Environ Monit Assess (2019) 191: 78
https://doi.org/10.1007/s10661-019-7205-x

Atmo-ecometabolomics: a novel atmospheric particle chemical characterization methodology for ecological research

Albert Rivas-Ubach · Yina Liu · Allison L. Steiner · Jordi Sardans · Malak M. Tfaily · Gourihar Kulkarni · Young-Mo Kim · Eric Bourrianne · Ljiljana Paša-Tolić · Josep Peñuelas · Alex Guenther

Received: 8 June 2018 / Accepted: 4 January 2019 / Published online: 16 January 2019

Abstract Aerosol particles play important roles in processes controlling the composition of the atmosphere and function of ecosystems. A better understanding of the composition of aerosol particles is beginning to be recognized as critical for ecological research to further comprehend the link between aerosols and ecosystems. While chemical characterization of aerosols has been practiced in the atmospheric science community, detailed methodology tailored to the needs of ecological research does not exist yet. In this study, we describe an efficient methodology (atmo-ecometabolomics), in step-by-step details, from the sampling to the data analyses, to characterize the chemical composition of aerosol particles, namely atmo-metabolome. This method employs mass spectrometry platforms such as liquid and gas chromatography mass spectrometries (MS) and Fourier transform ion cyclotron resonance MS (FT-ICR-MS). For methodology evaluation, we analyzed aerosol particles collected during two different seasons (spring and summer) in a low-biological-activity ecosystem. Additionally, to further validate our methodology, we analyzed aerosol particles collected in

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10661-019-7205-x) contains supplementary material, which is available to authorized users.

A. Rivas-Ubach (✉) · Y. Liu · M. M. Tfaily · L. Paša-Tolić
Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA 99354, USA
e-mail: albert.rivas.ubach@gmail.com

e-mail: albert.rivas.ubach@pnnl.gov

Y. Liu
Geochemical and Environmental Research Group, Texas A&M University, College Station, TX 77845, USA

A. L. Steiner
Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI 48109, USA

J. Sardans · J. Peñuelas
CREAF, Campus UAB, 08913 Cerdanyola del Vallès, Catalonia, Spain

J. Sardans · J. Peñuelas
Global Ecology Unit CREAF-CSIC, Campus UAB, 08913 Cerdanyola del Vallès, Catalonia, Spain

G. Kulkarni
Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, WA 99352, USA

Y.-M. Kim
Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99354, USA

E. Bourrianne
Faculté des Sicences et d'Ingénierie, Université de Toulouse III Paul Sabatier, 31400 Toulouse, France

A. Guenther
Department of Earth System Science, University of California, Irvine, Irvine, CA 92697, USA

a more biologically active ecosystem during the pollination peaks of three different representative tree species. Our statistical results showed that our sampling and extraction methods are suitable for characterizing the atmo-ecometabolomes in these two distinct ecosystems with any of the analytical platforms. Datasets obtained from each mass spectrometry instrument showed overall significant differences of the atmo-ecometabolomes between spring and summer as well as between the three pollination peak periods. Furthermore, we have identified several metabolites that can be attributed to pollen and other plant-related aerosol particles. We additionally provide a basic guide of the potential use ecometabolomic techniques on different mass spectrometry platforms to accurately analyze the atmo-ecometabolomes for ecological studies. Our method represents an advanced novel approach for future studies in the impact of aerosol particle chemical compositions on ecosystem structure and function and biogeochemistry.

Keywords Aerosol particles · Metabolomics · Ecosystems · Biomarkers · Mass spectrometry · FT-ICR

Introduction

Aerosols are solids and/or liquids suspended in the atmosphere that are derived from both anthropogenic and biogenic sources (Canagaratna et al. 2007). Primary biogenic aerosol particles (PBAP), which include pollen, microorganisms, and spores, as well as fragments from animal and plant debris, are directly released from biological systems (Després et al. 2012). Primary producers also generate large volumes of gas-phase volatile organic compounds (VOCs), which are emitted into the atmosphere and, together with anthropogenic gases, are oxidized and can condense to form secondary organic aerosols (SOA) (Després et al. 2012; Fuzzi et al. 2006; Pandis et al. 1992) (Fig. 1).

To date, most research has focused on the impact of aerosols in climate and atmospheric processes (Andreae and Crutzen 1997; Ayers and Gras 1991; Baustian et al. 2012; Carlton et al. 2010; Després et al. 2012; Jokinen et al. 2015; Ramanathan et al. 2001; Zhang et al. 2004). However, many components of the biosphere are constantly in contact with aerosol particles. This aerosol-biosphere interaction can play critical roles in the structure and function of aquatic and terrestrial ecosystems (Baker et al. 2003; Gu et al. 2002; Mahowald et al. 2005; Seco et al. 2007). For example, it is well known that plants can absorb deposited particles from the atmosphere (Fageria et al. 2009; Seco et al. 2007; Uzu et al. 2010; Wedding et al. 1975), although the effects of plant particle uptake are still not totally understood. Furthermore, most of such research was focused on trace metals (Achotegui-Castells et al. 2013; Feng et al. 2011; Uzu et al. 2010; Xiong et al. 2014) and other nutrients that can play significant roles in agroecosystems (Fernández and Brown 2013). In

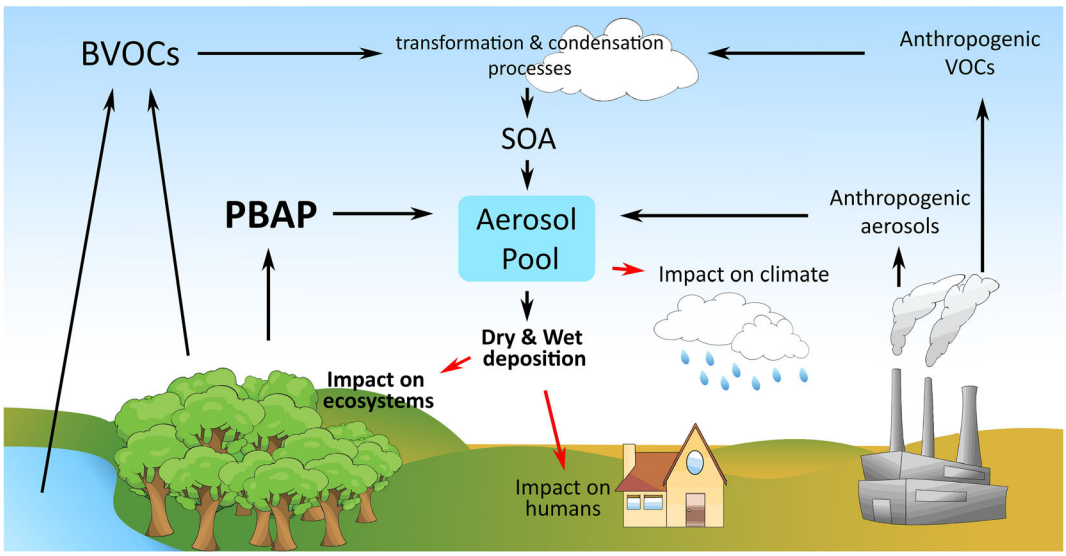

Fig. 1 Schematic diagram of the aerosol emissions and deposition on ecosystems

natural ecosystems, aerosol particles can serve as an important source of nutrients to diverse components of the biosphere. For example, the phyllosphere, which is the microbial community coexisting on plant leaves and have a close relationship with the physiology of plants (Arnold et al. 2000; Lindow and Brandl 2003; Vorholt 2012), can benefit from aerosol particles to acquire nutrients (e.g., nitrogen (N), phosphorus (P), sulfur (S), …). Variations in aerosol particle chemical composition could lead to significant changes in microbial abundance and diversity, affecting their host physiology, and, therefore, the ecosystem structure and function (Peñuelas and Terradas 2014; Whipps et al. 2008). In aquatic ecosystems, the nutritional effects of aerosol deposition on phytoplankton has been studied in great extent (Baker et al. 2003; Paerl 1997; Paytan et al. 2009; Wang et al. 2015), as it represents an important fraction of nutrients for several aquatic primary producers (e.g., cyanobacteria) (Baker et al. 2003; Wang et al. 2015). Furthermore, research in the field of ecological stoichiometry, the study on the effects of balance of essential elements on organisms and ecosystems, has proven that changes in nutrient proportions, mainly carbon (C), N, and P, can lead to substantial alterations in ecosystems (Elser et al. 1996; Sterner and Elser 2002). Therefore, changes in aerosol particle elemental composition may lead to significant modifications in terrestrial and aquatic ecosystems functioning (Carnicer et al. 2015; Peñuelas et al. 2012; Sardans et al. 2012). Understanding the elemental and molecular compositions of aerosol particles is thus crucial for comprehending processes occurring at the atmosphere-ecosystem interface and for understanding how aerosols can be related to ecosystemic changes.

Low molecular weight compounds (~ 80–1000 Da) have not generally identified or taken into account in aerosol particles and could play crucial roles in the ecosystem functioning, especially at the atmosphere-biosphere interface. Here, we describe in detail a novel and simple methodology to collect ambient aerosol particle samples and subsequently characterize their chemical composition using different mass spectrometry (MS) platforms. This methodology represents a useful tool to shed light on novel questions in plant physiology, ecophysiology, and ecology to understand deeper the atmosphere-ecosystem interface. In this manuscript, we first briefly define ecometabolomics and its potential as a technique to characterize the chemical composition of aerosol particles. We then provide a basic overview of the common MS techniques used in metabolomics. Finally, we discuss our methodologic approach and the different recent metabolomics technologies and, briefly, the obtained results from the samples collected and analyzed with the proposed methodology.

Atmo-ecometabolomics: metabolomics of aerosols

Metabolomics aims to study the metabolome of entire organisms or specific cells or tissues, and this technique includes all the procedures from sample collection, metabolite extraction, and analyses to data filtering, statistical analyses, and interpretation of the results (Fig. 2). A metabolome consists of the total set of small (< 1000 Da) compounds (metabolites) present in a sample at a given moment (Fiehn 2002). Such compounds are

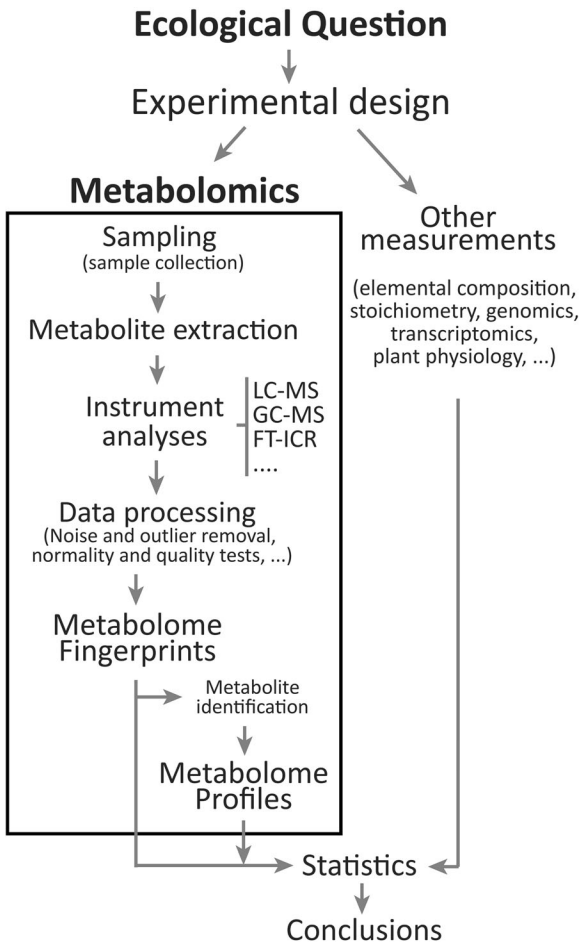

Fig. 2 Diagram of the main procedures of a general ecometabolomic study combined with complementary measurements

represented by the substrates and products from cellular primary metabolism such as organic acids, amino acids, and sugars as well as a vast diversity of compounds derived from the secondary metabolism of organisms such as polyphenolics and terpenes, typically synthetized by plants and fungi. All these metabolites are involved in diverse and complex chemical reactions to finally maintain organisms' homeostasis, reproduction, and growth (Peñuelas and Sardans 2009). Metabolomics has been widely used in human nutrition research (Gibney et al. 2005; Wishart 2008), medicine (Claudino et al. 2007; Walsh et al. 2006), plant physiology (Hirai et al. 2004; Kaplan et al. 2004), and more recently in the field of ecology (ecometabolomics) (Bundy et al. 2008; Gargallo-Garriga et al. 2014; Rivas-Ubach et al. 2012, 2016a, 2018a). Ecometabolomics has proven to be useful in understanding the plasticity of organisms' metabolomes under specific environmental conditions and/or stressor pressures (Gargallo-Garriga et al. 2014; Rivas-Ubach et al. 2016b, 2017; Sardans et al. 2011, 2014). However, ecometabolomic analyses of aerosol particles (atmo-ecometabolomics) have never been reported before and it is a step further to characterize the aerosol particle composition and to improve our understanding of the aerosol-biosphere interface.

Metabolomics can detect specific molecular signatures (biomarkers) directly related to organismal stress and the phenological status of ecosystems. Recent climate projections predict increments of extreme climatic events such as drought and warming that increase plant BVOC emissions (Peñuelas and Staudt 2010) and lead to shifts in the phenology of plants (Menzel et al. 2006; Parmesan 2006; Parmesan and Yohe 2003; Walther et al. 2002). Terrestrial plants have proven to show large overall metabolome shifts when exposed to stressors (Gargallo-Garriga et al. 2014, 2015; Leiss et al. 2009; Macedo 2012; Rivas-Ubach et al. 2012, 2014), and several stress biomarkers have been already identified (Glauser et al. 2008; Guy et al. 2007; Keltjens and van Beusichem 1998; Shulaev et al. 2008; Thompson et al. 2005). Significant changes in plant phenology have been detected during the last decades (Menzel et al. 2006; Parmesan 2006; Parmesan and Yohe 2003; Walther et al. 2002), and due to the direct link between PBAP and ecosystems, such phenological shifts should be detectable with atmo-ecometabolomics. Therefore, long-term atmo-ecometabolomics studies should provide valuble information on phenological changes and succession or recession of ecosystems. Atmo-ecometabolomics can also be applied to assess stress and phenological changes at ecosystem and regional scales through the characterization of atmo-ecometabolomes.

