Ultrahigh Performance LC/FT-MS Non-Targeted Screening for Biomass Burning Organic Aerosol with MZmine2 and MFAssignR

Thusitha Divisekara[1], Simeon Schum[1,2], and Lynn Mazzoleni[1]

[1]Department of Chemistry, Michigan Technological University, Houghton, MI, USA

[2] Chemical Advanced Resolution Methods Laboratory, Michigan Technological University, Houghton, MI, USA

Key words: Fourier transform mass spectrometry, Non-targeted screening, Ion mass extraction, Molecular formula assignment, Biomass burning organic aerosol

Highlights:

- A new non-targeted data processing strategy suitable for environmental complex mixtures is presented.
- Results of the LC/FT-MS approach are consistent with direct infusion FT-MS results.
- Liquid smoke is a potential surrogate for biomass burning organic aerosol.

Abstract

In recent years, ultrahigh performance liquid chromatography Fourier transform mass spectrometry (LC/FT-MS) based non-targeted screening (NTS) methods have become increasingly popular for comprehensive analysis of complex organic mixtures. However, applying these methods for environmental complex mixture analysis is challenging due to the extreme complexity of natural samples and a lack of standard samples or surrogates for environmental complex mixtures. Furthermore, limited molecular markers in the databases and insufficient data processing software workflows make the application of these methods more challenging for environmental complex mixtures. In this work, we implement a new NTS data processing workflow to process data collected from ultrahigh performance liquid chromatography and Fourier transform Orbitrap Elite Mass Spectrometry (LC/FT-MS) by combining MZmine2 and MFAssignR, two opensource data processing tools and commercial Mesquite liquid smoke as a surrogate for biomass burning organic aerosol. MZmine2.53 data extraction followed MFAssignR molecular formula assignment offered noise free and highly accurate 1733 individual molecular formulas presented in liquid smoke with 4906 molecular species, including isomers. The results of this new approach were consistent with the results of direct infusion FT-MS analysis confirming its reliability. Over 90% of the molecular formulas presented in mesquite liquid smoke were matched with the molecular formulas of ambient biomass burning organic aerosol. This suggests

the potential use of commercial liquid smoke as a surrogate for biomass burning organic aerosol research. The presented method significantly improves the identification of the molecular composition of biomass burning organic aerosol by successfully addressing some of the limitations related to the data analysis and giving a semi quantitative insight into the analysis.

1. Introduction

Natural and anthropogenic biomass combustion contributes a significant amount of organic aerosol to the atmosphere. Model estimates have indicated that the annual emission of organic carbon from biomass burning in 2008 ranged between 15.65 to 51.91 Tg (Pan et al., 2020). Smoke emissions from fires usually contain simple organic acids, sugars, phenolic compounds, polycyclic aromatic hydrocarbon (PAHs), and thousands of unknowns depending on the condition of the fire (Brege et al., 2018; Lin et al., 2016; Mazzoleni et al., 2007). Transportation of these fire plumes through the atmosphere can yield high levels of air pollution impacting human health as well as the Earth's climate (Kenward et al., 2013; Shiraiwa et al., 2017; Zhang et al., 2020). Unraveling the molecular composition of biomass burning organic aerosol is challenging due to its extreme complex nature (Brege et al., 2021). Although several analytical methods for a variety of organic aerosol analyses are available in the literature, none yet provide a comprehensive qualitative and quantitative assessment of the molecular composition (Duarte et al., 2021). Most LC/MS or GC/MS based quantitative studies in the literature use targeted approaches to analyze specific compounds (Duarte et al., 2021), which limits the analysis to the selected analytical standards available. In recent years, direct infusion Fourier transform mass spectrometry (DI-FT-MS), frequently with electrospray ionization (ESI; Brege et al., 2018, 2021; Ijaz et al., 2022; Lin et al., 2016; Schum et al., 2018), have been used technique for detailed analysis of organic aerosol. Such studies have demonstrated extreme molecular complexity ranging from saturated to highly unsaturated products of incomplete combustion with a wide variety of transformations from atmospheric oxidation processes (Brege et al., 2021; Ijaz et al., 2022; Lin et al., 2016; Schum et al., 2018). Although very powerful, direct infusion mass spectrometry cannot provide either structural or quantitative information of the studied analytes. The analysis can be subject to matrix effects and ionization suppression due to the competition for charge during the electrospray ionization (Chekmeneva et al., 2017). Therefore, relative intensities of direct infusion mass spectrometry may not directly proportional to the relative concentration of the analysts (Guillemant et al., 2021). The matrix

effects and ionization can be reduced using chromatographic separation to reduce the molecular complexity of species and their concentrations arriving at the ionization source (Forcisi et al., 2013). Due to the analogy of a molecular partitioning in the chromatographic column and within the ESI droplet, Cech et al., 2001 observed the variation in the ionization response with the chromatographic retention time. Such that, the fast eluent species can have low ESI response, and later eluant species have a high possibility of having high ESI response (Cech et al., 2001). At the same time, changing the mobile phase composition can lead to changes in the ionization response (Ghosh and Jones, 2015; Liigand et al., 2014). Due to these observations, compounds with similar ionization efficiencies are expected to elute with similar retention times. The reduced complexity and the lower concentration of analytes together minimize the detection artifacts caused by ionization competition (Forcisi et al., 2013; Kim et al., 2019; Lin et al., 2010). Therefore, the peak area values obtained from LC/MS analysis are more reliable for quantitative or semi-quantitative assessments than the relative ion abundance in mass spectra from direct infusion analysis (Forcisi et al., 2013; Lin et al., 2010).

In Fourier Transform Mass Spectrometry (FT-MS), the accuracy of the detected masses is directly proportional to a function of data acquisition time. On the other hand, in LC/FT-MS analysis, longer data acquisition times can result in a lower number of detection peaks (Kim et al., 2019; Michalski et al., 2012). Thus, balancing the two factors is crucial to obtain optimal results. In our study, we used ultrahigh performance LC to achieve maximum peak capacity and fast separation, which is compatible with the relatively fast data acquisition of ultrahigh resolution Fourier Transform Orbitrap Elite Mass Spectrometry (Michalski et al., 2012). Similar analytical approaches are widely used for non-targeted screening of metabolites and trace environmental contaminants (Hollender et al., 2019; Michalski et al., 2012; Zhang et al., 2020).

LC/FT-MS and high-resolution time of flight (ToF) mass spectrometry non targeted screening is a continuously developing technique in environmental complex mixture analysis (Hollender et al., 2017; Minkus et al., 2022). There are numerous open source and commercial software tools available for LC/MS post data processing in the literature. However, data post-processing remains as the main challenge when applying NTS approach in environmental complex mixture analysis (Hohrenk et al., 2020; Houriet et al., 2022; Minkus et al., 2022). Because LC/FT-MS data acquisition generates large quantities of data points within a short time frame. Collected data can be highly complexed due to the isotopes, adducts, and multicharged species detected and

sheer number of different molecules present in such samples (Hollender et al., 2017). The peak detection and mass extraction are highly dependent on the data processing software program used. For example, a study conducted by Hohrenk et al. observed a low level of coherence among the extracted masses extracted from different data processing software and algorithms (Hohrenk et al., 2020). Database based compound identification can lead to incorrect or uncertain molecular identifications due to the lack of observation molecular markers of the databases (Hollender et al., 2019). Further, signal threshold value also critical when considering the quality of the data, detection of low abundance species and the computational time (Hollender et al., 2019, 2017; Houriet et al., 2022). Therefore, data processing tools and parameters need to be carefully selected and standardized depending on the study (Hollender et al., 2017; Houriet et al., 2022).

