1 **Atmo-ecometabolomics: a novel atmospheric particle chemical characterization**

2 **methodology for ecological research**

3

4 Albert Rivas-Ubach[1], Yina Liu[1,2], Allison L. Steiner[3], Jordi Sardans[4,5], Malak M. Tfaily[1], Gourihar

5 Kulkarni[6], Young-Mo Kim[7], Eric Bourrianne[8], Ljiljana Paša-Tolić [1], Josep Peñuelas[4,5], Alex

6 Guenther[9].

7

8 1. Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland,

9 99354, WA, USA.

10 2. Geochemical and Environmental Research Group, Texas A&M University, College Station, 77845, TX,

11 USA.

12 3. Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor,

13 48109, MI, USA.

14 4. CREAF, Campus UAB, Cerdanyola del Vallès, 08913 Catalonia, Spain.

15 5. Global Ecology Unit CREAF-CSIC, Campus UAB, Cerdanyola del Vallès, 08913 Catalonia, Spain.

16 6. Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland,

17 99352, WA, USA.

18 7. Biological Sciences Division, Pacific Northwest National Laboratory, Richland, 99354, WA, USA.

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

20 9. Department of Earth System Science, University of California, Irvine, 92697, CA, USA.

21

22 **Author of correspondence:**

23 Albert Rivas-Ubach

24 Environmental Molecular Sciences Laboratory,

25 Pacific Northwest National Laboratory,

26 Richland, WA, USA, 99354

27 Tel: 971 319 5962

28 e-mail: albert.rivas.ubach@pnnl.gov // albert.rivas.ubach@gmail.com

29

30

31

32

33

34

35

36

37 **Abstract.**

38

39 Aerosol particles play important roles in processes controlling the composition of the

40 atmosphere and function of ecosystems. A better understanding on the composition of

41 aerosol particles is beginning to be recognized as critical for ecological research to further

42 comprehend the link between aerosols and ecosystems. Whilechemical characterization of

43 aerosols has been practiced in the atmospheric science community, detailed methodology

44 tailored to the needs of ecological research does not exist yet. In this study, we describe an

45 efficient methodology (atmo-ecometabolomics), in step-by-step details, from the sampling to

46 the data analyses, to characterize the chemical composition of aerosols particles, namely

47 atmo-metabolome. This method employs mass spectrometry platforms such as liquid and gas

48 chromatography mass spectrometries (MS), and Fourier transform ion cyclotron resonance MS

49 (FT-ICR-MS). For methodology evaluation, we analyzed aerosol particles collected during two

50 different seasons (spring and summer) in a low biological activity ecosystem. Additionally, to

51 further validate our methodology, we analyzed aerosol particlescollected in a more biological

52 active ecosystem during the pollination peaks of three different representative tree species.

53 Our statistical results showed that our sampling and extraction methods are suitable for

54 characterizing the atmo-ecometabolomes in these two distinct ecosystems with any of the

55 analytical platforms. Datasets obtained from each mass spectrometry instrument showed

56 overall significant differences of the atmo-ecometabolomes between spring and summer as

57 well asbetween the three polliation peak periods. Furthermore, we have identified several

58 metabolites that can be attributed to pollen and other plant-related aerosol particles. We

59 additionally provide a basic guide of the potential use ecometabolomics techniques on

60 different mass spectrometry platforms to accurately analyze the atmo-ecometabolomes for

61 ecological studies. Our method represents an advanced novel approach for future studies in

62 the impact of aerosol particle chemical compositions on ecosystem structure and function and

63 biogeochemistry.

64

65 Keywords: Aerosol particles, metabolomics, ecosystems, biomarkers, mass spectrometry, FT-

66 ICR

67

68

69

70

71

1. 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 include pollen, microorganisms, 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) (Figure 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 wellknown 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 is still not totally understood.Furthermore, most of such research were 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 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 plysiology of plants (Arnold et al. 2000; Lindow and Brandl 2003; Vorholt 2012), can benefit from aerosol particles to acquire nutrients (e.g. nitrogen (N), phosphorous (P), sulphur (S), …). Variations in aerosol particle chemical composition could lead to significant changes on 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 for phytoplantkon it has been studied in great extent (Baker et al. 2003; Paerl 1997; Paytan et al. 2009; Wang et al. 2015), as itrepresents 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 onthe essential elements balance on organisms and ecosystems have proven that changes in nutrient proportions, mainly carbon (C), N P, can lead to substantial alterations in ecosystems (Elser et al. 1996; Sterner and Elser

3

107   2002). Therefore, changes in aerosol particle elemental composition may lead to significant

108   modifications in terrestrial and aquatic ecosystems functioning (Carnicer et al. 2015; Peñuelas

109   et al. 2013; Sardans et al. 2012). Understanding the elemental and molecular compositions of

110   aerosol particles is thus crucial for comprehending processes occuring at the atmosphere-

111   ecosystem interface and for understanding how aerosols can be related to ecosystemic

112   changes.

113   Low molecular weight compounds (~80-1000 Dalton) have not generally identified or

114   taken into account in aerosol particles and could play crucial roles in the ecosystem

115   functioning, especially at the atmosphere-biosphere interface. Here, we describe in detail a

116   novel and simple methodology to collect ambient aerosol particle samples and subsequently

117   characterize their chemical composition using different mass spectrometry (MS) platforms.

118   This methodology represents a useful tool to shed light on novel questions in plant physiology,

119   ecophysiology, and ecology to understand deeper the atmosphere-ecosystem interface. In this

120   manuscript, we first briefly define ecometabolomics and its potential as a technique to

121   characterize the chemical composition of aerosol particles. We then provide a basic overview

122   of the common MS techniques used in metabolomics. Finally, we discuss our methodologic

123   approach and the different recent metabolomics technologies and, briefly, the obtained

124   results from the samples collected and analyzed with the proposed methodology.

125

126   **1.1 Atmo-ecometabolomics: metabolomics of aerosols.**

127   Metabolomics aims to study the metabolome of entire organisms or specific cells or tissues,

128   and this technique includes all the procedures from sample collection, metabolite extraction

129   and analyses to data filtering, statistical analyses, and interpretation of the results (Figure 2). A

130   metabolome consists of the total set of small (<1,000 Da) compounds (metabolites) present in

131   a sample at a given moment (Fiehn 2002). Such compounds are represented by the substrates

132   and products from cellular primary metabolism such as organic acids, amino acids, and sugars

133   as well as a vast diversity of compounds derived from the secondary metabolism of organisms

134   such as polyphenolics and terpenes, typically synthetized by plants and fungi. All these

135   metabolites are involved in diverse and complex chemical reactions to finally maintain

136   organisms' homeostasis, reproduction, and growth (Peñuelas and Sardans 2009).

137   Metabolomics has been widely used in human nutrition research (Gibney et al. 2005; Wishart

138   2008), medicine (Claudino et al. 2007; Walsh et al. 2006), plant physiology (Hirai et al. 2004;

139   Kaplan et al. 2004), and more recently in the field of ecology (ecometabolomics) (Bundy et al.

140   2008; Gargallo-Garriga et al. 2014; Rivas-Ubach et al. 2012, 2016a, 2018a). Ecometabolomics

141   has proven to be useful to understand the plasticity of organisms' metabolomes under specific

142 environmental conditions and/or stressor pressures (Gargallo-Garriga et al. 2014; Rivas-Ubach

143 et al. 2016b, 2017; Sardans et al. 2011, 2014). However, ecometabolomic analyses of aerosol

144 particles (atmo-ecometabolomics) have never been reported before and it is a step further to

145 characterize the aerosol particle composition and to improve our understanding of the

146 aerosol-biosphere interface.

147   Metabolomics can detect specific molecular signatures (biomarkers) directly related to

148 organismal stress and the phenological status of ecosystems. Recent climate projections

149 predict increments of extreme climatic events such as drought and warming that increases

150 plant BVOC emissions (Peñuelas and Staudt 2010) and lead to shifts in the phenology of plants

151 (Menzel et al. 2006; Parmesan 2006; Parmesan and Yohe 2003; Walther et al. 2002).

152 Terrestrial plants have proven to show large overall metabolome shifts when exposed to

153 stressors (Gargallo-Garriga et al. 2014, 2015; Leiss et al. 2009; Macedo 2012; Rivas-Ubach et al.

154 2012, 2014) and several stress biomarkers have been already identified (Glauser et al. 2008;

155 Guy et al. 2007; Keltjens and van Beusichem 1998; Shulaev et al. 2008; Thompson et al. 2005).

156 Significant changes in plant phenology have been detected during the last decades (Menzel et

157 al. 2006; Parmesan 2006; Parmesan and Yohe 2003; Walther et al. 2002) and due the directly

158 link between PBAP and ecosystems, such phenoylogical shifts should be detectable with atmo-

159 ecometabolomics. Therefore, long-term atmo-ecometabolomics studies should provide

160 valuble information of phenological changes and succession or recession of ecosystems. Atmo-

161 ecometabolomics can be also applied to assess stress and phenological changes at ecosystem

162 and regional scales through the characterization of atmo-ecometabolomes.

163

164 **1.2 Analytical instruments for MS-based metabolomics analyses.**

165 Metabolomic profiles are achieved through two main processes: i) obtaining the chemical

166 signature of the sample (metabolomic fingerprint) and ii) identification of metabolites through

167 the information previously obtained from the metabolomic fingerprints (e.g. exact mass,

168 retention time, fragmentation pattern, etc). MS coupled to liquid or gas chromatography (LC-

169 MS and GC-MS respectively) are currently two of the most common techniques used in

170 metabolomics (Fiehn 2002; Sardans et al. 2011; Zhang et al. 2012), demonstrating high

171 performance and sensitivity (Pan and Raftery 2007). LC-MS and GC-MS provide similar data

172 format; i.e. in both instruments, before compound detection, metabolites are separated by

173 chromatography (liquid or gas) producing finally two orthogonal and independent values;

174 retention time (RT) and mass-to-charge ratio (m/z) for each of the detected features. RT and

175 m/z, together with other additional parameters (e.g. fragmentation pattern), can be thus used

176 for metabolite identification (Sumner et al. 2007). Although there is a wide arrange of

177   compounds detectably GC-MS and LC-MS platforms, generally, GC-MS is more suitable for

178   detecting highly-volatile and non-polar compounds with mass weight typically <600 Da such as

179   fatty acids, carotenoids, and essential oils (Gullberg et al. 2004). Untargeted GC-MS analyses

180   require previous compound derivatization of the samples to allow polar molecular compounds

181   to be sufficiently volatile so they can elute at moderate temperature without disintegration.

182   Therefore, carbohydrates, organic acids and other polar analytes can be also detected by GC-

183   MS after derivatization (Gullberg et al. 2004). In comparison to LC, GC has proved to have

184   exceptional reproducibility showing little RT shifts between samples. However, derivatization

185   of the samples increases preparation time and provides indirect metabolite detection that

186   complicates the elucidation of unknown compounds. On the other hand, LC-MS can cover a

187   wider range of compounds than GC-MS and allows the detection of compounds with larger

188   molecular weight (typically up to 1200 Da for metabolomics). LC-MS allows analyzing non-

189   volatile and highly polar compounds ranging from primary metabolic compounds such as

190   organic acids, carbohydrates and amino acids to secondary metabolites such as alkaloids,

191   phenolic acids, flavonoids, or saponins (De Vos et al. 2007). LC-MS provides a direct detection

192   of molecular ions because derivatization is not required. However, LC-MS analyses commonly

193   show larger RT shifts between samples in comparison to GC-MS, especially in studies with

194   large number of samples (e.g. >300). Nonetheless, no single mass spectrometry instrument can

195   cover all molecular compound classes (Ding et al. 2007; Zhang et al. 2012), and combining

196   diverse platforms is a common approach for deeper metabolome characterization of the

197   samples (Hall 2006).

198   Mass resolving power of MS instruments is a critical factor to consider in MS-based

199   metabolomics. The high-resolution and mass accuracy (typically up to 250,000 and < 3 ppm,

200   respectively) of Orbitrap mass spectrometers has proven to be suitable for ecometabolomic

201   studies (White et al. 2017) reducing considerably the error of metabolite identification when

202   using high-resolution compound libraries (Rivas-Ubach et al. 2016c). However, Fourier

203   transform ion cyclotron resonance MS (FT-ICR-MS) is currently the MS platform achieving the

204   highest mass resolving power (up to $m/\Delta m_{50\%} > 2,700,000$ at m/z 400) and mass accuracy (<

205   1ppm mass error after internal calibration) (Smith et al. 2018). Although FT-ICR-MS can be

206   coupled to HPLC, the scan rate of FT-ICR-MS is commonly not rapid enough as to analyze

207   samples at ultra-high resolution (>400,000) when coupled to LC. This is one of the main

208   reasons why direct infusion (DI) has been the most typical approach to analyze samples with

209   the FT-ICR-MS in ultra-high resolution. Furthermore, DI-FT-ICR-MS acquisition time is

210   significantly shorter (typically between 5-15 minutes) than LC or GC methods which can take

211   over 40 minutes per sample. The ultra-high mass resolution and excellent mass accuracy of DI-

212   FT-ICR-MS enables the assignment of elemental formula of a wide portion of the detected

213   features based exclusively on their exact mass (Klein et al. 2006; Kujawinski 2002) providing,

214   thus, powerful means to understand the overall chemical characteristics of organic complex

215   samples (Kim et al. 2003; Reemtsma 2009; Roullier-Gall et al. 2014; Schmitt-Kopplin et al.

216   2012; Sleighter and Hatcher 2007; Tfaily et al. 2015). Assessing the diversity of molecular

217   compounds with nutrients such as P, N or S and understanding how the elemental composition

218   in aerosols shifts due environmental changes isof special interest for ecological stoichiometry

219   research and it is, therefore, feasible with DI-FT-ICR-MS.

