Supplement of Atmos. Chem. Phys., 25, 10965–10986, 2025
https://doi.org/10.5194/acp-25-10965-2025-supplement

Atmospheric Chemistry and Physics
Open Access
EGU

Supplement of

New water-soluble, toxic tracers of wood burning identified in fine brown carbon aerosol using a non-target approach

Vinh Nguyen et al.

Correspondence to: Bartłomiej Witkowski (bwitk@chem.uw.edu.pl)

S1. Materials and reagents

S1.1. LC/MS solvents and eluent additives

Acetonitrile, methanol, water, and formic acid (98%) were purchased from Sigma–Aldrich (Germany). Acetone was purchased from Fisher Scientific (Chemat, Poland).

S1.2. List of surrogate standards used for semi-quantification

Isatin (99%), succinic (≥99%), adipic (99%), and caffeic acid (99%) were purchased from Sigma – Aldrich, Germany; fenuron (99%) and isoproturon (99%) were from LGC, England; isoleucine and tryptophan were research grade from Serva, Germany; tetradecanedioic acid (98%) was purchased from Alfa Aesar, USA; sebacic acid (≥95%) was purchased from Honeywell Fluka, USA.

S1.3. Standards used for testing NTA workflows

4-nitrophenol (≥99%), 2-nitrophloroglucinol (99%), 4-nitrocatechol (≥96%), itaconic acid (≥99%), pimelic acid (98%), adipic acid (99%), azelaic acid (98%), suberic acid (98%), camphoric acid (99%), malic acid (≥99%), succinic acid (≥99%), citric acid (≥99.5%), sulfanilic acid (99%), 2-furoic acid (98%), 3-methyl-2-furoic acid (97%), 5-methyl-2-furoic acid (97%), 4-acetyl butyric acid (97%), ketopinic acid (99%), benzoic acid (≥99.5%), 2,6-dihydroxybenzoic acid (98%), salicylic acid (≥99%), gallic acid (97.5%), 1,2,4-benzenetriol (99%), 2,4-dihydroxybenzoic acid (97%), 2-phenyl butyric acid (98%), glutaric acid (99%), resorcinol (99%), orcinol (99%), palmitic acid (99%), stearic acid (98%), toluic acid (98%), 4-chlorobenzoic acid (99%), tran-3-hexenoic acid (≥98%) were purchased from Sigma-Aldrich, Germany. 2,5-furandicarboxylic acid (99.5%), 4-nitrobenzoic acid (98%), 3-hydroxy-4-nitrobenzoic acid (99.7%), pivalic acid (99.8%), homovanillic acid (95%), syringaldehyde (99.9%), benzyl succinic acid (99.97%), 2-methoxy benzoic (99.9%) were purchased from Ambeed Inc., USA. Tricarballylic acid (99%), 2-ketoglutaric acid (98%), 5-hexenoic acid (98%), and abietic acid (75%) were purchased from Acros Organic (Thermo Fisher Scientific), USA. 6-heptenoic acid (96%), tartaric acid (99%), and malonic acid (99%) were purchased from Alfa Aesar, USA. 3-furoic acid (98%), and 7-octenoic acid (99%) were purchased from Apollo Scientific, England. Sebaric acid (≥95%) and myristic acid (98%) were purchased from Honeywell Fluka, USA. 3,4-dihydroxyhydrocinnamic acid, ferulic acid, p-coumaric acid, cinnamic acid, caffeic acid, syringic acid, 3,4-dihydroxy benzoic acid were hyper-grade.

Table S1. CAS registry numbers and references justifying the use of selected standards to prepare representative water-soluble BrC mixture

Name	CAS no.	Reference
3,4-Dihydroxyhydrocinnamic acid	1078-61-1	(Moschos et al., 2024)
Ferulic acid	1135-24-6	(Laskin et al., 2025; Szeto et al., 2024)
p-Coumaric acid	501-98-4	(Laskin et al., 2025; Moschos et al., 2024; Li et al., 2021)
Cinnamic acid	140-10-3	(Moschos et al., 2024)
Caffeic acid	331-39-5	(Vicente et al., 2021; Smith et al., 2020a)
2-Phenylbutyric acid	90-27-7	(Evans et al., 2024a)
Benzylsuccinic acid	884-33-3	Surrogate for substituted aromatic acids
Itaconic acid	97-65-4	(Ibáñez and Bauer, 2014)
Pimelic acid	111-16-0	(Choudhary et al., 2023; Cao et al., 2017)
Sebacic acid	111-20-6	(Cao et al., 2017)
Adipic acid	124-04-9	(Burgay et al., 2023; Tomaz et al., 2018; Cao et al., 2017)
Azelaic acid	123-99-9	(Burgay et al., 2023; Sengupta et al., 2020; Cao et al., 2017)
Suberic acid	208-010-9	(Cao et al., 2017; Fan et al., 2016; Mazzoleni et al., 2007)
Camphoric acid	560-09-8	(Amarandei et al., 2023)
Malic acid	6915-15-7	(Hartikainen et al., 2020)
Succinic acid	110-15-6	(Hartikainen et al., 2020; Tomaz et al., 2018)
Citric acid	77-92-9	(Smith et al., 2020a; Ibáñez and Bauer, 2014)
Tricarballylic acid	99-14-9	(Ibáñez and Bauer, 2014)
Glutaric acid	110-94-1	(Burgay et al., 2023; Cao et al., 2017)
Malonic acid	141-82-2	(Cao et al., 2017; Li et al., 2014)
Tartaric acid	87-69-4	(Hems et al., 2021; Hems et al., 2020)
Pivalic acid	75-98-9	Surrogate of branched aliphatic acids, (Simoneit, 2002)
5-Hexenoic acid	1577-22-6	Surrogate of unsaturated acids, (Schauer et al., 2001)
6-Heptenoic acid	1119-60-4	Surrogate of unsaturated acids, (Schauer et al., 2001)
7-Octenoic acid	18719-24-9	Surrogate of unsaturated acids, (Schauer et al., 2001)
Palmitic acid	57-10-3	(Amarandei et al., 2023; Cao et al., 2017; Simoneit, 2002)
Stearic acid	57-11-4	(Cao et al., 2017; Simoneit, 2002)

Name	CAS no.	Reference
Myristic acid	544-63-8	(Simoneit, 2002; Mazzoleni et al., 2007; Cao et al., 2017)
tran-3-Hexanoic acid	1577-18-0	(Cao et al., 2017)
2-Ketoglutaric acid	328-50-7	(Kundu et al., 2010; Shen et al., 2022; Cao et al., 2017)
4-Acetylbutyric acid	3128-06-1	(Kundu et al., 2010; Shen et al., 2022; Cao et al., 2017)
Ketopinic acid	40724-67-2	(Kundu et al., 2010; Shen et al., 2022; Cao et al., 2017)
2-Furoic acid	88-14-2	(Tomaz et al., 2018; Koss et al., 2018; Hems and Abbatt, 2018; Ibáñez and Bauer, 2014; Iinuma et al., 2007)
3-Furoic acid	488-93-7	(Tomaz et al., 2018; Koss et al., 2018; Hems and Abbatt, 2018; Ibáñez and Bauer, 2014; Iinuma et al., 2007)
3-Methyl-2-furoic acid	4412-96-8	(Tomaz et al., 2018; Koss et al., 2018; Hems and Abbatt, 2018; Ibáñez and Bauer, 2014; Iinuma et al., 2007)
5-Methyl-2-furoic acid	1917-15-3	(Tomaz et al., 2018; Koss et al., 2018; Hems and Abbatt, 2018; Ibáñez and Bauer, 2014; Iinuma et al., 2007)
2,5-Furandicarboaylic acid	3238-40-2	(Tomaz et al., 2018; Koss et al., 2018; Hems and Abbatt, 2018; Ibáñez and Bauer, 2014; Iinuma et al., 2007)
4-nitrophenol	100-02-7	(Harrison et al., 2005; Laskin et al., 2025)
2-nitrophloroglucinol	16600-92-3	(Xie et al., 2019)
4-Nitrocatechol	3316-09-4	(Tomaz et al., 2018; Lin et al., 2018; Kroflič et al., 2018)
Benzoic acid	65-85-0	(Song et al., 2018)
2,6-Dihydroxybenzoic acid	303-07-1	(Li et al., 2021; Smith et al., 2020a)
Salicylic acid	69-72-7	(Smith et al., 2020a)
Syringic acid	530-57-4	(Moschos et al., 2024; Hartner et al., 2024b)
4-Nitrobenzoic acid	62-23-7	(Moschos et al., 2024; Yadav et al., 2023; Xie et al., 2019)
3-Hydroxy-4-nitrobenzoic acid	619-14-7	Moschos et al., 2024; Yadav et al., 2023; Xie et al., 2019)
Gallic acid	149-91-7	(Iinuma et al., 2007; Li et al., 2021; Kawamura et al., 2012; Wan et al., 2019; Graham et al., 2002a)
2,4-Dihydroxybenzoic acid	89-86-1	(Kawamura et al., 2012; Wan et al., 2019; Graham et al., 2002a)
3,4-Dihydroxybenzoic acid	99-50-3	(Kawamura et al., 2012; Wan et al., 2019;