The goal of this work is to develop an LC/FT-MS method to identify biomass burning organic aerosol components using commercial liquid smoke. Liquid smoke is a commercial food flavoring substance, created by dissolving the volatile components collected from the smoke produced during the slow pyrolysis of wood (Montazeri et al., 2013; Simon et al., 2005). This process resembles how organic aerosols are generated through pyrolysis during plant burning events (Sekimoto et al., 2018). However, the production of liquid smoke does not involve atmospheric oxidation and multiphase processing that would typically occur in ambient conditions. Laboratory oxidation and multiphase processing of liquid smoke to more comprehensively mimic ambient biomass burning organic aerosol is the subject of a forthcoming paper (Divisekera et al, in prep). Therefore, liquid smoke is a potential surrogate for biomass burning organic aerosol. In this study the obstacles in the LC/FT-MS data processing are addressed by combining the existing LC/MS data processing software MZmine2.53 (Pluskal et al., 2010) with MFAssignR (Schum et al., 2020) software for molecular formula assignments of environmental complex mixtures. MFAssignR was originally designed for the post-processing of direct infusion ultrahigh resolution mass spectra and has been adapted to handle LC-MS data as well. It calculates molecular formula using the expected chemical relationships between ions in a complex mixture, including CH_2, H_2O, O, H_2, and CH_2O which provides robust identification of unambiguous molecular formulas as described (Schum et al., 2020). In this approach, the exact masses of the detected ions and their peak areas were extracted using MZmine2.53 and then MFAssignR was used for molecular formula assignments, noise estimation, isotope filtration, and internal mass recalibration. The resulting data set contains assigned and noise-free unambiguous molecular formula along with the

recalibrated exact masses, chromatographic retention times, and chromatographic peak areas representing analyte abundance.

2. Material and methods

2.1 LC/FT-MS analysis of liquid smoke

Wright's commercial mesquite liquid smoke solution was filtered using 0.20 µm PTFE syringe filter and was diluted by two times using LC/MS grade water. Chromatographic separation was done using C18 column (Kinetex® 1.7 µm XB-C18 column) with an UltiMate 3000 UHPLC System (Thermo Fisher Scientific) operated with Chromeleon™ chromatography data system software. Compounds were separated using a solvent gradient at 35°C with the flowrate of 0.2 mL min^{-1} starting with 5% ACN and 95% water (both with 0.1% formic acid) and increased up to 70% ACN over 45 min with a linear gradient. Separated liquid smoke analytes were analyzed with both positive and negative electrospray ionization (ESI) modes. We used 0.1% formic acid to improve the chromatographic separation in both ionization modes. Hence, a decrease in the sensitivity of acidic and neutral compounds detection in the −ESI analysis might happen.

The ESI settings include: temperature of 200 °C, sheath gas flow rate of 20, auxiliary gas flow rate of 8, sweep gas flow rate of 1, spray voltage of 2.8 kV, capillary temperature of 300 °C and S-lens RF level of 60% were used in the –ESI analysis. In +ESI mode analysis, the same initial ESI settings were used except the sprayed voltage (3.8 kV) and S-lens RF level of 65%. Data acquisition was done using Fourier Transform Orbitrap Elite Hybrid Ion trap-Orbitrap Mass spectrometer (FT-OTE, Thermo Scientific™) at resolving power of 120,000 and 240,000 defined at m/z 400 Da. The instrument was operated with XCalibur 3.0 software in the range of 100-600 Da. MS/MS data were acquired by 15,000 resolutions at m/z 400 with the data dependent selected ion monitoring (SIM) for the 10 most abundant ions. Fragmentation was achieved with 25% and 35% of the normalized collision energy in the ion trap. Dynamic exclusion duration was disabled to obtain the fragments of the isomeric species. Triplicated analysis of the liquid smoke sample and the corresponding blank (LS/MS grade water) were performed using the discussed procedure.

2.1.2 Data extraction using MZmine2.

In the first step of ion extraction, Thermo raw data files were directly imported into the MZmine2.53 software platform without any file conversion. Baseline correction was done to remove the gradual shift of the baseline of the chromatogram. Then, the mass detection was carried

out by centroiding the profile masses using Exact mass detector algorithm (Pluskal et al., 2010) keeping an optimum noise intensity level for mass detection. The noise intensity level was kept at 250 during the entire retention time range to avoid filtering out the analyte peaks with low intensity. Shoulder peaks generated during the Fourier transformations were filtered out using the Lorentzian extended algorithm. Subsequently, chromatograms were constructed using the ADAP chromatogram builder (Myers et al., 2017) with a minimum group size of 3 scans, a group intensity threshold of 450, a minimum highest intensity of 600, and a mass error tolerance of 0.001 Da. Obtained chromatograms were then deconvoluted using the ADAP wavelet algorithm with a minimum feature height of 500, signal to noise ratio of 10, coefficient/area threshold of 1, peak duration of 0.00-1.6 minutes, and RT wavelet range of 0.2 in positive mode ion deconvolution and 0.3 in the negative mode ion deconvolution step. Then the deconvoluted ion masses, retention times and the peak area values were exported as a .csv file. The optimized data extraction steps are shown in Figure 1.

We used a small minimum group number of 3 scans and a low noise level of 250 during the ion mass extraction procedure. Keeping a low signal threshold is more helpful in detecting low intense and isotopic peaks (Houriet et al., 2022). At the same time, it can contribute to a higher chance of extracting false peaks. In this scenario, Kendrick mass defect (KMD) based noise estimation function in MFAssignR package (Schum et al., 2020) is useful to filter out the incorrectly extracted masses with smallest peak area values. Further, the isotope filtration function successfully identified the isotopic masses extracted. These two functions in the MFAssignR software package successfully adapted to processing LC/FT-MS data, yielding high accurate molecular formula assignment. (Further discussed in the supplementary section S1 and S2)

2.1.3 Molecular formula assignment using MFAssignR.

Molecular formulas for the extracted masses were assigned separately for the replicate liquid smoke and blank mass lists. The chromatographic peak area of each extracted mass was selected to represent the analyte abundance as opposed to the relative abundance in the mass spectra. Kendrick mass defect noise estimation and polyisotope filtration were done to avoid false assignments. During the formula assignment in $^{12}C_{1\text{-}c}$, $^{1}H_{0\text{-}h}$, $^{16}O_{0\text{-}o}$, $^{14}N_{0\text{-}2}$ and in +ESI, Na^{+} elements were selected within the following ranges: double bond equivalence minus number of oxygen atoms (DBE-O) (EqS4) between -13 to +13, $0.3 \leq H/C \leq 2.5$, $0.0 \leq O/C \leq 2$ de novo cut-off of 300

m/z, and mass error of ≤ 3 ppm using MFAssignR. the peak area of the Na^+ adducted species and H^+ adducted species with the same retention time and same molecular formulas were added together to yield the total peak area. Then all three replicate assignments were aligned according to the formula assignments and the retention time using a custom R script with a range of retention time ± 0.05 min. The molecular formulas in at least two out of the three data sets were retained. The same procedure was followed for the blank data sets. Subsequently, we subtracted the average peak area of the blank from that of the sample to eliminate any contaminants present. Then, all species with positive peak area values were selected for further analysis. Each peak area's average peak area and the corresponding relative standard deviation (RSD) were calculated.

