EUROPEAN RESPIRATORY *journal*

FLAGSHIP SCIENTIFIC JOURNAL OF ERS

Early View

Original research article

Indoor microbiome, microbial and plant metabolites, chemical compounds and asthma symptoms in junior high school students: a multicentre association study in Malaysia

Yu Sun, Mei Zhang, Zheyuan Ou, Yi Meng, Yang Chen, Ruqin Lin, Jamal Hisham Hashim, Zailina Hashim, Gunilla Wieslander, Qingsong Chen, Dan Norbäck, Xi Fu

Please cite this article as: Sun Y, Zhang M, Ou Z, *et al*. Indoor microbiome, microbial and plant metabolites, chemical compounds and asthma symptoms in junior high school students: a multicentre association study in Malaysia. *Eur Respir J* 2022; in press (https://doi.org/10.1183/13993003.00260-2022).

Indoor microbiome, microbial and plant metabolites, chemical compounds and asthma symptoms in junior high school students: a multicentre association study in Malaysia

Yu Sun [1,2,3*#], Mei Zhang [1,2,3#], Zheyuan Ou [1,2,3], Yi Meng [1,2,3], Yang Chen [1,2,3], Ruqin Lin [1,2,3], Jamal Hisham Hashim [5], Zailina Hashim [6], Gunilla Wieslander [7], Qingsong Chen [8,9], Dan Norbäck [7], Xi Fu [8,9*]

[1]Guangdong Provincial Key Laboratory of Protein Function and Regulation in Agricultural Organisms, College of Life Sciences, South China Agricultural University, Guangzhou 510642, Guangdong, P. R. China

[2]Guangdong Laboratory for Lingnan Modern Agriculture, Guangzhou 510642, Guangdong, P. R. China

[3]Key Laboratory of Zoonosis of the Ministry of Agriculture and Rural Affairs, South China Agricultural University, Guangzhou 510642, Guangdong, P. R. China

[5]Universiti Selangor, Shah Alam, Selangor, Malaysia.

[6]Department of Environmental and Occupational Health, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, UPM, Serdang, Selangor, Malaysia.

[7]Occupational and Environmental Medicine, Dept. of Medical Science, University Hospital, Uppsala University, 75237 Uppsala, Sweden.

[8]School of Public Health, Guangdong Pharmaceutical University, Guangzhou, 510006, P. R. China

[9]Guangdong Provincial Engineering Research Center of Public Health Detection and Assessment, Guangdong Pharmaceutical University, Guangzhou, P. R. China.

*To whom corresponds should be addressed: Xi Fu fuxi@gdpu.edu.cn Yu Sun sunyu@scau.edu.cn

[#]These authors contributed equally to this work.

Abstract

Indoor microbial exposure is associated with asthma, but the health effects of indoor metabolites and chemicals are not comprehensively assessed. Here, we collected classroom dust from 24 junior high schools in three geographically distanced areas in Malaysia, including Johor Bahru, Terengganu, and Penang, and conducted culture-independent high-throughput microbiome and untargeted metabolomics/chemical profilings. 1,290 students were surveyed for asthma symptoms (wheeze). In each center, we found significant variation in the prevalence of wheeze among schools, which cannot be explained by personal characteristics and air pollutants. Large-scale microbial variations were observed between three centers; the potential protective bacteria were mainly from phylum Actinobacteria in Johor Bahru, Cyanobacteria in Terengganu, and Proteobacteria in Penang. In total, 2,633 metabolites and chemicals were characterized. Many metabolites were enriched in low wheeze schools, including plant secondary metabolites flavonoids/isoflavonoids (isoliquiritigenin, formononetin, astragalin), indole and derivatives (indole, serotonin, 1H-indole-3-carboxaldehyde), and others (biotin, chavicol). A neural-network analysis showed that the indole derivatives were co-occurring with the potential protective microbial taxa, including *Actinomycetospora*, *Fischerella* and *Truepera*, suggesting these microorganisms may pose health effects by releasing indole metabolites. A few synthetic chemicals were enriched in high wheeze schools, including pesticide (2(3H)-benzothiazolethione), fragrance (2-aminobenzoic acid, isovaleric acid), detergent and plastic (phthalic acid), and industrial material (4,4-sulfonyldiphenol). This is the first association study between high-throughput indoor chemical profiling and asthma symptoms. The consistent results from three centers indicate that indoor metabolites/chemicals could be a better indicator than indoor microbiome for environmental and health assessments, providing new insights for asthma prediction, prevention, and control.

Introduction

Asthma is a common allergy-related chronic respiratory disease that affects more than 350 million patients worldwide. The prevalence of asthma symptoms (wheeze and whistling) is >30% in many countries, including Australia, Ireland, and the UK, posing a severe threat to public health [1]. Epidemiology studies show that the occurrence of asthma and allergies are mainly affected by

environmental exposure, including air pollution, environmental allergens, and microorganisms [2-4]. A striking phenomenon is that the prevalence of asthma is significantly lower in children growing up in the farm or rural areas than in the urban area [5]. Subsequent studies revealed that indoor microbial exposure is the driving factor for the variation [6]. Studies in Finland, Germany, Malaysia, and China also confirm the importance of the indoor microbiome in immune modulation and disease development [7-10]. However, it is challenging to transform the theoretical progress into practical applications, such as building an indoor microbiome indicator for environmental assessment and disease prediction. This is because the diversity of the indoor microbiome is extremely high. The total microbial species on earth is approximately 1 trillion [11]. Also, the indoor microbiome shows extremely high geographic diversity. Different sets of indoor microorganisms and health-associated microorganisms are characterized in different geographic regions [7-10]. It is almost impossible to find a consensus set of health-related species across the globe and make a solid indoor microbiome-health inference. Thus, an alternative environmental assessment indicator is needed.

Indoor metabolites and chemicals could be a potential alternative for environmental assessment. Each bacterial and fungal organism can release thousands of metabolites into the living environment per hour, affecting the health of occupants. The common health-related metabolites included lipopolysaccharide (LPS), muramic acid and microbial volatile organic compounds (MVOCs) [12, 13]. However, previous studies used culture-dependent or low-throughput approaches to characterize a small set of targeted chemical exposure from microorganisms. No study used a high-throughput untargeted approach to profile comprehensive indoor metabolites and chemicals. Thus, the whole picture of metabolites/chemicals in the indoor environment is still unclear. Also, no study conducted the multi-omic analysis between indoor microorganisms and metabolites to identify the potential microbial sources of metabolites.

In this study, we surveyed 1,290 junior high school students in Johor Bahru, Terengganu, and Penang, Malaysia, for asthma symptoms and collected classroom dust for culture-independent high-throughput microbiome and untargeted chemical profiling. We aim to characterize indoor metabolic/chemical exposure and uncover their relationships with health-related microorganisms. Also, we compared the environmental chemical patterns in multiple centers and tested whether it could be a better indicator than the indoor microbiome for exposure assessment.

Materials and methods

Study design and health data

We conducted classroom dust sampling and health surveys in three areas in Malaysia, including Johor Bahru, Terengganu, and Penang. The locations are displayed in Figure S1. In each center, 8 junior high schools (4 classrooms in each school) were randomly selected for dust sampling. In each class, the health questionnaires in Malay were sent to 15-20 students with age 14-15. Health questions were obtained from the International Study of Asthma and Allergies in Childhood (ISAAC) study [14], including an asthma symptom question - "In the last 12 months, have you had wheezing or whistling in the chest when you DID NOT have a cold or the flu?". Personal information was collected, including age and gender. The participants had no information regarding environmental data and samples collected. The study design and protocol were approved by the Medical Research and Ethics Committee of the National University of Malaysia, and informed consent was obtained from all participants.

Dust sampling and environmental characteristic measurements

We sampled classroom dust on floors, desks, chairs, bookshelves, and curtains with a vacuum cleaner in Johor Bahru and Terengganu [7, 15]. The vacuum procedure was maintained 4 minutes, 2 minutes on the floor and 2 minutes on other surfaces above the floor level, including student desks, chairs, bookshelves, and curtains. The dust was collected in a sampler (ALK Abello, Copenhagen, Denmark) with a filter pore size of 6 μm. In schools in Penang, we collected settled dust on the upper frame of the blackboard with a metal spoon. The vacuumed and settled dust was sieved to fine dust through a metal mesh screen (pore size 0.3mm). The fine dust was stored in -80˚C freezer.

Indoor relative humidity and CO_2 were measured by Q-TRACK IAQ (TSI, St. Paul, MN, USA). Indoor NO_2 was sampled by a diffusion sampler from IVL Swedish Environmental

Research Ltd. (Gothenburg, Sweden).

High-throughput amplicon sequencing

Bacterial and fungal amplicon sequencing was performed by using dust samples. In brief, total microbial DNA was extracted from 10 mg fine dust by E.Z.N.A soil DNA kit D5625-01 (Omega Bio-Tek, Inc., Norcross, GA, USA) and Fast DNA SPIN kit (MP Biomedicals, Santa Ana, CA, USA). DNA quality was assessed by NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and agarose gel electrophoresis. Bacterial 16S rRNA gene and fungal ITS (internal transcribed spacer) region were amplified, and sample-specific barcode sequences were added during the library preparation step. The amplicons were sequenced by Illumina MiSeq and PacBio platforms. Raw sequence data was deposited in QIITA (12875) and Genome Sequence Archive (CRA002825, CRA002876, CRA005646 and CRA005647) [16]. The absolute bacterial and fungal concentration was quantified by qPCR with universal primers [15].

Profiling chemical compounds in classroom dust

Chemical compounds in classroom dust were assessed by untargeted LC-MS (liquid chromatography-mass spectrometry) in BioNovoGene (Suzhou, China). 10 mg fine dust was added to 0.6 mL 2-chlorophenylalanine in methanol and centrifuged at 12,000 rpm at 4˚C for 10 min. 300 μL supernatant was filtered through a 0.22 μm membrane. Chromatographic separation was performed by an Acquity UPLC HSS T3 Columns (2.1 × 150 mm, 1.8 μm, Waters Corporation, Milford, Massachusetts, USA) at 40˚C at a flow rate of 0.25 mL/min. Vanquish HPLC systems Q Exactive HF-X Hybrid Quadrupole-Orbitrap Mass Spectrometer was used for LC-MS detection (Thermo Fisher Scientific, Waltham, Massachusetts, USA). Electrospray ionization mass spectrometric (ESI-MSn) experiments were performed with the spray voltage of 3.5 kV and -2.5 kV in positive and negative modes. Sheath and auxiliary gas and capillary temperature were set at 30 and 10 arbitrary units and 325 ˚C. The analyzer scanned a mass range of m/z 81-1000 at a mass resolution of 60,000. Chemicals were annotated by searching against mzCloud database (www.mzcloud.org), the Human Metabolome Database (www.hmdb.ca),

MoNA (mona. fiehnlab.ucdavis.edu), METLIN (metlin.scripps.edu), and MassBank (www.massbank.jp).

High-throughput data analysis

Microbiome data process and analysis were conducted in the QIIME2 platform [17]. Raw reads were assigned to samples according to the barcode information. Low quality and chimeric reads were removed. Sequence taxonomy was annotated by the Silva (release 115) and UNITE database (release 5) [18, 19]. LEfSe analysis was conducted to characterize the enriched microbial taxa in different groups (LDA > 2) [20]. Mmvec was used to estimate microbe-metabolite interactions co-occurrence probabilities [21], and the results were visualized by TBtools [22]. The key features of the chemical compounds and Globally Harmonized System (GHS) classification were obtained from the PubChem database [23].

Results

Prevalence of asthma symptoms in three centers of Malaysia

In total, 1,290 Junior high school students from Johor Bahru, Terengganu, and Penang, Malaysia, were surveyed for asthma symptoms/wheeze. Three centers were located in the south, northeast, and northwest Malaysia (Figure S1), and eight junior high schools were randomly selected in each center. We found large-scale variation in the prevalence of asthma symptoms among schools ($p = 0.0007$, 0.0004 and 0.002; Table 1). For example, the prevalence of wheeze was 12.7%, 20.0%, 14.3%, 29.9%, 8.9%, 8.7%, 5.6% and 9.1% in schools of Terengganu. In each center, the top four schools were defined as high wheeze schools, and the bottom four schools were defined as low wheeze schools.

We further explored the personal and environmental characteristics that could explain the variation of asthma symptoms among schools. Students' age and gender did not differ between high and low wheeze schools ($p > 0.05$, Mann-Whitney Test). The indoor environmental characteristics, including temperature, relative humidity, and CO_2 concentration, were similar

among the three centers (Table 2). Indoor NO_2 concentration was lower in Terengganu than in Johor Bahru and Penang. In all centers, these environmental characteristics did not differ between high and low wheeze schools ($p > 0.05$), failing to explain the variation of asthma symptoms.

