ARTICLE   OPEN

Check for updates

Extensive urban air pollution footprint evidenced by submicron organic aerosols molecular composition

Christian Mark Salvador[1,2,7], Charles C.-K. Chou [3 ✉], T.-T. Ho[3], I-Ting Ku[4], C.-Y. Tsai[3], T.-M. Tsao[5], M.-J. Tsai[5] and T.-C. Su[6]

Transport and transformation of urban air pollutants are among the major factors driving the changes in the atmospheric composition in the downwind rural/remote areas of a megacity. Here, we assess the impacts of urban air pollution in a subtropical forest through characterization of the organic markers in submicron aerosol particles. The aerosol samples were collected and analyzed using TD-PTR-ToF-MS, where 163 ions were detected. The concentration of these extracted ions accounts for 83% of the mass of submicron organic aerosols, which are accordingly characterized by a median formula of $C_7H_{10}O_2$. Molecular speciation indicates that urban and biomass burning pollution contributed substantially to the budget of organic aerosols, which were enhanced particularly by the liquid water content and acidity of the aerosols. Our results evidence that the footprint of urban air pollution was extended to its downwind forested areas and caused changes in the concentration and composition of submicron aerosols.

INTRODUCTION

The impact of anthropogenic emission of trace gases from urban areas to biogenic environment is still uncertain, particularly in terms of how urban air pollutants can modify the levels of reactive compounds and natural atmospheric chemical pathways[1]. The complex interaction between urban and biogenic species can influence atmospheric oxidation capacity[2] and modify the production and, thereby, ambient concentration of aerosols[3–6]. Understanding the changes in the natural atmospheric conditions due to the transport of pollutants from urban areas, particularly the megacities, is essential to accurately describe the influences of urbanization on regional atmospheric chemistry.

A typical approach of evaluating the perturbation of anthropogenic pollutants is to identify and account for the tracers associated with human endeavors. Such an approach is based on a hypothesis that the ambient concentration of the specific tracers increases with the strength of corresponding activities. Estimating the mass contribution of these tracers in a clean and biogenic environment where only natural processes partake may provide an ideal case to test the hypothesis. Several traditional tracers in gas and particle-phase have been effectively characterized and attributed to specific sources[7–10]. For instance, carbon monoxide (CO) is produced from the incomplete combustion of fossil/bio fuels, which makes it a typical tracer of urban plumes. However, the assessment of CO emission is associated with a large uncertainty due to variable sources of CO, which also include the natural biomass burning (BB) and oxidation of biogenic volatile organic compounds (BVOCs)[11]. Also, potassium ion (K^+) is customarily tagged as a marker of biomass burning (BB) but it also has an apparent contribution from sea spray in coastal areas and resuspension of soil from fertilized cropland in agricultural areas. Because of the relatively well-characterized emissions of volatile organic compounds (VOCs) from various urban sources, the VOCs

and their respective oxidation products in aerosols have been suggested as source-specific markers of urban air pollution in the atmosphere[12,13].

In this study, the perturbation of anthropogenic processes in a subtropical forest during the fall season of 2015 is evaluated through a comprehensive characterization of submicron organic aerosols, which includes analysis of source, size distribution, liquid water content (LWC), pH, and saturation concentration. Although several studies have previously shown the influence of urban emissions on the atmospheric conditions of forest environments[1,3,14], only a handful of observations were implemented in biogenic sites in the East Asia Region that is heavily impacted by both local and regional pollution events. The forested area in this study is located downwind of the major urban area in central Taiwan, where a substantial amount of air pollutants are emitted from sources like coal-fired power plants, vehicles, and industrial factories. The urban air pollutants are transported to the forested site with the aid of mountain-valley circulation[15]. Moreover, Taiwan is located in the downwind area of East Asian winter monsoons and receives transboundary atmospheric pollutants originating in the Chinese megacities during the cold seasons[16,17]. With the confounding sources of anthropogenic emissions, investigation at this forested site can provide direct evidences of the impact of human-related activities on the formation pathway of organic aerosols. The influence of anthropogenic activities in this study is assessed through the behavior of the submicron particle-phase organic tracers measured using Thermal-Desorption Proton-Transfer-Reaction Time-of-Flight Mass Spectrometry (TD-PTR-ToF-MS)[12]. This study involves the thorough identification of molecular ions, particularly the prominent organic markers, to explain the sources of submicron secondary organic aerosols (SOA). Even though both daytime and nighttime samples were collected, only daytime concentration of molecular tracers is analyzed in the study to focus on the anthropogenic pollutants'

[1]Department of Chemistry and Molecular Biology, Atmospheric Science, University of Gothenburg, SE-41296 Gothenburg, Sweden. [2]Balik Scientist Program, Department of Science and Technology - Philippine Council for Industry, Energy and Emerging Technology Research and Development, Bicutan, Taguig 1630, Philippines. [3]Research Centre for Environmental Changes, Academia Sinica, Taipei 11529, Taiwan. [4]Department of Atmospheric Science, Colorado State University, Fort Collins, CO, USA. [5]School of Forestry and Resource Conservation, National Taiwan University, Taipei, Taiwan. [6]Department of Internal Medicine and Cardiovascular Center, National Taiwan University Hospital, Taipei, Taiwan. [7]Present address: Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA. ✉email: ckchou@gate.sinica.edu.tw
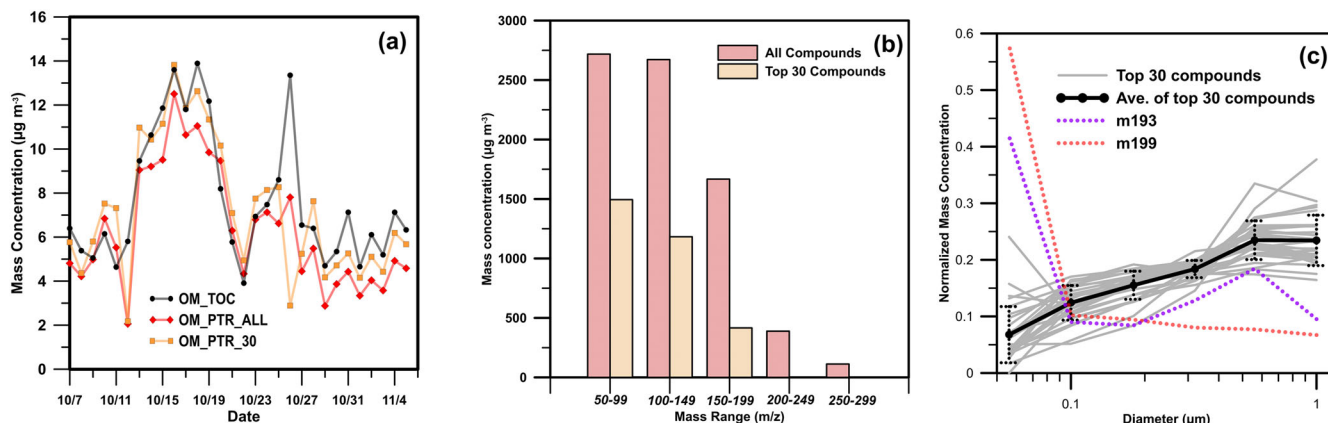
Fig. 1 Characteristics of the organic ions in the experimental forest. a Time series profile of the daytime mass concentration of submicron OM, the sum of all ions from the thermal desorption analysis (TD-PTR-ToF-MS, all), and the sum of mass concentration of the 30 major ions indicated in Table 1. **b** Breakdown of the contribution of different mass ranges to the overall mass concentration of all the ions from MS and top 30 ions. **c** Size distribution profile (<1.0 μm) of the normalized concentration to the sum of mass concentration of major organic ions. The strong bold line indicates the average of all the tracers. Included in the graph are the size distribution profiles of the ions at 193 and 199 (dotted lines). The error bars represent the range of ±1 standard deviation.

photochemical formation route and daytime transport of urban pollutants from nearby urban areas. The daytime concentrations of major ions typically have three times the levels measured during nighttime (see Supplementary Fig. 1 for comparison of daytime and nighttime).