Analytical instruments for MS-based metabolomics analyses

Metabolomic profiles are achieved through two main processes: (i) obtaining the chemical signature of the sample (metabolomic fingerprint) and (ii) identification of metabolites through the information previously obtained from the metabolomic fingerprints (e.g., exact mass, retention time, fragmentation pattern, etc). LC-MS and GC-MS are currently two of the most common techniques used in metabolomics (Fiehn 2002; Sardans et al. 2011; Zhang et al. 2012), demonstrating high performance and sensitivity (Pan and Raftery 2007). LC-MS and GC-MS provide similar data formats; i.e., in both instruments, before compound detection, metabolites are separated by chromatography (liquid or gas), producing finally two orthogonal and independent values, retention time (RT) and mass-to-charge ratio (m/z), for each of the detected features. RT and m/z, together with other additional parameters (e.g., fragmentation pattern), can be thus used for metabolite identification (Sumner et al. 2007). Although a wide range of compounds can be detected by GC-MS and LC-MS platforms, generally, GC-MS is more suitable for detecting highly volatile and non-polar compounds with mass weight typically < 600 Da such as fatty acids, carotenoids, and essential oils (Gullberg et al. 2004). Untargeted GC-MS analyses require previous compound derivatization of the samples to allow polar molecular compounds to be sufficiently volatile so they can elute at moderate temperature without disintegration. Therefore, carbohydrates, organic acids, and other polar analytes can be also detected by GC-MS after derivatization (Gullberg et al. 2004). In comparison to LC, GC has proved to have exceptional reproducibility showing little RT shifts between samples. However, derivatization of the samples increases preparation time and provides indirect metabolite detection that complicates the elucidation of unknown compounds. On the other hand, LC-MS can cover a wider range of compounds than GC-MS and allows the detection of compounds with larger molecular weight (typically up to 1200 Da for metabolomics). LC-MS allows analyzing non-volatile and highly polar compounds ranging from primary metabolic compounds such as organic acids, carbohydrates,

and amino acids to secondary metabolites such as alkaloids, phenolic acids, flavonoids, or saponins (De Vos et al. 2007). LC-MS provides a direct detection of molecular ions because derivatization is not required. However, LC-MS analyses commonly show larger RT shifts between samples in comparison to GC-MS, especially in studies with a large number of samples (e.g., > 300). Nonetheless, no single mass spectrometry instrument can cover all molecular compound classes (Ding et al. 2007; Zhang et al. 2012), and combining diverse platforms is a common approach for deeper metabolome characterization of the samples (Hall 2006).

Mass resolving power of MS instruments is a critical factor to consider in MS-based metabolomics. The high resolution and mass accuracy (typically up to 250,000 and < 3 ppm, respectively) of Orbitrap mass spectrometers have proven to be suitable for ecometabolomic studies (White et al. 2017), reducing considerably the error of metabolite identification when using high-resolution compound libraries (Rivas-Ubach et al. 2016c). However, Fourier transform ion cyclotron resonance MS (FT-ICR-MS) is currently the MS platform achieving the highest mass resolving power (up to $m/\Delta m_{50\%}$ > 2,700,000 at m/z 400) and mass accuracy (< 1 ppm mass error after internal calibration) (Smith et al. 2018). Although FT-ICR-MS can be coupled to HPLC, the scan rate of FT-ICR-MS is commonly not rapid enough as to analyze samples at ultra-high resolution (> 400,000) when coupled to LC. This is one of the main reasons why direct infusion (DI) has been the most typical approach to analyze samples with the FT-ICR-MS in ultra-high resolution. Furthermore, DI-FT-ICR-MS acquisition time is significantly shorter (typically between 5 and 15 min) than that of the LC or GC methods which can take over 40 min per sample. The ultra-high mass resolution and excellent mass accuracy of DI-FT-ICR-MS enable the assignment of an elemental formula of a wide portion of the detected features based exclusively on their exact mass (Klein et al. 2006; Kujawinski 2002) providing, thus, powerful means to understand the overall chemical characteristics of organic complex samples (Kim et al. 2003; Reemtsma 2009; Roullier-Gall et al. 2014; Schmitt-Kopplin et al. 2012; Sleighter and Hatcher 2007; Tfaily et al. 2015). Assessing the diversity of molecular compounds with nutrients such as P, N, or S and understanding how the elemental composition in aerosols shifts due to environmental changes are of special interest for ecological stoichiometry research, and they are, therefore, feasible with DI-FT-ICR-MS.

Visualization of DI-FT-ICR-MS data using van Krevelen diagrams (vK) (O:C vs. H:C ratios of the compounds assigned with an elemental formula) has been commonly used for analyzing complex organic matrices (Kim et al. 2003; Schmitt-Kopplin et al. 2012; van Krevelen 1950). vK diagrams provide information on the main chemical reactions such as demethylation, methylation, hydrogenation, condensation, hydration, reduction, or oxidation occurring in the samples (Kim et al. 2003). Additionally, plotting O:C vs. H:C ratios of the assigned compounds to elemental formulas can also provide an approximation of the main compound categories of the analyzed samples (e.g., lipids, protein, carbohydrates, etc.) (Kim et al. 2003; Minor et al. 2014; Sleighter and Hatcher 2007). However, several compounds in the environment can be transformed or degraded, thus changing their original O:C and H:C ratios (Rivas-Ubach et al. 2018b). Consequently, while this classification can still provide a general idea of the organic compound compositions in aerosol particles, any compound classification based on stoichiometric constraints should be used with caution and multidimensional stoichiometric compound classification approaches could be considered (Rivas-Ubach et al. 2018b). C number versus mass graph (CvM) is also commonly used to represent DI-FT-ICR-MS data in comparative studies providing important information on oxidation processes or molecular weight shifts (Reemtsma 2009). DI-FT-ICR-MS represents thus a useful tool to obtain high-resolution metabolomic fingerprints and to gain a better comprehension of potential chemical transformations of samples.

For more details in metabolomics technologies and their applications, several review articles with special focus on metabolomics technologies have been published (Rochford 2005; Fukusaki and Kobayashi 2005; Shulaev 2006; Lindon et al. 2007; Saito and Matsuda 2010; Lei et al. 2011; Zhang et al. 2012; White et al. 2017; Rivas-Ubach et al. 2018b; Azad and Shulaev 2018).

Testing metabolomics techniques in aerosol particle samples

The techniques to sample and characterize the gas phase of atmo-metabolomes have been already described elsewhere (Smith and Španěl 2011; Tholl et al. 2006). The main aim of this article is to describe a step-by-step method for sampling and characterizing the particle phase of the atmo-ecometabolomes through distinct

mass spectrometry techniques. We collected aerosol particles without size cutoff during two distinct seasons (spring and summer) in an ecosystem with low biological activity by following a simple methodology. Unlike traditional aerosol particle studies concerning atmospheric chemistry, size cutoff of particles was avoided in this method because aeresol particles have a broad range of sizes and size exclusion filtering methods employed in atmospheric studies will not yield informative results for ecology studies. This sampling allowed testing how our method performs in ecosystems with low biological activity by detecting significant changes in the composition of aerosols between seasons. Additionally, we further validate the robustness of our method by analyzing aerosol samples collected during three different peak periods of pollination within the same season in a more biologically active area.

In this manuscript, we describe a suitable protocol to sample aerosols, extract metabolites, and analyze them with (i) LC-MS, (ii) GC-MS, and (iii) DI-FT-ICR-MS. Data from each analytical instrument was thus analyzed following common statistical approaches for ecometabolomics and chemical characterization studies. The methods for aerosol particle sampling and metabolite extraction and the main results from the collected aerosols are discussed. The application of atmo-ecometabolomics techniques in natural ecosystems represents a novel approach to shed light on the understanding of the link between metabolic composition of aerosols and the ecosystem structure and function. This novel method applied in ecological sciences allows understanding of scientific questions related to ecosystem stress, the phyllosphere, ecological stoichiometry, ecosystem phenology, or global change.

Experimental details

In order to optimize the protocol for aerosol characterization, aerosol samples from different locations were collected and analyzed on multiple mass spectrometry platforms. Figure 3 summarizes the procedures implemented in each aerosol sampling campaign, detailing the sampling period, analytical instruments, the software used for generating the metabolomic datasets, and the tables/figures displaying the main results from the aerosol datasets.

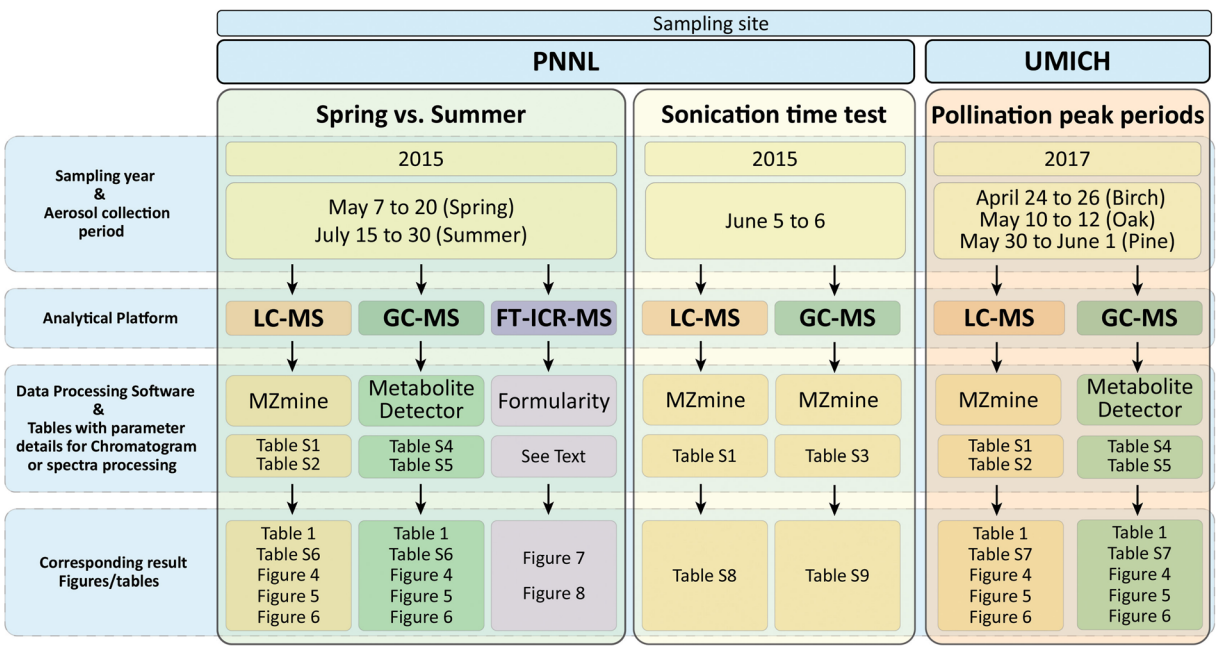

Fig. 3 Sampling year, collection period, and analytical instruments and software used for generating the datasets from each of the sampling sites and campaigns. Table numbers with the parameters used for dataset generation for each of the software are indicated. Table and figure numbers containing the main results derived from aerosols analyzed from each sampling campaign and analytical technique are also indicated

Study sites

The proposed method is demonstrated with the aerosol particle samples collected from two distinct seasons (spring and summer) in 2015 from the Pacific Northwest National Laboratory (PNNL) located in the north side of the city of Richland (WA, USA) (46° 34′ N, 119° 28′ W). Additional samples were collected in 2017 at the University of Michigan (UMICH; Ann Arbor, MI, USA) to validate further our methodology by characterizing aerosol particle signatures from different tree pollination peak periods sampled within a single spring season. UMICH is located in the northeast side of the city of Ann Arbor (MI, USA) (42° 29′ N, 83° 70′ W).

PNNL site (spring vs. summer) PNNL nearby landscape is represented by an extensive desert covered mainly by steppes, shrubs, and herbaceous species with *Purshia tridentate*, *Ericameria nauseosa*, *Grayia spinose*, *Chrysothamnus viscidiflorus*, *Artemisia tripartita*, *Salsola tragus*, *Sarcobatus vermiculatous*, and *Tamarix romosissima* as some of the most represented plant species. PNNL campus is covered by grass with planted non-autochthonous tree species such as *Platanus* sp. The population of the surrounding metropolitan area is about 250,000, and the economy and land use are mainly dominated by diverse crops and the nuclear reservation of Hanford. The climate at PNNL site is semi-arid desert with an averaged yearly rainfall ranging between 180 and 220 mm. Averaged maximum annual temperatures are around 32 °C, with sporadic peaks reaching 42–45 °C. Averaged minimum annual temperatures are around − 2 °C with sporadic temperatures reaching − 20 °C.

UMICH site (pollination peak periods) The surrounding areas are represented by mosaic landscapes mainly covered by urban, agricultural, and extended forested areas dominated by several species of maple trees (*Acer* sp.), oak (*Quercus* sp.), honey locust (*Gleditsia triacanthos*), *Betula alleghaniensis*, *Fagus grandiflora*, *Populus tremuloides*, and some populations of *Pinus strobus*. Ann Arbor is part of the metropolitan area of Detroit with a population of over 4 million. The climate is humid continental with strong influence from the Great Lakes with a mean annual precipitation of 950 mm per year. Average maximum annual temperature is around 28 °C, with peaks reaching up to 41 °C and the average minimum annual temperature is 4.7 °C.

Aerosol particle sampling

For the particle phase aerosol collection, we designed a portable and simple system that allows simultaneous collection of multiple aerosol particle samples (Fig. S1). For aerosol particle collection, we used high-purity quartz filters (Whatman QM-A 37 mm, Whatman International Ltd., Maidstone, UK) precombusted for 5 h at 450 °C for impurity elimination (Schmitt-Kopplin et al. 2012). Each sample was collected by placing a precombusted quartz filter into a filter cassette. Flexible PVC tubing of 0.6 cm diameter was used to connect the pump with the filter cassettes. For each site (PNNL and UMICH), the pump operated for 18 consecutive hours (09:00 a.m. to 03:00 a.m. of the following day) at a flow rate of 30 L per minute at the sampling point (filter). Filters were manually replaced daily at ~ 08:30 a.m. during the sampling periods. Filter samples were thus stored at − 80 °C until metabolite extraction.

PNNL site sampling (spring vs. summer) We collected aerosols for 14 consecutive days in spring 2015 (May 7 to May 20; spring samples) and for 16 consecutive days in summer 2015 (July 15 to July 30; summer samples) (Fig. 3). For each day, two samples were collected simultaneously at an 8-m tower at PNNL campus. According to the US National Weather Service at the Pasco airport (KPSC), daily averaged temperatures ranged from 11 to 21 °C (averaged maximum of 14–29 °C) for the spring aerosol collection period (May). Daily averaged temperatures for the summer aerosol collection period (July) ranged between 19 and 29 °C (averaged maximum of 28–40 °C). Humidity ranged from 49 to 78% (daily average) (averaged maximum of 72–100%) for the spring sampling period and from 35 to 50% (daily average) (averaged maximum of 57–86%) for the summer sampling period. Total precipitation for May and July sampling periods was 28.2 mm and 0 mm, respectively.