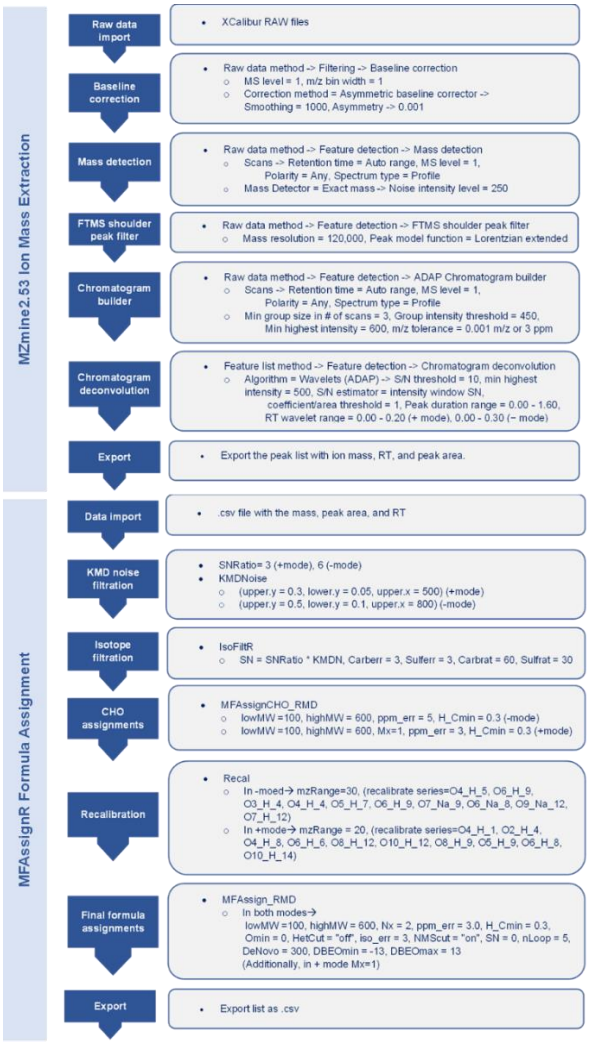

Figure 1: MZmine2.53 and MFAssignR data processing workflow and parameters used.

2.2 Direct infusion FT-MS and molecular formula assignment of liquid smoke

In direct infusion FT-MS analysis, a more diluted sample is typically used to reduce matrix effects, ion suppression, and interference, therefore the same mesquite liquid smoke sample diluted 10 times in 55% of ACN medium to best match the average chromatographic solvent matrix. Then the obtained exact masses were assigned using MFAssignR software using the same formulas assignment parameters discussed.

3. Results and Discussion

3.1 Ion mass extraction and molecular formula assignment

The C18 column and the linear gradient reasonably separated the liquid smoke complex mixture. Approximately 8000 and 7000 ion masses were extracted in both ionization modes. Figure 2 shows the extracted ion chromatogram generated by MZmine2 software (Figure 2A) and the extracted data plotted by peak height vs retention time and colored by observed ion mass in a reconstructed chromatogram for the +ESI analysis with a resolving power of 120,000 (Figure 2B). The similarity in peak intensities and distributions supports the observation that our approach accurately extracted the ions with their respective masses and retention times.

Accurate extraction of low intense noise peaks a challenge that we faced with a low noise level during the ion mass extraction. However, KMD noise estimation method (section S1) successfully removed the noise peak generated during the mass ion extraction (Figure S1) especially in the –ESI ionization mode analysis. According to the KMD plot generated with the +ESI mode analysis, very few noise peaks were extracted with the parameters used (Figure S2). MFAssignR isotope filtration function (IsoFiltR) (section S2) identified isotopic ion masses presented in the extracted mass list. For higher abundance species, ^{13}C isotopic masses, which have the same retention time as the monoisotopic masses, showed approximately equal isotopic abundance (peak area) ratio with the theoretical isotopic abundance (peak area) ratio (Figure S3). This behavior was more noticeable in the 120,000 resolving power than in the 240,000, consistent with the expected fewer species in the higher resolution.

MZmine2 ion mass extraction followed by MFAssignR formula assignment yielded about 3000 and 1500 unambiguous molecular species in both +ESI and –ESI ionization modes at 120,000 and 240,000 resolving power settings. The method showed good reproducibility, such that, the relative standard deviation of the peak area was below 20% for 90% of the aligned formulas that

were detected in at least two out of the tree replicate sets along with their retention times. Overall results obtained at 120,000 and 240,000 resolving power were similar in each ionization mode (Table S1). However, the higher number of total assignments obtained with 120,000 is likely due to the faster detection times at the lower resolving power allowing the instrument to detect more analyte signals (Michalski et al., 2012). The results obtained at 120,000 resolving power will be discussed for the remainder of this article.

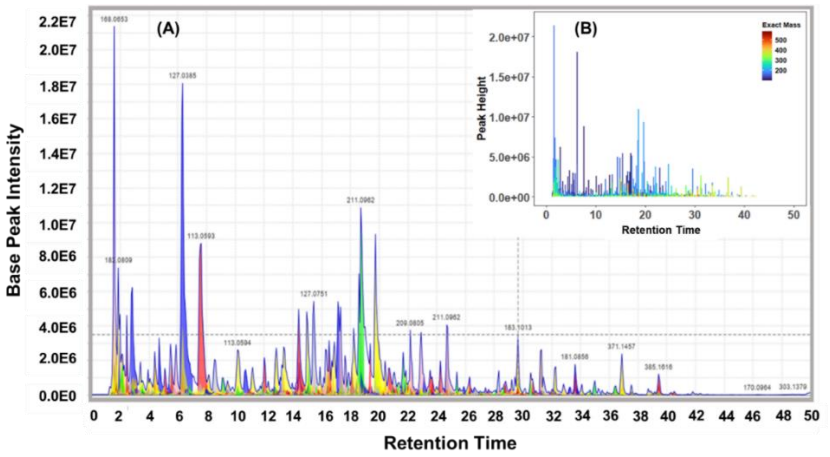

Figure 2: (A) Extracted ion chromatogram of liquid smoke sample and (B) retention time mass spectra of extracted masses with the RT and peak height. Both are positive ion data collected with a mass resolving power of 120,000.

It is known that MZMine2 identifies adduct ion species with two satisfactory conditions. (1) The adduct ion needs to have the same retention time as the protonated species, and (2) the difference between those two ions needs to be equal to the mass of the adduct (Pluskal et al., 2010). So, the analytes exclusively presented as Na^+ adducts (without protonated species) cannot be identified. Our method successfully solved the challenge of accurately identifying multiple adducted species without H adducts and we observed 338 molecular species with 185 molecular formulas exclusively presented as Na^+ adducted species in the +ESI analysis.

Most of the molecules detected in the liquid smoke samples were CHO species and represented 69% and 99% of the total peak area of the identified species in +ESI and –ESI analysis, respectively. Elemental ratios of the identified species follow the reversed phase chromatographic retention order, as the species with relatively low O/C and H/C values have longer retention times, whereas molecules with high O/C or higher H/C values elute early (Figure S5). A larger number of nitrogen-containing formulas were detected in positive ionization mode. These molecules may be valid CHNO/CHN species or ammonium adducts of the molecules in the liquid smoke sample.

So, the N-containing formulas were re-evaluated by removing NH_4^+ from the existing formula to check for non-adducted valid assignments. This led to 95 out of 1771 species with a matching protonated molecular formula at the same retention time suggesting that they were adducts while majority of (1676) species were valid CHN/CHNO assignments.

3.2 Comparison of LC/FT-MS with the direct infusion FT-MS analysis

According to the positive mode analysis 1098 individual molecular formulas, corresponding to 3209 molecular species, were successfully identified including structural isomers. Of these, 970 molecular formulas were common among the molecular formulas obtained from direct infusion mass spectrometric analysis of the same liquid smoke sample. This represents 99% of total peak area obtained from the molecular species identified by liquid chromatographic ion mass extraction and 95% with respect to the total abundance of the direct infusion mass spectrometric (DI-FT-MS) identified species (Figure 3A). These 970 common molecular formulas correspond to 3043 total molecular species due to the numerous isomers present. This is an inherent advantage of the LC/FT-MS because isomeric structures corresponding to only one molecular formula can be separated. Most of the commonly identified species were CHO species corresponding to about 70% of the total LC/FT-MS peak area and 86% of the direct infusion FT-MS peak abundance. The LC/FT-MS method uniquely detected 128 molecular formulas with 166 molecular species, and they were responsible for 1% of the total peak area of the LC/FT-MS identified species.