Name	CAS no.	Reference
		Graham et al., 2002a)
Toluic acid	99-94-5	Surrogate for substituted benzoic acids
4-Chlorobenzoic acid	74-11-3	Surrogate for chlorinated aromatics, (Kaivosoja et al., 2012)
Anisic acid	579-75-9	(Li et al., 2021; Smith et al., 2020a; Simoneit, 2002)
1,2,4-Benzenetriol	533-73-3	(Hems and Abbatt, 2018; Lin et al., 2017; Daneshvar et al., 2007; Fine et al., 2004)
Resorcinol	108-46-3	(Moschos et al., 2024; Hems et al., 2021; Smith et al., 2020a)
Orcinol	504-15-4	(Hatch et al., 2015; Moschos et al., 2024; Hems et al., 2021; Smith et al., 2020a)
Homovanillic acid	306-08-1	(Laskin et al., 2025; Moschos et al., 2024)
Syringaldehyde	134-96-3	(Moschos et al., 2024; Hartner et al., 2024b; Li et al., 2021)
Sulfanilic acid	121-57-3	Surrogate for sulfonated aromatics and chlorophenols, (Pratt et al., 2011; Wan et al., 2019)
Abietic acid	514-10-3	Representative resin acid, (Zhong et al., 2023; Vicente et al., 2021; Li et al., 2021; Sengupta et al., 2020)

The general characteristics of surrogate standards used to evaluate the workflows were derived as described in Section S4 and are listed in Table S2. Covering the high molecular diversity of the water-extractable BrC with commercially available standards is challenging. At the same time, values of Kendrick Mass and Kendrick Mass Defect, double-bond equivalent, O/C and H/C ratios, C+N, and C numbers for the selected standards are all within the range of the values presented in Fig. 6 in the main text, for the main components of water-extractable BrC.

Table S2. Elemental formulas and general properties of model compounds

Name	Elemental composition	O/C	H/C	Kendrick Mass	Kendrick Mass Defect	DBE	C+N	Oxidation State of carbon
3,4-Dihydroxyhydrocinnamic acid	$C_9H_{10}O_4$	0.4	1.1	181.9	-0.1	5.0	9.0	-0.2
Ferulic acid	$C_{10}H_{10}O_4$	0.4	1.0	193.8	-0.2	6.0	10.0	-0.2
p-Coumaric acid	$C_9H_8O_3$	0.3	0.9	163.9	-0.1	6.0	9.0	-0.2
Cinnamic acid	$C_9H_8O_2$	0.2	0.9	147.9	-0.1	6.0	9.0	-0.4
Caffeic acid	$C_9H_8O_4$	0.4	0.9	179.8	-0.2	6.0	9.0	0.0
2-Phenylbutyric acid	$C_{10}H_{12}O_2$	0.2	1.2	163.9	-0.1	5.0	10.0	-0.8
Benzylsuccinic acid	$C_{11}H_{12}O_4$	0.4	1.1	207.8	-0.2	6.0	11.0	-0.4

Name	Elemental composition	O/C	H/C	Kendrick Mass	Kendrick Mass Defect	DBE	C+N	Oxidation State of carbon
Itaconic acid	$C_5H_6O_4$	0.8	1.2	129.9	-0.1	3.0	5.0	0.4
Pimelic acid	$C_7H_{12}O_4$	0.6	1.7	159.9	-0.1	2.0	7.0	-0.6
Sebacic acid	$C_{10}H_{18}O_4$	0.4	1.8	201.9	-0.1	2.0	10.0	-1.0
Adipic acid	$C_6H_{10}O_4$	0.7	1.7	145.9	-0.1	2.0	6.0	-0.3
Azelaic acid	$C_9H_{16}O_4$	0.4	1.8	187.9	-0.1	2.0	9.0	-0.9
Suberic acid	$C_8H_{14}O_4$	0.5	1.8	173.9	-0.1	2.0	8.0	-0.8
Camphoric acid	$C_{10}H_{16}O_4$	0.4	1.6	199.9	-0.1	3.0	10.0	-0.8
Malic acid	$C_4H_6O_5$	1.3	1.5	133.9	-0.1	2.0	4.0	1.0
Succinic acid	$C_4H_6O_4$	1.0	1.5	117.9	-0.1	2.0	4.0	0.5
Citric acid	$C_6H_8O_7$	1.2	1.3	191.8	-0.2	3.0	6.0	1.0
Tricarballylic acid	$C_6H_8O_6$	1.0	1.3	175.8	-0.2	3.0	6.0	0.7
Glutaric acid	$C_5H_8O_4$	0.8	1.6	131.9	-0.1	2.0	5.0	0.0
Malonic acid	$C_3H_4O_4$	1.3	1.3	103.9	-0.1	2.0	3.0	1.3
Tartaric acid	$C_4H_6O_6$	1.5	1.5	149.8	-0.2	2.0	4.0	1.5
Pivalic acid	$C_5H_{10}O_2$	0.4	2.0	102.0	0.0	1.0	5.0	-1.2
5-Hexenoic acid	$C_6H_{10}O_2$	0.3	1.7	113.9	-0.1	2.0	6.0	-1.0
6-Heptenoic acid	$C_7H_{12}O_2$	0.3	1.7	127.9	-0.1	2.0	7.0	-1.1
7-Octenoic acid	$C_8H_{14}O_2$	0.3	1.8	141.9	-0.1	2.0	8.0	-1.3
Palmitic acid	$C_{16}H_{32}O_2$	0.1	2.0	256.0	0.0	1.0	16.0	-1.8
Stearic acid	$C_{18}H_{36}O_2$	0.1	2.0	284.0	0.0	1.0	18.0	-1.8
Myristic acid	$C_{14}H_{28}O_2$	0.1	2.0	228.0	0.0	1.0	14.0	-1.7
tran-3-Hexanoic acid	$C_6H_{10}O_2$	0.3	1.7	113.9	-0.1	2.0	6.0	-1.0
2-Ketoglutaric acid	$C_5H_6O_5$	1.0	1.2	145.9	-0.1	3.0	5.0	0.8
4-Acetylbutyric acid	$C_6H_{10}O_3$	0.5	1.7	129.9	-0.1	2.0	6.0	-0.7
Ketopinic acid	$C_{10}H_{14}O_3$	0.3	1.4	181.9	-0.1	4.0	10.0	-0.8
2-Furoic acid	$C_5H_4O_3$	0.6	0.8	111.9	-0.1	4.0	5.0	0.4
3-Furoic acid	$C_5H_4O_3$	0.6	0.8	111.9	-0.1	4.0	5.0	0.4
3-Methyl-2-furoic acid	$C_6H_6O_3$	0.5	1.0	125.9	-0.1	4.0	6.0	0.0
5-Methyl-2-furoic acid	$C_6H_6O_3$	0.5	1.0	125.9	-0.1	4.0	6.0	0.0
2,5-Furandicarboaylic acid	$C_6H_4O_5$	0.8	0.7	155.8	-0.2	5.0	6.0	1.0
4-nitrophenol	$C_6H_5NO_3$	0.5	0.8	138.9	-0.1	5.0	7.0	0.2
2-nitrophloroglucinol	$C_6H_5NO_5$	0.8	0.8	170.8	-0.2	5.0	7.0	0.8
4-Nitrocatechol	$C_6H_5NO_4$	0.7	0.8	154.8	-0.2	5.0	7.0	0.5
Benzoic acid	$C_7H_6O_2$	0.3	0.9	121.9	-0.1	5.0	7.0	-0.3

Name	Elemental composition	O/C	H/C	Kendrick Mass	Kendrick Mass Defect	DBE	C+N	Oxidation State of carbon
2,6-Dihydroxybenzoic acid	$C_7H_6O_4$	0.6	0.9	153.9	-0.1	5.0	7.0	0.3
Salicylic acid	$C_7H_6O_3$	0.4	0.9	137.9	-0.1	5.0	7.0	0.0
Syringic acid	$C_9H_{10}O_5$	0.6	1.1	197.8	-0.2	5.0	9.0	0.0
4-Nitrobenzoic acid	$C_7H_5NO_4$	0.6	0.7	166.8	-0.2	6.0	8.0	0.4
3-Hydroxy-4-nitrobenzoic acid	$C_7H_5NO_5$	0.7	0.7	182.8	-0.2	6.0	8.0	0.7
Gallic acid	$C_7H_6O_5$	0.7	0.9	169.8	-0.2	5.0	7.0	0.6
2,4-Dihydroxybenzoic acid	$C_7H_6O_4$	0.6	0.9	153.9	-0.1	5.0	7.0	0.3
3,4-Dihydroxybenzoic acid	$C_7H_6O_4$	0.6	0.9	153.9	-0.1	5.0	7.0	0.3
Toluic acid	$C_8H_8O_2$	0.3	1.0	135.9	-0.1	5.0	8.0	-0.5
4-Chlorobenzoic acid	$C_7H_5ClO_2$	0.3	0.7	120.9	-0.1	5.5	7.0	-0.1
Anisic acid	$C_8H_8O_3$	0.4	1.0	151.9	-0.1	5.0	8.0	-0.3
1,2,4-Benzenetriol	$C_6H_6O_3$	0.5	1.0	125.9	-0.1	4.0	6.0	0.0
Resorcinol	$C_6H_6O_2$	0.3	1.0	109.9	-0.1	4.0	6.0	-0.3
Orcinol	$C_7H_8O_2$	0.3	1.1	123.9	-0.1	4.0	7.0	-0.6
Homovanillic acid	$C_9H_{10}O_4$	0.4	1.1	181.9	-0.1	5.0	9.0	-0.2
Syringaldehyde	$C_9H_{10}O_4$	0.4	1.1	181.9	-0.1	5.0	9.0	-0.2
Sulfanilic acid	$C_6H_7NO_3S$	0.5	1.2	140.9	-0.1	4.0	7.0	-0.2
Abietic acid	$C_{20}H_{30}O_2$	0.1	1.5	301.9	-0.1	6.0	20.0	-1.3
Average	**C number 7.7**	**0.5**	**1.2**	**158.9**	**-0.1**	**3.8**	**7.8**	**-0.2**
2SD	**6.3**	**0.6**	**0.8**	**82.1**	**0.1**	**3.2**	**6.2**	**1.5**

S1.4. Biomass fuel

Mixed wood pellets with a diameter of 5 mm and a length of 5-20 mm and 8% water content were used as fuel.