220   Visualization of DI-FT-ICR-MS data using van Krevelen diagrams (vK) (O:C vs. H:C ratios

221   of the compounds assigned with an elemental formula) has been commonly used for analyzing

222   complex organic matrices (Kim et al. 2003; Schmitt-Kopplin et al. 2012; van Krevelen 1950). vK

223   diagrams provide information of the main chemical reactions such as demethylation,

224   methylation, hydrogenation, condensation, hydration, reduction or oxidation occuring in the

225   samples (Kim et al. 2003). Additionally, plotting O:C vs. H:C ratios of the assigned compounds

226   to elemental formulas can also provide an approximation of the main compound categories of

227   the analyzed samples (e.g. lipids, protein, carbohydrates, etc.) (Kim et al. 2003; Minor et al.

228   2014; Sleighter and Hatcher 2007). However, several compounds in the environment can be

229   transformed or degraded, and thus changing their original O:C and H:C ratios (Rivas-Ubach et

230   al. 2018b). Consequently, while this classification can still provide a general idea of the organic

231   compound compositions in aerosol particles, any compound classification based on

232   stoichiometric constraints should be used with caution and multidimensional stoichiometric

233   compound classification approaches could be considered (Rivas-Ubach et al. 2018b). C number

234   versus mass graph (CvM) is also commonly used to represent DI-FT-ICR-MS data in

235   comparative studies providing important information on oxidation processes or molecular

236   weight shifts (Reemtsma 2009). DI-FT-ICR-MS represents thus a useful tool to obtain high-

237   resolution metabolomic fingerprints and to gain a better comprehension of potential chemical

238   transformations of samples.

239   For more details in metabolomics technologies and their applications, several review

240   articles with special focus on metabolomics technologies have been published (Rochford 2005;

241   Fukusaki and Kobayashi 2005; Shulaev 2006; Lindon et al. 2007; Saito and Matsuda 2010; Lei et

242   al. 2011; Zhang et al. 2012; White et al. 2017; Rivas-Ubach et al. 2018b; Azad and Shulaev

243   2018).

244

245   **1.3 Testing metabolomics techniques in aerosol particle samples.**

246 The techniques to sample and characterize the gas phase of atmo-metabolomes have been
247 already described elsewhere (Smith and Španěl 2011; Tholl et al. 2006). The main aim of this
248 article is to describe a step-by-step method for sampling and characterizing the particle phase
249 of the atmo-ecometabolomes through distinct mass spectrometry techniques. We collected
250 aerosol particles without size cutoff during two distinct seasons (spring and summer) in an
251 ecosystem with low biological activity by following a simple methodology. Unlike traditional
252 aerosol particle studies concerning atmospheric chemistry, size cutoff of particles was avoided
253 in this method because aeresol particles have a broad range of sizes and size exclusion filtering
254 methods employed in atmospheric studies will not yield informative results for ecology
255 studies. This sampling allowed testing how our method perform in ecosystems with low
256 biological activity by detecting significant changes in the composition of aerosols between
257 seasons. Additionally, we further validate the robustness of our method by analyzing aerosols
258 samples collected during three different peak periods of pollination within the same season in
259 a more biologically active area.

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

270

271

272 **2. Experimental details.**

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

278

279 **2.1 Study sites.**

280 The proposed method is demostrated with the aerosol particle samples collected from two

281 distinct seasons (spring and summer) in 2015 from the Pacific Northwest National Laboratory

282 (PNNL) located in the north side of the city of Richland (Washington, USA) (46° 34´ N, 119° 28´

283 W). Additional samples were collected in 2017 at the University of Michigan (UMICH; Ann

284 Arbor, MI, USA) to validate further our methodology by characterizing aerosol particle

285 signatures from different tree pollination peak periods sampled within a single spring season.

286 UMICH is located in the north-east side of the city of Ann Arbor (Michigan, USA) (42° 29´ N, 83°

287 70´ W).

288 *PNNL site (spring vs. summer):* PNNL nearby landscape is represented by an extensive

289 desert covered mainly by steppes, shrubs and herbaceous species with *Purshia tridentate,*

290 *Ericameria nauseosa, Grayia spinose, Chrysothamnus viscidiflorus, Artemisia tripartita, Salsola*

291 *tragus, Sarcobatus vermiculatous*, and *Tamarix romosissima* as some of the most represented

292 plant species. PNNL campus is covered by grass with planted non-authoctonous tree species

293 such as *Platanus sp.* The population of the surrounding metropolitan area is of about 250,000

294 and the economy and land use are mainly dominated by diverse crops and the nuclear

295 reservation of Hanford. The climate at PNNL site is semi-arid desert with an averaged yearly

296 rainfall ranging between 180 and 220 mm. Averaged maximum annual temperatures are

297 around 32°C, with sporadic peaks reaching 42-45°C. Averaged minimum annual temperatures

298 are around -2°C with sporadic tempeatures reaching -20°C.

299 *UMICH site (pollination peak periods):* The surrounding areas are represented by

300 mosaic landscapes mainly covered by urban, agricultural and extended forested areas

301 dominated by several species of maple trees (*Acer sp.*), oak (*Quercus sp.*), honeylocust

302 (Gleditsia triacanthos) and, *Betula alleghaniensis, Fagus grandiflora, Populus tremuloides* and,

303 some populations of *Pinus strobus*. Ann Arbor is part of the metropolitan area of Detroit with a

304 population of over 4 million. The climate is humid continental with strong influence from the

305 Great Lakes with a mean annual precipitation of 950 mm per year. Average maximum annual

306 temperature is around 28°C, with peaks reaching up to 41°C and the average minimum annual

307 temperature is 4.7°C.

308

309 **2.2 Aerosol particle sampling.**

310 For the particle phase aerosol collection, we designed a portable and simple system that

311 allows simultaneous collection of multiple aerosol particle samples (Figure S1). For aerosol

312 particle collection, we used high-purity quartz filters (Whatman QM-A 37mm, Whatman

313 International Ltd, Maidstone, UK) precombusted for 5hrs at 450°C for impurity elimination

314 (Schmitt-Kopplin et al. 2012). Each sample was collected by placing a precombusted quartz

315   filter into a filter cassette. Flexible PVC tubing of 0.6 cm diameter was used to connect the

316   pump with the filter cassetes. For each site (PNNL and UMICH), the pump operated during 18

317   consecutive hours (09:00am to 03:00am of the following day) at a flow rate of 30 L per minute

318   at the sampling point (filter). Filters were maunally replaced daily at ~08:30am during the

319   sampling periods. Filter samples were thus stored at -80°C until metabolite extraction.

320   *PNNL site sampling (spring vs. summer):* We collected aerosols during 14 consective

321   days in spring 2015 (May 7 to May 20; spring samples) and for 16 consecutive days in summer

322   2015 (July 15 to July 30; summer samples) (Figure 3). For each day, two samples were collected

323   simultaneously at an 8-meter tower at PNNL campus. According to the US National Weather

324   Service at the Pasco airport (KPSC); daily averaged temperatures ranged from 11 to 21°C

325   (averaged maximum of 14-29°C) for the spring aerosol collection period (May). Daily averaged

326   temperatures for the summer aerosol collection period (July) ranged between 19 and 29°C

327   (averaged maximum of 28-40°C). Humidity ranged from 49 to 78% (daily average)(averaged

328   maximum of 72-100%) for spring sampling period and from 35 to 50% (daily average)(averaged

329   maximum of 57-86%) for summer sampling period. Total precipitation for May and July

330   sampling periods was 28.2 mm and 0 mm, respectively.

331   *UMICH site sampling (pollination peak periods):* Additional aerosol samples collected

332   at UMICH campus were collected from 24th to 26th April (birch pollination peak), from 10th to

333   12th May (oak pollination peak), and from 30th May to 1st June 2017 (pine pollination peak)

334   (Figure 3). One filter was collected daily during 3 consecutive days for each pollination peak

335   period from a rooftop location. According to weather conditions reported by the US National

336   Weather Service at the local airport (KARB), temperature averages were 15.5°C, 12.2°C, 15°C

337   for sampling periods corresponding to the birch, oak, and pine pollination peaks, respectively.

338   Average humidity ranged between 55-74%, 53-66%, and 59-74% for the sampling periods

339   corresponding to the birch, oak, and pine pollination peaks, respectively. Accumulated

340   recorded precipitation was 3.3 mm, 3.3 mm and 0.3 mm for the sampling sampling periods

341   corresponding to the birch, oak, and pine pollination peaks, respectively.

342   *Sampling for sonication time test (PNNL site):* The extraction of compounds from the

343   filters was sonication-based. Thus, we collected additional aerosol samples in spring 2015 at

344   PNNL campus for different sonication time testing in GC-MS and LC-MS platforms. Sonication

345   test was not performed in DI-FT-ICR-MS since peak intensity measured with such a method is

346   only semi-quantitative (Kujawinski 2002; Liu and Kujawinski 2015).

347   During two consecutive days, aerosols were simultaneously collected in 3 independent

348   filters (hereafter test-filters) at a flow rate of 30L per minute at the sampling point. The pump

349   operated daily from 09:00am to 03:00am of the following day (18 hours × 2 days = 36 hours

350 per filter). A total of 6 rounds of test-filters were collected (6 rounds × 3 filters = 18 filters in

351 total). Sampling of the test-filters was performed from June 5 to June 16 (12 consecutive days)

352 (Figure 3). Test-filters were stored at -80°C until extraction (sonication test details are shown in

353 the supplementary text of the supporting information).

354

355 **2.3 Metabolite extraction for mass spectrometry analysis.**

356 Three different tube sets were used; set of tubes A, B and C (Figure 4). The number of tubes in

357 each set was the same than the number of samples collected. Set A (8 mL glass tubes) were

358 used to perform the extractions. Set B (15 mL polypropylene centrifuge tubes) were to keep

359 the extracts. Set C tubes (2mL HPLC glass vials) was used to keep the concentrated extract.

360 Glass tubes were combusted at 450°C for 5 hours before use. Each sample (filters) was rolled

361 (Figure 4.1) and inserted into the corresponding tube of set A (Figure 4.2). Subsequently, each

362 tube from the set A received 5 mL of H_2O/MeOH (20:80) (this volume of extract may vary

363 depending on the used filter diameter and the diameter of the set A tubes; filters have to be

364 totally covered by the extract solution) (Figure 4.3). Samples were then sonicated at 24°C for

365 10 min (Figure 4.4) (10min was the established time after performing the sonication time tests,

366 see Supporting Information). Four mL of the extract from each tube of the set A were

367 transferred to the corresponding tube of set B (Figure 4.4.1). Two extractions were performed

368 to each sample by repeating the same procedures but adding 4 mL of H_2O/MeOH (20:80)

369 instead of 5mL as fresh extract for the second extraction. The resulting extracts from the

370 second extraction were then combined the first extracts (in tubes of set B) (Figure 4.5.1). All

371 extracts (in set B tubes) were evaporated using an ultra-high purity nitrogen evaporator

372 (Figure 4.6). Subsequently, 1 mL of fresh solvent (H_2O/MeOH (20:80)) was added to each set B

373 tube and tubes vortexed for 30 s for a proper resuspension of the dried extracts (Figure 4.7).

374 All concentrated extracts (in set B tubes) were centrifuged at 4,500 x g for 5 min (Figure 4.8)

375 and supernatants were transferred into the corresponding tubes of the set C (HPLC

376 vials)(Figure 4.9). Extracts were stored at -80 ºC until MS analyses (Figure 4.10). Experimental

377 blanks (not used combusted filters) were also extracted and analyzed with each of the MS

378 platforms.

379   Preliminary aerosol particle samples collected at PNNL for sonication tests were

380 analyzed by LC-MS and GC-MS. LC-MS and GC-MS provide accurately relatively quantitative

381 data making them suitable for this sonication testing. Spring and summer samples collected at

382 PNNL campus were analyzed by LC-MS, GC-MS and DI-FT-ICR-MS (Figures 3 and 4.11).

383 Additional samples collected at the UMICH campus to contrast different tree pollination peak

384 periods were analyzed exclusively by LC-MS and GC-MS (Figures 3 and 4.11). For DI-FT-ICR-MS

385 and LC-MS plaforms, extracts (~200 µL) were directly used for analyses (Figure 4.12). GC-MS

386 required a pre-treatment of the samples prior to the instrumental analyses.

387   For metabolite derivatization for GC-MS analyses, we followed a protocol described

388 elsewhere (Kim et al. 2015). Briefly, for each sample, 500 µL of extract from the set of tubes C

389 (Figure 4.10) was transferred into a clean HPLC vial and dried down in a vacuum evaporator

390 centrifuge. Then, each sample received 20 µL of methoxyamine in pyridine solution (30

391 mg/mL). Samples were vortexed for 30 s and subsequently incubated for 90 min at 37°C and

392 1,000 rpm in a Thermomixer (Eppendorf AG, Hamburg, Germany) to protect carbonyl groups.

393 After the first incubation, all vials were centrifuged for 15 s and 80 µL of N-methyl-N-

394 (trimethylsilyl)trifluoroacetamide (MSTFA) with 1% trimethylchlorosilane (TMCS) was added to

395 each sample. All vials were then vortexed for 10 s and incubated for 30 min at 37°C and 1,000

396 rpm to derivatize caboxyl, hydroxyl, and amine functional groups. After the second incubation,

397 all samples were centrifuged for 5 min at 4,500 x g and derivatized supernatants were

398 transferred into clean glass vials with 200 µL inserts by using glass Pasteur pipettes. All vials

399 were finally tightened with caps with septum.