Indoor microorganisms enriched in three centers and high/low wheeze schools

High-throughput sequencing was applied to characterize the abundance of the indoor microorganisms in Johor Bahru, Terengganu, and Penang (Table S1-8). We found strong geographic variation in indoor bacteria and fungi among centers. For example, bacterial genera *Ralstonia*, *Ochrobactrum* and *Craurococcus* were present in high abundance (7.8%, 2.7% and 1.3%) in Penang but were presented in low abundance in Johor Bahru and Terengganu (0%, 0.06%; 0.02%, 0.42%; 0.05%, 0.74%; Table S7). Fungal genera *Wallemia* and *Candida* were present in high abundance in Johor Bahru and Terengganu (6.6%, 12.2%; 1.3%, 1.8%) but in low abundance in Penang (0.06% and 0.03%)(Table S8). Common mold species, *Penicillium* and *Cladosporium*, were present in high abundance only in Johor Bahru (10.2% and 7.8%). *Fusarium* was present in high abundance in Terengganu but not in Johor Bahru and Penang (2.9% vs 0.3%, 0.009%). *Aspergillus* was present in high abundance in all centers (20.7%, 12.7%, and 12.1%).

We further explored the potential indoor microorganisms enriched in high/low wheeze schools in each center. In Johor Bahru, 15 bacterial taxa were enriched in low wheeze schools, and more than half were from phylum Actinobacteria, including *Rubrobacter*, *Actinomyces*, *Blastococcus*, *Janibacter*, *Actinomycetospora*, *Pseudokineococcus*, and *Marmoricola* (Figure 1 and S3). However, in Terengganu, the bacteria enriched in low wheeze schools were mainly from phylum Cyanobacteria, including *Chroococcidiopsis*, *Fischerella*, *Mastigocoleus*, and *Iphinoe*. In Penang, the bacteria enriched in low wheeze schools were mainly from phylum Proteobacteria, including *Caulobacter*, *Bosea*, *Acidovorax*, *Undibacterium*, and uncharacterized (uc) Sphingomonadaceae. The results indicate that different centers have a unique set of potential protective bacterial taxa. The bacterial taxa enriched in high wheeze schools also differed among centers. *Catellicoccus* and *Ignatzschineria* were enriched in high wheeze schools in Johor Bahru, *Methylobacterium*, *Bacillus*, *Sphingomonas*, and *Pantoea* were enriched in Terengganu, and

Roseomonas was enriched in Penang (Figure S3). The potential protective fungal taxa were from the phylum Ascomycota and Basidiomycota (Figure 1).

Quantitative PCR was also conducted in Johor Bahru and Terengganu (Table S9 and S10). The absolute concentration of indoor bacteria and fungi did not differ between high and low wheeze schools ($p>0.1$, t-test).

Potential protective metabolites and associations with microorganisms

In total, 2,633 chemicals were characterized by LC-MS. The geographic pattern was also observed for chemical compounds (Figure S2). Indoor chemical composition in Johor Bahru, Terengganu, and Penang is located on the right side, upper-left, and lower-left side of the ordination plot (Figure S2B). However, general rules were also observed. Three classes of metabolites were almost exclusively enriched in low wheeze schools ($p < 0.01$, FDR < 0.1, fold change > 2), including flavonoids, isoflavonoids, and indole and derivatives (Table 3 and Figure 1). Flavonoids and isoflavonoids are important classes of plant secondary metabolites, widely found in various plants, fruits, and vegetables. Three flavonoids and isoflavonoids were enriched in two or more centers, including isoliquiritigenin, formononetin, and 6-hydroxydaidzein. Other flavonoid metabolites, including baicalin, astragalin, tangeritin, daidzein, luteolin, and procyanidin B2, were enriched in one center (Table 3). Indole and derivatives is a class of common small signaling molecules in microorganisms, plants, and animals. Serotonin and indole-3-carboxaldehyde were enriched in two or more centers, and indole, L-tryptophan, 1H-indole-3-acetamide, and indolepyruvate were enriched in one center. The explicit enrichment of flavonoids and indole derivatives in the dust of low wheeze schools suggests their potential anti-inflammatory and anti-allergic effects. A neural network analysis showed that the indoles and derivatives were co-occurring with many potential protective microbial taxa, such as L-tryptophan with *Actinomycetospora* and uc Corynebacteriacea, N-acetylserotonin and indole with *Truepera*, indole with uc Pleosporaceae, 5-methoxyindoleacetate with *Fischerella*. The results suggest that these microorganisms may produce these metabolites. A literature search showed that *Actinomycetospora* and *Fischerella* were capable of producing indole derivatives [24, 25], supporting the in silico association results.

Other potential protective metabolites were also identified, including biotin, chavicol, ecgonine, dihydrocortisol, and so on (Figure 2), which belong to different metabolic classes. Biotin was closely associated with many protective bacteria, including *Actinomyces*, *Paracoccus*, and *Sphingomonas*. A literature search in laboratory experiments showed that these taxa could produce biotin [26-28], consistent with the co-occurring analysis.

Potential risk chemicals were all synthetic chemicals

Potential risk environmental chemicals were defined as chemicals significantly enriched in the high wheeze schools ($p < 0.01$, FDR < 0.1, fold change > 2). One (2(3H)-benzothiazolethione), one (2-aminobenzoic acid), and three (isovaleric acid, phthalic acid, and 4,4-sulfonyldiphenol) hazardous chemicals were characterized in Johor Bahru, Terengganu, and Penang. These chemicals included pesticides, cleaning detergents, perfumes, and industrial materials (Table 4). Globally Harmonized System (GHS) has classified them as hazardous chemicals, and adverse health effects include dermatitis, hypersensitivity, inflammation, and eye and respiratory tract irritation. More hazardous chemicals were detected in Penang, which may explain the overall high prevalence of asthma symptoms compared with Johor Bahru.

Discussion

This is the first study to use a high-throughput untargeted approach to profile indoor chemical exposure and asthma. Natural metabolites, including microbial and plant metabolites, were protective for asthma, and synthetic chemicals, including pesticides, detergents, and industrial solvents, were risk factors for asthma. Also, this is the first study to assess the interactions between indoor microbiome and metabolites, revealing that indoor microorganisms may produce protective metabolites. In addition, three centers in Malaysia with large geographic distances were surveyed to support our results and conclusion. We found large-scale indoor microbiome variation, indicating an overall different microbiome exposure in each center. However, general rules were observed for indoor chemical compounds. Plant metabolites from flavonoids, isoflavonoids, and microbial metabolites from indole derivatives showed potential protective effects, whereas

synthetic chemicals showed adverse health effects. The results suggest that indoor chemicals could be a more solid and consistent indicator in exposure assessments for asthma.

There are also limitations in this study. First, dust in Johor Bahru and Terengganu was collected from vacuuming dust from floors, desks, tables, bookshelves, and curtains, whereas dust in Penang was collected from the blackboard frame, which may produce sampling bias. However, we argue that the two approaches should be comparable. Bookshelves, curtains, and blackboard frames are seldom or never clean in these classrooms, and thus both approaches collect dust representing long-term exposure. Second, we applied second-generation amplicon sequencing in this study, only resolving taxonomic resolution at the genus level. Third, as only the marker gene is sequenced, the abundance and health associations for function genes cannot be assessed. However, our study profiled indoor microbial metabolites, which provide more direct evidence for metabolic exposures than functional gene assessment.

Flavonoids and isoflavonoids are plant secondary metabolites with a polyphenolic structure. Flavonoids are found in many plants and plant-derived food, and isoflavonoids are predominantly found in soybeans and leguminous plants [29]. Flavonoids and isoflavonoids have anti-inflammatory, anti-oxidative, anti-carcinogenic, and anti-allergic properties for humans and animals (Table 3). Several flavonoids identified have shown protective effects for asthma in previous studies. For example, isoliquiritigenin suppresses IL-4 and IL-5 production in a dose-dependent manner in vitro [30], and astragalin reduces IL-4, IL-5, and IL-13 levels and inhibits eosinophil infiltration in mice [31]. Also, formononetin alleviates lung inflammation and cytokine levels and reduces oxidative stress in an ovalbumin-sensitized mouse model [32]. However, the previous studies mainly reported the health effects of flavonoids and isoflavonoids in lab animals and cell lines. This is the first study to show their beneficial effects as inhaled exposure for human populations, providing a novel perspective on respiratory disease.

Indole and derivatives are a group of aromatic heterocyclic organic compounds widely distributed in bacteria, plants, and animals [33]. The health effects of indoles were mainly studied in the human gut, with few studies on environmental microorganisms. A large variety of gut microorganisms can produce indoles and derivatives, including *Clostridium novyi*, *Escherichia coli*, *Fusobacterium*, *Enterococcus faecalis*, and *Corynebacterium acnes* [34, 35]. Indole and derivatives can improve human intestinal epithelial barrier integrity and reduce gut inflammation

by decreasing the expression of proinflammatory cytokines NF-κB and increasing the anti-inflammatory IL-10 [36]. Indolepyruvate and indole-3-acetamide were identified as potential protective metabolites in our study. In human gut studies, these indole derivatives can activate the expression of aryl hydrocarbon receptor (AhR) gene [37], which has an anti-inflammatory role in blocking the proinflammatory T cells in asthma development [38]. The inhaled exposure of indole metabolites could have a similar mechanism by activating AhR receptors in the lung and respiratory tract.

Besides flavonoids and indoles, biotin and chavicol were enriched in low wheeze schools. Biotin, a B7 vitamin, is an essential nutrient for humans. A previous indoor metagenomics survey reported that a higher abundance of biotin metabolisms pathways was associated with a lower prevalence of sick-building syndrome [39]. Chavicol is a natural phenylpropene found in *Piper betle* and used in traditional herb medicine of China and India. Chavicol analog can attenuate IFNγ expression in T-help cells and modulate inflammation and immune responses [40], supporting their roles in reducing asthma symptoms.

Only a few chemicals were significantly enriched in high wheeze schools after removing drugs and common human and plant metabolites, and all of them were synthetic chemicals, including pesticides, paints, fragrances, and industrial solvents. Phthalate exposure is reported to associate with asthma. A meta-analysis of 43 studies reported that benzyl phthalate increased 39% to 41% odds of childhood asthma [41]. A home survey in China also reported that a high concentration of phthalic acid esters increased childhood diagnosed asthma [42]. The other potential risk chemicals are not reported to associate with asthma by scientific publications, but the GHS classification, developed by the United Nations, indicated that these chemicals might have adverse health effects, including dermatitis, hypersensitivity, and respiratory tract irritation. Thus, future environmental surveys and asthma epidemiology should also consider these chemicals.

Previous indoor metabolite studies mainly surveyed microbial by low throughput approaches. Araki and Choi et al. reported that many MVOCs were positively associated with asthma and rhinitis [12, 43]. In our study, only three MVOCs (bornyl acetate, 2-heptanone, and estragole) were detected in vacuum dust, and none of them were significantly enriched in high/low schools. It is likely that most volatile chemicals cannot be detected by vacuum dust sampling. The total

LPS concentration seems to be mainly protectively associated with asthma [44]. Muramic acid was reported to be positively or negatively associated with asthma [45, 46]. In this study, the untargeted LC-MS also detected muramic acid and LPS (tridecanoic acid, hydroxy hexadecanoic acid, myristic acid), but none reached significance after the FDR adjustment.

In this study, we found large-scale variation in the microbiome composition and health-related microorganisms in three centers of Malaysia. This could be due to the extremely high diversity of environmental microorganisms [11]. Thus, it is challenging to use the indoor microbial composition to build a universal reference catalog for health assessments and disease prediction. However, consistent associations were observed for indoor chemical compounds, suggesting they could be used as an environmental assessment indicator for disease prediction, providing new insights and strategies for disease prevention and control.

Acknowledgments

The study was funded by the Natural Science Foundation of Guangdong Province (2020A1515010845 and 2021A1515010492) and the Science and Technology Program of Guangzhou (202102080362).

Conflict of interest: Y. Sun declares no conflict of interest.

Conflict of interest: M. Zhang declares no conflict of interest.

Conflict of interest: Z. Ou declares no conflict of interest.

Conflict of interest: Y. Meng declares no conflict of interest.

Conflict of interest: Y. Chen declares no conflict of interest.