UMICH site sampling (pollination peak periods) Additional aerosol samples collected at UMICH campus were collected from 24th to 26th April (birch pollination peak), from 10th to 12th May (oak pollination peak), and from 30th May to 1st June 2017 (pine pollination peak) (Fig. 3). One filter was collected daily during three consecutive days for each pollination peak period from a rooftop location. According to weather conditions reported by the US National Weather

Service at the local airport (KARB), temperature averages were 15.5 °C, 12.2 °C, and 15 °C for sampling periods corresponding to the birch, oak, and pine pollination peaks, respectively. Average humidity ranged between 55–74%, 53–66%, and 59–74% for the sampling periods corresponding to the birch, oak, and pine pollination peaks, respectively. Accumulated recorded precipitation was 3.3 mm, 3.3 mm, and 0.3 mm for the sampling periods corresponding to the birch, oak, and pine pollination peaks, respectively.

Sampling for sonication time test (PNNL site) The extraction of compounds from the filters was sonication-based. Thus, we collected additional aerosol samples in spring 2015 at the PNNL campus for different sonication times testing in GC-MS and LC-MS platforms. Sonication test was not performed in DI-FT-ICR-MS since peak intensity measured with such a method is only semi-quantitative (Kujawinski 2002; Liu and Kujawinski 2015).

For two consecutive days, aerosols were simultaneously collected in three independent filters (hereafter test filters) at a flow rate of 30 L per minute at the sampling point. The pump operated daily from 09:00 a.m. to 03:00 a.m. of the following day (18 h × 2 days = 36 h per filter). A total of six rounds of test filters were collected (6 rounds × 3 filters = 18 filters in total). Sampling of the test filters was performed from June 5 to June 16 (12 consecutive days) (Fig. 3). Test filters were stored at − 80 °C until extraction (sonication test details are shown in the supplementary text of the supporting information).

Metabolite extraction for mass spectrometry analysis

Three different tube sets were used, sets of tubes A, B, and C (Fig. 4). The number of tubes in each set was the same as the number of samples collected. Set A (8-mL glass tubes) was used to perform the extractions. Set B (15-mL polypropylene centrifuge tubes) was to keep the extracts. Set C tubes (2-mL HPLC glass vials) were used to keep the concentrated extract. The glass tubes were combusted at 450 °C for 5 h before use. Each sample (filters) was rolled (Fig. 4, 4.1) and inserted into the corresponding tube of set A (Fig. 4, 4.2). Subsequently, each tube from set A received 5 mL of H_2O/MeOH (20:80) (this volume of extract may vary depending on the used filter diameter and the diameter of the set A tubes; filters have to be totally covered by the extract solution) (Fig. 4, 4.3). The samples were then sonicated at 24 °C for 10 min (Fig. 4, 4.4) (10 min was the established time after performing the sonication time tests; see Supporting Information). Four milliliters of the extract from each tube of set A was transferred to the corresponding tube of set B (Fig. 4, 4.4.1). Two extractions were performed to each sample by repeating the same procedures but adding 4 mL of H_2O/MeOH (20:80) instead of 5 mL as fresh extract for the second extraction. The resulting extracts from the second extraction were then combined with the first extracts (in tubes of set B) (Fig. 4, 4.5.1). All extracts (in set B tubes) were evaporated using an ultra-high purity nitrogen evaporator (Fig. 4, 4.6). Subsequently, 1 mL of fresh solvent (H_2O/MeOH (20:80)) was added to each set B tube and tubes vortexed for 30 s for a proper resuspension of the dried extracts (Fig. 4, 4.7). All concentrated extracts (in set B tubes) were centrifuged at 4500×*g* for 5 min (Fig. 4, 4.8), and supernatants were transferred into the corresponding tubes of set C (HPLC vials) (Fig. 4, 4.9). Extracts were stored at − 80 °C until MS analyses (Fig. 4, 4.10). Experimental blanks (not used combusted filters) were also extracted and analyzed with each of the MS platforms.

Preliminary aerosol particle samples collected at PNNL for sonication tests were analyzed by LC-MS and GC-MS. LC-MS and GC-MS provide accurately relatively quantitative data making them suitable for this sonication testing. Spring and summer samples collected at PNNL campus were analyzed by LC-MS, GC-MS, and DI-FT-ICR-MS (Figs. 3 and 4.11). Additional samples collected at the UMICH campus to contrast different tree pollination peak periods were analyzed exclusively by LC-MS and GC-MS (Figs. 3 and 4.11). For DI-FT-ICR-MS and LC-MS platforms, extracts (~ 200 μL) were directly used for analyses (Fig. 4, 4.12). GC-MS required a pre-treatment of the samples prior to the instrumental analyses.

For metabolite derivatization for GC-MS analyses, we followed a protocol described elsewhere (Kim et al. 2015). Briefly, for each sample, 500 μL of extract from the set of tubes C (Fig. 4, 4.10) was transferred into a clean HPLC vial and dried down in a vacuum evaporator centrifuge. Then, each sample received 20 μL of methoxyamine in pyridine solution (30 mg/mL). Samples were vortexed for 30 s and subsequently incubated for 90 min at 37 °C and 1000 rpm in a Thermomixer (Eppendorf AG, Hamburg, Germany) to protect carbonyl groups. After the first incubation, all vials were

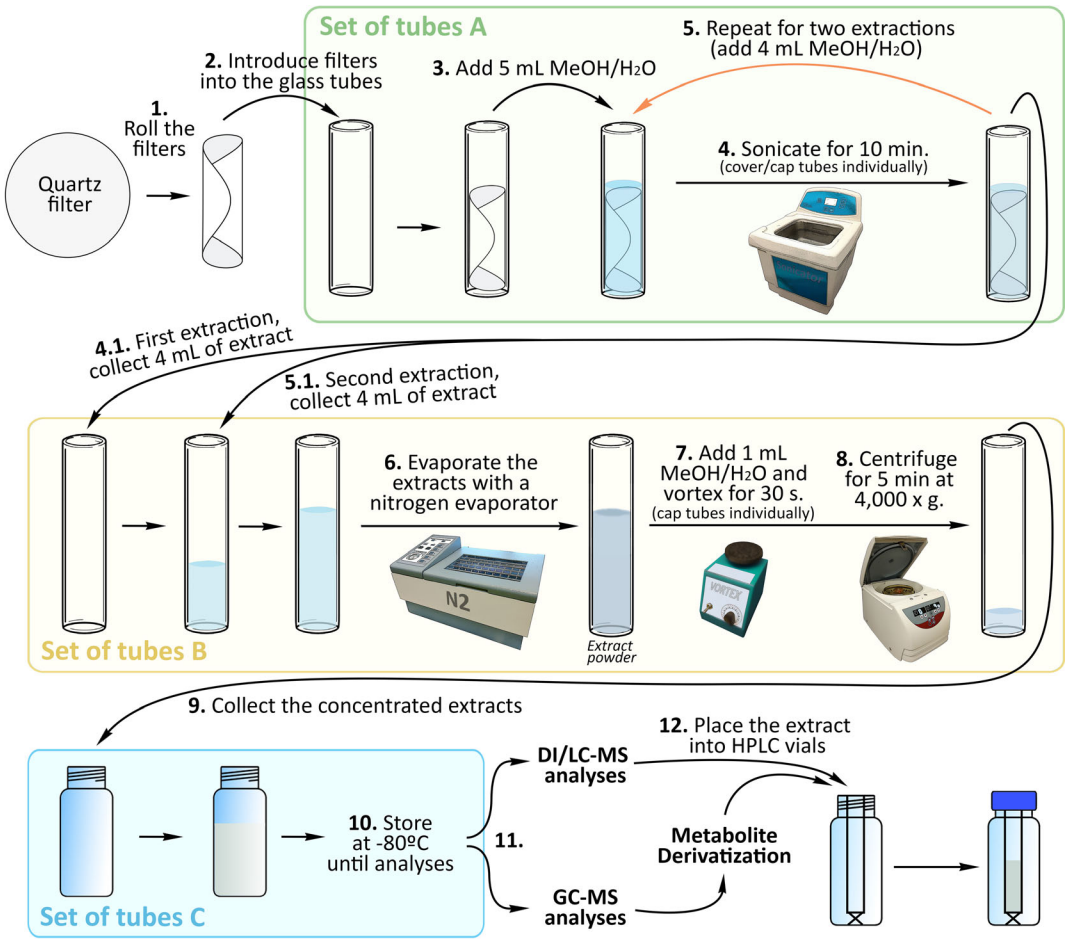

Fig. 4 Experimental procedures performed on filters to obtain the metabolomic extracts from aerosol samples for subsequent analyses with distinct mass spectrometry techniques

centrifuged for 15 s and 80 μL of N-methyl-N-(trimethylsilyl)trifluoroacetamide (MSTFA) with 1% trimethylchlorosilane (TMCS) was added to each sample. All vials were then vortexed for 10 s and incubated for 30 min at 37 °C and 1000 rpm to derivatize caboxyl, hydroxyl, and amine functional groups. After the second incubation, all samples were centrifuged for 5 min at 4500×*g* and derivatized supernatants were transferred into clean glass vials with 200 μL inserts by using glass Pasteur pipettes. All vials were finally tightened with caps with septum.

LC-MS analyses and chromatogram processing

A Vanquish ultra-high pressure liquid chromatographer (UHPLC) coupled to a high-resolution mass spectrometer LTQ Orbitrap Velos (HRMS) equipped with a heated electrospray ionization (HESI) source (Thermo Fisher Scientific, Waltham, MA, USA) was used to obtain the LC-MS chromatograms of aerosol samples. LC was performed with a reversed-phase C18 Hypersil gold column (150 × 2.1 mm, 3 μ particle size; Thermo Scientific, Waltham, MA, USA) operating at 30 °C and at a constant flow rate of 0.3 mL per minute. Chromatography mobile phases consisted of 0.1% formic acid in water (A) and acetonitrile:0.1% formic acid in water (90:10) (B). The injection volume was set at 5 μL. The elution gradient initiated at 90% A (10% B) and hold for 5 min, then the gradient ramped linearly to 10% A (90% B) during the next 15 min. Those conditions were maintained for 2 min before the initial mobile phase proportions (90% A; 10% B) were recovered over the next 2 min. The column was washed and stabilized for an additional 11 min at the initial chromatographic conditions (90% A; 10% B). All samples were analyzed in positive (+) and negative (−) electrospray ionization

modes. The HRMS operated in FTMS (Fourier transform mass spectrometry) and full-scan mode at high resolution (60,000) and acquired in the mass range of 50–1000 *m*/*z*. An experimental blank was injected every 15 samples. Blanks were used to determine the chromatographic background during the dataset filtering. A mixture of standards was injected every 25 samples to test for mass accuracy and sensitivity of the instrument. Two injections of fresh H_2O/MeOH (20:80) were analyzed right after the standard mixture for column washing purposes and to avoid any potential compound carry over.

MZmine 2.17 (Pluskal et al. 2010) was used to process the LC-MS RAW files. Chromatograms obtained in positive and negative modes were treated separately. Chromatograms were baseline corrected, deconvoluted, aligned, and "gap-filled," and metabolite identifications were assigned. Datasets were thus exported to CSV format files. The parameters used for generating the LC-MS datasets are detailed in Table S1. The same MZmine parameters were used to generate all datasets, spring vs. summer LC-MS dataset (PNNL) and pollination peak period LC-MS dataset (UMICH).

Metabolite assignation for LC-MS files was performed by matching the exact mass and RT of the detected features with the corresponding values of our metabolite library including more than 600 metabolites present in organisms commonly present in plants. According to Sumner et al. (2007), our LC-MS metabolite assignment is putative since it was based on the measured exact mass of metabolic features and their RT. However, the use of both RT and the high MS resolution achieved with Orbitrap technology reduces substantially the number of false-positive assignations. For more detailed information in relation to metabolite library matching, see Rivas-Ubach et al. (2016c). Metabolite matching information (RT and *m*/*z* values) for LC-MS files are detailed in Table S2.

GC-MS analyses and chromatogram processing

Derivatized samples were analyzed by an Agilent GC 7890A coupled with a MSD 5975C mass spectrometer (Agilent Technologies, Santa Clara, CA). An HP-5MS column (30 m × 0.25 mm × 0.25 μm; Agilent Technologies) was used for the GC. The injection port temperature was held at 250 °C. The injection mode was splitless. The oven with the column was maintained for 1 min at 60 °C and then ramped at 10 °C per minute to 325 °C (26.5 min ramp) and hold for 10 min. A mixture of fatty acid methyl esters (FAMEs; C8–C28) for retention index calculation (RI) was analyzed prior to sample injections. Experimental blanks were analyzed every 15 samples to determine the chromatographic background.

MZmine 2.17 (Pluskal et al. 2010) was exclusively used to generate the metabolomic fingerprint datasets for the test filters collected at PNNL to be consistent with the LC-MS datasets for the sonication test (Fig. 3). Parameters implemented to generate the datasets from the test filters with MZmine are described in Table S3. Metabolite Detector 2.5 (Hiller et al. 2009) was used to generate the GC-MS datasets for the ambient samples collected at PNNL (spring vs. summer) and UMICH (pollination peak periods) sites (Fig. 3). Agilent ".D" files directly obtained from the MS were converted to "netCDF" format with Agilent Chemstation software. Subsequently, Metabolite Detector converted "netCDF" files to "bin" files and chromatograms were thus deconvoluted, aligned according to the retention indices (RI), and metabolic features were assigned to metabolites. Metabolites were identified by matching measured mass spectra to a PNNL improved version of FiehnLib (Kind et al. 2009), containing validated RIs and mass spectra for over 850 metabolites. Metabolite matching probability threshold was set at 0.65. Datasets were finally exported to CSV format files. Metabolite identifications were manually validated by matching measured mass spectra with mass spectra from the NIST14 GC-MS library to avoid false identifications and reduce any alleged deconvolution error during data processing. Table S4 includes all the parameters used in the metabolite detector to generate the GC-MS datasets from the aerosols collected at PNNL and UMICH sites. Table S5 shows metabolite matching information for GC-MS.