400

401 **2.4 LC-MS analyses and chromatogram processing.**

402 A Vanquish ultra-high pressure liquid chromatographer (UHPLC) coupled to a high-resolution

403 mass spectrometer LTQ Orbitrap Velos (HRMS) equipped with a heated electrospray ionization

404 (HESI) source (Thermo Fisher Scientific, Waltham, Massachusetts, USA) was used to obtain the

405 LC-MS chromatograms of aerosol samples. LC was performed with a reversed-phase C18

406 Hypersil gold column (150 × 2.1 mm, 3µ particle size; Thermo Scientific, Waltham,

407 Massachusetts, USA) operating at 30 °C. and at a constant flow rate of 0.3 mL per minute.

408 Chromatography mobile phases consisted of 0.1% formic acid in water (A) and

409 acetonitrile:0.1% formic acid in water (90:10) (B). The injection volume was set at 5 µL.The

410 elution gradient initiated at 90% A (10% B) and hold for 5 min, then the gradient ramped

411 linearly to 10% A (90% B) during the next 15 min. Those conditions were maintained for 2 min

412 before the initial mobil phase proportions (90% A; 10% B) were recovered over the next 2 min.

413 The column was washed and stabilized for an additional 11 minutes at the initial

414 chromatographic conditions (90% A; 10% B). All samples were analyzed in positive (+) and

415 negative (-) electrospray ionization modes. The HRMS operated in FTMS (Fourier Transform

416 Mass Spectrometry) and full-scan mode at high resolution (60,000) and acquired in the mass

417 range of 50-1000 m/z. An experimental blank was injected every 15 samples. Blanks were used

418 to determine the chromatographic background during the dataset filtering. A mixture of

419 standards were injected every 25 samples to test for mass accuracy and sensitivity of the

420 instrument. Two injections of fresh H_2O/MeOH (20:80) were analyzed right after the standard

421 mixture for column washing purposes and avoid any potential compound carry over.

422 MZmine 2.17 (Pluskal et al. 2010) was used to process the LC-MS RAW files.

423 Chromatograms obtained in positive and negative modes were treated separately.

424 Chromatograms were baseline corrected, deconvoluted, aligned, and "gap-filled" and

425 metabolite identifications were assigned. Datasets were thus exported to CSV format files. The

426 parameters used for generating the LC-MS datasets are detailed in Table S1. The same MZmine

427 parameters were used to generate all datasets spring vs. summer LC-MS dataset (PNNL) and

428 pollination peak period LC-MS dataset (UMICH).

429 Metabolite assignation for LC-MS files was performed by matching the exact mass and

430 RT of the detected features with the corresponding values of our metabolite library including

431 more than 600 metabolites present in organisms commonly present in plants. According to

432 Sumner et al. (Sumner et al. 2007), our LC-MS metabolite assignment is putative since it was

433 based on the measured exact mass of metabolic features and their RT. However, the use of

434 both RT and the high MS resolution achieved with Orbitrap technology reduces substantially

435 the number of false positive assignations. For more detailed information in relation to

436 metabolite library matching see Rivas-Ubach et al. (Rivas-Ubach et al. 2016c). Metabolite

437 matching information (RT and m/z values) for LC-MS files are detailed in Table S2.

438

439 **2.5 GC-MS analyses and chromatogram processing.**

440 Derivatized samples were analyzed by an Agilent GC 7890A coupled with a MSD 5975C mass

441 spectrometer (Agilent Technologies, Santa Clara, CA). An HP-5MS column (30 m × 0.25 mm ×

442 0.25 µm; Agilent Technologies) was used for the GC. The injection port temperature was held

443 at 250 °C. The injecton mode was split-less. The oven with the column was maintained for 1

444 min at 60°C and then ramped at 10°C per minute to 325°C (26.5 min ramp) and hold for 10

445 min. A mixture of fatty acid methyl esters (FAMEs; C8-C28) for retention index calculation (RI)

446 was analyzed prior to sample injections. Experimental blanks were analyzed every 15 samples

447 to determine the chromatographic background.

448 MZmine 2.17 (Pluskal et al. 2010) was exclusively used to generate the metabolomic

449 fingerprint datasets for the test-filters collected at PNNL to be consistent with the LC-MS

450 datasets for the sonication test (Figure 3). Parameters implemented to generate the datasets

451 from the test-filters with MZmine are described in Table S3. Metabolite Detector 2.5 (Hiller et

452 al. 2009) was used to generate the GC-MS datasets for the ambient samples collected at PNNL

453 (spring vs. summer) and UMICH (pollination peak periods) sites (Figure 3). Agilent ".D" files

454 directly obtained from the MS were converted to "netCDF" format with Agilent Chemstation

13

455   software. Subsequently, Metabolite Detector converted "netCDF" files to "bin" files and

456   chromatograms were thus deconvoluted, aligned according to the retention indices (RI) and

457   metabolic features were were assigned to metabolites. Metabolites were identified by

458   matching measured mass spectra to a PNNL improved version of FiehnLib (Kind et al. 2009),

459   containing validated RIs and mass spectra for over 850 metabolites. Metabolite matching

460   probability threshold was set at 0.65. Datasets were finally exported to CSV format files.

461   Metabolite identifications were manually validated by matching measured mass spectra with

462   mass spectra from NIST14 GC-MS library to avoid false identifications and reduce any allegued

463   deconvolution errorduring data-processing. Table S4 includes all the parameters used in

464   Metabolite Detector to generate the GC-MS datasets from the aerosols collected at PNNL and

465   UMICH sites. Table S5 shows metabolite matching information for GC-MS.

466

467   **2.6 DI-FT-ICR-MS analyses and spectra processing.**

468   DI-FT-ICR-MS analyses were performed exclusively on the spring and summer samples

469   collected at PNNL. Extracts of aerosols were analyzed on a 12 Tesla Bruker SolariX FT-ICR-MS

470   (Bruker daltonics Inc, Billerica, MA, USA). At a flow rate of 3.0 µL/min, aerosol extracts were

471   directly infused into the FT-ICR-MS using an Agilent 1200 series pump (Agilent Technologies,

472   Santa Clara, CA, USA). Compounds were ionizated by a standard Bruker electrospray (ESI)

473   operating in negative mode and equipped with a fused silica tube (30 µm i.d.). Ion

474   accumulation time was optimized for each sample (0.1s). FT-ICR-MS operated at a resolution

475   of 4 MWord, equilavent to a resolving power of 400,000 ($m/\Delta m_{50\%}$ at m/z 400). Experimental

476   conditions were as follows: needle voltage was set at +4.4 kV; Q1 set to 50 m/z; and the

477   heated resistively coated glass capillary was maintained at 180 °C. Solvent blanks were run

478   every 10 samples.

479   A total of 144 individual scans from 100 to 1100 m/z were averaged to produce the

480   final DI-FT-ICR-MS spectrum for each sample. The instrument achieved mass measurement

481   accuracy with < 1 ppm error after external calibration with a standard mixture. All detected

482   ions in the spectra were singly charged as confirmed by the 1.0034 Da spacing between

483   isotopic forms of the same compound (i.e., between $^{12}C_n$ and $^{12}C_{n-1}$–$^{13}C_1$). Each spectrum raw

484   file was converted to a list of m/z values with DataAnalysis software (BrukerDaltonik version

485   4.2) with the following settings: an absolute intensity threshold of 100 and a signal to noise

486   (S/N) of 7, which is above the minimum detection limit for FT-ICR-MS for natural organic

487   matter (Riedel and Dittmar 2014). The exported peak lists were then internally calibrated using

488   an organic matter homologous series separated by 14 Da (–CH_2 groups) using the Formularity

489   software (Tolić et al. 2017). After internal calibration, the mass error was < 0.5 ppm. The

490   calibrated peak list was then used to assign with elemental formulas (C, H, O, N, S, and P) using

491   a module within Formularity, based on the compound identification algorithm described

492   elsewhere (Kujawinski and Behn 2006).

493

494   **2.7 Datasets filtering.**

495   A total of 7 distinct datasets were generated: 3 contrasting spring vs. summer aerosols (LC-MS,

496   GC-MS and DI-FT-ICR-MS; collected at PNNL), 2 comparing three different pollination peak

497   periods (LC-MS and GC-MS; collected at UMICH) and 2 for sonication tests (LC-MS and GC-MS;

498   collected at PNNL). LC-MS and GC-MS datasets were separately filtered through 5 main steps:

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

500   2) For each variable (metabolite feature), outlier values were determined for each season

501   separately and replaced for NA. Outliers were defined as:

502    $$Upper\ Outliers \rightarrow value > Q75 + 2.5 \times IQR$$

503    $$Lower\ Outliers \rightarrow value < Q25 - 2.5 \times IQR$$

504   where Q75 represents the third quartile, Q25 represents the first quartile and IQR is the

505   interquartile range (IQR = Q75-Q25).

506   3) Variables (metabolite features) with less than 50% of data within all factor levels of the

507   dataset (spring and summer for datasets from PNNL aerosol particles; birch, oak and

508   pine pollination peak periods from UMICH aerosols) were removed.

509   4) Variables (metabolite features) with signal to noise lower than 6 were removed from

510   the dataset. Noise level was determined by the experimental blanks run during the

511   analytical sequence.

512   5)   To avoid statistical artifacts in multivariate analyses due sample variability, the values

513   for each variable (metabolite feature) and sample were scaled by the total intensity of

514   its chromatogram (LC-MS and GC-MS). This scaling allows comparing the metabolic

515   function between groups of samples independently of the amount of organic matter.

516   It should be noted that DI-FT-ICR-MS data should go through rigorious data quality

517   check before further interpretation, to avoid false discovery. For DI-FT-ICR-MS dataset (PNNL;

518   spring vs. summer), mass features detected from solvent blanks were removed from all the

519   samples. Additionally mass features observed in less than 70% of replicates within all factor

520   levels were also be removed to avoid consideration of noise peaks. Finally, for robust data

521   interpration based on elemental formulas, we only used the assigned formulas with less than

522   0.3ppm of error to be conservative; as cutoff values up to 0.5ppm of error are typically used

523   (Osterholz et al. 2016; Rivas-Ubach et al. 2018b).

524   After filtering all databases, our spring vs. summer databases were finally composed

525   by 1,333, 148, and 3,567 metabolomic features detected by LC-MS, GC-MS, and DI-FT-ICR-MS,

526   respectively. For LC-MS and GC-MS datasets, a total of 18 and 15 metabolites were identified,

527   respectively. The pollination peak periods databases were finally composed by 7,832 and 221

528   metabolomic features detected by LC-MS and GC-MS, respectively, with a total of 45 and 20

529   identified metabolites, respectively.

530

531   **2.8 Statistical analyses.**

532   While the values obtained from the deconvoluted peaks in LC-MS and GC-MS do not represent

533   an absolute concentration (e.g., mg of metabolite per weight of sample), they represent the

534   relative abundance and are suitable for metabolomic comparative analyses as proven in

535   previous studies (Gargallo-Garriga et al. 2015; Lee and Fiehn 2013; Leiss et al. 2013; Mari et al.

536   2013; Rivas-Ubach et al. 2014, 2016a, 2018). We used then the term *relative abundance* along

537   the article as the relative concentration of metabolic features or metabolites. DI-FT-ICR-MS

538   data is not directly quantifiable (Wozniak et al. 2008), and although not as robust as GC-MS or

539   LC-MS techniques, using the feature intensity can still provide a valuable proxy of their relative

540   abundance (Kellerman et al. 2014; Spencer et al. 2015).

541   All LC-MS and GC-MS datasets, independently if generated from PNNL (spring vs.

542   summer) or UMICH (pollination peak periods) aerosols, were submitted to the same statistical

543   analyses. First, the aerosol metabolome fingerprints obtained from LC-MS and GC-MS were

544   submitted to PERMANOVAs using the Euclidean distance to test for overall metabolomic

545   differences between aerosols from distinct seasons (PNNL site), and between different

546   pollination peak periods (UMICH site) (Table 1). The permutations for the test were set at

547   10,000.

548   LC-MS and GC-MS metabolomic fingerprints were also subjected to principal

549   component analysis (PCA) to graphically represent the natural variability among the analyzed

550   samples reduced in two dimensions (Principal Component (PC) 1 vs. PC2) (Kim et al. 2010; van

551   den Berg et al. 2006) (Figure 5).

552   A heat map was plotted for each entire LC-MS and GC-MS metabolomic fingerprints

553   from the aerosols collected at PNNL and UMICH sites (Figure 6). Additional heat maps were

554   plotted only for the identified features from the spring vs. summer and for the pollination peak

555   periods datasets (Figure 7). All the identified LC-MS and GC-MS features for the spring vs.

556   summer dataset (PNNL) were submitted to Student t-test with season as the categorical factor

557   to assess for statistical significance between spring and summer seasons (Table S6). One-way

558   ANOVAs followed by Tukey HSD posthoc tests were performed to asses for significant

559   differences between pollination peak period samples (birch vs. oak vs. pine). (Table S7).

560   For the DI-FT-ICR-MS dataset generated from the PNNL aerosols (spring vs. summer),

561   we counted the proportions of formula classes (CHNO, CHO, CHOS, CHNOS, CHNOP, CHOSP,

562   CHNOSP, and CHOP) that were assigned to each sample. Proportions were then transformed

563   using *arcsin (rootsquare)* and subsequently submited to Student t-tests with season (spring

564   and summer) as the categorical factor to assess for statistical significance (Figure 8). For each

565   feature detected by DI-FT-ICR-MS, the intensity difference between spring and summer was

566   calculated and used in the C number of each feature vs. m/z (CvM) (Figure 9a) and in the O:C

567   ratio vs. m/z (Figure 9b) plots. Additionally, a Student t-test was performed on the O:C ratios of

568   the detected features with season as the categorical factor to determine whether the

569   oxidation status of the molecular compounds statistically changed significantly between spring

570   and summer aerosols (Figure 9c).

