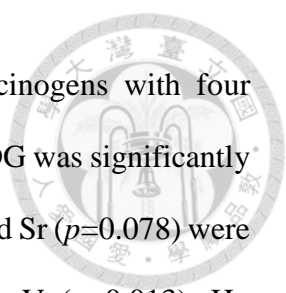
4.3.2 Discussion

This is the first study to use an untargeted lipidomics approach to examine changes in serum lipid profile after multiple industrial pollutants exposure that could be linked to early health effects in children and adolescents. Previous studies mostly used animal models and focused on single toxic exposure. Li et al. reported PAH benzo[a]pyrene exposure reduced levels of serum LPC(18:0) and LPC(22:0), serum PCs with a total of 36-40 carbons in the two acyl chains, and serum PIs in mice (Li et al. 2019). Low dose Cd exposure in rat model led to decreased serum PC(18:4/18:0) and increased LPC(20:0) (Hu et al. 2018). Serum analysis of mice chronically exposed to low dose inorganic arsenic showed increased levels of LPC(O-18:0), LPC(20:3), LPC(18:1), and LPC(22:6) (Chi et al. 2019). We observed similar trends in our study participants including increased levels of LPC(18:1), reduced levels of PCs with total of 36-38 carbons in the two acyl chains except for PC(18:2/20:5) which was increased, and reduced levels of PIs. Our findings suggest multiple industrial pollutants exposure induced deregulations of serum lipid profiles in children and adolescents.

Reduced levels of serum LPCs have been reported in chronic liver dysfunctions, including liver cirrhosis, non-alcoholic fatty liver disease, and hepatocellular carcinoma (Orešič et al. 2013; Zhou et al. 2012a; Zhou et al. 2012b). This reduction of serum LPCs might be due to reduced activities of liver enzymes involved in LPC metabolism and reduced hepatic secretion of LPCs (Sekas et al. 1985; Tahara et al. 1993). Krautbauer et al. found saturated LPCs LPC(16:0) and LPC(18:0) were significantly associated with residual liver functions, but association was not found for unsaturated LPCs (Krautbauer et al. 2016). Zhou et al. associated decreased LPCs, both saturated and unsaturated, with

chronic liver disease, liver cirrhosis, and hepatocellular carcinoma. Out of the 16 LPCs they identified in their study, LPC(22:6) was also identified in our study with similar down-regulated trends, while LPC(18:1) was decreased in their findings and increased in ours. In contrast to our results, they found significantly decreased levels of serum acylcarnitine C6 in chronic liver disease, liver cirrhosis, and hepatocellular carcinoma patients compared to normal control subjects. Interestingly, they also reported decreased levels of tryptophan in hepatocellular carcinoma patients, which was also identified in Part 1 of our study, along with the deregulation of tryptophan metabolism in children, linking multiple industrial pollutant exposure with all four oxidative stress biomarkers (Chen et al. 2017; Zhou et al. 2012a; Zhou et al. 2012b). In view of the increased ALT in high exposure group compared to low exposure group in our study participants that reached borderline statistical significance, it could be worthwhile to further investigate the possibility of decreased serum LPC(22:6) as an early indicator for liver function abnormities in children and adolescents exposed to multiple industrial pollutants.

5. Conclusion and Recommendation
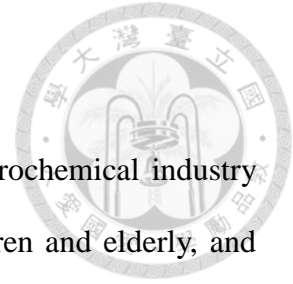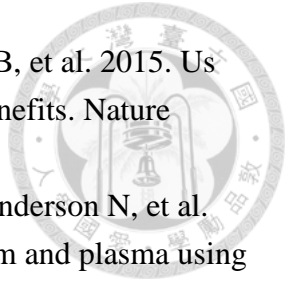
Public health exposome approach could be used in a large petrochemical industry influenced region to identify vulnerable populations such as children and elderly, and understand how multiple industrial pollutants exposure are affecting critical biological mechanisms, leading to early health effects that may be precursors to chronic and acute diseases. Urine metabolomics analyzed via GC-based method could be used to identify vulnerable populations such as children and elderly in regions influenced by a large petrochemical industry, and found age-dependent pathways linking multiple exposures to increased oxidative stress. Serum metabolomics analyzed via LC-based method could be used to find biological pathways affected by multiple industrial carcinogenic pollutants exposure in children and adolescents, that could be linked to cancer-related early health effects. Serum lipidomics analyzed via LC-based method could be used to identify in children and adolescents exposed to multiple industrial pollutants, lipid profile changes that have been implicated in liver dysfunctions. Based on our findings, we suggest significant reduction of petrochemical industrial emissions from the complex to decrease multiple pollutants exposure and metabolic abnormalities, and continued follow up on of residents' health. This dissertation also attests the application of exposomics as a public health research tool, in the investigation of current and potential health impacts of industrial pollution on nearby residents, providing information for future identification of novel personalized health indicators and exposure biomarkers, and establishment of individual risk index.

6. References
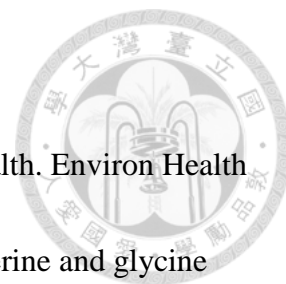
Adler T. 2003. Aging research: The future face of environmental health. Environ Health Perspect 111:A760-765.

Amelio I, Cutruzzola F, Antonov A, Agostini M, Melino G. 2014. Serine and glycine metabolism in cancer. Trends Biochem Sci 39:191-198.

Ayala A, Munoz MF, Arguelles S. 2014. Lipid peroxidation: Production, metabolism, and signaling mechanisms of malondialdehyde and 4-hydroxy-2-nonenal. Oxid Med Cell Longev 2014:360438.

Bartsch H, Nair J. 2005. Accumulation of lipid peroxidation-derived DNA lesions: Potential lead markers for chemoprevention of inflammation-driven malignancies. Mutat Res 591:34-44.

Bester AC, Roniger M, Oren YS, Im MM, Sarni D, Chaoat M, et al. 2011. Nucleotide deficiency promotes genomic instability in early stages of cancer development. Cell 145:435-446.

Blow N. 2008. Metabolomics: Biochemistry's new look. Nature 455:697-700.

Bouatra S, Aziat F, Mandal R, Guo AC, Wilson MR, Knox C, et al. 2013. The human urine metabolome. PLoS One 8:e73076.

Bowler RP, Crapo JD. 2002. Oxidative stress in allergic respiratory diseases. The Journal of allergy and clinical immunology 110:349-356.

Cantu-Medellin N, Kelley EE. 2013. Xanthine oxidoreductase-catalyzed reactive species generation: A process in critical need of reevaluation. Redox Biology 1:353-358.

Carpenter DO, Arcaro K, Spink DC. 2002. Understanding the human health effects of chemical mixtures. Environ Health Perspect 110 Suppl 1:25-42.

Chadeau-Hyam M, Athersuch TJ, Keun HC, De Iorio M, Ebbels TM, Jenab M, et al. 2011. Meeting-in-the-middle using metabolic profiling - a strategy for the identification of intermediate biomarkers in cohort studies. Biomarkers : biochemical indicators of exposure, response, and susceptibility to chemicals 16:83-88.

Chan CC, Shie RH, Chang TY, Tsai DH. 2006. Workers' exposures and potential health risks to air toxics in a petrochemical complex assessed by improved methodology. Int Arch Occup Environ Health 79:135-142.

Chan EC, Pasikanti KK, Nicholson JK. 2011. Global urinary metabolic profiling procedures using gas chromatography-mass spectrometry. Nature protocols 6:1483-1499.

Chen C-HS, Kuo T-C, Kuo H-C, Tseng YJ, Kuo C-H, Yuan T-H, et al. 2019. Metabolomics of children and adolescents exposed to industrial carcinogenic
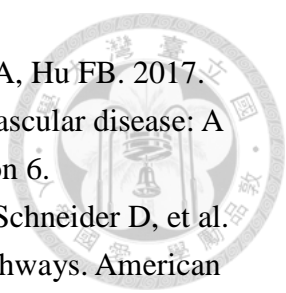
pollutants. Environ Sci Technol.

Chen CF, Chio CP, Yuan TH, Yeh YP, Chan CC. 2018. Increased cancer incidence of changhua residents living in taisi village north to the no. 6 naphtha cracking complex. Journal of the Formosan Medical Association = Taiwan yi zhi 117:1101-1107.

Chen CS, Yuan TH, Shie RH, Wu KY, Chan CC. 2017. Linking sources to early effects by profiling urine metabolome of residents living near oil refineries and coal-fired power plants. Environ Int 102:87-96.

Chen S, Kong H, Lu X, Li Y, Yin P, Zeng Z, et al. 2013. Pseudotargeted metabolomics method and its application in serum biomarker discovery for hepatocellular carcinoma based on ultra high-performance liquid chromatography/triple quadrupole mass spectrometry. Anal Chem 85:8326-8333.

Chen Y, Guillemin GJ. 2009. Kynurenine pathway metabolites in humans: Disease and healthy states. Int J Tryptophan Res 2:1-19.

Chen YM, Lin WY, Chan CC. 2014. The impact of petrochemical industrialisation on life expectancy and per capita income in taiwan: An 11-year longitudinal study. BMC Public Health 14:247.

Chi L, Tu P, Liu CW, Lai Y, Xue J, Ru H, et al. 2019. Chronic arsenic exposure induces oxidative stress and perturbs serum lysolipids and fecal unsaturated fatty acid metabolism. Chem Res Toxicol.

Chiang TY, Yuan TH, Shie RH, Chen CF, Chan CC. 2016. Increased incidence of allergic rhinitis, bronchitis and asthma, in children living near a petrochemical complex with so2 pollution. Environ Int 96:1-7.

Chio CP, Yuan TH, Shie RH, Chan CC. 2014. Assessing vanadium and arsenic exposure of people living near a petrochemical complex with two-stage dispersion models. Journal of Hazardous Materials 271:98-107.

Chong J, Soufan O, Li C, Caraus I, Li S, Bourque G, et al. 2018. Metaboanalyst 4.0: Towards more transparent and integrative metabolomics analysis. Nucleic Acids Res 46:W486-W494.

Chuang HC, Shie RH, Chio CP, Yuan TH, Lee JH, Chan CC. 2018. Cluster analysis of fine particulate matter (pm2.5) emissions and its bioreactivity in the vicinity of a petrochemical complex. Environ Pollut 236:591-597.

Ciprandi G, De Amici M, Tosca M, Fuchs D. 2010. Tryptophan metabolism in allergic rhinitis: The effect of pollen allergen exposure. Human Immunology 71:911-915.

Davis AP, Grondin CJ, Johnson RJ, Sciaky D, McMorran R, Wiegers J, et al. 2019. The comparative toxicogenomics database: Update 2019. Nucleic Acids Res 47:D948-D954.

Driscoll CT, Buonocore JJ, Levy JI, Lambert KF, Burtraw D, Reid SB, et al. 2015. Us power plant carbon standards and clean air and health co-benefits. Nature Climate Change 5:535-540.

Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, Anderson N, et al. 2011. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. Nature protocols 6:1060-1083.

Dybing E, Schwarze PE, Nafstad P, Victorin K, Penning TM. 2013. Chapter 7. Polyclyclic aromatic hydrocarbons in ambient air and cancer. France: International Agency for Research on Cancer.

Ellis JK, Athersuch TJ, Thomas LD, Teichert F, Perez-Trujillo M, Svendsen C, et al. 2012. Metabolic profiling detects early effects of environmental and lifestyle exposure to cadmium in a human population. BMC Med 10:61.

EMA. 2011. Guideline on bioanalytical method validation emea/chmp/ewp/192217/2009 rev. 1.

Ferroni P, Santilli F, Cavaliere F, Simeone P, Costarelli L, Liani R, et al. 2017. Oxidant stress as a major determinant of platelet activation in invasive breast cancer. Int J Cancer 140:696-704.

Fouque D, Holt S, Guebre-Egziabher F, Nakamura K, Vianey-Saban C, Hadj-Aissa A, et al. 2006. Relationship between serum carnitine, acylcarnitines, and renal function in patients with chronic renal disease. Journal of renal nutrition : the official journal of the Council on Renal Nutrition of the National Kidney Foundation 16:125-131.

Fu PP, Xia Q, Sun X, Yu H. 2012. Phototoxicity and environmental transformation of polycyclic aromatic hydrocarbons (pahs)-light-induced reactive oxygen species, lipid peroxidation, and DNA damage. J Environ Sci Health C Environ Carcinog Ecotoxicol Rev 30:1-41.

Gao X, Chen W, Li R, Wang M, Chen C, Zeng R, et al. 2012. Systematic variations associated with renal disease uncovered by parallel metabolomics of urine and serum. BMC Syst Biol 6 Suppl 1:S14.

George J, Masto RE, Ram LC, Das TB, Rout TK, Mohan M. 2015. Human exposure risks for metals in soil near a coal-fired power-generating plant. Archives of environmental contamination and toxicology 68:451-461.

Gostner JM, Becker K, Kofler H, Strasser B, Fuchs D. 2016. Tryptophan metabolism in allergic disorders. Int Arch Allergy Immunol 169:203-215.

Hastings J, de Matos P, Dekker A, Ennis M, Harsha B, Kale N, et al. 2013. The chebi reference database and ontology for biologically relevant chemistry: Enhancements for 2013. Nucleic Acids Res 41:D456-463.
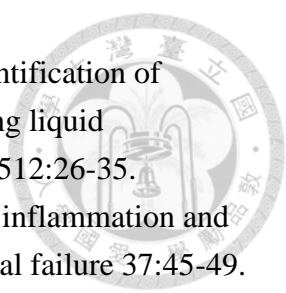
Hiraku Y. 2010. Formation of 8-nitroguanine, a nitrative DNA lesion, in inflammation-related carcinogenesis and its significance. Environmental health and preventive medicine 15:63-72.

Ho TJ, Kuo CH, Wang SY, Chen GY, Tseng YJ. 2013. True ion pick (tipick): A denoising and peak picking algorithm to extract ion signals from liquid chromatography/mass spectrometry data. Journal of mass spectrometry : JMS 48:234-242.

Hopps E, Noto D, Caimi G, Averna MR. 2010. A novel component of the metabolic syndrome: The oxidative stress. Nutrition, metabolism, and cardiovascular diseases : NMCD 20:72-77.

Hu L, Bo L, Zhang M, Li S, Zhao X, Sun C. 2018. Metabonomics analysis of serum from rats given long-term and low-level cadmium by ultra-performance liquid chromatography-mass spectrometry. Xenobiotica; the fate of foreign compounds in biological systems 48:1079-1088.

Hu SW, Chan YJ, Hsu HT, Wu KY, ChangChien GP, Shie RH, et al. 2011. Urinary levels of 1-hydroxypyrene in children residing near a coal-fired power plant. Environ Res 111:1185-1191.

Huang PC, Liu LH, Shie RH, Tsai CH, Liang WY, Wang CW, et al. 2016. Assessment of urinary thiodiglycolic acid exposure in school-aged children in the vicinity of a petrochemical complex in central taiwan. Environ Res 150:566-572.

IARC. 1989a. Occupational exposures in petroleum refining; crude oil and major petroleum fuels. Iarc working group on the evaluation of carcinogenic risks to humans. IARC monographs on the evaluation of carcinogenic risks to humans 45:1-322.

IARC. 1989b. Cyclohexanone. IARC monographs on the evaluation of carcinogenic risks to humans 47:157-169.

Jomova K, Valko M. 2011. Advances in metal-induced oxidative stress and human disease. Toxicology 283:65-87.

Juarez PD, Matthews-Juarez P, Hood DB, Im W, Levine RS, Kilbourne BJ, et al. 2014. The public health exposome: A population-based, exposure science approach to health disparities research. Int J Environ Res Public Health 11:12866-12895.

Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2014. Data, information, knowledge and principle: Back to metabolism in kegg. Nucleic Acids Res 42:D199-205.

Kell DB, Brown M, Davey HM, Dunn WB, Spasic I, Oliver SG. 2005. Metabolic footprinting and systems biology: The medium is the message. Nat Rev Microbiol 3:557-565.

Klupczynska A, Derezinski P, Dyszkiewicz W, Pawlak K, Kasprzyk M, Kokot ZJ.

2016a. Evaluation of serum amino acid profiles' utility in non-small cell lung cancer detection in polish population. Lung cancer (Amsterdam, Netherlands) 100:71-76.

Klupczynska A, Plewa S, Dyszkiewicz W, Kasprzyk M, Sytek N, Kokot ZJ. 2016b. Determination of low-molecular-weight organic acids in non-small cell lung cancer with a new liquid chromatography-tandem mass spectrometry method. Journal of pharmaceutical and biomedical analysis 129:299-309.

Klupczynska A, Derezinski P, Garrett TJ, Rubio VY, Dyszkiewicz W, Kasprzyk M, et al. 2017. Study of early stage non-small-cell lung cancer using orbitrap-based global serum metabolomics. Journal of cancer research and clinical oncology 143:649-659.

Ko JL, Cheng YJ, Liu GC, Hsin IL, Chen HL. 2017. The association of occupational metals exposure and oxidative damage, telomere shortening in fitness equipments manufacturing workers. Industrial health 55:345-353.

Krautbauer S, Eisinger K, Wiest R, Liebisch G, Buechler C. 2016. Systemic saturated lysophosphatidylcholine is associated with hepatic function in patients with liver cirrhosis. Prostaglandins & other lipid mediators 124:27-33.

Kumar A, Bachhawat AK. 2012. Pyroglutamic acid: Throwing light on a lightly studied metabolite. Current Science 102:288-297.

Lai YS, Chen WC, Kuo TC, Ho CT, Kuo CH, Tseng YJ, et al. 2015. Mass-spectrometry-based serum metabolomics of a c57bl/6j mouse model of high-fat-diet-induced non-alcoholic fatty liver disease development. J Agric Food Chem 63:7873-7884.

Li F, Xiang B, Jin Y, Li C, Li J, Ren S, et al. 2019. Dysregulation of lipid metabolism induced by airway exposure to polycyclic aromatic hydrocarbons in c57bl/6 mice. Environ Pollut 245:986-993.

Maiuolo J, Oppedisano F, Gratteri S, Muscoli C, Mollace V. 2016. Regulation of uric acid metabolism and excretion. Int J Cardiol 213:8-14.

Makri A, Stilianakis NI. 2008. Vulnerability to air pollution health effects. International journal of hygiene and environmental health 211:326-336.

Nadal M, Schuhmacher M, Domingo JL. 2004. Metal pollution of soils and vegetation in an area with petrochemical industry. Sci Total Environ 321:59-69.

Nadal M, Mari M, Schuhmacher M, Domingo JL. 2009. Multi-compartmental environmental surveillance of a petrochemical area: Levels of micropollutants. Environ Int 35:227-235.

Ni J, Xu L, Li W, Wu L. 2016. Simultaneous determination of thirteen kinds of amino acid and eight kinds of acylcarnitine in human serum by lc-ms/ms and its application to measure the serum concentration of lung cancer patients.

Biomedical chromatography : BMC 30:1796-1806.

O'Rourke D, Connolly S. 2003. Just oil? The distribution of environmental and social impacts of oil production and consumption.ANNUAL REVIEWS.

Orešič M, Hyötyläinen T, Kotronen A, Gopalacharyulu P, Nygren H, Arola J, et al. 2013. Prediction of non-alcoholic fatty-liver disease and liver fat content by serum molecular lipids. Diabetologia 56:2266-2274.

Orhan H, Vermeulen NP, Tump C, Zappey H, Meerman JH. 2004. Simultaneous determination of tyrosine, phenylalanine and deoxyguanosine oxidation products by liquid chromatography-tandem mass spectrometry as non-invasive biomarkers for oxidative damage. J Chromatogr B Analyt Technol Biomed Life Sci 799:245-254.

Pasikanti KK, Esuvaranathan K, Hong Y, Ho PC, Mahendran R, Raman Nee Mani L, et al. 2013. Urinary metabotyping of bladder cancer using two-dimensional gas chromatography time-of-flight mass spectrometry. Journal of proteome research 12:3865-3873.

Patti GJ, Yanes O, Siuzdak G. 2012. Innovation: Metabolomics: The apogee of the omics trilogy. Nat Rev Mol Cell Biol 13:263-269.

Pedley AM, Benkovic SJ. 2017. A new view into the regulation of purine metabolism: The purinosome. Trends in Biochemical Sciences 42:141-154.

Pennell K. 2016. Population screening for biological and environmental properties of the human metabolic phenotype: Implications for personalized medicine.

Penning TM, Drury JE. 2007. Human aldo-keto reductases: Function, gene regulation, and single nucleotide polymorphisms. Archives of Biochemistry and Biophysics 464:241-250.

Peter AL, Viraraghavan T. 2005. Thallium: A review of public health and environmental concerns. Environ Int 31:493-501.

Poli G. 2000. Pathogenesis of liver fibrosis: Role of oxidative stress. Molecular Aspects of Medicine 21:49-98.

Psychogios N, Hau DD, Peng J, Guo AC, Mandal R, Bouatra S, et al. 2011. The human serum metabolome. PLoS One 6:e16957.

Rappaport SM, Smith MT. 2010. Epidemiology. Environment and disease risks. Science 330:460-461.

Reuter S, Gupta SC, Chaturvedi MM, Aggarwal BB. 2010. Oxidative stress, inflammation, and cancer: How are they linked? Free Radic Biol Med 49:1603-1616.

Robertson DG, Watkins PB, Reily MD. 2011. Metabolomics in toxicology: Preclinical and clinical applications. Toxicological sciences : an official journal of the Society of Toxicology 120 Suppl 1:S146-170.

Ruiz-Canela M, Hruby A, Clish CB, Liang L, Martinez-Gonzalez MA, Hu FB. 2017. Comprehensive metabolomic profiling and incident cardiovascular disease: A systematic review. Journal of the American Heart Association 6.

Rutkowsky JM, Knotts TA, Ono-Moore KD, McCoin CS, Huang S, Schneider D, et al. 2014. Acylcarnitines activate proinflammatory signaling pathways. American journal of physiology Endocrinology and metabolism 306:E1378-1387.

Sekas G, Patton GM, Lincoln EC, Robins SJ. 1985. Origin of plasma lysophosphatidylcholine: Evidence for direct hepatic secretion in the rat. The Journal of laboratory and clinical medicine 105:190-194.

Semba RD, Trehan I, Li X, Moaddel R, Ordiz MI, Maleta KM, et al. 2017. Environmental enteric dysfunction is associated with carnitine deficiency and altered fatty acid oxidation. EBioMedicine 17:57-66.

Senghore T, Li YF, Sung FC, Tsai MH, Hua CH, Liu CS, et al. 2018. Biomarkers of oxidative stress associated with the risk of potentially malignant oral disorders. Anticancer research 38:5211-5216.

Shen S, Zhang R, Zhang J, Wei Y, Guo Y, Su L, et al. 2018. Welding fume exposure is associated with inflammation: A global metabolomics profiling study. Environmental health : a global access science source 17:68.

Shie RH, Chan CC. 2013. Tracking hazardous air pollutants from a refinery fire by applying on-line and off-line air monitoring and back trajectory modeling. J Hazard Mater 261:72-82.

Shie RH, Yuan TH, Chan CC. 2013. Using pollution roses to assess sulfur dioxide impacts in a township downwind of a petrochemical complex. J Air Waste Manag Assoc 63:702-711.

Simoni RE, Gomes LN, Scalco FB, Oliveira CP, Aquino Neto FR, de Oliveira ML. 2007. Uric acid changes in urine and plasma: An effective tool in screening for purine inborn errors of metabolism and other pathological conditions. Journal of inherited metabolic disease 30:295-309.

Sivakumar R, Babu PV, Shyamaladevi CS. 2011. Aspartate and glutamate prevents isoproterenol-induced cardiac toxicity by alleviating oxidative stress in rats. Exp Toxicol Pathol 63:137-142.

Smith MT, de la Rosa R, Daniels SI. 2015. Using exposomics to assess cumulative risks and promote health. Environ Mol Mutagen 56:715-723.

Stegemann C, Pechlaner R, Willeit P, Langley SR, Mangino M, Mayr U, et al. 2014. Lipidomics profiling and risk of cardiovascular disease in the prospective population-based bruneck study. Circulation 129:1821-1831.

Stoy N, Mackay GM, Forrest CM, Christofides J, Egerton M, Stone TW, et al. 2005. Tryptophan metabolism and oxidative stress in patients with huntington's

disease. J Neurochem 93:611-623.

Tahara D, Nakanishi T, Akazawa S, Yamaguchi Y, Yamamoto H, Akashi M, et al. 1993. Lecithin-cholesterol acyltransferase and lipid transfer protein activities in liver disease. Metabolism 42:19-23.

Toledo JB, Arnold M, Kastenmuller G, Chang R, Baillie RA, Han X, et al. 2017. Metabolic network failures in alzheimer's disease: A biochemical road map. Alzheimer's & dementia : the journal of the Alzheimer's Association 13:965-984.

Valavanidis A, Vlachogianni T, Fiotakis C. 2009. 8-hydroxy-2′-deoxyguanosine (8-ohdg): A critical biomarker of oxidative stress and carcinogenesis. Journal of Environmental Science and Health, Part C 27:120-139.

Valko M, Morris H, Cronin MT. 2005. Metals, toxicity and oxidative stress. Curr Med Chem 12:1161-1208.

Vineis P, van Veldhoven K, Chadeau-Hyam M, Athersuch TJ. 2013. Advancing the application of omics-based biomarkers in environmental epidemiology. Environ Mol Mutagen 54:461-467.

Wang CW, Liao KW, Chan CC, Yu ML, Chuang HY, Chiang HC, et al. 2019. Association between urinary thiodiglycolic acid level and hepatic function or fibrosis index in school-aged children living near a petrochemical complex. Environ Pollut 244:648-656.

Wang SY, Kuo CH, Tseng YJ. 2015. Ion trace detection algorithm to extract pure ion chromatograms to improve untargeted peak detection quality for liquid chromatography/time-of-flight mass spectrometry-based metabolomics data. Anal Chem 87:3048-3055.

Wenk MR. 2005. The emerging field of lipidomics. Nature reviews Drug discovery 4:594-610.

Wild CP. 2005. Complementing the genome with an "exposome": The outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 14:1847-1850.

Wild CP. 2009. Environmental exposure measurement in cancer epidemiology. Mutagenesis 24:117-125.

Williams PR, Patterson J, Briggs DW. 2006. Vccep pilot: Progress on evaluating children's risks and data needs. Risk Anal 26:781-801.

Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vazquez-Fresno R, et al. 2018. Hmdb 4.0: The human metabolome database for 2018. Nucleic Acids Res 46:D608-D617.

Wu C, Chen ST, Peng KH, Cheng TJ, Wu KY. 2016. Concurrent quantification of multiple biomarkers indicative of oxidative stress status using liquid chromatography-tandem mass spectrometry. Anal Biochem 512:26-35.

Xu G, Luo K, Liu H, Huang T, Fang X, Tu W. 2015. The progress of inflammation and oxidative stress in patients with chronic kidney disease. Renal failure 37:45-49.

Yang L, Li M, Shan Y, Shen S, Bai Y, Liu H. 2016. Recent advances in lipidomics for disease research. J Sep Sci 39:38-50.

Yuan TH, Chio CP, Shie RH, Pien WH, Chan CC. 2015a. The distance-to-source trend in vanadium and arsenic exposures for residents living near a petrochemical complex. J Expo Sci Environ Epidemiol.

Yuan TH, Shie RH, Chin YY, Chan CC. 2015b. Assessment of the levels of urinary 1-hydroxypyrene and air polycyclic aromatic hydrocarbon in pm2.5 for adult exposure to the petrochemical complex emissions. Environ Res 136:219-226.

Yuan TH, Chung MK, Lin CY, Chen ST, Wu KY, Chan CC. 2016. Metabolites of resident exposed to vanadium and pahs in the vicinity of a petrochemical complex. In: Science of the Total Environment.

Yuan TH, Shen YC, Shie RH, Hung SH, Chen CF, Chan CC. 2018. Increased cancers among residents living in the neighborhood of a petrochemical complex: A 12-year retrospective cohort study. International journal of hygiene and environmental health 221:308-314.

Zhao Y-Y, Vaziri ND, Lin R-C. 2015. Chapter six - lipidomics: New insight into kidney disease. In: Advances in clinical chemistry, Vol. 68, (Makowski GS, ed):Elsevier, 153-175.

Zhou L, Ding L, Yin P, Lu X, Wang X, Niu J, et al. 2012a. Serum metabolic profiling study of hepatocellular carcinoma infected with hepatitis b or hepatitis c virus by using liquid chromatography-mass spectrometry. Journal of proteome research 11:5433-5442.

Zhou L, Wang Q, Yin P, Xing W, Wu Z, Chen S, et al. 2012b. Serum metabolomics reveals the deregulation of fatty acids metabolism in hepatocellular carcinoma and chronic liver diseases. Anal Bioanal Chem 403:203-213.

江姿穎. 2015. 鄰近六輕工業區孩童之空氣汙染暴露與過敏性疾病及支氣管炎相關性研究:臺灣大學.

沈育正. 2014. 六輕石化工業區附近成人癌症發生之探討:臺灣大學.

柯昀君. 2017. 重金屬暴露與六輕石化工業區附近幼兒園孩童氧化壓力之相關及飲食中抗氧化物可能的保護效果:臺灣大學.

柯登元. 2016. 六輕石化工業區周界成人居民腎功能與慢性腎臟病之探討:臺灣大學.

孫稚翔. 2017. 鄰近六輕工業區成人血清中重金屬濃度與高血脂症、慢性腎臟病相

關性之研究:臺灣大學.

莊明潔. 2018. 彰化縣大城鄉居民慢性腎臟病與六輕工業區距離相關性之研究:臺灣大學.

陳其欣. 2019. 六輕石化工業區附近居民多重污染暴露與代謝體關係之暴露體學研究:臺灣大學.

陳俊霖. 2018. 六輕工業區北鄰之彰化縣大城鄉居民尿中硫代二乙酸與非侵襲性肝纖維指標的關係:臺灣大學.

詹長權. 2010. 雲林縣沿海地區空氣汙染物及環境健康世代研究計畫期末報告.

詹長權. 2011. 雲林縣沿海地區空氣汙染物及環境健康世代研究計畫期末報告.

詹長權. 2012. 雲林縣沿海地區空氣汙染物及環境健康世代研究計畫期末報告.

詹長權. 2013. 雲林縣沿海地區空氣汙染物及環境健康世代研究計畫期末報告.