S1.5. High-purity gasses

Nitrogen (≥99.999%) was supplied by Multax (Stare Babice, Poland).

S2. Semi-quantification using surrogate standards

The five surrogate standards (for each ionization mode) were chosen to cover the range of retention times of water-soluble organics in BrC. Initial tests were performed on a larger number of molecules, and the list was narrowed down, prioritizing the standards with good chromatographic properties under the LC/MS analysis conditions used – Table S3.

Table S3. Surrogate standards used for semi-quantitative analyses

Name	t_r (min)	t_r range covered	m/z	Ionization mode
Isoleucine	2.15	0-5	132.10249	ESI(+)
Trytophan	10.27	5-15	205.09740	
Isatin	17.52	15-25	148.03951	
Fenuron	20.38	25-35	165.10260	
Isoproturon	38.45	>35	207.14931	
Succinic acid	2.76	0-5	117.01790	ESI(-)
Adipic acid	9.79	5-15	129.01810	
Caffeic acid	17.94	15-25	179.03329	
Sebacic acid	26.28	25-35	201.11160	
Tetradecanedioic acid	39.33	>35	257.17499	

For the compounds detected in BrC_{aq}, the accuracy of the semi-quantitation method was assessed by comparison with the concentrations obtained using authentic standards – Table S4 (Evans et al., 2024b).

Table S4. Quantification results obtained using authentic and surrogate standards for the compounds detected in water-soluble BrC

No.	Ion mode	Compound detected in BrC	Concentration (mg/L)		$Bias = (\frac{C1 - C2}{C1}, \%)$
			Authentic (C1)	Surrogate (C2)	
1	(-)	Glutaric acid	5.0	7.4	44
2		Coumaric acid	15.1	18.9	25
3		Benzoic acid	2.2	1.1	-50
4		Azelaic acid	12.7	10.5	-17
5		Suberic acid	4.1	2.1	-49
6		Homovanillic acid	3.4	3.1	-8.8

No.	Ion mode	Compound detected in BrC	Concentration (mg/L)		$Bias = (\frac{C1 - C2}{C1}, \%)$
			Authentic (C1)	Surrogate (C2)	
7		Orcinol	3.3	4.0	21
8	(+)	Syringic acid	3.5	1.8	-49
9		Cinnamic acid	3.4	2.1	-38
10		Syringaldehyde	18.6	15.4	-17

Of the 10 identified compounds in biomass-burning aqueous extract, concentrations of coumaric acid, glutaric acid, and orcinol were overestimated, whereas the other 7 compounds were underestimated, resulting in an average deviation of -14 % compared to the data obtained with authentic standards - Table S4 (Evans et al., 2024b; Mccord et al., 2018).

The matrix effect for the surrogate standards were calculated with Eq. S1.

$$\%\ Matrix\ Effect = \frac{(Peak\ area_{BrC} - Peak\ area_{standard\ \ mix})}{Peak\ area_{standard\ \ mix}} * 100 \quad \text{(S1)}$$

In Eq. S1, Peak area$_{BrC}$ and Peak area$_{standard\ mix}$ are chromatographic peak areas for the surrogate standards obtained following LC/MS analysis of the spiked filter extracts and standard mixture, respectively. Table S5 summarizes the matrix effect values for surrogate compounds measured in both ionization modes for concentrations between 1 and 20 mg/L.

Table S5. Matrix effects for the selected surrogate standards across modes across different concentration levels obtained with Eq. S1

Concentration (mg/L)	Negative ion mode			Positive ion mode			
	Adipic acid	Caffeic acid	Sebacic acid	Trytophan	Isatin	Fenuron	Isoproturon
1	19	10	9.2	-20	-40	-20	-8.6
2	8.4	2.4	3.9	-13	-32	-19	-7.5
5	1.4	6.9	4.6	-8.2	-27	-20	-9.2
10	6.9	3.4	3.5	-8.8	-22	-14	-8.4
20	2.0	9.3	8.4	-11	-13	-11	-7.2

The results indicate that the matrix effect for all compounds remained below 20%, except isatin, which exhibited higher suppression ranging from -13% to -40%. According to the criteria defined in SANTE/11312/2021v2, these results indicate that matrix effects were generally within acceptable limits for most analytes in this workflow (Pihlström et al., 2021). The data presented in Table S5 strongly indicate that the effects of sample matrix, although non-negligible, are generally lower than the quantification uncertainty resulting from the use of surrogate standards.

S3. Analysis of MS and MS/MS data

Table S6. Databases used in the evaluated workflows

Software name	MS databases used	Number of spectra (or unique compounds)	Ref.
MS-DIAL MZmine MS-FINDER	Mass Bank of North America, consisting of spectra in both positive and negative ionization modes.	2,079,804 spectra	(Mona, 2024)
	MS-DIAL, MSMS_Public_ExpBioInsilico_NEG_VS19	53,337 records 15,100 unique compounds	(Ms-Dial, 2024)
	MS-DIAL, MSMS_Public_ExpBioInsilico_Pos_VS19	326,575 records 16,746 unique compounds	(Ms-Dial, 2024)
	Local databases for only MS-FINDER: human metabolites (HMDB, Urine, Saliva, Feces, Serum, CSF, SMPDB), lipids (LipidMAPS), yeast (YMDB), E.coli (ECMDB), bovine (BMDB), drug (Drugbank), food (FooDB), plants (PlantCyc), biomolecules (ChEBI, PubChem), toxins (T3DB), environmental compounds (STOFF), blood exposome (BLEXP), and natural products (NPA, NANPDB, COCONUT, KNApSAcK, UNPD).	45,181 formulas	(Vaniya et al., 2017)
GNPS	All available online databases from the GNPS web-based platform were used, including BERKELEY-LAB, Birmingham – UPLC-MS (NEG&POS), GNPS, HMDB, IQAMDB, LDB, LEAFBOT, Massbank, Massbank EU, MoNA, NEO-MSMS, PNNL-LIPIDS, PSU-MSMLS, RESPECT, SUMNER, TUEBINGEN (natural product), Xanthones-DB, MKDIR, etc.	1,164,393 spectra	(Gnps, 2024)
	Upload databases: + Mass Bank of North America, consisting of spectra in both positive and negative ionization modes + MS-DIAL, MSMS_Public_ExpBioInsilico_NEG_VS17 + MS-DIAL, MSMS_Public_ExpBioInsilico_Pos_VS17	2,079,804 spectra 15,100 unique compounds 16,746 unique compounds	(Mona, 2024) (Ms-Dial, 2024)
MetaboAnalyst	All databases, including HMDB experimental, HMDB predicted, GNPS, MINEs, LIPIDblast, MoNA, Massbank, RIKEN, ReSpect, MS-DIAL, and BMDMS were used for compound matching and annotation	10,420,215 MS2 records (1,551,012 unique compounds)	(Pang et al., 2024; Metaboanalyst, 2024a)
CFM-ID	All online databases were used, including ChEBI, Drugbank, DSSTox, ECMDB, FooDB, HMDB, KEGG, LIPID MAPS, Massbank JP/EU, MoNA, NP-MRD, STOFF-IDENT, YMDB) and experimental spectra (HMDB, Massbank JP/EU, MoNA, TMIC.	144,536 (experimental) 8,866,848 (predicted)	(Databases, 2024)

S3.1. Mass spectrometric data processing workflows tested

The final, best-performing NTA workflows, based on MS-DIAL and MS-FINDER, are described in the main text – section 2.3.

S3.1.1. CFM-ID

CFM-ID (v4.0) is an online platform employing Single Energy Competitive Fragmentation Modeling for compound annotation via MS-spectra processing (Wang et al., 2022; Wang et al., 2021). CFM-ID which integrates handwritten and machine-learned rules. Here, CFM-ID was utilized to identify model compounds (Table S2) using MS/MS spectra acquired at low, medium, and high CE. The model ranks candidates based on their similarity to the inputted spectra. The ionization method on the CFM-ID 4.0 webpage was set to ESI; $[M-H]^-$ and $[M+H]^+$ adducts were selected. The parent ion mass of each compound in the spectral dataset was entered, with a candidate mass tolerance of 50 ppm, a candidate limit of 20, and a dice score function.