571   All statistical analyses were performed with R (R Core Team 2013). PERMANOVAs were

572   conducted with the function *adonis* from the "vegan" package (Oksanen et al. 2013). PCAs

573   were plotted using the *PCA* function from "FactoMineR" package (Husson et al. 2016) with the

574   missing data from the dataset imputed using *imputePCA* function from package "missMDA"

575   (Husson and Josse 2015). The data matrix was scaled before the PCA by setting SCALE as TRUE

576   from the *PCA* function. Heat maps were plotted using the function *heatmap.2* from "gplots"

577   package (Warnes et al. 2016). Student t-tests and oneway ANOVAs were performed with the

578   functions *t.test* and *aov*, respectively, included in the "stats" package (R Core Team 2013).

579   Tukey HSD posthoc tests were performed with the *HSD.test* function in the "agricolae"

580   package (de Mendiburu 2015). All graphs were first plotted in R and subsequently treated with

581   Adobe Illustrator CS6. Analyses and results from the different sonication times during

582   metabolite extraction are detailed in the supporting information (Supplementary Text).

583

584   **3. Results and discussion.**

585   **3.1 Aerosol sampling.**

586   Optimal flow rate for aerosol particles collection is important, as low flow rate may not collect

587   enough aerosol particles and excessive flow rate may damage the filters. After coupling the

588   tubing with the pump and connected two filters, we achieved constant flow rates of 30 L/min

589   at the aerosol sampling point. However, flow rates of ~50 L/min at the aerosol collection point

590   performed properly in 37mm quartz filters without collapsing. Different filter materials, such

591   as polytetrafluoroethylene (PTFE), may present distinct resistances to high flow rates. It is

592   important to note that the internal air friction associated with the length of the tube and the

593   sampling of multiple filters simultaneouly can substantially decrease the flow rates at the

594   aerosol sampling point. Larger tube diameters (>0.6cm) could be used if higher flow rates at

595   the sampling point are necessary, especially when sampling in ecosystems with low biological

596   activity.

597   For statistical puproses, our simple sampling methodology allows collecting different

598   number of biological replicates at once. Furthermore, aerosol collection can be easily

599   performed at different heights by extending tubing if the pump performance is sufficient as to

600   ensure acceptable flow rates at the aerosol collection point. The experimental design (number

601   of replicates, filter material, length and diameter of tubing) and the pump performance are

602   thus key elements to consider in atmo-ecometabolomics research.

603   Aerosol particles collection was performed in two distant locations with distinct

604   environments; PNNL and UMICH. PNNL site is semi-urban area surrounded by diverse

605   croplands and a large desert with relatively low biological activity. For this reason we expected

606   to detect a complex variety of compounds which difficult the interpretation of the data at

607   regional scale. Even so, the results obtained contrasting spring vs. summer aerosols were

608   equally valid as to describe in detail the necessary steps to obtain accurately the atmo-

609   metabolomes and to test the sensitivity of each mass spectrometry platform (LC-MS, GC-MS,

610   DI-FT-ICR-MS) by assessing their potential for detecting statistically significant changes

611   between the atmo-ecometabolomes of two contrasted seasons of ecosystems with low

612   biological activity. At the same time, UMICH location provided additional data collected in a

613   more biologically diverse and active ecosystem to further evaluate the sampling and the

614   analytical methodologies.

615

616   **3.2 Metabolite extraction in organic solvents.**

617   Organic solvents in combination with water are commonly used for metabolite extractions

618   allowing a good extraction range of polar, semi-polar and non-polar metabolites (Kim et al.

619   2010; Lin et al. 2006; Rivas-Ubach et al. 2013; t'Kindt et al. 2008). Although different solvents

620   recover different matrices of compounds with distinct polarities, water/methanol (20:80)

621   solution has been widely used in metabolomics studies showing a wide recovery of polar and

622   semi-polar metabolites (t'Kindt et al. 2008). Solvents such as acetonitrile, methanol, and

623   chloroform could leach plastics especially during sonication, as such MS spectra may show

624   contaminant features derived from plastics (Figure S2). Therefore, the use of silanized glass

625   tubes is advised during sonication (Figure 4.4). Combusting glassware for 5-6 hours at 450ºC or

626   higher is also recommended to prevent organic contaminations. If plastic tubes are used for

627 metabolite extraction, especially during the sonication step (set of tubes A), initial tests are

628 recommended to identify any potential contaminant in the extracts. Due the relatively low

629 metabolite concentration in aerosol samples, each sample was extracted twice in

630 water/methanol (20:80) to increase metabolite recovery from aerosols (Böttcher et al. 2007;

631 Nikiforova et al. 2005; Rivas-Ubach et al. 2013) (Figure 4.5).

632   We did not detect overall significant variation in the relative abundance of the

633 observed features between test-filters under different sonication times (Tables S8 and S9) with

634 the water/methanol (20:80) extraction; only days 5th & 6th June showed significant differences

635 between different sonication times (F= 3.21, *P* = 0.02) in the GC-MS analyses (Table S9). We

636 thus considered that 10 minutes of sonication was enough as to extract the metabolites from

637 aerosol samples in water/methanol (20:80) solution.

638   Filter size is also a crucial factor to consider in atmo-ecoecometabolomic reseach,

639 especially during the extraction procedures. Extracts of aerosol particles will be more

640 concentrated as lower is the *filter size/pump flow rate* ratio. Furthermore, smaller filters are

641 easier to handle during extractions and allow higher recovery of the extracts. Quartz filters

642 absorb high volumes of solvent that cannot be easily recovered, even if samples are

643 centrifuged at high rpm. We could recover the 89% of the initial solvent volume with 37mm

644 diameter filters. Larger filters complicate significantly the extraction of metabolic compounds

645 (more filter handling, larger tubes and volumes of extract are required) and decrease

646 remarkably the extract recovery due to the large absorption of solvent. For studies using FT-

647 ICR-MS, a final dissolved organic carbon (DOC) concentration of 15 mg of C per L in 1:1

648 methanol/water in samples has been proven to provide suitable signal for detection of

649 compounds (Medeiros et al. 2017). However, higher concentration of C may be required for

650 other mass spectrometry platforms depending on their sensitivity. Analyses by LC-MS and GC-

651 MS techniques separate the analytes through chromatography; therefore, the amount of C

652 analyzed in each scan is diluted and higher amount of DOC allows a better detection of organic

653 compounds and metabolome characterization.

654

655 **3.3 Testing atmo-ecometabolomics contrasting two distinct seasons (PNNL site)**

656 Even collecting aerosols in low biologically productive ecosystems, PERMANOVAs showed

657 overall significant differences between spring and summer aerosols for the fingerprints

658 obtained with both LC-MS and GC-MS platforms (Pseudo-F = 13.8, *P* < 0.0001 and Pseudo-F =

659 9.69, *P* < 0.0001; respectively) (Table 1).

660   PCAs plotted for both LC-MS and GC-MS figerprtings also showed clear separation

661 between spring and summer aerosol samples (Figure 5a and 5b). The PC1 and PC2 of the PCA

662   performed with LC-MS dataset explained the 37.7% and 14.81% of the total variation among

663   samples, respectively. The PC1 and PC2 of the GC-MS PCA explained the 39.72% and 17.2%,

664   respectively, of the total variance. Both PCAs showed similar explained variability for the PC1

665   (37.7% for LC-MS and 39.72% for GC-MS) that primarily segregated spring from summer

666   samples. The results from PERMANOVAs and PCAs indicate, indeed, that the metabolomic

667   fingerprints from both seasons are significantly distinct. Even so, it is important to note that

668   each analytical technique (LC-MS and GC-MS) provide different complementary information

669   (Ding et al. 2007; Zhang et al. 2012).

670   Heat maps of the entire LC-MS and GC-MS metabolomic fingerprints showed large

671   metabolomic variability between aerosol samples within the same season (Figure 6a and 6b).

672   Even so, we still identified several clusters of metabolic features (marked in blue) presenting

673   higher relative abundances in spring or summer aerosol samples. Those heat maps are clearly

674   in accordance with the PERMANOVAs and the PCA by showing major metabolic differences

675   between spring and summer aerosols supporting thus our protocol as sensitive enough as to

676   differentiate the overall aerosol composition between two seasons in a complex and low

677   biologically active environment.

678   Several of the identified metabolites with LC-MS and GC-MS showed statistical

679   significance between seasons after Student t-tests (Figure 7a and Table S6). LC-MS analyses

680   showed spring aerosol particles with significantly higher relative abundance ($P < 0.05$) of α-

681   ketoglutaric acid, acacetin, adenosine, adonitol-ribitol, chrysin, citric acid, glutamine, hexoses,

682   isoleucine, shikimic acid and sorbitol-mannitol (Figure 7a). GC-MS dataset indicated that spring

683   aerosols presented significantly higher relative abundances of glucose and galactose ($P < 0.05$)

684   and summer aerosol particles showed significantly higher relative abundances of arachidic

685   acid, capric acid, fumaric acid, glyceric acid, linoleic acid, palmitic acid, stearic acid and uracil

686   (Figure 7a). Higher relative abundances of sugars and organic acids related to the Krebs cycle

687   such as citric and alpha-ketoglutaric can be realted to plant growth activity (Rivas-Ubach et al.

688   2012) and atmospheric pollination (Roulston and Cane 2000). Our results from the LC-MS and

689   GC-MS platfroms are in agreement with the DI-FT-ICR-MS dataset showing spring aerosols with

690   significantly ($P < 0.05$) higher proportions of CHOP and marginally significant ($P < 0.1$) CHNOSP

691   molecular formulas (Figure 8). High concentrations of sugars and phosphorus in biomass have

692   been previously associated to higher plant activity (Rivas-Ubach et al. 2012, 2014) although

693   sugars can also play multiple physiological functions in plants such as stress tolerance under

694   drought conditions (Ingram and Bartels 1996; Rivas-Ubach et al. 2014, 2016a). Several fatty

695   acid compounds, such as arachidic acid, capric acid, heptadecanoic acid, linoleic acid, oleic

696   acid, palmitic acid, and stearic acid, have been detected by GC-MS (Figure 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
(Figure 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 (Figure 8). In
addition, we generally observed higher relative intensities in high-mass features in summer
aerosols with respect to spring aerosols (Figure 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 similar range of carbon
numbers as those observed in spring, which were more abundant at lower molecular mass
(see region between dashed lines in Figure 9a). In addition, Student t-test on the O:C of the
formula-assigned features indicated that summer aerosol particles had significantlyhigher
relative intensities in features with higher O:C (more oxidized compounds) compared to spring
aerosols (Figure 9b, 9c). 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 (Figure 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 (Figure 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.

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

PERMANOVAs on the LC-MS and GC-MS fingerprintsshowed 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 (Figure 5c and 5d). In accordance with PERMANOVAs and PCAs, heatmaps of the

732   LC-MS and GC-MS aerosol fingerprints showed clear differences between pollination periods

733   (Figure 6c and 6d). These results indicate thus that the proposed atmo-ecometabolomics

734   methodology was sensitive enough as to accurately accurately extract, characterize and

735   distinguish aerosols collected during three different pollination peak periods (birch, oas and

736   pine) occurring within a continuous 40-day timeframe.

737   Compared to PNNL, the surroundings of UMICH have higher biological and human

738   activity and this difference was distinctly noticeable in the aerosol metabolomic fingerprints.

739   The number of features detected by both LC-MS and GC-MS instruments is considereably

740   larger in UMICH than PNNL aerosol fingerprints (1,333 vs. 7,832 for LC-MS fingerprints and 148

741   vs. 221 for GC-MS fingerprints) (Figure 6a vs. 6c and 6b vs. 6d). Furthermore, a larger number

742   of metabolomics features could be identified at the aerosols collected in UMICH (45 and 20 for

743   LC-MS and GC-MS, respetively) (Figure 7b). Then, those results suggest that the number of

744   detected and identified features in aerosol samples directly depends on the biological and

745   anthrophogenic activity of the surrondings. For this reason, it is crucial to keep relatively high

746   flow rates (≥25-30 L/min) at the aerosol collection point in order to characterize properly the

747   particle component of aerosol metabolomes, especially in ecosystems with low biological

748   activity.

749

750   **3.5 Conclusions and future perspectives.**

751   · Our methodology was sensitive enough as to detect significant overall differences between

752   atmo-ecometabolomes from aerosol particles sampled in different seasons (spring vs.

753   summer) in an ecosystem with low biological activity. Furthermoe, we detected clear changes

754   between the composition of aerosol particles collected during tree pollination periods. These

755   differences were apparent despite the high complexity of aerosol particle sources in the mix of

756   urban and vegetated areas.

757   · LC-MS and GC-MS showed similar accuracy in terms of sample clustering; both instruments

758   were able to detect clear overall differences between spring and summer aerosols and

759   between aerosols collected in different pollination peak seasons. Which instrument to use will

760   finally depends on the hypotheses of the study and whether a targeted approximation is

761   required. However, no analytical technique alone can still characterize the entire metabolome

762   fingerprint of a sample and combining both platforms is a common approach. In general, the

763   number of detected metabolic features is substantially higher in LC-MS analyses than GC-MS.

764   This is especially useful for overall multivariate analyses in non-targeted metabolomics studies

765   as LC-MS analyses provide a wider representation of the sample metabolome. Both

766   instruments detected several polar and semi-polar metabolites but GC-MS analyses allowed

22

767 the identification of some non-polar compounds in aerosols (e.g. linoleic acid, oleic acid,

768 palmitic acid, stearic acid). On the other hand, DI-FT-ICR-MS obtains ultra high-resolution

769 metabolomic data allowing the assignation of elemental formulas in the detected features in

770 aerosols. This data format allows performing different analyses on the distinct heteroatoms

771 (e.g. O, N, P) present in aerosols. This information is especially useful to study the nutrient

772 transport and deposition in ecosystems.