In negative mode ESI analysis, 1697 molecular species corresponding to 635 individual molecular formulas were obtained. Of these 578 molecular formulas corresponding to 1619 molecular species are common with the –ESI direct infusion analysis. This represents 98% of the total LC/FT-MS peak area and 96% of the total direct infusion FT-MS peak abundance (Figure 3B). CHO species accounted for 97% of the LC/FT-MS peak area of the commonly identified species. The summary of the molecular formulas detected in both methods is shown in Table1.

The direct infusion analysis identified a significant number of unique molecular formulas, relative to the liquid chromatography analysis. Specifically, 1119 and 1183, in +ESI and −ESI analysis, respectively (Figure 3). These direct infusion species were less intense than the majority of the common species (Figure S4) likely due to lower concentrations or lower ionization efficiencies. Many of the formulas uniquely identified species with direct infusion FT-MS methods

had relatively higher masses with relatively higher C and O content in comparison to the commonly detected species by both methods (Figure S4). While we tuned the Orbitrap to be as sensitive as possible to higher mass-to-charge values and liquid chromatography can improve ionization by minimizing the matrix effects, there is an increased sensitivity to lower mass-to-charge values as has been seen in (Hawkes et al., 2020). This observation in conjunction with potentially lower ionization efficiency, possible low species concentration, and the additional dilution encountered during LC/FT-MS analysis (due to 5 uL injection) likely explain why these species are not observed.

Although many of the compounds observed in direct infusion are not also observed in LC/FT-MS, those compounds unique to direct infusion make up only ~5% of the total abundance in direct infusion, highlighting that although we observe fewer compounds, we are still able to see the compounds that make up 95% of the total abundance in this sample and may be most important to the overall composition. The trade-off between more formulas and more specific information about the major species in a sample needs to be considered. Despite this, the LC/FT-MS analysis provided several additional pieces of information that were not obtained from direct infusion mass spectrometry, including: C18 column retention, isomeric complexity, the ability to automate fragmentation scans for structural analysis, and better semi-quantitative information, which allows us to assess the changes in compound concentration in a relative way (i.e., a compound has higher or lower concentration in one sample vs. another).

The proposed approach is not intended as a substitute for direct infusion mass spectrometry. Instead, this method enhances our understanding of the sample's composition with the relative quantities, isomers, and tentative structure. Therefore, by performing both methods, more comprehensive analysis of a complex sample can be achieved.

Table 1. Number of species detected by LC/FT-MS and direct infusion FT-MS methods.

	Number of Formulas							
	CHO		**CHNO**		**CHN**		**CH**	
	LC-MS	DI-MS	LC-MS	DI-MS	LC-MS	DI-MS	LC-MS	DI-MS
Positive	405 (1425)	952	605 (1495)	1329	83 (276)	116	5 (13)	50
Negative	579 (1619)	1428	56 (78)	330	-	-	-	-

The number of the species including isomers are in the parentheses.

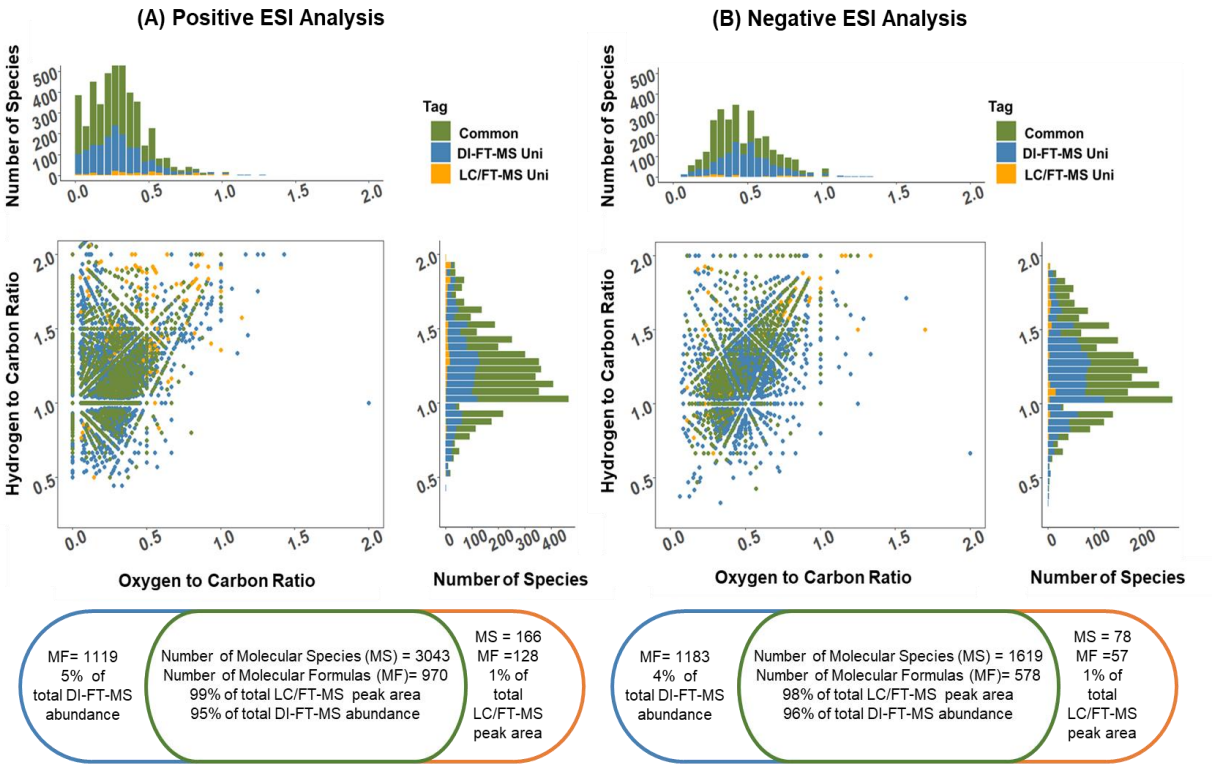

Figure 3: Comparison of the elemental ratios identified with LC/FT-MS analysis and direct infusion FT-MS analysis and the distribution of the identified species with O/C and H/C ratios, and comparison of the number of formulas, peak area, and peak abundances identified each method. (A) Positive ESI analysis and (B) Negative ESI analysis

Overall, the results obtained from LC/FT-MS, and direct infusion FT-MS agreed with each other in both ionization methods. This further highlights the reliability of using the novel method in environmental organic mixture analysis.

Although the spectral abundances obtained from direct infusion ESI-FT-MS are not necessarily quantitative representation of the concentration of species, relative quantitative information such as normalized peak intensity abundances and weighted averages are widely used in the literature for sample trend comparison in environment complex mixture analysis (Brege et al., 2018; Hawkins et al., 2018; Wozniak et al., 2014; Xu et al., 2020)

In our study, we observed an interesting correlation between the chromatographic peak area (cumulative peak area of its all isomers) and the direct infusion peak abundance in commonly detected species in both methods (Figure 4). According to the Pearson correlation coefficient value ($r=\sqrt{r^2}$), log (LC/FT-MS Peak Area) and log (DI-FT-MS Peak Intensity Abundance) shown a positive linear relationship ($r>0$). However, this is not sufficient to predict the concentration of the

species. Further method development with analytical standards is recommended and is beyond the scope of this study. Nonetheless, the correlation is valuable because it suggests that direct infusion peak intensities are not strongly impacted by matrix effects and thus may be related to the actual abundance of the analytes. This observation supports the concept of using peak intensity abundance for giving semi-quantitative insight of direct infusion peak intensities when comparing the environmental samples as has been done in several other studies in the literature (M. Brege et al., 2018; Hawkins et al., 2018; Wozniak et al., 2014; Xu et al., 2020).