DI-FT-ICR-MS analyses and spectra processing

DI-FT-ICR-MS analyses were performed exclusively on the spring and summer samples collected at PNNL. Extracts of aerosols were analyzed on a 12 Tesla Bruker SolariX FT-ICR-MS (Bruker Daltonics Inc., Billerica, MA, USA). At a flow rate of 3.0 μL/min, aerosol extracts were directly infused into the FT-ICR-MS using an Agilent 1200 series pump (Agilent Technologies, Santa Clara, CA, USA). Compounds were ionizated by a standard Bruker electrospray (ESI) operating in negative mode and

equipped with a fused silica tube (30 μm i.d.). Ion accumulation time was optimized for each sample (0.1 s). FT-ICR-MS operated at a resolution of 4 MWord, equivalent to a resolving power of 400,000 ($m/\Delta m_{50\%}$ at m/z 400). Experimental conditions were as follows: needle voltage was set at + 4.4 kV, Q1 set to 50 m/z, and the heated resistively coated glass capillary was maintained at 180 °C. Solvent blanks were run every 10 samples.

A total of 144 individual scans from 100 to 1100 m/z were averaged to produce the final DI-FT-ICR-MS spectrum for each sample. The instrument achieved mass measurement accuracy with < 1 ppm error after external calibration with a standard mixture. All detected ions in the spectra were singly charged as confirmed by the 1.0034 Da spacing between isotopic forms of the same compound (i.e., between $^{12}C_n$ and $^{12}C_{n-1}{}^{-13}C_1$). Each spectrum raw file was converted to a list of m/z values with DataAnalysis software (BrukerDaltonik version 4.2) with the following settings: an absolute intensity threshold of 100 and a signal to noise (S/N) of 7, which is above the minimum detection limit for FT-ICR-MS for natural organic matter (Riedel and Dittmar 2014). The exported peak lists were then internally calibrated using an organic matter homologous series separated by 14 Da ($-CH_2$ groups) using the Formularity software (Tolić et al. 2017). After internal calibration, the mass error was < 0.5 ppm. The calibrated peak list was then used to assign with elemental formulas (C, H, O, N, S, and P) using a module within Formularity, based on the compound identification algorithm described elsewhere (Kujawinski and Behn 2006).

Dataset filtering

A total of seven distinct datasets were generated: three contrasting spring vs. summer aerosols (LC-MS, GC-MS, and DI-FT-ICR-MS; collected at PNNL), two comparing three different pollination peak periods (LC-MS and GC-MS; collected at UMICH), and two for sonication tests (LC-MS and GC-MS; collected at PNNL). LC-MS and GC-MS datasets were separately filtered through five main steps:

(1) All zero values of the dataset were replaced for missing data (NA).

(2) For each variable (metabolite feature), outlier values were determined for each season separately and replaced for NA. Outliers were defined as:

$$\text{Upper outliers} \rightarrow \text{value} > Q75 + 2.5 \times IQR$$

$$\text{Lower outliers} \rightarrow \text{value} < Q25 - 2.5 \times IQR$$

where Q75 represents the third quartile, Q25 represents the first quartile, and IQR is the interquartile range (IQR = Q75–Q25).

(3) Variables (metabolite features) with less than 50% of data within all factor levels of the dataset (spring and summer for datasets from PNNL aerosol particles; birch, oak, and pine pollination peak periods from UMICH aerosols) were removed.

(4) Variables (metabolite features) with signal to noise lower than 6 were removed from the dataset. Noise level was determined by the experimental blanks run during the analytical sequence.

(5) To avoid statistical artifacts in multivariate analyses due sample variability, the values for each variable (metabolite feature) and sample were scaled by the total intensity of its chromatogram (LC-MS and GC-MS). This scaling allows comparing the metabolic function between groups of samples independently of the amount of organic matter.

It should be noted that DI-FT-ICR-MS data should go through rigorous data quality check before further interpretation, to avoid false discovery. For DI-FT-ICR-MS dataset (PNNL; spring vs. summer), mass features detected from solvent blanks were removed from all the samples. Additionally, mass features observed in less than 70% of replicates within all factor levels were also removed to avoid consideration of noise peaks. Finally, for robust data interpretation based on elemental formulas, we only used the assigned formulas with less than 0.3 ppm of error to be conservative, as cutoff values up to 0.5 ppm of error are typically used (Osterholz et al. 2016; Rivas-Ubach et al. 2018b).

After filtering all databases, our spring vs. summer databases were finally composed by 1333, 148, and 3567 metabolomic features detected by LC-MS, GC-MS, and DI-FT-ICR-MS, respectively. For LC-MS and GC-MS datasets, a total of 18 and 15 metabolites were identified, respectively. The pollination peak periods

databases were finally composed of 7832 and 221 metabolomic features detected by LC-MS and GC-MS, respectively, with a total of 45 and 20 identified metabolites, respectively.

Statistical analyses

While the values obtained from the deconvoluted peaks in LC-MS and GC-MS do not represent an absolute concentration (e.g., mg of metabolite per weight of sample), they represent the relative abundance and are suitable for metabolomic comparative analyses as proven in previous studies (Gargallo-Garriga et al. 2015; Lee and Fiehn 2013; Leiss et al. 2013; Mari et al. 2013; Rivas-Ubach et al. 2014, 2016a, 2019). We used then the term *relative abundance* along the article as the relative concentration of metabolic features or metabolites. DI-FT-ICR-MS data is not directly quantifiable (Wozniak et al. 2008), and although not as robust as GC-MS or LC-MS techniques, using the feature intensity can still provide a valuable proxy of their relative abundance (Kellerman et al. 2014; Spencer et al. 2015).

All LC-MS and GC-MS datasets, independently if generated from PNNL (spring vs. summer) or UMICH (pollination peak periods) aerosols, were submitted to the same statistical analyses. First, the aerosol metabolome fingerprints obtained from LC-MS and GC-MS were submitted to PERMANOVAs using the Euclidean distance to test for overall metabolomic differences between aerosols from distinct seasons (PNNL site), and between different pollination peak periods (UMICH site) (Table 1). The permutations for the test were set at 10,000.

LC-MS and GC-MS metabolomic fingerprints were also subjected to principal component analysis (PCA) to graphically represent the natural variability among the analyzed samples reduced in two dimensions (Principal Component (PC) 1 vs. PC2) (Kim et al. 2010; van den Berg et al. 2006) (Fig. 5).

A heat map was plotted for each entire LC-MS and GC-MS metabolomic fingerprints from the aerosols collected at PNNL and UMICH sites (Fig. 6). Additional heat maps were plotted only for the identified features from the spring vs. summer and for the pollination peak period datasets (Fig. 7). All the identified LC-MS and

Table 1 PERMANOVAs of the atmo-metabolome fingerprints generated by LC-MS and GC-MS instruments for overall metabolome comparison between seasons

Spring vs. summer dataset (aerosols collected at PNNL)					
LC-MS					
		Sum of squares	Mean square	*F*	*P*
Season	1	0.12	0.12	13.8	< 0.0001
Residuals	28	0.24	0.009		
Total	29	0.36			
GC-MS					
		Sum of squares	Mean square	*F*	*P*
Season	1	0.027	0.027	9.69	< 0.0001
Residuals	28	0.079	0.0028		
Total	29	0.11			
Pollination peak periods dataset (aerosols collected at UMICH)					
LC-MS					
		Sum of squares	Mean square	*F*	*P*
Season	2	0.12	0.06	9.54	0.0039
Residuals	6	0.037	0.006		
Total	8	0.16			
GC-MS					
		Sum of squares	Mean square	*F*	*P*
Season	2	0.14	0.07	7.63	0.0041
Residuals	6	0.05	0.009		
Total	8	0.19			

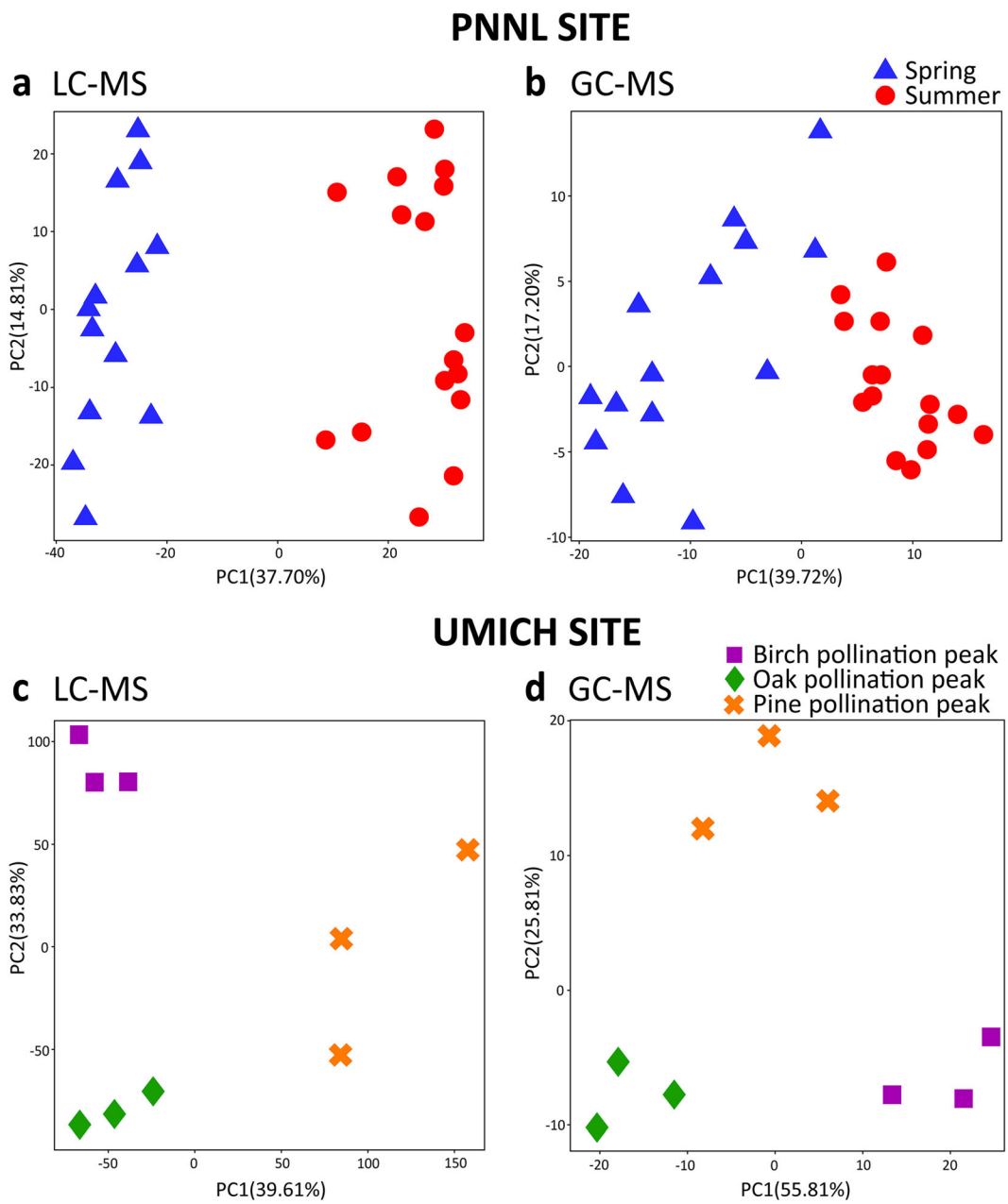

Fig. 5 Case plots of the principal component (PC) 1 versus the PC2 of the PCAs conducted on the aerosol metabolomic fingerprints collected at PNNL site contrasting spring vs. summer aerosols (**a** for LC-MS; **b** for GC-MS) and collected at UMICH site contrasting different pollination peak periods (**c** for LC-MS; **d** for GC-MS). Each filter sample corresponds to a point for each of the case plots. For PNNL site (spring vs. summer), spring aerosol samples are represented by blue triangles and summer aerosol samples are represented by red circles. For UMICH site (pollination peak periods), samples collected during birch, oak, and pine pollination peaks are represented by violet squares, green diamonds, and orange crosses, respectively

GC-MS features for the spring vs. summer dataset (PNNL) were submitted to Student *t* test with season as the categorical factor to assess for statistical significance between spring and summer seasons (Table S6). One-way ANOVAs followed by Tukey HSD post hoc tests were performed to asses for significant differences between pollination peak period samples (birch vs. oak vs. pine). (Table S7).

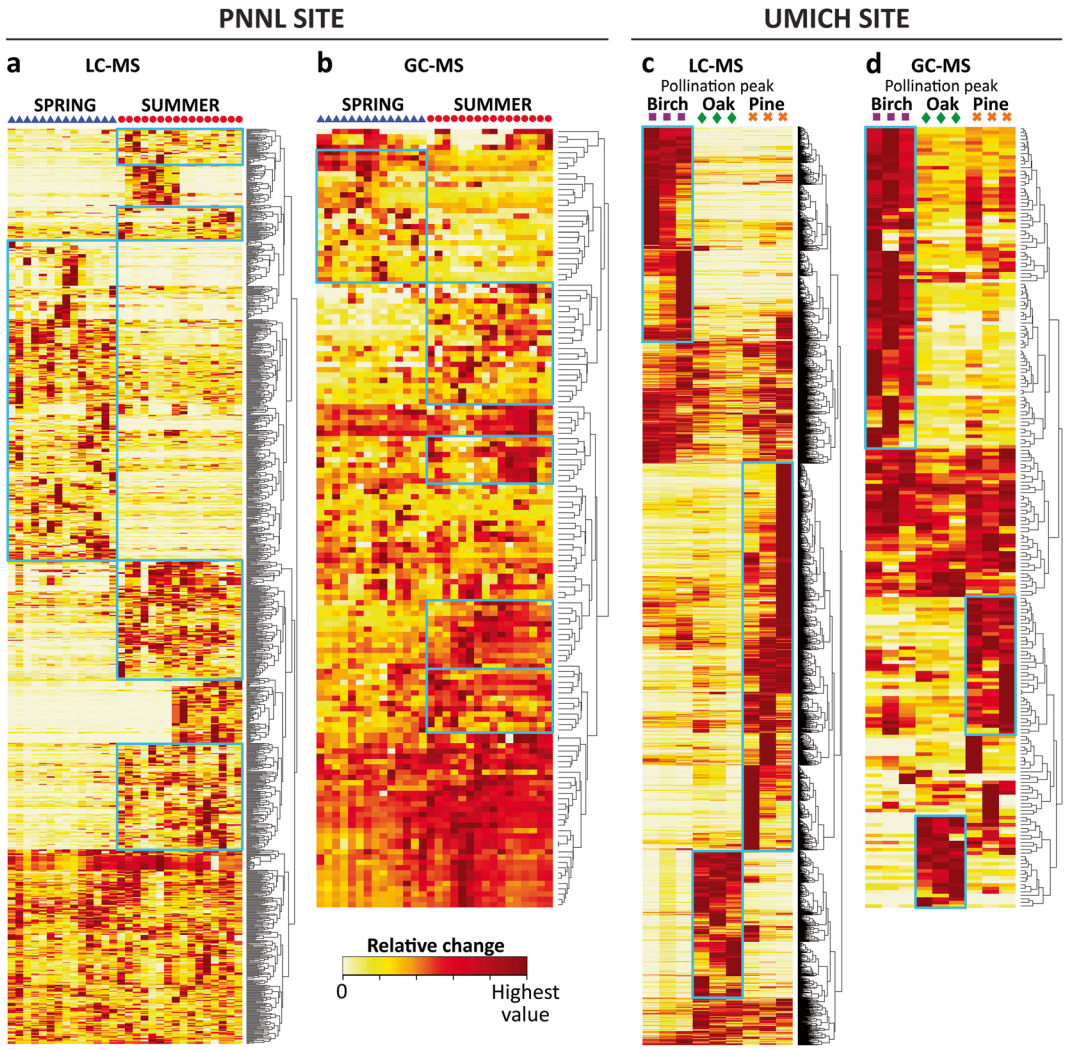

Fig. 6 Heatmaps of the metabolomic fingerprints obtained from LC-MS and GC-MS for PNNL metabolomics datasets (spring vs. summer) (a and b) and for the UMICH metabolomics datasets (pollination peak periods) (c and d). Cluster dendrogram for the variables is shown for each heatmap. Blue rectangles point cluster of variables for the specific group of samples having higher overall relative abundance values. The color gradient represents the relative abundance of the metabolomic feature between samples. Color red represents the highest relative abundance