773 · Coupling environmental variables (e.g., temperature, wind, precipitation, etc.) with atmo-

774 ecometabolomics would allow a more precise interpretation of the aerosol-biosphere

775 interface.

776 · Long term atmo-ecometabolomic research in natural ecosystems would significantly improve

777 our understanding of interannual and seasonal changes of aerosol composition, directly linking

778 the composition of aerosols with plant phenology and physiology, along environmental

779 changes and/or natural spatial gradients.

780 · The use of atmo-ecometabolomic techniques in ecological sciences could improve the

781 detection, identification and quantification of molecular compounds directly related with

782 environmental stressors (biomarkers), providing thus important information of the general

783 status of the ecosystems. A good description of such biomarkers would allow generating

784 libraries of compounds that may serve to asses the ecosystem status or as an environmental

785 monitoring tool.

786 · A better understanding of the direct impacts of aerosols on biological surfaces (e.g.

787 phyllosphere) and on the overall C:N:P stoichiometry of ecosystems can be improved

788 significantly by the overall characterization of the aerosol chemical and molecular

789 composition.

790 · New modern high-resolution mass spectrometry instruments coupled to separation

791 techniques should be implemented in atmo-ecometabolomic research to enable high

792 performance for both mass accuracy and resolution, and retention time. Furthermore, recent

793 advances in metabolomics instruments, such as mass spectrometers coupled to Ion Mobility

794 Spectrometry (IMS-MS), could substantially enhance the number of detected metabolic

795 features in aerosol samples from tens and hundreds to thousands.

796

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825 **References.**

826 Achotegui-Castells, A., Sardans, J., Ribas, À., & Peñuelas, J. (2013). Identifying the origin of
827   atmospheric inputs of trace elements in the Prades Mountains (Catalonia) with
828   bryophytes, lichens, and soil monitoring. Environmental Monitoring and Assessment,
829   185(1), 615–629.

830 Andreae, M. 0, & Crutzen, P. J. (1997). Atmospheric Aerosols: Biogeochemical Sources and
831   Role in Atmospheric Chemistry. Science, 276(5315), 1052–1058.

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

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

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

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

840 Baustian, K. J., Cziczo, D. J., Wise, M. E., Pratt, K. A., Kulkarni, G., Hallar, A. G., & Tolbert, M. A.
841   (2012). Importance of aerosol composition, mixing state, and morphology for
842   heterogeneous ice nucleation: A combined field and laboratory approach. Journal of
843   Geophysical Research: Atmospheres, 117, D06217.

844 Böttcher, C., Roepenack-Lahaye, E. v., Willscher, E., Scheel, D., & Clemens, S. (2007). Evaluation
845   of Matrix Effects in Metabolite Profiling Based on Capillary Liquid Chromatography
846   Electrospray Ionization Quadrupole Time-of-Flight Mass Spectrometry. Analytical
847   Chemistry, 79(4), 1507–1513.

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

850 Canagaratna, M. R., Jayne, J. T., Jimenez, J. L., Allan, J. D., Alfarra, M. R., Zhang, Q., et al. (2007).
851   Chemical and microphysical characterization of ambient aerosols with the aerodyne
852   aerosol mass spectrometer. Mass Spectrometry Reviews, 26(2), 185–222.

853   Carlton, A. G., Pinder, R. W., Bhave, P. V., & Pouliot, G. A. (2010). To What Extent Can Biogenic
854   SOA be Controlled? Environmental Science & Technology, 44(9), 3376–3380.

855   Carnicer, J., Sardans, J., Stefanescu, C., Ubach, A., Bartrons, M., Asensio, D., & Peñuelas, J.
856   (2015). Global biodiversity, stoichiometry and ecosystem function responses to
857   human-induced C–N–P imbalances. Journal of Plant Physiology, 172, 82–91.

858   Claudino, W. M., Quattrone, A., Biganzoli, L., Pestrin, M., Bertini, I., & Di Leo, A. (2007).
859   Metabolomics: Available Results, Current Research Projects in Breast Cancer, and
860   Future Applications. Journal of Clinical Oncology, 25(19), 2840–2846.

861   de Mendiburu, F. (2015). agricolae: Statistical Procedures for Agricultural Research. R package
862   version 1.2-3, http://CRAN.R-project.org/package=agricolae.

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

866   Després, V. R., Alex Huffman, J., Burrows, S. M., Hoose, C., Safatov, A. S., Buryak, G., et al.
867   (2012). Primary biological aerosol particles in the atmosphere: a review. Tellus B,
868   64(15598), 1–58.

869   Ding, J., Sorensen, C. M., Zhang, Q., Jiang, H., Jaitly, N., Livesay, E. A., et al. (2007). Capillary LC
870   Coupled with High-Mass Measurement Accuracy Mass Spectrometry for Metabolic
871   Profiling. Analytical Chemistry, 79(16), 6081–6093.

872   Elser, J. J., Dobberfuhl, D. R., MacKay, N. A., & Schampel, J. H. (1996). Organism Size, Life
873   History, and N:P Stoichiometry. BioScience, 46(9), 674-684.

874   Fageria, N. K., Filho, M. P. B., Moreira, A., & Guimarães, C. M. (2009). Foliar Fertilization of
875   Crop Plants. Journal of Plant Nutrition, 32(6), 1044–1064.

876   Feng, J., Wang, Y., Zhao, J., Zhu, L., Bian, X., & Zhang, W. (2011). Source attributions of heavy
877   metals in rice plant along highway in Eastern China. Journal of Environmental Sciences,
878   23(7), 1158–1164.

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

881 Fiehn, O. (2002). Metabolomics - the link between genotypes and phenotypes. Plant molecular
882   biology, 48(1–2), 155–171.

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

885 Fuzzi, S., Andreae, M. O., Huebert, B. J., Kulmala, M., Bond, T. C., Boy, M., et al. (2006). Critical
886   assessment of the current state of scientific knowledge, terminology, and research
887   needs concerning the role of organic aerosols in the atmosphere, climate, and global
888   change. Atmospheric Chemistry and Physics, 6(7), 2017–2038.

889 Gargallo-Garriga, A., Sardans, J., Pérez-Trujillo, M., Oravec, M., Urban, O., Jentsch, A., et al.
890   (2015). Warming differentially influences the effects of drought on stoichiometry and
891   metabolomics in shoots and roots. New phytologist, 207(3), 591–603.

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

895 Gibney, M. J., Walsh, M., Brennan, L., Roche, H. M., German, B., & van Ommen, B. (2005).
896   Metabolomics in human nutrition: opportunities and challenges. The American journal
897   of clinical nutrition, 82(3), 497–503.

898 Glauser, G., Guillarme, D., Grata, E., Boccard, J., Thiocone, A., Carrupt, P.-A., et al. (2008).
899   Optimized liquid chromatography–mass spectrometry approach for the isolation of
900   minor stress biomarkers in plant extracts and their identification by capillary nuclear
901   magnetic resonance. Journal of Chromatography A, 1180(1), 90–98.

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

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

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

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

913 Hiller, K., Hangebrauk, J., Jäger, C., Spura, J., Schreiber, K., & Schomburg, D. (2009).
914   MetaboliteDetector: Comprehensive Analysis Tool for Targeted and Nontargeted
915   GC/MS Based Metabolome Analysis. Analytical Chemistry, 81(9), 3429–3439.

916 Hirai, M. Y., Yano, M., Goodenowe, D. B., Kanaya, S., Kimura, T., Awazuhara, M., et al. (2004).
917   From The Cover: Integration of transcriptomics and metabolomics for understanding
918   of global responses to nutritional stresses in Arabidopsis thaliana. Proceedings of the
919   National Academy of Sciences, 101(27), 10205–10210.

920 Husson, F., & Josse, J. (2015). missMDA: Handling Missing Values with Multivariate Data
921   Analysis. R package version 1.9., http://CRA.

922 Husson, F., Josse, J., Le, S., & Mazet, J. (2016). FactoMineR: Multivariate Exploratory Data
923   Analysis and Data Mining. R package version 1.32., http://CRA.

924 Ingram, J., & Bartels, D. (1996). The molecular basis of dehydration tolerance in plants. Annual
925   review of plant biology, 47(1), 377–403.

926 Jokinen, T., Berndt, T., Makkonen, R., Kerminen, V.-M., Junninen, H., Paasonen, P., et al.
927   (2015). Production of extremely low volatile organic compounds from biogenic
928   emissions: Measured yields and atmospheric implications. Proceedings of the National
929   Academy of Sciences, 112(23), 7123–7128.

930 Kaplan, F., Kopka, J., Haskell, D. W., Zhao, W., Schiller, K. C., Gatzke, N., et al. (2004). Exploring
931   the Temperature-Stress Metabolome of Arabidopsis. Plant physiology, 136(4), 4159–
932   4168.

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

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

939 Kim, H. K., Choi, Y. H., & Verpoorte, R. (2010). NMR-based metabolomic analysis of plants.
940   Nature protocols, 5(3), 536–49.

941 Kim, S., Kramer, R. W., & Hatcher, P. G. (2003). Graphical Method for Analysis of Ultrahigh-
942   Resolution Broadband Mass Spectra of Natural Organic Matter, the Van Krevelen
943   Diagram. Analytical Chemistry, 75(20), 5336–5344.

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

947 Kind, T., Wohlgemuth, G., Lee, D. Y., Lu, Y., Palazoglu, M., Shahbaz, S., & Fiehn, O. (2009).
948   FiehnLib: Mass Spectral and Retention Index Libraries for Metabolomics Based on
949   Quadrupole and Time-of-Flight Gas Chromatography/Mass Spectrometry. Analytical
950   Chemistry, 81(24), 10038–10048.

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

954 Kujawinski, E. (2002). Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass
955   Spectrometry (ESI FT-ICR MS): Characterization of Complex Environmental Mixtures.
956   Environmental Forensics, 3(3), 207–216.

957 Kujawinski, E. B., & Behn, M. D. (2006). Automated Analysis of Electrospray Ionization Fourier
958   Transform Ion Cyclotron Resonance Mass Spectra of Natural Organic Matter. Analytical
959   Chemistry, 78(13), 4363–4373.

960 Lee, D. Y., & Fiehn, O. (2013). Metabolomic response of Chlamydomonas reinhardtii to the
961   inhibition of target of rapamycin (TOR) by rapamycin. Journal of microbiology and
962   biotechnology, 23(7), 923–31. L

963 Lei, Z., Huhman, D., & Sumner, LL. (2011). Mass spectrometry strategies in metabolomics.
964   Journal of Biological Chemistry. doi: 10.1074/jbc.R111.238691

965 Leiss, K. A., Choi, Y. H., Abdel-Farid, I. B., Verpoorte, R., & Klinkhamer, P. G. L. (2009). NMR
966   metabolomics of thrips (Frankliniella occidentalis) resistance in Senecio hybrids.
967   Journal of chemical ecology, 35(2), 219–29.

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

971   Lin, C. Y., Viant, M. R., & Tjeerdema, R. S. (2006). Metabolomics: Methodologies and

972   applications in the environmental sciences. Journal of Pesticide Science, 31(3), 245–

973   251.

974   Lindon, J.C., Nicholson, J.K. & Holmes, E. (eds.) *The Handbook of Metabonomics and*

975   *Metabolomics* (Elsevier, Amsterdam, 2007), 55-201.

976   Lindow, S. E., & Brandl, M. T. (2003). Microbiology of the phyllosphere. Applied and

977   environmental microbiology, 69(4), 1875–83.

978   Liu, Y., & Kujawinski, E. B. (2015). Chemical Composition and Potential Environmental Impacts

979   of Water-Soluble Polar Crude Oil Components Inferred from ESI FT-ICR MS. PLOS ONE,

980   10(9), e0136376.

981   Macedo, A. F. (2012). Abiotic Stress Responses in Plants: Metabolism to Productivity. In Abiotic

982   Stress Responses in Plants (pp. 41–61). New York, NY: Springer New York.

983   Mahowald, N. M., Artaxo, P., Baker, A. R., Jickells, T. D., Okin, G. S., Randerson, J. T., &

984   Townsend, A. R. (2005). Impacts of biomass burning emissions and land use change on

985   Amazonian atmospheric phosphorus cycling and deposition. Global Biogeochemical

986   Cycles, 19, GC4030.

987   Mari, A., Lyon, D., Fragner, L., Montoro, P., Piacente, S., Wienkoop, S., et al. (2013).

988   Phytochemical composition of Potentilla anserina L. analyzed by an integrative GC-MS

989   and LC-MS metabolomics platform. Metabolomics : Official journal of the Metabolomic

990   Society, 9(3), 599–607.

991   Medeiros, P. M., Babcock-Adamos, L., Seidel, M., Castelao, R. M., Di Iorio, D., Hollibaugh, J. T.,

992   & Dittmar, T. (2017). Export of terrigenous dissolved organic matter in a broad

993   continental shelf. Limnology and Oceanography, 62, 1718-1731.

994   Menzel, A., Sparks, T. H., Estrella, N., Koch, E., Aasa, A., Ahas, R., et al. (2006). European

995   phenological response to climate change matches the warming pattern. Global Change

996   Biology, 12(10), 1969–1976.

997   Minor, E. C., Swenson, M. M., Mattson, B. M., & Oyler, A. R. (2014). Structural characterization

998   of dissolved organic matter: a review of current techniques for isolation and analysis.