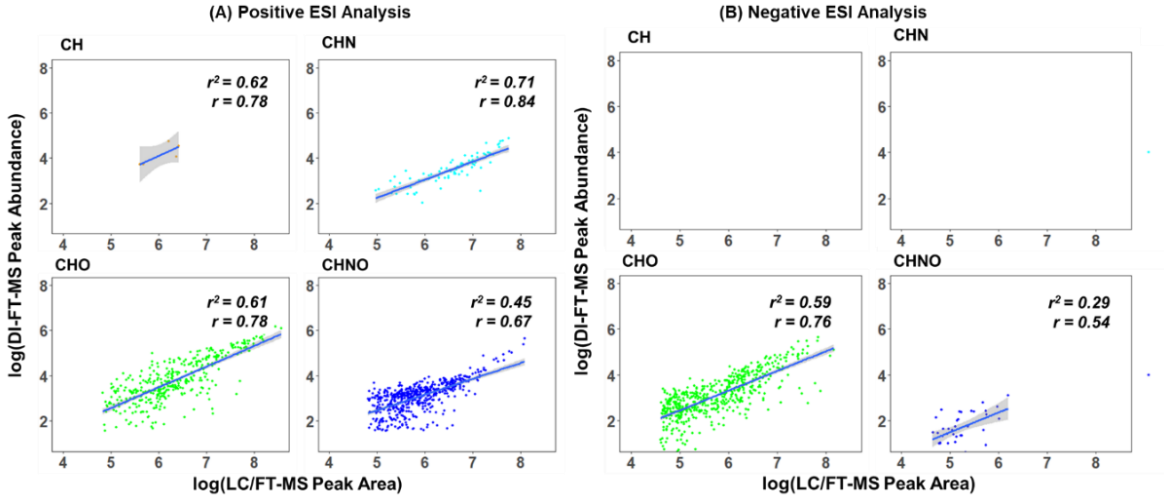

Figure 4: Variation of the logarithmic value of direct infusion-FT-MS peak abundance with logarithmic value of LC/FT-MS peak area with their corresponding trendlines, r^2 values and, Pearson correlation coefficient r. (A) Positive ESI analysis and (B) Negative ESI analysis

3.3 Comparison of two ionization modes

Pairwise comparison of the molecules obtained from both ionization mode detected 344 molecular formulas in common. These common species are detected between H/C ratio of 0.75 -1.50 and O/C ratio of 0.0 - 0.5 (Figure 5) and they might correspond to the phenolic compounds obtained the decomposition of lignin which are expected to be present in smoke (Rivas-Ubach et al., 2018). In accordance with the literature, the trend of compounds uniquely identified using the –ESI analysis are the

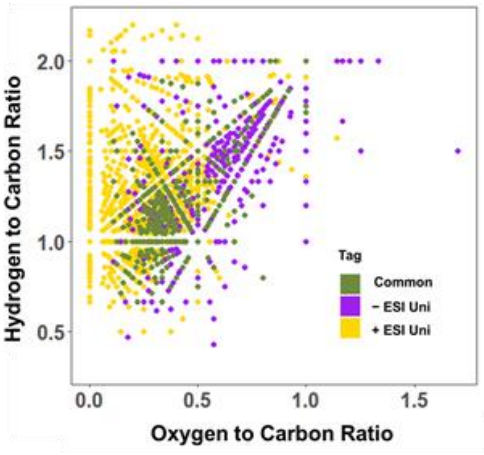

Figure 5: Elemental ratios of the molecular formulas presented in the liquid smoke sample detected by the new proposed LC/FT-MS analysis in both positive and negative ESI mode.

relatively highly oxygenated species with high O/C ratio values and most of the formulas unique to the positive mode are consist with CHN and CHNO species present in the sample. Mesquite is a leguminous tree and has the ability of fixing nitrogen in the environment (Lyons and Rector, 2015) That may be the reason for having noticeable amount of CHN and CHNO molecules (~20% of the total peak area abundance) in the liquid smoke sample in the +ESI mode analysis. Additionally, these N species include 8 out of the 18 pyrazine-based chromophores present in biomass burning organic aerosol reported in the literature (Hawkins et al., 2018). These molecules can be produced with the Maillard type reaction of amino acids and the reducing sugars during the pyrolysis of mesquite plants. Among the 1389 identified molecular formulas from both +ESI and –ESI mode analysis 95 molecular formulas with 155 possible structures are matched with Phenol Explorer database (Rothwell et al., 2013). Of these, 21 structures belonged to the class of flavonoids, 25 belonged to lignans. 54, 50, and 5 are classified as other phenols, phenolic acids, and stilbenes, respectively.

3.4 Detection of isomeric species

Environmental complex mixtures can contain several isomeric species with the same molecular formula. LC/MS analysis is valuable in this scenario to separate the isomeric species to be independently identified. We observed 689 (in +ESI mode) and 341 (–ESI mode) molecular formulas present in at least two different retention times, supporting the existence of the isomers in the liquid smoke sample (Figure S6). MS/MS fragmentation patterns allow for the tentative structural identification of these isomeric molecules. For example, $C_6H_6O_3$ appeared in three different retention times in the +ESI chromatogram (Figure 6). Three of these peaks elute very close to each other with close retention time of 4.4 min, 5.5 min and 6.3 min. Hence, they are likely the isomeric species of $C_6H_6O_3$. The corresponding MS/MS pattern of these peaks showed similarities to the characteristic's fragmentations of benzenetriol reported in the literature (Fahmey et al., 2001) Consequently, the suspected molecules are likely phloroglucinol, hydroxyquinol, and pyrogallol. Predicted octanol to water partition coefficients (logP) from Molinspiration interactive logP calculator, (Molinspiration.com, n.d.) log P values are 0.43, 0.49 and 0.73 respectively. When considering chromatographic affinity of the molecules in reversed phase chromatography, those peaks could be identified as phloroglucinol (at 4.4 min), hydroxyquinol (at 5,5 min), and pyrogallol (at 6.3 min) as shown in Figure 6. Fragmentation of trihydroxybenzene molecules result significant peaks at 81 (MH^+-H_2O-CO), 98 (MH^+-CHO) with the 25% percent normalized collision energy.

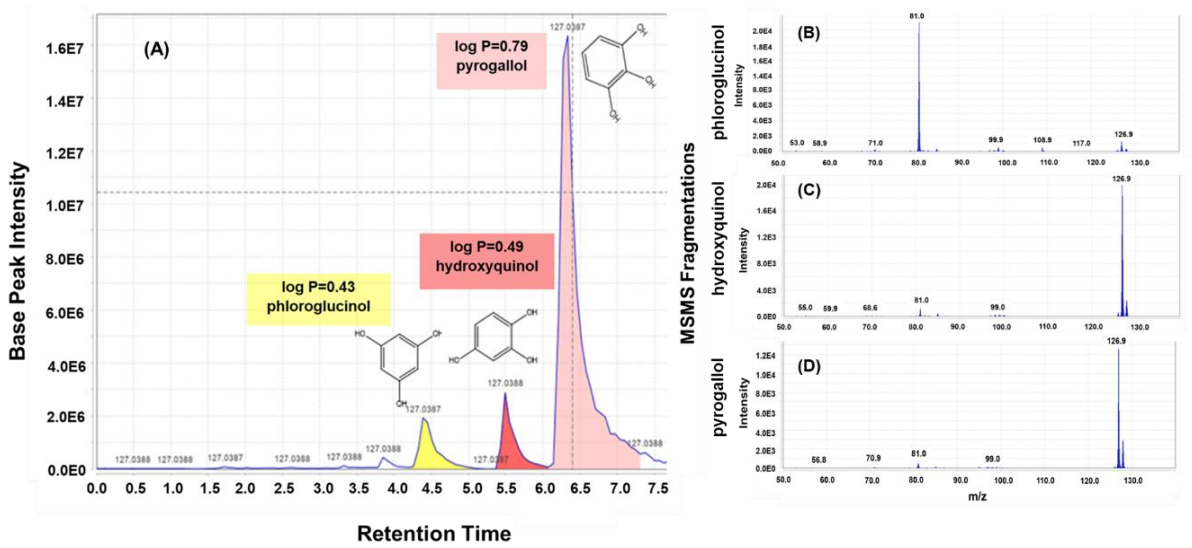

Figure 6: (A) Extracted ion chromatogram of $C_6H_6O_3$ species and MS/MS fragmented spectra of the three isomers with highest intense peaks, (B) phloroglucinol (mz 81.0, 126.9, 99.9, 108.9), (C) phloroglucinol (mz 126.9, 81.0, 99.0, 68.6), (D) pyrogallol (mz 126.9, 81.0, 70.9)