For the DI-FT-ICR-MS dataset generated from the PNNL aerosols (spring vs. summer), we counted the proportions of formula classes (CHNO, CHO, CHOS, CHNOS, CHNOP, CHOSP, CHNOSP, and CHOP) that were assigned to each sample. Proportions were then transformed using *arcsin (rootsquare)* and subsequently submitted to Student *t* tests with season (spring and summer) as the categorical factor to assess for statistical significance (Fig. 8). For each feature detected by DI-FT-ICR-MS, the intensity difference between spring and summer was calculated and used in the C number of each feature vs. *m*/*z* (CvM) (Fig. 9a) and in the O:C ratio vs. *m*/*z* (Fig. 9b) plots. Additionally, a Student *t* test was performed on the O:C ratios of the detected features with season as the categorical factor to determine whether the oxidation status of the molecular compounds statistically changed significantly between spring and summer aerosols (Fig. 9c).

All statistical analyses were performed with R (R Core Team 2013). PERMANOVAs were conducted with the function *adonis* from the "vegan" package (Oksanen et al. 2013). PCAs were plotted using the *PCA* function from "FactoMineR" package (Husson et al. 2016) with the missing data from the dataset

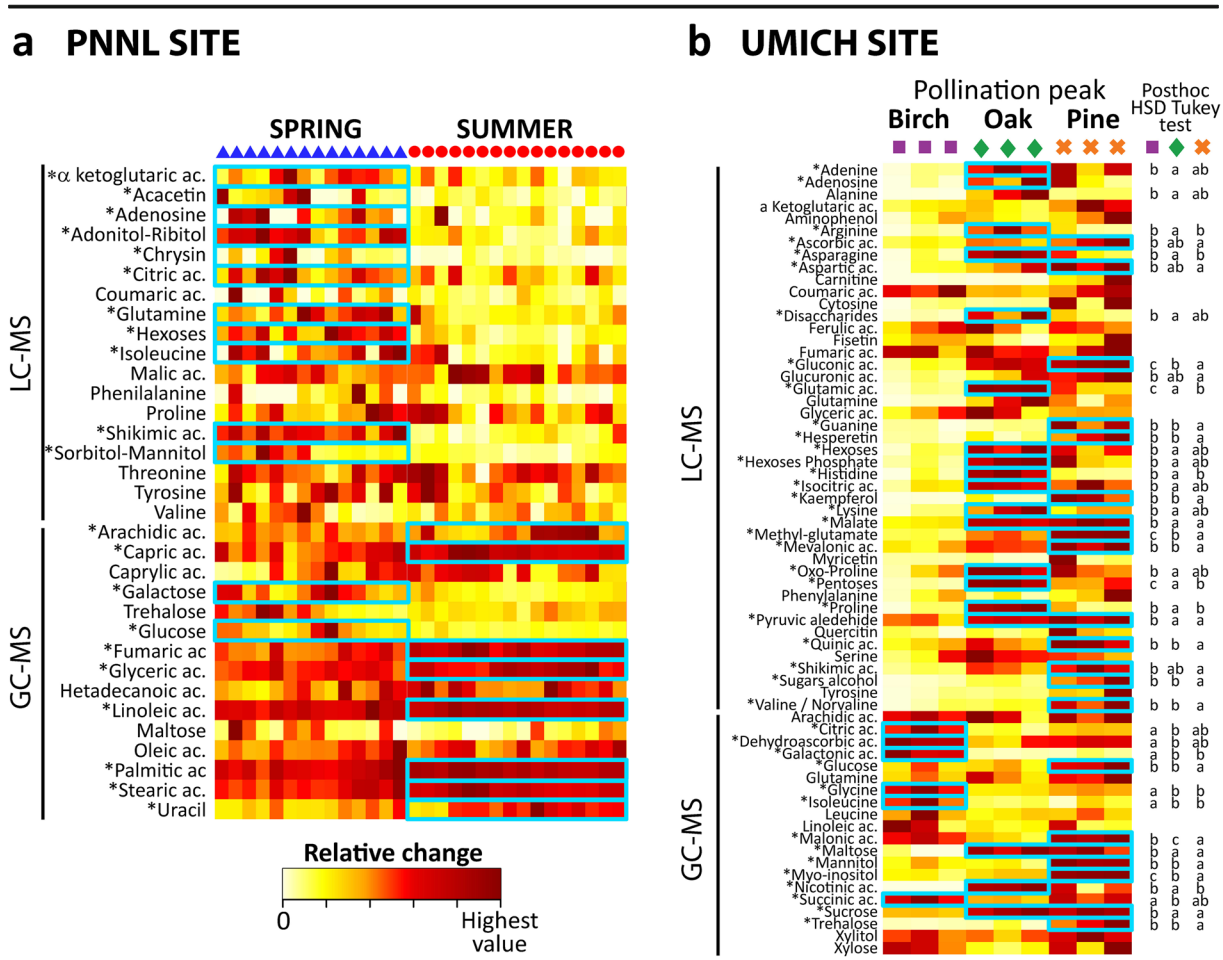

Fig. 7 Heat maps of the identified metabolomic features from LC-MS and GC-MS for the PNNL metabolomics dataset (spring vs. summer) (**a**) and for the UMICH metabolomics dataset (pollination peak period dataset) (**b**). For each site, heatmaps of the identified features were plotted using LC-MS and GC-MS data combined. The color gradient represents the relative abundance of the metabolomic feature between samples. Color red represents the highest relative abundance. Variables that presented significance difference ($P < 0.05$) after *t* test (**a**) or after one-way ANOVA (**b**) are marked by an asterisk. Blue rectangles indicate the group of samples having higher average relative abundance value for each specific variable that presented statistical significance. For the pollination peak period (**b**), different letters next to the heatmap indicate significant differences between groups after Tukey HSD post hoc test

imputed using *imputePCA* function from package "missMDA" (Husson and Josse 2016). The data matrix was scaled before the PCA by setting SCALE as TRUE from the *PCA* function. Heat maps were plotted using the function *heatmap.2* from "gplots" package (Warnes et al. 2016). Student *t* tests and one-way ANOVAs were performed with the functions *t.test* and *aov*, respectively, included in the "stats" package (R Core Team 2013). Tukey HSD post hoc tests were performed with the *HSD.test* function in the "agricolae" package (de Mendiburu 2015). All graphs were first plotted in R and subsequently treated with Adobe Illustrator CS6. Analyses and results from the different sonication times during metabolite extraction are detailed in the supporting information (Supplementary Text).

Results and discussion

Aerosol sampling

The optimal flow rate for aerosol particle collection is important, as low flow rate may not collect enough aerosol particles and excessive flow rate may damage the filters. After coupling the tubing with the pump and connected two filters, we achieved constant flow rates of

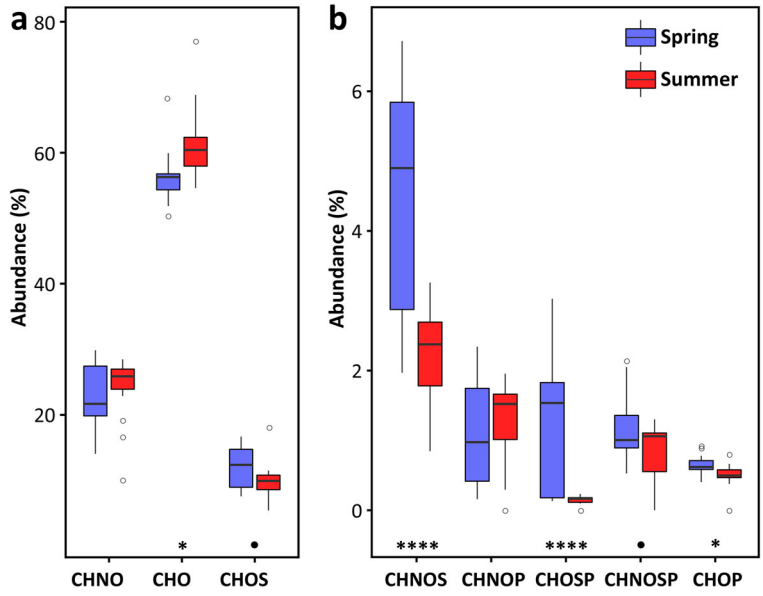

Fig. 8 Box plots for the proportion (%) of different formula classes (CHNO, CHO, and CHOS (**a**) and CHNOS, CHNOP, CHOSP, and CHOP (**b**)) for PNNL site DI-FT-ICR-MS dataset (spring vs. summer). Boxes show the median value for each formula class and season. Open dots correspond to extreme values. Asterisks indicate statistical significance between aerosols collected spring and summer for each formula class ($P < 0.05$ (*); $P < 0.0001$ (****)). Black dots indicate marginal significance between spring and summer aerosols for each formula class ($P < 0.1$)

30 L/min at the aerosol sampling point. However, flow rates of ~50 L/min at the aerosol collection point performed properly in 37 mm quartz filters without collapsing. Different filter materials, such as polytetrafluoroethylene (PTFE), may present distinct resistances to high flow rates. It is important to note that the internal air friction associated with the length of the tube and the sampling of multiple filters simultaneously can substantially decrease the flow rates at the aerosol sampling point. Larger tube diameters (>0.6 cm) could be used if higher flow rates at the sampling point are necessary, especially when sampling in ecosystems with low biological activity.

For statistical purposes, our simple sampling methodology allows collecting different number of biological replicates at once. Furthermore, aerosol collection can be easily performed at different heights by extending tubing if the pump performance is sufficient as to ensure acceptable flow rates at the aerosol collection point. The experimental design (number of replicates, filter material, length and diameter of tubing) and the pump performance are thus key elements to consider in atmoecometabolomics research.

Aerosol particles collection was performed in two distant locations with distinct environments; PNNL and UMICH. PNNL site is semi-urban area surrounded by diverse croplands and a large desert with relatively low biological activity. For this reason we expected to detect a complex variety of compounds which difficult the interpretation of the data at regional scale. Even so, the results obtained contrasting spring vs. summer aerosols were equally valid as to describe in detail the necessary steps to obtain accurately the atmometabolomes and to test the sensitivity of each mass spectrometry platform (LC-MS, GC-MS, DI-FT-ICR-MS) by assessing their potential for detecting statistically significant changes between the atmoecometabolomes of two contrasted seasons of ecosystems with low biological activity. At the same time, UMICH location provided additional data collected in a more biologically diverse and active ecosystem to further evaluate the sampling and the analytical methodologies.

Metabolite extraction in organic solvents

Organic solvents in combination with water are commonly used for metabolite extractions, allowing a good extraction range of polar, semi-polar, and non-polar metabolites (Kim et al. 2010; Lin et al. 2006; Rivas-Ubach et al. 2013; t'Kindt et al. 2008). Although different solvents recover different matrices of compounds with distinct polarities, water/methanol (20:80) solution has been widely used in metabolomics studies, showing a wide recovery of polar and semi-polar metabolites (t'Kindt et al. 2008). Solvents such as acetonitrile, methanol, and chloroform could leach plastics especially during sonication, as such MS spectra may show contaminant features derived from plastics (Fig. S2). Therefore, the use of silanized glass tubes is advised during

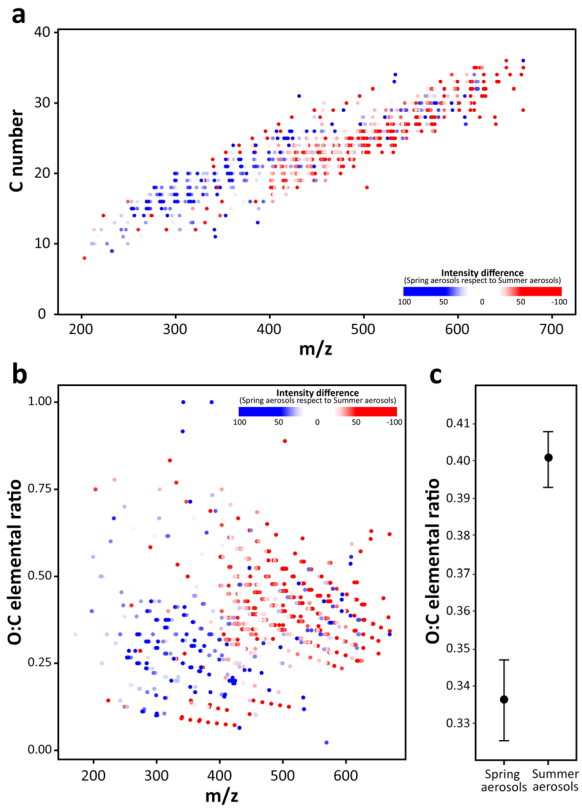

Fig. 9 Carbon number versus *m/z* (CvM) (**a**) and oxygen/carbon ratio versus mass (**b**) diagrams plotted from DI-FT-ICR-MS datasets (PNNL site). Different colors of dots represent the relative intensity of spring relative to summer for each metabolic feature; darker red dots represent higher relative intensity in summer and darker blue dots represent higher relative intensity in spring. Mean (±SE) of oxygen/carbon of the metabolic features detected in aerosols collected in spring and summer (**c**). Statistic-*t* and *P* value are indicated in the graph

sonication (Fig. 4, 4.4). Combusting glassware for 5–6 h at 450 °C or higher is also recommended to prevent organic contaminations. If plastic tubes are used for metabolite extraction, especially during the sonication step (set of tubes A), initial tests are recommended to identify any potential contaminant in the extracts. Due the relatively low metabolite concentration in aerosol samples, each sample was extracted twice in water/methanol (20:80) to increase metabolite recovery from aerosols (Böttcher et al. 2007; Nikiforova et al. 2005; Rivas-Ubach et al. 2013) (Fig. 4, 4.5).