劉力瑄. 2014. 以尿液中硫代二乙酸評估雲林縣麥寮鄉六輕工業區附近國小學童氯乙烯單體之暴露:臺灣大學.

謝瑞豪. 2014. 六輕石化工業區營運及意外排放對於周遭社區空氣品質的影響評估:臺灣大學.

謝億廷. 2019. 六輕工業區周圍居民無機砷暴露之研究:臺灣大學.

邊瑋緒. 2011. 六輕離島工業區周界之懸浮微粒及附近居民尿中重金屬濃度之評估研究:臺灣大學.

7. Appendix

7.1 Appendix 1: Urine metabolite profiles in children and elderly participants and the association with multiple exposures and oxidative stress.

HMDB	Metabolite identification	Comp. 1[a]	Comp. 2[a]	Comp. 3[a]	p value[b]	q value[c]	8-OHDG[d]	HNE-MA[d]	8-isoPGF$_{2\alpha}$[d]	8-NO$_2$Gua[d]
Children										
HMDB00001	1-Methylhistidine	0.69	0.90	0.81	0.229	0.295	0.0440	0.0163	0.1127	0.0798
HMDB00005	2-Ketobutyric acid	0.78	0.89	0.84	0.174	0.253	0.0643	0.0083	0.0032	0.6519
HMDB00008	2-Hydroxybutyric acid	0.57	0.60	0.57	0.321	0.356	0.0114	0.0477	0.2722	0.4985
HMDB00017	4-Pyridoxic Acid	1.25	1.11	1.04	0.029	0.086	0.7092	0.8297	0.3291	0.3974
HMDB00019	3-Methyl-2-oxobutanoic acid	0.73	0.67	0.88	0.201	0.275	0.1920	0.0997	0.1654	0.5379
HMDB00020	p-Hydroxyphenylacetic acid	0.09	0.20	0.30	0.871	0.599	0.9495	0.7759	0.4618	0.7956
HMDB00023	(S)-3-Hydroxyisobutyric acid	1.06	0.93	1.19	0.063	0.139	0.4287	0.0018	0.1099	0.0016
HMDB00034	Adenine	0.29	0.34	0.59	0.617	0.515	0.3377	0.0118	0.5669	0.9848
HMDB00073	dopamine	0.21	0.47	0.56	0.717	0.552	0.0107	0.0054	0.1578	0.4685
HMDB00076	Dihydrouracil	0.19	0.18	0.60	0.736	0.559	0.0505	0.0020	0.0070	0.1801
HMDB00087	Dimethylamine	1.23	1.30	1.30	0.031	0.091	0.9582	0.1662	0.8406	0.0335
HMDB00094	Citric acid	0.37	0.35	0.61	0.516	0.470	0.1577	0.1837	0.8283	0.8961
HMDB00112	γ-Aminobutyric acid	1.37	1.64	1.46	0.016	0.060	0.1857	0.1540	0.1716	0.5158
HMDB00115	Glycolic acid	0.43	0.39	0.42	0.458	0.441	0.5068	0.9831	0.5593	0.1402
HMDB00118	Homovanillic acid	0.30	0.37	0.80	0.602	0.508	0.0310	0.0044	0.0436	0.1395
HMDB00123	Glycine	0.31	0.30	0.48	0.586	0.502	0.1258	0.0858	0.3628	0.5183
HMDB00126	Glycerol 3-phosphate	1.61	1.41	1.37	0.004	0.026	0.1783	0.4889	0.1133	0.9593
HMDB00131	Glycerol	0.17	0.52	0.83	0.771	0.570	0.5323	0.1881	0.9915	0.2525
HMDB00134	Fumaric acid	1.39	1.31	1.37	0.014	0.056	0.8539	0.6930	0.2614	0.7829
HMDB00139	Glyceric acid	0.05	0.10	0.99	0.928	0.614	0.0084	0.0001	0.0941	0.0529
HMDB00143	Hexose	0.58	0.60	0.58	0.312	0.349	0.0004	0.0012	0.0254	0.0718
HMDB00148	Glutamic acid	1.06	1.24	1.11	0.063	0.139	0.0768	0.9826	0.8382	0.0003
HMDB00149	Ethanolamine	0.26	0.63	0.94	0.647	0.527	0.7265	0.5501	0.3066	0.0001
HMDB00157	Hypoxanthin	0.20	1.14	1.01	0.724	0.554	0.0190	0.0833	0.6005	0.0172
HMDB00158	DL-Tyrosine	0.90	1.00	0.89	0.116	0.197	0.1751	0.0013	0.1675	0.8839
HMDB00159	L-Phenylalanine	1.29	1.38	1.24	0.024	0.077	0.0199	0.0050	0.2633	0.0575
HMDB00161	Alanine	0.94	0.83	1.11	0.102	0.185	0.0006	0.0001	0.0006	0.0001
HMDB00162	L-Proline	0.64	0.62	0.76	0.265	0.321	0.2509	0.0760	0.1695	0.1085
HMDB00167	L-Threonine	0.49	0.53	0.71	0.399	0.407	0.0952	0.0074	0.0034	0.0365
HMDB00168	L-Asparagine	0.18	1.26	1.16	0.750	0.563	0.1553	0.0421	0.0338	0.0315
HMDB00172	Isoleucine	0.60	1.22	1.12	0.299	0.342	0.0718	0.0653	0.0256	0.0043

I

HMDB	Metabolite identification	Comp. 1[a]	Comp. 2[a]	Comp. 3[a]	p value[b]	q value[c]	8-OHDG[d]	HNE-MA[d]	8-isoPGF$_{2\alpha}$[d]	8-NO$_2$Gua[d]
HMDB00177	L-Histidine	1.30	1.46	1.38	0.022	0.074	0.4554	0.2771	0.7251	0.0002
HMDB00181	L-Dopa	0.83	0.73	0.77	0.149	0.231	0.9420	0.8303	0.3828	0.1601
HMDB00182	Lysine	0.17	1.12	1.01	0.764	0.568	0.0582	0.0006	0.0716	0.0135
HMDB00187	Serine	0.92	0.82	0.93	0.108	0.189	0.1163	0.0204	0.0650	0.4575
HMDB00191	L-Aspartic acid	1.29	1.16	1.08	0.023	0.076	0.0779	0.5078	0.6788	0.0332
HMDB00193	Isocitric acid	0.28	0.36	0.56	0.627	0.518	0.0814	0.8679	0.1454	0.4741
HMDB00197	Indole-3-acetic acid	0.72	0.91	0.97	0.212	0.283	0.0004	0.0008	0.0724	0.4912
HMDB00202	Methylmalonic acid	0.60	0.88	0.81	0.295	0.340	0.0151	0.0162	0.0987	0.4682
HMDB00207	Oleic acid	0.45	0.65	0.82	0.435	0.428	0.1783	0.0555	0.9065	0.3623
HMDB00209	Phenylacetic acid	0.07	0.54	0.53	0.898	0.607	0.0778	0.3924	0.2551	0.1631
HMDB00210	Pantothenic acid	1.17	1.06	1.20	0.041	0.107	0.0119	0.0033	0.0668	0.2908
HMDB00214	DL-Ornithine	0.63	0.57	0.81	0.273	0.326	0.0043	0.0837	0.0145	0.3811
HMDB00220	Palmitic acid	1.14	1.01	1.14	0.046	0.116	0.7072	0.0037	0.0186	0.0004
HMDB00228	Phenol	0.75	0.70	1.05	0.193	0.268	0.0931	0.0008	0.0777	0.0033
HMDB00232	Quinolinic acid	1.47	1.47	1.38	0.009	0.043	0.0092	0.1153	0.0705	0.0227
HMDB00243	Pyruvic acid	0.67	0.88	0.87	0.245	0.307	0.4451	0.1970	0.1975	0.0652
HMDB00254	Succinic acid	1.50	1.31	1.29	0.008	0.040	0.0639	0.0116	0.0117	0.0001
HMDB00262	Thymine	0.87	0.76	1.11	0.131	0.213	0.0120	0.0082	0.1593	0.7355
HMDB00267	Pyroglutamic acid	0.73	0.89	0.79	0.205	0.277	0.1298	0.7975	0.6290	0.2787
HMDB00283	D-Ribose	0.21	0.60	0.56	0.715	0.551	0.2261	0.5603	0.5959	0.7229
HMDB00289	Uric Acid	1.12	1.01	0.94	0.051	0.123	0.3396	0.0213	0.2499	0.8346
HMDB00291	Vanillylmandelic acid	0.19	0.84	0.75	0.748	0.562	0.0760	0.0646	0.2335	0.9453
HMDB00294	Urea	0.61	0.55	0.82	0.289	0.336	0.0015	0.0018	0.0211	0.2406
HMDB00300	Uracil	0.81	0.92	0.82	0.158	0.240	0.2274	0.0364	0.8859	0.6657
HMDB00306	Tyramine	0.92	0.90	0.90	0.108	0.190	0.0137	0.0137	0.1833	0.8889
HMDB00321	2-Hydroxyadipic acid	0.67	0.60	0.79	0.245	0.307	0.0217	0.0016	0.4425	0.9396
HMDB00337	3,4-Dihydroxybutanoic acid	0.36	0.36	0.94	0.527	0.475	0.8925	0.1228	0.8532	0.0537
HMDB00345	3-Hydroxyadipic acid	0.87	0.80	0.73	0.131	0.213	0.0833	0.4990	0.1914	0.5975
HMDB00354	a-Methyl-b-hydroxybutyric acid	0.64	0.68	0.84	0.268	0.322	0.0092	0.0020	0.0142	0.3817
HMDB00355	3-Hydroxymethylglutaric acid	1.19	1.09	1.27	0.037	0.100	0.0194	0.0053	0.0973	0.9696
HMDB00360	3-Deoxytetronic acid	0.54	0.47	0.75	0.351	0.376	0.9609	0.2056	0.2195	0.0005
HMDB00375	m-hydroxy-Hydrocinnamic acid	0.86	0.97	0.88	0.135	0.217	0.0718	0.0001	0.0524	0.2419
HMDB00393	(E)-3-Hexenedioic acid	1.19	1.05	0.99	0.037	0.100	0.9362	0.0322	0.0182	0.0134
HMDB00396	2-(hydroxymethyl)butanoic acid	0.22	0.37	0.85	0.698	0.545	0.0331	0.0008	0.0329	0.1269
HMDB00407	2-hydroxyisovaleric acid	0.55	0.53	0.84	0.339	0.368	0.7225	0.7318	0.4729	0.2492
HMDB00426	Citramalic acid	0.39	0.39	0.91	0.496	0.460	0.0003	0.0004	0.0106	0.2283

II

HMDB	Metabolite identification	Comp. 1[a]	Comp. 2[a]	Comp. 3[a]	p value[b]	q value[c]	8-OHDG[d]	HNE-MA[d]	8-isoPGF$_{2a}$[d]	8-NO$_2$Gua[d]
HMDB00439	2-Furoylglycine	0.05	0.17	0.26	0.930	0.615	0.0663	0.0027	0.5899	0.9080
HMDB00440	3-Hydroxyphenylacetic acid	0.03	0.34	0.45	0.958	0.622	0.7121	0.0821	0.0998	0.5005
HMDB00448	Adipic acid	0.53	0.81	0.82	0.359	0.381	0.0271	0.0257	0.0240	0.5811
HMDB00452	L-Alpha-aminobutyric acid	1.59	1.78	1.58	0.005	0.028	0.3788	0.0696	0.0062	0.4033
HMDB00479	3-Methylhistidine	0.67	0.63	0.61	0.242	0.305	0.8421	0.1122	0.0316	0.4721
HMDB00491	a-Oxo-b-methylvaleric acid	0.42	0.37	0.51	0.466	0.445	0.2321	0.0065	0.0025	0.8995
HMDB00498	4-Deoxyerythronic acid	1.34	1.18	1.21	0.018	0.065	0.0244	0.0001	0.0028	0.0008
HMDB00500	4-Hydroxybenzoic acid	1.25	1.17	1.07	0.028	0.086	0.7414	0.0114	0.1052	0.1230
HMDB00522	3-Methylglutaconic acid	0.93	1.11	1.02	0.103	0.186	0.0957	0.1231	0.0080	0.0590
HMDB00525	5-Hydroxyhexanoic acid	0.29	0.43	0.61	0.617	0.514	0.0396	0.0168	0.0968	0.2256
HMDB00539	Pentonic acid	0.30	0.26	0.65	0.602	0.508	0.2073	0.2004	0.0355	0.0670
HMDB00555	B-methyladipic acid	0.05	0.46	0.41	0.928	0.615	0.3892	0.0258	0.2011	0.6693
HMDB00568	L-Arabitol	0.22	0.65	0.60	0.698	0.545	0.3610	0.2673	0.0151	0.5636
HMDB00574	L-Cysteine	0.06	0.89	0.79	0.923	0.613	0.4249	0.1854	0.0585	0.8218
HMDB00613	Erythronic acid	0.25	0.72	0.76	0.670	0.535	0.4436	0.9609	0.4253	0.6431
HMDB00622	Ethylmalonic acid	0.04	0.30	0.78	0.946	0.619	0.0302	0.0001	0.0307	0.0628
HMDB00625	Gluconic acid	1.48	1.37	1.22	0.009	0.043	0.6666	0.0799	0.2655	0.0280
HMDB00630	Cytosine	0.13	0.13	0.32	0.822	0.585	0.0121	0.4637	0.2241	0.7450
HMDB00638	Dodecanoic acid	0.53	0.80	0.72	0.354	0.378	0.4392	0.0838	0.7246	0.5010
HMDB00639	Galactaric acid	0.77	1.07	0.98	0.181	0.259	0.0455	0.0369	0.4514	0.1356
HMDB00640	Levoglucosan	0.71	0.74	0.67	0.219	0.288	0.1238	0.4841	0.3785	0.5002
HMDB00661	Glutaric acid	0.15	0.54	0.89	0.789	0.576	0.0238	0.0033	0.0156	0.3870
HMDB00663	Glucaric acid	0.25	0.33	0.33	0.662	0.532	0.6329	0.3982	0.2553	0.0577
HMDB00687	Leucine	0.48	0.77	0.69	0.407	0.411	0.0085	0.0001	0.0003	0.8341
HMDB00691	Malonic acid	0.06	0.51	0.58	0.915	0.611	0.0761	0.0080	0.7159	0.2199
HMDB00694	2-Hydroxyglutaric acid	0.02	0.23	0.56	0.976	0.626	0.7884	0.0160	0.8872	0.8361
HMDB00696	L-Methionine	0.07	0.98	0.87	0.898	0.607	0.1048	0.2215	0.1979	0.0287
HMDB00700	Hydroxypropionic acid	0.55	0.49	0.45	0.335	0.365	0.0836	0.0215	0.2592	0.0354
HMDB00710	γ-Hydroxybutyric acid	0.25	0.34	0.55	0.660	0.531	0.0032	0.0533	0.0344	0.5773
HMDB00714	Hippuric acid	1.40	1.26	1.26	0.014	0.055	0.2672	0.0022	0.4725	0.0001
HMDB00715	Kynurenic acid	0.67	1.03	1.00	0.242	0.305	0.0518	0.0390	0.0050	0.0843
HMDB00719	Homoserine	0.07	0.68	0.74	0.897	0.607	0.3593	0.8113	0.0088	0.0756
HMDB00725	4-Hydroxyproline	0.82	0.74	0.88	0.154	0.235	0.1314	0.0165	0.0436	0.3403
HMDB00729	2-hydroxyisobutyrate	0.98	0.87	1.05	0.086	0.168	0.3185	0.0029	0.4410	0.1579
HMDB00744	Malic Acid	0.35	0.32	0.68	0.545	0.483	0.2363	0.0328	0.0514	0.3730
HMDB00749	Mesaconic acid	0.66	0.62	1.16	0.251	0.311	0.2923	0.0724	0.0660	0.9855

III

HMDB	Metabolite identification	Comp. 1[a]	Comp. 2[a]	Comp. 3[a]	p value[b]	q value[c]	8-OHDG[d]	HNE-MA[d]	8-isoPGF$_{2\alpha}$[d]	8-NO$_2$Gua[d]
HMDB00750	3-hydroxymandelic acid	0.96	0.86	1.14	0.093	0.175	0.3117	0.0635	0.0181	0.8415
HMDB00752	3-Methylglutaric acid	0.59	0.84	0.95	0.304	0.345	0.0151	0.0059	0.0417	0.6094
HMDB00754	3-Hydroxyisovaleric acid	0.24	0.86	1.22	0.674	0.537	0.9975	0.1292	0.8230	0.3429
HMDB00755	Hydroxyphenyllactic acid	1.39	1.30	1.27	0.015	0.057	0.0006	0.0038	0.2942	0.2202
HMDB00763	5-Hydroxyindoleacetate	0.16	0.91	0.82	0.787	0.575	0.0003	0.0188	0.0009	0.6347
HMDB00784	Azelaic acid	1.30	1.18	1.09	0.022	0.074	0.2172	0.0014	0.0122	0.0506
HMDB00806	Myristic acid	1.64	1.43	1.45	0.004	0.024	0.6178	0.1222	0.8404	0.0020
HMDB00827	Stearic acid	1.71	1.50	1.51	0.002	0.017	0.5580	0.0176	0.1446	0.0001
HMDB00847	Nonanoic acid	0.38	0.35	0.65	0.513	0.468	0.7702	0.0190	0.7709	0.0147
HMDB00849	Rhamnose	1.77	1.68	1.50	0.002	0.013	0.5008	0.0002	0.0003	0.1051
HMDB00857	Pimelic acid	0.25	0.77	0.74	0.662	0.532	0.0620	0.0001	0.0035	0.7340
HMDB00863	Isopropanol	1.04	1.07	1.23	0.070	0.149	0.0891	0.2632	0.3955	0.1026
HMDB00870	Histamine	0.74	0.80	0.80	0.196	0.271	0.3309	0.4760	0.0857	0.6366
HMDB00873	4-methylcatechol	0.74	1.08	0.98	0.199	0.273	0.0350	0.0502	0.0104	0.7782
HMDB00881	Xanthurenic acid	0.98	1.15	1.09	0.088	0.170	0.2789	0.5305	0.2253	0.0401
HMDB00883	L-Valine	0.60	0.53	0.62	0.298	0.341	0.1026	0.0866	0.0304	0.0277
HMDB00893	Suberic acid	0.84	1.14	1.02	0.144	0.226	0.0672	0.0005	0.0013	0.0669
HMDB00904	Citrulline	0.73	1.04	0.93	0.205	0.277	0.8052	0.2061	0.0047	0.1108
HMDB00929	DL-Tryptophan	1.21	1.15	1.18	0.033	0.094	0.0403	0.0056	0.0045	0.0579
HMDB00943	Threonic acid	1.51	1.32	1.49	0.008	0.038	0.3774	0.0003	0.0684	0.0015
HMDB00956	(±)-Tartaric acid	0.14	0.13	0.46	0.809	0.582	0.3570	0.0023	0.0677	0.2412
HMDB00957	Catechol	0.93	0.96	0.89	0.104	0.186	0.2317	0.0001	0.1553	0.3635
HMDB00958	trans-Aconitic acid	0.61	0.54	0.52	0.288	0.336	0.0950	0.1974	0.0512	0.4480
HMDB00959	Tiglylglycine	1.45	1.27	1.15	0.011	0.047	0.0070	0.5525	0.9227	0.5987
HMDB01051	Glyceraldehyde	0.93	0.84	1.16	0.105	0.187	0.1011	0.1952	0.0448	0.6522
HMDB01123	Anthranilic acid	0.64	0.64	0.76	0.263	0.319	0.0927	0.0049	0.9862	0.5410
HMDB01138	N-acetylglutamic acid	1.30	1.30	1.27	0.023	0.075	0.0058	0.0027	0.0069	0.3324
HMDB01147	Aminomalonic acid	1.22	1.16	1.12	0.032	0.092	0.1244	0.0039	0.0906	0.1012
HMDB01266	L-Sorbose	1.35	1.30	1.18	0.017	0.063	0.3007	0.0255	0.7336	0.6782
HMDB01352	Hydroxypyruvic acid	1.24	1.10	1.10	0.030	0.089	0.3353	0.2338	0.9247	0.3273
HMDB01398	Guaiacol	0.12	0.62	0.62	0.840	0.591	0.5525	0.1124	0.5061	0.2912
HMDB01414	putrescine	0.35	0.64	0.57	0.547	0.484	0.0454	0.0559	0.7566	0.3257
HMDB01470	Tiglic acid	2.10	1.88	1.73	0.000	0.003	0.4229	0.0024	0.4547	0.0209
HMDB01476	3-Hydroxyanthranilic acid	0.24	0.27	0.29	0.675	0.537	0.0917	0.1476	0.4709	0.6269
HMDB01488	Niacin	0.74	0.67	0.93	0.199	0.273	0.1357	0.0099	0.0824	0.2741
HMDB01490	Vanylglycol	0.93	0.89	0.82	0.106	0.188	0.4773	0.3149	0.1774	0.6565

IV

HMDB	Metabolite identification	Comp. 1[a]	Comp. 2[a]	Comp. 3[a]	p value[b]	q value[c]	8-OHDG[d]	HNE-MA[d]	8-isoPGF$_{2\alpha}$[d]	8-NO$_2$Gua[d]
HMDB01568	2-Octenoic acid	0.04	0.20	0.62	0.946	0.619	0.0899	0.0244	0.5234	0.1824
HMDB01624	2-Hydroxycaproic acid	0.40	0.40	0.87	0.490	0.457	0.3948	0.2284	0.4148	0.1826
HMDB01644	D-Xylulose	0.71	0.78	0.71	0.214	0.284	0.3024	0.7665	0.3548	0.0981
HMDB01870	Benzoic acid	0.37	0.53	1.20	0.523	0.473	0.5032	0.0097	0.8861	0.0001
HMDB01881	(±)-1,2-propanediol	0.99	0.98	0.88	0.085	0.166	0.0945	0.8404	0.0459	0.5980
HMDB01882	Dihydroxyacetone	0.97	0.88	0.89	0.089	0.172	0.5758	0.4433	0.9946	0.0385
HMDB01954	3-Hydroxyoctanoic acid	1.58	1.44	1.40	0.005	0.030	0.6057	0.4830	0.7270	0.1981
HMDB01988	4-Hydroxycyclohexanecarboxylic acid	1.00	1.23	1.10	0.082	0.163	0.0339	0.0014	0.0950	0.1962
HMDB02039	2-Pyrrolidinone	1.08	1.07	0.99	0.059	0.134	0.1725	0.0396	0.0977	0.4809
HMDB02048	3-Methylphenol	0.65	0.57	0.59	0.257	0.315	0.0267	0.0250	0.2141	0.0106
HMDB02078	Cyanic acid	1.81	1.60	1.44	0.001	0.011	0.6682	0.3236	0.3152	0.0577
HMDB02142	Phosphoric acid	0.90	0.79	0.86	0.116	0.197	0.8642	0.1482	0.2721	0.0895
HMDB02199	Desaminotyrosine	0.47	0.83	0.84	0.414	0.416	0.2519	0.0203	0.0001	0.1492
HMDB02243	Picolinic acid	0.25	0.26	0.99	0.669	0.535	0.0009	0.0001	0.0119	0.2248
HMDB02329	Oxalic acid	0.37	0.62	0.56	0.515	0.470	0.1755	0.2804	0.7028	0.4792
HMDB02432	Sumiki's acid	2.48	2.24	2.08	0.000	0.000	0.6420	0.6038	0.6635	0.0112
HMDB02434	Hydroquinone	0.06	0.35	0.54	0.917	0.612	0.1063	0.0663	0.1019	0.1106
HMDB02643	3-(3-Hydroxyphenyl)-3-hydroxypropionic acid	0.07	0.22	0.67	0.907	0.609	0.4256	0.0786	0.0400	0.0428
HMDB02649	Erythrose	0.14	0.39	0.44	0.812	0.583	0.7474	0.2271	0.0002	0.2563
HMDB02712	1,5-Anhydrosorbitol	0.84	0.73	0.75	0.144	0.226	0.2706	0.6744	0.3467	0.1182
HMDB03070	Shikimic acid	0.93	0.81	0.75	0.104	0.186	0.3140	0.1341	0.4608	0.9940
HMDB03156	2,3-Butanediol	0.64	0.57	0.61	0.264	0.320	0.7369	0.6161	0.9880	0.9961
HMDB03219	Hept-2-ulose	0.12	0.59	0.52	0.835	0.589	0.4438	0.1842	0.0137	0.6108
HMDB03243	Acetoin	0.67	0.63	0.63	0.243	0.305	0.8184	0.9876	0.6718	0.6420
HMDB03315	Cyclohexanone	2.47	2.17	1.95	0.000	0.000	0.3576	0.0230	0.5663	0.0016
HMDB03466	L-Gulonolactone	1.85	1.63	1.51	0.001	0.009	0.0205	0.2218	0.3242	0.2336
HMDB03903	2-Hydroxyethanesulfonate	1.14	1.04	0.94	0.046	0.116	0.0253	0.1590	0.2821	0.0001
HMDB03911	3-Aminoisobutanoic acid	1.13	1.07	0.96	0.048	0.118	0.0583	0.0063	0.0185	0.0516
HMDB04136	D-Threitol	1.90	1.78	1.59	0.001	0.008	0.4647	0.0001	0.0180	0.1679
HMDB04230	Pyrrole-2-carboxylic acid	0.07	0.34	0.52	0.902	0.608	0.0254	0.0130	0.0708	0.5859
HMDB04437	Diethanolamine	1.07	1.01	0.92	0.062	0.137	0.7468	0.4748	0.6296	0.3304
HMDB04812	2,5-Furandicarboxylic acid	0.21	0.46	0.47	0.720	0.553	0.7220	0.1982	0.1434	0.1672
HMDB04983	Dimethyl sulfone	0.47	0.69	1.03	0.417	0.417	0.0690	0.0302	0.2881	0.0498
HMDB05802	Trans-isoeugenol	0.37	1.33	1.31	0.525	0.474	0.2057	0.3179	0.1778	0.0035
HMDB06116	3-Hydroxyhippuric acid	1.54	1.39	1.24	0.006	0.034	0.1413	0.7647	0.0365	0.7520

V

HMDB	Metabolite identification	Comp. 1[a]	Comp. 2[a]	Comp. 3[a]	p value[b]	q value[c]	8-OHDG[d]	HNE-MA[d]	8-isoPGF$_{2\alpha}$[d]	8-NO$_2$Gua[d]
HMDB06899	Alanylglycine	0.01	0.19	0.28	0.991	0.630	0.5757	0.8154	0.6383	0.1249
HMDB10719	trans-Hex-2-enoic acid	0.47	0.41	0.60	0.419	0.419	0.8913	0.2402	0.3737	0.0728
HMDB11676	3,4,5-Trihydroxytetrahydro-2H-pyran-2-one	0.44	0.89	0.80	0.448	0.435	0.5023	0.0043	0.0687	0.9361
HMDB11732	Hex-2-ulosonic acid	1.50	1.36	1.26	0.008	0.040	0.3151	0.6691	0.1114	0.5028
HMDB13231	Ethylamine	0.75	0.66	0.72	0.194	0.269	0.2593	0.1907	0.0480	0.5094
HMDB13248	Mono(2-ethylhexyl)phthalate [MEHP]	1.09	0.99	1.33	0.057	0.132	0.1244	0.0025	0.0030	0.0065
HMDB13674	Pyrogallol	0.18	0.47	0.43	0.751	0.564	0.7921	0.4113	0.0131	0.0843
HMDB13716	Norvaline	0.49	0.67	0.60	0.393	0.403	0.1631	0.2940	0.0051	0.4414
HMDB14328	Amphetamine	0.34	0.91	0.86	0.561	0.491	0.0155	0.0827	0.0201	0.1257
HMDB29635	4-Methylbenzoic acid	0.79	0.74	0.79	0.170	0.249	0.6427	0.0500	0.6794	0.8868
HMDB29739	1H-Indole-3-acetamide	2.04	1.78	1.58	0.000	0.004	0.5837	0.6450	0.7874	0.0066
HMDB29942	Pentose	0.37	0.36	0.79	0.519	0.471	0.1930	0.0513	0.1132	0.3322
HMDB31213	2-Ethoxyethanol	0.81	1.03	0.92	0.157	0.238	0.2194	0.3289	0.6778	0.1500
HMDB31320	1,3-butanediol	0.37	0.41	0.50	0.518	0.471	0.6085	0.0474	0.0719	0.0239
HMDB31404	Cyclohexylamine	0.02	0.18	0.43	0.977	0.627	0.3575	0.5249	0.2950	0.5059
HMDB31444	Dodecane	1.36	1.21	1.37	0.017	0.062	0.2155	0.0019	0.8948	0.0001
HMDB31445	Undecane	0.99	0.88	0.89	0.085	0.167	0.7239	0.0208	0.2802	0.0003
HMDB31450	Decane	1.54	1.35	1.37	0.007	0.035	0.0651	0.0082	0.4617	0.0003
HMDB31511	Diacetone alcohol	3.42	3.06	2.71	0.000	0.000	0.4636	0.0072	0.0427	0.0031
HMDB31602	4-Pentenoic acid	0.59	0.61	0.62	0.307	0.347	0.8826	0.4206	0.2104	0.4134
HMDB31626	2-phenylpropanal	1.41	1.24	1.38	0.013	0.053	0.3094	0.0366	0.1853	0.0183
HMDB32037	1,3-Benzenediol	1.12	0.99	0.94	0.051	0.123	0.4365	0.1662	0.5088	0.0483
HMDB32619	(±)-1-phenylethanol	0.12	0.79	0.71	0.829	0.588	0.4221	0.0524	0.0029	0.9497
HMDB33244	Dibutyl phthalate	1.07	1.13	1.01	0.061	0.136	0.4438	0.4441	0.2783	0.6122
HMDB33958	2-Deoxy-L-ribono-1,4-lactone	1.85	1.69	1.52	0.001	0.009	0.0118	0.5731	0.0767	0.2674
HMDB34220	Inositol	0.82	0.74	1.26	0.154	0.236	0.0052	0.0001	0.0446	0.0001
HMDB34284	Tridecane	1.59	1.40	1.35	0.005	0.029	0.0075	0.0014	0.9111	0.0315
HMDB34778	(2R*,3R*)-1,2,3-Butanetriol	0.37	0.33	0.53	0.521	0.473	0.2073	0.1011	0.0615	0.2719
HMDB34976	Borneol	0.38	1.03	0.92	0.509	0.467	0.1093	0.0821	0.0569	0.1180
HMDB35056	1-(3-Hydroxy-4-methoxyphenyl)-1,2-ethanediol	0.58	0.53	0.60	0.310	0.348	0.8048	0.4010	0.3802	0.3733
HMDB35227	Tartronic acid	0.27	0.58	0.53	0.645	0.526	0.0220	0.0002	0.3718	0.6764
HMDB37050	o-Cymene	2.25	1.97	2.01	0.000	0.001	0.0862	0.0001	0.1274	0.0001
HMDB41486	1-Deoxy-D-ribitol	0.35	0.36	0.55	0.541	0.482	0.6410	0.3979	0.1194	0.6476
HMDB41861	Cyanuric acid	2.15	1.89	1.68	0.000	0.002	0.4612	0.0995	0.0338	0.0167
HMDB41932	Methylephedrine	0.40	0.35	0.54	0.491	0.458	0.0904	0.1314	0.2761	0.5704