S3.1.2. MetaboAnalyst

MetaboAnalyst is a web-based, metabolomics platform (Metaboanalyst, 2024b; Pang et al., 2024). 20 MS/MS spectra were used at a time to ensure the timely completion of the search and any spectra lacking MS/MS information were excluded. Precursor and fragment mass tolerances were 10 ppm and 20 ppm, respectively. The ion modes (negative and positive) were selected. “Neutral loss” ion was enabled to mainly characterize the unknown compounds. The dot product matching function was selected for spectral similarity scoring.

S3.1.3. GNPS

GNPS is a web-based platform for MS and MS/MS data mining. In this work, the MS/MS spectra of model compounds were compared against a library of MS/MS spectra generated from structurally characterized metabolites (2024; Wang et al., 2016). The workflow began by setting up basic parameters, including search options, advanced search settings, and filtering options. Input data, comprising MS/MS spectra files and a feature quantification table exported from MS-DIAL, were processed with the precursor ion mass tolerance set at 0.020 Da and the fragment ion mass tolerance at 0.05 Da. The minimum pairs score was set at 0.5, with at least six matched fragment ions, and a minimum of one common fragment ions matched with library data. The score threshold (cut-off) was set at 0.5, search analogs were disabled, and the minimum peak intensity was set to 0.0. Both the filter precursor and filter peak were selected (Nothias et al., 2020).

S3.1.4. MZmine

MZmine (v4.3.0) is a Java-based MS data-mining program (Schmid et al., 2023). Here, the Shimadzu data files in a ".lcd" format were converted into an "mzML" file using the MSconvert tool from ProteoWizard package – Fig. 5 (Chambers et al., 2012). Workflow parameters were set up with MZwizard, including sample introduction (UHPLC), MS analysis type (QTOF), and data acquisition (DDA) (Heuckeroth et al., 2024). In the UHPLC settings section, retention time was cropped to 2-50 minutes, with a maximum peak count of 5000, and minimum scan number of 4. The full width at haft maximum (FWHM) was set at 0.05, with retention time tolerances of 0.05 minutes for intra-sample and 50 min for sample-to-sample comparison. For the QTOF mass analyzer, the absolute intensity was chosen; and m/z tolerance was set at 0.05 Da for scan-to-scan, 0.025 Da for intra-sample, and 0.05 Da for sample-to-sample. Noise threshold and minimum peak height were set to the same values used during the feature detection stage (Fig. 5). In the annotation phase, spectral libraries in ".msp" format, similar to those used in MS-DIAL (see section 2.4.1), were imported. The annotation report was configured to export the identified compound list as a Microsoft Excel file.

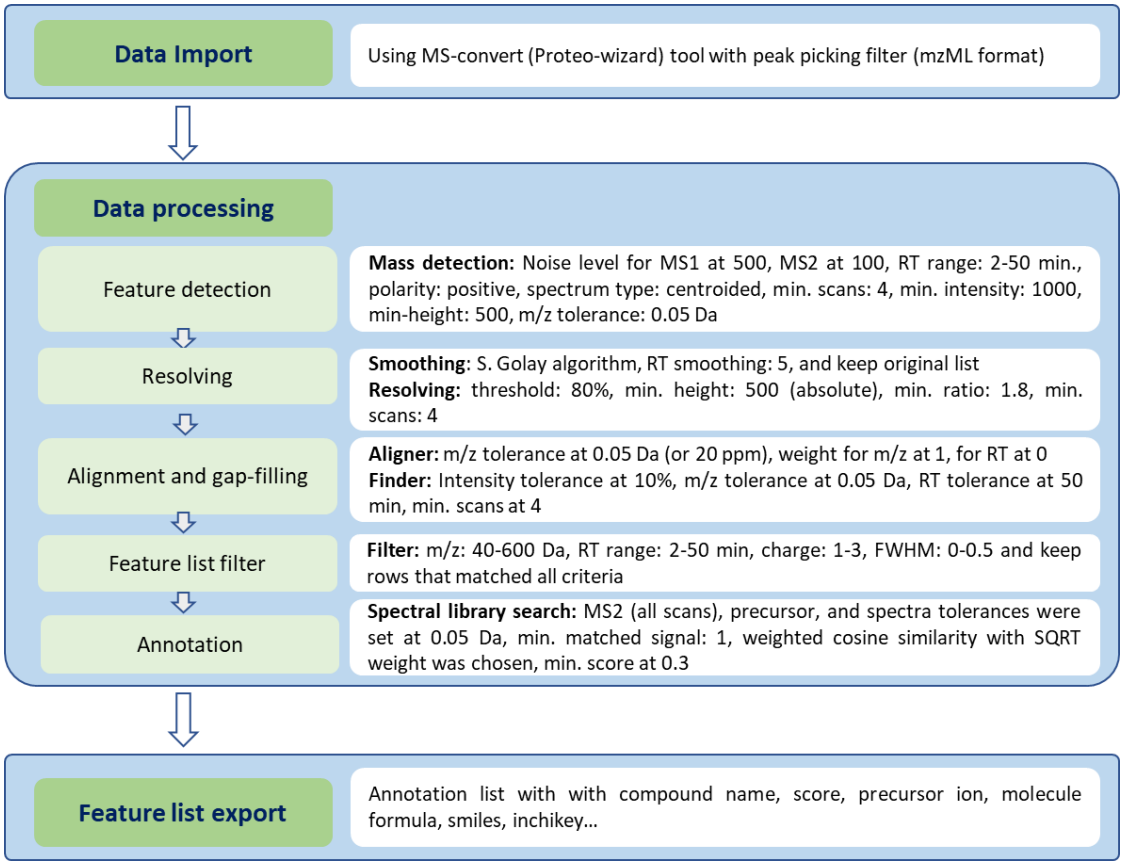

Figure S1. MZmine workflow

S3.1.5. Annotation based on MS1 and MS2 spectra in MS-DIAL

The annotation workflow was adapted from the MS-DIAL scoring approach, which combines MS1 similarity, MS2 similarity, retention time (RT) similarity, and isotope ratio similarity (Tsugawa et al., 2015)

$$Score\_MS-DIAL = \frac{MS2\ similarity\ + MS1\ similarity\ + RT\ similarity\ + 0.5\ isotope\ ratio\ similarity}{3.5} \times 100 \quad \text{(S2)}$$

However, in this work, retention time and isotope ratio information were unavailable. Therefore, the scoring function was simplified to include only MS1 and MS2 similarities

$$Score\_this\ work = \frac{MS2\ similarity\ + MS1\ similarity}{2} \times 100 \quad \text{(S3)}$$

For cases with MS1-only data (no MS2 spectra acquired or no match found in MS2 library), structure suggestions were based solely on high-resolution mass matching (Tsugawa et al., 2015)

$$Score\_MS1\ only = MS1\ similarity = exp\left\{-0.5 \times \left(\frac{Mass_{sample} - Mass_{library}}{\delta_{mass}}\right)^2\right\} \quad \text{(S4)}$$

where δ_{mass} is mass tolerance parameter, set to 0.01 Da in this study.

S4. General characteristics derived from the elemental compositions

Double Bond Equivalents (DBE)

The number of double bonds and ring structures associated with the formula is defined as DBE and is calculated as described in the previous study (Koch and Dittmar, 2006).

$$\mathrm{DBE} = \frac{1}{2}(2 * \mathrm{n_C} - \mathrm{n_H} + \mathrm{n_N} - n_{halogen} + 2) \quad \text{(S5)}$$

Carbon oxidation state (OS_c)

The carbon oxidation state was calculated following Eq. S6 (Moschos et al., 2024).

$$\mathrm{OS}_c = (2 \times \mathrm{n_O} - \mathrm{n_H} + 3 \times \mathrm{n_N})/n_C \quad \text{(S6)}$$

Kendrick mass and Kendrick mass defect (Hughey et al., 2001).

$$\text{Kendrick mass - IUPAC mass} \times (14/14.01565) \quad \text{(S7)}$$

$$\text{Kendrick mass defect = nominal Kendrick mass - exact Kendrick mass} \quad \text{(S8)}$$

Table S7. H/C and O/C ratio values in Van Krevelen Diagrams for each chemical class (Laszakovits and Mackay, 2021; Bhatia et al., 2010)

Chemical class		Con. hydrocarbon	Carbohydrate	Lignin	Lipid	Peptide	Tannin
H/C	Min.	< 1.1	1.53	0.86	1.34	1.34	0.70
	Max.		2.20	1.34	2.18	2.18	1.01
O/C	Min.	< 0.2	0.56	0.21	0.01	0.17	0.16
	Max.		1.23	0.44	0.35	0.48	0.84
Contribution (This work, %)		4.5	3.2	33.1	30.5	8.8	19.8