Conflict of interest: R. Lin declares no conflict of interest.

Conflict of interest: J. Hashim declares no conflict of interest.

Conflict of interest: Z. Hashim declares no conflict of interest.

Conflict of interest: G. Wieslander declares no conflict of interest.

Conflict of interest: Q. Chen declares no conflict of interest.

Conflict of interest: D. Norbäck declares no conflict of interest.

Conflict of interest: X. Fu declares no conflict of interest.

References

1. Asher MI, Montefort S, Bjorksten B, Lai CK, Strachan DP, Weiland SK, Williams H. Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC Phases One and Three repeat multicountry cross-sectional surveys. *Lancet* 2006: 368(9537): 733-743.

2. von Mutius E, Smits HH. Primary prevention of asthma: from risk and protective factors to targeted strategies for prevention. *The Lancet* 2020: 396(10254): 854-866.

3. Lu C, Norbäck D, Li Y, Deng Q. Early-life exposure to air pollution and childhood allergic diseases: an update on the link and its implications. *Expert review of clinical immunology* 2020: 16(8): 813-827.

4. Lu C, Norbäck D, Zhang Y, Li B, Zhao Z, Huang C, Zhang X, Qian H, Sun Y, Wang J, Liu W, Sundell J, Deng Q. Furry pet-related wheeze and rhinitis in pre-school children across China: Associations with early life dampness and mould, furry pet keeping, outdoor temperature, PM(10) and PM(2.5). *Environ Int* 2020: 144: 106033.

5. von Mutius E, Vercelli D. Farm living: effects on childhood asthma and allergy. *Nature Reviews Immunology* 2010: 10: 861.

6. Ege MJ, Mayer M, Normand AC, Genuneit J, Cookson WO, Braun-Fahrlander C, Heederik D, Piarroux R, von Mutius E, Group GTS. Exposure to environmental microorganisms and childhood asthma. *The New England journal of medicine* 2011: 364(8): 701-709.

7. Fu X, Norbäck D, Yuan Q, Li Y, Zhu X, Hashim JH, Hashim Z, Ali F, Zheng Y-W, Lai X-X,

Spangfort MD, Deng Y, Sun Y. Indoor microbiome, environmental characteristics and asthma among junior high school students in Johor Bahru, Malaysia. *Environment International* 2020: 138: 105664.

8. Fu X, Li Y, Meng Y, Yuan Q, Zhang Z, Wen H, Deng Y, Norbäck D, Hu Q, Zhang X, Sun Y. Derived habitats of indoor microbes are associated with asthma symptoms in Chinese university dormitories. *Environmental research* 2021: 194: 110501.

9. Kirjavainen PV, Karvonen AM, Adams RI, Täubel M, Roponen M, Tuoresmäki P, Loss G, Jayaprakash B, Depner M, Ege MJ, Renz H, Pfefferle PI, Schaub B, Lauener R, Hyvärinen A, Knight R, Heederik DJJ, von Mutius E, Pekkanen J. Farm-like indoor microbiota in non-farm homes protects children from asthma development. *Nature medicine* 2019: 25(7): 1089-1095.

10. Fu X, Ou Z, Zhang M, Meng Y, Li Y, Wen J, Hu Q, Zhang X, Norbäck D, Deng Y, Zhao Z, Sun Y. Indoor bacterial, fungal and viral species and functional genes in urban and rural schools in Shanxi Province, China–association with asthma, rhinitis and rhinoconjunctivitis in high school students. *Microbiome* 2021: 9(1): 138.

11. Locey KJ, Lennon JT. Scaling laws predict global microbial diversity. *Proceedings of the National Academy of Sciences* 2016: 113(21): 5970.

12. Araki A, Kanazawa A, Kawai T, Eitaki Y, Morimoto K, Nakayama K, Shibata E, Tanaka M, Takigawa T, Yoshimura T, Chikara H, Saijo Y, Kishi R. The relationship between exposure to microbial volatile organic compound and allergy prevalence in single-family homes. *The Science of the total environment* 2012: 423: 18-26.

13. Norbäck D, Markowicz P, Cai G-H, Hashim Z, Ali F, Zheng Y-W, Lai X-X, Spangfort MD, Larsson L, Hashim JH. Endotoxin, Ergosterol, Fungal DNA and Allergens in Dust from Schools

in Johor Bahru, Malaysia- Associations with Asthma and Respiratory Infections in Pupils. *PLOS ONE* 2014: 9(2): e88303.

14. Beasley R. Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and atopic eczema: ISAAC. *The Lancet* 1998: 351(9111): 1225-1232.

15. Fu X, Meng Y, Li Y, Zhu X, Yuan Q, Ma'pol A, Hashim JH, Hashim Z, Wieslander G, Zheng Y-W, Lai X-X, Spangfort MD, Wu J, Mu P, Wang J, Norbäck D, Sun Y. Associations between species-level indoor microbiome, environmental characteristics, and asthma in junior high schools of Terengganu, Malaysia. *Air Quality, Atmosphere & Health* 2021.

16. Wang Y, Song F, Zhu J, Zhang S, Yang Y, Chen T, Tang B, Dong L, Ding N, Zhang Q, Bai Z, Dong X, Chen H, Sun M, Zhai S, Sun Y, Yu L, Lan L, Xiao J, Fang X, Lei H, Zhang Z, Zhao W. GSA: Genome Sequence Archive. *Genomics, proteomics & bioinformatics* 2017: 15(1): 14-18.

17. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu Y-X, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML,

Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature biotechnology* 2019: 37(8): 852-857.

18. Quast C, Pruesse E, Gerken J, Peplies J, Yarza P, Yilmaz P, Schweer T, Glöckner FO. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 2012: 41(D1): D590-D596.

19. Koljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AF, Bahram M, Bates ST, Bruns TD, Bengtsson-Palme J, Callaghan TM, Douglas B, Drenkhan T, Eberhardt U, Duenas M, Grebenc T, Griffith GW, Hartmann M, Kirk PM, Kohout P, Larsson E, Lindahl BD, Lucking R, Martin MP, Matheny PB, Nguyen NH, Niskanen T, Oja J, Peay KG, Peintner U, Peterson M, Poldmaa K, Saag L, Saar I, Schussler A, Scott JA, Senes C, Smith ME, Suija A, Taylor DL, Telleria MT, Weiss M, Larsson KH. Towards a unified paradigm for sequence-based identification of fungi. *Mol Ecol* 2013: 22(21): 5271-5277.

20. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. *Nature methods* 2012: 9(8): 811-814.

21. Morton JT, Marotz C, Washburne A, Silverman J, Zaramela LS, Edlund A, Zengler K, Knight R. Establishing microbial composition measurement standards with reference frames.

Nature communications 2019: 10(1): 2719.

22. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Molecular plant* 2020: 13(8): 1194-1202.

23. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH. PubChem Substance and Compound databases. *Nucleic acids research* 2016: 44(D1): D1202-D1213.

24. Saito S, Oku N, Igarashi Y. Mycetoindole, an N-acyl dehydrotryptophan with plant growth inhibitory activity from an actinomycete of the genus Actinomycetospora. *The Journal of Antibiotics* 2022: 75(1): 44-47.

25. Hillwig ML, Zhu Q, Liu X. Biosynthesis of ambiguine indole alkaloids in cyanobacterium Fischerella ambigua. *ACS chemical biology* 2014: 9(2): 372-377.

26. Saito II, Honda H, Kawabe T, Mukumoto F, Shimizu M, Kobayashi T. Comparison of biotin production by recombinant Sphingomonas sp. under various agitation conditions. *Biochemical engineering journal* 2000: 5(2): 129-136.

27. Feng Y, Kumar R, Ravcheev DA, Zhang H. Paracoccus denitrificans possesses two BioR homologs having a role in regulation of biotin metabolism. *MicrobiologyOpen* 2015: 4(4): 644-659.

28. Akishina RI, Voronkova VV, Pomortseva NV. [Effect of the nutrient medium composition on biotin biosynthesis by an Actinomyces species 313-152 culture]. *Prikladnaia biokhimiia i mikrobiologiia* 1982: 18(3): 339-342.

29. Panche AN, Diwan AD, Chandra SR. Flavonoids: an overview. *Journal of nutritional*

science 2016: 5: e47.

30. Yang N, Patil S, Zhuge J, Wen MC, Bolleddula J, Doddaga S, Goldfarb J, Sampson HA, Li XM. Glycyrrhiza uralensis flavonoids present in anti-asthma formula, ASHMI™, inhibit memory Th2 responses in vitro and in vivo. *Phytotherapy research : PTR* 2013: 27(9): 1381-1391.

31. Liu J, Cheng Y, Zhang X, Zhang X, Chen S, Hu Z, Zhou C, Zhang E, Ma S. Astragalin Attenuates Allergic Inflammation in a Murine Asthma Model. *Inflammation* 2015: 38(5): 2007-2016.

32. Yi L, Cui J, Wang W, Tang W, Teng F, Zhu X, Qin J, Wuniqiemu T, Sun J, Wei Y, Dong J. Formononetin Attenuates Airway Inflammation and Oxidative Stress in Murine Allergic Asthma. *Frontiers in Pharmacology* 2020: 11.

33. Kumar P, Lee J-H, Lee J. Diverse roles of microbial indole compounds in eukaryotic systems. *Biological Reviews* 2021: 96(6): 2522-2545.

34. Lee J-H, Lee J. Indole as an intercellular signal in microbial communities. *FEMS Microbiology Reviews* 2010: 34(4): 426-444.

35. Yano JM, Yu K, Donaldson GP, Shastri GG, Ann P, Ma L, Nagler CR, Ismagilov RF, Mazmanian SK, Hsiao EY. Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis. *Cell* 2015: 161(2): 264-276.

36. Bansal T, Alaniz RC, Wood TK, Jayaraman A. The bacterial signal indole increases epithelial-cell tight-junction resistance and attenuates indicators of inflammation. *Proc Natl Acad Sci U S A* 2010: 107(1): 228-233.

37. Vyhlídalová B, Krasulová K, Pečinková P, Marcalíková A, Vrzal R, Zemánková L, Vančo

J, Trávníček Z, Vondráček J, Karasová M, Mani S, Dvořák Z. Gut Microbial Catabolites of Tryptophan Are Ligands and Agonists of the Aryl Hydrocarbon Receptor: A Detailed Characterization. *Int J Mol Sci* 2020: 21(7).

38. Poulain-Godefroy O, Bouté M, Carrard J, Alvarez-Simon D, Tsicopoulos A, de Nadai P. The Aryl Hydrocarbon Receptor in Asthma: Friend or Foe? *Int J Mol Sci* 2020: 21(22).

39. Fu X, Ou Z, Zhang M, Meng Y, Li Y, Chen Q, Jiang J, Zhang X, Norbäck D, Zhao Z, Sun Y. Classroom microbiome, functional pathways and sick-building syndrome (SBS) in urban and rural schools - Potential roles of indoor microbial amino acids and vitamin metabolites. *The Science of the total environment* 2021: 795: 148879.

40. Min HJ, Nam J-W, Yu ES, Hong J-H, Seo E-K, Hwang ES. Effect of naturally occurring hydroxychavicol acetate on the cytokine production in T helper cells. *International Immunopharmacology* 2009: 9(4): 448-454.

41. Li MC, Chen CH, Guo YL. Phthalate esters and childhood asthma: A systematic review and congener-specific meta-analysis. *Environ Pollut* 2017: 229: 655-660.

42. Zhang J, Sun C, Lu R, Zou Z, Liu W, Huang C. Associations between phthalic acid esters in household dust and childhood asthma in Shanghai, China. *Environmental research* 2021: 200: 111760.

43. Choi H, Schmidbauer N, Bornehag CG. Volatile organic compounds of possible microbial origin and their risks on childhood asthma and allergies within damp homes. *Environment International* 2017: 98: 143-151.

44. Williams LK, Ownby DR, Maliarik MJ, Johnson CC. The role of endotoxin and its receptors in allergic disease. *Annals of Allergy, Asthma & Immunology* 2005: 94(3): 323-332.

45. van Strien RT, Engel R, Holst O, Bufe A, Eder W, Waser M, Braun-Fahrländer C, Riedler J, Nowak D, von Mutius E, the AST. Microbial exposure of rural school children, as assessed by levels of N-acetyl-muramic acid in mattress dust, and its association with respiratory health. *Journal of Allergy and Clinical Immunology* 2004: 113(5): 860-867.