HMDB	Metabolite identification	Comp. 1[a]	Comp. 2[a]	Comp. 3[a]	p value[b]	q value[c]	8-OHDG[d]	HNE-MA[d]	8-isoPGF2a[d]	8-NO2Gua[d]
HMDB42032	Thiodiacetic acid	0.76	0.77	1.16	0.183	0.261	0.5880	0.0008	0.6909	0.0027
HMDB59740	4-Methyl-1,2-pentanediol	1.05	0.99	0.91	0.066	0.143	0.0290	0.0593	0.1334	0.2426
HMDB59873	3-Ethylphenol	0.24	0.43	0.69	0.676	0.538	0.3726	0.1300	0.4243	0.1118
HMDB59889	Isohexanol	0.25	0.65	0.60	0.666	0.534	0.7032	0.2639	0.3867	0.1875
HMDB59901	Hemimellitene	1.01	0.88	0.90	0.079	0.160	0.8022	0.2888	0.1880	0.2892
HMDB61927	5-Hydroxypentanoic acid	0.12	0.23	0.63	0.831	0.588	0.3040	0.0407	0.1920	0.8063
HMDB61933	1H-Indole-2,3-dione	0.62	0.66	0.66	0.283	0.332	0.8686	0.1735	0.1107	0.2638

Elderly

HMDB	Metabolite identification	Comp. 1[a]	Comp. 2[a]	Comp. 3[a]	p value[b]	q value[c]	8-OHDG[d]	HNE-MA[d]	8-isoPGF2a[d]	8-NO2Gua[d]
HMDB02712	1,5-Anhydrosorbitol	0.79	1.02	1.02	0.274	0.318	0.6182	0.7174	0.4820	0.1918
HMDB00001	1-Methylhistidine	1.10	0.89	0.96	0.119	0.215	0.8247	0.0034	0.0032	0.3412
HMDB00005	2-Ketobutyric acid	0.48	0.54	0.59	0.148	0.233	0.2159	0.0114	0.0063	0.6355
HMDB00008	2-Hydroxybutyric acid	0.35	0.46	0.43	0.307	0.338	0.8730	0.0067	0.8935	0.2809
HMDB00017	4-Pyridoxic Acid	0.17	0.71	0.85	0.361	0.367	0.7156	0.3355	0.7537	0.4912
HMDB00019	Alpha-ketoisovaleric acid	1.15	1.14	1.11	0.077	0.184	0.6481	0.8662	0.8770	0.3414
HMDB00020	p-Hydroxyphenylacetic acid	1.13	0.89	0.74	0.284	0.324	0.0711	0.0006	0.1474	0.3868
HMDB00023	(S)-3-Hydroxyisobutyric acid	1.60	1.40	1.25	0.162	0.242	0.0009	0.0307	0.1150	0.0256
HMDB00034	Adenine	0.90	0.70	0.92	0.191	0.260	0.1757	0.0123	0.1135	0.6974
HMDB00073	dopamine	0.09	0.56	0.46	0.059	0.165	0.0977	0.0349	0.1377	0.7260
HMDB00076	Dihydrouracil	1.04	0.93	0.73	0.027	0.118	0.1190	0.2560	0.6192	0.0879
HMDB00087	Dimethylamine	0.89	0.68	0.54	0.889	0.553	0.1898	0.4215	0.1821	0.8437
HMDB00094	Citric acid	0.55	1.25	1.48	0.086	0.192	0.0123	0.1210	0.0317	0.0272
HMDB00115	Glycolic acid	0.86	0.70	0.65	0.070	0.178	0.0562	0.0662	0.7560	0.2809
HMDB00118	Homovanillic acid	0.18	1.20	0.98	0.320	0.346	0.1645	0.1066	0.0285	0.8256
HMDB00123	Glycine	1.01	0.73	0.66	0.686	0.488	0.4619	0.0001	0.1802	0.2830
HMDB00126	Glycerol 3-phosphate	0.90	0.96	0.76	0.665	0.482	0.0285	0.0019	0.2009	0.0764
HMDB00131	Glycerol	0.46	0.59	0.46	0.371	0.372	0.6011	0.0163	0.1432	0.6561
HMDB00134	Fumaric acid	0.82	0.72	1.26	0.400	0.385	0.9472	0.0009	0.2056	0.5292
HMDB00139	Glyceric acid	2.00	1.52	1.19	0.029	0.122	0.0756	0.0040	0.0545	0.0425
HMDB00143	Hexose	0.30	0.40	0.35	0.743	0.508	0.4979	0.0833	0.1686	0.1751
HMDB00148	Glutamic acid	0.67	0.74	1.04	0.238	0.294	0.8426	0.0124	0.0113	0.8274
HMDB00149	Ethanolamine	0.64	1.16	1.64	0.219	0.281	0.0064	0.5809	0.5264	0.0043
HMDB00157	Hypoxanthin	1.38	1.39	1.27	0.483	0.419	0.4215	0.6340	0.2227	0.0037
HMDB00158	DL-Tyrosine	0.56	0.62	0.81	0.094	0.198	0.2288	0.0064	0.0367	0.7003
HMDB00159	Phenylalanine	1.17	1.38	1.08	0.499	0.426	0.9325	0.0717	0.2180	0.2623
HMDB00161	Alanine	1.19	1.39	1.35	0.178	0.253	0.1315	0.0039	0.5709	0.0021
HMDB00162	L-Proline	0.76	0.58	1.12	0.470	0.414	0.9275	0.9278	0.0039	0.6532

VII

HMDB	Metabolite identification	Comp. 1[a]	Comp. 2[a]	Comp. 3[a]	p value[b]	q value[c]	8-OHDG[d]	HNE-MA[d]	8-isoPGF$_{2a}$[d]	8-NO$_2$Gua[d]
HMDB00167	L-Threonine	1.47	1.24	1.23	0.713	0.498	0.2666	0.1126	0.7332	0.0001
HMDB00168	L-Asparagine	1.22	1.55	1.60	0.169	0.247	0.7446	0.0669	0.8659	0.5794
HMDB00172	Isoleucine	1.01	1.11	1.42	0.098	0.201	0.8423	0.0245	0.0015	0.7229
HMDB00177	L-Histidine	1.03	0.77	1.08	0.841	0.539	0.3380	0.0156	0.0127	0.2477
HMDB00182	Lysine	0.49	0.47	0.71	0.512	0.430	0.8799	0.1963	0.0936	0.5330
HMDB00187	Serine	2.21	1.55	1.23	0.536	0.439	0.0008	0.0184	0.1950	0.0001
HMDB00191	L-Aspartic acid	1.39	1.27	1.30	0.751	0.511	0.4233	0.0199	0.0722	0.3050
HMDB00193	Isocitric acid	0.92	0.89	0.78	0.035	0.133	0.5489	0.3137	0.4277	0.2156
HMDB00197	Indole-3-acetic acid	0.54	1.04	0.81	0.103	0.204	0.3160	0.0032	0.1050	0.6093
HMDB00202	Methylmalonic acid	1.09	0.99	0.91	0.597	0.460	0.0468	0.0007	0.0367	0.7975
HMDB00207	Oleic acid	1.09	0.78	0.66	0.057	0.162	0.3464	0.3084	0.0009	0.1330
HMDB00209	Phenylacetic acid	1.11	0.83	0.69	0.189	0.259	0.1318	0.0057	0.1005	0.0406
HMDB00210	Pantothenic acid	0.97	1.04	0.94	0.576	0.453	0.4603	0.0611	0.2308	0.6218
HMDB00214	DL-Ornithine	0.95	0.96	0.76	0.097	0.200	0.0946	0.0001	0.0048	0.0452
HMDB00220	Palmitic acid	1.98	1.62	1.42	0.433	0.399	0.2270	0.2854	0.2632	0.0563
HMDB00228	Phenol	1.94	1.62	1.28	0.574	0.453	0.0276	0.0321	0.3199	0.0001
HMDB00232	Quinolinic acid	0.34	1.03	1.25	0.540	0.440	0.5849	0.0041	0.0172	0.8265
HMDB00243	Pyruvic acid	0.89	1.13	0.92	0.001	0.021	0.0684	0.0311	0.0322	0.0003
HMDB00254	Succinic acid	1.30	1.05	0.99	0.231	0.289	0.1925	0.0090	0.3105	0.3223
HMDB00262	Thymine	0.98	1.02	1.16	0.079	0.186	0.0234	0.0008	0.0009	0.1630
HMDB00267	Pyroglutamic acid	0.10	0.10	0.92	0.052	0.156	0.7392	0.1115	0.5222	0.4735
HMDB00283	D-Ribose	0.20	0.21	0.76	0.649	0.477	0.9580	0.0153	0.1168	0.1307
HMDB00291	Vanillylmandelic acid	1.00	1.00	1.21	0.070	0.177	0.6950	0.3219	0.9840	0.9744
HMDB00294	Urea	0.60	1.59	1.39	0.710	0.497	0.0770	0.0010	0.0024	0.9817
HMDB00300	Uracil	1.50	1.09	0.91	0.088	0.194	0.1557	0.0420	0.1898	0.0730
HMDB00306	Tyramine	0.62	0.67	0.96	0.430	0.398	0.6558	0.0062	0.2630	0.9065
HMDB00321	2-Hydroxyadipic acid	0.94	0.70	1.26	0.003	0.029	0.1839	0.0001	0.0061	0.1597
HMDB00337	3,4-Dihydroxybutanoic acid	0.99	1.46	1.32	0.628	0.470	0.0014	0.0006	0.0259	0.0043
HMDB00345	3-Hydroxyadipic acid	0.59	0.92	0.82	0.185	0.257	0.5512	0.1742	0.2444	0.4895
HMDB00354	a-Methyl-b-hydroxybutyric acid	0.36	0.89	0.71	0.184	0.256	0.2027	0.0002	0.0509	0.7253
HMDB00355	3-Hydroxymethylglutaric acid	0.91	0.94	1.01	0.143	0.231	0.0844	0.0013	0.1344	0.9682
HMDB00360	2,4-Dihydroxybutanoic acid	1.70	1.33	1.05	0.039	0.138	0.0690	0.0002	0.0543	0.0898
HMDB00375	m-hydroxy-Hydrocinnamic acid	0.48	0.93	0.75	0.176	0.251	0.7618	0.0310	0.1751	0.4624
HMDB00393	(E)-3-Hexenedioic acid	0.93	0.74	0.58	0.163	0.243	0.6711	0.1484	0.3376	0.4313
HMDB00396	2-Ethylhydracrylic acid	2.11	1.51	1.31	0.214	0.277	0.3003	0.0001	0.0638	0.0981
HMDB00407	2-Hydroxy-3-methylbutyric acid	1.28	0.90	0.90	0.788	0.523	0.6070	0.3382	0.1752	0.2080

VIII

HMDB	Metabolite identification	Comp. 1[a]	Comp. 2[a]	Comp. 3[a]	p value[b]	q value[c]	8-OHDG[d]	HNE-MA[d]	8-isoPGF$_{2a}$[d]	8-NO$_2$Gua[d]
HMDB00426	Citramalic acid	0.28	1.41	1.12	0.491	0.423	0.1454	0.0044	0.0502	0.1909
HMDB00439	2-Furoylglycine	0.01	0.62	0.56	0.430	0.398	0.4129	0.7058	0.9346	0.2920
HMDB00440	3-Hydroxyphenylacetic acid	0.53	0.74	0.67	0.011	0.074	0.8772	0.3609	0.4890	0.1614
HMDB00448	Adipic acid	0.69	0.79	1.01	0.053	0.157	0.5674	0.0916	0.0808	0.4045
HMDB00452	L-Alpha-aminobutyric acid	1.05	0.75	0.79	0.417	0.392	0.0599	0.0001	0.0001	0.0226
HMDB00479	3-Methylhistidine	0.22	0.56	0.46	0.346	0.360	0.1746	0.0772	0.0146	0.8489
HMDB00491	a-Oxo-b-methylvaleric acid	0.63	0.76	0.70	0.388	0.380	0.0786	0.2004	0.5202	0.1418
HMDB00498	4-Deoxyerythronic acid	1.45	1.55	1.54	0.104	0.204	0.0512	0.0063	0.3277	0.0100
HMDB00522	3-Methylglutaconic acid	0.53	1.46	1.15	0.422	0.395	0.7721	0.0256	0.1246	0.3039
HMDB00525	5-Hydroxyhexanoic acid	1.29	0.90	1.50	0.406	0.388	0.7569	0.5628	0.4757	0.9073
HMDB00539	Arabinonic acid	1.20	1.05	0.88	0.000	0.012	0.5187	0.1849	0.2093	0.5542
HMDB00555	B-methyladipic acid	0.28	0.53	0.44	0.189	0.259	0.8033	0.0912	0.9028	0.0721
HMDB00568	L-Arabitol	0.83	0.59	0.47	0.749	0.510	0.4990	0.6823	0.6695	0.5299
HMDB00574	L-Cysteine	0.61	0.87	1.18	0.182	0.255	0.4669	0.0577	0.0210	0.1544
HMDB00613	Erythronic acid	0.19	0.34	1.17	0.678	0.486	0.9376	0.4881	0.3588	0.0593
HMDB00617	2-Furoic acid	0.22	0.98	0.77	0.000	0.012	0.9703	0.1505	0.0303	0.4797
HMDB00622	Ethylmalonic acid	1.12	1.09	1.05	0.001	0.019	0.4414	0.0354	0.1769	0.8848
HMDB00625	Gluconic acid	0.28	0.23	1.09	0.049	0.152	0.5178	0.1182	0.0980	0.0983
HMDB00630	Cytosine	1.44	1.12	1.04	0.347	0.360	0.0199	0.0001	0.0021	0.0626
HMDB00638	Dodecanoic acid	0.42	0.88	0.72	0.382	0.377	0.8705	0.5933	0.5287	0.3826
HMDB00639	Galactaric acid	0.95	1.09	0.87	0.369	0.371	0.3643	0.3197	0.0950	0.3085
HMDB00640	Levoglucosan	0.36	0.80	0.64	0.383	0.378	0.5014	0.2806	0.0483	0.2123
HMDB00661	Glutaric acid	2.11	1.42	1.31	0.150	0.235	0.1154	0.0001	0.1274	0.0001
HMDB00687	Leucine	1.53	1.08	1.29	0.648	0.477	0.2004	0.0026	0.0062	0.0501
HMDB00691	Malonic acid	0.83	0.59	0.54	0.061	0.167	0.1170	0.0538	0.1258	0.9130
HMDB00694	2-Hydroxyglutaric acid	2.45	1.81	1.60	0.023	0.111	0.0242	0.0001	0.4960	0.0182
HMDB00696	L-Methionine	0.53	0.43	0.82	0.969	0.574	0.4514	0.0005	0.0149	0.4297
HMDB00700	Hydroxypropionic acid	0.81	0.91	0.71	0.842	0.539	0.9811	0.6882	0.4692	0.0854
HMDB00710	γ-Hydroxybutyric acid	0.04	1.26	1.00	0.943	0.567	0.8160	0.4047	0.6418	0.5773
HMDB00714	Hippuric acid	0.84	1.15	1.27	0.656	0.479	0.5217	0.0015	0.8150	0.0284
HMDB00715	Kynurenic acid	0.31	0.76	0.60	0.949	0.569	0.6291	0.2579	0.2523	0.3431
HMDB00719	Homoserine	0.85	1.14	0.90	0.677	0.486	0.3386	0.0302	0.0837	0.2451
HMDB00725	4-Hydroxyproline	0.67	0.60	0.83	0.000	0.015	0.0072	0.0001	0.0008	0.1811
HMDB00729	2-hydroxyisobutyrate	0.99	1.42	1.42	0.107	0.206	0.1083	0.0001	0.0692	0.0196
HMDB00744	Malic Acid	0.88	0.62	0.89	0.719	0.500	0.1673	0.0087	0.6677	0.8917
HMDB00749	Mesaconic acid	1.07	1.22	0.95	0.647	0.477	0.0495	0.0001	0.0245	0.2361

HMDB	Metabolite identification	Comp. 1[a]	Comp. 2[a]	Comp. 3[a]	p value[b]	q value[c]	8-OHDG[d]	HNE-MA[d]	8-isoPGF$_{2\alpha}$[d]	8-NO$_2$Gua[d]
HMDB00750	3-hydroxymandelic acid	0.18	0.66	1.11	0.001	0.020	0.9565	0.0005	0.0108	0.8233
HMDB00752	3-Methylglutaric acid	0.05	1.34	1.12	0.026	0.117	0.1076	0.0023	0.0429	0.9024
HMDB00754	3-Hydroxyisovaleric acid	0.49	0.83	0.90	0.947	0.568	0.1953	0.0239	0.1522	0.0807
HMDB00755	b-(4-Hydroxyphenyl)lactic acid	0.47	0.73	0.99	0.499	0.426	0.8814	0.0203	0.2361	0.0152
HMDB00763	5-Hydroxyindoleacetic acid	1.69	1.34	1.58	0.602	0.462	0.3941	0.0001	0.0134	0.6607
HMDB00784	Azelaic acid	0.07	0.71	0.74	0.129	0.222	0.5159	0.0017	0.4336	0.3267
HMDB00806	Myristic acid	2.18	1.67	1.39	0.107	0.206	0.0211	0.0001	0.0001	0.0092
HMDB00827	Stearic acid	2.52	1.85	1.54	0.540	0.440	0.0143	0.0024	0.0261	0.0001
HMDB00847	Nonanoic acid	0.61	0.82	0.93	0.137	0.227	0.5676	0.0634	0.0298	0.1061
HMDB00849	Rhamnose	0.67	0.67	0.57	0.026	0.116	0.3766	0.0914	0.1407	0.1256
HMDB00857	Pimelic acid	0.04	0.85	0.78	0.002	0.022	0.6134	0.0010	0.1502	0.1220
HMDB00863	Isopropanol	0.99	0.70	0.75	0.386	0.379	0.9804	0.7295	0.8718	0.6418
HMDB00873	4-methylcatechol	1.11	1.09	0.98	0.684	0.488	0.9926	0.9670	0.2501	0.3788
HMDB00881	Xanthurenic acid	0.53	0.39	0.84	0.100	0.202	0.5480	0.7783	0.0317	0.8167
HMDB00883	L-Valine	1.49	1.08	0.96	0.020	0.102	0.3558	0.0344	0.3560	0.0010
HMDB00893	Suberic acid	0.59	0.68	0.88	0.218	0.280	0.3927	0.0400	0.1230	0.6432
HMDB00904	Citrulline	0.73	0.79	0.91	0.182	0.255	0.5990	0.4857	0.4024	0.3911
HMDB00929	L-Tryptophan	0.31	0.59	0.67	0.286	0.326	0.8445	0.1631	0.9042	0.6756
HMDB00943	Threonic acid	2.31	1.64	1.40	0.144	0.231	0.0016	0.0134	0.0336	0.0001
HMDB00956	(±)-Tartaric acid	0.17	0.70	0.99	0.036	0.133	0.3990	0.7274	0.9548	0.4623
HMDB00957	Catechol	2.15	1.68	1.32	0.004	0.034	0.0195	0.0200	0.3246	0.0002
HMDB00958	trans-Aconitic acid	0.44	0.31	0.31	0.002	0.022	0.4744	0.5271	0.0935	0.8356
HMDB00959	Tiglylglycine	0.91	0.64	1.09	0.244	0.298	0.1174	0.0098	0.0516	0.2229
HMDB01051	Glyceraldehyde	0.84	1.03	0.81	0.116	0.213	0.3107	0.0003	0.2475	0.1176
HMDB01123	Anthranilic acid	0.19	0.22	1.31	0.285	0.325	0.3150	0.0001	0.0073	0.7095
HMDB01138	N-acetylglutamic acid	0.96	0.82	1.73	0.173	0.250	0.5401	0.0003	0.0049	0.7253
HMDB01147	Aminomalonic acid	0.97	1.15	0.90	0.325	0.349	0.2399	0.0001	0.0428	0.1381
HMDB01266	L-Sorbose	1.46	1.15	0.92	0.351	0.362	0.0118	0.0006	0.1703	0.0159
HMDB01352	Hydroxypyruvic acid	0.15	0.24	0.33	0.795	0.525	0.2944	0.0026	0.4271	0.5827
HMDB01398	Guaiacol	0.90	1.18	1.10	0.001	0.016	0.3729	0.0229	0.2835	0.1772
HMDB01414	putrescine	0.28	0.20	0.65	0.513	0.431	0.5134	0.0442	0.0098	0.3321
HMDB01476	3-Hydroxyanthranilic acid	1.12	0.78	0.98	0.713	0.498	0.4712	0.1188	0.1071	0.1534
HMDB01488	Niacin	0.76	0.70	1.19	0.161	0.241	0.0478	0.0078	0.0425	0.6665
HMDB01490	Vanylglycol	1.15	0.93	0.73	0.259	0.309	0.2777	0.6107	0.0970	0.3368
HMDB01568	2-Octenoic acid	1.31	1.03	0.86	0.372	0.373	0.0711	0.5574	0.3834	0.0171
HMDB01624	2-Hydroxycaproic acid	0.64	0.86	0.73	0.093	0.197	0.0079	0.0271	0.0060	0.2960

X

HMDB	Metabolite identification	Comp. 1[a]	Comp. 2[a]	Comp. 3[a]	p value[b]	q value[c]	8-OHDG[d]	HNE-MA[d]	8-isoPGF$_{2a}$[d]	8-NO$_2$Gua[d]
HMDB01644	D-Xylulose	0.48	0.95	0.77	0.395	0.383	0.8651	0.1073	0.2452	0.6233
HMDB01847	Caffeine	0.14	0.71	0.58	0.012	0.077	0.7374	0.0951	0.2387	0.4943
HMDB01858	p-Cresol	0.05	0.31	0.25	0.950	0.569	0.1335	0.9970	0.4159	0.1525
HMDB01870	Benzoic acid	0.97	1.15	1.12	0.040	0.138	0.0181	0.0006	0.1404	0.0002
HMDB01881	(±)-1,2-propanediol	0.46	0.78	0.71	0.111	0.209	0.2043	0.4515	0.3616	0.2797
HMDB01954	3-Hydroxyoctanoic acid	0.47	0.75	0.60	0.821	0.533	0.8597	0.0699	0.4750	0.8168
HMDB01988	4-Hydroxycyclohexanecarboxylic acid	0.27	1.02	0.82	0.810	0.530	0.7832	0.0658	0.1873	0.8469
HMDB02039	2-Pyrrolidinone	0.09	0.82	0.77	0.180	0.254	0.3286	0.0092	0.1892	0.1768
HMDB02048	m-Cresol	1.41	0.99	0.83	0.123	0.218	0.1848	0.2472	0.5508	0.1409
HMDB02078	Cyanic acid	0.41	0.83	0.87	0.468	0.413	0.1782	0.0358	0.0266	0.2756
HMDB02142	Phosphoric acid	1.85	2.33	1.82	0.315	0.343	0.1353	0.5977	0.5968	0.3591
HMDB02199	Desaminotyrosine	0.25	1.60	1.31	0.229	0.287	0.0236	0.0019	0.0456	0.1045
HMDB02243	Picolinic acid	1.98	1.61	1.69	0.774	0.518	0.1494	0.0147	0.0600	0.2762
HMDB02259	Heptadecanoic acid	1.27	0.90	0.75	0.561	0.448	0.2663	0.0403	0.2003	0.7449
HMDB02329	Oxalic acid	0.90	0.63	0.73	0.098	0.201	0.7952	0.1584	0.6175	0.9816
HMDB02432	5-(Hydroxymethyl)furoic acid	0.85	1.12	1.01	0.005	0.041	0.6035	0.7636	0.1915	0.7088
HMDB02434	Hydroquinone	0.84	1.38	1.09	0.308	0.339	0.7926	0.0393	0.6851	0.5842
HMDB02643	3-(3-Hydroxyphenyl)-3-hydroxypropionic acid	0.25	0.52	0.41	0.003	0.031	0.2997	0.8750	0.3743	0.6031
HMDB02649	Erythrose	0.42	0.35	0.49	0.011	0.074	0.3075	0.6382	0.2586	0.0450
HMDB03070	Shikimic acid	0.29	0.52	0.75	0.032	0.128	0.5554	0.0657	0.0819	0.4349
HMDB03156	2,3-Butanediol	1.27	0.92	1.03	0.041	0.139	0.5977	0.2487	0.9722	0.9525
HMDB03243	Acetoin	1.42	0.99	0.81	0.004	0.033	0.9255	0.0546	0.2071	0.0563
HMDB03315	Cyclohexanone	2.50	1.76	1.38	0.378	0.375	0.0398	0.0164	0.0796	0.0047
HMDB03466	L-Gulonolactone	0.22	0.81	1.19	0.074	0.182	0.7533	0.1409	0.0059	0.4745
HMDB03903	2-Hydroxyethanesulfonate	0.80	0.98	0.87	0.003	0.031	0.6995	0.0600	0.2051	0.1217
HMDB03911	3-Aminoisobutanoic acid	0.31	0.75	0.59	0.215	0.278	0.1089	0.0409	0.9278	0.1504
HMDB04136	D-Threitol	1.32	0.93	0.88	0.004	0.034	0.0437	0.0592	0.0351	0.0385
HMDB04230	Pyrrole-2-carboxylic acid	0.16	1.45	1.40	0.055	0.161	0.0223	0.0192	0.4029	0.0039
HMDB04437	Diethanolamine	1.74	1.45	1.13	0.023	0.110	0.2987	0.0007	0.8195	0.4339
HMDB04812	2,5-Furandicarboxylic acid	0.79	1.56	1.28	0.596	0.460	0.1831	0.7411	0.5927	0.0220
HMDB04983	Dimethyl sulfone	2.13	1.69	1.50	0.001	0.021	0.0007	0.0001	0.0129	0.0001
HMDB05802	Trans-isoeugenol	0.91	0.74	1.48	0.210	0.274	0.9084	0.2942	0.0369	0.5517
HMDB06116	m-Hydroxyhippuric acid	0.14	0.19	0.99	0.000	0.012	0.0629	0.1852	0.6364	0.4120
HMDB06899	Alanylglycine	0.44	1.15	1.19	0.477	0.417	0.0824	0.2477	0.2131	0.0155
HMDB10719	trans-Hex-2-enoic acid	0.38	1.12	1.03	0.123	0.218	0.3323	0.4104	0.4761	0.0440

XI

HMDB	Metabolite identification	Comp. 1[a]	Comp. 2[a]	Comp. 3[a]	p value[b]	q value[c]	8-OHDG[d]	HNE-MA[d]	8-isoPGF$_{2\alpha}$[d]	8-NO$_2$Gua[d]
HMDB11676	D-Xylono-1,5-lactone	2.00	1.43	1.58	0.679	0.486	0.0324	0.0014	0.0036	0.0176
HMDB11732	Hex-2-ulosonic acid	1.19	1.23	1.06	0.958	0.571	0.8867	0.3316	0.6890	0.5702
HMDB13231	Ethylamine	0.46	0.91	0.71	0.536	0.439	0.9827	0.1952	0.2369	0.6486
HMDB13248	Mono(2-ethylhexyl)phthalate [MEHP]	1.70	1.45	1.42	0.119	0.215	0.0011	0.0506	0.1521	0.0082
HMDB13674	1,2,3-Trihydroxybenzene	1.19	1.51	1.19	0.143	0.230	0.1850	0.7337	0.6522	0.5163
HMDB13716	Norvaline	0.42	0.54	0.44	0.077	0.185	0.8619	0.0541	0.2283	0.4500
HMDB14328	Amphetamine	0.60	0.68	1.07	0.343	0.358	0.5672	0.0255	0.0097	0.7313
HMDB29635	4-Methylbenzoic acid	1.01	0.89	0.70	0.617	0.467	0.4232	0.1444	0.0192	0.0425
HMDB29739	1H-Indole-3-acetamide	1.22	0.92	1.19	0.006	0.045	0.2003	0.9194	0.1897	0.0362
HMDB29942	Arabinose	1.50	1.28	1.00	0.770	0.517	0.0371	0.0025	0.0146	0.0077
HMDB31213	2-Ethoxyethanol	1.13	0.90	0.71	0.154	0.236	0.4991	0.0383	0.0031	0.0762
HMDB31320	1,3-butanediol	1.38	0.98	0.83	0.025	0.114	0.7117	0.0622	0.2327	0.0195
HMDB31404	Cyclohexylamine	0.68	0.88	0.80	0.433	0.399	0.3059	0.4540	0.0330	0.2994
HMDB31444	Dodecane	1.88	1.47	1.56	0.212	0.276	0.0120	0.0001	0.1099	0.0065
HMDB31445	Undecane	0.72	1.19	1.35	0.295	0.332	0.6644	0.5070	0.9094	0.0018
HMDB31450	Decane	1.57	1.14	0.89	0.986	0.578	0.1120	0.0011	0.0029	0.6343
HMDB31602	4-Pentenoic acid	0.71	0.66	0.59	0.101	0.202	0.2216	0.8091	0.0553	0.5298
HMDB31626	2-phenylpropanal	1.12	1.12	1.33	0.031	0.125	0.0184	0.3475	0.7970	0.1510
HMDB32037	1,3-Benzenediol	0.38	0.28	0.72	0.003	0.030	0.0606	0.3639	0.0720	0.3702
HMDB32619	(±)-1-phenylethanol	0.43	0.57	0.47	0.026	0.117	0.7552	0.0142	0.7730	0.7111
HMDB33244	Dibutyl phthalate	0.58	0.43	0.77	0.208	0.273	0.8265	0.0446	0.3435	0.6329
HMDB33958	Dihydro-4-hydroxy-5-hydroxymethyl-2(3H)-furanone	0.39	1.40	1.32	0.010	0.066	0.2138	0.2657	0.0458	0.5398
HMDB34220	Inositol	1.50	1.53	1.44	0.880	0.550	0.1996	0.0003	0.1427	0.0001
HMDB34284	Tridecane	1.95	1.37	1.16	0.445	0.404	0.1738	0.0120	0.5312	0.0051
HMDB34778	(2R*,3R*)-1,2,3-Butanetriol	0.69	0.57	0.46	0.779	0.520	0.4161	0.0006	0.0084	0.1226
HMDB34976	Borneol	1.53	1.63	1.28	0.150	0.234	0.7223	0.7451	0.0537	0.4094
HMDB35056	1-(3-Hydroxy-4-methoxyphenyl)-1,2-ethanediol	1.04	1.16	0.91	0.136	0.226	0.1501	0.7574	0.6302	0.6822
HMDB35227	Tartronic acid	0.04	0.86	0.67	0.983	0.578	0.5803	0.1195	0.2741	0.5077
HMDB37050	o-Cymene	1.93	1.63	1.75	0.806	0.528	0.0022	0.0033	0.1216	0.0021
HMDB41486	1-Deoxy-D-ribitol	0.90	0.63	0.50	0.152	0.236	0.4176	0.1673	0.4732	0.1467
HMDB41861	Cyanuric acid	0.58	0.47	0.93	0.136	0.227	0.0782	0.0143	0.3659	0.5550
HMDB41932	Methylephedrine	0.72	0.75	0.60	0.568	0.451	0.0864	0.2468	0.3534	0.0780
HMDB42032	Thiodiacetic acid	1.51	1.07	1.02	0.798	0.526	0.1012	0.0033	0.1419	0.0422
HMDB59873	3-Ethylphenol	1.16	0.86	0.69	0.914	0.560	0.1045	0.0205	0.3326	0.3190