S5. Results of NTA of water-soluble organic compounds detected in BrC_{aq}

Table S8. List of tentative candidates from ESI(-) mode

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
Matched from MS-DIAL and confirmed authentic standards							
1	2.31	117.0178	Succinic acid	$C_4H_6O_4$	1.01	1	(Sengupta et al., 2020)
2	3.52	145.0488	Adipic acid	$C_6H_{10}O_4$	0.55	1	(Sengupta et al., 2020)
3	5.07	131.0333	Glutaric acid	$C_5H_8O_4$	1.16	1	(Sengupta et al., 2020)
4	14.41	123.0436	Orcinol	$C_7H_8O_2$	1.50	1	(López-Caravaca et al., 2024)
5	14.67	163.0383	p-Coumaric acid	$C_9H_8O_3$	1.80	1	(Oros et al., 2006)
6	16.69	121.0279	Benzoic acid	$C_7H_6O_2$	1.60	1	(Fleming et al., 2018)
7	18.52	181.0488	Homovanillic acid	$C_9H_{10}O_4$	1.83	1	(Fleming et al., 2018), (Moschos et al., 2024)
8	20.70	173.0800	Suberic acid	$C_8H_{14}O_4$	1.89	1	(Sengupta et al., 2020)
9	24.04	187.0960	Azelaic acid	$C_9H_{16}O_4$	1.61	1	(Sengupta et al., 2020)
10	28.72	201.1114	Sebacic acid	$C_{10}H_{18}O_4$	1.30	1	(Sengupta et al., 2020)
Matched from MS-DIAL							
1	7.09	153.0176	2,3-Dihydroxybenzoic acid	$C_7H_6O_4$	0.81	2	(Graham et al., 2002b)
2	13.21	137.0228	4-hydroxybenzoic acid	$C_7H_6O_3$	1.60	2	(Divisekara, 2023)
3	14.41	167.0332	Orsellinic acid	$C_8H_8O_4$	0.80	2	
4	15.71	167.0332	vanillic acid	$C_8H_8O_4$	1.80	2	(Fleming et al., 2018; Moschos et al., 2024)
5	15.96	151.0387	(R)-(-)-mandelic acid	$C_8H_8O_3$	0.84	2	(Pietrogrande et al., 2013)
6	16.82	137.0227	3-hydroxybenzoic acid	$C_7H_6O_3$	1.69	2	(Moschos et al., 2024; Oros and Simoneit, 2001b)
7	16.84	177.0174	Daphnetin (7,8-Dihydroxycoumarin)	$C_9H_6O_4$	1.41	2	
8	16.90	195.0644	Ethyl vanillate	$C_{10}H_{12}O_4$	1.89	2	
9	17.71	177.0173	Esculetin (6,7-Dihydroxycoumarin)	$C_9H_6O_4$	2.03	2	(Jen et al., 2019)
10	18.98	147.0434	p-Coumaraldehyde	$C_9H_8O_2$	1.75	2	(Chan et al., 2020)
11	19.28	161.0226	4-Hydroxycoumarin	$C_9H_6O_3$	1.80	2	(Moschos et al., 2024)
12	19.36	151.0384	Vanillin	$C_8H_8O_3$	0.59	2	(Hartner et al., 2024a; Moschos et al., 2024)

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
13	20.06	167.0331	3,4-dihydroxyphenylacetate	$C_8H_8O_4$	0.90	2	
14	20.12	193.0485	trans-Ferulic acid	$C_{10}H_{10}O_4$	2.14	2	(Fleming et al., 2018; Moschos et al., 2024)
15	20.61	181.0488	Methyl vanillate	$C_9H_{10}O_4$	0.87	2	(Oros et al., 2006)
16	20.94	339.0704	Esculin	$C_{15}H_{16}O_9$	1.71	2	(Divisekara, 2023)
17	21.30	179.0695	Propylparaben	$C_{10}H_{12}O_3$	1.90	2	
18	21.47	165.0539	Ethylparaben	$C_9H_{10}O_3$	1.80	2	
19	21.66	189.0174	3-(Trifluoromethyl)benzoic acid	$C_8H_5F_3O_2$	0.49	2	
20	21.75	195.0644	3',5'-Dimethoxy-4'-hydroxyacetophenone (acetosyringone)	$C_{10}H_{12}O_4$	1.87	2	(Hartner et al., 2024a)
21	22.76	177.0539	2-Methoxycinnamic acid	$C_{10}H_{10}O_3$	1.84	2	(Oros et al., 2006)
22	23.54	162.0304	6-Nitrobenzimidazole	$C_7H_5N_3O_2$	0.50	2	
23	23.54	177.0539	4-Hydroxy-3-methoxycinnamaldehyde	$C_{10}H_{10}O_3$	1.30	2	
24	23.73	207.0643	Sinapoyl aldehyde	$C_{11}H_{12}O_4$	1.98	2	(Chan et al., 2020)
25	25.81	193.0850	Butylparaben	$C_{11}H_{14}O_3$	2.00	2	
26	26.29	223.0592	Sinapic acid	$C_{11}H_{12}O_5$	2.00	2	(Moschos et al., 2024; Divisekara, 2023)
27	37.71	215.1267	Undecanedioic acid	$C_{11}H_{20}O_4$	2.22	2	(Sengupta et al., 2020)
28	38.55	229.1424	Dodecanedioic acid	$C_{12}H_{22}O_4$	2.10	2	(Sengupta et al., 2020)
29	40.63	233.1528	CCMSLIB00004691910	$C_{15}H_{22}O_2$	1.89	2	
30	41.03	311.1992	dl-Norgestrel	$C_{21}H_{28}O_2$	2.50	2	
31	41.16	249.1477	Compound NP-008382	$C_{15}H_{22}O_3$	2.09	2	
32	41.16	299.2571	12(13)-EpOME-[d4]	$C_{18}H_{32}O_3$	4.12	2	
Suggested from MS-DIAL							
1	2.14	199.0207	2-Methyl-4-Chlorophenoxyacetic Acid (MCPA)	$C_9H_9ClO_3$	4.02	3	
2	2.31	257.0260	7-chloro-4,8-dihydroxy-6-methoxy-3-methyl-3,4-dihydroisochromen-1-one	$C_{11}H_{11}ClO_5$	2.99	3	
3	2.49	300.9949	Ellagic acid	$C_{14}H_6O_8$	4.09	3	
4	2.49	139.0021	3-Fluorobenzoate	$C_7H_5FO_2$	2.11	3	
5	2.57	191.0544	D-(-)-quinic acid	$C_7H_{12}O_6$	1.71	3	(Jen et al., 2019)
6	3.58	418.9476	6-O-Methylarthothelin	$C_{15}H_9Cl_3O_5$	1.41	3	
7	3.58	282.9728	[5-hydroxy-3-(hydroxymethyl)-2-oxo-6-propan-2-ylcyclohex-3-en-1-	$C_{15}H_{24}O_5$	4.39	3	

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
			yl] 3-methylbutanoate				
8	3.60	273.0572	Pesticide3_Neburon	$C_{12}H_{16}Cl_2N_2O$	0.18	3	
9	3.78	143.0333	4-Ethoxy-4-oxobut-2-enoic acid	$C_6H_8O_4$	1.71	3	
10	4.16	261.0598	4-Thiouridine	$C_9H_{12}N_2O_5S$	6.10	3	
11	4.93	111.0437	2-ethyl-4-methylimidazole	$C_6H_8O_2$	0.05	3	
12	5.49	109.0279	Catechol	$C_6H_6O_2$	1.60	3	(Divisekara, 2023)
13	5.71	203.0001	Tryptophan	$C_{11}H_{12}N_2O_2$	0.11	3	(Bianco et al., 2016)
14	6.06	255.0468	3-(5-phenylthiophen-2-yl)prop-2-ynyl Acetate	$C_{15}H_{12}O_2S$	1.72	3	
15	7.01	221.0048	isofraxidin	$C_{11}H_{10}O_5$	0.17	3	
16	8.35	201.0750	Diethyloxalpropionate	$C_9H_{14}O_5$	1.40	3	
17	9.22	181.0488	Everninic Acid	$C_9H_{10}O_4$	2.00	3	
18	9.30	211.0230	Zonisamide	$C_8H_8N_2O_3S$	4.69	3	
19	10.40	183.0279	3,4-Dihydroxymandelic acid	C_8H_8O5	1.41	3	
20	10.50	125.0591	Thymine	$C_5H_6N_2O_2$	0.90	3	
21	14.41	235.0204	Harmalol hydrochloride	$C_{12}H_{13}ClN_2O$	6.10	3	
22	14.67	277.0311	2-(4-chlorophenyl)-1-(2,4,6-trihydroxyphenyl)ethanone	$C_{14}H_{11}ClO_4$	3.79	3	
23	14.75	167.0696	Norharman	$C_{11}H_8N_2$	8.12	3	
24	15.11	165.0543	Ethionamide	$C_8H_{10}N_2S$	5.12	3	
25	15.47	165.0176	Phthalic acid	$C_8H_6O_4$	0.81	3	(Sengupta et al., 2020; Moschos et al., 2024)
26	15.71	152.0097	1-Chlorobenzotriazole	$C_6H_4ClN_3$	7.59	3	
27	16.28	205.0851	4-hydroxy-3-(3-methylbut-2-enyl)benzoic acid	$C_{12}H_{14}O_3$	1.92	3	
28	16.69	162.0179	2H-3,1-Benzoxazine-2,4(1H)-dione	$C_8H_5NO_3$	1.8	3	
29	16.80	272.9973	Galacturonate 1-phosphate	$C_6H_{11}O_{10}P$	2.69	3	
30	17.52	233.0801	(2R)-5-methoxy-2-methyl-2,3,8,9-tetrahydrofuro[2,3-h]chromen-4-one	$C_{13}H_{14}O_4$	1.83	3	
31	17.79	181.0488	4-O-Methylphloracetophenone	$C_9H_{10}O_4$	1.71	3	
32	17.92	461.0000	Luteolin 7-O-glucuronide	$C_{21}H_{18}O_{12}$	0	3	
33	17.92	309.0595	Sulfadimethoxine	$C_{12}H_{14}N_4O_4S$	6.71	3	
34	18.02	149.0225	2-Mercaptobenzimidazole	$C_7H_6N_2S$	4.61	3	
35	18.37	167.0331	"2,4,6-Trihydroxyacetophenone"	$C_8H_8O_4$	0.9	3	
36	18.52	317.0225	Fragilin	$C_{16}H_{11}ClO_5$	2.5	3	
37	19.04	233.0048	Riluzole	$C_8H_5F_3N_2OS$	4.81	3	