46. Karvonen AM, Hyvarinen A, Rintala H, Korppi M, Taubel M, Doekes G, Gehring U, Renz H, Pfefferle PI, Genuneit J, Keski-Nisula L, Remes S, Lampi J, von Mutius E, Pekkanen J. Quantity and diversity of environmental microbial exposure and development of asthma: a birth cohort study. *Allergy* 2014: 69(8): 1092-1101.

47. Remy S, Verstraelen S, Van Den Heuvel R, Nelissen I, Lambrechts N, Hooyberghs J, Schoeters G. Gene expressions changes in bronchial epithelial cells: markers for respiratory sensitizers and exploration of the NRF2 pathway. *Toxicology in vitro : an international journal published in association with BIBRA* 2014: 28(2): 209-217.

48. Marchesi JR, Holmes E, Khan F, Kochhar S, Scanlan P, Shanahan F, Wilson ID, Wang Y. Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. *Journal of proteome research* 2007: 6(2): 546-551.

49. Azevedo LF, Porto Dechandt CR, Cristina de Souza Rocha C, Hornos Carneiro MF, Alberici LC, Barbosa F, Jr. Long-term exposure to bisphenol A or S promotes glucose intolerance and changes hepatic mitochondrial metabolism in male Wistar rats. *Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association* 2019: 132: 110694.

Table 1. Prevalence of asthma symptoms (wheeze) among junior high school students in three centers of Malaysia. In each center, four schools were defined as schools with high prevalence of asthma symptoms, and four schools were defined as schools with low prevalence of asthma symptoms. P values were calculated by chi-square test.

		Number of students	Prevalence (%)	Total number of students	Prevalence (%)	p	Male (%)	Female (%)	p
Johor Bahru	Schools with high prevalence	57	15.8	236	14.8	0.0007	45.5	54.5	0.051
		60	13.3						
		59	11.9						
		60	18.3						
	Schools with low prevalence	56	7.1	226	5.3		54.5	45.5	
		54	1.9						
		56	3.6						
		60	8.3						
Terengganu	Schools with high prevalence	63	12.7	241	19.5	0.0004	40.7	59.3	0.41
		55	20.0						
		56	14.3						
		67	29.9						
	Schools with low prevalence	45	8.9	222	8.1		36.9	63.1	
		46	8.7						
		54	5.6						
		77	9.1						
Penang	Schools with high prevalence	56	23.2	158	21.6	0.002	49.4	50.6	0.21
		35	20.0						
		22	18.2						
		44	22.7						
	Schools with low prevalence	52	13.5	210	10.0		42.9	57.1	
		47	10.6						
		42	9.5						
		67	7.5						

Table 2. Environmental characteristics in junior high schools in three centers of Malaysia.

Environmental characteristics	Johor Bahru	Terengganu	Penang
	Mean (min, max)	Mean (min, max)	Mean (min, max)
Indoor temperature (℃)	29.0 (27.3, 30.9)	29.5 (28.3, 32.8)	28.1 (25.4, 30.3)
Indoor relative humidity (%)	69.9 (62.9, 76.9)	72.4 (63.8, 79,9)	78.9 (69.7, 88.2)
Indoor CO_2 (ppm)	492 (376, 689)	432 (397, 600)	425 (360, 720)
Indoor NO_2 (μg/m^3)	23.4 (13.0, 41.7)	8.2 (2.8, 14.6)	22.9 (20.5, 29.9)

Table 3. Flavonoids and isoflavonoids enriched in schools with low asthma prevalence in three centers of Malaysia. The potential protective flavonoids and isoflavonoids were defined as enriched in schools with low asthma prevalence compared with schools with high asthma prevalence ($p < 0.01$, FDR < 0.1, fold change > 2). The representative plants and suggested health effects were mainly acquired from PubChem. It should be noted that these are only representative plants, not all plants that can produce these flavonoids and isoflavonoids.

Enriched center	Classification	Name	Molecular formula	Representative plants	Suggested health effects
Commonly enriched in Johor Bahru, Terengganu and Penang	Flavonoids	Isoliquiritigenin	$C_{15}H_{12}O_4$	*Glycyrrhiza uralensis*, *Mongolian glycyrrhiza*, *Glycyrrhiza glabra*	antioxidant, anti-inflammatory, GABA (gamma-aminobutyric acid) modulator
	Isoflavonoids	Formononetin	$C_{16}H_{12}O_4$	*Trifolium pratense*	antioxidant
		6-Hydroxydaidzein	$C_{15}H_{10}O_5$	soybean	anti-inflammatory
Johor Bahru	Flavonoids	Afzelechin	$C_{15}H_{14}O_5$	*Bergenia ligulata*	
		Amentoflavone	$C_{30}H_{18}O_{10}$	*Ginkgo biloba*, *Hypericum perforatum*	anti-allergic
		Astragalin	$C_{21}H_{20}O_{11}$	*Phytolacca americana*, *Phegopteris connectilis*	anti-inflammatory, antioxidant, anti-atopic dermatitis
		Baicalin	$C_{21}H_{18}O_{11}$	Chinese herbal medicine *Scutellaria baicalensis*	anti-allergic, attenuated serum IgE and effector T cells
		Diosmetin	$C_{16}H_{12}O_6$	citrus	anti-allergic
		Epicatechin	$C_{15}H_{14}O_6$	*Pentace burmanica*	antioxidant, reduces the resistance to insulin
		Tangeritin	$C_{20}H_{20}O_7$	citrus, tangerines	antioxidant

	Isoflavonoids	Glycitein	$C_{16}H_{12}O_5$	soybean	improve sleep quality
Terengganu	Flavonoids	(2S)-Liquiritigenin	$C_{15}H_{12}O_4$	*Glycyrrhizae uralensis*	antioxidant
	Isoflavonoids	Daidzein	$C_{15}H_{10}O_4$	soy plants	anti-inflammatory
Penang	Flavonoids	Luteolin	$C_{15}H_{10}O_6$	Pteridophyta, Bryophyta, Magnoliophyta, Pinophyta	anti-inflammatory, immune system modulator
		Procyanidin B2	$C_{30}H_{26}O_{12}$	*Cinchona pubescens*, *Cinnamomum verum*, *Crataegus monogyna*, *Uncaria guianensis*, *Vitis vinifera*, *Litchi chinensis*, *Ecdysanthera utilis*	therapeutic for acute lung injury
		Quercetin 3-(6-malonyl-glucoside)	$C_{24}H_{22}O_{15}$	*Lactuca sativum*	
	Isoflavonoids	Ononin	$C_{22}H_{22}O_9$	*Trifolium pratense*	

Table 4. Potential risk environmental chemicals in three centers of Malaysia. The potential risk environmental chemicals were defined as chemicals enriched in schools with high asthma prevalence compared with schools with low asthma prevalence ($p < 0.01$, FDR < 0.12, fold change > 2) in each center. The characteristics information of chemicals were searched in the PubChem database from NCBI. Chemicals annotated as drug or common human, animal or plant metabolites were not included in the table. GHS classification, including GHS hazard code, hazard statements and notified classification ratio, was presented in the table. GHS hazard code between H310 and H336 were presented in the table as these hazard statements were mainly related to allergic irritations.

	Environmental chemicals	Molecular formula	Key features and manufacturing products	GHS hazard statements	Symptoms, disorders and diseases	Experimental evidence
Johor Bahru	2(3H)-benzothiazolethione	$C_7H_5NS_2$	industrial materials, paints, pesticide	H317: May cause an allergic skin reaction	dermatitis, hypersensitivity	[47]
Terengganu	2-aminobenzoic acid	$C_7H_7NO_2$	dyes, perfumes	H318 (42.55%): Causes serious eye damage H319 (57.45%): Causes serious eye irritation	nausea, allergic reaction, skin irritation, itching and dermatitis	
Penang	isovaleric acid	$C_5H_{10}O_2$	fragrance, perfumes	H314 (100%): Causes severe skin burns and eye damage H318 (84.07%): Causes serious eye damage	inflammatory bowel disease, autism	[48]
	phthalic acid	$C_8H_6O_4$	cleaning products, laundry,	H315 (96.4%): Causes skin irritation	eyes, skin, respiratory tract and mucous membrane irritation	

			pesticides, plastic products	H319 (94.96%): Causes serious eye irritation H335 (95.68%): May cause respiratory irritation		
	4,4-sulfonyldiphenol	$C_{12}H_{10}O_4S$	electroplating solvent, washfastening agent, metabolite and endocrine disruptor	H319 (15.38%): Causes serious eye irritation	obesity, glucose intolerance	[49]

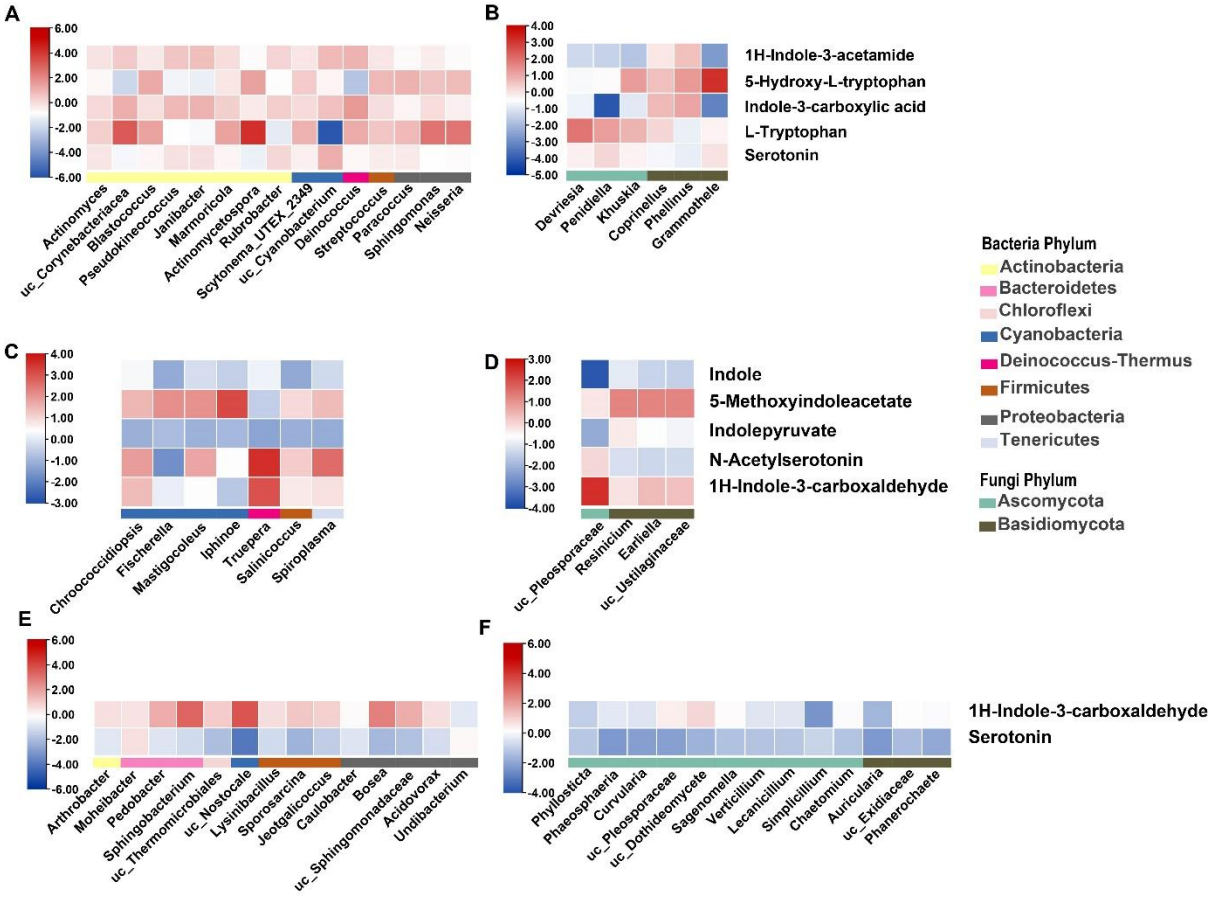

Figure 1. Co-occurrence probability of potential protective microbial taxa and indoles and derivatives in three centers of Malaysia. The potential protective microbial taxa were presented at the x-axis of the heat plot, and the potential protective metabolites were presented at the y-axis. The potential protective taxa were defined as bacterial and fungal taxa enriched in schools with low asthma prevalence (LDA > 3 in Johor Bahru and LDA > 2 in Terengganu and Penang). The potential protective metabolites were defined as indoles and derivatives enriched in schools with low asthma prevalence ($p < 0.01$, FDR < 0.1, fold change > 2). Co-occurrence probability of potential protective (A) bacterial and (B) fungal taxa and metabolites in Johor Bahru. Co-occurrence probability of potential protective (C) bacterial and (D) fungal taxa and metabolites in Terengganu. Co-occurrence probability of potential protective (E) bacterial and (F) fungal taxa and metabolites in Penang.