HMDB	Metabolite identification	Comp. 1[a]	Comp. 2[a]	Comp. 3[a]	p value[b]	q value[c]	8-OHDG[d]	HNE-MA[d]	8-isoPGF$_{2a}$[d]	8-NO$_2$Gua[d]
HMDB59889	Isohexanol	0.52	0.53	0.56	0.684	0.487	0.3668	0.7734	0.2347	0.0789
HMDB59901	Hemimellitene	0.41	1.07	1.13	0.487	0.421	0.2106	0.2099	0.8353	0.9239
HMDB61927	5-Hydroxypentanoic acid	1.06	0.76	0.76	0.261	0.310	0.7495	0.0618	0.6087	0.1083
HMDB61933	1H-Indole-2,3-dione	0.03	0.90	0.71	0.323	0.347	0.1488	0.1902	0.1407	0.6546

[a] VIP score of components 1, 2, and 3 from PLS-DA analysis

[b] p values from Student's t test analysis of peak area between high and low exposure groups

[c] FDR adjusted q value from Student's t test p value

[d] p values from Pearson's correlation analysis of peak area and urine oxidative stress biomarkers
concentration

7.2 Appendix 2: Identified potential metabolite features in serum sample of 107 subjects using metabolomics.

Name	HMDB	Average VIP score	ANCOVA p value	Trend
Carnitine	HMDB0000062	2.47	<0.05	Up
Oxoglutaric acid	HMDB0000208	2.33	<0.05	Down
Pyroglutamic acid	HMDB0000267	2.31	<0.05	Down
Adenosine monophosphate	HMDB0000045	1.85	<0.05	Down
Inosinic acid	HMDB0000175	1.77	<0.05	Down
Malic Acid	HMDB0000744	1.61	<0.05	Down
Ketoleucine	HMDB0000695	1.58	<0.05	Up
Aspartic acid	HMDB0000191	1.54	<0.05	Up
Pyroglutamic acid	HMDB0000267	1.50	<0.05	Down
Octenoyl-L-carnitine	HMDB0013324	1.44	<0.05	Up
Isovalerylcarnitine	HMDB0000688	1.28	<0.05	Up
(R)-3-Hydroxybutyric acid	HMDB0000011	0.17	1.00	Down
Octenoyl-L-carnitine	HMDB0013324	0.49	0.99	Up
Dehydroascorbic acid	HMDB0001264	0.50	0.97	Up
L-Kynurenine	HMDB0000684	0.28	0.96	Down
Octenoyl-L-carnitine	HMDB0013324	0.42	0.95	Down
Indolelactic acid	HMDB0000671	0.16	0.94	Down
L-Tryptophan	HMDB0000929	0.32	0.93	Up
L-Alanine	HMDB0000161	0.31	0.92	Up
Undecanoyl-L-carnitine	NA	0.32	0.82	Up
Decenoyl-L-carnitine	NA	0.50	0.81	Down
cis-Aconitate	HMDB0000072	0.21	0.79	Down
Propionyl-L-carnitine	HMDB0000824	0.50	0.78	Up
3-Methyladenine	HMDB0011600	0.23	0.77	Up
Tetradecenoyl-L-carnitine	HMDB0013329	0.51	0.66	Up
Deoxycholic acid	HMDB0000626	0.34	0.65	Down
Tyrosine	HMDB0000158	0.62	0.65	Up
Methionine	HMDB0000696	0.66	0.62	Up
Glycerophosphocholine	HMDB0000086	0.39	0.62	Up
1-Methyladenosine	HMDB0003331	0.96	0.60	Up
Tiglyl-L-carnitine	HMDB0002366	0.67	0.58	Up
cis-Aconitate	HMDB0000072	0.72	0.57	Down
L-Acetylcarnitine	HMDB0000201	0.56	0.56	Down
Uridine	HMDB0000296	0.26	0.56	Up

Name	HMDB	Average VIP score	ANCOVA p value	Trend
D-Glucurono-6,3-lactone	HMDB0006355	0.42	0.54	Down
L-Alloisoleucine	HMDB0000557	0.56	0.51	Up
L-Acetylcarnitine	HMDB0000201	0.76	0.51	Up
L-Alloisoleucine	HMDB0000557	0.77	0.46	Down
3-Methyladenine	HMDB0011600	0.84	0.44	Down
Deoxycholic acid	HMDB0000626	0.57	0.44	Up
Pipecolic acid	HMDB0000070	0.48	0.44	Up
N-Acetyl-L-alanine	HMDB0000766	0.32	0.42	Up
5-Methylcytidine	HMDB0000982	0.69	0.42	Down
L-Norleucine	HMDB0001645	0.54	0.42	Up
Dodecenoyl-L-carnitine	HMDB0013326	0.83	0.42	Down
citric acid	HMDB0000094	0.99	0.42	Down
Dehydroascorbic acid	HMDB0001264	0.80	0.40	Down
Methionine	HMDB0000696	0.85	0.37	Up
N6-Acetyl-L-lysine	HMDB0000206	0.86	0.36	Down
Ascorbic Acid	HMDB0000044	0.54	0.35	Down
Levulinic acid	HMDB0000720	0.77	0.33	Up
(R)-3-Hydroxybutyric acid	HMDB0000011	0.59	0.31	Up
Carnitine	HMDB0000062	0.78	0.27	Up
Cholic Acid	HMDB0000619	0.90	0.27	Up
Pyruvic acid	HMDB0000243	0.78	0.27	Down
Phenylalanine	HMDB0000159	1.04	0.26	Down
Hexanoylcarnitine	HMDB0000705	0.77	0.25	Up
Taurodeoxycholic acid	HMDB0000896	0.84	0.23	Down
Hypoxanthine	HMDB0000157	0.97	0.21	Down
Hypoxanthine	HMDB0000157	0.68	0.21	Down
D-threo-Isocitric acid	HMDB0001874	0.79	0.21	Down
2-Ketohexanoic acid	HMDB0001864	0.94	0.21	Up
5-Aminolevulinic acid	HMDB0001149	0.88	0.21	Up
Tetradecadienoyl-L-carnitine	NA	1.07	0.20	Down
Decanoylcarnitine	HMDB0000651	0.65	0.18	Up
2-Hydroxy-2-methylbutyric acid	HMDB0001987	0.69	0.17	Up
Butyrylcarnitine	HMDB0002013	0.91	0.16	Up
3-Hydroxyisovaleric acid	HMDB0000754	0.76	0.15	Up
2-Hydroxybutyric acid	HMDB0000008	1.03	0.14	Up
Pipecolic acid	HMDB0000070	0.94	0.14	Up

XV

Name	HMDB	Average VIP score	ANCOVA p value	Trend
Decenoyl-L-carnitine	NA	0.89	0.14	Up
3-Hydroxydodecanoyl-L-carnitine	HMDB0061638	1.02	0.13	Down
Octanoyl-L-carnitine	HMDB0000791	0.92	0.12	Up
Malic Acid	HMDB0000744	1.04	0.12	Down
Dopamine	HMDB0000073	1.29	0.11	Up
L-alpha-Aminobutyric acid	HMDB0000452	1.02	0.10	Up
D-threo-Isocitric acid	HMDB0001874	1.14	0.10	Down
Isovalerylcarnitine	HMDB0000688	1.25	0.09	Up
L-Valine	HMDB0000883	1.43	0.09	Down
Creatine	HMDB0000064	1.22	0.09	Up
Creatinine	HMDB0000562	1.46	0.08	Up
Oxoglutaric acid	HMDB0000208	1.03	0.06	Down
Hydroxyphenyllactic acid	HMDB0000755	1.05	0.06	Down
N-Acetyl-L-alanine	HMDB0000766	0.93	0.06	Up

7.3 Appendix 3: Identified potential lipid species in serum sample of 107 subjects using lipidomics.
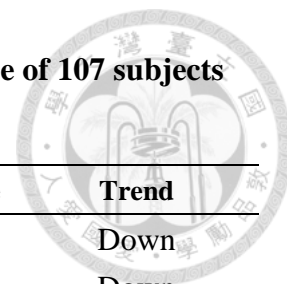
Name	Average VIP score	ANCOVA p value	Trend
CB (d18:1/18:1)	0.605	0.122	Down
CB (d18:1/22:1)	0.145	0.996	Down
CB (d18:1/24:1)	0.979	0.484	Down
CB (d18:1/25:0)	0.749	0.862	Down
CB (d18:1/26:1)	0.580	0.222	Down
Cer (d18:1/22:1)	0.605	0.345	Down
LCP (17:0)	1.278	0.039	Down
LCP (19:0)	1.299	0.026	Down
LPC (14:0)	0.434	0.103	Down
LPC (15:0)	0.981	0.060	Down
LPC (16:0)	0.910	0.350	Down
LPC (17:1)	0.432	0.731	Down
LPC (20:1)	0.593	0.045	Down
LPC (20:4)	0.896	0.344	Down
LPC (22:5)	0.655	0.309	Down
LPC (22:6)	1.036	0.090	Down
PC (16:0/16:0)	0.159	0.350	Down
PC (18:0/16:0)	0.393	0.527	Down
PC (18:0/18:0)	1.402	0.004	Down
PC (18:0/22:4)	0.751	0.444	Down
PC (18:2/20:2)	1.811	0.054	Down
PC (18:3/16:0)	0.580	0.189	Down
PC (18:3/18:2)	1.509	0.017	Down
PC (20:1/18:0)	1.487	0.000	Down
PC (20:4/16:0)	1.288	0.210	Down
PC (20:4/17:0)	0.698	0.645	Down
PI (32:1)	1.765	0.051	Down
PI (34:3)	1.622	0.812	Down
SM (d18:1/12:0)	0.800	0.707	Down
SM (d18:1/17:0)	0.422	0.956	Down
SM (d18:1/20:0)	0.858	0.731	Down
SM (d18:1/21:0)	1.334	0.372	Down
SM (d18:1/23:1)	0.697	0.797	Down
SM(d18:1/17:1)	0.506	0.954	Down

Name	Average VIP score	ANCOVA p value	Trend
CB (d18:1/22:0)	0.236	0.331	Up
CB (d18:1/23:0)	0.418	0.061	Up
Cer (d18:1/24:1)	0.741	0.079	Up
LPC (16:1)	0.726	0.427	Up
LPC (18:1)	1.101	0.003	Up
LPC (18:2)	0.541	0.046	Up
LPC (18:3)	0.633	0.436	Up
PC (16:0/14:0)	0.680	0.216	Up
PC (16:0/17:1)	0.501	0.164	Up
PC (16:0/20:1)	2.604	0.000	Up
PC (16:1/14:0)	0.877	0.018	Up
PC (17:0/18:2)	0.887	0.014	Up
PC (18:0/22:5)	0.929	0.111	Up
PC (18:0/22:6)	0.198	0.852	Up
PC (18:1/17:0)	0.301	0.238	Up
PC (18:2/14:0)	1.711	0.000	Up
PC (18:2/17:1)	1.312	0.001	Up
PC (18:2/20:5)	1.281	0.002	Up
PC (20:4/14:0)	0.713	0.622	Up
PC (22:6/17:0)	0.447	0.491	Up
SM (d18:1/14:0)	1.056	0.105	Up
SM (d18:1/15:0)	0.885	0.117	Up
SM (d18:1/18:0)	0.700	0.050	Up
SM (d18:1/19:1)	1.057	0.119	Up
SM (d18:1/22:0)	1.303	0.012	Up
SM (d18:1/23:0)	0.626	0.006	Up
SM (d18:1/24:1)	0.822	0.023	Up
SM (d18:1/25:0)	1.215	0.009	Up
SM (d18:1/25:1)	0.484	0.381	Up
SM(d18:1/26:1)	0.536	0.204	Up

CB: cerebroside; CER: ceramide; LPC: lysophosphatidylcholine;

PC: phosphatidylcholine; PI: phosphatidylinositol; SM: sphingomyelin

Linking sources to early effects by profiling urine metabolome of residents living near oil refineries and coal-fired power plants☆,☆☆

Chi-Hsin Sally Chen [a], Tzu-Hsuen Yuan [a], Ruei-Hao Shie [b], Kuen-Yuh Wu [a], Chang-Chuan Chan [a,*]

[a] Institute of Occupational Medicine and Industrial Hygiene, College of Public Health, National Taiwan University, Taiwan
[b] Green Energy and Environment Research Laboratories, Industrial Technology Research Institute of Taiwan, Taiwan

ARTICLE INFO

ABSTRACT

Background: This study aims at identifying metabolic changes linking external exposure to industrial air toxics with oxidative stress biomarkers.

Methods: We classified 252 study subjects as 111 high vs. 141 low exposure subjects by the distance from their homes to the two main emission sources, oil refineries and coal-fired power plants. We estimated individual's external exposure to heavy metals and polycyclic aromatic hydrocarbons (PAHs) by dispersion and kriging models, respectively. We measured urinary levels of heavy metals and 1-hydroxypyrene (1-OHP) as biomarkers of internal exposure, and 8-OHdG, HNE-MA, 8-isoPGF$_{2\alpha}$, and 8-NO$_2$Gua as biomarkers of early health effects. We used two-dimensional gas chromatography time-of-flight mass spectrometry to identify urine metabolomics. We applied "meet-in-the-middle" approach to identify potential metabolites as putative intermediate biomarkers linking multiple air toxics exposures to oxidative stress with plausible exposures-related pathways.

Results: High exposure subjects showed elevated ambient concentrations of vanadium and PAHs, increased urine concentrations of 1-OHP, vanadium, nickel, copper, arsenic, strontium, cadmium, mercury, and thallium, and higher urine concentrations of all four urine oxidative stress biomarkers compared to low exposure subjects. We identified a profile of putative intermediate biomarkers that were associated with both exposures and oxidative stress biomarkers in participants. Urine metabolomics identified age-dependent biological pathways, including tryptophan metabolism and phenylalanine metabolism in children subjects (aged 9–11), and glycine, serine, and threonine metabolism in elderly subjects (aged > 55), that could associate multiple exposures with oxidative stress.

Conclusion: By profiling urine biomarkers and metabolomics in children and elderly residents living near a petrochemical complex, we can link their internal exposure to oxidative stress biomarkers through biological pathways associated with common complex chronic diseases and allergic respiratory diseases. The internal exposure may possibly be traced to multiple air toxics emitted from specific sources of oil refineries and coal-fired power plants.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

The concept of the exposome was first proposed in 2005 (Wild, 2005), and characterized as the comprehensive evaluation of all exposures and their contribution to disease causation or progression (Rappaport and Smith, 2010). Recently, a branching paradigm, the public health exposome, focuses on the impact of exposures on the overall health of a population within a particular region, with the intention of identifying vulnerable populations with higher risks of chronic illnesses (Juarez et al., 2014; Smith et al., 2015). The use of omics methods has been recommended in exposomics studies to identify the links between exposures and health outcomes (Vineis et al., 2013). Untargeted metabolomics allow unbiased global profiling of metabolites, the endpoint of biological processes, and could best reflect the biochemical effects of exposure and easier correlated with phenotypes (Patti et al., 2012). So far metabolomics have been applied in pharmacology, clinical disease diagnosis, nutritional, and environmental studies (Robertson et al., 2011). However, most environmental metabolomics studies have focused on the assessment of single exposure (Ellis et al., 2012; Gao et al., 2014).

Petrochemical complex is usually a consortium of high-pollution facilities such as coal-fired power plants and oil refineries that emit multiple air toxics including sulfur dioxide (SO$_2$), nitrogen oxides (NO$_x$), polycyclic aromatic hydrocarbons (PAHs), heavy metals, and volatile

organic compounds (VOCs) which can induce common complex diseases (Chan et al., 2006; Driscoll et al., 2015; Nadal et al., 2004). Epidemiological studies have shown increased risks of diseases such as asthma and cancer in residents living near petrochemical complexes, without linking health effects to emissions from specific sources (Pan et al., 1994; Wichmann et al., 2009). The complexity of petrochemical-related air toxics and potential health effects calls for a comprehensive evaluation on nearby communities.

This study selected high exposure subjects who lived in close proximity to a predominant emission source, the No. 6 Naphtha Cracking Plant, the largest petrochemical complex in Taiwan, which houses 64 plants including a coal-fired power plant and three oil refineries (Shie et al., 2013), and low exposure subjects that lived further away. We aimed to clarify the effects of air toxics exposures on biological pathways that could lead to increased oxidative stress by measuring and comparing multiple air toxics exposures and possible health effect oxidative stress in the high and low exposure subjects, and using two-dimensional gas chromatography time-of-flight mass spectrometry (GC × GC — TOFMS)-based untargeted urine metabolomics to identify the link between exposures and oxidative stress. We specifically focused on children (aged 9–15) and elderly (aged > 55) residents whom are more susceptible to environmental exposures. Children are more susceptible due to immature physical development, and higher inhalation of air per unit time, while elderly may have compromised immune responses and underlying health conditions (Adler, 2003; Makri and Stilianakis, 2008). Since age has been reported to affect the urine metabolite profile, in this study we conducted the statistical analysis of the metabolomics results for children and elderly subjects separately (Slupsky et al., 2007; Thevenot et al., 2015).

2. Materials and methods

2.1. Study area and subjects

Our study area surrounded the No. 6 Naphtha Cracking Plant, Taiwan's largest petrochemical complex, which is located in Yunlin County on the western coast of central Taiwan (Fig. 1A). The complex

began major operations in 2000, covering a total area of 2603 ha. As shown in Fig. 1B, the complex houses 64 plants including one coal-fired power plant that generates 1800 MW of power, three oil refineries with a total capacity of 450,000 to 540,000 barrels of crude oil per day, three co-generation plants that generate 2820 MW of power, three naphtha cracking plants that produce 2.9 million tons of ethylene per year, and other downstream plants (Shie et al., 2013). Our study subjects were selected from a prospective cohort of 3230 residents (aged 5–88) living in the three townships closer to the complex (pink) and the other seven townships relatively farther away from the complex (yellow) as shown in Fig. 1B. All of them have completed questionnaire surveys on key factors related to exposure and provided one morning spot urine sample to measure biomarkers of exposure, early effects, and metabolomics. Among them, we had 257 cohort members who lived in the three townships closest to the complex, with urine concentrations of previously established exposure biomarkers vanadium (V) and PAHs metabolite, 1-hydroxypyrene (1-OHP), in the top 60% of the 3230 residents (Yuan et al., 2015a; Yuan et al., 2015b), and another 337 cohort members who lived in townships further away, with urine concentrations of V and 1-OHP in the bottom 40%. We then randomly select 40 children (aged 9–15) and 71 elderly (aged > 55) subjects from the 257 cohort members as our high exposure subjects, and 70 children and 71 elderly participants from the 337 cohort members as our low exposure subjects. Study subjects' home locations were shown in Fig. 1B. This study was approved by the Research Ethics Committee of the National Taiwan University Hospital, and informed consent was obtained for each participant.

2.2. External exposures

Geographic coordinates for each participant's home address were determined, and geological information system (GIS) software (ArcGIS version 10.1) was used to calculate the distances from each home address to previously identified main emission points of coal-fired power plant and oil refineries (Shie et al., 2013), respectively. GIS software was also used to measure road area surrounding homes, in order to estimate traffic contribution on air toxics levels. Ambient
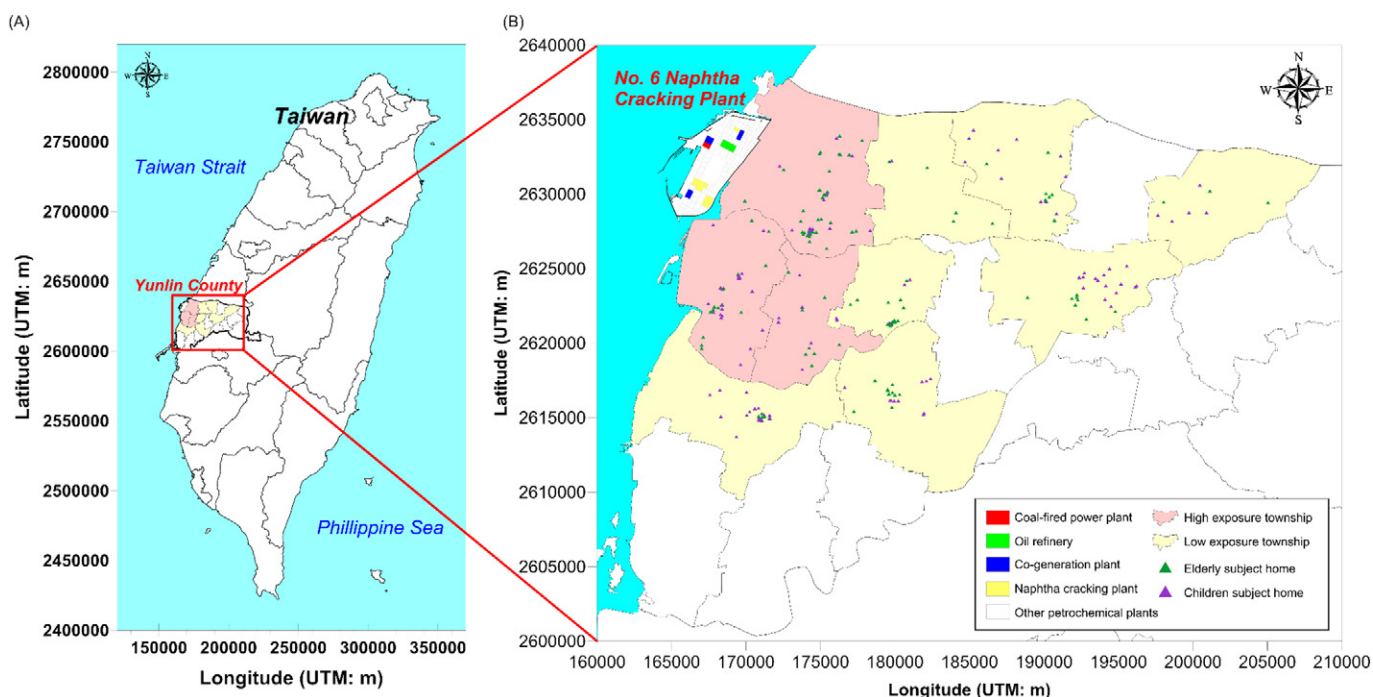

Fig. 1. GIS map of (A) Yunlin county in Taiwan, and (B) the locations of petrochemical plants, study area, and study subjects' homes.

concentrations of V and five PAHs including pyrene, fluoranthene, dibenzo[*a,h*]anthracene, benzo[*k*]fluoranthene, and benzo[*a*]anthracene were calculated using previously established two-stage dispersion model and kriging method model, respectively (see Fig. S1) (Chio et al., 2014; Yuan et al., 2015b). Briefly, the two-stage dispersion model of ambient V was established using 192 air samples obtained from 19 sampling sessions at 14 sampling sites from 2009 to 2012. In the first stage of modeling, a backward fitting approach was used to estimate V emissions from the main oil refinery emission point in the complex. At the second stage, model-derived V emission rates and meteorological parameters such as ambient temperature, wind direction, and wind speed were fed to the Industrial Source Complex, version 3 models to estimate the spatial distribution of V. Kriging interpolation method and shift weighting average were then used to estimate ambient V concentrations at study subjects' home addresses in the simulation domain of 50 km × 40 km. 1424 residents from the same cohort as the present study were used in a multiple regression model, which confirmed the association between urinary vanadium and estimated air levels at home address after adjusting for age, gender, education level, smoking, alcohol intake, betel nut intake, fish consumption, and drinking water source (Yuan et al., 2015a). Ambient PAHs models were established using air samples collected at 10 sampling locations in Yunlin at three different time points during downwind season in 2011 (May, June, and November). Kriging interpolation method was used to estimate the spatial distribution of the PAHs in the simulation domain of 100 m × 100 m. 781 residents from the same cohort as the present study were used to determine the association between the estimated ambient PAHs levels at their home addresses and their individual urinary 1-OHP concentrations by applying a multiple linear regression model. After adjusting for confounders including age, gender, education level, smoking habit, grilled food consumption, incense burning frequency, and proximity of residence to a road, results showed that urinary 1-OHP levels significantly correlated with the five PAHs we cited in this study (pyrene, benzo[*a*]anthracene, benzo[*k*]fluoranthene, fluoranthene, and dibenzo[*a,h*]anthracene) (Yuan et al., 2015b).

2.3. Urine exposure biomarkers

Urine concentrations of V and 1-OHP were analyzed using inductively coupled plasma mass spectrometry (ICP-MS) and high-performance liquid chromatography (HPLC) methods, respectively, as previously reported (Yuan et al., 2015a; Yuan et al., 2015b). Additional analysis of urine chromium (Cr), nickel (Ni), copper (Cu), arsenic (As), strontium (Sr), cadmium (Cd), mercury (Hg), thallium (Tl) and lead (Pb) were performed using the same ICP-MS instrument and method as V analysis (see Supplemental Material, p. 19). Urinary creatinine analysis was performed on all the urinary samples for adjustment of urinary exposure biomarker levels. Urine creatinine analysis was conducted using enzyme-linked immunosorbent assay at National Taiwan University Hospital medical diagnosis laboratory. All the subjects selected in this study had urinary creatinine concentrations between 30 and 300 mg/dL.

2.4. Urine oxidative stress biomarkers

Urine concentrations of four oxidative stress biomarkers 8-hydroxy-2′-deoxyguanosine (8-OHdG), 4-hydroxy-2-nonenal-mercapturic acid (HNE-MA), 8-isoprostaglandin $F_{2\alpha}$ (8-isoPF$_{2\alpha}$), and 8-nitroguanine (8-NO$_2$Gua) were analyzed using a liquid chromatography-tandem mass spectrometry (LC-MS/MS) method in previously published literature (Wu et al., 2016). Urinary oxidative stress biomarkers levels were adjusted using urinary creatinine concentrations.

2.5. Urine metabolomics

Samples were prepared and analyzed following previously published protocol (Chan et al., 2011), using a Pegasus 4D GC × GC —

TOFMS (LecoCorp., St. Joseph, MI, USA) for analysis (see Supplemental Material, p. 20 and Fig. S2). After data cleaning and preprocessing, acquired data were outputted as peak area for NIST library (version 08, National Institute of Standards and Technology, Gaithersburg, MD, USA) identified potential metabolite peaks, and we selected those with mass spectrum matched library spectrum for >60% (similarity score > 600 out of 1000). After removing artifacts and peaks missing in >50% of the individuals in children or elderly groups, we obtained 405 and 391 potential metabolite peaks for each individual in children and elderly subjects, respectively.

2.6. Biological information search

NIST library match showed the derivatized form of potential metabolite peaks, which we converted to underivatized forms by replacing methoxyamine and trimethylsilyl groups with carbonyl and hydroxyl functional groups, respectively. ChemSpider was used for the identification of the underivatized forms of potential metabolite peaks (Royal Society of Chemistry, London, UK). All potential metabolites peaks were then put through an online repository (http://cts.fiehnlab.ucdavis.edu/) and searched under three online databases: the Human Metabolome Database (HMDB) (Wishart et al., 2013), Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2014), and Chemical Entities of Biological Interest (ChEBI) (Hastings et al., 2013), for identification of known metabolites, chemical class, and involved biological pathways.

2.7. Meet-in-the-middle approach

For identifying exposures-related potential metabolite peaks, peak area data were normalized by sum of total peak area so that the cumulative peak area summed up to 1 in each subject, log transformed, and autoscaled before performing partial least squares discrimination analysis (PLS-DA) between high and low exposure groups in children and elderly subjects, respectively. PLS-DA models were validated using permutation test and cross-validation test. We further used Student's *t*-test to compare the peak area of each potential metabolite peak between high and low exposure groups. For identifying oxidative stress-related potential metabolite peaks, peak area data were adjusted using creatinine concentration for each subject, before using Pearson's correlation test to assess the association with each of the four oxidative stress biomarkers' creatinine-adjusted urine concentration. By comparing the two lists of potential metabolite peaks related to exposures or oxidative stress, we identified intercepting potential metabolite peaks as putative intermediate biomarkers associating multiple exposures with oxidative stress (Chadeau-Hyam et al., 2011).

2.8. Pathway analysis

Pathway analysis was performed using Metaboanalyst 3.0 (The Metabolomics Innovation Center, Edmonton, Alberta, Canada), which currently supports 80 pathways in the *Homo sapiens* pathway library (Xia et al., 2015). HMDB ID and peak area values were used as input, and were first processed by log transformation, normalization by sum and autoscaling (mean-centered and divided by the standard deviation of each variable). The method "Globaltest" was used for pathway enrichment analysis, and "betweenness centrality" for pathway topology analysis.

2.9. Statistical analysis

For comparison of basic characteristics and external exposure levels between high and low exposure groups, we used Student's *t*-test or Wilcoxon Mann Whitney test to analyze continuous variables, and Chi-squared test or Fisher's exact test for discrete variables. We compared log-transformed urine concentrations of exposure and oxidative

stress biomarkers between high and low exposure groups using analysis of covariance (ANCOVA) test, adjusting for age, gender, smoking, alcohol consumption, betel nut chewing, fish consumption, and source of drinking water with a post comparison by Scheffe test. Betel nut chewing is a unique lifestyle factor in Yunlin area. Previous studies have shown that betel nut chewing may increase the risks of arsenic, cadmium, and lead exposure, and could also induce oxidative stress (Al-Rmalli et al., 2011; Shih et al., 2010; Wang et al., 2007). Student's *t*-test, Wilcoxon Mann Whitney test, Chi-squared test, Fisher's exact test, ANCOVA test, and Pearson's correlation test were performed using SAS 9.2 for Windows. PLS-DA was performed using Metaboanalyst 3.0. FDR q value was measured using fdrtool package in R-3.1.3 for Windows.