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
38	19.74	191.0331	4-Methyldaphnetin	$C_{10}H_8O_4$	2	3	
39	20.61	166.0252	6-Thioguanine	$C_5H_5N_5S$	5.91	3	
40	20.70	263.0499	2',2'-Difluoro-2'-deoxyuridine	$C_9H_{10}F_2N_2O_5$	1.41	3	
41	20.70	195.0626	2-Hydroxyphenanzine	$C_{12}H_8N_2O$	6.21	3	
42	20.79	347.0363	Inosine-5'-monophosphate	$C_{10}H_{13}N_4O_8P$	1	3	
43	20.92	177.0538	4-Methoxycinnamic acid	$C_{10}H_{10}O_3$	1.93	3	(Oros et al., 2006)
44	20.94	153.0540	4-Hydroxy-3-methoxybenzyl alcohol	$C_8H_{10}O_3$	0.99	3	
45	21.90	247.0958	Olivetonide	$C_{14}H_{16}O_4$	1.21	3	
46	22.09	175.0381	7-Hydroxy-4-methylcoumarin	$C_{10}H_8O_3$	2	3	
47	22.96	135.0433	3-Methylbenzoic acid	$C_8H_8O_2$	1.9	3	(Smith et al., 2020b)
48	23.23	263.0542	4-hydroxy-3-[(4-hydroxy-6-methyl-2-oxopyran-3-yl)methyl]-6-methylpyran-2-one	$C_{13}H_{12}O_6$	1.92	3	
49	23.67	161.0589	1,2-Dihydro-1,2-naphthalenediol	$C_{10}H_{10}O_2$	0.9	3	
50	24.04	277.0651	3-(5,7-dimethoxy-4-oxochromen-2-yl)propanoic acid	$C_{14}H_{14}O_6$	4.79	3	
51	24.04	271.0480	Alternariol monomethyl ether	$C_{15}H_{12}O_5$	5.64	3	
52	24.04	435.1295	1-[2,4-dihydroxy-3-[(2S,3R,4R,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]phenyl]-2-hydroxy-3-(4-hydroxyphenyl)propan-1-one	$C_{21}H_{24}O_{10}$	0	3	
53	24.04	419.1622	NCGC00380324-01	$C_{23}H_{29}ClO_5$	0.79	3	
54	24.04	209.0776	Divaricatinic acid	$C_{11}H_{14}O_4$	4.39	3	
55	25.22	217.0486	Chlorophene	$C_{13}H_{11}ClO$	5.49	3	
56	26.58	237.0207	Nitrofurantoin	$C_8H_6N_4O_5$	5.79	3	
57	27.53	245.0800	columbianetin	$C_{14}H_{14}O_4$	1.89	3	
58	31.18	201.0538	2-Naphthoxyacetic acid	$C_{12}H_{10}O_3$	1.89	3	
59	33.00	259.0956	Peucenin	$C_{15}H_{16}O_4$	1.99	3	
60	33.24	329.1374	1,7-bis(3,4-dihydroxyphenyl)heptan-3-one	$C_{19}H_{22}O_5$	2.9	3	
61	33.86	231.0643	Goniothalenol	$C_{13}H_{12}O_4$	1.98	3	
62	36.62	225.1477	Methyl Dihydrojasmonate	$C_{13}H_{22}O_3$	1.38	3	
63	37.63	301.1059	1-(2,4-dihydroxyphenyl)-2-(3,5-dimethoxyphenyl)propan-1-one	$C_{17}H_{18}O_5$	3.51	3	
64	38.44	329.1007	2',4-Dihydroxy-3,4',6'-Trimethoxychalcone	$C_{15}H_{12}O_3$	1.98	3	

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
65	38.98	243.1581	Tridecanedioic acid	$C_{13}H_{24}O_4$	2.1	3	
66	39.00	270.2055	N-Dodecanoyl-N-methylglycine	$C_{15}H_{29}NO_3$	1.98	3	
67	39.09	399.2341	bonactin	$C_{21}H_{36}O_7$	4.39	3	
68	39.09	331.2469	2'-Deoxyinosine-5'-monophosphate sodium salt	$C_{10}H_{13}N_4O_7P$	3.11	3	
69	39.27	361.1610	Thyrotropin releasing hormone	$C_{16}H_{22}N_6O_4$	0.98	3	
70	39.48	377.1567	NCGC00380298-01	$C_{19}H_{24}O_5$	3.29	3	
71	39.48	309.1692	C12-AE1S (tentative)	$C_{14}H_{30}O_5S$	3.81	3	
72	39.55	273.1112	Heteropeucenin, Methyl Ether	$C_{16}H_{18}O_4$	2.01	3	
73	39.76	249.1480	Gemfibrozil	$C_{15}H_{22}O_3$	1.61	3	
74	39.99	315.0492	Pannaric acid	$C_{16}H_{12}O_7$	0.18	3	
75	40.17	313.0334	Haematommone	$C_{16}H_{10}O_7$	0.4	3	
76	40.38	299.2210	Roccellic acid	$C_{17}H_{32}O_4$	0.97	3	
77	40.48	239.0654	2',4'-Dihydroxychalcone	$C_{15}H_{12}O_3$	5.6	3	
78	40.65	291.1581	Curvularin	$C_{16}H_{20}O_5$	5.37	3	
79	40.82	345.1299	Gibberellic acid	$C_{19}H_{22}O_6$	1.92	3	
80	40.99	243.1946	beta-Hydroxymyristic acid	$C_{14}H_{28}O_3$	1.31	3	
81	41.47	353.1981	(E)-3-(4-acetyloxy-2,3-dihydroxy-2,5,5,8a-tetramethyl-3,4,4a,6,7,8-hexahydro-1H-naphthalen-1-yl)prop-2-enoic acid	$C_{19}H_{30}O_6$	1.13	3	
82	41.51	421.1856	[(6Z,10Z)-6-(acetyloxymethyl)-10-(hydroperoxymethyl)-3-methylidene-2-oxo-3a,4,5,8,9,11a-hexahydrocyclodeca[b]furan-4-yl] 2-methylbutanoate	$C_{22}H_{30}O_8$	1.16	3	
83	41.65	277.1425	Di-n-butyl phthalate	$C_{16}H_{22}O_4$	2.01	3	
84	42.01	271.2259	16-Hydroxyhexadecanoic acid	$C_{16}H_{32}O_3$	1.44	3	
85	44.35	387.1540	Cetirizine	$C_{21}H_{25}ClN_2O_3$	5.89	3	
86	46.03	469.1085	Prunin	$C_{21}H_{22}O_{10}$	6.16	3	
87	46.03	401.1207	NCGC00381273-01!3-hydroxy-7-(4-hydroxyphenyl)-2-(2-hydroxypropan-2-yl)-2,3,6,7-tetrahydrofuro[3,2-g]chromen-5-one	$C_{20}H_{20}O_6$	3.29	3	
88	46.03	333.1333	Difucol Hexamethyl Ether	$C_{18}H_{22}O_6$	1.07	3	
Suggested from MS-FINDER							
1	2.49	257.0051	Ethoxzolamide	$C_9H_{10}N_2O_3S_2$	0.91	3	