RESULTS

Molecular characterization of the submicron aerosols

The thermal desorption of the aerosol samples resulted in the extraction of 163 ions in the PTR-ToF-MS. The average mass spectra is illustrated in the Supplementary Fig. 2. The molecular formulas of the ions are identified and listed in Supplementary Table 1. These ions have a molecular composition containing 2–16 carbons, 2–33 hydrogens, 0–7 oxygens, 0–2 nitrogens, with a median formula of $C_7H_{10}O_2$. The median oxygen to carbon ratio (O:C) for these particulate components is 0.25 (Min:0.067; Max: 1.5), agreeing with the results of previous aerosol chemical characterization utilizing PTR-ToF-MS[13]. Furthermore, the average (median) molecular weight (MW) per carbon weight (CW) of all the ions (MW/CW) is 1.65 ± 0.46 (1.52), similar to the factor widely used to estimate particulate organic matter mass for urban aerosols[18]. This mass factor is thus justified and applied in this study. The O:C ratio is significantly higher than the values expected for primary vehicle exhausts (0.022–0.15)[19], agreeing with the general concept that primary air pollutants will be oxidized along the transport process and become more abundant in oxygen in the downwind areas. However, the oxygen content observed in this study is lower than the results derived from measurements of the Aerodyne Aerosol Mass Spectrometry[20]. The low oxygen content is due partly to the limited capacity of the PTR-ToF-MS to measure highly oxygenated compounds (O > 4) compared to other ionization techniques (e.g., $CH_3C(O)O^-$ and $I(H_2O)O^-)$[21,22]. Besides the ionic fragmentation, thermal desorption and dehydration of the parent compounds will reduce the oxygen content of these detected ions[23,24]. Assuming dehydration impacted all detected ions and several carboxylic acids formed acylium ions, we estimate that the median O:C ratios of the original organic aerosol constituents could be as high as 0.36 (Min: 0.067; Max: 2). The correction was based on the loss of water molecule upon protonation of the parent molecule. One oxygen and one hydrogen were added to the dehydrated ions detected by the PTR-MS to retrieve the chemical formula of the parent molecule.

Figure 1a shows the time series of the sum of the mass concentration of the ions measured by the TD-PTR-ToF-MS (OM_{PTR_All}). Also shown in the figure is the time series of the mass concentration of total OM measured using a conventional thermal/optical carbon analyzer (OM_{TOC}). The OM_{PTR_All} correlates well with the total OM_{TOC} ($r^2 = 0.768$), indicating that molecular level identification of submicron organic aerosol can appropriately explain the variability of the OM. The absolute total mass concentration of the OM from the identified ions underestimates the OM_{TOC} by 17% (1.34 μg m^{-3}). This deviation is due partly to thermal degradation and fragmentation leading to loss of carboxyl, carbonyl, and water groups which cannot be accounted by the PTR-ToF-MS. Limited sensitivity of the PTR-ToF-MS to highly oxygenated compounds could also have contributed to the lower mass concentration of OM provided by the TD-PTR-ToF-MS.

Most of the ions detected in this study have a mass to charge ratio (m/z) of less than 300, comparing to other chemical ionization techniques using iodide(I^-) or nitrate (NO_3^-) as reagent ion that can effectively detect ions beyond 500 m/z[24,25]. Nevertheless, the limited difference between OM_{PTR_All} and OM_{TOC} suggests that the submicron organic matter consisted mostly of compounds with m/z < 300. Figure 1b shows the distribution of the concentration of the observed ions in different mass ranges. Low molecular weight (LMW) ions at the mass range between 50 and 150 m/z contributed 70% of the mass concentration. This is coherent with a previous study where the main fraction of ions (75%) measured by a TD-PTR-ToF-MS appeared below 150 m/z[13]. In such study, TD-PTR-ToF-MS was utilized to investigate SOA formation from the ozonolysis of several biogenic VOCs associated with real plant emissions under a controlled reaction chamber. The similarity of effective molecular range in both studies even with different analytes and atmospheric conditions highlights the effective MW range of TD-PTR-ToF-MS in the chemical characterization of organic aerosols. To provide a more in-depth analysis of the sources of the organic compounds in the submicron aerosols, the top 30 ions that dominated the mass spectra for all the filter samples were characterized. Table 1 lists the major ions, including their suggested molecular formulas and average mass. All the ions in consideration were observed in all the 30 aerosol filters and the sum of the mass concentration of these ions (OM_{PTR_30}) comprised 42% of the total OM $_{PTR\_All}$. Figure 1a illustrates that these ions can also effectively estimate the variability of the submicron OM, indicating that these ions are suitable representatives in assessing the sources of submicron OM and its chemical characteristics. The distribution of the mass ranges and their concentration in Fig. 1b

Table 1. Major ions retrieved from the submicron particles collected during daytime in this study[a].

Experimental m/z	Possible MF	Theor. m/z	mDa diff	Ave. MC	Mass Burden (%)[b]	Tmax region[c]	log (C*) µg m^{-3}
57.033	$C_3H_5O^+$	57.0340	0.66	97.66	1.30	4	8.90
59.049	$C_3H_7O^+$	59.049	0.67	108.60	1.45	2	8.90
60.045	$C_2H_6NO^+$	60.044	0.119	83.70	1.12	2	NA
61.029	$C_2H_5O_2^+$	61.028	0.234	352.15	4.70	2	8.73
71.049	$C_4H_7O^+$	71.049	0.63	65.86	0.88	2	8.34
73.028	$C_3H_5O_2^+$	73.028	0.934	100.64	1.34	3	7.89
83.05	$C_5H_7O^+$	83.049	0.18	92.27	1.23	2	7.80
85.03	$C_4H_5O_2^+$	85.028	−0.896	104.70	1.40	3	7.18
87.046	$C_4H_7O_2^+$	87.044	−1.786	85.76	1.14	2	7.18
97.03	$C_5H_5O_2^+$	97.028	−1.276	227.76	3.04	3	6.53
99.009	$C_4H_3O_3^+$	99.008	−0.601	175.26	2.34	4	6.29
101.022	$C_4H_5O_3^+$	101.023	1.469	207.98	2.78	2	6.29
111.044	$C_6H_7O_2^+$	111.044	0.604	86.04	1.15	2	5.93
113.023	$C_5H_5O_3^+$	113.023	0.969	102.03	1.36	2	5.53
115.041	$C_5H_7O_3^+$	115.039	−1.581	63.42	0.85	2	5.53
125.058	$C_7H_9O_2^+$	125.060	1.855	73.28	0.98	3	5.35
127.038	$C_6H_7O_3^+$	127.039	1.909	92.48	1.23	2	4.83
137.105	$C_9H_{13}O$	137.096	−0.826	24.56	0.33	2	5.78
139.075	$C_8H_{11}O_2^+$	139.075	0.805	682.48	0.91	2	4.80
141.095	$C_8H_{13}O_2^+$	141.091	−2.945	44.01	0.59	2	4.80
143.105	$C_8H_{15}O_2^+$	143.107	2.005	35.40	0.47	2	4.80
147.078	$C_{10}H_{11}O^+$	147.081	3.43	43.92	0.59	2	5.29
149.024	$C_8H_5O_3^+$	149.023	−0.031	340.74	4.55	2	3.55
151.111	$C_{10}H_{15}O^+$	151.112	−1.51	40.05	0.53	2	5.29
153.093	$C_9H_{13}O_2^+$	153.091	−0.945	63.23	0.84	2	4.25
155.108	$C_9H_{15}O_2^+$	155.107	−0.495	62.46	0.83	2	4.25
157.089	$C_8H_{13}O_3^+$	157.086	−2.731	48.22	0.64	2	3.55
163.042	$C_9H_7O_3^+$	163.039	−2.881	73.09	0.98	1	2.95
167.108	$C_{10}H_{15}O_2^+$	167.107	0.80	47.829	0.64	2	3.73
169.087	$C_9H_{13}O_3^+$	169.086	−0.231	81.274	1.08	2	2.95

MF molecular formula; MC mass concentration; C* saturation concentration.
[a]Comparison of daytime and nighttime concentration of the major compounds is given in Supplementary Fig. 1.
[b]Mass contribution to total submicron organic matter measured using OC/EC analyzer.
[c]Temperature max (T_{max}) regions: (1) RT-100 °C; (2) 100–200 °C; (3) 200–350 °C; (4) 350–550 °C.

is consistent with the overall profile for all the extracted ions. In both the entire list of ions and the top 30 ions, the majority of the ions have a low mass to charge ratio (<150 m/z) with a consistent decreasing trend of abundance as the m/z increases (see Supplementary Fig. 2 for average mass spectra). This further reinforces the suitability of these 30 ions as representative of the whole organic ions measured by the TD-PTR-ToF-MS.