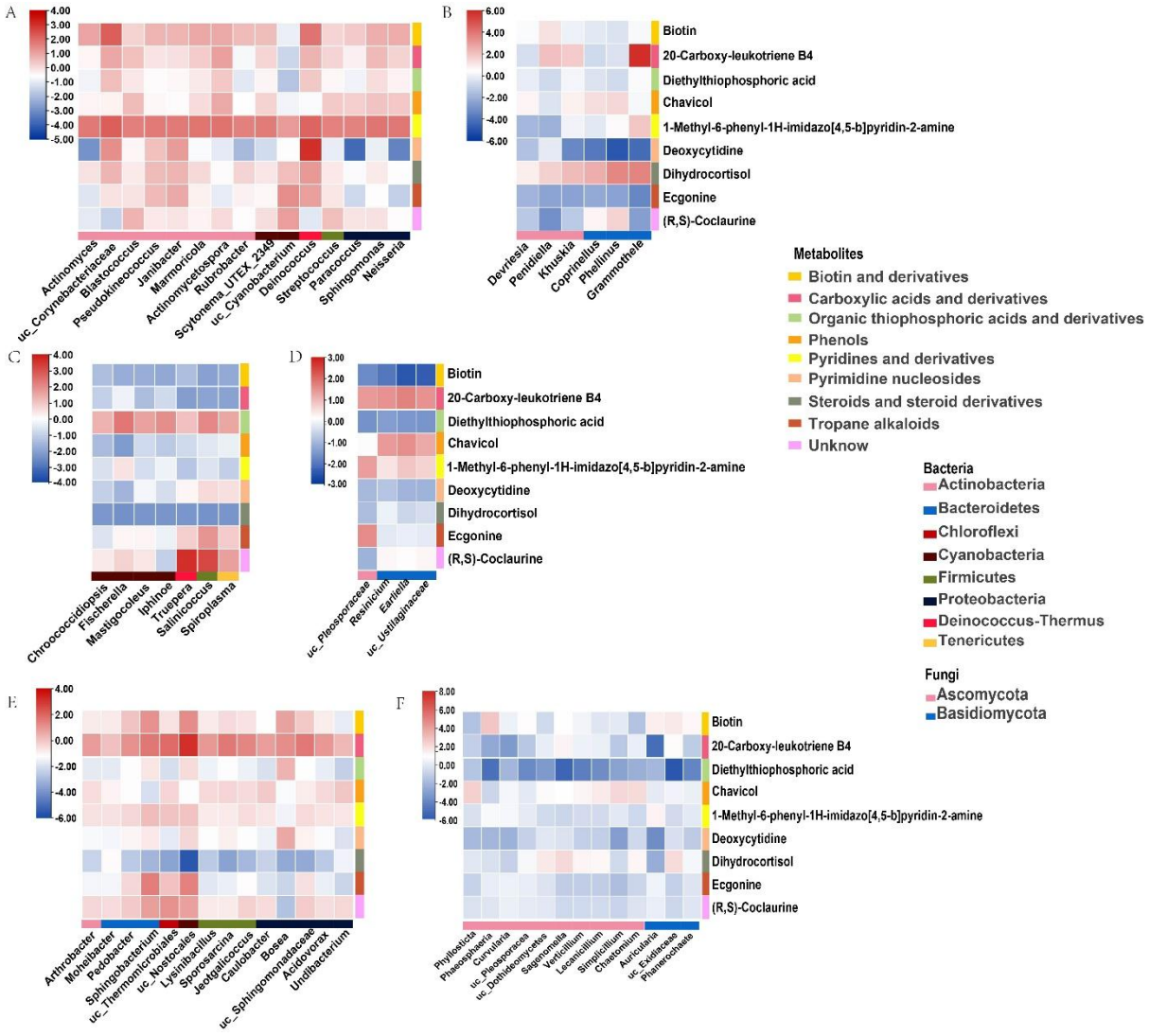

Figure 2. Co-occurrence probability of potential protective microbial taxa and other protective metabolite classes in three centers of Malaysia. The potential protective microbial taxa were presented at the x-axis of the heat plot, and the potential protective metabolites were presented at the y-axis. The potential protective taxa were defined as bacterial and fungal taxa enriched in schools with low asthma prevalence (LDA > 3 in Johor Bahru and LDA > 2 in Terengganu and Penang). The potential protective microbial metabolites were defined as metabolites enriched in schools with low asthma prevalence compared with schools with high asthma prevalence ($p < 0.01$, FDR < 0.2, fold change > 2) in each center. The potential protective metabolites were categorized into HMDB class, and the metabolite classes were also presented in the table. Co-occurrence probability of potential protective (A) bacterial and (B) fungal taxa and metabolites in Johor Bahru. Co-occurrence probability of potential protective (C) bacterial and (D) fungal taxa and metabolites in Terengganu. Co-occurrence probability of potential protective (E) bacterial and (F) fungal taxa and metabolites in Penang.

Supplementary Figures

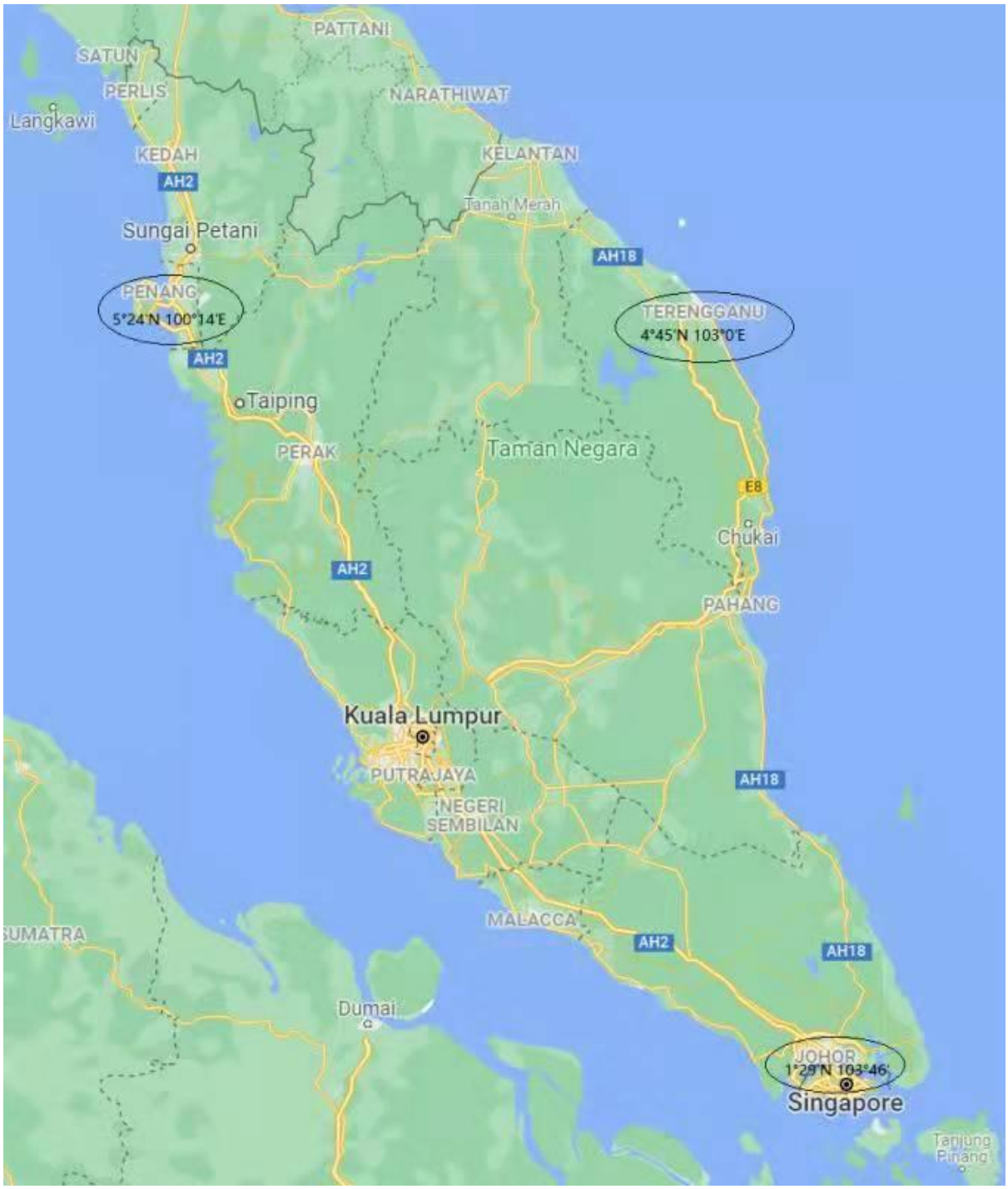

Figure S1. Geographic locations of Johor Bahru, Terengganu and Penang.

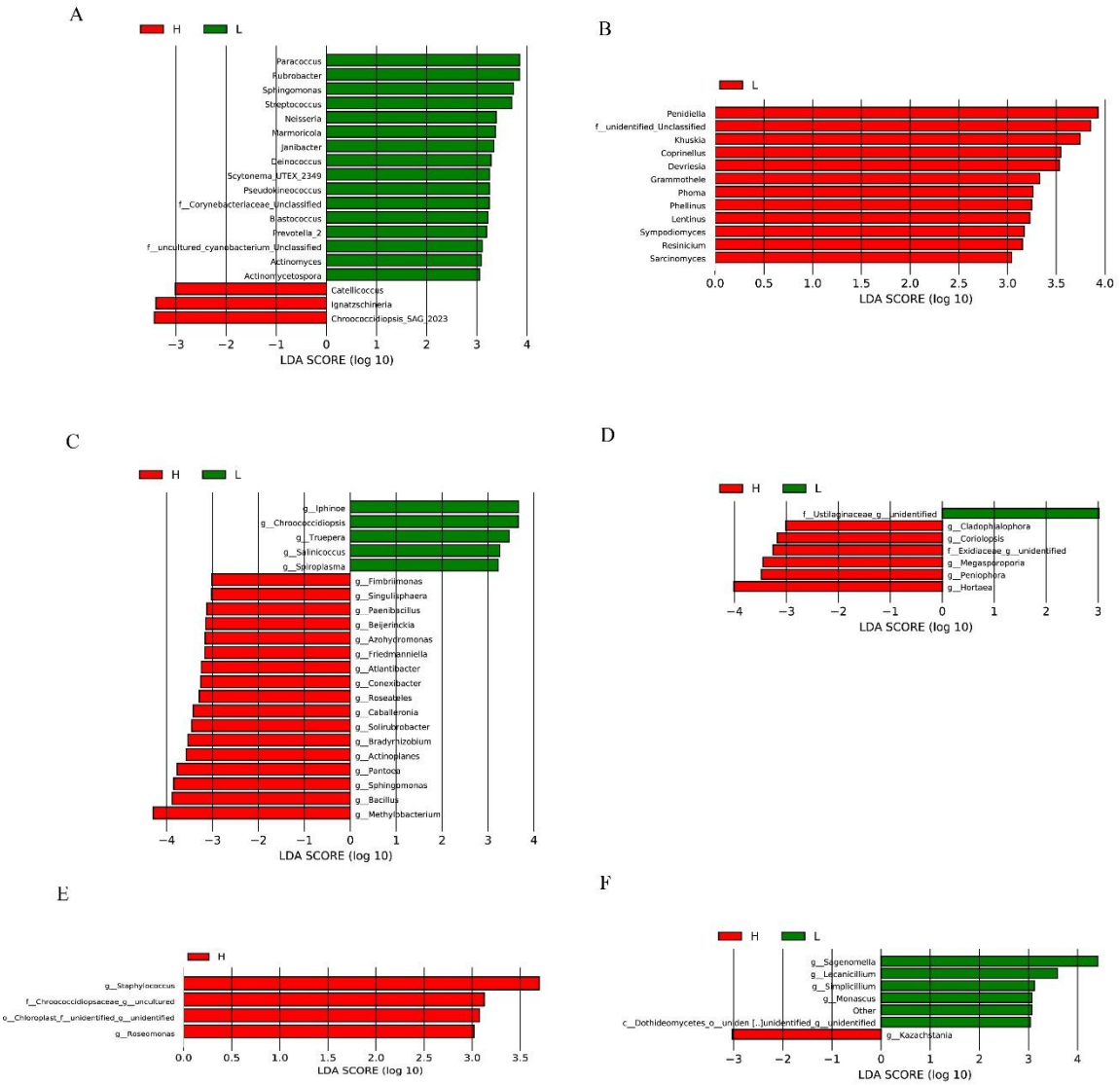

Figure S2. Microbial taxa enriched in schools with low asthma prevalence (L) and schools with high asthma prevalence (H). The analysis was calculated by LEfSe (linear discriminant analysis effect size), and only taxa with LDA > 3 were presented in the figure. Enriched (A) bacterial and (B) fungal taxa in Johor Bahru. Enriched (C) bacterial and (D) fungal taxa in Terengganu. Enriched (E) bacterial and (F) fungal taxa in Penang.