999   Environ. Sci.: Processes Impacts, 16(9), 2064–2079.

1000   Nikiforova, V. J., Kopka, J., Tolstikov, V., Fiehn, O., Hopkins, L., Hawkesford, M. J., et al. (2005).
1001   Systems Rebalancing of Metabolism in Response to Sulfur Deprivation, as Revealed by
1002   Metabolome Analysis of Arabidopsis Plants. Plant physiology, 138(1), 304–318.

1003   Obee, T. N., & Hay, S. O. (1997). Effects of Moisture and Temperature on the Photooxidation of
1004   Ethylene on Titania. Environmental Science & Technology, 31(7), 2034–2038.

1005   Oksanen, J., Guillaume-Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., et al.
1006   (2013). vegan: Community Ecology Package. R package version 2.0-9, http ://CRAN.R-
1007   project.org/package=vegan

1008   Osterholz, H., Singer, G., Wemheuer, B., Daniel, R., Simon, M., Niggemann, J., & Dittmar, T.
1009   (2016). Deciphering associations between dissolved organic molecules and bacterial
1010   communities in a pelagic marine system. The ISME Journal, 10(7), 1717-1730.

1011   Paerl, H. W. (1997). Coastal eutrophication and harmful algal blooms: Importance of
1012   atmospheric deposition and groundwater as "new" nitrogen and other nutrient
1013   sources. Limnology and Oceanography, 42(5part2), 1154–1165.

1014   Pan, Z., & Raftery, D. (2007). Comparing and combining NMR spectroscopy and mass
1015   spectrometry in metabolomics. Analytical and Bioanalytical Chemistry, 387(2), 525–
1016   527.

1017   Pandis, S. N., Harley, R. A., Cass, G. R., & Seinfeld, J. H. (1992). Secondary organic aerosol
1018   formation and transport. Atmospheric Environment. Part A. General Topics, 26(13),
1019   2269–2282.

1020   Parmesan, C. (2006). Ecological and Evolutionary Responses to Recent Climate Change. Annual
1021   Review of Ecology, Evolution, and Systematics, 37(1), 637–669.

1022   Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts
1023   across natural systems. Nature, 421(6918), 37–42.

1024   Paytan, A., Mackey, K. R. M., Chen, Y., Lima, I. D., Doney, S. C., Mahowald, N., et al. (2009).
1025   Toxicity of atmospheric aerosols on marine phytoplankton. Proceedings of the
1026   National Academy of Sciences of the United States of America, 106(12), 4601–5.

1027   Peñuelas, J., & Sardans, J. (2009). Ecological metabolomics. Chemistry and Ecology, 25(4), 305–
1028   309.

1029 Peñuelas, J., Sardans, J., Rivas-ubach, A., & Janssens, I. A. (2012). The human-induced
1030   imbalance between C, N and P in Earth's life system. Global Change Biology, 18(1), 3-6.

1031 Peñuelas, J., & Staudt, M. (2010). BVOCs and global change. Trends in plant science, 15(3),
1032   133–44.

1033 Peñuelas, J., & Terradas, J. (2014). The foliar microbiome. Trends in Plant Science, 19(5), 278-
1034   280.

1035 Pluskal, T., Castillo, S., Villar-Briones, A., & Orešič, M. (2010). MZmine 2: modular framework
1036   for processing, visualizing, and analyzing mass spectrometry-based molecular profile
1037   data. BMC bioinformatics, 11(1), 395.

1038 R Core Team. (2013). R: A language and environment for statistical computing. Vienna.

1039 Ramanathan, V., Crutzen, P. J., Kiehl, J. T., & Rosenfeld, D. (2001). Aerosols, climate, and the
1040   hydrological cycle. Science, 294(5549), 2119–24.

1041 Reemtsma, T. (2009). Determination of molecular formulas of natural organic matter
1042   molecules by (ultra-) high-resolution mass spectrometry: Status and needs. Journal of
1043   Chromatography A, 1216(18), 3687–3701.

1044 Riedel, T., & Dittmar, T. (2014). A Method Detection Limit for the Analysis of Natural Organic
1045   Matter via Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. Analytical
1046   Chemistry, 86(16), 8376–8382.

1047 Rivas-Ubach, A., Barbeta, A., Sardans, J., Guenther, A., Ogaya, R., Oravec, M., et al. (2016b).
1048   Topsoil depth substantially influences the responses to drought of the foliar
1049   metabolomes of Mediterranean forests. Perspectives in Plant Ecology, Evolution and
1050   Systematics, 21.

1051 Rivas-Ubach, A., Gargallo-Garriga, A., Sardans, J., Oravec, M., Mateu-Castell, L., Pérez-Trujillo,
1052   M., et al. (2014). Drought enhances folivory by shifting foliar metabolomes in Quercus
1053   ilex trees. New Phytologist, 202(3).

1054 Rivas-Ubach, A., Hódar, J. A., Sardans, J., Kyle, J. E., Kim, Y.-M., Oravec, M., et al. (2016a). Are
1055   the metabolomic responses to folivory of closely related plant species linked to
1056   macroevolutionary and plant–folivore coevolutionary processes? Ecology and
1057   Evolution, 6(13), 4372-4386.

1058   Rivas-Ubach, A., Liu, Y., Bianchi, T. S., Tolić, N., Jansson, C., & Paša-Tolić, L. (2018). Moving
1059   beyond the van Krevelen Diagram: A New Stoichiometric Approach for Compound
1060   Classification in Organisms. Analytical Chemistry, acs.analchem.8b00529.
1061   doi:10.1021/acs.analchem.8b00529

1062   Rivas-Ubach, A., Pérez-Trujillo, M., Sardans, J., Gargallo-Garriga, A., Parella, T., & Peñuelas, J.
1063   (2013). Ecometabolomics: Optimized NMR-based method. Methods in Ecology and
1064   Evolution, 4(5), 464-473.

1065   Rivas-Ubach, A., Poret-Peterson, A. T., Peñuelas, J., Sardans, J., Pérez-Trujillo, M., Legido-
1066   Quigley, C., et al. (2018). Coping with iron limitation: a metabolomic study of
1067   Synechocystis sp. PCC 6803. Acta Physiologiae Plantarum, 40(2), 28.

1068   Rivas-Ubach, A., Sardans, J., Hódar, J. A., Garcia-Porta, J., Guenther, A., Oravec, M., et al.
1069   (2016c). Similar local, but different systemic, metabolomic responses of closely related
1070   pine subspecies to folivory by caterpillars of the processionary moth. Plant Biology,
1071   18(3), 484–494.

1072   Rivas-Ubach, A., Sardans, J., Hódar, J. A., Garcia-Porta, J., Guenther, A., Paša-Tolić, L., et al.
1073   (2017). Close and distant: Contrasting the metabolism of two closely related
1074   subspecies of Scots pine under the effects of folivory and summer drought. Ecology
1075   and Evolution, 7(21), 8976–8988.

1076   Rivas-Ubach, A., Sardans, J., Peŕez-Trujillo, M., Estiarte, M., & Penũelas, J. (2012). Strong
1077   relationship between elemental stoichiometry and metabolome in plants. Proceedings
1078   of the National Academy of Sciences of the United States of America, 109(11), 4181-
1079   4186.

1080   Rochford, S. (2005). Metabolomics reviewed: A new "omics" platform technology for systems
1081   biology and implications for natural products research. Journal of Natural Products, 68,
1082   1813-1820.

1083   Roullier-Gall, C., Boutegrabet, L., Gougeon, R. D., & Schmitt-Kopplin, P. (2014). A grape and
1084   wine chemodiversity comparison of different appellations in Burgundy: Vintage vs
1085   terroir effects. Food Chemistry, 152, 100–107.

1086   Roulston, T. H., & Cane, J. H. (2000). Pollen nutritional content and digestibility for animals.
1087   Plant Systematics and Evolution, 222(1–4), 187–209.

1088   Saito, K., & Matsuda, F. (2010). Metabolomics for functional genomics, systems biology, and
1089   biotechnology. Annual Review of Plant Biology, 361, 463-489.

1090   Sardans, J., Gargallo-Garriga, A., Pérez-Trujillo, M., Parella, T. J., Seco, R., Filella, I., & Peñuelas,
1091   J. (2014). Metabolic responses of Quercus ilex seedlings to wounding analysed with
1092   nuclear magnetic resonance profiling. Plant biology, 16(2), 395–403.

1093   Sardans, J., Peñuelas, J., & Rivas-Ubach, A. (2011). Ecological metabolomics: overview of
1094   current developments and future challenges. Chemoecology, 21(4), 191–225.

1095   Sardans, J., Rivas-Ubach, A., & Peñuelas, J. (2012a). The elemental stoichiometry of aquatic and
1096   terrestrial ecosystems and its relationships with organismic lifestyle and ecosystem
1097   structure and function: A review and perspectives. Biogeochemistry, 111(1–3).

1098   Schmitt-Kopplin, P., Liger-Belair, G., Koch, B. P., Flerus, R., Kattner, G., Harir, M., et al. (2012).
1099   Dissolved organic matter in sea spray: a transfer study from marine surface water to
1100   aerosols. Biogeosciences, 9(4), 1571–1582.

1101   Shulaev, V. (2006). Metabolomics technology and bioinformatics. Briefings in Bioinformatics,
1102   7(2), 128-139.

1103   Seco, R., Peñuelas, J., & Filella, I. (2007). Short-chain oxygenated VOCs: Emission and uptake by
1104   plants and atmospheric sources, sinks, and concentrations. Atmospheric Environment,
1105   41(12), 2477–2499.

1106   Shulaev, V., Cortes, D., Miller, G., & Mittler, R. (2008). Metabolomics for plant stress response.
1107   Physiologia Plantarum, 132(2), 199–208.

1108   Sleighter, R. L., & Hatcher, P. G. (2007). The application of electrospray ionization coupled to
1109   ultrahigh resolution mass spectrometry for the molecular characterization of natural
1110   organic matter. Journal of mass spectrometry : JMS, 42(5), 559–74.

1111   Smith, D. F., Podgorski, D. C., Rodgers, R. P., Blakney, G. T., & Hendrickson, C. L. (2018). 21
1112   Tesla FT-ICR Mass Spectrometer for Ultrahigh-Resolution Analysis of Complex Organic
1113   Mixtures. Analytical Chemistry, doi:10.1021/acs.analchem.7b04159

1114   Smith, D., & Španěl, P. (2011). Direct, rapid quantitative analyses of BVOCs using SIFT-MS and
1115   PTR-MS obviating sample collection. TrAC Trends in Analytical Chemistry, 30(7), 945–
1116   959.

1117   Solberg, Y., & Remedios, G. (1980). Chemical composition of pure and Bee-Collected pollen.

1118   Meldinger fra Norges landbrukshogskole, 59, 1–13.

1119   Spencer, R. G. M., Mann, P. J., Dittmar, T., Eglinton, T. I., McIntyre, C., Holmes, R. M., et al.

1120   (2015). Detecting the signature of permafrost thaw in Arctic rivers. Geophysical

1121   Research Letters, 42(8), 2830–2835.

1122   Sterner, R., & Elser, J. (2002). Ecological Stoichiometry: The Biology of Elements from

1123   Molecules to the Biosphere. Princetion University Press.

1124   Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., et al. (2007).

1125   Proposed minimum reporting standards for chemical analysis. Metabolomics, 3(3),

1126   211–221.

1127   t'Kindt, R., De Veylder, L., Storme, M., Deforce, D., & Van Bocxlaer, J. (2008). LC-MS metabolic

1128   profiling of Arabidopsis thaliana plant leaves and cell cultures: optimization of pre-LC-

1129   MS procedure parameters. Journal of chromatography. B, Analytical technologies in

1130   the biomedical and life sciences, 871(1), 37–43.

1131   Tfaily, M. M., Chu, R. K., Tolić, N., Roscioli, K. M., Anderton, C. R., Paša-Tolić, L., et al. (2015).

1132   Advanced Solvent Based Methods for Molecular Characterization of Soil Organic

1133   Matter by High-Resolution Mass Spectrometry. Analytical Chemistry, 87(10), 5206–

1134   5215.

1135   Tholl, D., Boland, W., Hansel, A., Loreto, F., Röse, U. S. R., & Schnitzler, J.-P. (2006). Practical

1136   approaches to plant volatile analysis. The Plant Journal, 45(4), 540–560.

1137   Thompson, H., Heimendinger, J., Gillette, C., Sedlacek, S., Haegele, A., O'Neill, C., & Wolfe, P.

1138   (2005). In Vivo Investigation of Changes in Biomarkers of Oxidative Stress Induced by

1139   Plant Food Rich Diets. Journal of Agricultural and food Chemistry, 53(15), 6126–6132.

1140   Tolić, N., Liu, Y., Liyu, A., Shen, Y., Tfaily, M. M., Kujawinski, E. B., et al. (2017). Formularity:

1141   Software for Automated Formula Assignment of Natural and Other Organic Matter

1142   from Ultrahigh-Resolution Mass Spectra. Analytical Chemistry,

1143   doi:10.1021/acs.analchem.7b03318

1144   Uzu, G., Sobanska, S., Sarret, G., Muñoz, M., & Dumat, C. (2010). Foliar Lead Uptake by Lettuce

1145   Exposed to Atmospheric Fallouts. Environmental Science & Technology, 44(3), 1036–

1146   1042.

1147   van den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K., & van der Werf, M. J.

1148   (2006). Centering, scaling, and transformations: improving the biological information

1149   content of metabolomics data. BMC Genomics, 7(1), 142.

1150   van Krevelen, D. (1950). Graphical-statistical method for the study of structure and reaction

1151   processes of coal. Fuel, 29, 269–284.

1152   Vorholt, J. A. (2012). Microbial life in the phyllosphere. Nature Reviews Microbiology, 10(12),

1153   828–840.