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
2	2.86	218.0440	8-Methoxykynurenate	$C_{11}H_9NO_4$	1.88	3	
3	3.28	233.0649	1-Isopropyl citrate	$C_9H_{14}O_7$	1.78	3	
4	3.30	327.0314	7-amino-3-([(1-methyl-1H-1,2,3,4-tetrazol-5-yl)sulfanyl]methyl)-8-oxo-5-thia-1-azabicyclo[4.2.0]oct-2-ene-2-carboxylic acid	$C_{10}H_{12}N_6O_3S_2$	2.55	3	
5	3.60	205.0700	Diethyl tartrate	$C_8H_{14}O_6$	1.76	3	
6	3.72	125.0231	Trihydroxybenzene	$C_6H_6O_3$	1.32	3	(Yee et al., 2013)
7	6.04	299.0364	Cinnavalininate	$C_{14}H_8N_2O_6$	5.44	3	
8	6.22	181.0124	stipitatic acid	$C_8H_6O_5$	1.85	3	
9	6.95	299.0365	Tazobactam	$C_{10}H_{12}N_4O_5S$	9.10	3	
10	7.35	209.0073	benzene-1,3,5-tricarboxylic acid	$C_9H_6O_6$	1.86	3	
11	11.68	251.0154	UNPD186460	$C_{11}H_8O_7$	4.33	3	
12	12.09	139.0384	5-methoxybenzene-1,3-diol	$C_7H_8O_3$	1.67	3	
13	13.29	153.0176	2-Pyrocatechuic acid	$C_7H_6O_4$	1.73	3	
14	14.33	325.0545	Fertaric acid	$C_{14}H_{14}O_9$	2.01	3	
15	15.09	166.0573	Pyridoxal	$C_8H_9NO_3$	6.33	3	
16	15.96	265.0313	2-O-p-Coumaroyltartronic acid	$C_{12}H_{10}O_7$	4.08	3	
17	17.23	190.0127	5,6-Indolequinone-2-carboxylic acid	$C_9H_5NO_4$	1.88	3	
18	17.92	513.0198	1-[bis(2,4-dinitrophenyl)methyl]-2,4-dinitrobenzene	$C_{19}H_{10}N_6O_{12}$	8.59	3	
19	17.92	377.0469	UNPD24980	$C_{17}H_{14}O_{10}$	4.52	3	
20	17.92	445.0337	4-O-Demethylbarbamide	$C_{19}H_{21}Cl_3N_2O_2S$	2.04	3	
21	18.23	139.0746	2-Propyl-2,4-pentadienoic acid	$C_8H_{12}O_2$	1.85	3	
22	22.22	251.0906	Oxcarbazepine	$C_{15}H_{12}N_2O_2$	8.00	3	
23	23.54	178.0571	Hippuric acid	$C_9H_9NO_3$	6.13	3	
24	24.04	339.0352	Oxadiargyl	$C_{15}H_{14}Cl_2N_2O_3$	4.33	3	
25	24.04	407.0229	Olivanic acid	$C_{13}H_{16}N_2O_9S_2$	0.45	3	
26	24.04	441.1463	1-O-E-Cinnamoyl-(6-arabinosylglucose)	$C_{20}H_{26}O_{11}$	6.07	3	
27	27.14	279.0855	N-Despyridinyl rosiglitazone	$C_{13}H_{16}N_2O_3S$	4.61	3	
28	37.65	294.1689	Esmolol	$C_{16}H_{25}NO_4$	2.19	3	
29	38.15	242.1741	N-Undecanoylglycine	$C_{13}H_{25}NO_3$	2.07	3	
30	39.29	236.1034	UNPD224128	$C_{16}H_{15}NO$	4.69	3	
31	39.35	395.2415	Australin A	$C_{22}H_{36}O_6$	2.41	3	
32	39.48	445.1439	Bipinnatin K	$C_{23}H_{26}O_9$	6.51	3	
33	39.84	266.1014	Mortivinacin B	$C_{13}H_{17}NO_5$	1.99	3	

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
34	40.17	339.0853	Dihydro-O-methylsterigmatocystin	$C_{19}H_{16}O_6$	2.11	3	
35	40.75	403.1893	4'-ethyl 5'-methyl 2-hydroxy-2'-(2-methoxy-2-methylpropyl)spiro[indole-3,3'-pyrrolidine]-4',5'-dicarboxylate	$C_{21}H_{28}N_2O_6$	7.26	3	
36	41.13	250.1497	Furmecyclox	$C_{14}H_{21}NO_3$	4.83	3	
37	41.63	356.1848	(S)-Laudanosine	$C_{21}H_{27}NO_4$	1.93	3	
38	42.77	367.2623	(R)-3,4-Dihydro-2-(4,8,12-trimethyl-3,7,11-tridecatrienyl)-2H-1-benzopyran-6-ol	$C_{25}H_{36}O_2$	1.95	3	
39	44.00	396.2157	Drotaverine	$C_{24}H_{31}NO_4$	2.33	3	
40	44.61	441.2509	Prostaglandin G2 2-glyceryl Ester	$C_{23}H_{38}O_8$	1.51	3	
41	46.22	304.9119	2,3,4,5-Tetrachloro-4'-biphenylol	$C_{12}H_6Cl_4O$	1.90	3	

Table S9. List of tentative candidates from ESI(+) mode

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
Matched from MS-DIAL and confirmed with authentic standards							
1	17.9	199.0602	Syringic acid	$C_9H_{10}O_5$	0.2	1	(Moschos et al., 2024; Divisekara et al., 2023)
2	19.0	149.0598	Cinnamic acid	$C_9H_8O_2$	0.2	1	(Oros and Simoneit, 2001a)
3	20.6	183.0655	Syringaldehyde	$C_9H_{10}O_4$	0.32	1	(Oros et al., 2006; Moschos et al., 2024)
Matched from MS-DIAL							
1	4.6	193.0473	5,7-Dihydroxy-4-methylcoumarin	$C_{10}H_8O_4$	2.7	2	
2	5.0	113.0598	1,3 Cyclohexanedione	$C_6H_8O_2$	0.1	2	
3	5.3	217.1050	1,2,3,4-tetrahydro-6-methoxy-1-oxo-beta-carboline	$C_{12}H_{12}N_2O_2$	4.99	2	
4	5.5	155.0340	Protocatechuic acid	$C_7H_6O_4$	4	2	(Divisekara et al., 2023)
5	7.3	219.1131	(Z)-2-((Z)-1-(hydroxyimino)ethyl)-6,6-dimethylbicyclo[3.1.0]hexan-3-one oxime	$C_{10}H_{16}N_2O_2$	3.1	2	
6	8.8	218.0813	5,8-dimethoxyquinoline-2-	$C_{12}H_{11}NO_3$	1.3	2	

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
			carbaldehyde				
7	10.0	293.0636	Cardamonin	$C_{16}H_{14}O_4$	6.41	2	
8	10.6	127.0755	Melamine	$C_3H_6N_6$	2.8	2	
9	10.7	285.0374	4,5-dihydroxy-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid	$C_{15}H_8O_6$	2.6	2	
10	11.2	191.0681	2-hydroxy-4-methyl-3H-benzo[e][1,4]diazepin-5(4H)-one	$C_{10}H_{10}N_2O_2$	1.91	2	
11	11.2	261.1313	2-(4-oxoquinazolin-3(4H)-yl)ethyl isobutyrate	$C_{14}H_{16}N_2O_3$	1.29	2	
12	11.8	216.1021	Atrazine	$C_8H_{14}ClN_5$	1.1	2	(Lamnoi et al., 2024)
13	13.3	139.0391	4-hydroxy-benzoic acid	$C_7H_6O_3$	0.4	2	(Oros and Simoneit, 2001a)
14	14.3	206.1177	Dehydrosalsolidine	$C_{12}H_{15}NO_2$	0.16	2	
15	14.6	203.0680	7-hydroxy-2,3-dihydrocyclopenta[c]chromen-4(1H)-one	$C_{12}H_{10}O_3$	2.02	2	
16	14.6	221.0786	3,5-dimethyl-7-oxo-6-oxabicyclo[3.2.1]octane-2-carboxylic acid	$C_{10}H_{14}O_4$	1.4	2	
17	14.9	197.1286	NCGC00381364-01!3-propan-2-yl-2,3,6,7,8,8a-hexahydropyrrolo[1,2-a]pyrazine-1,4-dione	$C_{10}H_{16}N_2O_2$	0.59	2	
18	15.7	305.1576	Miltirone	$C_{19}H_{22}O_2$	7.6	2	
19	16.0	153.0547	Gaultherin	$C_{19}H_{26}O_{12}$	0.29	2	
20	16.3	229.0836	p-methoxycinnamic acid ethyl ester	$C_{12}H_{14}O_3$	3.6	2	
21	16.6	169.0497	vanillic acid	$C_8H_8O_4$	0.3	2	(Moschos et al., 2024; Divisekara et al., 2023)
22	17.0	197.0809	1-(2-hydroxy-4,6-dimethoxyphenyl)ethanone	$C_{10}H_{12}O_4$	0.9	2	
23	17.0	205.0836	Angelic anhydride	$C_{10}H_{14}O_3$	3.6	2	
24	17.6	257.0786	Pinocembrin	$C_{15}H_{12}O_4$	3.39	2	
25	17.6	163.0390	Umbelliferone	$C_9H_6O_3$	0.99	2	(Moschos et al., 2024; Fleming et al., 2020)
26	18.3	393.2103	Glabrol	$C_{25}H_{28}O_4$	0.31	2	
27	19.2	437.2367	Nonaethylene glycol	$C_{18}H_{38}O_{10}$	1	2	