We did not detect overall significant variation in the relative abundance of the observed features between test-filters under different sonication times (Tables S8 and S9) with the water/methanol (20:80) extraction; only days 5th & 6th June showed significant differences between different sonication times ($F = 3.21$, $P = 0.02$) in the GC-MS analyses (Table S9). We thus considered that 10 min of sonication was enough as to extract the metabolites from aerosol samples in water/methanol (20:80) solution.

Filter size is also a crucial factor to consider in atmo-ecometabolomic research, especially during the extraction procedures. Extracts of aerosol particles will be more concentrated as lower is the *filter size/pump flow rate* ratio. Furthermore, smaller filters are easier to handle during extractions and allow higher recovery of the extracts. Quartz filters absorb high volumes of solvent that cannot be easily recovered, even if samples are centrifuged at high rpm. We could recover the 89% of the initial solvent volume with 37 mm diameter filters. Larger filters complicate significantly the extraction of metabolic compounds (more filter handling, larger tubes and volumes of extract are required) and decrease remarkably the extract recovery due to the large absorption of solvent. For studies using FT-ICR-MS, a final dissolved organic carbon (DOC) concentration of 15 mg of C per L in 1:1 methanol/water in samples has been proven to provide suitable signal for detection of compounds (Medeiros et al. 2017). However, higher concentration of C may be required for other mass spectrometry platforms depending on their sensitivity. Analyses by LC-MS and GC-MS techniques separate the analytes through chromatography; therefore, the amount of C analyzed in each scan is diluted and higher amount of DOC allows a better detection of organic compounds and metabolome characterization.

Testing atmo-ecometabolomics contrasting two distinct seasons (PNNL site)

Even collecting aerosols in low biologically productive ecosystems, PERMANOVAs showed overall significant differences between spring and summer aerosols for the fingerprints obtained with both LC-MS and GC-MS platforms (pseudo-$F = 13.8$, $P < 0.0001$, and pseudo-$F = 9.69$, $P < 0.0001$, respectively) (Table 1).

PCAs plotted for both LC-MS and GC-MS fingerprints also showed clear separation between spring and summer aerosol samples (Fig. 5a, b). The PC1 and PC2 of the PCA performed with LC-MS dataset explained the 37.7% and 14.81% of the total variation among samples, respectively. The PC1 and PC2 of the GC-MS PCA explained the 39.72% and 17.2%, respectively, of the total variance. Both PCAs showed similar

explained variability for the PC1 (37.7% for LC-MS and 39.72% for GC-MS) that primarily segregated spring from summer samples. The results from PERMANOVAs and PCAs indicate, indeed, that the metabolomic fingerprints from both seasons are significantly distinct. Even so, it is important to note that each analytical technique (LC-MS and GC-MS) provides different complementary information (Ding et al. 2007; Zhang et al. 2012).

Heat maps of the entire LC-MS and GC-MS metabolomic fingerprints showed large metabolomic variability between aerosol samples within the same season (Fig. 6a, b). Even so, we still identified several clusters of metabolic features (marked in blue) presenting higher relative abundances in spring or summer aerosol samples. Those heat maps are clearly in accordance with the PERMANOVAs and the PCA by showing major metabolic differences between spring and summer aerosols supporting thus our protocol as sensitive enough as to differentiate the overall aerosol composition between two seasons in a complex and low biologically active environment.

Several of the identified metabolites with LC-MS and GC-MS showed statistical significance between seasons after Student *t* tests (Fig. 7a and Table S6). LC-MS analyses showed spring aerosol particles with significantly higher relative abundance ($P < 0.05$) of α-ketoglutaric acid, acacetin, adenosine, adonitol-ribitol, chrysin, citric acid, glutamine, hexoses, isoleucine, shikimic acid, and sorbitol-mannitol (Fig. 7a). The GC-MS dataset indicated that spring aerosols presented significantly higher relative abundances of glucose and galactose ($P < 0.05$) and summer aerosol particles showed significantly higher relative abundances of arachidic acid, capric acid, fumaric acid, glyceric acid, linoleic acid, palmitic acid, stearic acid, and uracil (Fig. 7a). Higher relative abundances of sugars and organic acids related to the Krebs cycle such as citric and alpha-ketoglutaric can be related to plant growth activity (Rivas-Ubach et al. 2012) and atmospheric pollination (Roulston and Cane 2000). Our results from the LC-MS and GC-MS platforms are in agreement with the DI-FT-ICR-MS dataset showing spring aerosols with significantly ($P < 0.05$) higher proportions of CHOP and marginally significant ($P < 0.1$) CHNOSP molecular formulas (Fig. 8). High concentrations of sugars and phosphorus in biomass have been previously associated to higher plant activity (Rivas-Ubach et al. 2012, 2014) although sugars can also play multiple physiological functions in plants such as stress tolerance under drought conditions (Ingram and Bartels 1996; Rivas-Ubach et al. 2014, 2016a). Several fatty acid compounds, such as arachidic acid, capric acid, heptadecanoic acid, linoleic acid, oleic acid, palmitic acid, and stearic acid, have been detected by GC-MS (Fig. 7a). Fatty acids are well represented in cells and can represent up to 20% of dry weight in pollen and other plant material (Roulston and Cane 2000). For example, arachidic acid and linoleic acid, among others, are typically found in pollen (Solberg and Remedios 1980). Our GC-MS analyses indicated that summer aerosols had higher relative abundances of most identified fatty acids (Fig. 7a) coinciding with the agricultural production peak of the area, the most extensive of the Pacific Northwest of USA.

DI-FT-ICR-MS analyses showed that aerosol particles collected in summer aerosols had significantly ($P < 0.05$) higher proportions of CHO features than spring aerosols (Fig. 8). In addition, we generally observed higher relative intensities in high-mass features in summer aerosols with respect to spring aerosols (Fig. 9a). In a CvM plot, at a given carbon number, the increase of nominal mass is contributed by heteroatoms (e.g., N, S, and O). We observed that the higher relative intensities at high-mass features in summer aerosol particles were contributed by heteroatoms as these features appear to possess a similar range of carbon numbers as those observed in spring, which were more abundant at lower molecular mass (see region between dashed lines in Fig. 9a). In addition, Student *t* test on the O:C of the formula-assigned features indicated that summer aerosol particles had significantly higher relative intensities in features with higher O:C (more oxidized compounds) compared to spring aerosols (Fig. 9b, c). This result is in accordance with the larger number of high-molecular weight compounds for a same *C*-number found in aerosol particles collected in summer compared to those from spring (Fig. 9a), suggesting that summer aerosol particles experienced higher oxidation rates. This trend could be related to higher levels of photochemical oxidants associated with hot sunny conditions of the area during summer (Obee and Hay 1997). Additionally, higher relative intensities in compounds over 500 Da were also found in summer (Fig. 9a) suggesting higher rates of aerosol condensation and polymerization. These measurements point to a major challenge in atmo-ecometabolomics research: understanding deeply the atmospheric processing of the initial biogenic emissions.

Atmo-ecometabolomics for differentiate three different pollination peak periods (UMICH site)

PERMANOVAs on the LC-MS and GC-MS fingerprints showed overall significant differences between the atmo-metabolomes collected during different pollination peak periods (pseudo-$F = 9.54$, $P < 0.05$, and pseudo-$F = 7.63$, $P < 0.05$, respectively) (Table 1). Additionally, PCAs showed clear clustering of the different sample groups along PC1 and PC2, explaining the 39.61% and the 33.83% for LC-MS and the 55.81% and 25.81% for GC-MS analyses, respectively (Fig. 5c, d). In accordance with PERMANOVAs and PCAs, heatmaps of the LC-MS and GC-MS aerosol fingerprints showed clear differences between pollination periods (Fig. 6c, d). These results indicate thus that the proposed atmo-ecometabolomics methodology was sensitive enough as to accurately accurately extract, characterize, and distinguish aerosols collected during three different pollination peak periods (birch, oak, and pine) occurring within a continuous 40-day timeframe.

Compared to PNNL, the surroundings of UMICH have higher biological and human activity and this difference was distinctly noticeable in the aerosol metabolomic fingerprints. The number of features detected by both LC-MS and GC-MS instruments is considerably larger in UMICH than PNNL aerosol fingerprints (1333 vs. 7832 for LC-MS fingerprints and 148 vs. 221 for GC-MS fingerprints) (Fig. 6a vs. c and b vs. d). Furthermore, a larger number of metabolomics features could be identified at the aerosols collected in UMICH (45 and 20 for LC-MS and GC-MS, respectively) (Fig. 7b). Then, those results suggest that the number of detected and identified features in aerosol samples directly depends on the biological and anthropogenic activity of the surroundings. For this reason, it is crucial to keep relatively high flow rates (≥ 25–30 L/min) at the aerosol collection point in order to characterize properly the particle component of aerosol metabolomes, especially in ecosystems with low biological activity.

Conclusions and future perspectives

Our methodology was sensitive enough to detect significant overall differences between atmo-ecometabolomes from aerosol particles sampled in different seasons (spring vs. summer) in an ecosystem with low biological activity. Furthermore, we detected clear changes between the compositions of aerosol particles collected during tree pollination periods. These differences were apparent despite the high complexity of aerosol particle sources in the mix of urban and vegetated areas.

LC-MS and GC-MS showed similar accuracy in terms of sample clustering; both instruments were able to detect clear overall differences between spring and summer aerosols and between aerosols collected in different pollination peak seasons. Which instrument to use will finally depend on the hypotheses of the study and whether a targeted approximation is required. However, no analytical technique alone can still characterize the entire metabolome fingerprint of a sample and combining both platforms is a common approach. In general, the number of detected metabolic features is substantially higher in LC-MS analyses than GC-MS. This is especially useful for overall multivariate analyses in non-targeted metabolomics studies as LC-MS analyses provide a wider representation of the sample metabolome. Both instruments detected several polar and semi-polar metabolites, but GC-MS analyses allowed the identification of some non-polar compounds in aerosols (e.g., linoleic acid, oleic acid, palmitic acid, stearic acid). On the other hand, DI-FT-ICR-MS obtains ultra high-resolution metabolomic data allowing the assignation of elemental formulas in the detected features in aerosols. This data format allows performing different analyses on the distinct heteroatoms (e.g., O, N, P) present in aerosols. This information is especially useful to study the nutrient transport and deposition in ecosystems.

Coupling environmental variables (e.g., temperature, wind, precipitation, etc.) with atmo-ecometabolomics would allow a more precise interpretation of the aerosol-biosphere interface.

Long-term atmo-ecometabolomic research in natural ecosystems would significantly improve our understanding of interannual and seasonal changes of aerosol composition, directly linking the composition of aerosols with plant phenology and physiology, along environmental changes and/or natural spatial gradients.

The use of atmo-ecometabolomic techniques in ecological sciences could improve the detection, identification, and quantification of molecular compounds directly related with environmental stressors (biomarkers), providing thus important information of the general status of the ecosystems. A good description of such biomarkers would allow generating libraries of compounds that may serve to assess the ecosystem status or as an environmental monitoring tool.

A better understanding of the direct impacts of aerosols on biological surfaces (e.g., phyllosphere) and on the overall C:N:P stoichiometry of ecosystems can be improved significantly by the overall characterization of the aerosol chemical and molecular composition.

New modern high-resolution mass spectrometry instruments coupled to separation techniques should be implemented in atmo-ecometabolomic research to enable high performance for both mass accuracy and resolution, and retention time. Furthermore, recent advances in metabolomics instruments, such as mass spectrometers coupled to ion mobility spectrometry (IMS-MS), could substantially enhance the number of detected metabolic features in aerosol samples from tens and hundreds to thousands.

Acknowledgements The authors thank Rosalie Chu and Therese Clauss for their laboratory support. This research was performed using EMSL, a DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research at Pacific Northwest National Laboratory. JS and JP were supported by the European Research Council Synergy grant SyG-2013-610028 IMBALANCE-P, the Spanish Government projects CGL2013-48074-P, and the Catalan Government project SGR 2014-274. ALS was supported in part by National Science Foundation grant AGS 0952659.

References

Achotegui-Castells, A., Sardans, J., Ribas, À., & Peñuelas, J. (2013). Identifying the origin of atmospheric inputs of trace elements in the Prades Mountains (Catalonia) with bryophytes, lichens, and soil monitoring. *Environmental Monitoring and Assessment, 185*(1), 615–629.

Andreae, M. O., & Crutzen, P. J. (1997). Atmospheric aerosols: biogeochemical sources and role in atmospheric chemistry. *Science, 276*(5315), 1052–1058.

Arnold, A. E., Maynard, Z., Gilbert, G. S., Coley, P. D., & Kursar, T. A. (2000). Are tropical fungal endophytes hyperdiverse? *Ecology Letters, 3*(4), 267–274.

Ayers, G. P., & Gras, J. L. (1991). Seasonal relationship between cloud condensation nuclei and aerosol methanesulphonate in marine air. *Nature, 353*(6347), 834–835.

Azad K. R., & Shulaev, V. (2018). Metabolomics technology and bioinformatics for precision medicine. Briefings in bioinformatics, bxx170, https://doi.org/10.1093/bib/bbx170.

Baker, A. R., Kelly, S. D., Biswas, K. F., Witt, M., & Jickells, T. D. (2003). Atmospheric deposition of nutrients to the Atlantic Ocean. *Geophysical Research Letters, 30*(24).

Baustian, K. J., Cziczo, D. J., Wise, M. E., Pratt, K. A., Kulkarni, G., Hallar, A. G., & Tolbert, M. A. (2012). Importance of aerosol composition, mixing state, and morphology for heterogeneous ice nucleation: a combined field and laboratory approach. *Journal of Geophysical Research: Atmospheres, 117*, D06217.

Böttcher, C., Roepenack-Lahaye, E. v., Willscher, E., Scheel, D., & Clemens, S. (2007). Evaluation of matrix effects in metabolite profiling based on capillary liquid chromatography electrospray ionization quadrupole time-of-flight mass spectrometry. *Analytical Chemistry, 79*(4), 1507–1513.

Bundy, J. G., Davey, M. P., & Viant, M. R. (2008). Environmental metabolomics: a critical review and future perspectives. *Metabolomics, 5*(1), 3–21.