The size distribution of major organic markers was assessed to determine their participation constituent in gas to particle partitioning and particle formation processes. Figure 1c shows the normalized concentration of the top 30 organic ions across size distribution below 1.0 µm. A common pattern was observed among major organic ions, in which most of the ions were abundant in the larger particles with a steady reduction in their mass concentration as the particle diameter decreases. Half of the concentration of the dominant organic ions persisted in particles with an aerodynamic diameter between 1.0 µm and 320 nm. The observed trend is consistent with the portioning of the gas-phase compound to the condensed phase, where condensation is more evident on particles predominant in the size distribution of surface area. Interestingly, a few ions not included in the top 30 ions show a reverse trend, with Aitken particles ($D_p < 0.1$ µm) containing at least 50% of the mass concentration. These ions contribute to the enhanced OA mass for particles with a diameter less than 56 nm. For instance, the ions at 199.170 ($C_{12}H_{23}O_2^+$) and 193.154 ($C_{13}H_{21}O^+$) persist mostly in small particles ($D_p < 0.056$ µm) even though these particles typically contribute insignificantly to the overall PM mass concentration. The high concentration of these long-chain carbon compounds in small particles indicates that these compounds could participate in the early nucleation and particle growth processes.

Volatility of the submicron organic aerosols

Based on the proposed molecular formulas, the saturation concentrations (C^*) of the top 30 ions were calculated according to the number of carbon, oxygen, and nitrogen atoms[17,26]. The saturation concentration of a compound dictates the volatility of a compound and its contribution to both gas and particle phases. C^* (µg m^{-3}) values, expressed as log(C^*), were calculated using the Eq. (1):

$$log(C^*) = (n_{O*} - n_C)b_C - (n_O - 3n_N)b_O - 2\frac{(n_O - 3n_N)n_C}{(n_C + n_O - 3n_N)}b_{CO} - n_N b_N \quad (1)$$
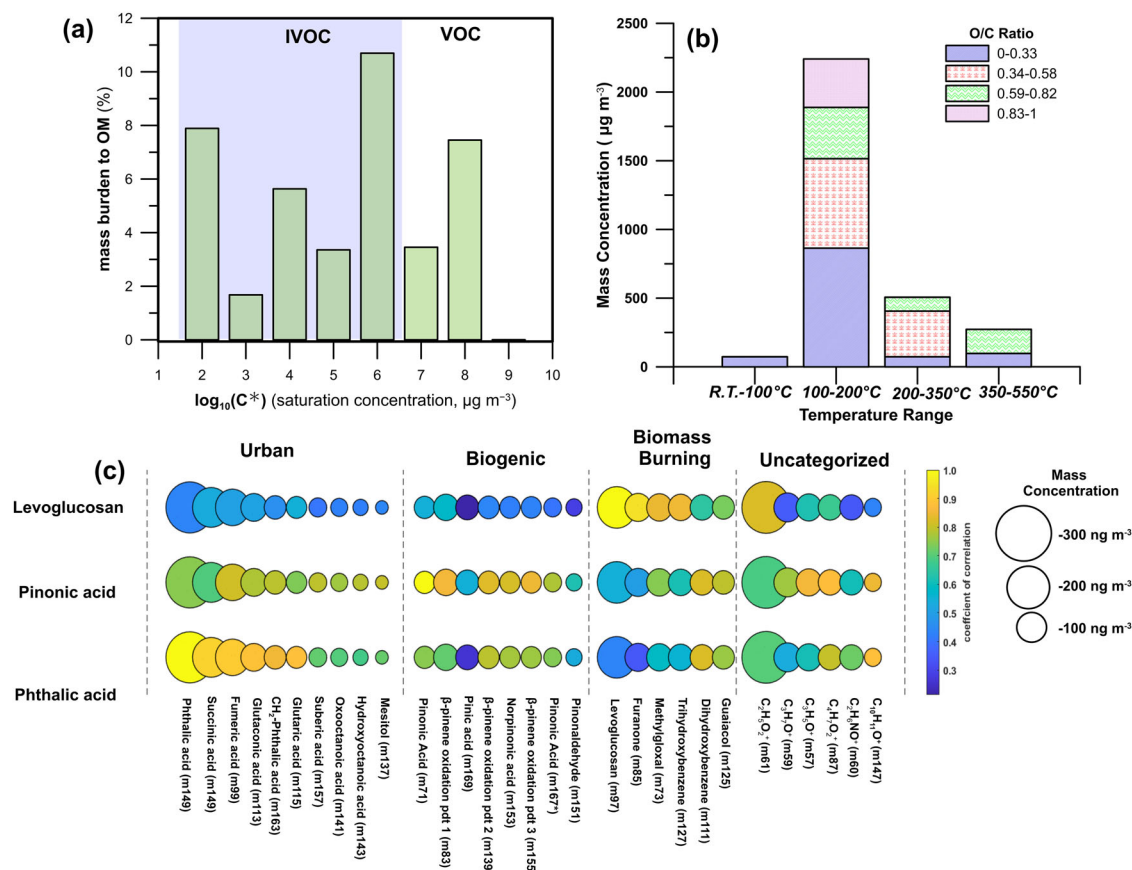
Fig. 2   Volatility, thermal profile, and sources of the major ions. a Distribution profile of the major organic ions binned into a volatility basis set (VBS). Note that the saturation concentration (C^*) of each ion was corrected, assuming that the major ions detected were produced from dehydration (-H_2O) after protonation of parent compounds. The ions with a $\log(C^*) > 7\,\mu g\,m^{-3}$ were deemed a thermal decomposition fragment of large molecules. The uncorrected saturation concentrations of the major ions are presented in Supplementary Fig. 3. **b** Histogram profile of the top 30 ions based on temperature range with the inset of oxygen to carbon ratio (O/C). **c** Clustering of the organic tracers based on their sources. The size of the circles is proportional to the total measured mass concentrations (average over the 1-month measurement). The color of circles indicates the coefficient of correlation of each compound to phthalic acid (urban), pinonic acid (biogenic), and levoglucosan (BB) primary tracers.

where $n_{0*} = 25$, $b_C = 0.475$, $b_O = 0.2$, $b_{CO} = 0.9$, and $b_N = 2.5$. These five constants are based on prior volatility predictions and represented elemental interaction (e.g., carbon–carbon) and carbon–oxygen nonideality. More information is referred to Donahue et al.[26] and Mohr et al.[17]. The terms n_C, n_O, and n_N are the number of carbon, oxygen, and nitrogen, respectively. The calculated saturated concentration of each ion is given in Table 1. Using the calculated C^*, the mass contribution of each compound to OM was binned into a volatility-basis set (VBS). Initially, the theoretical calculation of saturation concentration placed the 30 ions in the intermediate (IVOC) and volatile organic compounds (VOC) ranges (as shown in Supplementary Fig. 3). Almost 25% of OM_{PTR_30} belonged to a highly volatile compound range ($\log(C^*) > 7\,\mu g\,m^{-3}$), which is counterintuitive considering that these compounds were measured in the particle phase. However, T_{max} regions of these compounds occurred mostly above 100 °C (Fig. 2b), which indicates that these small compounds (MW < 100) are fragments of larger compounds produced through collision-induced dissociation (CID) and/or thermal decomposition. The estimated real $\log(C^*)$ values of the compounds should be at least 1.0 $\mu g\,m^{-3}$ lower, considering that majority of the compounds in this study were fragments of dehydration after protonation of OVOCs. For instance, the fragment of phthalic acid at 149.025 ($C_8H_5O_3^+$) has a $\log(C^*)$ of 3.55 $\mu g\,m^{-3}$, which decreases to 2.48 $\mu g\,m^{-3}$ upon the addition of missing OH. This shifts some of the compounds into the IVOC region (as shown in Fig. 2a),

consistent with the saturation concentration range of typical particle-phase tracers[26]. Indeed there are still some low MW ions with a $\log(C^*) > 7\,\mu g\,m^{-3}$, which are deemed a thermal decomposition fragment of large molecules because all the ions are associated with a $T_{max} > 100$ °C. Parameterization of the volatility of the compounds desorbed from the particle phase and measured by the PTR-ToF-MS should always include temperature profile to prevent overestimation of saturation concentration.