3. Results

Table 1 showed the comparison between high and low exposure groups in basic characteristics, external exposure at each subject's home locations, and internal exposures as urine biomarkers concentrations. Overall, high exposure subjects lived 10.07 ± 2.43 km away from the main emission point of the coal-fired power plant and 9.35 ± 2.65 km away from the main emission point of oil refineries, while low exposure subjects lived 21.64 ± 5.19 and 20.69 ± 5.00 km away from the two main emission points, respectively. High and low exposure subjects in children (13.76 ± 0.93 years old) and elderly

(65.88 ± 6.92 years old) age groups showed no significant difference in gender distribution, smoking, drinking, and betel nut chewing history, body mass index, and systolic blood pressure. Ambient concentrations of V, and three PAHs pyrene, fluoranthene, and dibenzo[*a,h*]anthracene were significantly higher at the home locations of high exposure subjects when compared to low exposure subjects, for both children and elderly participants. Another PAH, benzo[*k*]fluoranthene, was significantly increased in high exposure group in elderly subjects, but showed no difference between high and low exposure groups in children subjects. Benzo[*a*]anthracene was decreased in high exposure groups for both children and elderly subjects. Road areas surrounding participants' homes, which we used to represent traffic contribution of air toxics levels, showed no difference between high and low exposure groups for both children and elderly residents at 25 m buffer. When we increased the buffer to 500 m, elderly subjects in the low exposure group had larger road areas surrounding their homes than those in the high exposure group. Urine concentrations of 1-OHP, V, Ni, Cu, As, Sr, Cd, Hg, and Tl were significantly increased in high exposure groups compared to low exposure groups for both children and elderly subjects. The difference between high and low exposure groups was most profound for V, 1-OHP, and Tl, followed by Sr. Pearson's correlation analysis results showed significant correlation between ambient and urinary V levels; and between ambient pyrene, fluoranthene, and dibenzo[*a,h*]anthracene and urine 1-OHP concentrations for both children and elderly subjects (see Table S1).

Table 1
Comparison of basic characteristics and exposure levels in 252 study subjects.

	Children					Elderly				
	High (n = 40)		Low (n = 70)		p^a	High (n = 71)		Low (n = 71)		p^a
Basic characteristics										
Age (years), mean ± SD	13.78	±0.93	13.83	±0.89	0.88	66.23	±6.54	66.36	±7.47	0.76
Male, n (%)	22	(55.0)	38	(54.3)	0.94	28	(39.4)	35	(49.3)	0.24
Smoking history, n (%)	3	(7.5)	3	(4.3)	0.67	3	(4.2)	7	(9.9)	0.33
Drinking history, n (%)	2	(5.0)	1	(1.4)	0.30	8	(11.3)	7	(9.9)	0.79
Betel nut chewing history, n (%)	1	(2.5)	0	(0)	0.60	4	(5.6)	3	(4.2)	1.00
BMI (kg/m^2), mean ± SD	21.13	±3.20	20.05	±3.43	0.10	26.30	±3.89	26.36	±3.35	0.93
SBP (mmHg), mean ± SD	117.63	±13.02	115.79	±14.26	0.50	140.76	±20.74	141.76	±18.95	0.77
External exposures at study subjects' homes[b], mean ± SD										
Distance to coal-fired power plant	10.31	±2.50	22.66	±10.31	<0.05	9.94	±2.39	20.62	±4.72	<0.05
Distance to oil refinery	9.80	±2.63	21.73	±5.20	<0.05	9.09	±2.65	19.66	±4.59	<0.05
Road area surrounding homes										
25 m buffer	241.08	±279.23	279.23	±226.64	0.38	269.41	±222.79	304.52	±237.50	0.37
500 m buffer	68,088.07	±61,248.67	61,248.67	±20,502.19	0.18	71,269.64	±22,808.14	82,235.11	±27,217.15	<0.05
Ambient concentrations										
Vanadium	8.60	±1.39	5.75	±1.08	<0.05	8.97	±1.63	6.22	±0.96	<0.05
Polycyclic aromatic hydrocarbons[c]										
Pyrene	0.026	±0.005	0.023	±0.003	<0.05	0.030	±0.005	0.022	±0.003	<0.05
Fluoranthene	0.028	±0.001	0.026	±0.003	<0.05	0.027	±0.001	0.024	±0.003	<0.05
Dibenzo[*a,h*]anthracene	0.013	±0.002	0.011	±0.001	<0.05	0.014	±0.002	0.011	±0.001	<0.05
Benzo[*k*]fluoranthene	0.017	±0.003	0.017	±0.002	0.98	0.019	±0.003	0.018	±0.002	<0.05
Benzo[*a*]anthracene	0.017	±0.001	0.019	±0.001	<0.05	0.017	±0.001	0.020	±0.002	<0.05
Internal exposures[d], mean ± SD										
1-Hydroxypyrene	0.25	±0.31	0.03	±0.01	<0.05	0.42	±0.70	0.03	±0.01	<0.05
Vanadium	2.34	±1.53	0.23	±0.10	<0.05	4.02	±2.23	0.17	±0.08	<0.05
Chromium	3.89	±4.56	2.06	±1.65	0.11	5.32	±7.33	2.98	±2.62	0.09
Nickel	10.41	±16.62	3.70	±2.89	<0.05	11.28	±15.34	8.33	±29.64	<0.05
Copper	16.38	±14.94	11.22	±7.50	<0.05	22.87	±24.33	17.36	±30.23	<0.05
Arsenic	62.28	±42.02	39.47	±29.46	<0.05	119.60	±205.04	64.92	±51.73	<0.05
Strontium	170.77	±249.93	70.47	±64.55	<0.05	211.26	±176.33	86.53	±55.23	<0.05
Cadmium	0.34	±0.34	0.19	±0.15	<0.05	1.30	±1.11	0.87	±0.73	<0.05
Mercury	3.30	±3.15	1.92	±1.81	<0.05	2.59	±2.35	1.49	±1.27	<0.05
Thallium	2.13	±3.92	0.21	±0.11	<0.05	1.60	±3.67	0.12	±0.08	<0.05
Lead	0.77	±1.08	0.65	±0.59	0.71	1.79	±2.30	1.09	±1.21	0.17

BMI: Body Mass Index; SBP: Systolic Blood Pressure.

[a] Comparison of basic characteristics between the high and low exposure groups for continuous variables was performed using Student's *t*-test, and for discrete variables, Chi-squared test or Fisher's exact test. Urinary exposure biomarker concentrations are log-transformed, high and low exposure groups compared by ANCOVA test adjusting age, gender, smoking, alcohol consumption, betel nut chewing, fish consumption, and source of drinking water with a post comparison by Scheffe test.

[b] Distance to source: Average of home-to-coal-fired power plant and home-to-oil refinery distance, unit: km; Road area surrounding homes unit: m^2; Ambient V unit: ng/m^3; Polycyclic Aromatic Hydrocarbons unit: ng/m^3.

[c] Children low exposure group n = 35; Elderly low exposure group n = 50.

[d] For 1-hydroxypyrene, unit: μmol/mol-creatinine; for heavy metals, unit: μg/g-creatinine.

Fig. 2 compared log-transformed urine concentrations of four oxidative stress biomarkers between high and low exposure groups in children and elderly subjects. In children, urine concentrations of 8-OHdG were 3.1 ± 2.52 and 2.59 ± 2.78 μg/g-creatinine in high and low exposure groups, respectively, with no significant statistical difference. For elderly subjects, 8-OHdG urine levels were 6.61 ± 20.34 and 3.16 ± 4.07 μg/g-creatinine for high and low exposure groups, respectively, with p value of 0.006 (Fig. 2A). Urine levels of HNE-MA were significantly increased in high exposure groups compared to low exposure groups for both children and elderly subjects (Fig. 2B), with urine concentrations 2.16 ± 2.7 and 1.4 ± 2.3 μg/g-creatinine for high and low exposure groups, respectively, in children, and 2.59 ± 3.16 and 1.82 ± 3.66 μg/g-creatinine for high and low exposure groups, respectively, in elderly subjects. 8-isoPGF$_{2\alpha}$ was increased when comparing high (3.22 ± 3.4 μg/g-creatinine) to low exposure group (2.06 ± 2.14 μg/g-creatinine) in children, but the difference was not as significant in elderly subjects, with urine concentrations of high exposure group at 2.88 ± 2.94 μg/g-creatinine, and low exposure group at 2.47 ± 4.32 μg/g-creatinine (Fig. 2C). Fig. 2D showed 8-NO$_2$Gua was significantly increased when comparing high and low exposure groups for both children and elderly subjects. Urine concentrations were 6.88 ± 11.93 and 2.43 ± 2.97 μg/g-creatinine for the high and low exposure groups, respectively,

in children, while elderly subjects had 7.44 ± 15.21 and 3.19 ± 3.34 μg/g-creatinine for high and low exposure groups, respectively.

Urine metabolomics identified 405 potential metabolite peaks in each individual's urine sample in children, and 391 in elderly participants, 216 and 209 of which were identified as known human metabolites by HMDB in children and elderly participants, respectively (see Table S2). PLS-DA results showed separation between the urine metabolite profiles of children and elderly subjects, as well as high and low exposure groups, while no significant difference was found between metabolite profiles of different genders (see Table S3). Separate analysis of urine metabolite profiles in children and elderly participants showed clear separation of urine metabolite profiles between high and low exposure groups in both children (Fig. 3A) and elderly subjects (Fig. 3B). Permutation test confirmed the validity of the PLS-DA models (see Fig. S4A and S4B), and cross-validation test results showed that for both children and elderly subjects, best performance of the PLS-DA models were acquired after applying three components (see Fig. S4C and S4D). The variable importance in the projection (VIP) value of each potential metabolite peak in first three components was listed in Table S2. Potential metabolite peaks with Component 1 VIP score > 1 were considered responsible for the separation between high and low exposure groups, 45 of these exposure-related potential metabolites
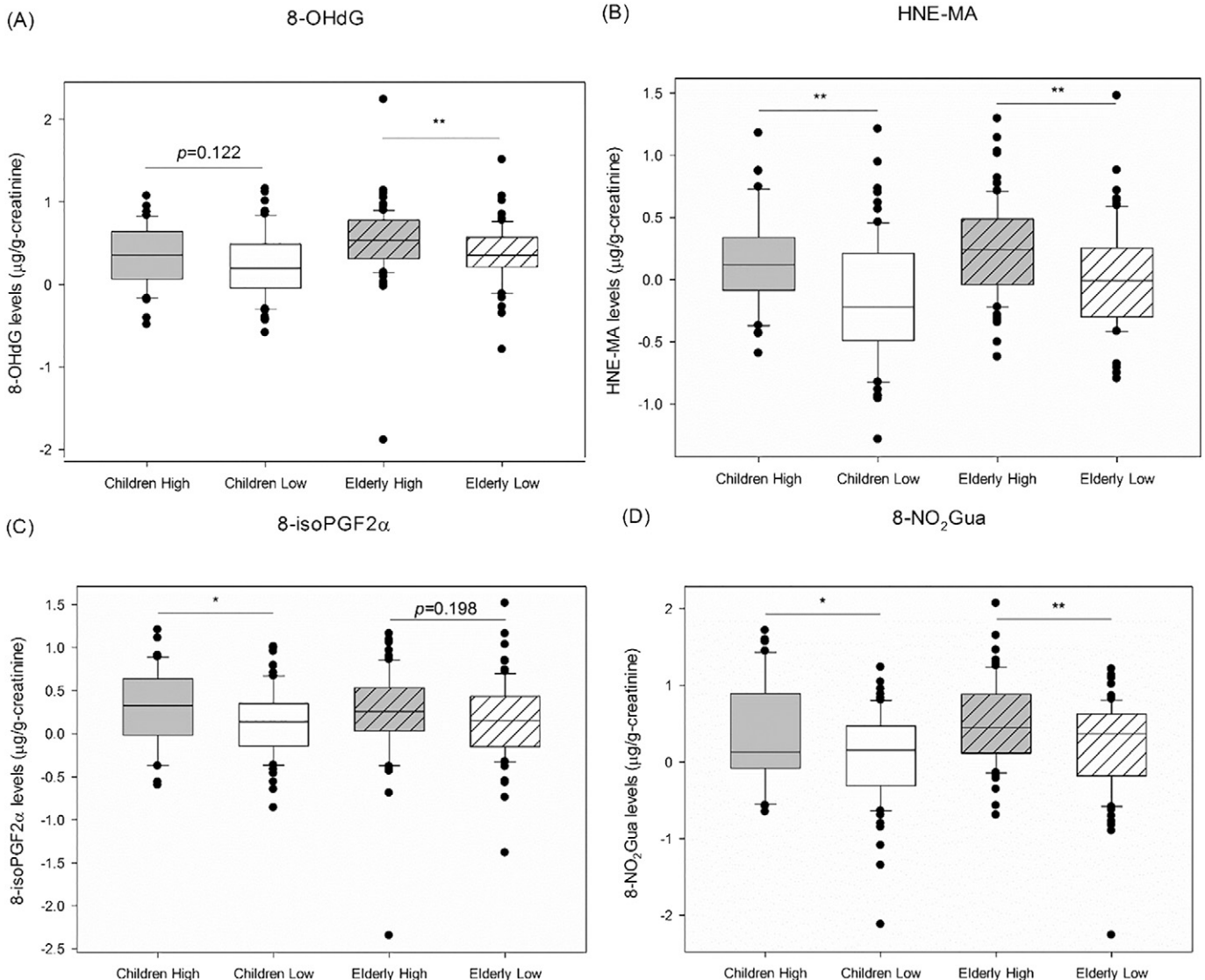

Fig. 2. Urine concentrations of (A) oxidative DNA damage biomarker 8-OHdG (B) (C) lipid peroxidation biomarkers HNE-MA and 8-isoPGF$_{2\alpha}$ and (D) nitrative DNA damage biomarker 8-NO$_2$Gua levels between high and low exposure groups in children and elderly study subjects. $*p < 0.05$, $**p < 0.01$.
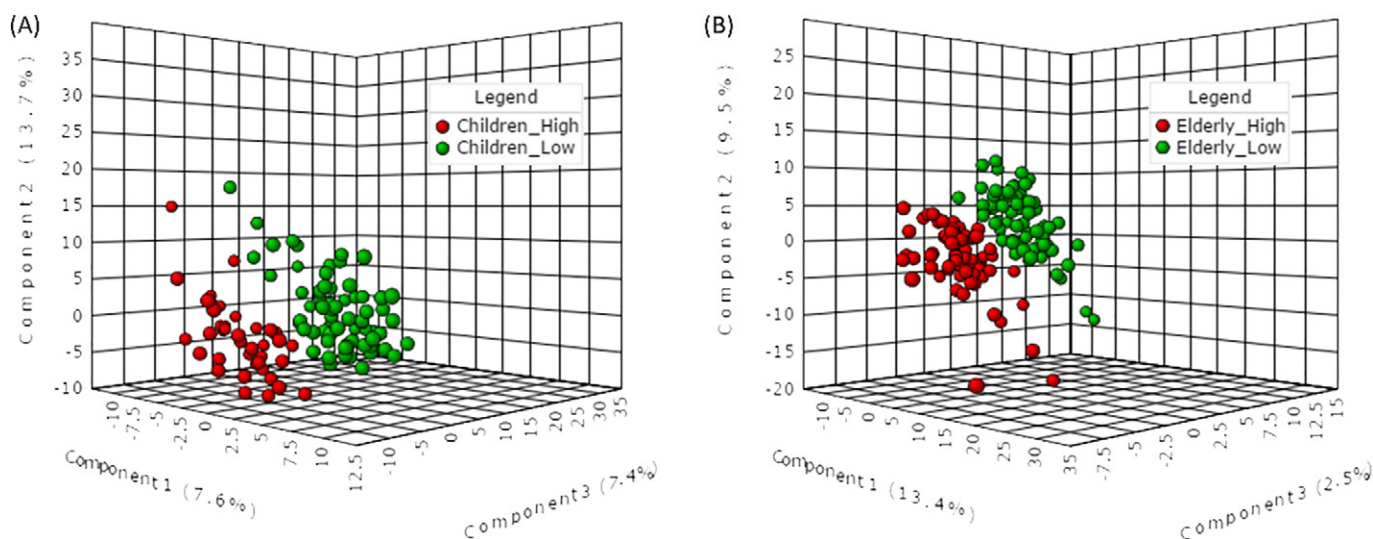
Fig. 3. PLS-DA score plots of urinary metabolite peaks from (A) children ($R^2 = 0.826$, $Q^2 = 0.494$, Permutation $p = 0.04$) and (B) elderly study subjects ($R^2 = 0.823$, $Q^2 = 0.278$, Permutation $p = 0.02$).

were identified in children (fdr q < 0.1, $p \leq 0.033$), and 42 in elderly subjects (fdr q < 0.15, $p \leq 0.040$), of which only 11 were found in both children and elderly subjects. These 76 potential metabolites were classed by HMDB as 9 benzenoids, 3 hydrocarbons, 10 lipids and lipid-like molecules, 19 organic acids and derivatives, 14 organoheterocyclic compounds, 2 organonitrogen compounds, 15 organooxygen compounds, 1 organosulfur compound, 2 phenylpropanoids and polyketides, and 1 homogeneous non-metal compound (Table 2).

Pathway enrichment and pathway topology analysis on the 45 exposure-related potential metabolites in children identified three main biological pathways potentially affected by multiple exposures: alanine, aspartate, and glutamate metabolism, phenylalanine metabolism, and tryptophan metabolism. For the 42 potential metabolites found in elderly participants, pathway analysis showed glycine, serine, and threonine metabolism, alanine, aspartate, and glutamate metabolism, as well as aminoacyl-tRNA biosynthesis were the most important pathways disrupted by multiple exposures (Impact value > 0.1, FDR adjusted $p < 0.05$) (see Table S4). Through the Comparative Toxicogenomics Database (CTD), we found that pyrene, fluoranthene, As, Cu, Cd, and Ni were significantly associated with tryptophan metabolism, As, Cd, and Ni with phenylalanine metabolism, dibenzo[a,h]anthracene, As, Cu, Cd, Ni, and Hg with alanine, aspartate, and glutamate metabolism, and As, Cu, Cd, and Ni with glycine, serine, and threonine metabolism (Bonferroni adjusted $p < 0.01$) (Fig. 4) (Davis et al., 2015).

We also identified 163 oxidative stress-related potential metabolites in children subjects, and 144 in elderly subjects that were associated with at least one of the four oxidative stress biomarkers ($p < 0.05$) (see Table S2). In children, we found 54 potential metabolites associated with 8-OHdG, 108 with HNE-MA, 69 with 8-isoPGF$_{2\alpha}$, and 56 with 8-NO$_2$Gua, while in elderly subjects, 33 were correlated with 8-OHdG, 113 with HNE-MA, 65 with 8-isoPGF$_{2\alpha}$, and 49 with 8-NO$_2$Gua. Identified oxidative stress-related potential metabolites showed more overlap (113 out of 194) than multiple exposure-related potential metabolites (11 out of 76) between children and elderly subjects.

Through the "Meet-in-the-middle" approach, we identified a profile of putative intermediate biomarkers that were associated with both exposures and oxidative stress biomarkers in children participants: 10 for 8-OHdG, 23 for HNE-MA, 11 for 8-isoPGF$_{2\alpha}$, and 22 for 8-NO$_2$Gua. We also identified another profile of such biomarkers in elderly subjects: 17 for 8-OHdG, 32 for HNE-MA, 10 for 8-isoPGF$_{2\alpha}$, and 26 for 8-NO$_2$Gua (see Table S5). Fig. 4 showed that these putative intermediate

biomarkers could be located in previously identified exposures-related pathways. In children, tryptophan and indole-3-acetamide could be located in tryptophan metabolism pathway, phenylalanine, hippuric acid, and 4-hydroxy benzoic acid in phenylalanine metabolism pathway, and succinic acid in both phenylalanine metabolism and alanine, aspartate, and glutamate metabolism pathways. In elderly participants, threonine, serine, and glyceric could be located in glycine, serine, and threonine metabolism. Aspartic acid, which was identified in both children and elderly subjects, is involved in alanine, aspartate, and glutamate metabolism.

4. Discussion

Here we demonstrate a urine metabolomic approach to evaluate how a mixture of air toxics from an industrial emission source can increase the external and internal exposures of nearby residents in a distance-to-source-related manner, inducing age-dependent responses in children and elderly residents that led to elevated oxidative stress. The strength of this study was that we comprehensively evaluated multiple air toxics, and successfully used untargeted metabolomics to find the association between exposures and early health effect oxidative stress through the "Meet-in-the-middle" approach (Fig. 4).

External and internal exposure results suggest petrochemical complex is the major source of both external and internal exposures to multiple air toxics for our study subjects (Table 1). High exposure groups lived closer to the emission sources, had higher levels of ambient V and PAHs at the location of their home addresses, with increased urinary concentrations of heavy metals and PAHs metabolite 1-OHP, of which PAHs, V, Ni, As, and Cu exposures have been reported near both coal-fired power plants and oil refineries, and Cd, Hg, Sr, and Tl near coal-fired power plants (Dybing et al., 2013; George et al., 2015; IARC, 1989a; O'Rourke and Connolly, 2003; Peter and Viraraghavan, 2005). Lack of difference in surrounding road area at home addresses of high and low exposure groups suggest limited traffic influence on exposures.

Urine oxidative stress biomarkers were elevated in high exposure groups, including 8-OHdG, a biomarker for oxidative DNA damage, HNE-MA, produced by lipid peroxidation, 8-isoPGF$_{2\alpha}$, a metabolite from arachidonic acid peroxidation, and 8-NO$_2$Gua, potential biomarker for nitrative DNA damage (Fig. 2) (Wu et al., 2016). Previous studies have reported exposure to PAHs and heavy metals V, Ni, Cu, As, Cd, and Hg can induce oxidative stress through increment of reactive

Table 2

Urine metabolic profiling of multiple exposures from refineries and coal-fired power plants in children and elderly subjects using GCxGC-TOFMS analysis.

Metabolite identification[a]	Involved pathway[b]	Trend[c]
Children[d]		
Benzenoids		
2-Phenylpropanal	–	Up
3-Hydroxyhippuric acid	–	Down
4-Hydroxybenzoic acid	Phenylalanine metabolism	Down
Hippuric acid	Phenylalanine metabolism	Down
o-Cymene	–	Up
Hydrocarbons		
Decane	–	Up
Dodecane	None	Up
Tridecane	None	Up
Lipids and lipid-like molecules		
Azelaic acid	None	Down
Glycerol 3-phosphate	Glycerolipid metabolism	Up
Myristic acid	Fatty acid biosynthesis	Down
Stearic acid	Fatty acid biosynthesis	Up
Tiglic acid	None	Up
Organic acids and derivatives		
3-Hydroxyoctanoic acid	cAMP signaling pathway	Down
Aminomalonic acid	None	Up
Fumaric acid	Citrate cycle (TCA cycle)	Up
Hydroxypyruvic acid	Glycine, serine, and threonine metabolism	Up
L-Alpha-aminobutyric acid	Cysteine and methionine metabolism	Up
L-Aspartic acid	Alanine, aspartate, and glutamate metabolism	Down
L-Histidine	Histidine metabolism	Down
N-acetylglutamic acid	Arginine biosynthesis	Down
Succinic acid	Citrate cycle (TCA cycle)	Up
Tiglylglycine	–	Down
γ-Aminobutyric acid	Alanine, aspartate, and glutamate metabolism	Down
Organoheterocyclic compounds		
1H-Indole-3-acetamide	Tryptophan metabolism	Up
2-Deoxy-L-ribono-1,4-lactone	–	Down
4-Pyridoxic Acid	Vitamin B6 metabolism	Down
Cyanuric acid	None	Up
DL-Tryptophan	Tryptophan metabolism	Down
L-Gulonolactone	Tryptophan metabolism	Down
Quinolinic acid	Ascorbate and aldarate metabolism	Down
Sumiki's acid	Tryptophan metabolism	Down
Organonitrogen compounds		
Dimethylamine	Methane metabolism	Up
Organooxygen compounds		
4-Deoxyerythronic acid	–	Up
Cyanic acid	Nitrogen metabolism	Down
Cyclohexanone	Caprolactam degradation	Up
Diacetone alcohol	–	Up
D-Threitol	None	Up
Gluconic acid	Pentose phosphate pathway	Down
Hex-2-ulosonic acid	None	Down
L-Sorbose	None	Up
Rhamnose	Fructose and mannose metabolism	Down
Threonic acid	Ascorbate and aldarate metabolism	Down
Phenylpropanoids and polyketides		
Hydroxyphenyllactic acid	Tyrosine metabolism	Up
L-Phenylalanine	Phenylalanine metabolism	Up
Elderly[e]		
Benzenoids		
Catechol	Chlorocyclohexane and chlorobenzene degradation	Up
m-Cresol	Toluene degradation	Up
Mono(2-ethylhexyl)phthalate [MEHP]	None	Up
o-Cymene	–	Up
Phenol	Tyrosine metabolism	Up
Homogeneous non-metal compounds		
Phosphoric acid	Oxidative phosphorylation	Up
Hydrocarbons		
Decane	–	Up
Dodecane	None	Up

Table 2 (continued)

Metabolite identification[a]	Involved pathway[b]	Trend[c]
Tridecane	None	Up
Lipids and lipid-like molecules		
2,4-Dihydroxybutanoic acid	–	Up
2-Hydroxyglutaric acid	None	Up
Borneol	None	Down
L-Threonine	Glycine, serine and threonine metabolism	Up
Myristic acid	Fatty acid biosynthesis	Up
Palmitic acid	Fatty acid biosynthesis	Up
Stearic acid	Fatty acid biosynthesis	Up
Organic acids and derivatives		
(S)-3-Hydroxyisobutyric acid	Valine, leucine and isoleucine degradation	Up
2-Ethylhydracrylic acid	–	Up
Alanine	Alanine, aspartate and glutamate metabolism	Up
Glutaric acid	Fatty acid degradation	Up
L-Aspartic acid	Alanine, aspartate and glutamate metabolism	Down
Leucine	Valine, leucine and isoleucine degradation	Up
L-Valine	Valine, leucine and isoleucine degradation	Up
Serine	Glycine, serine and threonine metabolism	Up
Thiodiacetic acid	Metabolism of xenobiotics by cytochrome P450	Up
Organoheterocyclic compounds		
5-Hydroxyindoleacetic acid	Tryptophan metabolism	Down
Cytosine	Pyrimidine metabolism	Up
D-Xylono-1,5-lactone	Pentose and glucuronate interconversions	Up
Hypoxanthin	Purine metabolism	Down
Picolinic acid	Tryptophan metabolism	Up
Uracil	Pyrimidine metabolism	Up
Organonitrogen compounds		
Diethanolamine	Glycerophospholipid metabolism	Up
Organooxygen compounds		
1,3-Butanediol	–	Up
4-Deoxyerythronic acid	–	Up
Acetoin	None	Up
Arabinose	–	Up
Cyclohexanone	Caprolactam degradation	Up
Glyceric acid	Pentose phosphate pathway	Up
Inositol	–	Up
L-Sorbose	None	Up
Threonic acid	Ascorbate and aldarate metabolism	Up
Organosulfur compounds		
Dimethyl sulfone	Sulfur metabolism	Up

[a] Urine metabolites detected by GCxGC-TOFMS, identified using NIST library, and found in and classified via Human Metabolome Database.

[b] Metabolites were searched in KEGG database for involved biological pathways. If metabolite is involved in multiple pathways, only one is shown. None: Found in KEGG database but with no known involved pathways. —: Not found in KEGG database.

[c] The up- or downregulation of metabolites in high exposure group compared to low exposure group.

[d] VIP > 1, FDR q < 0.1.

[e] VIP > 1, FDR q < 0.15.

oxygen species (ROS) and/or reduction of anti-oxidants (Jomova and Valko, 2011; Penning and Drury, 2007). Our findings suggest that exposure to multiple air toxics induces oxidative stress, an early health effect that contributes to numerous common complex chronic diseases such as cancer, cardiovascular disease, diabetes, and neurodegenerative diseases, as well as acute respiratory diseases such as allergic rhinitis and asthma (Bowler and Crapo, 2002; Reuter et al., 2010).

Potential metabolites responsible for the separation between high and low exposure groups were different for children and elderly subjects (Table 2), and pathway analysis results showed three different and one common biological pathway affected by exposures in the two age groups (see Table S4). "Meet-in-the-middle" approach found putative intermediate biomarkers that were involved in these four pathways (see Table S5), suggesting that exposures can disrupt diverse biological
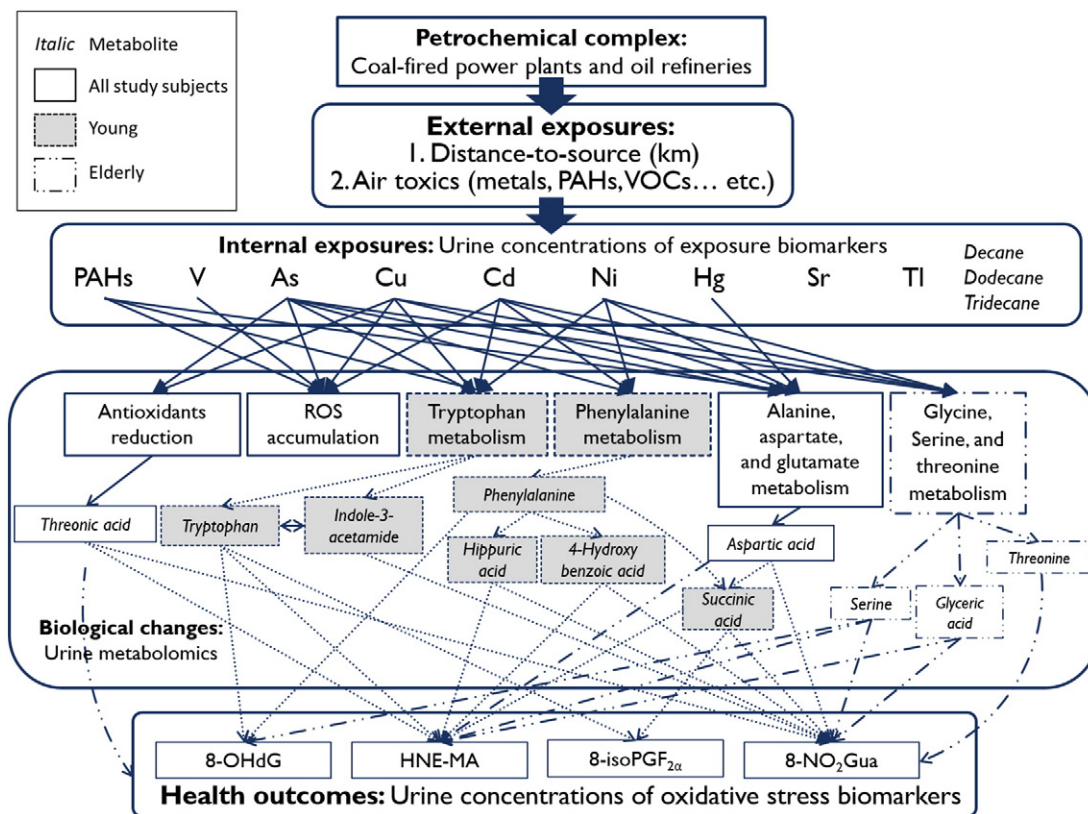
Fig. 4. Exposure pathways of petrochemical air pollution and the effects on urine metabolic profile changes and increased oxidative stress.

mechanisms in different age groups, inducing common early health effect oxidative stress (Fig. 4).