1154   Walsh, M. C., Brennan, L., Malthouse, J. P. G., Roche, H. M., & Gibney, M. J. (2006). Effect of

1155   acute dietary standardization on the urinary, plasma, and salivary metabolomic

1156   profiles of healthy humans. The American journal of clinical nutrition, 84(3), 531–9.

1157   Walther, G. R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., et al. (2002).

1158   Ecological responses to recent climate change. Nature, 416(6879), 389–395.

1159   Wang, R., Balkanski, Y., Bopp, L., Aumont, O., Boucher, O., Ciais, P., et al. (2015). Influence of

1160   anthropogenic aerosol deposition on the relationship between oceanic productivity

1161   and warming. Geophysical Research Letters, 42(24), 10745–10754.

1162   Warnes, G. R., Bolker, B., Bonebakker, L., Gentleman, R., Huber Andy Liaw, W., Lumley, T., et

1163   al. (2016). gplots: various R programing tools for plotting data. R package version 3.0.1,

1164   http://CRAN.R-project.org/package=gplots

1165   Wedding, J. B., Carlson, R. W., Stukel, J. J., & Bazzaz, F. A. (1975). Aerosol deposition on plant

1166   leaves. Environmental Science & Technology, 9(2), 151–153.

1167   Whipps, J. M., Hand, P., Pink, D., & Bending, G. D. (2008). Phyllosphere microbiology with

1168   special reference to diversity and plant genotype. Journal of Applied Microbiology,

1169   105(6), 1744–1755.

1170   White, R. A., Rivas-Ubach, A., Borkum, M. I., Köberl, M., Bilbao, A., Colby, S. M., et al. (2017).

1171   The state of rhizospheric science in the era of multi-omics: A practical guide to omics

1172   technologies. Rhizosphere, 3, 212–221. doi:10.1016/J.RHISPH.2017.05.003

1173   Wishart, D. S. (2008). Metabolomics: applications to food science and nutrition research.

1174   Trends in Food Science & Technology, 19(9), 482–493.

1175 Wozniak, A. S., Bauer, J. E., Sleighter, R. L., Dickhut, R. M., & Hatcher, P. G. (2008). Technical
1176   Note: Molecular characterization of aerosol-derived water soluble organic carbon
1177   using ultrahigh resolution electrospray ionization Fourier transform ion cyclotron
1178   resonance mass spectrometry. Atmospheric Chemistry and Physics, 8(17), 5099–5111.

1179 Xiong, T.-T., Leveque, T., Austruy, A., Goix, S., Schreck, E., Dappe, V., et al. (2014). Foliar uptake
1180   and metal(loid) bioaccessibility in vegetables exposed to particulate matter.
1181   Environmental Geochemistry and Health, 36(5), 897–909.

1182 Zhang, A., Sun, H., Wang, P., Han, Y., & Wang, X. (2012). Modern analytical techniques in
1183   metabolomics analysis. The Analyst, 137(2), 293–300.

1184 Zhang, Q., Stanier, C. O., Canagaratna, M. R., Jayne, J. T., Worsnop, D. R., Pandis, S. N., &
1185   Jimenez, J. L. (2004). Insights into the Chemistry of New Particle Formation and Growth
1186   Events in Pittsburgh Based on Aerosol Mass Spectrometry. Environmental Science &
1187   Technology, 38(18), 4797–4809.

1188

1189

1190

1191

1192

1193

1194

1195

1196

1197

1198

1199

1200

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			

1203

1204

1205

1206

1207

1208

1209

1210

1211

1212

1213

1214

1215 **Figure with captions.**

1216

1217
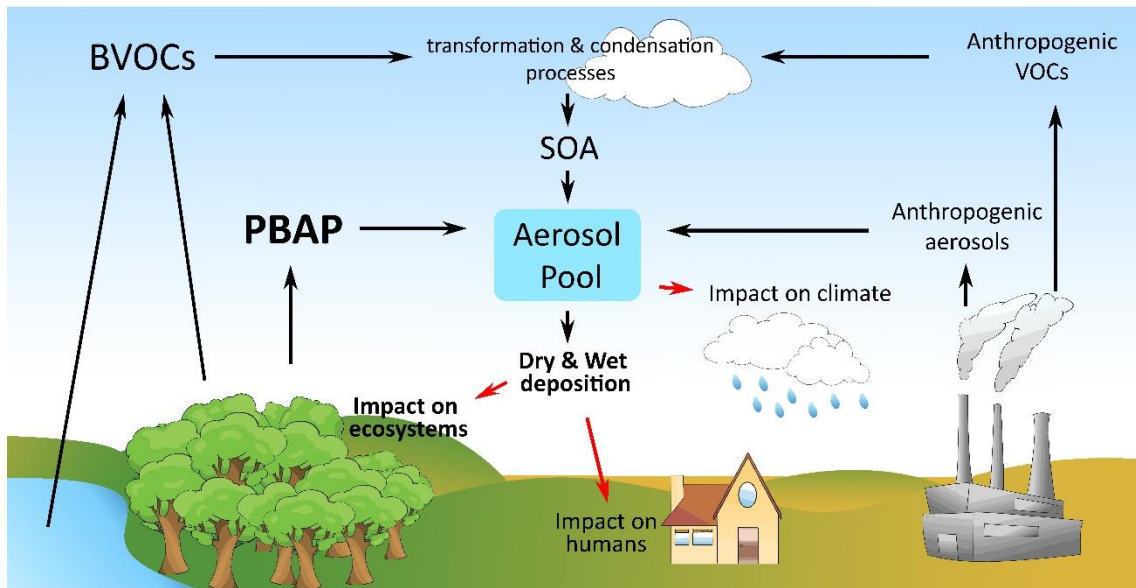
1218

1219

1220 **Figure 1.** Schematic diagram of the aerosol emissions and deposition on ecosystems.

1221

1222

1223

1224

1225

1226

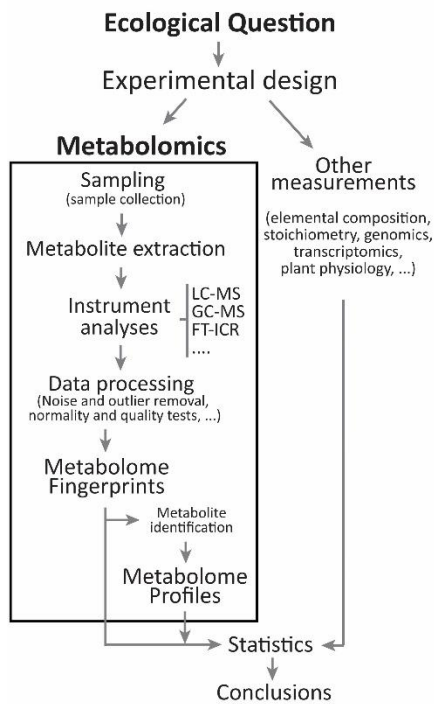1227

1228

1229

1230

1231

1232

1233

1234

1235

1236

1237

Figure 2. Diagram of the main procedures of a general ecometabolomic study combined with complementary measurements.
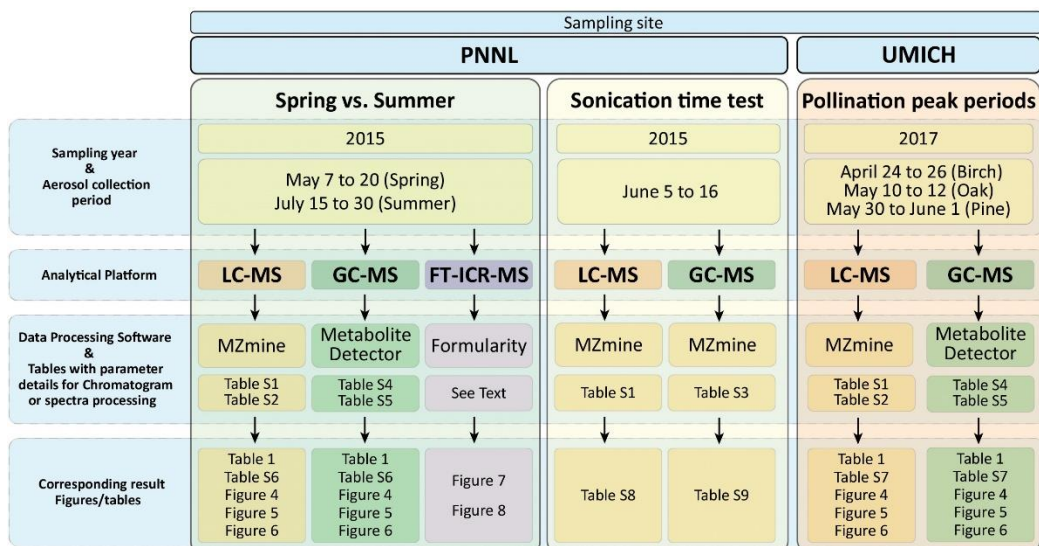
Figure 3. Sampling year, collection period, and analytical instruments and software used for generating the datasets from each of the sampling site and campaign. Table numbers with the parameters used for dataset generation for each of the software are indicated. Table and figure numbers containing the main results derived from aerosols analyzed from each sampling campaign and analytical technique are also indicated.

Set of tubes A

1. Roll the filters

Quartz filter

2. Introduce filters into the glass tubes

3. Add 5 mL MeOH/H$_2$O

4. Sonicate for 10 min. (cover/cap tubes individually)

5. Repeat for two extractions (add 4 mL MeOH/H$_2$O)

4.1. First extraction, collect 4 mL of extract

5.1. Second extraction, collect 4 mL of extract

Set of tubes B

6. Evaporate the extracts with a nitrogen evaporator

N2

Extract powder

7. Add 1 mL MeOH/H$_2$O and vortex for 30 s. (cap tubes individually)

8. Centrifuge for 5 min at 4,000 x g.

9. Collect the concentrated extracts

Set of tubes C

10. Store at -80ºC until analyses

11.

DI/LC-MS analyses

GC-MS analyses

12. Place the extract into HPLC vials

Metabolite Derivatization

1284

1285

1286 **Figure 4.** Experimental procedures performed on filters to obtain the metabolomic extracts

1287 from aerosol samples for subsequent analyses with distinct mass spectrometry techniques.