Source attribution of the submicron organic matter

Some of the ions in Table 1 also appeared in the measurements of our previous urban study, where most of the species were attributed to human activities[12]. For instance, the ion at m/z 149.023 had the highest concentration for both biogenic and urban sites. This ion corresponds to the acylium fragment (RCO^+) of phthalic acid, which is a product of secondary oxidation of naphthalene. The attribution of the ion at m/z 149.023 to phthalic acid was confirmed using authentic standards in a prior study[12]. The formation of acylium ions inside the drift tube of PTR-MS is a product of the protonation of carboxylic acids ($RCOOH$)[27,28].

$$H_3O^+ + RCOOH \rightarrow RCO^+ + 2H_2O \tag{2}$$
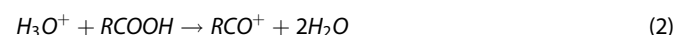
Moreover, a recent study also proposed that phthalate esters and phthalic acid undergo an α-cleavage of the parent ion, which results in the loss of OH group, and an intermolecular McLafferty

rearrangement occurs to form the protonated anhydride[29]. These reaction mechanisms are valuable in the identification of the molecular identity of the major ions as most of the organic compounds in the aerosol phase are organic acids[19,20]. The ions at 71.049 and 97.026 m/z were previously reported as major fragments of pinonic acid and levoglucosan, which are biogenic and biomass burning markers[21,22]. Moreover, the ion with m/z of 61.029 was earlier attributed to protonated acetic acid. Such ion also had a substantial mass burden (4.70%) in the PM in Xitou forest, similar to the measurements in a remote mountain observatory in the Austrian Alps[23] and in a subtropical megacity in Taiwan[12]. However, given the desorption temperature (T_{max}) of the largest signal of such ion (100–200 °C), it was inferred that the ion at m/z 61.029 was a product of fragmentation of larger molecules in submicron organic aerosols. Other ions in the top 30 list were identified based on their association with phthalic acid (urban), pinonic acid (biogenic), and levoglucosan (biomass burning). Figure 2c shows the clustering of the 30 major ions based on their sources, together with their average mass concentration in the submicron aerosols. In the succeeding sections, the possible sources and production pathways (i.e. urban and biomass burning) of some of the compounds will be discussed, together with comparisons with prior studies. Note that this study is mostly addressing the impacts of anthropogenic activities (i.e. urban pollution and biomass burning), whereas the biogenic tracers were discussed in a previous companion paper[5]. Furthermore, source apportionment analysis of the major ions using PMF (see Supplementary Discussion) was utilized to compare the empirically derived cluster based on major tracers (e.g. phthalic and pinonic acid) with the emission sources/factors identified by the PMF model.

Urban tracers
In Table 1, the ion with a mass to charge ratio of 149.024 tops the list of the major ions in terms of mass concentration. This ion was attributed to the acylium ion of phthalic acid. This diacid is a typical urban tracer, primarily generated from the oxidation of mono/polyaromatic hydrocarbons[24]. The mass concentration of phthalic acid reached as high as 600 ng m^{-3} during the daytime, with field campaign average concentration of 340 ng m^{-3} (4.5% of submicron OM). Interestingly, the mass burden of phthalic acid was comparable with the observation from an urban city that was highly influenced by anthropogenic activities[12], thus, hinting the dominant participation of urban-related activities in the experimental forest. The mass concentration of phthalic acid in this study was elevated as comparing to measurements in a coastal urban city in Greece (74 ng m^{-3})[25], in 4 sites in Southern China (36 ng m^{-3})[24], in tropical India (6.0 ng m^{-3})[30] and Beijing (78 ng m^{-3})[31]. Most of these sites had severe polluted cases (e.g., Beijing), thus indicating the adverse influence of anthropogenic activities at our study site.

The source of the precursor/s of the phthalic acid may shed some insights on the elevated concentration of this urban tracer. Most studies indicated that among the PAHs, the oxidation of naphthalene significantly produces phthalic acid[24,32]. Here, a moderate correlation ($r = 0.563$, $p < 0.01$) was found between the concentrations of naphthalene and phthalic acid, indicating a possible precursor-product relationship. Naphthalene is one of the most volatile 2-ring PAHs and is considered as one of the urban air pollutants found at a significant level from automobile exhausts. Its source also includes gasoline and oil combustion, tobacco smoking, and fumigants[33]. The reaction of OH radical with naphthalene results in the cleavage of one of the aromatic rings, which leads to the formation of dicarbonyl products such as phthalic acid. The average daytime naphthalene concentration for this study was 0.1 ppb, which was on the high end of estimates for naphthalene in urban and suburban sites[33]. The oddly high level

of naphthalene in the forest was suspected to be predominantly transported from an urban city (Taichung). Wind direction analysis showed that wind during daytime mostly originated from north north west (NNW, see Supplementary Fig. 4), consistent with the location of Taichung city, which has substantial vehicular and industrial activities. The human-related tracers and their precursors were suspected to be transported from this area, which altered the atmospheric processes such as aerosol formation in the forest. Moreover, the contributions of phthalate esters (PAEs), which are ester congeners of phthalic acid, were not discounted in this study. PAEs are typically used in industrial and medical applications, particularly as plasticizers. Prior studies have shown substantial levels of PAEs in atmospheric aerosols, particularly at industrial sites[34–36]. The impact of PAEs on m/z 149.024 will be further accounted in succeeding studies.

Among the urban tracers in Fig. 2c, the ions at m/z 101.022 and 99.009 were also significantly contributing to the total mass concentration of OM in submicron aerosols, accounting as much as 2.7 and 2.3%. This ion at 101 was assigned as the acylium fragment of succinic acid, a prevalent linear C_4 dicarboxylic acid in ambient particles. Succinic acid, together with oxalic (C_2) and malonic acid (C_3), is the dominant species of the low molecular weight dicarboxylic acids[37]. Dicarboxylic acids like succinic acid induce new particle formation by interacting with nucleating precursors (e.g., sulfuric acid) via hydrogen bonding of the two carboxylic acid moieties[38,39]. Like malonic and oxalic acid, succinic acid is a product of the oxidation of non-methane volatile organic compounds such as cyclic alkenes[40]. Oxalic acid was not accounted in this study due to weak signals of the expected ions attributed to such diacid. A prior study indicated that the major PTR-MS signal of oxalic acid occured at m/z 46[41]. Ions below m/z 50 were not included in this study to limit the possible contribution and interference of small gaseous compounds. Succinic acid observed in this study was thrice the level observed in different urban and rural sites in China[39,42]. This is likely due to the high oxidation capacity (Ozone$_{av}$ = 53 ppb) and water content (RH > 90%) in the forest site that induced the gas phase production of diacids and their better partitioning into the particle phase. With regards to ion at 99, this ion is exactly two hydrogens less (2.015 Da) of the fragment of succinic acid at m/z 101. This difference may indicate that such ion is maleic/fumaric acid, which is the unsaturated C_4 analog of succinic acid. Maleic/fumaric acid is also an urban tracer, formed through the oxidation of aromatic VOCs such as benzene and toluene[43].

Using the same approach to identify ions at m/z 99.009 and 101.022, a homologous series of seven low molecular weight saturated (C_5–C_{11}) and two unsaturated fatty acids (C_4–C_5) (see Supplementary Table 2 for structures) were extracted from the mass spectra. Some of the ions identified were already included in the major ions listed in Table 1, such as m/z 113.023 and 115.041 which corresponds to the acylium ions of protonated glutaric ($C_5H_8O_4$) and glutaconic ($C_5H_6O_4$) acid. LMW diacids with less than 9 carbons were anthropogenic tracers generated from the oxidation of cyclic olefins. Longer diacids such as azelaic acid (C_9) are photochemical oxidation products of unsaturated biogenic VOCs with a double bond at C_9 position[37]. Figure 3 shows the time series profiles of 7 urban-related diacids, which were coherent with the variability showed by phthalic acid. This indicates that the source of diacids was also human-related activities, like phthalic acid. Correlation analysis of the diacids with elemental carbon (EC) and phthalic acid also revealed that the 7 dicarboxylic acids were primarily formed through the secondary production pathway.

Overall, the urban tracers identified in the biogenic site account for 16% of submicron OM, with as much as 1.2 µg m^{-3} mass concentration contribution to submicron aerosols. This shows that the formation of submicron organic aerosols was highly influenced by human-related activities based on the prominent
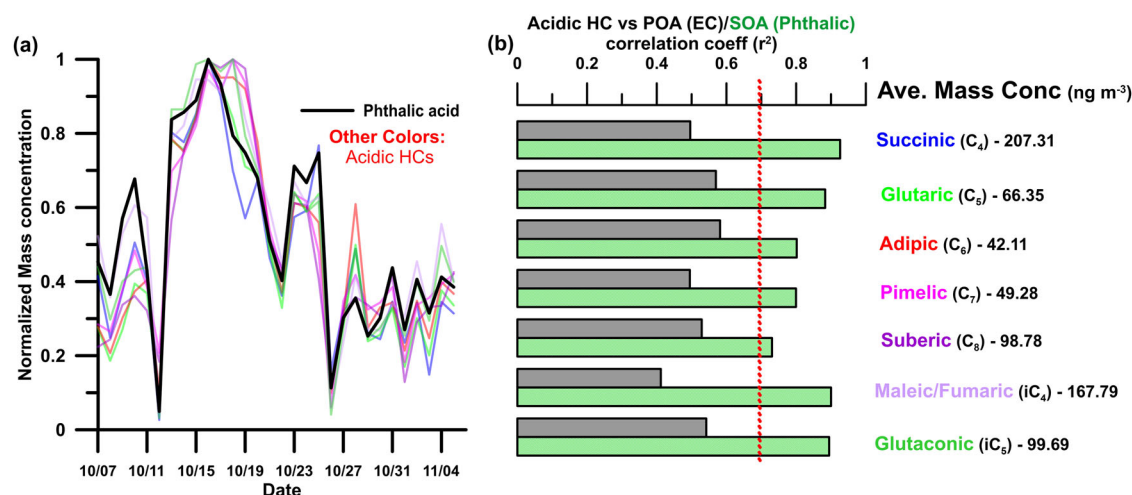
Fig. 3 Profile of the urban tracers in submicron aerosols. a Time series of the normalized mass concentration of phthalic acid and other dicarboxylic acids. The black line is for the phthalic acid while the other colors are for the other urban tracers. **b** Production pathway analysis of urban dicarboxylic acid based on their relationship with elemental carbon and phthalic acid, which are indicators for primary and secondary origins. The numbers in the legend are the mean mass concentration of the urban tracers.

presence of the urban tracers, particularly the acidic hydrocarbons. Diagnostic mass ratios of dicarboxylic acids also revealed the mixed sources in the forest. The ratio of phthalic acid (anthropogenic) to azelaic acid (biogenic) for this case was 5.0 ± 1.8, comparable with the observations in urban continental measurements[44], thus demonstrating the strong influence of anthropogenic activities in the forest.