Tryptophan metabolism was one of the exposures-related pathways identified in children subjects (see Table S4). Tryptophan can be metabolized through different pathways, the most important route, kynurenine pathway, is often upregulated by activated immune responses, leading to the depletion of tryptophan, and has shown involvement in increased oxidative stress as well as numerous diseases including cancer, neurodegenerative diseases, and allergic disorders such as rhinitis and asthma (Chen and Guillemin, 2009; Ciprandi et al., 2010; Gostner et al., 2016; Stoy et al., 2005). Our previous study in the same area also showed increased incidence of allergic rhinitis, bronchitis, and asthma in children living near the petrochemical complex (Chiang et al., 2016). Tryptophan was downregulated in high exposure group compared to low exposure group in children subjects (Table 2), identified as a putative intermediate biomarker and correlated with 8-OHdG, HNE-MA, and 8-isoPGF$_{2\alpha}$ (see Table S5). Downstream metabolite 1H-indole-3-acetamide from another tryptophan metabolism route was also identified and associated with 8-NO$_2$Gua. These findings suggest exposures could affect at least one tryptophan metabolism pathway, inducing multiple oxidative stress outcomes. Given the mediating roles of tryptophan metabolism and oxidative stress in allergic respiratory diseases (Bowler and Crapo, 2002; Ciprandi et al., 2010; Gostner et al., 2016), our findings may provide further information on the biological mechanisms deregulated by petrochemical industry exposures that led to increased risks in children. Phenylalanine metabolism was also identified in children subjects (see Table S4). Phenylalanine, which have been used as an oxidative stress biomarker (Orhan et al., 2004), was significantly correlated with exposures and oxidative stress biomarkers 8-OHdG and HNE-MA, and with HNE-MA, 8-isoPGF$_{2\alpha}$, and 8-NO$_2$Gua through its downstream metabolites hippuric acid, 4-hydroxy benzoic acid, and succinic acid in children participants (see Table S5). Since the increase of 8-OHdG levels in high exposure group for children participants increased but did not reach statistical

significance (Fig. 2A), our findings suggest that urinary tryptophan and phenylalanine could be a potential intermediate biomarkers of exposure-induced oxidative stress in children study subjects, before DNA damage became significant. In elderly subjects, glycine, serine, and threonine metabolism was identified (see Table S4), with related-compounds threonine associating with 8-NO$_2$Gua, serine with 8-OHdG, HNE-MA, and 8-NO$_2$Gua, and glyceric acid with HNE-MA and 8-NO$_2$Gua (see Table S5). The disruption of this pathway is closely related to oncogenic transformation, and the biosynthesis of antioxidant glutathione (Amelio et al., 2014). Alanine, aspartate, and glutamate metabolism was identified in both children and elderly participants (see Table S4), with aspartic acid downregulated in high exposure subjects of both age groups (Table 2), associated with 8-NO$_2$Gua in children participants, and with HNE-MA in elderly participants (see Table S5). Previous studies have shown that aspartic acid could increase glutathione levels and decrease lipid peroxidation in animal models (Sivakumar et al., 2011). Threonic acid was also identified in both age groups as a putative intermediate biomarker, indicating deregulation of its precursor, antioxidant ascorbic acid (Gao et al., 2012), associating multiple exposures with HNE-MA and 8-NO$_2$Gua in children and all four oxidative stress biomarkers in elderly participants (see Table S5). We can draw from these results a complicated web showing the relation between exposures and different oxidative stress induced health effects through age-dependent diverse biological pathways (Fig. 4).

Accidentally but not unexpectedly, we also found some exposure biomarkers in our untargeted urine analysis of metabolomes. Decane, dodecane, and tridecane were elevated in high exposure groups for both children and elderly subjects (Table 2). These compounds are intermediates in petrochemical industrial productions, and were previously reported as potential health risks (IARC, 1989b; Williams et al., 2006). We did not locate decane, dodecane, or tridecane in human biological pathway analysis, suggesting these compounds are from external sources, and can be used as exposure biomarkers of petrochemical emissions (Fig. 4).

doi:10.6342/NTU201901261

There are limitations to the present study. First, biomarkers measured in one spot urine may have daily variability, and may not be as stable in reflecting exposure to heavy metals as those measured in other biospecimens such as hair, which unfortunately was not collected in this study. Nevertheless, we believe this would not systemically bias our exposure classification because distance-to-source and model-based ambient concentrations all showed the same pattern as urine biomarkers. In addition, the petrochemical complex has had continuous emissions since operation started in 1999, and our study subjects have lived in the area for at least five years. Secondly, the air toxics in this study were limited to PAHs and heavy metals for which we performed biomonitoring. SO_2, NO_x, black carbon, and other toxics whose biomarkers were not measured in this study may also contribute to the increase of oxidative stress. Thirdly, we could not rule out routes of exposure other than air that could induce oxidative stress, such as water and food. Lastly, we could not rule out the possibility of inaccurate metabolite identification by library match.

5. Conclusion

Urine exposure biomarkers of PAHs and heavy metals V, Ni, As, Cu, Sr, Cd, Hg, and Tl were elevated in children and elderly residents living near a petrochemical complex. These internal exposures were associated with model-estimated ambient concentrations at residential addresses, and could possibly be traced to air toxics emitted by oil refineries and coal-fired power plants within the complex. Both children and elderly residents living in the pollution-affected area with higher levels of urine exposure biomarkers showed changes in urine metabolite profiles which could be linked to increased oxidative stress, including oxidative and nitrative DNA damage, and lipid peroxidation. We conclude that urine metabolomics could possibly serve as the link to trace multiple air toxics exposure to oxidative stress through age-specific biological pathways including tryptophan metabolism and phenylalanine metabolism in children subjects, glycine, serine, and threonine metabolism in elderly subjects, and alanine, aspartate, and glutamate metabolism in both age groups. The identified exposures and metabolic pathways will improve risk assessments on developing common complex chronic diseases, such as cancers and cardiovascular diseases, as well as allergic respiratory diseases, such as allergic rhinitis and asthma, for the residents living near the petrochemical complex if air toxics exposure continues in the future. We recommend to significantly reduce air toxics emissions from the petrochemical complex to lower residents' health risks. Our findings also warrant a follow-up study on residents who continue to be affected by petrochemical pollution.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.envint.2017.02.003.

References

Adler, T., 2003. Aging research: the future face of environmental health. Environ. Health Perspect. 111, A760–A765.

Al-Rmalli, S.W., Jenkins, R.O., Haris, P.I., 2011. Betel quid chewing elevates human exposure to arsenic, cadmium and lead. J. Hazard. Mater. 190, 69–74.

Amelio, I., Cutruzzola, F., Antonov, A., Agostini, M., Melino, G., 2014. Serine and glycine metabolism in cancer. Trends Biochem. Sci. 39, 191–198.

Bowler, R.P., Crapo, J.D., 2002. Oxidative stress in allergic respiratory diseases. J. Allergy Clin. Immunol. 110, 349–356.

Chadeau-Hyam, M., Athersuch, T.J., Keun, H.C., De Iorio, M., Ebbels, T.M., Jenab, M., et al., 2011. Meeting-in-the-middle using metabolic profiling—a strategy for the identification of intermediate biomarkers in cohort studies. Biomarkers 16, 83–88.

Chan, C.C., Shie, R.H., Chang, T.Y., Tsai, D.H., 2006. Workers' exposures and potential health risks to air toxics in a petrochemical complex assessed by improved methodology. Int. Arch. Occup. Environ. Health 79, 135–142.

Chan, E.C., Pasikanti, K.K., Nicholson, J.K., 2011. Global urinary metabolic profiling procedures using gas chromatography-mass spectrometry. Nat. Protoc. 6, 1483–1499.

Chen, Y., Guillemin, G.J., 2009. Kynurenine pathway metabolites in humans: disease and healthy states. Int. J. Tryptophan. Res. 2, 1–19.

Chiang, T.Y., Yuan, T.H., Shie, R.H., Chen, C.F., Chan, C.C., 2016. Increased incidence of allergic rhinitis, bronchitis and asthma, in children living near a petrochemical complex with SO_2 pollution. Environ. Int. 96, 1–7.

Chio, C.P., Yuan, T.H., Shie, R.H., Chan, C.C., 2014. Assessing vanadium and arsenic exposure of people living near a petrochemical complex with two-stage dispersion models. J. Hazard. Mater. 271, 98–107.

Ciprandi, G., De Amici, M., Tosca, M., Fuchs, D., 2010. Tryptophan metabolism in allergic rhinitis: the effect of pollen allergen exposure. Hum. Immunol. 71, 911–915.

Davis, A.P., Grondin, C.J., Lennon-Hopkins, K., Saraceni-Richards, C., Sciaky, D., King, B.L., et al., 2015. The comparative toxicogenomics database's 10th year anniversary: update 2015. Nucleic Acids Res. 43, D914–D920.

Driscoll, C.T., Buonocore, J.J., Levy, J.I., Lambert, K.F., Burtraw, D., Reid, S.B., et al., 2015. Us power plant carbon standards and clean air and health co-benefits. Nat. Clim. Chang. 5, 535–540.

Dybing, E., Schwarze, P.E., Nafstad, P., Victorin, K., Penning, T.M., 2013. Polyclyclic Aromatic Hydrocarbons in Ambient Air and Cancer. (Chapter 7). International Agency for Research on Cancer, France.

Ellis, J.K., Athersuch, T.J., Thomas, L.D., Teichert, F., Perez-Trujillo, M., Svendsen, C., et al., 2012. Metabolic profiling detects early effects of environmental and lifestyle exposure to cadmium in a human population. BMC Med. 10, 61.

Gao, X., Chen, W., Li, R., Wang, M., Chen, C., Zeng, R., et al., 2012. Systematic variations associated with renal disease uncovered by parallel metabolomics of urine and serum. BMC Syst. Biol. 6 (Suppl. 1), S14.

Gao, Y., Lu, Y., Huang, S., Gao, L., Liang, X., Wu, Y., et al., 2014. Identifying early urinary metabolic changes with long-term environmental exposure to cadmium by mass-spectrometry-based metabolomics. Environ. Sci. Technol. 48, 6409–6418.

George, J., Masto, R.E., Ram, L.C., Das, T.B., Rout, T.K., Mohan, M., 2015. Human exposure risks for metals in soil near a coal-fired power-generating plant. Arch. Environ. Contam. Toxicol. 68, 451–461.

Gostner, J.M., Becker, K., Kofler, H., Strasser, B., Fuchs, D., 2016. Tryptophan metabolism in allergic disorders. Int. Arch. Allergy Immunol. 169, 203–215.

Hastings, J., de Matos, P., Dekker, A., Ennis, M., Harsha, B., Kale, N., et al., 2013. The chebi reference database and ontology for biologically relevant chemistry: enhancements for 2013. Nucleic Acids Res. 41, D456–D463.

IARC, 1989a. Occupational exposures in petroleum refining; crude oil and major petroleum fuels. Iarc working group on the evaluation of carcinogenic risks to humans. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans/World Health Organization. 45. International Agency for Research on Cancer, pp. 1–322.

IARC, 1989b. Cyclohexanone. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans/World Health Organization. 47. International Agency for Research on Cancer, pp. 157–169.

Jomova, K., Valko, M., 2011. Advances in metal-induced oxidative stress and human disease. Toxicology 283, 65–87.

Juarez, P.D., Matthews-Juarez, P., Hood, D.B., Im, W., Levine, R.S., Kilbourne, B.J., et al., 2014. The public health exposome: a population-based, exposure science approach to health disparities research. Int. J. Environ. Res. Public Health 11, 12866–12895.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., Tanabe, M., 2014. Data, information, knowledge and principle: back to metabolism in kegg. Nucleic Acids Res. 42, D199–D205.

Makri, A., Stilianakis, N.I., 2008. Vulnerability to air pollution health effects. Int. J. Hyg. Environ. Health 211, 326–336.

Nadal, M., Schuhmacher, M., Domingo, J.L., 2004. Metal pollution of soils and vegetation in an area with petrochemical industry. Sci. Total Environ. 321, 59–69.

Orhan, H., Vermeulen, N.P., Tump, C., Zappey, H., Meerman, J.H., 2004. Simultaneous determination of tyrosine, phenylalanine and deoxyguanosine oxidation products by liquid chromatography-tandem mass spectrometry as non-invasive biomarkers for oxidative damage. J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 799, 245–254.

O'Rourke, D., Connolly, S., 2003. Just oil? The distribution of environmental and social impacts of oil production and consumption. Annu. Rev. Environ. Resour. 23, 587–617.

Pan, B.J., Hong, Y.J., Chang, G.C., Wang, M.T., Cinkotai, F.F., Ko, Y.C., 1994. Excess cancer mortality among children and adolescents in residential districts polluted by petrochemical manufacturing plants in Taiwan. J. Toxicol. Environ. Health 43, 117–129.

Patti, G.J., Yanes, O., Siuzdak, G., 2012. Innovation: metabolomics: the apogee of the omics trilogy. Nat. Rev. Mol. Cell Biol. 13, 263–269.

Penning, T.M., Drury, J.E., 2007. Human aldo-keto reductases: function, gene regulation, and single nucleotide polymorphisms. Arch. Biochem. Biophys. 464, 241–250.

Peter, A.L., Viraraghavan, T., 2005. Thallium: a review of public health and environmental concerns. Environ. Int. 31, 493–501.

Rappaport, S.M., Smith, M.T., 2010. Epidemiology. Environment and disease risks. Science 330, 460–461.

Reuter, S., Gupta, S.C., Chaturvedi, M.M., Aggarwal, B.B., 2010. Oxidative stress, inflammation, and cancer: how are they linked? Free Radic. Biol. Med. 49, 1603–1616.

Robertson, D.G., Watkins, P.B., Reily, M.D., 2011. Metabolomics in toxicology: preclinical and clinical applications. Toxicol. Sci. 120 (Suppl. 1), S146–S170.

Shie, R.H., Yuan, T.H., Chan, C.C., 2013. Using pollution roses to assess sulfur dioxide impacts in a township downwind of a petrochemical complex. J. Air Waste Manage. Assoc. 63, 702–711.

Shih, Y.T., Chen, P.S., Wu, C.H., Tseng, Y.T., Wu, Y.C., Lo, Y.C., 2010. Arecoline, a major alkaloid of the areca nut, causes neurotoxicity through enhancement of oxidative stress and suppression of the antioxidant protective system. Free Radic. Biol. Med. 49, 1471–1479.

Sivakumar, R., Babu, P.V., Shyamaladevi, C.S., 2011. Aspartate and glutamate prevents isoproterenol-induced cardiac toxicity by alleviating oxidative stress in rats. Exp. Toxicol. Pathol. 63, 137–142.

Slupsky, C.M., Rankin, K.N., Wagner, J., Fu, H., Chang, D., Weljie, A.M., et al., 2007. Investigations of the effects of gender, diurnal variation, and age in human urinary metabolomic profiles. Anal. Chem. 79, 6995–7004.

Smith, M.T., de la Rosa, R., Daniels, S.I., 2015. Using exposomics to assess cumulative risks and promote health. Environ. Mol. Mutagen. 56, 715–723.

Stoy, N., Mackay, G.M., Forrest, C.M., Christofides, J., Egerton, M., Stone, T.W., et al., 2005. Tryptophan metabolism and oxidative stress in patients with huntington's disease. J. Neurochem. 93, 611–623.

Thevenot, E.A., Roux, A., Xu, Y., Ezan, E., Junot, C., 2015. Analysis of the human adult urinary metabolome variations with age, body mass index, and gender by implementing a comprehensive workflow for univariate and opls statistical analyses. J. Proteome Res. 14, 3322–3335.

Vineis, P., van Veldhoven, K., Chadeau-Hyam, M., Athersuch, T.J., 2013. Advancing the application of omics-based biomarkers in environmental epidemiology. Environ. Mol. Mutagen. 54, 461–467.

Wang, C.C., Liu, T.Y., Wey, S.P., Wang, F.I., Jan, T.R., 2007. Areca nut extract suppresses t-cell activation and interferon-gamma production via the induction of oxidative stress. Food Chem. Toxicol. 45, 1410–1418.

Wichmann, F.A., Muller, A., Busi, L.E., Cianni, N., Massolo, L., Schlink, U., et al., 2009. Increased asthma and respiratory symptoms in children exposed to petrochemical pollution. J. Allergy Clin. Immunol. 123, 632–638.

Wild, C.P., 2005. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research. 14. Cosponsored by the American Society of Preventive Oncology, pp. 1847–1850.

Williams, P.R., Patterson, J., Briggs, D.W., 2006. Vccep pilot: progress on evaluating children's risks and data needs. Risk Analysis: An Official Publication of the Society for Risk Analysis. 26, pp. 781–801.

Wishart, D.S., Jewison, T., Guo, A.C., Wilson, M., Knox, C., Liu, Y., et al., 2013. Hmdb 3.0—the human metabolome database in 2013. Nucleic Acids Res. 41, D801–D807.

Wu, C., Chen, S.T., Peng, K.H., Cheng, T.J., Wu, K.Y., 2016. Concurrent quantification of multiple biomarkers indicative of oxidative stress status using liquid chromatography-tandem mass spectrometry. Anal. Biochem. 512, 26–35.

Xia, J., Sinelnikov, I.V., Han, B., Wishart, D.S., 2015. Metaboanalyst 3.0—making metabolomics more meaningful. Nucleic Acids Res.

Yuan, T.H., Chio, C.P., Shie, R.H., Pien, W.H., Chan, C.C., 2015a. The distance-to-source trend in vanadium and arsenic exposures for residents living near a petrochemical complex. J. Expo. Sci. Environ. Epidemiol.

Yuan, T.H., Shie, R.H., Chin, Y.Y., Chan, C.C., 2015b. Assessment of the levels of urinary 1-hydroxypyrene and air polycyclic aromatic hydrocarbon in pm2.5 for adult exposure to the petrochemical complex emissions. Environ. Res. 136, 219–226.

Metabolomics of Children and Adolescents Exposed to Industrial Carcinogenic Pollutants

Chi-Hsin S. Chen,[†] Tien-Chueh Kuo,[‡,§] Han-Chun Kuo,[‡] Yufeng J. Tseng,[§,∥] Ching-Hua Kuo,[⊥] Tzu-Hsuen Yuan,[†] and Chang-Chuan Chan*,[†]

[†]Institute of Occupational Medicine and Industrial Hygiene, College of Public Health, National Taiwan University, No. 17, Xu-Zhou Road, Taipei 10055, Taiwan

[‡]The Metabolomics Core Laboratory, Center of Genomic Medicine, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan

[§]Graduate Institute of Biomedical Electronics and Bioinformatics, College of Electrical Engineering and Computer Science, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan

[∥]Department of Computer Science and Information Engineering, College of Electrical Engineering and Computer Science, National Taiwan University., No. 1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan

[⊥]School of Pharmacy, College of Medicine, National Taiwan University, No. 33, Linsen S. Road, Taipei 10055, Taiwan

Ⓢ Supporting Information

ABSTRACT: Studies on metabolomes of carcinogenic pollutants among children and adolescents are limited. We aim to identify metabolic perturbations in 107 children and adolescents (aged 9−15) exposed to multiple carcinogens in a polluted area surrounding the largest petrochemical complex in Taiwan. We measured urinary concentrations of eight carcinogen exposure biomarkers (heavy metals and polycyclic aromatic hydrocarbons (PAHs) represented by 1-hydroxypyrene), and urinary oxidative stress biomarkers and serum acylcarnitines as biomarkers of early health effects. Serum metabolomics was analyzed using a liquid chromatography mass spectrometry-based method. Pathway analysis and "meet-in-the-middle" approach were applied to identify potential metabolites and biological mechanisms linking carcinogens exposure with early health effects. We found 10 potential metabolites possibly linking increased exposure to IARC group 1 carcinogens (As, Cd, Cr, Ni) and group 2 carcinogens (V, Hg, PAHs) with elevated oxidative stress and deregulated serum acylcarnitines, including inosine monophosphate and adenosine monophosphate (purine metabolism), malic acid and oxoglutaric acid (citrate cycle), carnitine (fatty acid metabolism), and pyroglutamic acid (glutathione metabolism). Purine metabolism was identified as the possible mechanism affected by children and adolescents' exposure to carcinogens. These findings contribute to understanding the health effects of childhood and adolescence exposure to multiple industrial carcinogens during critical periods of development.
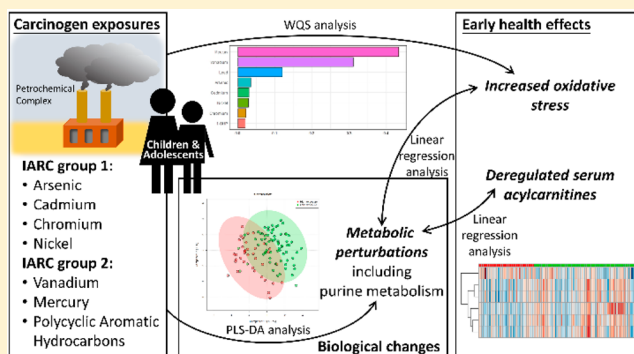
INTRODUCTION

The petrochemical industrial complex is a consortium of high-pollution facilities, such as oil refineries and coal-fired power plants, that emit multiple carcinogenic pollutants including group 1 carcinogens: arsenic (As), cadmium (Cd), chromium (Cr), nickel (Ni), and vinyl chloride (VCM), and group 2 carcinogens: vanadium(V), mercury (Hg), lead (Pb), and polycyclic aromatic hydrocarbons (PAHs), most of which are associated with lung cancer risk.[1−9] Cumulative exposure to such complex chemical mixtures may have synergistic effects on health.[10] Previous studies found residents living near petrochemical industrial complexes have a higher risk of lung cancer compared to those living farther away, and 972 lung cancer cases are attributable to residential exposure to petrochemical industrial complexes each year in 22 EU countries.[11] In Taiwan, we conducted extensive environmental and epidemiological studies near the No. 6 Naphtha Cracking Plant, the largest petrochemical complex in Taiwan, and found a distance-to-source trend in elevated carcinogenic exposures and increased incidence of all cancers in adult residents living in the vicinity of the complex.[12,13]

In 2016, the United Nations released a special report reiterating the systemic problem of childhood exposure to industrial chemicals all around the world, leading to a "silent pandemic" of diseases and disabilities that impacts children's
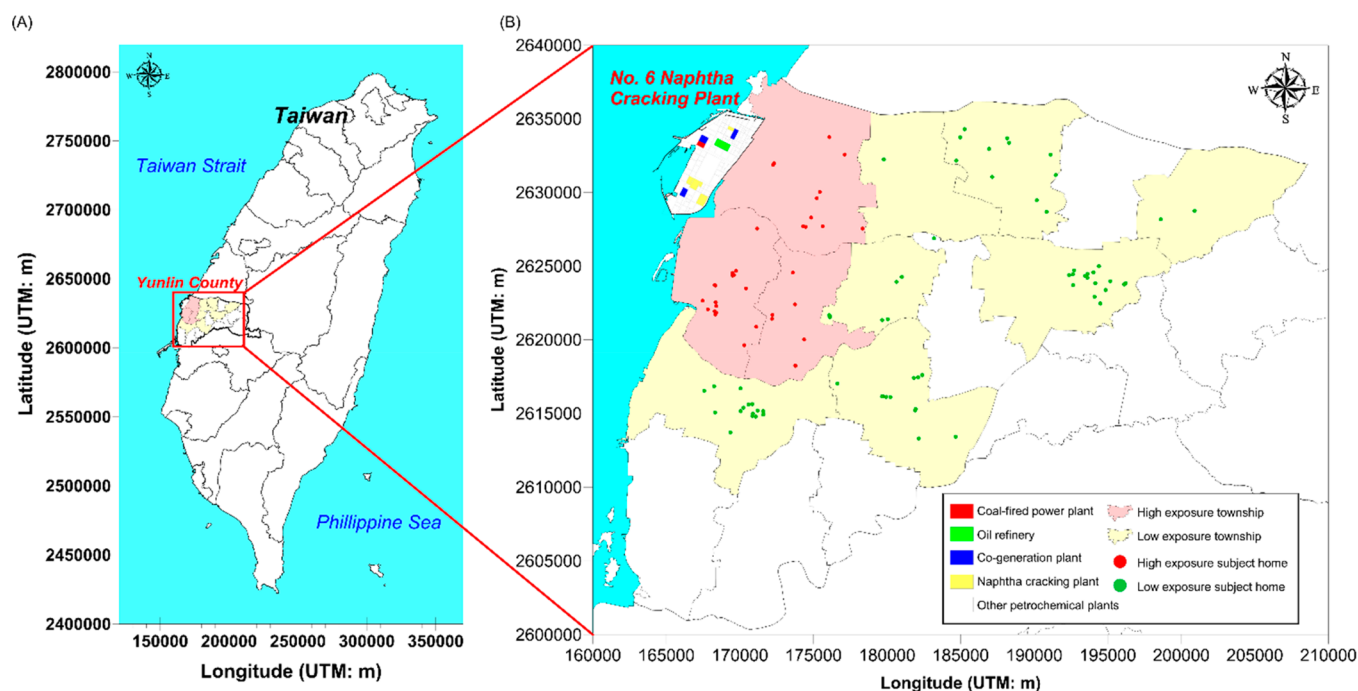
Figure 1. GIS map of (A) Yunlin County in central Taiwan and (B) location of study area, petrochemical plants, and 107 study subjects' homes.

health, development, and life.[14] Children are more vulnerable to environmental exposures due to lifestyle and physiological conditions compared to adults, and they have time for chronic adverse health effects, such as cancer, to manifest.[15] Juvenile animal bioassays also reported earlier onset and increased incidence of cancer in response to carcinogen exposure compared to mature animal models.[16] Previous studies showed children and adolescents (aged 11−14) living near the No. 6 Naphtha Cracking Plant had increased incidence of respiratory diseases associated with SO_2 exposure and subclinically abnormal liver functions associated with elevated VCM exposure.[17,18] However, there are limited studies on cancer risks of children and adolescents living near petrochemical complexes. Pan et al. found excess cancer mortality in 10−19 year old children and adolescents that potentially have been exposed to petrochemical pollutants since birth.[19] Weng et al. showed 0−19 year old children and adolescents who lived in areas with higher petrochemical air pollution exposure index had statistically increased risk of leukemia deaths compared to those living in areas with a lower index.[20] Comprehensive evaluation of children and adolescents living near petrochemical industrial areas is warranted, including their personal carcinogens exposures and related biological disruptions that may accumulate and contribute to cancer development later in life.

The concept of exposome was first proposed in 2005 and refers to all exposures throughout a lifetime and the contribution to disease causation or progression.[21,22] Recently, metabolomics was proposed to quantitatively measure exogenous chemicals and biological responses in order to provide "a snapshot measure of an individual's exposome".[23] Vineis et al. also proposed using metabolomics to investigate the intermediate biomarkers and biological pathways linking exposures to disease risk.[24] Metabolites are the end point of biochemical activities and could best reflect the effects of exposures and correlate with phenotype.[25] The metabolomics approach has also been widely used in clinical cancer research

to investigate the pathogenesis of cancer with limited studies focusing on children and adolescents.[26]

We had previously analyzed urine metabolomics of children and adolescents living near the No. 6 Naphtha Cracking Plant, and identified three biological pathways: tryptophan metabolism, phenylalanine metabolism, and alanine, aspartate, and glutamate metabolism, that linked increased multiple air toxics exposure with elevated oxidative stress.[27] Deregulations in tryptophan metabolism has been reported in lung cancer, colorectal cancer, and breast cancer.[28] This study analyzed serum metabolomics of children and adolescents (aged 9−15 years) who lived in the same petrochemical industrial area for more than five years to identify biological perturbations linking multiple carcinogen exposure with cancer-related early health effects including oxidative stress and serum acylcarnitines. Oxidative stress interacts with all three stages of the cancer process: cancer initiation, cancer promotion, and cancer progression through reactive oxygen species (ROS) and nitrogen reactive species (RNS) induced DNA damage, lipid peroxidation, and protein damage.[29,30] Serum acylcarnitines are involved in transporting fatty acids into the mitochondria for β-oxidation and production of energy.[31] Deregulations in serum acylcarnitines can activate inflammatory signaling pathways, and have been associated with chronic diseases including cancer.[32,33]

■ MATERIAL AND METHODS

Study Area and Subjects. Our study area surrounded the largest petrochemical complex in Taiwan, No. 6 Naphtha Cracking Plant, located in Yunlin County on the west coast of central Taiwan. The complex began major operations in 1999. To date, the complex covers a total area of 2603 ha, housing 64 plants including one coal-fired power plant that generates 1800 MW of power, three oil refineries that processes 450 000 barrels of crude oil every day, two naphtha cracking plants that produce 160 million tons of ethylene per year, and three cogeneration plants that generate 2820 MW of power.[34] Our

subjects were selected from a prospective cohort of 680 children and adolescents (aged 9−15 years) who have lived in three townships in close vicinity to the complex (pink) and seven other townships further away (yellow) as shown in Figure 1 for more than five years. All 680 have completed interview-administered questionnaire surveys including key factors related to exposure, provided one morning spot urine sample, and one fasting blood sample. Urine samples were stored at −20 °C, serum samples were extracted from coagulated blood samples using a centrifuge and stored at −80 °C. All 680 participants' urine samples have been analyzed for V and PAHs exposure biomarker 1-hydroxypyrene (1-OHP). We had previously established urinary V and 1-OHP as exposure biomarkers specific to petrochemical industrial emissions in our study area, with statistically significant association to ambient concentrations of V and PAHs, respectively.[35,36] We used urine concentrations of V and 1-OHP as well as residential address to identify among the 680 participants 49 who lived in the three townships closer to the complex, with urine concentrations of V and 1-OHP in the top 60% of the cohort as high exposure group, and 71 who lived in townships further away, with urine concentrations of the two exposure biomarkers in the bottom 40% of the cohort as low exposure group. We set these two criteria in order to get sufficient number of subjects in both high and low exposure groups with significant contrast in external and internal exposure levels. After checking for serum sample availability, we confirmed 37 children and adolescents in the high exposure group and 70 in the low exposure group as subjects for this study. This study was approved by the Research Ethics Committee of National Taiwan University Hospital, and informed consent was obtained from each participant.