Canagaratna, M. R., Jayne, J. T., Jimenez, J. L., Allan, J. D., Alfarra, M. R., Zhang, Q., Onasch, T. B., Drewnick, F., Coe, H., Middlebrook, A., Delia, A., Williams, L. R., Trimborn, A. M., Northway, M. J., DeCarlo, P. F., Kolb, C. E., Davidovits, P., & Worsnop, D. R. (2007). Chemical and microphysical characterization of ambient aerosols with the aerodyne aerosol mass spectrometer. *Mass Spectrometry Reviews, 26*(2), 185–222.

Carlton, A. G., Pinder, R. W., Bhave, P. V., & Pouliot, G. A. (2010). To what extent can biogenic SOA be controlled? *Environmental Science & Technology, 44*(9), 3376–3380.

Carnicer, J., Sardans, J., Stefanescu, C., Ubach, A., Bartrons, M., Asensio, D., & Peñuelas, J. (2015). Global biodiversity, stoichiometry and ecosystem function responses to human-induced C–N–P imbalances. *Journal of Plant Physiology, 172*, 82–91.

Claudino, W. M., Quattrone, A., Biganzoli, L., Pestrin, M., Bertini, I., & Di Leo, A. (2007). Metabolomics: available results, current research projects in breast cancer, and future applications. *Journal of Clinical Oncology, 25*(19), 2840–2846.

de Mendiburu, F. (2015). Agricolae: statistical procedures for agricultural research. *R package version, 1*, 2–3 http://CRAN.R-project.org/package=agricolae.

De Vos, R. C., Moco, S., Lommen, A., Keurentjes, J. J., Bino, R. J., & Hall, R. D. (2007). Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nature Protocols, 2*(4), 778–791.

Després, V. R., Alex Huffman, J., Burrows, S. M., Hoose, C., Safatov, A. S., Buryak, G., et al. (2012). Primary biological aerosol particles in the atmosphere: a review. *Tellus B, 64*(15598), 1–58.

Ding, J., Sorensen, C. M., Zhang, Q., Jiang, H., Jaitly, N., Livesay, E. A., Shen, Y., Smith, R. D., & Metz, T. O. (2007). Capillary LC coupled with high-mass measurement accuracy mass spectrometry for metabolic profiling. *Analytical Chemistry, 79*(16), 6081–6093.

Elser, J. J., Dobberfuhl, D. R., MacKay, N. A., & Schampel, J. H. (1996). Organism size, life history, and N:P stoichiometry. *BioScience, 46*(9), 674–684.

Fageria, N. K., Filho, M. P. B., Moreira, A., & Guimarães, C. M. (2009). Foliar fertilization of crop plants. *Journal of Plant Nutrition, 32*(6), 1044–1064.

Feng, J., Wang, Y., Zhao, J., Zhu, L., Bian, X., & Zhang, W. (2011). Source attributions of heavy metals in rice plant along

highway in eastern China. *Journal of Environmental Sciences, 23*(7), 1158–1164.

Fernández, V., & Brown, P. H. (2013). From plant surface to plant metabolism: the uncertain fate of foliar-applied nutrients. *Frontiers in Plant Science, 4*, 289.

Fiehn, O. (2002). Metabolomics - the link between genotypes and phenotypes. *Plant Molecular Biology, 48*(1–2), 155–171.

Fukusaki, E., & Kobayashi, A. (2005). Plant metabolomics: potential for practical operation. *Journal of Bioscience and Bioengineering, 100*(4), 347–354.

Fuzzi, S., Andreae, M. O., Huebert, B. J., Kulmala, M., Bond, T. C., Boy, M., Doherty, S. J., Guenther, A., Kanakidou, M., Kawamura, K., Kerminen, V. M., Lohmann, U., Russell, L. M., & Pöschl, U. (2006). Critical assessment of the current state of scientific knowledge, terminology, and research needs concerning the role of organic aerosols in the atmosphere, climate, and global change. *Atmospheric Chemistry and Physics, 6*(7), 2017–2038.

Gargallo-Garriga, A., Sardans, J., Pérez-Trujillo, M., Oravec, M., Urban, O., Jentsch, A., Kreyling, J., Beierkuhnlein, C., Parella, T., & Peñuelas, J. (2015). Warming differentially influences the effects of drought on stoichiometry and metabolomics in shoots and roots. *New Phytologist, 207*(3), 591–603.

Gargallo-Garriga, A., Sardans, J., Pérez-Trujillo, M., Rivas-Ubach, A., Oravec, M., Vecerova, K., et al. (2014). Opposite metabolic responses of shoots and roots to drought. *Scientific Reports, 4*, 6829.

Gibney, M. J., Walsh, M., Brennan, L., Roche, H. M., German, B., & van Ommen, B. (2005). Metabolomics in human nutrition: opportunities and challenges. *The American Journal of Clinical Nutrition, 82*(3), 497–503.

Glauser, G., Guillarme, D., Grata, E., Boccard, J., Thiocone, A., Carrupt, P.-A., Veuthey, J. L., Rudaz, S., & Wolfender, J. L. (2008). Optimized liquid chromatography–mass spectrometry approach for the isolation of minor stress biomarkers in plant extracts and their identification by capillary nuclear magnetic resonance. *Journal of Chromatography A, 1180*(1), 90–98.

Gu, L., Baldocchi, D., Verma, S. B., Black, T. A., Vesala, T., Falge, E. M., & Dowty, P. R. (2002). Advantages of diffuse radiation for terrestrial ecosystem productivity. *Journal of Geophysical Research: Atmospheres, 107*(D6), 4050.

Gullberg, J., Jonsson, P., Nordström, A., Sjöström, M., & Moritz, T. (2004). Design of experiments: an efficient strategy to identify factors influencing extraction and derivatization of Arabidopsis thaliana samples in metabolomic studies with gas chromatography/mass spectrometry. *Analytical Biochemistry, 331*(2), 283–295.

Guy, C., Kaplan, F., Kopka, J., Selbig, J., & Hincha, D. K. (2007). Metabolomics of temperature stress. *Physiologia Plantarum, 132*(2), 220–235.

Hall, R. D. (2006). Plant metabolomics: from holistic hope, to hype, to hot topic. *New Phytologist, 169*(3), 453–468.

Hiller, K., Hangebrauk, J., Jäger, C., Spura, J., Schreiber, K., & Schomburg, D. (2009). MetaboliteDetector: comprehensive analysis tool for targeted and nontargeted GC/MS based metabolome analysis. *Analytical Chemistry, 81*(9), 3429–3439.

Hirai, M. Y., Yano, M., Goodenowe, D. B., Kanaya, S., Kimura, T., Awazuhara, M., Arita, M., Fujiwara, T., & Saito, K. (2004). From the cover: integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences, 101*(27), 10205–10210.

Husson, F., & Josse, J. (2016). missMDA: A package for handling missing values in multivariate data analysis. *Journal of Statistical Software, 70*(1), 1–31.

Husson, F., Josse, J., Le, S., & Mazet, J. (2016). FactoMineR: multivariate exploratory data analysis and data mining. *R package version, 1.32,* https://CRAN.R-project.org/package=FactoMineR. Accessed 5 Feb 2018.

Ingram, J., & Bartels, D. (1996). The molecular basis of dehydration tolerance in plants. *Annual Review of Plant Biology, 47*(1), 377–403.

Jokinen, T., Berndt, T., Makkonen, R., Kerminen, V.-M., Junninen, H., Paasonen, P., Stratmann, F., Herrmann, H., Guenther, A. B., Worsnop, D. R., Kulmala, M., Ehn, M., & Sipilä, M. (2015). Production of extremely low volatile organic compounds from biogenic emissions: measured yields and atmospheric implications. *Proceedings of the National Academy of Sciences, 112*(23), 7123–7128.

Kaplan, F., Kopka, J., Haskell, D. W., Zhao, W., Schiller, K. C., Gatzke, N., et al. (2004). Exploring the temperature-stress metabolome of Arabidopsis. *Plant Physiology, 136*(4), 4159–4168.

Kellerman, A. M., Dittmar, T., Kothawala, D. N., & Tranvik, L. J. (2014). Chemodiversity of dissolved organic matter in lakes driven by climate and hydrology. *Nature Communications, 5*, 3804.

Keltjens, W. G., & van Beusichem, M. L. (1998). Phytochelatins as biomarkers for heavy metal stress in maize (Zea mays L.) and wheat (Triticum aestivum L.): combined effects of copper and cadmium. *Plant and Soil, 203*(1), 119–126.

Kim, H. K., Choi, Y. H., & Verpoorte, R. (2010). NMR-based metabolomic analysis of plants. *Nature Protocols, 5*(3), 536–549.

Kim, S., Kramer, R. W., & Hatcher, P. G. (2003). Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the Van Krevelen diagram. *Analytical Chemistry, 75*(20), 5336–5344.

Kim, Y.-M., Nowack, S., Olsen, M. T., Becraft, E. D., Wood, J. M., Thiel, V., et al. (2015). Diel metabolomics analysis of a hot spring chlorophototrophic microbial mat leads to new hypotheses of community member metabolisms. *Frontiers in Microbiology, 6*, 209.

Kind, T., Wohlgemuth, G., Lee, D. Y., Lu, Y., Palazoglu, M., Shahbaz, S., & Fiehn, O. (2009). FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Analytical Chemistry, 81*(24), 10038–10048.

Klein, G. C., Rodgers, R. P., & Marshall, A. G. (2006). Identification of hydrotreatment-resistant heteroatomic species in a crude oil distillation cut by electrospray ionization FT-ICR mass spectrometry. *Fuel, 85*, 2071–2080.

Kujawinski, E. (2002). Electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS): characterization of complex environmental mixtures. *Environmental Forensics, 3*(3), 207–216.

Kujawinski, E. B., & Behn, M. D. (2006). Automated analysis of electrospray ionization Fourier transform ion cyclotron

resonance mass spectra of natural organic matter. *Analytical Chemistry, 78*(13), 4363–4373.

Lee, D. Y., & Fiehn, O. (2013). Metabolomic response of Chlamydomonas reinhardtii to the inhibition of target of rapamycin (TOR) by rapamycin. Journal of Microbiology and Biotechnology, 23(7), 923–931. L.

Lei, Z., Huhman, D., & Sumner, L. L. (2011). Mass spectrometry strategies in metabolomics. *Journal of Biological Chemistry., 286*, 25435–25442. https://doi.org/10.1074/jbc.R111.238691.

Leiss, K. A., Choi, Y. H., Abdel-Farid, I. B., Verpoorte, R., & Klinkhamer, P. G. L. (2009). NMR metabolomics of thrips (Frankliniella occidentalis) resistance in Senecio hybrids. *Journal of Chemical Ecology, 35*(2), 219–229.

Leiss, K. A., Cristofori, G., van Steenis, R., Verpoorte, R., & Klinkhamer, P. G. L. (2013). An eco-metabolomic study of host plant resistance to Western flower thrips in cultivated, biofortified and wild carrots. *Phytochemistry, 93*, 63–70.

Lin, C. Y., Viant, M. R., & Tjeerdema, R. S. (2006). Metabolomics: methodologies and applications in the environmental sciences. *Journal of Pesticide Science, 31*(3), 245–251.

Lindon, J.C., Nicholson, J.K. & Holmes, E. (eds.) The handbook of metabonomics and metabolomics (Elsevier, Amsterdam, 2007), 55–201.

Lindow, S. E., & Brandl, M. T. (2003). Microbiology of the phyllosphere. *Applied and Environmental Microbiology, 69*(4), 1875–1883.

Liu, Y., & Kujawinski, E. B. (2015). Chemical composition and potential environmental impacts of water-soluble polar crude oil components inferred from ESI FT-ICR MS. *PLoS One, 10*(9), e0136376.

Macedo, A. F. (2012). Abiotic stress responses in plants: metabolism to productivity. In *Abiotic stress responses in plants* (pp. 41–61). New York, NY: Springer New York.

Mahowald, N. M., Artaxo, P., Baker, A. R., Jickells, T. D., Okin, G. S., Randerson, J. T., & Townsend, A. R. (2005). Impacts of biomass burning emissions and land use change on Amazonian atmospheric phosphorus cycling and deposition. *Global Biogeochemical Cycles, 19*, GC4030.

Mari, A., Lyon, D., Fragner, L., Montoro, P., Piacente, S., Wienkoop, S., Egelhofer, V., & Weckwerth, W. (2013). Phytochemical composition of Potentilla anserina L. analyzed by an integrative GC-MS and LC-MS metabolomics platform. *Metabolomics : Official journal of the Metabolomic Society, 9*(3), 599–607.

Medeiros, P. M., Babcock-Adamos, L., Seidel, M., Castelao, R. M., Di Iorio, D., Hollibaugh, J. T., & Dittmar, T. (2017). Export of terrigenous dissolved organic matter in a broad continental shelf. *Limnology and Oceanography, 62*, 1718–1731.

Menzel, A., Sparks, T. H., Estrella, N., Koch, E., Aasa, A., Ahas, R., et al. (2006). European phenological response to climate change matches the warming pattern. *Global Change Biology, 12*(10), 1969–1976.

Minor, E. C., Swenson, M. M., Mattson, B. M., & Oyler, A. R. (2014). Structural characterization of dissolved organic matter: a review of current techniques for isolation and analysis. *Environ. Sci Processes Impacts, 16*(9), 2064–2079.

Nikiforova, V. J., Kopka, J., Tolstikov, V., Fiehn, O., Hopkins, L., Hawkesford, M. J., et al. (2005). Systems rebalancing of metabolism in response to sulfur deprivation, as revealed by metabolome analysis of Arabidopsis plants. *Plant Physiology, 138*(1), 304–318.

Obee, T. N., & Hay, S. O. (1997). Effects of moisture and temperature on the photooxidation of ethylene on Titania. *Environmental Science & Technology, 31*(7), 2034–2038.

Oksanen, J., Guillaume-Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., et al. (2013). vegan: community ecology package. R package version 2.0–9, http ://CRAN.R-project.org/package=vegan

Osterholz, H., Singer, G., Wemheuer, B., Daniel, R., Simon, M., Niggemann, J., & Dittmar, T. (2016). Deciphering associations between dissolved organic molecules and bacterial communities in a pelagic marine system. *The ISME Journal, 10*(7), 1717–1730.

Paerl, H. W. (1997). Coastal eutrophication and harmful algal blooms: importance of atmospheric deposition and groundwater as "new" nitrogen and other nutrient sources. *Limnology and Oceanography, 42*(5part2), 1154–1165.

Pan, Z., & Raftery, D. (2007). Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics. *Analytical and Bioanalytical Chemistry, 387*(2), 525–527.