Unlike the other two isomeric species of trihydroxybenzene, the significant fragmentation of pyrogallol (RT at 6.3 min) was not observed with the 25% normalized collision energy CID fragmentation. This could be due to the proximity of the H atom to OH group present in the benzene ring of the molecule. The similar way we could use our method for tentative structural isomeric species identification in biomass burning organic aerosol. This approach allows for the tentative identification of isomeric species in a complex mixture and is significant because there is only very limited database information. Reference standards are presented for detailed identifications. However, none of the tentative structures were confirmed.

3.5 Liquid smoke as a surrogate for biomass burning organic aerosol.

All detected molecular formulas in liquid smoke continuously spanning from C_3 to C_{32} with a DBE (Eq S3) range of 1-15 were observed (Figure S8) and the species' retention time increases with the C number and the DBE value (Figure S8B). Most identified molecular species are defined as olefins based on the modified aromaticity index AI_{mod} (Eq S5) (Figure S7, Figure S8C). Aromatic species were the next most common class detected related to +ESI ionization (Figure S7C, Figure S8). While aliphatic species were the second most common class observed in the –ESI analysis. This may be due to the ionization preference for relatively oxygenated functional groups (Figure S7).

Many molecular formulas identified in our study matched the molecular formulas of known biomass burning organic aerosol chemical fingerprints found in the literature, including levoglucosan, methoxy phenols, and organic acids (Bertrand et al., 2018; Brege et al., 2018; Mazzoleni et al., 2007; Qi et al., 2020). Molecular formulas presented in the liquid smoke sample detected by LC/FT-MS study were compared with an ambient wildfire influenced organic aerosol sample, namely BBO6 (Brege et al., 2021). According to the pairwise comparison with the ambient biomass burning organic aerosol sample, and liquid smoke, 97% of the total peak area (with 1036 out of 1098 total individual formulas) and 78% of the total peak area (with 510 out of 612 total individual formula) matched with each other in +ESI and −ESI analysis respectively (Figure S9). The carbon, DBE continuum, and degree of unsaturation of liquid smoke were analogous to those trends of the ambient biomass burning aerosol in the literature (Brege et al., 2021). Though, unlike the ambient sample the maximum frequency cluster of species were observed with relatively low C number and DBE values (C_8 - C_{12} and DBE 5-6). The C number above 32 or DBE above 15 formulas were not detected in the liquid smoke. This may be due to the removal of water insoluble aromatics or condensed aromatics during liquid smoke production (Montazeri et al., 2013). The species common between liquid smoke and the ambient biomass burning aerosol sample correspond to only 38% and 48% of the ambient biomass burning organic aerosol peak abundances in the positive mode and negative mode respectively (Figure S9). Still, ambient biomass burning organic aerosol sample was more complex than liquid smoke. This could be due to the different levels of oxidation, transformations, and the addition of other pollutants occurring in the natural environment. Sample extracting solvent, and direct infusion analysis of the ambient biomass burning organic aerosol sample also contribute to the relatively low percentage match of the common species with respect to it.

In the research literature, most of the chemical reactions and transformations of organic aerosol are carried out using single compounds or simple mixtures of species related to biomass burning organic aerosol, such as methoxy phenols, vanillin, and levoglucosan-like compounds. (Li et al., 2014; Rodriguez et al., 2022; Zhao et al., 2014). Even though these molecules are ideal for targeted studies, the complex environmental behavior cannot be observed. Currently, no standard reference material is available for the research and development of methods focused on organic aerosol. According to our analysis, over 90% of the molecular formulas detected in the LC/FT-MS analysis of liquid smoke were also detected in the ambient sample. Furthermore, the

characteristics and trends observed in the commercial liquid smoke sample show good agreement with the biomass burning organic aerosol. As a result, commercial liquid smoke may be a valuable substitute for organic aerosol research and method development. Liquid smoke is inexpensive and easily obtainable in large quantities, making it an excellent candidate for further studies, including inter-laboratory comparisons and characterization. These efforts may eventually lead to its development as a standard reference material for biomass-burning organic aerosols.

4. Conclusion

LC/FT-MS nontargeted screening is a powerful semi-quantitative technique for environmental complex mixture analysis. Adapting existing techniques, matching molecular markers in databases, and finding representative samples for method development are some key challenges. To overcome these problems, we present a novel approach to separate and identify the compounds in biomass related organic aerosol using the surrogate mesquite liquid smoke. LC/FT-MS data were processed, and molecular formulas were assigned using the combination of MZmine2.53 and MFAssignR software tools.

Using the new approach, more than 1700 compounds in mesquite liquid smoke have been identified with their isomers. Formula obtained from this approach showed a high consistency with direct infusion LC/FT-MS results, confirming the accuracy of it. A linear correlation between the logarithmic values of the chromatographic mass spectrometric peak area and direct infusion mass spectrometric peak abundance was observed in the study favoring the use of relative quantitative insight from direct infusion mass spectrometric peak intensity abundance. Over 90% of the molecular formulas identified in mesquite liquid smoke matched with the ambient biomass burning organic aerosol sample. So, commercial liquid smoke can be used as a suitable surrogate for organic aerosol research and method developments. Despite the complexity of biomass burning organic aerosol found in the literature, our study foresees the further complexity of organic aerosol with their isomers. Thus, the method significantly improves the molecular identification of biomass combustion aerosol components and obtains some semi quantitative insight. Further, MS/MS fragmentation information can be used for tentative structural identification of the existing compounds, but further studies need to be done for the structural confirmation.

Acknowledgements: The authors thank Dr. Matthew Brege for ambient biomass burning organic aerosol datasets and data interpretation. This research was supported with funds from the US Department of Energy (DE-SC0021168).

References

Bertrand, A., Stefenelli, G., Jen, C.N., Pieber, S.M., Bruns, E.A., Ni, H., Temime-Roussel, B., Slowik, J.G., Goldstein, A.H., El Haddad, I., Baltensperger, U., Prévôt, A.S.H., Wortham, H., Marchand, N., 2018. Evolution of the chemical fingerprint of biomass burning organic aerosol during aging. Atmos Chem Phys 18, 7607–7624. https://doi.org/10.5194/acp-18-7607-2018

Brege, M., Paglione, M., Gilardoni, S., Decesari, S., Cristina Facchini, M., Mazzoleni, L.R., Facchini, M.C., Mazzoleni, L.R., Cristina Facchini, M., Mazzoleni, L.R., 2018. Molecular insights on aging and aqueous-phase processing from ambient biomass burning emissions-influenced Po Valley fog and aerosol. Atmos Chem Phys 18, 13197–13214. https://doi.org/10.5194/acp-18-13197-2018

Brege, M.A., China, S., Schum, S., Zelenyuk, A., Mazzoleni, L.R., 2021. Extreme Molecular Complexity Resulting in a Continuum of Carbonaceous Species in Biomass Burning Tar Balls from Wildfire Smoke. ACS Earth Space Chem 5, 2729–2739. https://doi.org/10.1021/acsearthspacechem.1c00141

Cech, N.B., Krone, J.R., Enke, C.G., 2001. Predicting Electrospray Response from Chromatographic Retention Time. Anal Chem 73, 208–213. https://doi.org/10.1021/ac0006019

Chekmeneva, E., dos Santos Correia, G., Chan, Q., Wijeyesekera, A., Tin, A., Young, J.H., Elliott, P., Nicholson, J.K., Holmes, E., 2017. Optimization and Application of Direct Infusion Nanoelectrospray HRMS Method for Large-Scale Urinary Metabolic Phenotyping in Molecular Epidemiology. J Proteome Res 16, 1646–1658. https://doi.org/10.1021/acs.jproteome.6b01003

Divisekara, DMR Thusitha, " UHPLC/FT-MS NON-TARGETED SCREENING APPROACH FOR BIOMASS BURNING ORGANIC AEROSOL AND LIQUID SMOKE AS BIOMASS BURNING ORGANIC AEROSOL SURROGATE", Dissertation, Michigan Technological University, 2023.