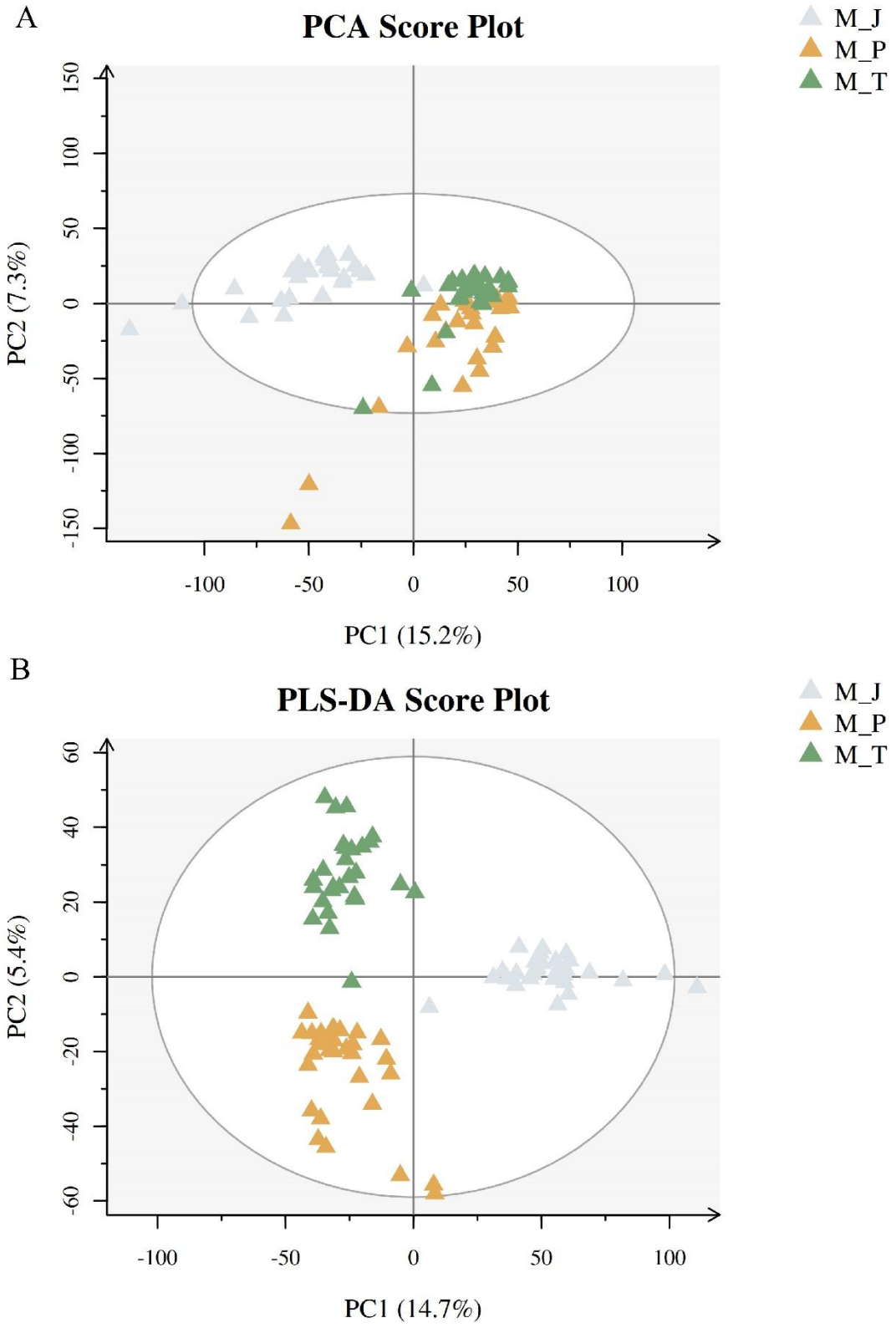

Figure S3. Ordination plot of environmental chemicals and metabolites in classroom dust in three centers of Malaysia. (A) Ordination plot by PCA (principal component analysis). (B) Ordination plot by PLS-DA (partial least squares discriminant analysis).

Table 1. Abundance of bacterial genera in Johor Bahru. Genera with abundance > 0.2% were shown in this table.

Taxon	average abundance
Bacillus	4.220476
Staphylococcus	3.371429
Paracoccus	3.131905
Sphingomonas	2.760476
Saccharopolyspora	2.5
Acinetobacter	2.21
Enterobacter	2.209048
Lactobacillus	1.956667
f__Chroococcidiopsaceae_Unclassified	1.909524
Corynebacterium_1	1.77381
Bacteroides	1.670952
Rubellimicrobium	1.654286
Pantoea	1.56381
Streptococcus	1.45
Kocuria	1.43
Brevibacterium	1.419048
Methylobacterium	1.32619
Rubrobacter	1.116667
Haemophilus	1.013333
Pseudonocardia	0.963333
Deinococcus	0.949524
Chroococcidiopsis_SAG_2023	0.938571
Enterococcus	0.921429
Micrococcus	0.819048
Calothrix_PCC-6303	0.808571
Pseudomonas	0.801429
Massilia	0.76381
Klebsiella	0.747143
Nephrolepis_biserrata_var._furcans	0.720476
Brachybacterium	0.709524
Skermanella	0.706667
Parabacteroides	0.661905
Janibacter	0.644286
Roseomonas	0.640476
Actinomycetospora	0.589524
Escherichia-Shigella	0.575714
Truepera	0.544286
MN_122.2a	0.54381
Geodermatophilus	0.53381
Curtobacterium	0.508095
Scytonema_UCFS19	0.491429
Arthrobacter	0.484762
Chroococcidiopsis_PCC_7203	0.461429
Nocardioides	0.457143
Leptolyngbya_PCC-6306	0.455238
Nesterenkonia	0.435714
CENA359	0.431905
Marmoricola	0.428571
Clostridium_sensu_stricto_1	0.420952
Brevundimonas	0.407619
Blastococcus	0.406667
Enhydrobacter	0.392857
Exiguobacterium	0.374286
1174-901-12	0.369048
Scytonema_VB-61278	0.367143
Gordonia	0.359048
Hymenobacter	0.356667
Dapisostemonum_CCIBt_3536	0.34619
Lachnoclostridium	0.324286
Pseudokineococcus	0.322381
Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium	0.32
Neisseria	0.291905
f__uncultured_Unclassified	0.283333
Ignatzschineria	0.254286
Weissella	0.253333
Kurthia	0.249524
Stenotrophomonas	0.243333
Mycobacterium	0.244286
Proteus	0.241429
Romboutsia	0.241905
Glutamicibacter	0.231429
Bryum_argenteum_var._argenteum	0.225238
Scytonema_UTEX_2349	0.221905
f__uncultured_cyanobacterium_Unclassified	0.217619
Rothia	0.205238
Mastigocladopsis_PCC-10914	0.20381
Shimwellia	0.203333
Quadrisphaera	0.202381

Table 2. Abundance of bacterial genera in Terengganu. Genera with abundance > 0.2% were shown in this table.

Taxon	average
k__Bacteria;p__Deinococcus-Thermus;c__Deinococci;o__Deinococcales;f__Deinococcaceae;g__Deinococcus	6.537443336
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Acinetobacter	5.292570074
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Pseudonocardiales;f__Pseudonocardiaceae;g__Saccharopolyspora	4.461683511
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Methylobacteriaceae;g__Methylobacterium	3.765105015
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Massilia	3.618960111
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Sphingomonas	3.185104162
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Paracoccus	3.046318174
k__Bacteria;p__Cyanobacteria;c__unidentified;o__Chroococcidiopsidales;f__Chroococcidiopsidaceae;g__Aliterella	2.737348837
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus	2.438727302
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Staphylococcaceae;g__Staphylococcus	2.330050814
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacterales;f__Enterobacteriaceae;g__Klebsiella	2.322404537
k__Bacteria;p__Cyanobacteria;c__unidentified;o__Nostocales;f__Scytonemataceae;g__Brasilonema	1.966929381
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__unidentified;g__Exiguobacterium	1.823037326
k__Bacteria;p__Cyanobacteria;c__unidentified;o__Nostocales;f__Symphyonemataceae;g__Loriellopsis	1.721634528
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas	1.67150108
k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Solirubrobacterales;f__Solirubrobacteraceae;g__Solirubrobacter	1.468153909
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacterales;f__Erwiniaceae;g__Pantoea	1.191780912
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Rubellimicrobium	1.161815735
k__Bacteria;p__Cyanobacteria;c__unidentified;o__Nostocales;f__Stigonemataceae;g__Iphinoe	1.016295225
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides	0.945630541
k__Bacteria;p__Cyanobacteria;c__unidentified;o__Nostocales;f__Nostocaceae;g__Cylindrospermum	0.918926255
k__Bacteria;p__Cyanobacteria;c__unidentified;o__Chroococcidiopsidales;f__Chroococcidiopsidaceae;g__Chroococcidiopsis	0.9045298
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacterales;f__Enterobacteriaceae;g__Enterobacter	0.849928678
k__Bacteria;p__Cyanobacteria;c__unidentified;o__Nostocales;f__Rivulariaceae;g__Calothrix	0.838409994
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Micrococcales;f__Dermabacteraceae;g__Brachybacterium	0.804444275
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Acetobacteraceae;g__Roseomonas	0.748089065
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Acetobacteraceae;g__Craurococcus	0.741618863
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Micrococcales;f__Micrococcaceae;g__Kocuria	0.690270394
k__Bacteria;p__Deinococcus-Thermus;c__Deinococci;o__Deinococcales;f__Trueperaceae;g__Truepera	0.688062627
k__Bacteria;p__Cyanobacteria;c__unidentified;o__Synechococcales;f__Chamaesiphonaceae;g__Chamaesiphon	0.657603128
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Parabacteroides	0.575715058
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae;g__Pelomonas	0.57129543
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Pseudonocardiales;f__Pseudonocardiaceae;g__Pseudonocardia	0.562592028

k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Neisseriales;f__Neisseriaceae;g__Neisseria	0.552882876
k__Bacteria;p__Actinobacteria;c__Rubrobacteria;o__Rubrobacterales;f__Rubrobacteraceae;g__Rubrobacter	0.5471647
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Enterococcaceae;g__Enterococcus	0.535397028
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Micromonosporales;f__Micromonosporaceae;g__Actinoplanes	0.526771806
k__Bacteria;p__Chloroflexi;c__Ktedonobacteria;o__Thermogemmatisporales;f__Thermogemmatisporaceae;g__Thermogemmatispora	0.513728512
k__Bacteria;p__Actinobacteria;c__Thermoleophilia;o__Solirubrobacterales;f__Conexibacteraceae;g__Conexibacter	0.503019063
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__Stenotrophomonas	0.495094483
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Bradyrhizobiaceae;g__Bradyrhizobium	0.494072543
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Planococcaceae;g__Kurthia	0.473385722
k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__Archangiaceae;g__Cystobacter	0.463162202
k__Bacteria;p__Acidobacteria;c__Blastocatellia;o__unidentified;f__unidentified;g__Aridibacter	0.45518448
k__Bacteria;p__Cyanobacteria;c__unidentified;o__Pleurocapsales;f__Dermocarpellaceae;g__Stanieria	0.434673561
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus	0.434545077
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacterales;f__Enterobacteriaceae;g__Escherichia	0.430508454
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Planococcaceae;g__Sporosarcina	0.42988234
k__Bacteria;p__Cyanobacteria;c__unidentified;o__Nostocales;f__Chlorogloeopsidaceae;g__Chlorogloeopsis	0.427386389
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Pseudonocardiales;f__Pseudonocardiaceae;g__Actinomycetospora	0.419646501
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Brucellaceae;g__Ochrobactrum	0.416855389
k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cytophagaceae;g__Larkinella	0.403917253
k__Bacteria;p__Cyanobacteria;c__unidentified;o__Synechococcales;f__Heteroleibleiniaceae;g__Tapinothrix	0.398779015
k__Bacteria;p__Cyanobacteria;c__unidentified;o__Nostocales;f__Nostocaceae;g__Anabaena	0.390035411
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Micrococcales;f__Microbacteriaceae;g__Microbacterium	0.385483173
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__unidentified;g__Roseateles	0.377555818
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Micrococcales;f__Intrasporangiaceae;g__Janibacter	0.375522586
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Micrococcales;f__Micrococcaceae;g__Micrococcus	0.37450469
k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__Flavobacteriaceae;g__Chryseobacterium	0.363220942
k__Bacteria;p__Chloroflexi;c__Ktedonobacteria;o__Ktedonobacterales;f__Ktedonobacteraceae;g__Ktedonobacter	0.354106442
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Propionibacteriales;f__Nocardioidaceae;g__Nocardioides	0.320856501
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacterales;f__Enterobacteriaceae;g__Yokenella	0.311385784
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Beijerinckiaceae;g__Beijerinckia	0.306785685
k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Cytophagaceae;g__Spirosoma	0.30476205
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__Luteimonas	0.304743082
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacterales;f__Yersiniaceae;g__Serratia	0.303949514
k__Bacteria;p__Acidobacteria;c__Blastocatellia;o__unidentified;f__unidentified;g__Blastocatella	0.301344557