1288

1289

1290

1291

1292

1293

1294

1295

1296

1297

1298

PNNL SITE

a LC-MS   **b** GC-MS

UMICH SITE

c LC-MS   **d** GC-MS
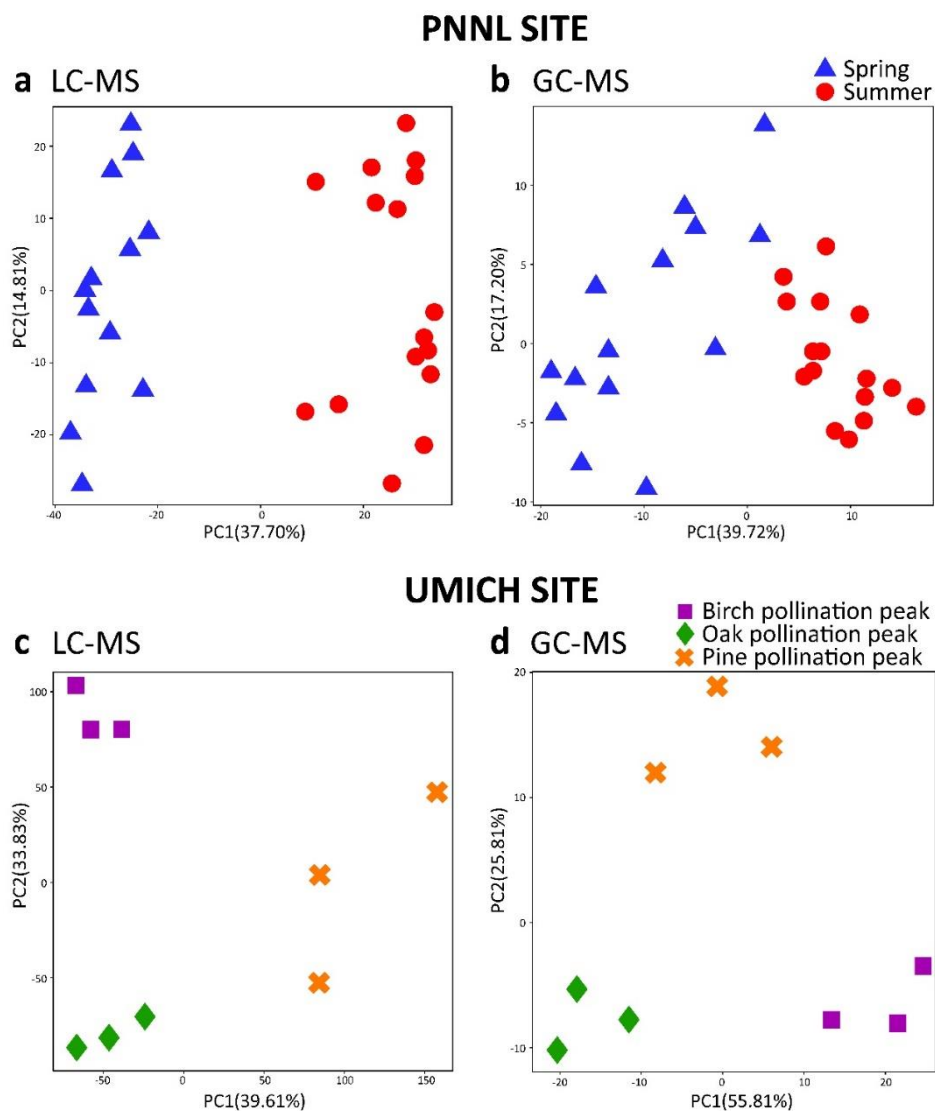
Figure 5. Case plots of the Principal Component (PC) 1 versus the PC2 of the PCAs conducted on the aerosol metabolomic fingerprints collected at PNNL site contrasting spring vs. summer aerosols (**a** for LC-MS; **b** for GC-MS) and collected at UMICH site contrasting different pollination peak periods (**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.
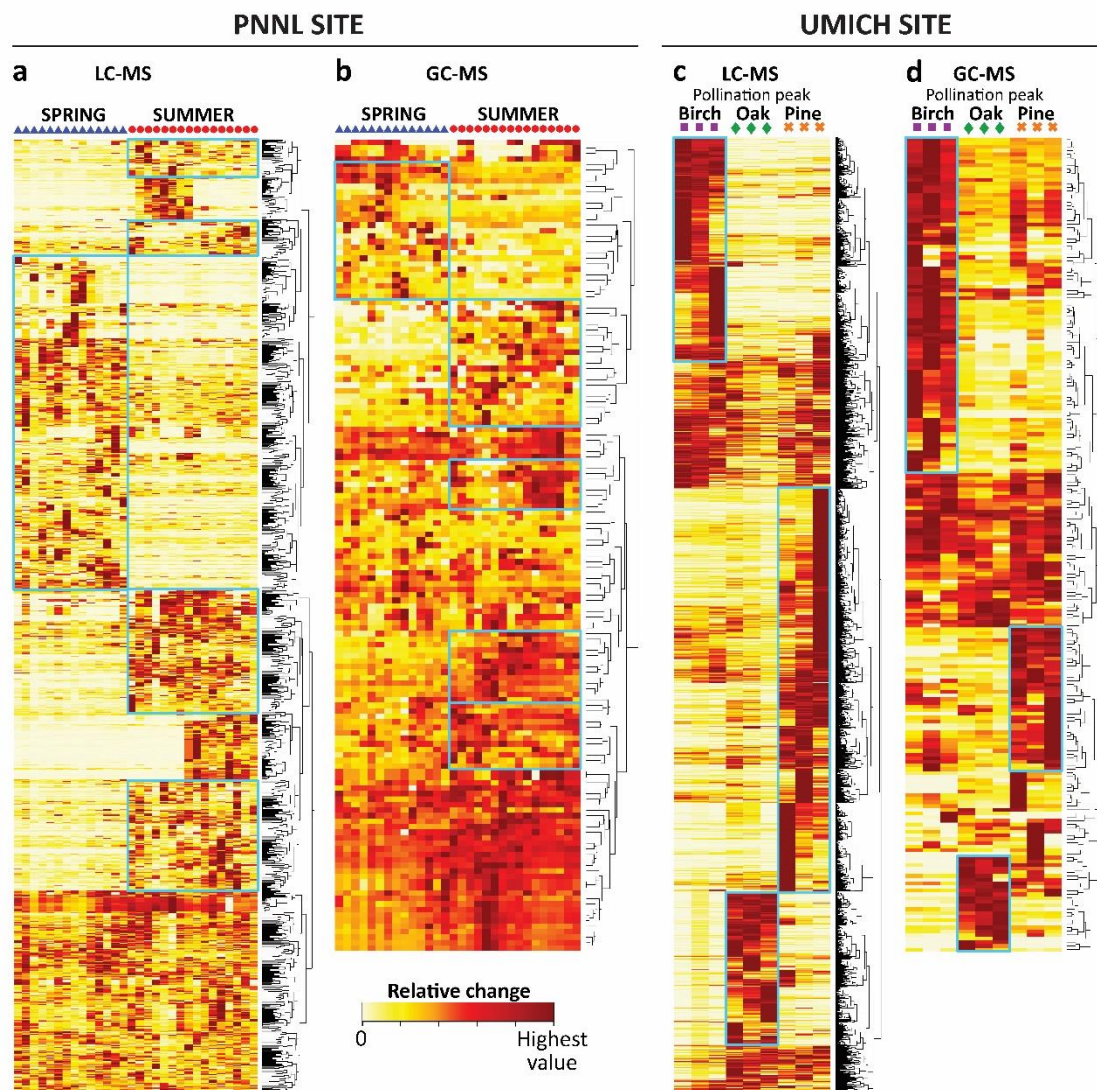
Figure 6. Heatmaps of the metabolomic fingerprints obtained from LC-MS and GC-MS for PNNL metabolomics datasets (spring vs. summer) (**a** and **b**) and for the UMICH metabolomics datasets (pollination peak periods) (**c** and **d**). Cluster dendrogram for the variables is shown for each heatmap. Blue rectangles point cluster of variables for the specific group of samples having higher overall relative abundance values. The color gradient represent the relative abundance of the metabolomic feature between samples. Color red represents the highest relative abundance.
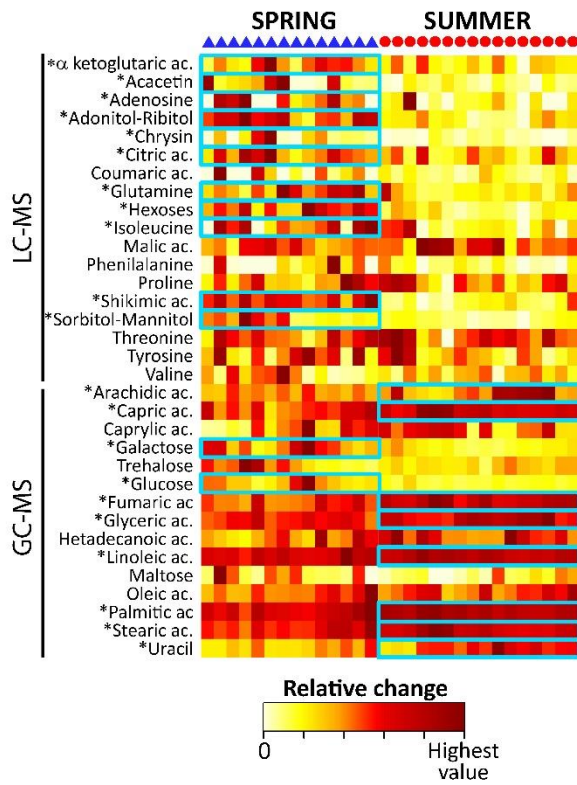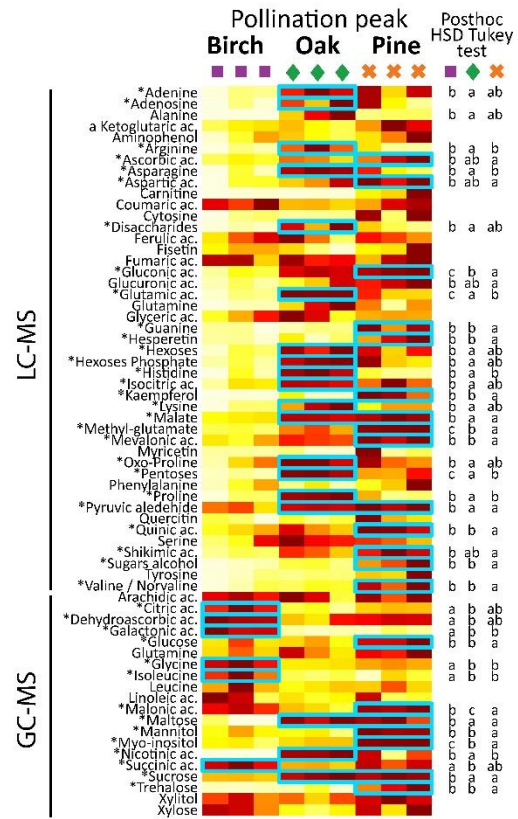
Figure 7. Heat maps of the identified metabolomic features from LC-MS and GC-MS for the PNNL metabolomics dataset (spring vs. summer) (**a**) and for the UMICH metabolomics dataset (pollination peak periods dataset)(b). For each site, heatmaps of the identified features were plotted using LC-MS and GC-MS data combined. The color gradient represent 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 oneway 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 posthoc test.
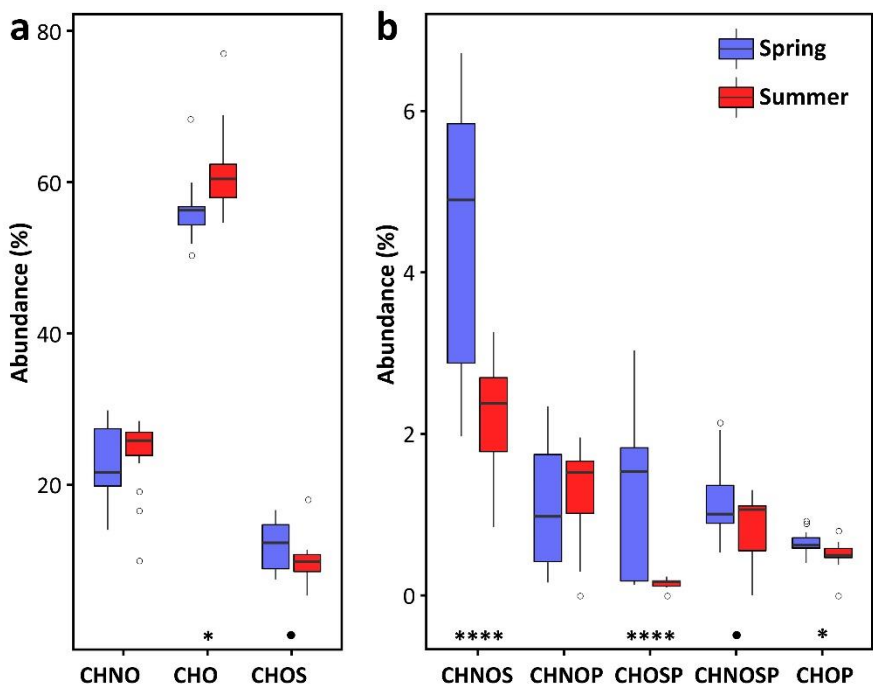
Figure 8. Box plots for the proportion (%) of different formula classes (CHNO, CHO and CHOS (**a**) and CHNOS, CHNOP, CHOSP, CHNOSP and CHOP (**b**)) for PNNL site DI-FT-ICR-MS dataset (spring vs. summer). Boxes show the median value for each formula class and season. Open dots correspond to extreme 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 ($P < 0.1$).

1365

1366

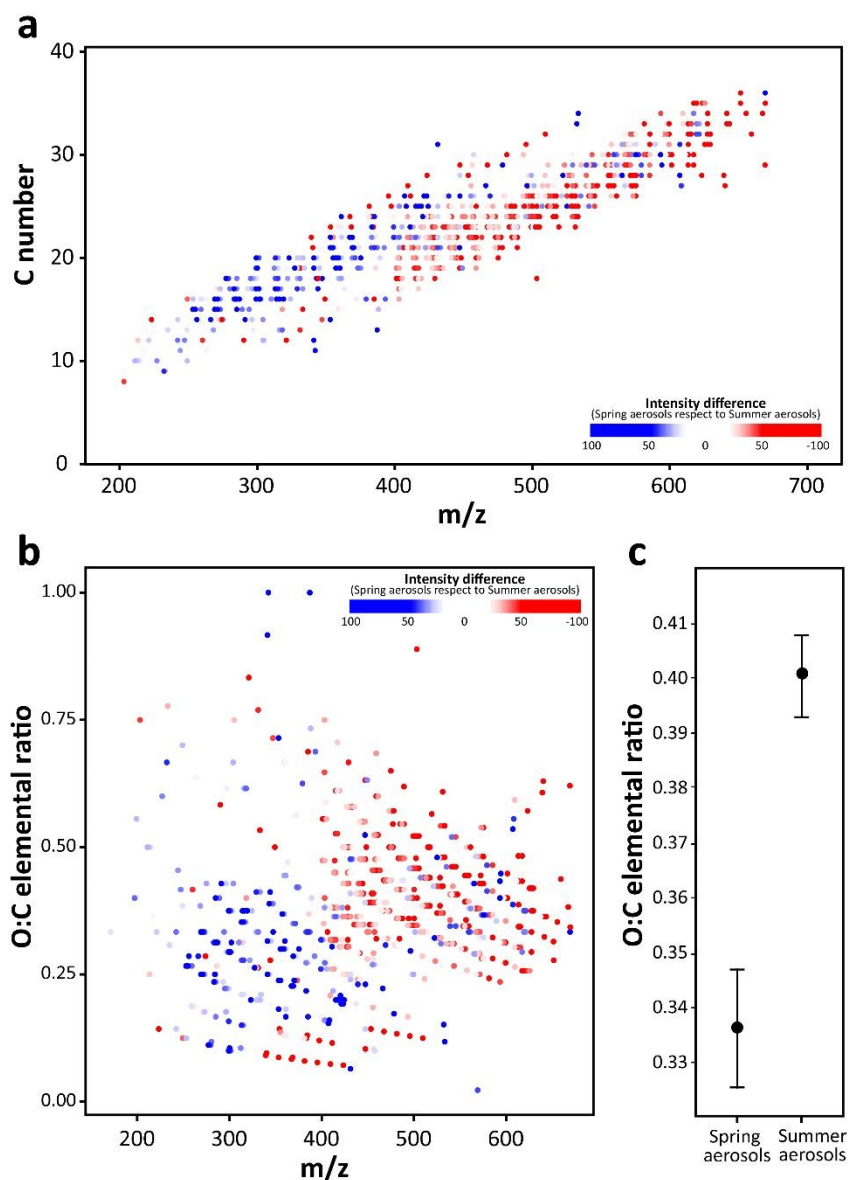**Figure 9.** Carbon number versus m/z (CvM) (**a**) and Oxygen/Carbon ratio versus mass (**b**)

diagrams plotted from DI-FT-ICR-MS datasets (PNNL site). Different color 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.

1367

1368

1369

1370

1371

1372

1373