Duarte, R.M.B.O., Matos, J.T. v., Duarte, A.C., 2021. Multidimensional Analytical Characterization of Water-Soluble Organic Aerosols: Challenges and New Perspectives. Applied Sciences 11, 2539. https://doi.org/10.3390/app11062539

Fahmey, M.A., Zayed, M.A., Keshk, Y.H., 2001. Comparative study on the fragmentation of some simple phenolic compounds using mass spectrometry and thermal analyses. Thermochim Acta 366, 183–188. https://doi.org/10.1016/S0040-6031(00)00724-3

Forcisi, S., Moritz, F., Kanawati, B., Tziotis, D., Lehmann, R., Schmitt-Kopplin, P., 2013. Liquid chromatography–mass spectrometry in metabolomics research: Mass analyzers in ultra high pressure liquid chromatography coupling. J Chromatogr A 1292, 51–65. https://doi.org/10.1016/j.chroma.2013.04.017

Ghosh, B., Jones, A.D., 2015. Dependence of negative-mode electrospray ionization response factors on mobile phase composition and molecular structure for newly-authenticated neutral acylsucrose metabolites. Analyst 140, 6522–6531. https://doi.org/10.1039/C4AN02124J

Guillemant, J., Lacoue-Nègre, M., Berlioz-Barbier, A., Albrieux, F., de Oliveira, L.P., Joly, J.F., Duponchel, L., 2021. Towards a new pseudo-quantitative approach to evaluate the ionization response of

nitrogen compounds in complex matrices. Sci Rep 11. https://doi.org/10.1038/s41598-021-85854-7

Hawkes, J.A., D'Andrilli, J., Agar, J.N., Barrow, M.P., Berg, S.M., Catalán, N., Chen, H., Chu, R.K., Cole, R.B., Dittmar, T., Gavard, R., Gleixner, G., Hatcher, P.G., He, C., Hess, N.J., Hutchins, R.H.S., Ijaz, A., Jones, H.E., Kew, W., Khaksari, M., Palacio Lozano, D.C., Lv, J., Mazzoleni, L.R., Noriega-Ortega, B.E., Osterholz, H., Radoman, N., Remucal, C.K., Schmitt, N.D., Schum, S.K., Shi, Q., Simon, C., Singer, G., Sleighter, R.L., Stubbins, A., Thomas, M.J., Tolic, N., Zhang, S., Zito, P., Podgorski, D.C., 2020. An international laboratory comparison of dissolved organic matter composition by high resolution mass spectrometry: Are we getting the same answer? Limnol Oceanogr Methods 18, 235–258. https://doi.org/10.1002/lom3.10364

Hawkins, L., Welsh, H.G., Alexander, M. V., 2018. Evidence for pyrazine-based chromophores in cloud water mimics containing methylglyoxal and ammonium sulfate. Atmos Chem Phys 18, 12413–12431. https://doi.org/10.5194/acp-18-12413-2018

Hohrenk, L.L., Itzel, F., Baetz, N., Tuerk, J., Vosough, M., Schmidt, T.C., 2020. Comparison of Software Tools for Liquid Chromatography–High-Resolution Mass Spectrometry Data Processing in Nontarget Screening of Environmental Samples. Anal Chem 92, 1898–1907. https://doi.org/10.1021/acs.analchem.9b04095

Hollender, J., Schymanski, E.L., Singer, H.P., Ferguson, P.L., 2017a. Nontarget Screening with High Resolution Mass Spectrometry in the Environment: Ready to Go? Environ Sci Technol 51, 11505–11512. https://doi.org/10.1021/acs.est.7b02184

Hollender, J., van Bavel, B., Dulio, V., Farmen, E., Furtmann, K., Koschorreck, J., Kunkel, U., Krauss, M., Munthe, J., Schlabach, M., Slobodnik, J., Stroomberg, G., Ternes, T., Thomaidis, N.S., Togola, A., Tornero, V., 2019. High resolution mass spectrometry-based non-target screening can support regulatory environmental monitoring and chemicals management. Environ Sci Eur 31, 42. https://doi.org/10.1186/s12302-019-0225-x

Houriet, J., Vidar, W.S., Manwill, P.K., Todd, D.A., Cech, N.B., 2022. How Low Can You Go? Selecting Intensity Thresholds for Untargeted Metabolomics Data Preprocessing. Anal Chem. https://doi.org/10.1021/acs.analchem.2c04088

Ijaz, A., Kew, W., China, S., Schum, S.K., Mazzoleni, L.R., 2022. Molecular Characterization of Organophosphorus Compounds in Wildfire Smoke Using 21-T Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry. Anal Chem 94, 14537–14545. https://doi.org/10.1021/acs.analchem.2c00916

Kenward, A., Smith, Deninis.A., Raja, U., 2013. Wildfires and Air Pollution. Climate and Central.

Kim, Sunghwan, Kim, D., Kim, Sungjune, Son, S., Jung, M.J., 2019. Application of Online Liquid Chromatography 7 T FT-ICR Mass Spectrometer Equipped with Quadrupolar Detection for Analysis of Natural Organic Matter. Anal Chem 91, 7690–7697. https://doi.org/10.1021/acs.analchem.9b00689

Li, Y.J., Huang, D.D., Cheung, H.Y., Lee, A.K.Y., Chan, C.K., 2014. Aqueous-phase photochemical oxidation and direct photolysis of vanillin - A model compound of methoxy phenols from biomass burning. Atmos Chem Phys 14, 2871–2885. https://doi.org/10.5194/acp-14-2871-2014

Liigand, J., Kruve, A., Leito, I., Girod, M., Antoine, R., 2014. Effect of mobile phase on electrospray ionization efficiency. J Am Soc Mass Spectrom 25, 1853–1861. https://doi.org/10.1007/s13361-014-0969-x

Lin, L., Yu, Q., Yan, X., Hang, W., Zheng, J., Xing, J., Huang, B., 2010. Direct infusion mass spectrometry or liquid chromatography mass spectrometry for human metabonomics? A serum metabonomic study of kidney cancer. Analyst 135, 2970. https://doi.org/10.1039/c0an00265h

Lin, P., Aiona, P.K., Li, Y., Shiraiwa, M., Laskin, J., Nizkorodov, S.A., Laskin, A., 2016. Molecular Characterization of Brown Carbon in Biomass Burning Aerosol Particles. Environ Sci Technol 50, 11815–11824. https://doi.org/10.1021/acs.est.6b03024

Lyons, R.K., Rector, B., 2015. Mesquite Ecology and Management. US Department of Agriculture.