External Exposures. Distance from 107 individual study subject's home locations to two previously identified main emission points: coal-fired power plant and oil refinery plants,[34] respectively, and road area surrounding homes were calculated using geological information system (GIS) software (ArcGIS version 10.1).

Internal Exposures. Urine concentrations of heavy metals As, Cd, Cr, Ni, Pb, V, and Hg in 107 study subjects were analyzed using previously reported inductively coupled plasma mass spectrometry (ICP−MS) method,[27,35] and 1-OHP was analyzed using previously reported high performance liquid chromatography (HPLC) method.[36] For heavy metals, spikes were examined to confirm measurement stability, and standard reference materials (SRM) for each metal were analyzed to assess accuracy. For 1-OHP analysis, detection limit was 0.01 ng/mL with an 89.6% recovery rate and a 4.0% coefficient of variation for repeated measurements. Urine concentration of exposure biomarkers below the method detection limit (MDL) was replaced by half of the MDL. Yuan et al. had correlated individual urine concentration of 1-OHP with ambient levels of group 2 carcinogens benzo[a]anthrancene, benzo[k]-fluoranthene, and dibenzo[a,h]anthracene at study subject's home locations in our study area.[9,36,37] Industrial emissions always contain multiple pollutants, especially petrochemical industries.[1,2,6,7] While we used V and 1-OHP as biological tracers for exposure to petrochemical industrial emissions, we analyzed other petrochemical industrial-related carcinogenic pollutants in order to comprehensively evaluate the synergistic early health effects, and to clarify the contribution of each carcinogenic pollutant to early health effects. Urinary creatinine analysis was conducted using enzyme-linked

immunosorbent assay at National Taiwan University Hospital medical diagnosis laboratory and used for adjustment of urinary exposure biomarker levels.

Metabolomics. Metabolic profiles for 107 study subjects were analyzed. Sample preparation and analytical method using Agilent 1290 UHPLC system coupled with 6540-QTOF (UHPLC-QTOF) (Agilent Technologies, Santa Clara, CA) were performed following previously reported protocols.[38] For quality control (QC), blanks and pooled QC samples were analyzed at the beginning of each batch, and every five samples within each batch. Principle component analysis (PCA) showed clear clustering of pooled QC on the score plots shown in Supporting Information (SI) Figure S1. Synthetic samples containing 40 chemical standards (QC standard) were analyzed at the beginning of each batch to check instrument performance. Three repeated analysis was performed for each sample and total ion chromatogram was manually checked for technical replicates. Acquired data was preprocessed using True Ion Pick (TIPick) algorithm for background subtraction and peak picking.[39] Peak identification was conducted by matching m/z to an established in-house database: the National Taiwan University MetaCore Metabolomics Chemical Standard Library. Obtained potential metabolite features were preprocessed by removing those with >50% missing values, and replacing the missing values of the remaining features with half of the minimum positive value in the original data. Preprocessed data were normalized by sum of total peak area, log transformed, and autoscaled (mean-centered and divided by the standard deviation of each variable) prior to further statistical analysis.

Biomarkers of Early Health Effects. Urine concentrations of four oxidative stress biomarkers 8-hydroxy-2′-deoxyguanosine (8-OHdG), 4-hydroxy-2-nonenal-mercapturic acid (HNE−MA), 8-isoprostaglandin $F_{2\alpha}$ (8-isoPF$_{2\alpha}$), and 8-nitroguanine (8-NO$_2$Gua) were analyzed using previously published methods and adjusted with urinary creatinine concentrations.[40] QC was conducted following European Medicines Agency guidelines.[41] The four oxidative stress biomarkers applied in this study represent the different effects of oxidative stress, and all four biomarkers participate in the process of carcinogenesis. 8-OHdG is the most used biomarker for free radical induced DNA damage, and it has been reported as a good biomarker for risk assessment of cancer.[30] HNE−MA is a metabolite of lipid peroxidation product 4-hydroxy-2-nonenal (HNE), a cytotoxic and mutagenic signaling molecule that regulates cell cycle and forms DNA adducts leading to DNA damage.[30,42,43] Urine levels of 8-isoPF$_{2\alpha}$ is a biomarker for arachidonic acid peroxidation, and has been associated with increased risk of potential malignant oral disorders and breast cancer progression.[44,45] 8-NO$_2$Gua is formed from DNA damage induced by RNS generated under inflammatory conditions, and reported to participate in carcinogenesis.[46] Serum levels of 31 acylcarnitines in 107 study subjects were analyzed using UHPLC−MS. Four hundred μL of methanol (Scharlau, Sentmenat, Spain) was added into 100 μL of human serum to extract metabolites. The extraction was performed on Geno/Grinder2010 (SPEX, Metuchen, NJ) at 1000 rpm for 2 min followed by centrifugation at 15 000 rcf for 5 min at 4 °C. Supernatant was collected and evaporated using EYELA CVE-200D Centrifugal Evaporator (TOKYO RIKAKIKAI CO., Tokyo, Japan) until dry. Dried extracts were reconstituted with 200 μL of 10% methanol and centrifuged at 15 000 rcf for 5 min. The supernatant was then filtered with 0.2 μm of

Table 1. Comparison of Basic Characteristics, Carcinogens Exposure Levels, and Oxidative Stress Biomarker Levels in 107 Study Subjects

	high ($n = 37$)	low ($n = 70$)	p^a
Basic Characteristics			
age, mean ± SD	13.67 ± 0.92	13.70 ± 0.90	0.84
male, n(%)	20 (54.1)	38 (54.3)	0.98
systolic blood pressure (SBP), mean ± SD	118.7 ± 12.96	116.2 ± 13.73	0.37
body Mass Index (BMI), mean ± SD	21.67 ± 3.41	20.15 ± 3.48	0.04
smoke history, n(%)	5 (13.5)	5 (7.1)	0.28
drink history, n(%)	5 (13.5)	3 (4.3)	0.12
betelnut history, n(%)	1 (2.7)	3 (4.4)	1.00
External Exposures[b], Mean ± SD			
distance to coal-fired power plant	10.57 ± 2.52	21.81 ± 5.71	<0.0001
distance to oil refinery	10.02 ± 2.73	20.91 ± 5.44	<0.0001
road area surrounding homes			
25 m buffer	304.1 ± 211.4	329.4 ± 204.7	0.58
500 m buffer	70 938.4 ± 26 594.8	64 120.1 ± 20 016.4	0.18
Internal Exposures[c], Mean ± SD			
Group 1 Carcinogen			
arsenic	60.27 ± 42.16	39.62 ± 30.18	0.01
cadmium	0.34 ± 0.34	0.19 ± 0.15	0.02
chromium	3.24 ± 2.96	2.14 ± 1.63	0.10
nickel	6.69 ± 8.72	3.89 ± 2.85	0.31
Group 2A Carcinogen			
lead	0.64 ± 0.64	0.66 ± 0.65	0.80
1-OHP	0.19 ± 0.14	0.03 ± 0.01	<0.0001
Group 2B Carcinogen			
vanadium	2.46 ± 1.64	0.24 ± 0.10	<0.0001
mercury	3.13 ± 2.89	1.86 ± 1.88	0.04
Oxidative Stress[d], Mean ± SD			
8-OHdG	3.01 ± 2.15	2.63 ± 2.86	0.21
HNE−MA	2.00 ± 2.58	1.30 ± 2.20	0.006
8-isoPF$_{2\alpha}$	3.27 ± 3.54	2.06 ± 2.18	0.09
8-NO$_2$Gua	6.49 ± 11.77	2.54 ± 3.05	0.11

[a]Comparison of basic characteristics between the high and low exposure groups for continuous variables was performed using Student's t test, and for discrete variables, Chi-squared test or Fisher's exact test. Urinary exposure biomarker concentrations are log-transformed, high and low exposure groups compared by ANCOVA test adjusting for age, sex, smoking, alcohol consumption, betel nut chewing, fish consumption, and BMI with a post comparison by Scheffe test. Urinary oxidative stress biomarker concentrations are log-transformed, high and low exposure groups compared by Student's t test. [b]Distance to source: Average of home-to-coal-fired power plant and home-to-oil refinery distance, unit: km; Road area surrounding homes unit: m². [c]For urine 1-OHP, unit: μmol/mol-creatinine; for urine heavy metals, unit: μg/g-creatinine [d]Urine oxidative stress biomarkers unit: μg/g-creatinine. High exposure group N = 34, low exposure group N = 65.

Ministart RC 4 filter (Sartorius, Goettingen, Germany). All aliquots were transferred to glass inserts prior to UHPLC−MS analysis. Analysis was performed on Agilent 1290 UHPLC coupled with an Agilent 6460 triple quadrupole mass spectrometer (Agilent Technologies). The MS parameters were 325 °C, 325 °C, 500 V, and 3500 V for the drying gas temperature, sheath gas temperature, nozzle voltage, and capillary voltage, respectively. The dry gas flow and sheath gas flow were 7 and 11 L min⁻¹, respectively. The nebulizer was set at 45 psi. HSS T3 column (100 × 2.1 mm, 1.8 μm, Waters, Milford, MA) was employed and the column temperature was set at 40 °C. The mobile phase was composed of solvent A (0.1% formic acid in DI water) and solvent B (0.1% formic acid in ACN) (J.T. Baker, Phillipsburg, NJ). The gradient elution program was as followed: 0−1.5 min: 2% B; 1.5−9 min: linear gradient from 2 to 50% B; 9−14 min: linear gradient from 50 to 95% B; and 3 min maintenance in 95% B with the flow rate of 0.3 mL min⁻¹. A 3 min equilibrium was used before next injection. All analytes were monitored in positive MRM mode. All the peaks were integrated with MassHunter Quantitative Analysis software (Agilent Tech-

nologies). Pooled QC sample was analyzed every 20 samples and calculated for relative standard deviation (RSD). Out of the 31 analyzed acylcarnitines, 29 had RSD < 20%, and was used for statistical analysis of samples.

Meet-in-the-Middle. Partial least-squares discrimination analysis (PLS-DA) was performed using Metaboanalyst 4.0 (The Metabolomics Innovation Center, Edmonton, Alberta, Canada) to identify exposures-related potential metabolites.[47] PLS-DA models were validated using permutation test and cross-validation test. We further used ANCOVA to compare the peak area of each potential metabolite between high and low exposure groups adjusting for age, sex, and body mass index (BMI). Linear regression analysis was conducted for identifying early health effect-related potential metabolites adjusting for age, sex, and BMI.

Pathway Analysis. Pathway analysis was performed using Metaboanalyst 4.0, which currently supports 80 pathways in the *Homo sapiens* pathway library.[47] HMDB ID number and normalized peak area values were used as input. The method "Globaltest" was used for pathway enrichment analysis, and "betweenness centrality" for pathway topology analysis.
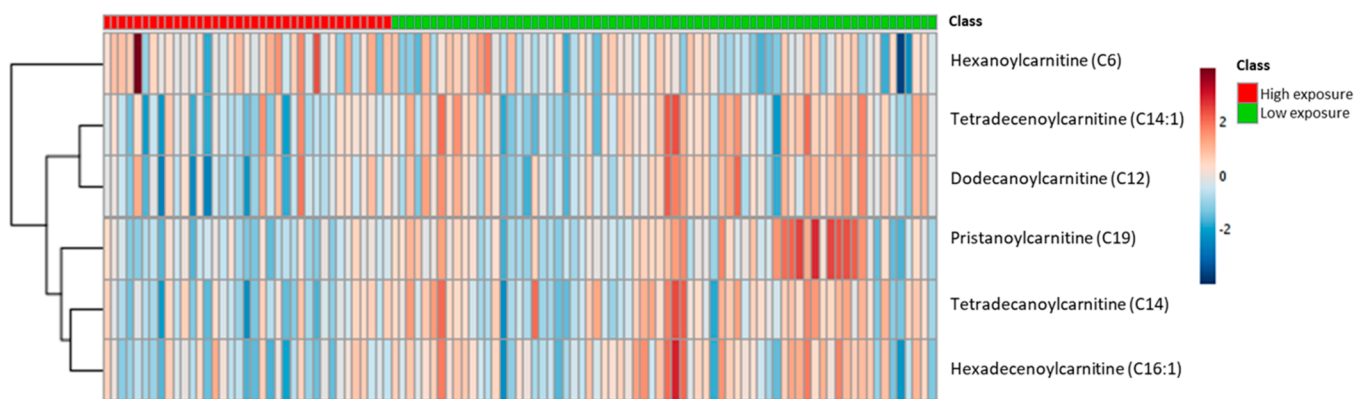
Figure 2. Heat map of serum acylcarnitine levels in 107 study subjects.
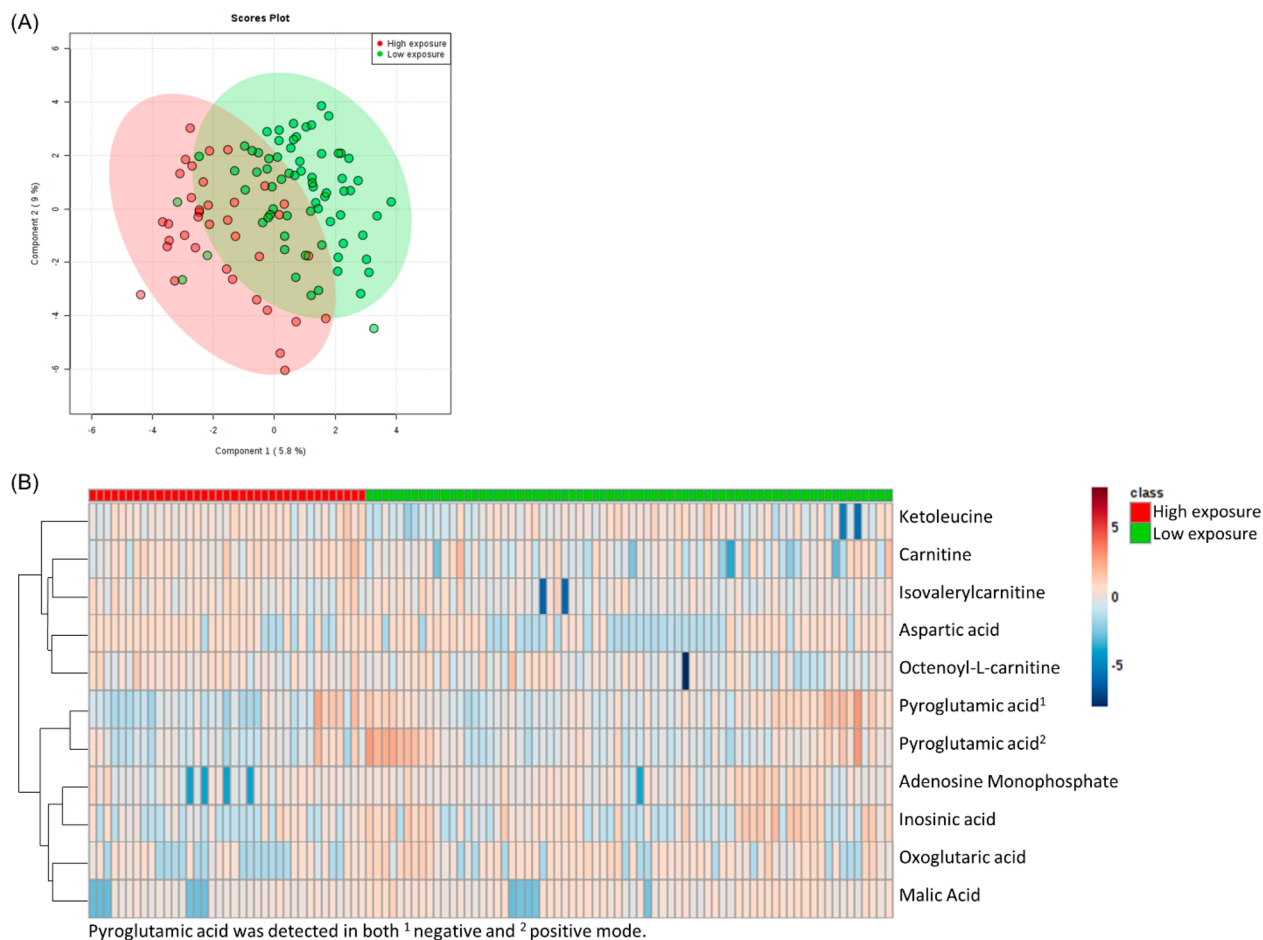

Figure 3. Comparison of serum metabolic profile in 107 study subjects using (A) PLS-DA score plot (accuracy = 0.78, R^2 = 0.53, Q^2 = 0.23, permutation p = 0.01) and (B) heat map of exposure-related potential metabolite levels (average VIP score >1, ANOVA $p < 0.05$ are shown adjusted for sex, age, and BMI).

Statistical Analysis. For comparison of basic characteristics and external exposure levels between high and low exposure groups, we used Student's t test to analyze continuous variables, and Chi-squared test or Fisher's exact test for discrete variables. Urine concentrations of internal exposure biomarkers were adjusted using urine creatinine concentrations and log transformed before comparing between high and low exposure groups using analysis of covariance (ANCOVA) adjusting for age, sex, smoking, alcohol consumption, betel nut chewing, fish consumption, and BMI

with a post comparison by Scheffe test. Oxidative stress biomarkers were adjusted using urine creatinine concentrations and log transformed before comparing between high and low exposure groups using Student's t test. Serum levels of potential metabolite features and acylcarnitines were normalized before comparing between high and low exposure groups using ANCOVA, adjusting for age, sex, and BMI with a post comparison by Scheffe test. Individual association between eight carcinogens and four oxidative stress biomarkers were analyzed using linear regression analysis, whereas association of

Table 2. Individual Association between Urine Carcinogens and Oxidative Stress Biomarkers in 99 Study Subjects[a]

	8-OHdG			HNE−MA			8-isoPF$_{2\alpha}$			8-NO$_2$Gua		
	estimate	95% CI	p value	estimate	95% CI	p value	estimate	95% CI	p value	estimate	95% CI	p value
Group 1 Carcinogen												
arsenic	0.049	(−0.165, 0.263)	0.653	0.092	(−0.163, 0.347)	0.476	−0.096	(−0.321, 0.130)	0.402	0.016	(−0.324, 0.356)	0.925
cadmium	−0.025	(−0.252, 0.202)	0.831	0.097	(−0.173, 0.366)	0.480	−0.141	(−0.379, 0.097)	0.243	0.391	(0.040, 0.743)	0.029
chromium	−0.186	(−0.380, 0.007)	0.059	−0.016	(−0.251, 0.219)	0.892	−0.061	(−0.269, 0.147)	0.562	0.120	(−0.192, 0.432)	0.448
nickel	0.018	(−0.172, 0.209)	0.848	0.033	(−0.194, 0.260)	0.775	0.126	(−0.073, 0.326)	0.213	0.224	(−0.075, 0.523)	0.140
Group 2A Carcinogen												
lead	0.115	(−0.112, 0.341)	0.318	0.158	(−0.112, 0.429)	0.249	−0.135	(−0.374, 0.104)	0.265	−0.033	(−0.397, 0.331)	0.858
1-OHP	0.094	(−0.057, 0.244)	0.219	0.146	(−0.032, 0.325)	0.107	0.073	(−0.087, 0.232)	0.369	0.188	(−0.050, 0.425)	0.121
Group 2B Carcinogen												
vanadium	0.132	(−0.013, 0.276)	0.073	0.254	(0.087, 0.422)	0.003	0.176	(0.025, 0.327)	0.023	0.330	(0.107, 0.554)	0.004
mercury	0.246	(0.064, 0.427)	0.008	0.279	(0.062, 0.496)	0.012	0.119	(−0.078, 0.316)	0.233	0.005	(−0.293, 0.304)	0.973

[a]Linear regression analysis adjusted for sex, age, and BMI.

eight combined carcinogen exposures with four oxidative stress biomarkers were analyzed using weighted quantile sum (WQS) regression, both adjusted for age, sex, and BMI. The weighted contribution of quantile-scored exposures were derived based on bootstrap sampling ($n = 100$). Student's t test, Chi-squared test, Fisher's exact test, ANCOVA test, and linear regression analysis were performed using SAS 9.4 for Windows. PCA and PLS-DA were performed using Metaboanalyst 4.0.[47] WQS regression analysis was conducted using the gWQS package in R 3.5.1.[48]

RESULTS

Table 1 showed the comparison of basic characteristics, external and internal exposure levels, and urine oxidative stress biomarker levels between high and low exposure groups. The high exposure group lived 10.57 ± 2.52 km and 10.02 ± 2.73 km away from the main emission points of coal-fired power plant and oil refineries, respectively, whereas the low exposure group lived 21.81 ± 5.71 km and 20.91 ± 5.44 km away, respectively. The high and low exposure groups showed no significant difference in age, sex distribution, systolic blood pressure (SBP), smoking history, alcohol history, and betelnut history. However, the high exposure group had higher BMI compared to the low exposure group. Road area surrounding homes showed no significant difference between the two exposure groups at either a 25 or 500 m buffer. Urine concentrations of exposure biomarkers As, Cd, Cr, Ni, 1-OHP, V, and Hg were increased in the high exposure group compared to the low exposure group, with As, Cd, 1-OHP, V, and Hg reaching statistical significance. The difference was most profound in 1-OHP and V. Urine concentrations of oxidative stress biomarkers showed all four biomarkers were increased in the high exposure group compared to the low exposure group, but the differences were more statistically significant for lipid peroxidation biomarkers HNE−MA ($p = 0.006$) and 8-isoPF$_{2\alpha}$ ($p = 0.09$) than DNA damage biomarkers 8-OHdG ($p = 0.21$) and 8-NO$_2$Gua ($p = 0.11$). Only 99 out of the 107 subjects had available data for urine oxidative stress biomarkers concentrations.

Figure 2 showed serum levels of six acylcarnitines that were significantly different in the high exposure group compared to the low exposure group in 107 study subjects. Samples are in columns and arranged according to high exposure (red) and low exposure (green) groups. Acylcarnitines are in rows and were arranged according to hierarchical clustering using Euclidean distance measure and Ward algorithm. The colors vary from deep blue to dark brown to indicate data values change from down-regulation (blue) to up-regulation (brown). We found long-chain acylcarnitines were clustered together and down-regulated in high exposure group compared to low exposure group (dodecanoylcarnitine, C12; tetradecanoylcarnitine, C14; tetradecenoylcarnitine, C14:1; hexadecenoylcarnitine, C16:1; pristanoylcarnitine, C19), whereas short-chain acylcarnitine (hexanoylcarnitine, C6) was up-regulated in the high exposure group compared to the low exposure group.

Metabolomics identified 84 potential metabolite features in study subjects serum samples after removing features missing in more than 50% of the samples, 80 of which had available HMDB ID number as shown in SI Table S1.[49] 84 potential metabolite features were put through PLS-DA analysis, and results showed metabolic profiles between the high and low exposure groups could be significantly separated by two components that accounted for 5.8% and 9.0% of variability of metabolic profiles between the high and low exposure groups, respectively (accuracy = 0.78, $R^2 = 0.53$, $Q^2 = 0.23$) (Figure 3A). Permutation test was performed to confirm the validity of PLS-DA model ($p = 0.01$). PLS-DA and ANCOVA analysis adjusting for age, sex, and BMI found 11 exposure-related potential metabolite features (average variable importance in projection (VIP) score >1, ANCOVA $p < 0.05$), which through in house library search was identified as 10 potential metabolites. Two potential metabolites, one detected under positive mode and one under negative mode of UHPLC−qTOFMS analysis, were both identified as pyroglutamic acid. Figure 3B showed the up- and down-regulation of exposure-related potential metabolites in the high and low exposure groups. Samples are in columns arranged according to high exposure (red) and low exposure (green) groups. Potential metabolites are in rows and were arranged according to
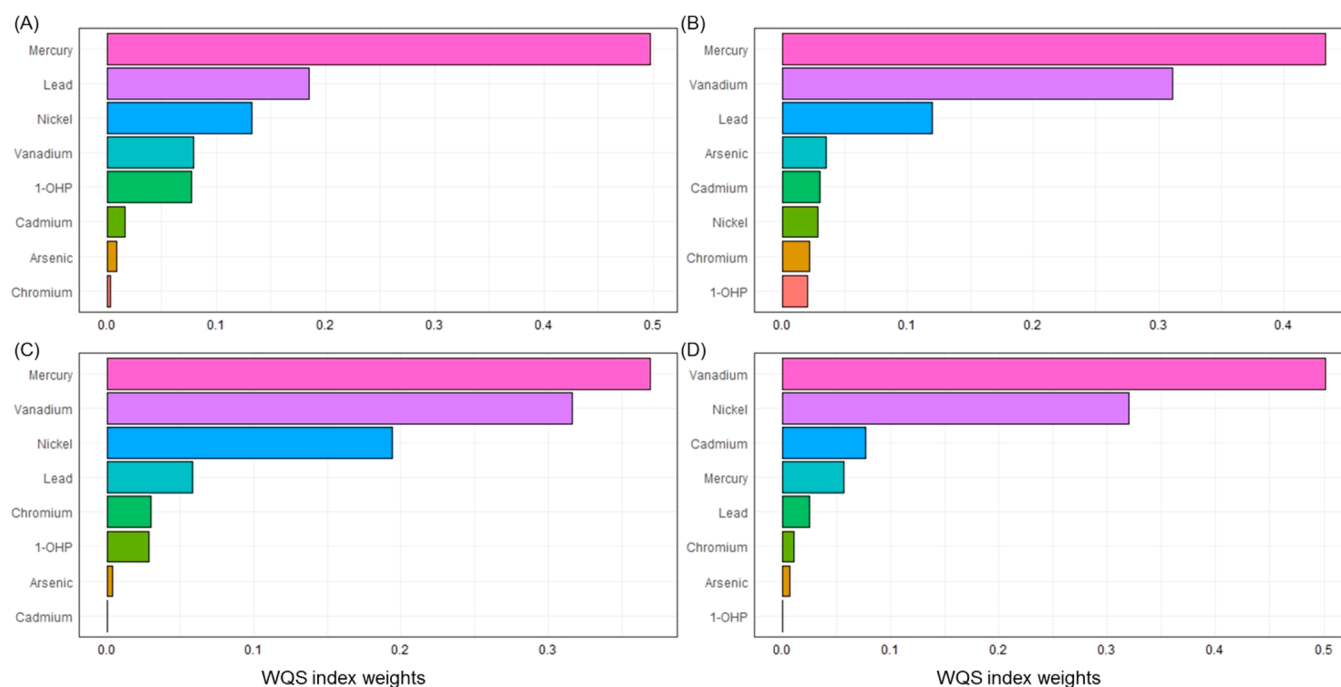
Figure 4. Combined associations between internal exposure levels and (A) 8-OHdG ($p = 0.002$), (B) HNE−MA ($p = 0.0006$), (C) 8-isoPF$_{2\alpha}$ ($p = 0.08$), and (D) 8-NO$_2$Gua ($p = 0.0001$) levels based on weighted quantile sum (WQS) regression analysis in 99 study subjects. (Adjusted for sex, age, and BMI.)

hierarchical clustering using Euclidean distance measure and Ward algorithm. The colors vary from deep blue to dark brown to indicate data values change from down-regulation (blue) to up-regulation (brown). We found potential metabolites up-regulated in high exposure group compared to low exposure group were clustered together, including ketoleucine, carnitine, isovalerylcarnitine, aspartic acid, and octenoyl-L-carnitine, whereas down-regulated potential metabolites were also clustered together, including pyroglutamic acid, adenosine monophosphate (AMP), inosinic acid (inosine monophosphate, IMP), oxoglutaric acid, and malic acid (Figure 3B). Pathway analysis results showed purine metabolism was the main biological pathway affected by multiple exposures ($p < 0.05$, impact >0.1) (data not shown). We identified two exposure-related potential metabolites involved in purine metabolism, nucleotides AMP and IMP[50] (Figure 3B). Through the Comparative Toxicogenomics Database (CTD), we found that group 1 carcinogens, As, Cd, Cr, and Ni, were significantly associated with purine metabolism pathway (Bonferroni adjusted $p < 0.01$).[51]

Table 2 showed the association between eight individual carcinogens with four oxidative stress biomarkers in 99 study subjects. Individually, among four group 1 carcinogens, urinary levels of Cd was positively associated with urinary concentrations of 8-NO$_2$Gua ($p = 0.029$). The two group 2A carcinogens was not associated with any of the four oxidative stress biomarkers. For the two group 2B carcinogens, V was associated with HNE−MA ($p = 0.003$), 8-isoPF$_{2\alpha}$ ($p = 0.023$), and 8-NO$_2$Gua ($p = 0.004$), and Hg was associated with 8-OHdG ($p = 0.008$) and HNE−MA ($p = 0.012$). Figure 4A to 4D showed the WQS regression analysis of the association of combined eight carcinogens exposure with four oxidative stress biomarkers, respectively. Association with all four oxidative stress biomarkers were positive and statistically significant with 8-OHdG, HNE−MA, and 8-NO$_2$Gua, whereas association

with 8-isoPF$_{2\alpha}$ was borderline significant. For 8-OHdG, group 2B carcinogen Hg predominated in the mixture index (49.7%) and group 1 carcinogens Ni, As, and Cd also contributed to the association ($p = 0.002$) (Figure 4A). Figure 4B showed group 2B carcinogens Hg (43.3%) and V (31.1%) contributed to over half of the mixture index positively associated with HNE−MA levels ($p = 0.0006$), and group 1 carcinogens As, Cd, Ni, and Cr also showed contribution. Associations with 8-isoPF$_{2\alpha}$ was predominated by group 2B carcinogens Hg (36.9%) and V (31.6%), followed by group 1 carcinogen Ni (19.4%), with contributions from Cr and As ($p = 0.08$) (Figure 4C). In Figure 4D, we can see in the mixture index positively associated with 8-NO$_2$Gua ($p = 0.0001$), group 2B carcinogen V contributed to half of the association (50.1%), followed by group 1 carcinogen Ni (32.0%), with contributions from Cd, Cr, and As.