Pandis, S. N., Harley, R. A., Cass, G. R., & Seinfeld, J. H. (1992). Secondary organic aerosol formation and transport. *Atmospheric Environment Part A. General Topics, 26*(13), 2269–2282.

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics, 37*(1), 637–669.

Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature, 421*(6918), 37–42.

Paytan, A., Mackey, K. R. M., Chen, Y., Lima, I. D., Doney, S. C., Mahowald, N., Labiosa, R., & Post, A. F. (2009). Toxicity of atmospheric aerosols on marine phytoplankton. *Proceedings of the National Academy of Sciences of the United States of America, 106*(12), 4601–4605.

Peñuelas, J., & Sardans, J. (2009). Ecological metabolomics. *Chemistry and Ecology, 25*(4), 305–309.

Peñuelas, J., Sardans, J., Rivas-Ubach, A., & Janssens, I. A. (2012). The human-induced imbalance between C, N and P in Earth's life system. *Global Change Biology, 18*(1), 3–6.

Peñuelas, J., & Staudt, M. (2010). BVOCs and global change. *Trends in Plant Science, 15*(3), 133–144.

Peñuelas, J., & Terradas, J. (2014). The foliar microbiome. *Trends in Plant Science, 19*(5), 278–280.

Pluskal, T., Castillo, S., Villar-Briones, A., & Orešič, M. (2010). MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics, 11*(1), 395.

R Core Team. (2013). R: a language and environment for statistical computing. Vienna.

Ramanathan, V., Crutzen, P. J., Kiehl, J. T., & Rosenfeld, D. (2001). Aerosols, climate, and the hydrological cycle. *Science, 294*(5549), 2119–2124.

Reemtsma, T. (2009). Determination of molecular formulas of natural organic matter molecules by (ultra-) high-resolution mass spectrometry: status and needs. *Journal of Chromatography A, 1216*(18), 3687–3701.

Riedel, T., & Dittmar, T. (2014). A method detection limit for the analysis of natural organic matter via Fourier transform ion

cyclotron resonance mass spectrometry. *Analytical Chemistry, 86*(16), 8376–8382.

Rivas-Ubach, A., Barbeta, A., Sardans, J., Guenther, A., Ogaya, R., Oravec, M., Urban, O., & Peñuelas, J. (2016b). Topsoil depth substantially influences the responses to drought of the foliar metabolomes of Mediterranean forests. *Perspectives in Plant Ecology, Evolution and Systematics, 21*, 41–54.

Rivas-Ubach, A., Gargallo-Garriga, A., Sardans, J., Oravec, M., Mateu-Castell, L., Pérez-Trujillo, M., Parella, T., Ogaya, R., Urban, O., & Peñuelas, J. (2014). Drought enhances folivory by shifting foliar metabolomes in Quercus ilex trees. *New Phytologist, 202*(3), 874–885.

Rivas-Ubach, A., Hódar, J. A., Sardans, J., Kyle, J. E., Kim, Y.-M., Oravec, M., Urban, O., Guenther, A., & Peñuelas, J. (2016a). Are the metabolomic responses to folivory of closely related plant species linked to macroevolutionary and plant–folivore coevolutionary processes? *Ecology and Evolution, 6*(13), 4372–4386.

Rivas-Ubach, A., Liu, Y., Bianchi, T. S., Tolić, N., Jansson, C., & Paša-Tolić, L. (2018a). Moving beyond the van Krevelen diagram: a new stoichiometric approach for compound classification in organisms. Analytical Chemistry, acs.Analchem.8b00529. doi:https://doi.org/10.1021/acs.analchem.8b00529.

Rivas-Ubach, A., Pérez-Trujillo, M., Sardans, J., Gargallo-Garriga, A., Parella, T., & Peñuelas, J. (2013). Ecometabolomics: optimized NMR-based method. *Methods in Ecology and Evolution, 4*(5), 464–473.

Rivas-Ubach, A., Poret-Peterson, A. T., Peñuelas, J., Sardans, J., Pérez-Trujillo, M., Legido-Quigley, C., et al. (2018b). Coping with iron limitation: a metabolomic study of Synechocystis sp. PCC 6803. *Acta Physiologiae Plantarum, 40*(2), 28.

Rivas-Ubach, A., Sardans, J., Hódar, J. A., Garcia-Porta, J., Guenther, A., Oravec, M., Urban, O., & Peñuelas, J. (2016c). Similar local, but different systemic, metabolomic responses of closely related pine subspecies to folivory by caterpillars of the processionary moth. *Plant Biology, 18*(3), 484–494.

Rivas-Ubach, A., Sardans, J., Hódar, J. A., Garcia-Porta, J., Guenther, A., Paša-Tolić, L., Oravec, M., Urban, O., & Peñuelas, J. (2017). Close and distant: contrasting the metabolism of two closely related subspecies of Scots pine under the effects of folivory and summer drought. *Ecology and Evolution, 7*(21), 8976–8988.

Rivas-Ubach, A., Sardans, J., Peŕez-Trujillo, M., Estiarte, M., & Penũelas, J. (2012). Strong relationship between elemental stoichiometry and metabolome in plants. *Proceedings of the National Academy of Sciences of the United States of America, 109*(11), 4181–4186.

Rochford, S. (2005). Metabolomics reviewed: a new "omics" platform technology for systems biology and implications for natural products research. *Journal of Natural Products, 68*, 1813–1820.

Roullier-Gall, C., Boutegrabet, L., Gougeon, R. D., & Schmitt-Kopplin, P. (2014). A grape and wine chemodiversity comparison of different appellations in Burgundy: vintage vs terroir effects. *Food Chemistry, 152*, 100–107.

Roulston, T. H., & Cane, J. H. (2000). Pollen nutritional content and digestibility for animals. *Plant Systematics and Evolution, 222*(1–4), 187–209.

Saito, K., & Matsuda, F. (2010). Metabolomics for functional genomics, systems biology, and biotechnology. *Annual Review of Plant Biology, 361*, 463–489.

Sardans, J., Gargallo-Garriga, A., Pérez-Trujillo, M., Parella, T. J., Seco, R., Filella, I., & Peñuelas, J. (2014). Metabolic responses of Quercus ilex seedlings to wounding analysed with nuclear magnetic resonance profiling. *Plant Biology, 16*(2), 395–403.

Sardans, J., Peñuelas, J., & Rivas-Ubach, A. (2011). Ecological metabolomics: overview of current developments and future challenges. *Chemoecology, 21*(4), 191–225.

Sardans, J., Rivas-Ubach, A., & Peñuelas, J. (2012). The elemental stoichiometry of aquatic and terrestrial ecosystems and its relationships with organismic lifestyle and ecosystem structure and function: a review and perspectives. *Biogeochemistry, 111*(1–3), 1–39.

Schmitt-Kopplin, P., Liger-Belair, G., Koch, B. P., Flerus, R., Kattner, G., Harir, M., Kanawati, B., Lucio, M., Tziotis, D., Hertkorn, N., & Gebefügi, I. (2012). Dissolved organic matter in sea spray: a transfer study from marine surface water to aerosols. *Biogeosciences, 9*(4), 1571–1582.

Shulaev, V. (2006). Metabolomics technology and bioinformatics. *Briefings in Bioinformatics, 7*(2), 128–139.

Seco, R., Peñuelas, J., & Filella, I. (2007). Short-chain oxygenated VOCs: emission and uptake by plants and atmospheric sources, sinks, and concentrations. *Atmospheric Environment, 41*(12), 2477–2499.

Shulaev, V., Cortes, D., Miller, G., & Mittler, R. (2008). Metabolomics for plant stress response. *Physiologia Plantarum, 132*(2), 199–208.

Sleighter, R. L., & Hatcher, P. G. (2007). The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter. *Journal of mass spectrometry : JMS, 42*(5), 559–574.

Smith, D. F., Podgorski, D. C., Rodgers, R. P., Blakney, G. T., & Hendrickson, C. L. (2018). 21 tesla FT-ICR mass spectrometer for ultrahigh-resolution analysis of complex organic mixtures. *Analytical Chemistry, 90*, 2041–2047. https://doi.org/10.1021/acs.analchem.7b04159.

Smith, D., & Španěl, P. (2011). Direct, rapid quantitative analyses of BVOCs using SIFT-MS and PTR-MS obviating sample collection. *TrAC Trends in Analytical Chemistry, 30*(7), 945–959.

Solberg, Y., & Remedios, G. (1980). Chemical composition of pure and bee-collected pollen. *Meldinger fra Norges landbrukshogskole, 59*, 1–13.

Spencer, R. G. M., Mann, P. J., Dittmar, T., Eglinton, T. I., McIntyre, C., Holmes, R. M., Zimov, N., & Stubbins, A. (2015). Detecting the signature of permafrost thaw in Arctic rivers. *Geophysical Research Letters, 42*(8), 2830–2835.

Sterner, R., & Elser, J. (2002). *Ecological stoichiometry: the biology of elements from molecules to the biosphere*. Princetion University Press.

Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., Fan, T. W. M., Fiehn, O., Goodacre, R., Griffin, J. L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A. N., Lindon, J. C., Marriott, P., Nicholls, A. W., Reily, M. D., Thaden, J. J., & Viant, M. R. (2007). Proposed minimum reporting standards for chemical analysis. *Metabolomics, 3*(3), 211–221.

t'Kindt, R., De Veylder, L., Storme, M., Deforce, D., & Van Bocxlaer, J. (2008). LC-MS metabolic profiling of Arabidopsis thaliana plant leaves and cell cultures: optimization of pre-LC-MS procedure parameters. *Journal of chromatography B, Analytical technologies in the biomedical and life sciences, 871*(1), 37–43.

Tfaily, M. M., Chu, R. K., Tolić, N., Roscioli, K. M., Anderton, C. R., Paša-Tolić, L., Robinson, E. W., & Hess, N. J. (2015). Advanced solvent based methods for molecular characterization of soil organic matter by high-resolution mass spectrometry. *Analytical Chemistry, 87*(10), 5206–5215.

Tholl, D., Boland, W., Hansel, A., Loreto, F., Röse, U. S. R., & Schnitzler, J.-P. (2006). Practical approaches to plant volatile analysis. *The Plant Journal, 45*(4), 540–560.

Thompson, H., Heimendinger, J., Gillette, C., Sedlacek, S., Haegele, A., O'Neill, C., & Wolfe, P. (2005). In vivo investigation of changes in biomarkers of oxidative stress induced by plant food rich diets. *Journal of Agricultural and Food Chemistry, 53*(15), 6126–6132.

Tolić, N., Liu, Y., Liyu, A., Shen, Y., Tfaily, M. M., Kujawinski, E. B., Longnecker, K., Kuo, L. J., Robinson, E. W., Paša-Tolić, L., & Hess, N. J. (2017). Formularity: software for automated formula assignment of natural and other organic matter from ultrahigh-resolution mass spectra. *Analytical Chemistry, 89*, 12659–12665. https://doi.org/10.1021/acs.analchem.7b03318.

Uzu, G., Sobanska, S., Sarret, G., Muñoz, M., & Dumat, C. (2010). Foliar lead uptake by lettuce exposed to atmospheric fallouts. *Environmental Science & Technology, 44*(3), 1036–1042.

van den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K., & van der Werf, M. J. (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics, 7*(1), 142.

van Krevelen, D. (1950). Graphical-statistical method for the study of structure and reaction processes of coal. *Fuel, 29*, 269–284.

Vorholt, J. A. (2012). Microbial life in the phyllosphere. *Nature Reviews Microbiology, 10*(12), 828–840.

Walsh, M. C., Brennan, L., Malthouse, J. P. G., Roche, H. M., & Gibney, M. J. (2006). Effect of acute dietary standardization on the urinary, plasma, and salivary metabolomic profiles of healthy humans. *The American Journal of Clinical Nutrition, 84*(3), 531–539.

Walther, G. R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., Fromentin, J. M., Hoegh-Guldberg, O., & Bairlein, F. (2002). Ecological responses to recent climate change. *Nature, 416*(6879), 389–395.

Wang, R., Balkanski, Y., Bopp, L., Aumont, O., Boucher, O., Ciais, P., Gehlen, M., Peñuelas, J., Ethé, C., Hauglustaine, D., Li, B., Liu, J., Zhou, F., & Tao, S. (2015). Influence of anthropogenic aerosol deposition on the relationship between oceanic productivity and warming. *Geophysical Research Letters, 42*(24), 10745–10754.

Warnes, G. R., Bolker, B., Bonebakker, L., Gentleman, R., Huber Andy Liaw, W., Lumley, T., et al. (2016). gplots: various R programing tools for plotting data. R package version 3.0.1, http://CRAN.R-project.org/package=gplots

Wedding, J. B., Carlson, R. W., Stukel, J. J., & Bazzaz, F. A. (1975). Aerosol deposition on plant leaves. *Environmental Science & Technology, 9*(2), 151–153.

Whipps, J. M., Hand, P., Pink, D., & Bending, G. D. (2008). Phyllosphere microbiology with special reference to diversity and plant genotype. *Journal of Applied Microbiology, 105*(6), 1744–1755.

White, R. A., Rivas-Ubach, A., Borkum, M. I., Köberl, M., Bilbao, A., Colby, S. M., et al. (2017). The state of rhizospheric science in the era of multi-omics: a practical guide to omics technologies. *Rhizosphere, 3*, 212–221. https://doi.org/10.1016/J.RHISPH.2017.05.003.

Wishart, D. S. (2008). Metabolomics: applications to food science and nutrition research. *Trends in Food Science & Technology, 19*(9), 482–493.

Wozniak, A. S., Bauer, J. E., Sleighter, R. L., Dickhut, R. M., & Hatcher, P. G. (2008). Technical note: molecular characterization of aerosol-derived water soluble organic carbon using ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Atmospheric Chemistry and Physics, 8*(17), 5099–5111.

Xiong, T.-T., Leveque, T., Austruy, A., Goix, S., Schreck, E., Dappe, V., Sobanska, S., Foucault, Y., & Dumat, C. (2014). Foliar uptake and metal(loid) bioaccessibility in vegetables exposed to particulate matter. *Environmental Geochemistry and Health, 36*(5), 897–909.

Zhang, A., Sun, H., Wang, P., Han, Y., & Wang, X. (2012). Modern analytical techniques in metabolomics analysis. *The Analyst, 137*(2), 293–300.

Zhang, Q., Stanier, C. O., Canagaratna, M. R., Jayne, J. T., Worsnop, D. R., Pandis, S. N., & Jimenez, J. L. (2004). Insights into the chemistry of new particle formation and growth events in Pittsburgh based on aerosol mass spectrometry. *Environmental Science & Technology, 38*(18), 4797–4809.