Mazzoleni, L.R., Zielinska, B., Moosmüller, H., 2007. Emissions of levoglucosan, methoxy phenols, and organic acids from prescribed burns, laboratory combustion of wildland fuels, and residential wood combustion. Environ Sci Technol 41, 2115–2122. https://doi.org/10.1021/es061702c

Michalski, A., Damoc, E., Lange, O., Denisov, E., Nolting, D., Müller, M., Viner, R., Schwartz, J., Remes, P., Belford, M., Dunyach, J.J., Cox, J., Horning, S., Mann, M., Makarov, A., 2012. Ultra high resolution linear ion trap orbitrap mass spectrometer (orbitrap elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. Molecular and Cellular Proteomics 11, 1–11. https://doi.org/10.1074/mcp.O111.013698

Minkus, S., Bieber, S., Letzel, T., 2022. Spotlight on mass spectrometric non-target screening analysis: Advanced data processing methods recently communicated for extracting, prioritizing and quantifying features. Analytical Science Advances 3, 103–112. https://doi.org/10.1002/ansa.202200001

Molinspiration.com, n.d. Molinspiration Cheminformatics free web services [WWW Document]. Molinspiration Cheminformatics free web services. URL https://molinspiration.com

Montazeri, N., Oliveira, A.C.M., Himelbloom, B.H., Leigh, M.B., Crapo, C.A., 2013. Chemical characterization of commercial liquid smoke products. Food Sci Nutr 1, 102–115. https://doi.org/10.1002/fsn3.9

Myers, O.D., Sumner, S.J., Li, S., Barnes, S., Du, X., 2017. One Step Forward for Reducing False Positive and False Negative Compound Identifications from Mass Spectrometry Metabolomics Data: New Algorithms for Constructing Extracted Ion Chromatograms and Detecting Chromatographic Peaks. Anal Chem 89, 8696–8703. https://doi.org/10.1021/acs.analchem.7b00947

Pan, X., Ichoku, C., Chin, M., Bian, H., Darmenov, A., Colarco, P., Ellison, L., Kucsera, T., Da Silva, A., Wang, J., Oda, T., Cui, G., 2020. Six global biomass burning emission datasets: Intercomparison and application in one global aerosol model. Atmos Chem Phys 20, 969–994. https://doi.org/10.5194/acp-20-969-2020

559 Pluskal, T., Castillo, S., Villar-Briones, A., Orešič, M., 2010. MZmine 2: Modular framework for processing,
560 visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics 11,
561 395. https://doi.org/10.1186/1471-2105-11-395

562 Qi, L., Vogel, A., Esmaeilirad, S., Cao, L., Zheng, J., Jaffrezo, J.-L., Fermo, P., Kasper-Giebl, A., Daellenbach,
563 K., Chen, M., Ge, X., Baltensperger, U., Prévôt, A.S., Slowik, J., 2020. One-year characterization of
564 organic aerosol composition and sources using an extractive electrospray ionization time-of-flight
565 mass spectrometer (EESI-TOF). Atmospheric Chemistry and Physics Discussions 1–34.
566 https://doi.org/10.5194/acp-2019-1165

567 Rivas-Ubach, A., Liu, Y., Bianchi, T.S., Tolić, N., Jansson, C., Paša-Tolić, L., 2018. Moving beyond the van
568 Krevelen Diagram: A New Stoichiometric Approach for Compound Classification in Organisms. Anal
569 Chem 90, 6152–6160. https://doi.org/10.1021/acs.analchem.8b00529

570 Rodriguez, A.A., Rafla, M.A., Welsh, H.G., Pennington, E.A., Casar, J.R., Hawkins, L.N., Jimenez, N.G., De
571 Loera, A., Stewart, D.R., Rojas, A., Tran, M.K., Lin, P., Laskin, A., Formenti, P., Cazaunau, M., Pangui,
572 E., Doussin, J.F., De Haan, D.O., 2022. Kinetics, Products, and Brown Carbon Formation by
573 Aqueous-Phase Reactions of Glycolaldehyde with Atmospheric Amines and Ammonium Sulfate.
574 Journal of Physical Chemistry A 126, 5375–5385. https://doi.org/10.1021/acs.jpca.2c02606

575 Rothwell, J.A., Perez-Jimenez, J., Neveu, V., Medina-Remon, A., M'Hiri, N., Garcia-Lobato, P., Manach, C.,
576 Knox, C., Eisner, R., Wishart, D.S., Scalbert, A., 2013. Phenol-Explorer 3.0: a major update of the
577 Phenol-Explorer database to incorporate data on the effects of food processing on polyphenol
578 content. Database 2013, bat070–bat070. https://doi.org/10.1093/database/bat070

579 Schum, S.K., Brown, L.E., Mazzoleni, L.R., 2020. MFAssignR: Molecular formula assignment software for
580 ultrahigh resolution mass spectrometry analysis of environmental complex mixtures. Environ Res
581 191, 110114. https://doi.org/10.1016/j.envres.2020.110114

582 Schum, S.K., Zhang, B., Džepina, K., Fialho, P., Mazzoleni, C., Mazzoleni, L.R., 2018. Molecular and
583 physical characteristics of aerosol at a remote free troposphere site: implications for atmospheric
584 aging. Atmos Chem Phys 18, 14017–14036. https://doi.org/10.5194/acp-18-14017-2018

585 Sekimoto, K., Koss, A.R., Gilman, J.B., Selimovic, V., Coggon, M.M., Zarzana, K.J., Yuan, B., Lerner, B.M.,
586 Brown, S.S., Warneke, C., Yokelson, R.J., Roberts, J.M., De Gouw, J., 2018. High-and low-
587 temperature pyrolysis profiles describe volatile organic compound emissions from western US
588 wildfire fuels. Atmos Chem Phys 18, 9263–9281. https://doi.org/10.5194/acp-18-9263-2018

589 Shiraiwa, M., Ueda, K., Pozzer, A., Lammel, G., Kampf, C.J., Fushimi, A., Enami, S., Arangio, A.M.,
590 Fröhlich-Nowoisky, J., Fujitani, Y., Furuyama, A., Lakey, P.S.J., Lelieveld, J., Lucas, K., Morino, Y.,
591 Pöschl, U., Takahama, S., Takami, A., Tong, H., Weber, B., Yoshino, A., Sato, K., 2017. Aerosol
592 Health Effects from Molecular to Global Scales. Environ Sci Technol 51, 13545–13567.
593 https://doi.org/10.1021/acs.est.7b04417

594 Simon, R., de la Calle, B., Palme, S., Meier, D., Anklam, E., 2005. Composition and analysis of liquid
595 smoke flavouring primary products. J Sep Sci 28, 871–882. https://doi.org/10.1002/jssc.200500009

596 Wozniak, A.S., Willoughby, A.S., Gurganus, S.C., Hatcher, P.G., 2014. Distinguishing molecular
597 characteristics of aerosol water soluble organic matter from the 2011 trans-North Atlantic US
598 GEOTRACES cruise. Atmos Chem Phys 14, 8419–8434. https://doi.org/10.5194/acp-14-8419-2014

599 Xu, J., Hettiyadura, A.P.S., Liu, Y., Zhang, X., Kang, S., Laskin, A., 2020. Regional Differences of Chemical
600 Composition and Optical Properties of Aerosols in the Tibetan Plateau. Journal of Geophysical
601 Research: Atmospheres 125. https://doi.org/10.1029/2019JD031226

602 Zhang, X.W., Li, Q.H., Xu, Z. di, Dou, J.J., 2020. Mass spectrometry-based metabolomics in health and
603 medical science: A systematic review. RSC Adv 10, 3092–3104. https://doi.org/10.1039/c9ra08985c

604 Zhao, R., Mungall, E.L., Lee, A.K.Y., Aljawhary, D., Abbatt, J.P.D., 2014. Aqueous-phase photooxidation of
605 levoglucosan – A mechanistic study using aerosol time-of-flight chemical ionization mass
606 spectrometry (Aerosol ToF-CIMS). Atmos Chem Phys 14, 9695–9705. https://doi.org/10.5194/acp-
607 14-9695-2014

608