k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Micrococcales;f__Brevibacteriaceae;g__Brevibacterium	0.293343704
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__Lysobacter	0.292612367
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Caballeronia	0.29155493
k__Bacteria;p__Bacteroidetes;c__Cytophagia;o__Cytophagales;f__Hymenobacteraceae;g__Hymenobacter	0.283713867
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Micrococcales;f__Microbacteriaceae;g__Curtobacterium	0.282228101
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Undibacterium	0.278684879
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacterales;f__Enterobacteriaceae;g__Kluyvera	0.275502421
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Rhodospirillaceae;g__Skermanella	0.275270584
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Enterococcaceae;g__Catellicoccus	0.269073884
k__Bacteria;p__Chloroflexi;c__Ktedonobacteria;o__Ktedonobacterales;f__Thermosporotrichaceae;g__Thermosporothrix	0.252599794
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Clostridium	0.250065654
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Amaricoccus	0.246686184
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacterales;f__Enterobacteriaceae;g__Leclercia	0.246382702
k__Bacteria;p__Bacteroidetes;c__Chitinophagia;o__Chitinophagales;f__Chitinophagaceae;g__Cnuella	0.241844628
k__Bacteria;p__Cyanobacteria;c__unidentified;o__Synechococcales;f__Leptolyngbyaceae;g__Trichocoleus	0.237963604
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Propionibacteriales;f__Nocardioidaceae;g__Friedmanniella	0.235277854
k__Bacteria;p__Bacteroidetes;c__Chitinophagia;o__Chitinophagales;f__Chitinophagaceae;g__Flavisolibacter	0.228151504
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacterales;f__Enterobacteriaceae;g__Cedecea	0.22662697
k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Gemmatimonadales;f__Gemmatimonadaceae;g__Gemmatirosa	0.226111045
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Macellibacteroides	0.219605521
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Rhodospirillaceae;g__Nivispirillum	0.215605749
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Methylobacteriaceae;g__Microvirga	0.207863604
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Moraxella	0.206799677
k__Bacteria;p__Tenericutes;c__Mollicutes;o__Entomoplasmatales;f__Spiroplasmataceae;g__Spiroplasma	0.203618704

Table 3. Abundance of bacterial genera in Penang. Genera with abundance > 0.2% were shown in this table.

Taxon	average
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Betaproteobacteriales;f__Burkholderiaceae;g__Pelomonas	18.92102117
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Betaproteobacteriales;f__Burkholderiaceae;g__Ralstonia	7.753735523
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas	4.854018676
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Rhizobiaceae;g__Ochrobactrum	2.674530295
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Acinetobacter	2.15594403
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Sphingomonas	2.116254487
k__Bacteria;p__Deinococcus-Thermus;c__Deinococci;o__Thermales;f__Thermaceae;g__Thermus	1.757560717
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Betaproteobacteriales;f__Burkholderiaceae;g__Cupriavidus	1.64675641
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__unidentified	1.487206986
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Betaproteobacteriales;f__Burkholderiaceae;g__Paucibacter	1.475461753
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Acetobacterales;f__Acetobacteraceae;g__Craurococcus	1.319640009
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Pseudonocardiales;f__Pseudonocardiaceae;g__Amycolatopsis	1.278841454
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Serratia	0.856200939
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Sphingobium	0.853209534
k__Bacteria;p__Deinococcus-Thermus;c__Deinococci;o__Deinococcales;f__Deinococcaceae;g__Deinococcus	0.750203931
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Staphylococcaceae;g__Staphylococcus	0.677614512
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Rhizobiaceae;g__Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium	0.655342744
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Salinisphaerales;f__Salinisphaeraceae;g__Salinisphaera	0.634812188
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Beijerinckiaceae;g__Methylobacterium	0.54702058
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Psychrobacter	0.540748458
k__Bacteria;p__Deinococcus-Thermus;c__Deinococci;o__Deinococcales;f__Trueperaceae;g__Truepera	0.525798997
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides	0.509026195
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Xanthobacteraceae;g__Bradyrhizobium	0.485658344
k__Bacteria;p__Gemmatimonadetes;c__Gemmatimonadetes;o__Gemmatimonadales;f__Gemmatimonadaceae;g__Gemmatirosa	0.4734531
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Caulobacterales;f__Caulobacteraceae;g__Brevundimonas	0.446256335
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Betaproteobacteriales;f__Burkholderiaceae;g__Burkholderia-Caballeronia-Paraburkholderia	0.435826758
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Blastomonas	0.401100108
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Chitinophagales;f__Chitinophagaceae;g__unidentified	0.380749342

k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Rubellimicrobium	0.366002455
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Frankiales;f__Geodermatophilaceae;g__Geodermatophilus	0.359630569
k__Bacteria;p__Cyanobacteria;c__Oxyphotobacteria;o__Chloroplast;f__unidentified;g__unidentified	0.354129718
k__Bacteria;p__Chlamydiae;c__Chlamydiae;o__Chlamydiales;f__uncultured_organism;g__uncultured_organism	0.335992819
k__Bacteria;p__Cyanobacteria;c__Oxyphotobacteria;o__Nostocales;f__Chroococcidiopsaceae;g__uncultured	0.332063882
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae_1;g__Clostridium_sensu_stricto_1	0.331780145
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus	0.330944923
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Pseudonocardiales;f__Pseudonocardiaceae;g__Pseudonocardia	0.323239769
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__Paracoccus	0.282922217
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Lachnoclostridium_5	0.264166998
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Pseudonocardiales;f__Pseudonocardiaceae;g__Saccharopolyspora	0.258157337
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Pseudonocardiales;f__Pseudonocardiaceae;g__Actinomycetospora	0.241160592
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Caulobacterales;f__Caulobacteraceae;g__Caulobacter	0.231461954
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Micrococcales;f__Micrococcaceae;g__Kocuria	0.219524621
k__Bacteria;p__Actinobacteria;c__Rubrobacteria;o__Rubrobacterales;f__Rubrobacteriaceae;g__Rubrobacter	0.218029983
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Beijerinckiaceae;g__1174-901-12	0.214829595
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Betaproteobacteriales;f__Burkholderiaceae;g__Aquabacterium	0.214326214
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Micrococcales;f__Micrococcaceae;g__Tersicoccus	0.210499609
k__Bacteria;p__Chloroflexi;c__Anaerolineae;o__SBR1031;f__anaerobic_digester_metagenome;g__anaerobic_digester_metagenome	0.210086212
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Betaproteobacteriales;f__Burkholderiaceae;g__Noviherbaspirillum	0.202442665
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus	0.201945435
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Acetobacterales;f__Acetobacteraceae;g__Roseomonas	0.200069227

Table 4. Abundance of fungi genera in Johor Bahru. Genera with abundance > 0.2% were shown in this table.

Taxon	average abundance
Aspergillus	16.59381
Penicillium	10.15476
Hortaea	7.98
Cladosporium	7.811429
Wallemia	6.569048
Emericella	4.059048
f__unidentified_	3.18381
Penidiella	3.02381
Khuskia	1.647619
Candida	1.256667
Grammothele	0.598095
Sagenomella	0.540476
Nigrospora	0.488095
Schizophyllum	0.411429
Bipolaris	0.377619
Eurotium	0.362381
Mycosphaerella	0.342857
Fusarium	0.339048
Coprinellus	0.328571
Trichosporon	0.309524
Resinicium	0.30381
Corynespora	0.302381
Letendraea	0.291429
Trichomonascus	0.277619
Kodamaea	0.251429
Choanephora	0.240952
Phellinus	0.238095
Devriesia	0.23619
Aureobasidium	0.214286
Gloeotinia	0.209048
Auricularia	0.208571

Table 5. Abundance of fungi genera in Terengganu. Genera with abundance > 0.2% were shown in this table.

Taxon	average
k__Fungi;p__Ascomycota;c__Eurotiomycetes;o__Eurotiales;f__Aspergillaceae;g__Aspergillus	12.74965657
k__Fungi;p__Basidiomycota;c__Wallemiomycetes;o__Wallemiales;f__Wallemiales_fam_Incertae_sedis;g__Wallemia	12.23770145
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Pleosporales;f__unidentified;g__unidentified	7.080850243
k__Fungi;p__Basidiomycota;c__Agaricostilbomycetes;o__Agaricostilbales;f__Agaricostilbaceae;g__Sterigmatomyces	5.822782511
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Capnodiales;f__Teratosphaeriaceae;g__Eupenidiella	5.301634439
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Valsariales;f__Valsariaceae;g__Bambusaria	5.227068667
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Dothideales;f__Dothideales_fam_Incertae_sedis;g__Hortaea	3.326543969
k__Fungi;p__Ascomycota;c__Sordariomycetes;o__Hypocreales;f__Nectriaceae;g__Fusarium	2.893140899
k__Fungi;p__Ascomycota;c__Eurotiomycetes;o__Eurotiales;f__unidentified;g__unidentified	2.643221802
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Capnodiales;f__Neodevriesiaceae;g__Neodevriesia	2.073555375
k__Fungi;p__Basidiomycota;c__Agaricomycetes;o__Hymenochaetales;f__Hymenochaetaceae;g__Asterodon	1.879529575
k__Fungi;p__Ascomycota;c__Saccharomycetes;o__Saccharomycetales;f__Saccharomycetales_fam_Incertae_sedis;g__Candida	1.774365778
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Pleosporales;f__Pleosporaceae;g__Curvularia	1.550263012
k__Fungi;p__unidentified;c__unidentified;o__unidentified;f__unidentified;g__unidentified	1.452285268
k__Fungi;p__Ascomycota;c__Sordariomycetes;o__unidentified;f__unidentified;g__unidentified	1.409097218
k__Fungi;p__Basidiomycota;c__Agaricomycetes;o__unidentified;f__unidentified;g__unidentified	1.360330776
k__Fungi;p__Ascomycota;c__Ascomycota_cls_Incertae_sedis;o__Ascomycota_ord_Incertae_sedis;f__Ascomycota_fam_Incertae_sedis;g__Engyodontium	1.291404311
k__Fungi;p__Ascomycota;c__unidentified;o__unidentified;f__unidentified;g__unidentified	1.246606611
k__Fungi;p__Ascomycota;c__Eurotiomycetes;o__Eurotiales;f__Elaphomycetaceae;g__Elaphomyces	1.242093911
k__Fungi;p__Ascomycota;c__Eurotiomycetes;o__Eurotiales;f__Trichocomaceae;g__Sagenomella	1.231864943
k__Fungi;p__Basidiomycota;c__Agaricomycetes;o__Russulales;f__Peniophoraceae;g__Peniophora	1.112779804
k__Fungi;p__Ascomycota;c__Eurotiomycetes;o__Eurotiales;f__Aspergillaceae;g__unidentified	0.919948251
k__Fungi;p__Basidiomycota;c__Agaricomycetes;o__Polyporales;f__Schizophyllaceae;g__Schizophyllum	0.896211339
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Capnodiales;f__Cladosporiaceae;g__Cladosporium	0.851576203
k__Fungi;p__Entomophthoromycota;c__Basidiobolomycetes;o__Basidiobolales;f__Basidiobolaceae;g__Basidiobolus	0.808500746
k__Fungi;p__Basidiomycota;c__Agaricomycetes;o__Polyporales;f__unidentified;g__unidentified	0.793676727
k__Fungi;p__Ascomycota;c__Sordariomycetes;o__Chaetosphaeriales;f__Chaetosphaeriaceae;g__Dinemasporium	0.752888129
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Capnodiales;f__Teratosphaeriaceae;g__Teratosphaeria	0.681607476
k__Fungi;p__Basidiomycota;c__Tremellomycetes;o__Trichosporonales;f__Trichosporonaceae;g__Trichosporon	0.659035341
k__Fungi;p__Ascomycota;c__Eurotiomycetes;o__Chaetothyriales;f__unidentified;g__unidentified	0.621900877
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Dothideales;f__Aureobasidiaceae;g__Aureobasidium	0.619216144
k__Fungi;p__Basidiomycota;c__Agaricomycetes;o__Auriculariales;f__Auriculariaceae;g__Auricularia	0.548657272