Biomass burning related compounds

Emission from anthropogenic biomass burning (e.g., burning of agricultural waste and domestic fuelwood) reduces air quality and poses a threat to human health conditions. The ion at m/z 97.030 was attributed to levoglucosan, based on a previous standard analysis[12]. Levoglucosan is an ideal tracer of BB since it is only produced at combustion processes at more than 300 °C while traditional combustion plume tracers such as potassium ion (K^+) and black carbon (BC) have sources other than BB[45]. A weak relationship between K^+ and the fragment of levoglucosan suggests that K^+ had a contribution from other sources, e.g. seawater, soil resuspension, and fertilizers[46]. In prior work, arabitol, a C_5 sugar alcohol, also had the same dominant fragment at 97.030 m/z ($C_5H_5O_2^+$)[12], which can overestimate the mass concentration of levoglucosan. However, arabitol is a bioaerosol primarily emitted by fungal spores[7], thus the contribution of arabitol to the OM should be negligible for this case since the aerosols collected here have a 1.0 μm cut-off and bioaerosol markers like arabitol exist mostly in coarse particles[47,48]. Strong persistence of the ion at m/z 97 in submicron aerosol, as shown in its size distribution profile (see Fig. 1c), clearly identifies this ion as levoglucosan instead of arabitol. Anion exchange analysis of monosaccharides using ion chromatography also indicated that levoglucosan accounted for more than 60% of the total sugars in submicron aerosols while arabitol contributed only less than 3% (see Supplementary Fig. 5).

Levoglucosan ranked third among the major compounds and contributed 230 ng m^{-3} to the total submicron organic aerosol mass (4% of OM). Typical concentrations of levoglucosan in other forest sites were below 100 ng m^{-3}, particularly in the absence of forest fires[49–51]. The concentration of levoglucosan, which reached as much as 500 ng m^{-3}, was intermediate between typical measurements obtained from episodic and non-episodic events of BB. In a forest fire in Indochina, levoglucosan reached 1.8 μg m^{-3}, with a mean concentration of 1.16 μg m^{-3}[46]. Similar

mass burdens were calculated both in Indochina and in our forest site, indicating that a strong source of combustion processes influenced the air quality in this study. Local source rather than long-range transport of plume was accounted for the elevated mixing ratio of levoglucosan. Episodic burning of agricultural residues during harvest season occurs in Central and Southern Taiwan which spans from late autumn to early winter[52,53]. Likewise, the HYSPLIT trajectory model indicated that the air mass arriving our study site did not pass through western Indochina, where BB events are prominent.

Figure 2c shows that the ions at m/z of 111, 125, and 127 were strongly associated with levoglucosan, which suggests that their corresponding parent compounds were also produced during pyrolysis. These ions were assigned as dihydroxybenzene/catechol (DHB) ($C_6H_6O_2$), guaiacol ($C_7H_8O_2$), and trihydroxy benzene (THB) ($C_6H_6O_3$) which are substituted phenolic compounds formed during the combustion of lignin[54]. The three BB-aromatic compounds (BBAC) contributed 3% of the OM (BBAC$_{ave}$ = 250 ng m^{-3}). Several studies showed the presence of these BB tracers in the particle phase[55,56] with a significant contribution of as much as 45% to the total aerosol mass during wood combustion. Other than direct emission, dihydroxybenzene can be produced through OH radical initiated oxidation of benzene and phenol in the gas phase[57]. Also, oxidation of other BB intermediates such as guaiacol and syringol under low NO$_x$ conditions generated a substantial amount of catechol[58]. Elevated mass concentration of catechol during daytime (85 ng m^{-3}) compared to nighttime (25 ng m^{-3}) may indicate that the photochemical production significantly contributed to the observed DHB concentration in the biogenic site. The same pattern was observed for THB and guaiacol. The mountain-valley breeze circulation also transported the biomass-burning plumes from the nearby metropolis area.

Like the urban tracers, the BB markers in the forest enhanced the mass concentration of the submicron particles. Levoglucosan, DHB, and other combustion tracers all together constituted 9% of submicron OM. The elevated mass concentration of prominent biomass burning tracers in a biogenic site indicates the strong influence of human-related activities in the forest. Beyond the enhanced concentration of atmospheric pollutants, combustion activities can also impact the hydrological cycle[59], which in turn could alter the biological growth due to a reduction in overall moisture level.
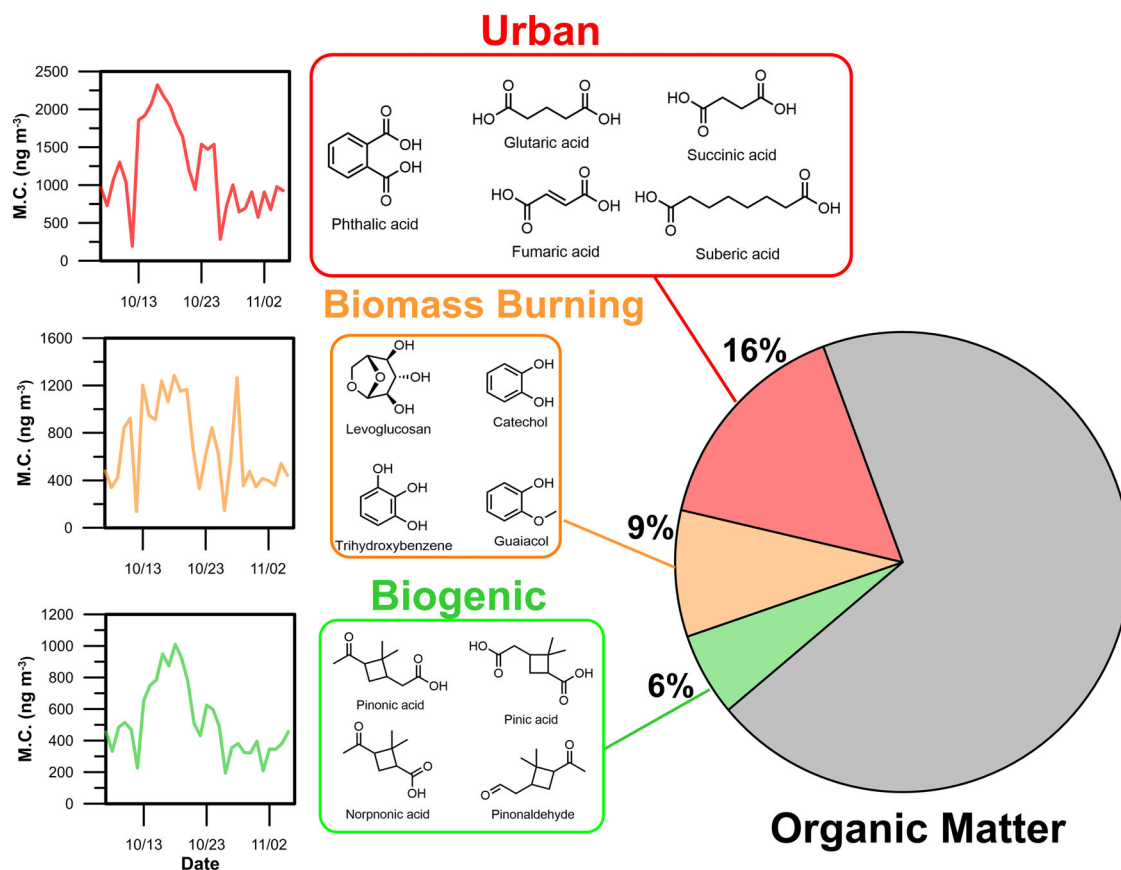
Fig. 4 Time series profile and contribution to OM of the major organic markers, clustered according to their primary source.