The "meet-in-the middle" approach identified eight potential metabolites that were both carcinogen exposure-related and associated with biomarkers of early health effects. Table 3 showed the level of association between carcinogen exposure-related potential metabolites (in rows) and biomarkers of early health effects (in columns). For oxidative stress biomarkers, 8-OHDG was significantly associated with pyroglutamic acid and inosinic acid, HNE−MA, was significantly associated with ketoleucine, octenoyl-L-carnitine, pyroglutamic acid, AMP, and IMP, 8-isoPF$_{2\alpha}$ was significantly associated with octenoyl-L-carnitine, and 8-NO$_2$Gua was not significantly associated with any exposure-related potential metabolites. Long-chain acyl-carnitines C14 and C19 were associated with the most number of exposure-related potential metabolites, five for C14 (carnitine, octenoyl-L-carnitine, pyroglutamic acid detected in both positive and negative modes, IMP) and five for C19 (carnitine, octenoyl-L-carnitine, pyroglutamic acid, AMP, IMP). C16:1 was associated with carnitine and octenoyl-L-carnitine. Short-chain acylcarnitine C6 was associated with four

Table 3. Association between Exposure-Related Potential Metabolites and Biomarkers of Early Health Effects in Study Subjects[a]

	8-OHDG[b]	HNE–MA[b]	8-isoPF$_{2\alpha}$[b]	8-NO2Gua[b]	C6[c]	C14.1[c]	C12[c]	C19[c]	C14[c]	C16.1[c]
ketoleucine	0.03 (−0.05, 0.11)	0.10[f] (2.19×10^{-03}, 0.19)	0.05 (−0.04, 0.13)	0.09 (−0.03, 0.22)	0.31[g] (0.12, 0.49)	−0.01 (−0.20, 0.17)	0.01 (−0.18, 0.19)	−0.04 (−0.23, 0.15)	−0.04 (−0.23, 0.15)	4.66×10^{-03} (−0.19, 0.20)
carnitine	2.79×10^{-03} (−0.08, 0.09)	0.05 (−0.05, 0.14)	0.01 (−0.08, 0.10)	−0.02 (−0.15, 0.11)	0.11 (−0.09, 0.31)	−0.13 (−0.32, 0.05)	−0.16 (−0.34, 0.02)	−0.19[f] (−0.38, −3.39×10^{-03})	−0.25[f] (−0.43, −0.06)	−0.23[f] (−0.42, −0.04)
isovalerylcarnitine	−0.05 (−0.13, 0.03)	−0.01 (−0.11, 0.09)	0.03 (−0.05, 0.12)	0.03 (−0.10, 0.16)	0.20[f] (1.39×10^{-03}, 0.40)	0.08 (−0.11, 0.28)	0.02 (−0.17, 0.21)	−0.09 (−0.29, 0.10)	−0.14 (−0.34, 0.05)	−0.09 (−0.29, 0.11)
aspartic acid	−0.03 (−0.11, 0.04)	−0.03 (−0.12, 0.07)	2.84×10^{-04} (−0.08, 0.08)	0.11 (−0.01, 0.24)	0.04 (−0.15, 0.24)	−0.09 (−0.27, 0.09)	−0.14 (−0.31, 0.04)	−0.02 (−0.21, 0.16)	−0.10 (−0.28, 0.11)	−0.08 (−0.27, 0.11)
octenoyl-L-carnitine	0.05 (−0.03, 0.13)	0.09[f] (2.00×10^{-03}, 0.18)	0.08[f] (4.28×10^{-03}, 0.17)	0.11 (−0.01, 0.23)	0.13 (−0.06, 0.32)	−0.12 (−0.30, 0.06)	−0.12 (−0.30, 0.05)	−0.32[h] (−0.49, −0.14)	−0.21[f] (−0.39, −0.03)	−0.19[f] (−0.38, −0.01)
pyroglutamic acid[d]	−0.04 (−0.12, 0.04)	−0.09 (−0.19, 1.89×10^{-03})	0.02 (−0.07, 0.10)	−0.07 (−0.20, 0.05)	−0.39[h] (−0.57, −0.32)	0.17 (−0.01, 0.36)	0.13 (−0.05, 0.32)	0.27[g] (0.09, 0.46)	0.21[f] (0.02, 0.39)	0.19 (−0.01, 0.38)
pyroglutamic acid[e]	−0.08[f] (−0.16, −3.36×10^{-03})	−0.09[f] (−0.19, −3.48×10^{-04})	0.01 (−0.07, 0.09)	−0.04 (−0.16, 0.09)	−0.32[g] (−0.50, −0.13)	0.13 (−0.05, 0.31)	0.07 (−0.11, 0.25)	0.17 (−0.01, 0.35)	0.20[f] (0.02, 0.39)	0.14 (−0.05, 0.33)
adenosine monophosphate	−0.05 (−0.13, 0.03)	−0.11[f] (−0.20, −0.02)	0.02 (−0.07, 0.10)	−0.11 (−0.24, 0.01)	0.03 (−0.17, 0.22)	0.08 (−0.10, 0.27)	0.15 (−0.02, 0.33)	0.34[h] (0.17, 0.52)	0.18 (−2.75×10^{-03}, 0.36)	0.09 (−0.10, 0.28)
inosinic acid	−0.08[f] (−0.16, −9.14×10^{-04})	−0.13[g] (−0.23, −0.04)	−0.04 (−0.12, 0.05)	−0.08 (−0.21, 0.05)	−0.06 (−0.25, 0.13)	0.01 (−0.17, 0.19)	0.05 (−0.13, 0.23)	0.48[h] (0.31, 0.64)	0.20[f] (0.02, 0.38)	0.05 (−0.15, 0.24)
oxoglutaric acid	−0.04 (−0.12, 0.04)	−0.01 (−0.11, 0.08)	1.48×10^{-03} (−0.08, 0.09)	−0.09 (−0.22, 0.03)	0.05 (−0.14, 0.25)	0.13 (−0.05, 0.32)	0.08 (−0.10, 0.26)	0.13 (−0.06, 0.31)	0.09 (−0.09, 0.28)	0.12 (−0.07, 0.31)
malic acid	−0.01 (−0.09, 0.07)	−0.03 (−0.12, 0.07)	0.05 (−0.03, 0.14)	0.01 (−0.11, 0.14)	0.06 (−0.14, 0.25)	0.14 (−0.04, 0.32)	0.14 (−0.04, 0.32)	0.17 (−0.01, 0.36)	0.15 (−0.03, 0.34)	0.08 (−0.11, 0.27)

[a]Estimates of linear regression analysis are shown. 95% CI are in brackets. Note: Linear regression model adjusted for age, sex, and BMI. [b]N = 99. [c]N = 107. [d]Detected in positive mode. [e]Detected in negative mode. [f]$p < 0.05$. [g]$p < 0.01$. [h]$p < 0.001$.
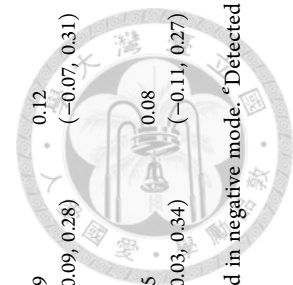
exposure-related potential metabolites including ketoleucine, isovalerylcarnitine, and pyroglutamic acid detected in both positive and negative modes. Overall, for the exposure-related potential metabolites, octenoyl-L-carnitine and pyroglutamic acid were associated with the most number of biomarkers of early health effects. Octenoyl-L-carnitine was associated with two oxidative stress biomarkers and three long-chain acylcarnitines, and pyroglutamic acid was associated with two oxidative stress biomarkers, short-chain acylcarnitine, and two long-chain acylcarnitines. Aspartic acid, oxoglutaric acid, and malic acid were not significantly associated with any of the biomarkers of early health effects.

■ DISCUSSION

Previous studies have reported exposure to individual carcinogens As, Cd, Cr, Ni, Pb, PAHs, V, and Hg can induce oxidative stress through production of reactive radicals and/or depletion of antioxidants.[52−54] However, these studies mostly focused on the association between single carcinogen exposure and oxidative stress, and only occupational exposure studies in adults reported the association between multiple heavy metals exposure and oxidative stress.[55] Our subjects were exposed to multiple carcinogens, and therefore, it was difficult to find a one-to-one association between specific carcinogens and oxidative stress. The level of which each carcinogen induced oxidative stress may also vary, especially in a mixture. The strength of our study is that we applied WQS regression analysis and showed in, children and adolescents, exposure to a mixture of eight environmental carcinogens was positively associated with four oxidative stress biomarkers, and both group 1 and group 2 carcinogens contributed to this association.

Our study is the first to report multiple carcinogens exposure could be associated with alterations in serum acylcarnitine levels in children and adolescents. Our findings support a previous study of adults occupationally exposed to metal-containing welding fumes who had a significant decrease in short- and long-chain acylcarnitines.[56] Previous studies have suggested acylcarnitines to be suitable candidates for cancer diagnosis.[57] Interestingly, the up- and down-regulation of short- and long-chain acylcarnitines vary by cancer type and study. Ni et al. reported both short-chain and long-chain acylcarnitines were significantly increased in lung cancer patients compared to healthy control subjects.[57] Another study showed a significant decrease of short-chain acylcarnitines in early stage nonsmall cell lung cancer patients.[58] In hepatocellular carcinoma patients, short-chain acylcarnitines were decreased and long-chain acylcarnitines were increased compared to control subjects.[32,59] We identified serum acylcarnitine deregulations in children and adolescents exposed to multiple carcinogens that have been reported in cancer patients, which may imply the possibility of increased cancer risk.

Deregulation in purine metabolism has been associated with early stage cancer development and cancer progression.[60] Purine metabolism is involved in energy production and signal transduction, and the enzymes and metabolites from this pathway can mediate oxidative stress through reactive species and antioxidant productions.[61−63] Our findings suggest multiple carcinogens exposure can induce perturbations in purine metabolism and link to increased oxidative stress and altered serum acylcarnitine levels.

Multiple carcinogens exposure was also associated with several potential metabolites in this study which could not be summarized in pathway analysis, but are involved in important biological mechanisms and have been reported in cancer studies. These potential metabolites included aspartic acid, an amino acid that has been reported to be involved in oxidative stress regulations.[64] Carnitine and citrate cycle-related metabolites malic acid and oxoglutaric acid were also identified, and carnitine was associated with acylcarnitines. Carnitine cooperates with acylcarnitines transporting fatty acids into mitochondria for β-oxidation, forming acetyl-CoA that enters the citrate cycle.[31] These findings suggest multiple carcinogens exposure in children and adolescents may affect fatty acid oxidation and energy production mechanisms leading to deregulation of acylcarnitines. Multiple carcinogens exposure in children and adolescents also affected pyroglutamic acid, an intermediate metabolite of antioxidant glutathione, and was linked to oxidative stress biomarkers and acylcarnitines.[65] Interestingly, four of the exposure-related potential metabolites we identified also showed similar patterns of alteration in early stage nonsmall cell carcinoma patients, including increased serum aspartic acid and carnitine, and decreased serum malic acid and pyroglutamic acid.[58,66,67] Our findings suggest multiple carcinogens exposure may have diverse effects on children and adolescents, causing disruptions in various biological mechanisms such as fatty acid oxidation, energy production, oxidative stress, and amino acid metabolism.

In this study, we found children and adolescents living near a petrochemical complex had increased exposure to multiple carcinogens, which induced metabolic changes associated with early health effects including increased oxidative stress and altered serum acylcarnitines, both of which may lead to increased cancer risk. Our findings may provide an explanation for increased cancer incidence among adult residents living near the same petrochemical complex reported in previous studies.[12,13]

There are limitations to this study. First, we analyzed metabolomics using a single analytical platform which limited the number of potential metabolite features detected, and cannot provide a comprehensive view of the metabolome. Second, we applied in-house library match using m/z for metabolite identification and, therefore, could not rule out the possibility of inaccurate metabolite identification and could not provide exact quantification of potential metabolites. Third, our sample size was limited, which could possibly explain why three of the four oxidative stress biomarkers were increased but did not reach a statistically significant difference between high and low exposure groups. Lastly, this is a cross-sectional study using single urine and serum samples, and therefore we could not confirm biomarker stability and could not be certain if the potential metabolites we identified can serve as life-long indicators of increased cancer risk.

Our findings imply multiple carcinogens exposure during critical periods of childhood and adolescence development induce metabolic perturbations in children and adolescents linking to early health effects that may contribute to cancer risk later in life. This indicates a significant reduction of toxic emissions from the complex could decrease multiple carcinogens exposure and metabolic abnormalities, which may potentially reduce cancer risks in children and adolescents living nearby. We recommend longitudinal epidemiological

studies in this area to follow up on children and adolescents' health if carcinogen emission continues in the near future.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.9b00392.

> Table S1, identified potential metabolite features in serum sample of 107 subjects using metabolomics; Figure S1, principle component analysis score plot of 61 pooled quality control samples data from 11 batches detected under (A) negative mode and (B) positive mode of UHPLC-qTOFMS metabolomics analysis (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: ccchan@ntu.edu.tw.

ORCID Ⓘ

Yufeng J. Tseng: 0000-0002-8461-6181
Ching-Hua Kuo: 0000-0001-5722-0360
Chang-Chuan Chan: 0000-0002-7518-5236

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) George, J.; Masto, R. E.; Ram, L. C.; Das, T. B.; Rout, T. K.; Mohan, M. Human exposure risks for metals in soil near a coal-fired power-generating plant. *Arch. Environ. Contam. Toxicol.* **2015**, *68* (3), 451−61.

(2) Nadal, M.; Schuhmacher, M.; Domingo, J. L. Metal pollution of soils and vegetation in an area with petrochemical industry. *Sci. Total Environ.* **2004**, *321* (1−3), 59−69.

(3) IARC. Arsenic, metals, fibres, and dusts. In *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*, Part C; International Agency for Research on Cancer, World Health Organization, 2012; Vol. *100*, pp 11−465.

(4) IARC. Cadmium and cadmium compounds. In *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*; International Agency for Research on Cancer, World Health Organization, 1993; Vol. *58*, pp 119−237.

(5) IARC. Chemical agents and related occupations. In *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*, Part F; International Agency for Research on Cancer, World Health Organization, 2012; Vol/ *100*, pp 9−562.

(6) Nadal, M.; Mari, M.; Schuhmacher, M.; Domingo, J. L. Multi-compartmental environmental surveillance of a petrochemical area: levels of micropollutants. *Environ. Int.* **2009**, *35* (2), 227−35.

(7) Hu, S. W.; Chan, Y. J.; Hsu, H. T.; Wu, K. Y.; ChangChien, G. P.; Shie, R. H.; Chan, C. C. Urinary levels of 1-hydroxypyrene in children residing near a coal-fired power plant. *Environ. Res.* **2011**, *111* (8), 1185−91.

(8) IARC. Cobalt in hard metals and cobalt sulfate, gallium arsenide, indium phosphide and vanadium pentoxide. In *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans* International Agency for Research on Cancer, World Health Organization, 2006; Vol. *86*, pp 1−294.

(9) IARC. Some non-heterocyclic polycyclic aromatic hydrocarbons and some related exposures. In *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*; International Agency for Research on Cancer, World Health Organization, 2010; Vol. *92*, pp 1−853.

(10) Carpenter, D. O.; Arcaro, K.; Spink, D. C. Understanding the human health effects of chemical mixtures. *Environ. Health Perspect.* **2002**, *110* (Suppl 1), 25−42.

(11) Lin, C. K.; Hsu, Y. T.; Christiani, D. C.; Hung, H. Y.; Lin, R. T. Risks and burden of lung cancer incidence for residential petrochemical industrial complexes: A meta-analysis and application. *Environ. Int.* **2018**, *121* (Pt 1), 404−414.

(12) Chen, C. F.; Chio, C. P.; Yuan, T. H.; Yeh, Y. P.; Chan, C. C. Increased cancer incidence of Changhua residents living in Taisi Village north to the No. 6 Naphtha Cracking Complex. *J. Formosan Med. Assoc.* **2018**, *117* (12), 1101−1107.

(13) Yuan, T. H.; Shen, Y. C.; Shie, R. H.; Hung, S. H.; Chen, C. F.; Chan, C. C. Increased cancers among residents living in the neighborhood of a petrochemical complex: A 12-year retrospective cohort study. *Int. J. Hyg. Environ. Health* **2018**, *221* (2), 308−314.

(14) UN. *Report of the Special Rapporteur on the Implications for Human Rights of the Environmentally Sound Management and Disposal of Hazardous Substances and Wastes*, A/HRC/33/41; Human Rights Council, 2016.

(15) Wild, C. P.; Kleinjans, J. Children and increased susceptibility to environmental carcinogens: evidence or empathy? In *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, 2003; Vol. *12* (12), pp 1389−94.

(16) Ginsberg, G. L. Assessing cancer risks from short-term exposures in children. *Risk Anal* **2003**, *23* (1), 19−34.

(17) Chiang, T. Y.; Yuan, T. H.; Shie, R. H.; Chen, C. F.; Chan, C. C. Increased incidence of allergic rhinitis, bronchitis and asthma, in children living near a petrochemical complex with SO2 pollution. *Environ. Int.* **2016**, *96*, 1−7.

(18) Wang, C. W.; Liao, K. W.; Chan, C. C.; Yu, M. L.; Chuang, H. Y.; Chiang, H. C.; Huang, P. C. Association between urinary thiodiglycolic acid level and hepatic function or fibrosis index in school-aged children living near a petrochemical complex. *Environ. Pollut.* **2019**, *244*, 648−656.

(19) Pan, B. J.; Hong, Y. J.; Chang, G. C.; Wang, M. T.; Cinkotai, F. F.; Ko, Y. C. Excess cancer mortality among children and adolescents in residential districts polluted by petrochemical manufacturing plants in Taiwan. *J. Toxicol. Environ. Health* **1994**, *43* (1), 117−29.

(20) Weng, H. H.; Tsai, S. S.; Chiu, H. F.; Wu, T. N.; Yang, C. Y. Association of childhood leukemia with residential exposure to petrochemical air pollution in taiwan. *Inhalation Toxicol.* **2008**, *20* (1), 31−6.

(21) Wild, C. P. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol., Biomarkers Prev.* **2005**, *14* (8), 1847−50.

(22) Rappaport, S. M.; Smith, M. T. Epidemiology. Environment and disease risks. *Science* **2010**, *330* (6003), 460−1.

(23) Pennell, K., Population screening for biological and environmental properties of the human metabolic phenotype: Implications for personalized medicine. In *Metabolic Phenotyping in Personalized and Public Healthcare*; Academic Press, 2016; pp 168−206, .

(24) Vineis, P.; van Veldhoven, K.; Chadeau-Hyam, M.; Athersuch, T. J. Advancing the application of omics-based biomarkers in environmental epidemiology. *Environ. Mol. Mutagen* **2013**, *54* (7), 461−7.

(25) Patti, G. J.; Yanes, O.; Siuzdak, G. Innovation: Metabolomics: the apogee of the omics trilogy. *Nat. Rev. Mol. Cell Biol.* **2012**, *13* (4), 263−9.

(26) Wishart, D. S. Emerging applications of metabolomics in drug discovery and precision medicine. *Nat. Rev. Drug Discovery* **2016**, *15* (7), 473−84.

(27) Chen, C. S.; Yuan, T. H.; Shie, R. H.; Wu, K. Y.; Chan, C. C. 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.

(28) Chen, Y.; Guillemin, G. J. Kynurenine pathway metabolites in humans: disease and healthy States. *Int. J. Tryptophan Res.* 2009, 2, 1−19.

(29) Reuter, S.; Gupta, S. C.; Chaturvedi, M. M.; Aggarwal, B. B. Oxidative stress, inflammation, and cancer: how are they linked? *Free Radical Biol. Med.* 2010, 49 (11), 1603−16.

(30) Valavanidis, A.; Vlachogianni, T.; Fiotakis, C. 8-hydroxy-2′-deoxyguanosine (8-OHdG): A Critical Biomarker of Oxidative Stress and Carcinogenesis. *Journal of Environmental Science and Health, Part C* 2009, 27 (2), 120−139.

(31) Semba, R. D.; Trehan, I.; Li, X.; Moaddel, R.; Ordiz, M. I.; Maleta, K. M.; Kraemer, K.; Shardell, M.; Ferrucci, L.; Manary, M. Environmental Enteric Dysfunction is Associated with Carnitine Deficiency and Altered Fatty Acid Oxidation. *EBioMedicine* 2017, 17, 57−66.

(32) Zhou, L.; Wang, Q.; Yin, P.; Xing, W.; Wu, Z.; Chen, S.; Lu, X.; Zhang, Y.; Lin, X.; Xu, G. Serum metabolomics reveals the deregulation of fatty acids metabolism in hepatocellular carcinoma and chronic liver diseases. *Anal. Bioanal. Chem.* 2012, 403 (1), 203−13.

(33) Rutkowsky, J. M.; Knotts, T. A.; Ono-Moore, K. D.; McCoin, C. S.; Huang, S.; Schneider, D.; Singh, S.; Adams, S. H.; Hwang, D. H. Acylcarnitines activate proinflammatory signaling pathways. *American journal of physiology. Endocrinology and metabolism* 2014, 306 (12), E1378−87.

(34) Shie, R. H.; Yuan, T. H.; Chan, C. C. Using pollution roses to assess sulfur dioxide impacts in a township downwind of a petrochemical complex. *J. Air Waste Manage. Assoc.* 2013, 63 (6), 702−11.

(35) Yuan, T. H.; Chio, C. P.; Shie, R. H.; Pien, W. H.; Chan, C. C. The distance-to-source trend in vanadium and arsenic exposures for residents living near a petrochemical complex. *J. Exposure Sci. Environ. Epidemiol.* 2016, 26 (3), 270−6.

(36) Yuan, T. H.; Shie, R. H.; Chin, Y. Y.; Chan, C. C. Assessment of the levels of urinary 1-hydroxypyrene and air polycyclic aromatic hydrocarbon in PM2.5 for adult exposure to the petrochemical complex emissions. *Environ. Res.* 2015, 136, 219−26.

(37) IARC. Dibenz[a,h]anthracene. In *IARC Monographs on the Evaluation of Carcinogenic Risks of Chemicals to Humans*; International Agency for Research on Cancer, World Health Organization, 1983; Vol. 32, pp 299−308.

(38) Liu, C. T.; Raghu, R.; Lin, S. H.; Wang, S. Y.; Kuo, C. H.; Tseng, Y. J.; Sheen, L. Y. Metabolomics of ginger essential oil against alcoholic fatty liver in mice. *J. Agric. Food Chem.* 2013, 61 (46), 11231−40.

(39) Ho, T. J.; Kuo, C. H.; Wang, S. Y.; Chen, G. Y.; Tseng, Y. J. True ion pick (TIPick): a denoising and peak picking algorithm to extract ion signals from liquid chromatography/mass spectrometry data. *J. Mass Spectrom.* 2013, 48 (2), 234−42.

(40) Wu, C.; Chen, S. T.; Peng, K. H.; Cheng, T. J.; Wu, K. Y. Concurrent quantification of multiple biomarkers indicative of oxidative stress status using liquid chromatography-tandem mass spectrometry. *Anal. Biochem.* 2016, 512, 26−35.

(41) European Medicines Agency. *EMEA/CHMP/EWP/192217/2009 Rev. 1 - Guideline on bioanalytical method validation*; European Medicines Agency (EMA), European Union, 2011; https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-bioanalytical-method-validation_en.pdf.

(42) Bartsch, H.; Nair, J. Accumulation of lipid peroxidation-derived DNA lesions: potential lead markers for chemoprevention of inflammation-driven malignancies. *Mutat. Res., Fundam. Mol. Mech. Mutagen.* 2005, 591 (1−2), 34−44.

(43) Ayala, A.; Munoz, M. F.; Arguelles, S. Lipid peroxidation: production, metabolism, and signaling mechanisms of malondialdehyde and 4-hydroxy-2-nonenal. *Oxid. Med. Cell. Longevity* 2014, 2014, 360438.

(44) Senghore, T.; Li, Y. F.; Sung, F. C.; Tsai, M. H.; Hua, C. H.; Liu, C. S.; Huang, R. J.; Yeh, C. C. Biomarkers of Oxidative Stress Associated with the Risk of Potentially Malignant Oral Disorders. *Anticancer Res.* 2018, 38 (9), 5211−5216.

(45) Ferroni, P.; Santilli, F.; Cavaliere, F.; Simeone, P.; Costarelli, L.; Liani, R.; Tripaldi, R.; Riondino, S.; Roselli, M.; Davi, G.; Guadagni, F. Oxidant stress as a major determinant of platelet activation in invasive breast cancer. *Int. J. Cancer* 2017, 140 (3), 696−704.

(46) Hiraku, Y. Formation of 8-nitroguanine, a nitrative DNA lesion, in inflammation-related carcinogenesis and its significance. *Environ. Health Prev. Med.* 2010, 15 (2), 63−72.

(47) Chong, J.; Soufan, O.; Li, C.; Caraus, I.; Li, S.; Bourque, G.; Wishart, D. S.; Xia, J. MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* 2018, 46 (W1), W486−W494.

(48) Carrico, C.; Gennings, C.; Wheeler, D. C.; Factor-Litvak, P. Characterization of Weighted Quantile Sum Regression for Highly Correlated Data in a Risk Analysis Setting. *Journal of Agricultural, Biological, and Environmental Statistics* 2015, 20 (1), 100−120.

(49) Wishart, D. S.; Feunang, Y. D.; Marcu, A.; Guo, A. C.; Liang, K.; Vazquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; Sayeeda, Z.; Lo, E.; Assempour, N.; Berjanskii, M.; Singhal, S.; Arndt, D.; Liang, Y.; Badran, H.; Grant, J.; Serra-Cayuela, A.; Liu, Y.; Mandal, R.; Neveu, V.; Pon, A.; Knox, C.; Wilson, M.; Manach, C.; Scalbert, A. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* 2018, 46 (D1), D608−D617.

(50) Simoni, R. E.; Gomes, L. N.; Scalco, F. B.; Oliveira, C. P.; Aquino Neto, F. R.; de Oliveira, M. L. Uric acid changes in urine and plasma: an effective tool in screening for purine inborn errors of metabolism and other pathological conditions. *J. Inherited Metab. Dis.* 2007, 30 (3), 295−309.

(51) Davis, A. P.; Grondin, C. J.; Johnson, R. J.; Sciaky, D.; McMorran, R.; Wiegers, J.; Wiegers, T. C.; Mattingly, C. J. The Comparative Toxicogenomics Database: update 2019. *Nucleic Acids Res.* 2019, 47 (D1), D948−D954.

(52) Jomova, K.; Valko, M. Advances in metal-induced oxidative stress and human disease. *Toxicology* 2011, 283 (2−3), 65−87.

(53) Fu, P. P.; Xia, Q.; Sun, X.; Yu, H. Phototoxicity and environmental transformation of polycyclic aromatic hydrocarbons (PAHs)-light-induced reactive oxygen species, lipid peroxidation, and DNA damage. *J. Environ. Sci. Health C Environ. Carcinog Ecotoxicol Rev.* 2012, 30 (1), 1−41.

(54) Valko, M.; Morris, H.; Cronin, M. T. Metals, toxicity and oxidative stress. *Curr. Med. Chem.* 2005, 12 (10), 1161−208.

(55) Ko, J.-L.; Cheng, Y.-J.; Liu, G.-C.; Hsin, I. L.; Chen, H.-L. The association of occupational metals exposure and oxidative damage, telomere shortening in fitness equipments manufacturing workers. *Ind. Health* 2017, 55 (4), 345−353.

(56) Shen, S.; Zhang, R.; Zhang, J.; Wei, Y.; Guo, Y.; Su, L.; Chen, F.; Christiani, D. C. Welding fume exposure is associated with inflammation: a global metabolomics profiling study. *Environ. Health* 2018, 17 (1), 68.

(57) Ni, J.; Xu, L.; Li, W.; Wu, L. Simultaneous determination of thirteen kinds of amino acid and eight kinds of acylcarnitine in human serum by LC-MS/MS and its application to measure the serum concentration of lung cancer patients. *Biomed. Chromatogr.* 2016, 30 (11), 1796−1806.

(58) Klupczynska, A.; Derezinski, P.; Garrett, T. J.; Rubio, V. Y.; Dyszkiewicz, W.; Kasprzyk, M.; Kokot, Z. J. Study of early stage non-small-cell lung cancer using Orbitrap-based global serum metabolomics. *J. Cancer Res. Clin. Oncol.* 2017, 143 (4), 649−659.

(59) Chen, S.; Kong, H.; Lu, X.; Li, Y.; Yin, P.; Zeng, Z.; Xu, G. Pseudotargeted metabolomics method and its application in serum biomarker discovery for hepatocellular carcinoma based on ultra high-performance liquid chromatography/triple quadrupole mass spectrometry. *Anal. Chem.* 2013, 85 (17), 8326−33.

(60) Bester, A. C.; Roniger, M.; Oren, Y. S.; Im, M. M.; Sarni, D.; Chaoat, M.; Bensimon, A.; Zamir, G.; Shewach, D. S.; Kerem, B. Nucleotide deficiency promotes genomic instability in early stages of cancer development. *Cell* 2011, 145 (3), 435−46.

(61) Pedley, A. M.; Benkovic, S. J. A New View into the Regulation of Purine Metabolism: The Purinosome. *Trends Biochem. Sci.* **2017**, *42* (2), 141−154.

(62) Cantu-Medellin, N.; Kelley, E. E. Xanthine oxidoreductase-catalyzed reactive species generation: A process in critical need of reevaluation. *Redox Biol.* **2013**, *1* (1), 353−358.

(63) Maiuolo, J.; Oppedisano, F.; Gratteri, S.; Muscoli, C.; Mollace, V. Regulation of uric acid metabolism and excretion. *Int. J. Cardiol.* **2016**, *213*, 8−14.

(64) Sivakumar, R.; Babu, P. V.; Shyamaladevi, C. S. Aspartate and glutamate prevents isoproterenol-induced cardiac toxicity by alleviating oxidative stress in rats. *Exp. Toxicol. Pathol.* **2011**, *63* (1−2), 137−42.

(65) Kumar, A.; Bachhawat, A. K. Pyroglutamic acid: throwing light on a lightly studied metabolite. *Curr. Sci.* **2012**, *102* (2), 288−297.

(66) Klupczynska, A.; Derezinski, P.; Dyszkiewicz, W.; Pawlak, K.; Kasprzyk, M.; Kokot, Z. J. Evaluation of serum amino acid profiles' utility in non-small cell lung cancer detection in Polish population. *Lung cancer (Amsterdam, Netherlands)* **2016**, *100*, 71−76.

(67) Klupczynska, A.; Plewa, S.; Dyszkiewicz, W.; Kasprzyk, M.; Sytek, N.; Kokot, Z. J. Determination of low-molecular-weight organic acids in non-small cell lung cancer with a new liquid chromatography-tandem mass spectrometry method. *J. Pharm. Biomed. Anal.* **2016**, *129*, 299−309.