k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Capnodiales;f__Teratosphaeriaceae;g__unidentified	0.483802697
k__Fungi;p__Ascomycota;c__Leotiomycetes;o__Helotiales;f__unidentified;g__unidentified	0.459510923
k__Fungi;p__Ascomycota;c__Eurotiomycetes;o__Eurotiales;f__Aspergillaceae;g__Penicillium	0.452513646
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Pleosporales;f__Periconiaceae;g__Periconia	0.441451605
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__unidentified;f__unidentified;g__unidentified	0.430814401
k__Fungi;p__Basidiomycota;c__Agaricomycetes;o__Agaricales;f__Psathyrellaceae;g__Coprinopsis	0.393541738
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Pleosporales;f__Didymellaceae;g__Didymella	0.373271684
k__Fungi;p__Basidiomycota;c__Agaricomycetes;o__Agaricales;f__Omphalotaceae;g__Gymnopus	0.36830361
k__Fungi;p__Ascomycota;c__Sordariomycetes;o__Hypocreales;f__Nectriaceae;g__Gibberella	0.3634815
k__Fungi;p__Ascomycota;c__Sordariomycetes;o__Xylariales;f__Xylariales_fam_Incertae_sedis;g__Robillarda	0.357884157
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Capnodiales;f__unidentified;g__unidentified	0.35152437
k__Fungi;p__Basidiomycota;c__Agaricomycetes;o__Agaricales;f__Entolomataceae;g__Clitopilus	0.349107382
k__Fungi;p__Basidiomycota;c__Agaricomycetes;o__Auriculariales;f__Exidiaceae;g__Exidia	0.346211157
k__Fungi;p__Basidiomycota;c__Agaricomycetes;o__Polyporales;f__Coriolaceae;g__Megasporoporia	0.321349768
k__Fungi;p__Basidiomycota;c__Agaricomycetes;o__Agaricales;f__Marasmiaceae;g__Marasmius	0.305058236
k__Fungi;p__Basidiomycota;c__Malasseziomycetes;o__Malasseziales;f__Malasseziaceae;g__Malassezia	0.300460937
k__Fungi;p__Basidiomycota;c__Agaricomycetes;o__Auriculariales;f__Exidiaceae;g__unidentified	0.290100792
k__Fungi;p__Basidiomycota;c__Ustilaginomycetes;o__Ustilaginales;f__Ustilaginaceae;g__unidentified	0.288005157
k__Fungi;p__Ascomycota;c__Sordariomycetes;o__Xylariales;f__Xylariaceae;g__Obolarina	0.285703601
k__Fungi;p__Basidiomycota;c__Agaricomycetes;o__Cantharellales;f__Ceratobasidiaceae;g__unidentified	0.268338697
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Dothideales;f__unidentified;g__unidentified	0.268155639
k__Fungi;p__Basidiomycota;c__Malasseziomycetes;o__Malasseziales;f__unidentified;g__unidentified	0.253805013
k__Fungi;p__Basidiomycota;c__Tremellomycetes;o__Tremellales;f__Trimorphomycetaceae;g__Saitozyma	0.231372863
k__Fungi;p__Ascomycota;c__Sordariomycetes;o__Hypocreales;f__unidentified;g__unidentified	0.214191504

Table 6. Abundance of fungi genera in Penang. Genera with abundance > 0.2% were shown in this table.

Taxon	average
k__Fungi;p__Ascomycota;c__Eurotiomycetes;o__Eurotiales;f__unidentified;g__unidentified	30.07270898
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Dothideales;f__Dothideales_fam_Incertae_sedis;g__Hortaea	19.41027526
k__Fungi;p__Ascomycota;c__Eurotiomycetes;o__Eurotiales;f__Aspergillaceae;g__Aspergillus	12.08773069
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Capnodiales;f__Teratosphaeriaceae;g__Eupenidiella	11.39295905
k__Fungi;p__Ascomycota;c__Arthoniomycetes;o__Arthoniales;f__Opegraphaceae;g__Combea	4.942647838
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Capnodiales;f__Neodevriesiaceae;g__Neodevriesia	3.921659015
k__Fungi;p__unidentified;c__unidentified;o__unidentified;f__unidentified;g__unidentified	3.416929154
k__Fungi;p__Ascomycota;c__Eurotiomycetes;o__Eurotiales;f__Trichocomaceae;g__Sagenomella	2.586698505
k__Fungi;p__Ascomycota;c__Sordariomycetes;o__Trichosphaeriales;f__Trichosphaeriaceae;g__Nigrospora	1.233168283
k__Fungi;p__Ascomycota;c__Sordariomycetes;o__Hypocreales;f__Clavicipitaceae;g__Metarhizium	1.115084962
k__Fungi;p__Ascomycota;c__Eurotiomycetes;o__Eurotiales;f__Aspergillaceae;g__unidentified	1.059992781
k__Fungi;p__Basidiomycota;c__Agaricomycetes;o__Polyporales;f__Schizophyllaceae;g__Schizophyllum	0.80998952
k__Fungi;p__Ascomycota;c__Leotiomycetes;o__Helotiales;f__unidentified;g__unidentified	0.478033655
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Capnodiales;f__Cladosporiaceae;g__Toxicocladosporium	0.46029299
k__Fungi;p__Ascomycota;c__Sordariomycetes;o__unidentified;f__unidentified;g__unidentified	0.41716135
k__Fungi;p__Ascomycota;c__Sordariomycetes;o__Hypocreales;f__Cordycipitaceae;g__Lecanicillium	0.404738863
k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Capnodiales;f__Teratosphaeriaceae;g__unidentified	0.37327374
k__Fungi;p__Ascomycota;c__Eurotiomycetes;o__Eurotiales;f__Aspergillaceae;g__Phialosimplex	0.290026658
k__Fungi;p__Entomophthoromycota;c__Basidiobolomycetes;o__Basidiobolales;f__Basidiobolaceae;g__Basidiobolus	0.215787851

Table 7. Relative abundance of bacterial genera in Johor Bahru, Terengganu and Penang, Malaysia.

Genus	Johor Bahru	Terengganu	Penang	average
Acinetobacter	**2.21**	**5.293**	**2.156**	3.22
Deinococcus	0.95	**6.537**	0.75	2.746
Sphingomonas	**2.76**	**3.185**	**2.116**	2.687
Ralstonia	0	0.062	**7.754**	2.605
Pseudomonas	0.801	**1.672**	**4.854**	2.442
Saccharopolyspora	**2.5**	**4.462**	0.258	2.407
Bacillus	**4.22**	**2.439**	0.331	2.33
Paracoccus	**3.132**	**3.046**	0.283	2.154
Staphylococcus	**3.371**	**2.33**	0.678	2.126
Methylobacterium	**1.326**	**3.765**	0.547	1.879
Massilia	0.764	**3.619**	0.115	1.499
Bacteroides	**1.671**	0.946	0.509	1.042
Enterobacter	**2.209**	0.85	0.066	1.042
Ochrobactrum	0.017	0.417	**2.675**	1.036
Aliterella	0.169	**2.737**	0.028	0.978
Lactobacillus	**1.957**	0.161	0.202	0.773
Corynebacterium	**1.774**	0.166	0.178	0.706
Craurococcus	0.048	0.742	**1.32**	0.703
Thermus	0	0.021	**1.758**	0.593
Cupriavidus	0.003	0.004	**1.647**	0.551
Paucibacter	0	0	**1.475**	0.492

Table 8. Relative abundance of fungal genera in Johor Bahru, Terengganu and Penang, Malaysia.

Genus	Johor Bahru	Terengganu	Penang	average
Aspergillus	**20.6528**	**12.74966**	**12.08773**	15.1634
Hortaea	**7.98**	**3.326544**	**19.41028**	10.23894
Wallemia	**6.569048**	**12.2377**	0.05993	6.288893
Eupenidiella	0	**5.301634**	**11.39296**	5.564864
Penicillium	**10.15476**	0.452514	0.034084	3.54712
Cladosporium	**7.811429**	0.851576	0.178032	2.947012
Neodevriesia	0	2.073555	**3.921659**	1.998405
Sterigmatomyces	0.004762	**5.822783**	0.008128	1.945224
Bambusaria	0	**5.227069**	0	1.742356
Sagenomella	0.540476	**1.231865**	**2.586699**	1.453013
Fusarium	0.339048	**2.893141**	0.009219	1.080469
Penidiella	**3.02381**	0.093623	0	1.039144
Candida	**1.256667**	**1.774366**	0.028391	1.019808
Schizophyllum	0.411429	0.896211	0.80999	0.705876
Asterodon	0	**1.87953**	0	0.62651
Nigrospora	0.488095	0	**1.233168**	0.573755
Khuskia	**1.647619**	0	0	0.549206
Metarhizium	0	0	**1.115085**	0.371695
Grammothele	0.598095	0.066493	0.036654	0.233748
Toxicocladosporium	0	0.176842	0.460293	0.212378

Table 9. Absolute bacterial concentration in Johor Bahru.

samples	Cq	logX	cell/g
950001259	24.07447908	6.208805449	161735535
950001261	24.99428302	5.938115651	86719278
950001262	22.12511086	6.782486504	606019368
950001263	23.09582257	6.496815018	313917132
950001265	23.3185505	6.431268246	269940623
950001266	24.08144793	6.206754582	160973572
950001268	22.97324925	6.532887214	341104315
950001269	22.01326223	6.815402521	653736180
950001270	22.61940208	6.637021166	433532007
950001271	23.30078778	6.43649565	273209408
950001276	23.49246072	6.380088077	239931946
950001281	23.54453266	6.364763786	231613455
950001282	23.90472756	6.258761754	181451998
950001283	22.37096929	6.710132638	513018042
950001285	24.34949176	6.127871761	134236853
950001286	25.36567029	5.828819807	67424822
950001287	24.09190443	6.20367733	159837004
950001288	24.05112762	6.21567757	164315136
950001289	23.73395389	6.309018868	203713058
950001290	24.94510473	5.952588366	89657859
950001264	22.65410009	6.626809863	423457533

Table 10. Absolute bacterial concentration in Terengganu.

samples	Cq	logX	cell/g
T1-1	25.839	6.126138706	133702247
T1-2	28.05074	5.442344592	13845690
T1-3	25.83267	6.128095592	19186579
T2-1	19.11841	8.203922801	1066182481
T2-2	22.22608	7.243133674	175038537
T2-3	17.14896	8.812810456	6498460092
T2-4	25.56481	6.210910373	162521332
T3-1	25.08055	6.360628172	45883672
T3-2	27.53814	5.600822402	8863595
T3-3	22.1223	7.27522146	188460987
T3-4	26.365	5.963519011	9194307
T4-1	18.22756	8.47934568	2010270134
T4-2	22.83147	7.055968865	227509146
T4-3	22.42929	7.180307726	100975606
T4-4	20.80402	7.682789241	1926855879
T5-1	25.86948	6.116716992	26166581
T5-2	25.94377	6.093747384	124093029
T5-3	27.32403	5.667017313	23226690
T6-1	20.3017	7.83808843	459195022
T6-2	19.08758	8.213455188	1634764462
T6-3	15.79801	9.230480456	5667078147
T6-4	20.63858	7.733938385	1083847999
T7-1	16.76146	8.932615212	3425115363
T7-2	16.73011	8.942304905	9216824105
T7-3	21.04883	7.607100134	2023345921
T7-4	20.05154	7.915431074	411529600
T8-1	25.88282	6.112593669	5183865
T8-2	26.94743	5.783448869	6073638
T8-3	21.7547	7.388868549	244832208
T8-4	27.22792	5.696733621	3316213