Discussion and atmospheric implications

Significant quantities of organic tracers were measured using a TD-PTR-ToF-MS in an elevated subtropical forest where the influence of anthropogenic activities was substantial. Molecular characterization of the submicron aerosols showed that the organic compounds can suitably explain the variability of OM. Furthermore, the top 30 ions extracted from MS, which accounted for ~50% of submicron OM, mirrored the variability of all the ions from TD-PTR-ToF-MS and total OM, allowing these ions to be appropriate representative of the organic aerosols. These particle-phase bound compounds showed relatively low $O{:}C$ ratio, with evident partitioning to larger particles (~1.0 µm). Volatility analysis based on molecular formula alone placed the 30 compounds mostly in IVOC and VOC concentration saturation range, but max desorption temperature indicated otherwise that the VOC ions were fragments of larger organics with low volatility. From the thirty major ions, eight ions were identified and associated with biogenic processes, ten were urban-related and six were accounted as BB tracers (see Supplementary Table 2 for some ions not discussed in the main text). Figure 4 shows the contribution of respective pollution categories, which were defined by the molecular tracers identified in this study, where a significant portion of the total submicron organic matter (69%) was not yet linked explicitly to a specific source. This signifies the need to expand the identification process of the ions from the mass spectra, which may involve increasing the number of major ions (e.g., >30 ions) to be examined, developing a comprehensive library of fragmentation patterns of identified tracers, or utilizing a different ionization technique (e.g., NO_3^- and I^-) that detect highly oxygenated compounds. Iodide-ToF-CIMS can effectively identify organonitrates, monomers, and dimers with 3–6 oxygens while Nitrate CI-ToF-MS can measure similar compounds with

more than 6 oxygens[60]. Nevertheless, the unaccounted mass of submicron OM had a significant relationship with urban ($r^2 = 0.54$) and biogenic ($r^2 = 0.60$) tracers, as comparing to BB tracer ($r^2 = 0.32$), which may indicate the possible sources of the organic compounds (see Supplementary Fig. 6 for correlation analysis). Moreover, in addition to the consistency in time series as shown in Fig. 1a, correlation analysis indicates a strong linear correlation ($r^2 = 0.8124$) between the total mass of OM and that of the top 30 ions. This result supports that the major ions were suitable representatives in assessing the sources of submicron OM and its chemical characteristics, and that the unaccounted portion of OM were still relevant to urban, biomass, or biogenic emissions.

The mass concentrations of most tracers (i.e., phthalic and pinonic acid) reported here were higher compared with other sampling sites. The possible impact of isomeric and isobaric isomers are not discounted in this study, particularly the contribution of products from fragmentation due to thermal desorption and collision induced dissociation (CID). More importantly, the elevated relative humidity and photochemical activities also justify the high mixing ratios of particle-phase bound tracers. A chamber study indicated that the measured fraction in the particle phase and R.H. showed an apparent relationship indicating the importance of such meteorological conditions in the formation of the aerosol tracers[61]. Though the controlled study only constrained such observation for pinonaldehyde, it may as well be extended to the formation of the oxygenated compounds. Likewise, the high R.H. and aerosol water liquid content (~91% and 30 µg m^{-3}, daytime average, shown in Supplementary Fig. 7) during the field campaign favored the partitioning of polar compounds, which perchance intensified the mass concentration of some of the semi-volatile organic tracers in the aerosols. Moreover, model simulation using ISORROPIA II[62]

indicated that PM in Xitou forest was acidic, with *pH* of 2.14 (min:1.45; max: 2.65, shown in Supplementary Fig. 7). The acidic condition of PM induced the partitioning of organic compounds in the particle phase, further justifying the elevated concentration of most tracers compared to the other sites. This is coherent with a previous chamber study that showed an increase of mass of particle phase organics by 40% when utilizing acidic aerosols[63].

METHODS

Site description and instrumentation
A 30-day measurement was conducted at the Xitou Experimental Forest of National Taiwan University, located in Nantou, Taiwan during the fall season of 2015. At 1120 meters above sea level, the forest with 2500 total hectares is surrounded by three different mountains. During the field measurement, the mean temperature and relative humidity were 20.5 °C and 91%, More information regarding the forest can be found somewhere else[4,15].

The instruments used in this study were housed in a wooden cabin with proper ventilation. The VOCs were measured using a proton transfer reaction - time of flight mass spectrometer (PTR-ToF-MS 8000, IONICON Analytik) at 132 E/N ratio (Townsend). Description of the general technique of the PTR-ToF-MS can be found elsewhere. The PTR-ToF-MS was calibrated daily using a 110-ppb mixture of gases (isoprene, limonene, benzene, toluene, ethylbenzene, dichlorobenzene, trichlorobenzene, and trimethyl-benzene, Restek Corp). An advanced Liquid Calibration Unit (LCU-a, IONICON Analytik) was utilized to dilute the gaseous standard (110 ppb) for calibration range of 1.0 to 20 ppb with high precision.

Submicron aerosols ($PM_{1.0}$) was collected using a pair of US federal reference method samplers equipped with sharp-cut cyclones with 1 µm cut-off diameter. The $PM_{1.0}$ samplers were loaded and replaced daily with a Teflon filter and a double-layered quartz filter (QBQ set-up), respectively. The sampling duration of the $PM_{1.0}$ samplers was from 9:00 to 18:00 ("daytime") and from 20:00 to 06:00 ("nighttime), local time. Most of the data presented in this study used the daytime samples unless indicated otherwise. Size resolved collection of submicron aerosols was also done in 21 selected days of the field measurement using a micro-orifice uniform deposit impactor (MOUDI 110, MSP corporation).

Characterization of the particle-bound organic compounds and water-soluble ions
A modified thermal desorption unit (DRI Model 2001A, Atmoslytic, Inc.) was utilized to evaporate the organic compounds from a $0.528 \, cm^2$ punch quartz filter. The desorption unit was designed to evaporate the organic compounds in a stepwise manner with a five-minute interval each step. The four target temperatures were 100 °C, 200 °C, 350 °C, and 550 °C. Ultra-pure nitrogen (99.9995%, Air Products) was used as the carrier gas and a heated transfer line (95–100 °C) served as the connector between the desorption unit and the PTR-ToF-MS. At the same time, the inlet temperature of the PTR-ToF-MS was set at 100 °C, to minimize condensation loss. A detailed description of the whole TD-PTR-ToF-MS tandem unit can be found in prior studies[4,12].

The mass concentration of the identified ions from the thermal desorption was calculated using the Eq. (3).

$$n_{aer,M} = \frac{C_{ave,M} \times MW_M \times F_{N2} \times t_{meas}}{22.4 \times F_{col} \times t_{col} \times 0.001 \times f_{filter}} \quad (3)$$

Here, $n_{aer,M}$ is the mass concentration of an organic compound (M) in ng m^{-3}, $C_{ave,M}$ is the average of the mixing ratio (ppb) measured by the PTR-ToF-MS at a given time of measurement (t_{meas}) and flow rate of the carrier gas (F_{N2}), which were 20 min and 0.095 LPM, respectively. The F_{col} and t_{col} are the flow rate and sampling time duration of the aerosol sampler (PQ-200), which

were 16.7 LPM and 9 h for daytime and 10 h for nighttime. F_{filter} is the fraction of the quartz filter introduced to the TD-PTR-ToF-MS unit using a customized platinum boat. The mass concentrations of the identified ions were background corrected by subtracting the signal measured from field blank filters.

Gravimetric analyses of the Teflon filters were done using a microbalance (Model XP6, Mettler Toledo). Particle bound water-soluble ions (Na^+, NH_4^+, K^+, Ca^{2+}, Mg^{2+}, Cl^-, NO_3^-, and SO_4^{2-}) were analyzed using ion chromatography (ICS-1000, Dionex). Sugar-related organic compounds such as levoglucosan (1,6-anhydro-β-D-glucopyranose), galactosan (1,6-anhydro-β-D-galacto-pyranose), and arabitol were analyzed using ion chromatography. Filter samples were cut into 17 mm round portions, each punched filter was placed in a 4 mL amber glass screw thread vial and extracted with 3 mL deionized ultra-pure water under ultrasonic agitation for 60 min. The extracts were then filtered using a syringe filter (25 mm filter with 0.45 µm PTFE membrane, Pall Corporation, East Hills, NY, USA) to remove insoluble materials. The extracts were analyzed using high-performance anion-exchange chromatography (HPAEC) with pulsed amperometric detection on a Dionex ICS-5000 Ion Chromatograph, using a Dionex Carbopac MA1 guard column and a Dionex Carbopac MA1 analytical column ($4 \times 250 \, mm$) with sodium hydroxide solution (400 mM, 0.4 mL min^{-1}) as eluent. Organic Carbon/Elemental Carbon (OC/EC) analysis of quartz filters was done using DRI Model2001A Thermal/Optical Carbon Analyzer (Atmoslytic, USA) with IMPROVE_A method. Particulate organic matter (OM) was determined using 1.6 as a conversion factor[18]. The value of 1.6 was assigned for urban aerosols while non-urban areas such as the Xitou forest site should utilize 2.1. However, we believe that the forest site was heavily impacted by urban processes, thus utilizing the 1.6 factor appears to be reasonable for this case. This was corroborated by the determined mean molecular weight per carbon weight of the ions measured by the PTR-TOF-MS (1.65 ± 0.46).

Positive matrix factorization for source apportionment
Source apportionment model such as the positive matrix factorization (PMF) provides quantification of the contribution of the sources to the pollutant concentration. The PMF model is described in detail elsewhere[64]. We applied the EPA PMF model[65] version 5.0.14 to the top 30 major ions extracted from the daytime samples, including the mass concentration of OM measured by using OC/EC analysis. The data matrix, therefore, had dimensions of 31 analytes × 30 filter samples. The model was run 20 times with a randomly chosen starting point for each run.

Calculation of pH and liquid water content of PM using ISORROPIA II
The acidity and LWC of submicron PM in Xitou forest were simulated using ISORROPIA II, a thermodynamic equilibrium model using inorganic ions (e.g. K^+, Ca^{2+}, NO_3^- and SO_4^{2-}) as primary inputs[62]. ISORROPIA provides the hydronium ion concentration and water uptake using relative humidity and temperature as critical parameters beside the mass concentration of water soluble inorganic ions. In Xitou forest, the daily average temperatures and RH range between 291.3–296.1 K and 81–97%, respectively. Simulation was performed in the forward mode and for aerosols in the metastable conditions. Time series profile of aerosol LWC and pH are provided in Supplementary Fig. 7.

DATA AVAILABILITY
Field measurement and lab study datasets are available from the corresponding author (C.C.-K.C.).

REFERENCES

1. Shrivastava, M. et al. Urban pollution greatly enhances formation of natural aerosols over the Amazon rainforest. *Nat. Commun.* **10**, 1046 (2019).
2. Liu, Y. et al. Isoprene photo-oxidation products quantify the effect of pollution on hydroxyl radicals over Amazonia. *Sci. Adv.* **4**, eaar2547 (2018).
3. de Sá, S. S. et al. Urban influence on the concentration and composition of submicron particulate matter in central Amazonia. *Atmos. Chem. Phys.* **18**, 12185–12206 (2018).
4. Salvador, C. M. et al. Measurements of submicron organonitrate particles: Implications for the impacts of NOx pollution in a subtropical forest. *Atmos. Res.* **245**, 105080 (2020).
5. Salvador, C. M. C. et al. Contribution of terpenes to ozone formation and secondary organic aerosols in a subtropical forest impacted by urban pollution. *Atmosphere* **11**, 1232 (2020).
6. Setyan, A. et al. Chemistry of new particle growth in mixed urban and biogenic emissions - insights from CARES. *Atmos. Chem. Phys.* **14**, 6477–6494 (2014).
7. Bauer, H. et al. Arabitol and mannitol as tracers for the quantification of airborne fungal spores. *Atmos. Environ.* **42**, 588–593 (2008).
8. Ding, X., He, Q. F., Shen, R. Q., Yu, Q. Q. & Wang, X. M. Spatial distributions of secondary organic aerosols from isoprene, monoterpenes, beta-caryophyllene, and aromatics over China during summer. *J. Geophys. Res.-Atmos.* **119**, 11877–11891 (2014).
9. Fu, P. Q. et al. Diurnal variations of organic molecular tracers and stable carbon isotopic composition in atmospheric aerosols over Mt. Tai in the North China Plain: an influence of biomass burning. *Atmos. Chem. Phys.* **12**, 8359–8375 (2012).
10. Pietrogrande, M. C., Bacco, D., Visentin, M., Ferrari, S. & Poluzzi, V. Polar organic marker compounds in atmospheric aerosol in the Po Valley during the Supersito campaigns — Part 1: Low molecular weight carboxylic acids in cold seasons. *Atmos. Environ.* **86**, 164–175 (2014).
11. Turnbull, J. C. et al. Atmospheric observations of carbon monoxide and fossil fuel CO_2 emissions from East Asia. *J. Geophys. Res.: Atmos.* **116** (2011).
12. Salvador, C. M. et al. Characterization of the organic matter in submicron urban aerosols using a Thermo-Desorption Proton-Transfer-Reaction Time-of-Flight Mass Spectrometer (TD-PTR-TOF-MS). *Atmos. Environ.* **140**, 565–575 (2016).
13. Gkatzelis, G. I. et al. Comparison of three aerosol chemical characterization techniques utilizing PTR-ToF-MS: a study on freshly formed and aged biogenic SOA. *Atmos. Meas. Tech.* **11**, 1481–1500 (2018).
14. Zhou, S. et al. Influences of upwind emission sources and atmospheric processing on aerosol chemistry and properties at a rural location in the Northeastern U.S. *J. Geophys. Res.: Atmos.* **121**, 6049–6065 (2016).
15. Lee, C. S. L. et al. Seasonal variation of chemical characteristics of fine particulate matter at a high-elevation subtropical forest in East Asia. *Environ. Pollut.* **246**, 668–677 (2019).
16. Griffith, S. M. et al. Long-range air pollution transport in East Asia during the first week of the COVID-19 lockdown in China. *Sci. Total Environ.* **741**, 140214 (2020).
17. Mohr, C. et al. Molecular identification of organic vapors driving atmospheric nanoparticle growth. *Nat. Commun.* **10**, 4442 (2019).
18. Turpin, B. J. & Lim, H.-J. Species contributions to PM2.5 mass concentrations: revisiting common assumptions for estimating organic mass. *Aerosol Sci. Technol.* **35**, 602–610 (2001).
19. Li, M., Chen, H., Yang, X., Chen, J. & Li, C. Direct quantification of organic acids in aerosols by desorption electrospray ionization mass spectrometry. *Atmos. Environ.* **43**, 2717–2720 (2009).
20. Sorooshian, A. et al. On the source of organic acid aerosol layers above clouds. *Environ. Sci. Technol.* **41**, 4647–4654 (2007).
21. Eddingsaas, N. C. et al. α-pinene photooxidation under controlled chemical conditions Part 2: SOA yield and composition in low- and high-NOx environments. *Atmos. Chem. Phys.* **12**, 7413–7427 (2012).
22. Simoneit, B. R. T. et al. Levoglucosan, a tracer for cellulose in biomass burning and atmospheric particles. *Atmos. Environ.* **33**, 173–182 (1999).
23. Holzinger, R., Kasper-Giebl, A., Staudinger, M., Schauer, G. & Röckmann, T. Analysis of the chemical composition of organic aerosol at the Mt. Sonnblick observatory using a novel high mass resolution thermal-desorption proton-transfer-reaction mass-spectrometer (hr-TD-PTR-MS). *Atmos. Chem. Phys.* **10**, 10111–10128 (2010).
24. He, X. et al. Abundance and sources of phthalic acids, benzene-tricarboxylic acids, and phenolic acids in PM2.5 at urban and suburban sites in Southern China. *ACS Earth Space Chem.* **2**, 147–158 (2018).
25. Balla, D., Voutsa, D. & Samara, C. Study of polar organic compounds in airborne particulate matter of a coastal urban city. *Environ. Sci. Pollut. Res.* **25**, 12191–12205 (2018).
26. Donahue, N. M., Epstein, S. A., Pandis, S. N. & Robinson, A. L. A two-dimensional volatility basis set: 1. organic-aerosol mixing thermodynamics. *Atmos. Chem. Phys.* **11**, 3303–3318 (2011).
27. Hartungen, E. V. et al. Proton-transfer-reaction mass spectrometry (PTR-MS) of carboxylic acids: Determination of Henry's law constants and axillary odour investigations. *Int. J. Mass Spectrom.* **239**, 243–248 (2004).
28. Romano, A. & Hanna, G. B. Identification and quantification of VOCs by proton transfer reaction time of flight mass spectrometry: An experimental workflow for the optimization of specificity, sensitivity, and accuracy. *J. Mass Spectrom.* **53**, 287–295 (2018).
29. Olivenza-León, D., Mayhew, C. A. & González-Méndez, R. Proton transfer reaction mass spectrometry investigations of phthalate esters via direct headspace sampling. *Int. J. Mass Spectrom.* **461**, 116497 (2021).
30. Fu, P. Q., Kawamura, K., Pavuluri, C. M., Swaminathan, T. & Chen, J. Molecular characterization of urban organic aerosol in tropical India: contributions of primary emissions and secondary photooxidation. *Atmos. Chem. Phys.* **10**, 2663–2689 (2010).
31. Ho, K. F. et al. Dicarboxylic acids, ketocarboxylic acids, α-dicarbonyls, fatty acids, and benzoic acid in urban aerosols collected during the 2006 Campaign of Air Quality Research in Beijing (CAREBeijing-2006). *J. Geophys. Res. Atmos.* **115**, D19312 (2010).
32. Kleindienst, T. E., Jaoui, M., Lewandowski, M., Offenberg, J. H. & Docherty, K. S. The formation of SOA and chemical tracer compounds from the photooxidation of naphthalene and its methyl analogs in the presence and absence of nitrogen oxides. *Atmos. Chem. Phys.* **12**, 8711–8726 (2012).
33. Jia, C. & Batterman, S. A critical review of naphthalene sources and exposures relevant to indoor and outdoor air. *Int J. Environ. Res Public Health* **7**, 2903 (2010).
34. Zhang, X. et al. PM2.5 bound phthalates in four metropolitan cities of China: Concentration, seasonal pattern and health risk via inhalation. *Sci. Total Environ.* **696**, 133982 (2019).
35. He, Y., Wang, Q., He, W. & Xu, F. Phthalate esters (PAEs) in atmospheric particles around a large shallow natural lake (Lake Chaohu, China). *Sci. Total Environ.* **687**, 297–308 (2019).
36. Kong, S. et al. Spatial and temporal variation of phthalic acid esters (PAEs) in atmospheric PM10 and PM2.5 and the influence of ambient temperature in Tianjin, China. *Atmos. Environ.* **74**, 199–208 (2013).
37. Kawamura, K. & Bikkina, S. A review of dicarboxylic acids and related compounds in atmospheric aerosols: Molecular distributions, sources and transformation. *Atmos. Res.* **170**, 140–160 (2016).
38. Lin, Y. et al. Interaction between succinic acid and sulfuric acid–base clusters. *Atmos. Chem. Phys.* **19**, 8003–8019 (2019).
39. Fang, X. et al. Observational evidence for the involvement of dicarboxylic acids in particle nucleation. *Environ. Sci. Technol. Lett.* **7**, 388–394 (2020).
40. Hsieh, L. Y., Kuo, S. C., Chen, C. L. & Tsai, Y. I. Origin of low-molecular-weight dicarboxylic acids and their concentration and size distribution variation in suburban aerosol. *Atmos. Environ.* **41**, 6648–6661 (2007).
41. Hanson, D. R., Koppes, M., Stoffers, A., Harsdorf, R. & Edelen, K. Proton transfer mass spectrometry at 11hPa with a circular glow discharge: Sensitivities and applications. *Int. J. Mass Spectrom.* **282**, 28–37 (2009).
42. Ho, K. F. et al. Dicarboxylic acids, ketocarboxylic acids, and dicarbonyls in the urban atmosphere of China. *J. Geophys. Res. Atmos.* **112** (2007).
43. Cong, Z., Kawamura, K., Kang, S. & Fu, P. Penetration of biomass-burning emissions from South Asia through the Himalayas: new insights from atmospheric organic acids. *Sci. Rep.* **5**, 9580 (2015).
44. Zhang, Y. L. et al. Aircraft observations of water-soluble dicarboxylic acids in the aerosols over China. *Atmos. Chem. Phys.* **16**, 6407–6419 (2016).
45. Urban, R. C. et al. Use of levoglucosan, potassium, and water-soluble organic carbon to characterize the origins of biomass-burning aerosols. *Atmos. Environ.* **61**, 562–569 (2012).
46. Chuang, M.-T. et al. Characterization of aerosol chemical properties from near-source biomass burning in the northern Indochina during 7-SEAS/Dongsha experiment. *Atmos. Environ.* **78**, 72–81 (2013).
47. Fu, P. Q. et al. Organic molecular compositions and temporal variations of summertime mountain aerosols over Mt. Tai, North China Plain. *J. Geophys. Res.-Atmos.* **113**, 20 (2008).
48. Claeys, M. et al. Polar organic marker compounds in atmospheric aerosols during the LBA-SMOCC 2002 biomass burning experiment in Rondônia, Brazil: sources and source processes, time series, diel variations and size distributions. *Atmos. Chem. Phys.* **10**, 9319–9331 (2010).
49. Pio, C. A. et al. Chemical composition of atmospheric aerosols during the 2003 summer intense forest fire period. *Atmos. Environ.* **42**, 7530–7543 (2008).

50. Fu, P. & Kawamura, K. Diurnal variations of polar organic tracers in summer forest aerosols: a case study of a Quercus and Picea mixed forest in Hokkaido, Japan. *Geochem. J.* **45**, 297–308 (2011).
51. García, M. I., van Drooge, B. L., Rodríguez, S. & Alastuey, A. Speciation of organic aerosols in the Saharan Air Layer and in the free troposphere westerlies. *Atmos. Chem. Phys.* **17**, 8939–8958 (2017).
52. Chio, C. P., Cheng, M. T. & Wang, C. F. Source apportionment to PM10 in different air quality conditions for Taichung urban and coastal areas, Taiwan. *Atmos. Environ.* **38**, 6893–6905 (2004).
53. Chou, C. C. K. et al. Seasonal variation and spatial distribution of carbonaceous aerosols in Taiwan. *Atmos. Chem. Phys.* **10**, 9563–9578 (2010).
54. Simoneit, B. R. T. et al. Lignin pyrolysis products, lignans, and resin acids as specific tracers of plant classes in emissions from biomass combustion. *Environ. Sci. Technol.* **27**, 2533–2541 (1993).
55. Shakya, K. M., Louchouarn, P. & Griffin, R. J. Lignin-derived phenols in houston aerosols: implications for natural background sources. *Environ. Sci. Technol.* **45**, 8268–8275 (2011).
56. Zangrando, R. et al. Molecular markers of biomass burning in arctic aerosols. *Environ. Sci. Technol.* **47**, 8565–8574 (2013).
57. Finewax, Z., de Gouw, J. A. & Ziemann, P. J. Identification and quantification of 4-nitrocatechol formed from oh and no3 radical-initiated reactions of catechol in air in the presence of NOx: implications for secondary organic aerosol formation from biomass burning. *Environ. Sci. Technol.* **52**, 1981–1989 (2018).
58. Yee, L. D. et al. Secondary organic aerosol formation from biomass burning intermediates: phenol and methoxyphenols. *Atmos. Chem. Phys.* **13**, 8019–8043 (2013).
59. Hodnebrog, Ø., Myhre, G., Forster, P. M., Sillmann, J. & Samset, B. H. Local biomass burning is a dominant cause of the observed precipitation reduction in southern Africa. *Nat. Commun.* **7**, 11236 (2016).
60. Riva, M. et al. Evaluating the performance of five different chemical ionization techniques for detecting gaseous oxygenated organic species. *Atmos. Meas. Tech.* **12**, 2403–2421 (2019).
61. Zhao, Y. et al. Insights into secondary organic aerosol formation mechanisms from measured gas/particle partitioning of specific organic tracer compounds. *Environ. Sci. Technol.* **47**, 3781–3787 (2013).
62. Fountoukis, C. & Nenes, A. ISORROPIA II: a computationally efficient thermodynamic equilibrium model for K+;Ca2+;Mg2+;NH4+;Na+;SO42-;NO3-;Cl;H2O aerosols. *Atmos. Chem. Phys.* **7**, 4639–4659 (2007).
63. Iinuma, Y., Böge, O., Gnauk, T. & Herrmann, H. Aerosol-chamber study of the α-pinene/O3 reaction: influence of particle acidity on aerosol yields and products. *Atmos. Environ.* **38**, 761–773 (2004).
64. Paatero, P., Hopke, P. K., Hoppenstock, J. & Eberly, S. I. Advanced factor analysis of spatial distributions of PM2.5 in the Eastern United States. *Environ. Sci. Technol.* **37**, 2460–2476 (2003).
65. Norris, G., Duvall, R., Brown, S. & Bai, S. *EPA Positive Matrix Factorization (PMF) 5.0 Fundamentals and User Guide*. (U.S. Environmental Protection Agency, EPA/600/R-14/108 (NTIS PB2015-105147), 2014).

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41612-022-00314-x.

Correspondence and requests for materials should be addressed to Charles C.-K. Chou.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.