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
28	19.2	217.0474	Isobergapten	$C_{12}H_8O_4$	2.61	2	(Divisekara et al., 2023)
29	19.3	215.0680	Benzophenone-1	$C_{13}H_{10}O_3$	2.29	2	
30	19.4	153.0546	Vanillin	$C_8H_8O_3$	5.4	2	(Oros et al., 2006; Moschos et al., 2024)
31	19.8	193.0497	1H-2-benzopyran-1-one, 6,8-dihydroxy-3-methyl-	$C_{10}H_8O_4$	0.6	2	
32	19.9	165.0911	Eugenol	$C_{10}H_{12}O_2$	0.09	2	
33	19.9	481.2630	Decaethylene glycol	$C_{20}H_{42}O_{11}$	1.1	2	
34	20.7	391.1945	Licoflavone B	$C_{25}H_{26}O_4$	4.49	2	
35	20.8	303.0479	2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-4H-chromen-4-one	$C_{15}H_{10}O_7$	2.08	2	
36	21.2	149.0599	3,4-Dihydrocoumarin	$C_9H_8O_2$	0.2	2	
37	21.2	231.0630	Ethyl caffeate	$C_{11}H_{12}O_4$	3	2	
38	21.8	197.0811	3',5'-Dimethoxy-4'-hydroxyacetophenone	$C_{10}H_{12}O_4$	0.26	2	
39	23.3	287.0531	Kaempferol	$C_{15}H_{10}O_6$	3.11	2	(Moschos et al., 2024; Fleming et al., 2020)
40	23.3	349.1263	Citrusin	$C_{16}H_{22}O_7$	3.69	2	
41	23.6	179.0706	Coniferaldehyde	$C_{10}H_{10}O_3$	0.30	2	(Moschos et al., 2024; Fleming et al., 2020)
42	23.7	209.0810	Sinapoyl aldehyde	$C_{11}H_{12}O_4$	0.15	2	(Moschos et al., 2024; Fleming et al., 2020)
43	24.0	233.0779	sinapyl alcohol	$C_{11}H_{14}O_4$	2.10	2	(Oros et al., 2006)
44	29.3	275.0895	Methyl 3,4,5-trimethoxycinnamate	$C_{13}H_{16}O_5$	0.49	2	
45	33.9	357.0951	Byakangelicin	$C_{17}H_{18}O_7$	0.09	2	
46	37.5	419.1687	2,6-dihydroxy-3-(2-methoxyphenyl)-5-(2-methyl-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indol-1-yl)pyrimidin-4(3H)-one	$C_{23}H_{22}N_4O_4$	1.31	2	
47	37.6	341.0793	Esculin	$C_{15}H_{16}O_9$	0.67	2	(Divisekara et al., 2023)
48	37.7	427.1736	NCGC00384974-01	$C_{22}H_{28}O_7$	0.61	2	
49	38.6	505.3363	Polyporenic acid C	$C_{31}H_{46}O_4$	6.31	2	

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
50	38.6	274.2746	Lauryldiethanolamine	$C_{16}H_{35}NO_2$	0.49	2	
51	38.6	339.1211	Licoflavone C	$C_{20}H_{18}O_5$	1.09	2	
52	38.8	250.1781	N-Desmethyltramadol	$C_{15}H_{23}NO_2$	2.11	2	
53	39.2	409.1631	Eudesmin	$C_{22}H_{26}O_6$	3.09	2	
54	39.5	333.1679	N-Methylnuciferine	$C_{20}H_{24}NO_2$	2.10	2	
55	39.8	453.1742	5-O-methylvisammioside	$C_{22}H_{28}O_{10}$	4.18	2	
56	39.9	539.2115	2-((6-((6,7-dimethoxy-3,4-dihydroisoquinolin-2(1H)-yl)methyl)-4-oxo-4H-pyran-3-yl)oxy)-N-(3,4-dimethoxyphenethyl)acetamide	$C_{29}H_{34}N_2O_8$	1.47	2	
57	40.1	337.1054	4'-hydroxy-2',4,6'-trimethoxychalcone	$C_{18}H_{18}O_5$	4.58	2	
58	40.2	299.1833	methyl 2-(8-methyl-3a,4,5,6-tetrahydro-1H-pyrazino[3,2,1-jk]carbazol-3(2H)-yl)acetate	$C_{18}H_{22}N_2O_2$	3.30	2	
59	41.0	279.1596	Dibutylphthalate	$C_{16}H_{22}O_4$	0.52	2	
60	41.2	345.2383	Darutigenol	$C_{20}H_{34}O_3$	1.68	2	
61	41.8	485.1139	Hispiduloside	$C_{22}H_{22}O_{11}$	3.90	2	
62	41.9	398.2334	Tuberostemonine	$C_{22}H_{33}NO_4$	3.39	2	
63	42.6	337.2356	NCGC00381425-01	$C_{18}H_{34}O_4$	0.39	2	
Suggested from MS-DIAL							
1	2.3	198.0074	Selenomethionine	$C_5H_{11}NO_2$	4.39	3	
2	2.6107	205.0711	Galactitol	$C_6H_{14}O_6$	3.12	2	
3	2.6	187.0603	L-Rhamnose	$C_6H_{12}O_5$	2.51	3	(Holmes, 2008)
4	2.9	178.0976	Pelletierine Hydrochloride	$C_8H_{16}ClNO$	1.7	3	
5	2.9	220.0606	methyl 2-(2-amino-4-hydroxy-6-methylpyrimidin-5-yl)acetate	$C_8H_{11}N_3O_3$	9.42	3	
6	3.0	189.0394	Plumbagin	$C_{11}H_8O_3$	5.94	3	
7	3.4	330.0503	Deoxycytidine monophosphate	$C_9H_{14}N_3O_7P$	4.3	3	
8	3.7	204.1021	1h-indole-3-butanoic acid	$C_{12}H_{13}NO_2$	2.09	3	
9	4.9	187.0003	D-(-)-3-phosphoglyceric acid	$C_3H_7O_7P$	0.31	3	
10	5.1	212.0232	6-Ethoxy-2-mercaptobenzothiazole	$C_9H_9NOS_2$	3.38	3	
11	5.6	197.0422	Shikimic Acid	$C_7H_{10}O_5$	0.17	3	(Jen et al., 2019)
12	6.4	137.0597	2-Hydroxyacetophenone	$C_8H_8O_2$	0.09	3	
13	6.4	232.0971	Propoxur	$C_{11}H_{15}NO_3$	3.12	3	
14	7.0	209.0786	3,5-Dimethoxycinnamic acid	$C_{11}H_{12}O_4$	2.24	3	(Moschos et al.,

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
							2024)
15	7.5	177.0524	4-Methylumbelliferone	$C_{10}H_8O_3$	2.2	3	(Moschos et al., 2024)
16	7.6	193.0472	6-Hydroxy-7-methoxycoumarin	$C_{10}H_8O^4$	2.8	3	(Arndt et al., 2020)
17	8.0	195.1130	6-Hydroxypseudooxynicotine	$C_{10}H_{14}N_2O_2$	0.68	3	
18	8.2	201.1097	Harmalol	$C_{12}H_{12}N_2O$	7.47	3	
19	8.5	337.0899	Vanilloside	$C_{14}H_{18}O_8$	0.51	3	
20	8.8	189.0524	3-phenyllactic acid	$C_9H_{10}O_3$	0.18	3	
21	9.2	172.0757	Metronidazole	$C_6H_9N_3O_3$	4	3	
22	9.7	303.0392	Morin	$C_{15}H_{10}O_7$	3.78	3	
23	9.8	215.0681	4-oxo-5-phenylpentanoic acid	$C_{11}H_{12}O_3$	0.24	3	
24	9.9	276.1233	MMV688469	$C_{16}H_{13}N_5$	0.89	3	
25	10.1	171.0418	2-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetic acid	$C_6H_6N_2O_4$	1.8	3	
26	10.3	187.0966	Gabapentin Related Compound E	$C_9H_{14}O_4$	0.1	3	
27	10.7	269.0636	columbianetin	$C_{14}H_{14}O_4$	6.41	3	
28	10.9	239.0529	6-hydroxyflavone	$C_{15}H_{10}O_3$	0.77	3	
29	10.9	351.0692	bergenin	$C_{14}H_{16}O_9$	4.7	3	
30	11.2	169.0860	2-(3-oxobutyl)cyclopentane-1,3-dione	$C_9H_{12}O_3$	4	3	
31	11.8	139.0390	Salicyclic acid	$C_7H_6O_3$	0.99	3	(Fleming et al., 2020)
32	13.0	202.0864	Carbaryl	$C_{12}H_{11}NO_2$	0.1	3	
33	13.9	221.0916	3-Acetylphenanthrene	$C_{16}H_{12}O$	4.5	3	
34	14.0	292.0795	Crufomate	$C_{12}H_{19}ClNO_3P$	6.90	3	
35	14.2	231.1200	Hydroxypestalotin	$C_{11}H_{18}O_5$	0	3	
36	14.4	227.0892	Genipin	$C_{11}H_{14}O_5$	0.80	3	
37	14.7	251.0529	Demethoxyyangonin	$C_{14}H_{12}O_3$	7.1	3	
38	15.0	288.1234	lycorine	$C_{16}H_{17}NO_4$	3.41	3	
39	15.1	293.0636	1-methyl-9H-pyrido[3,4-b]indol-7-ol hydrochloride dihydrate	$C_{12}H_{15}ClN_2O_3$	6.41	3	
40	15.2	167.0704	Paeonol	$C_9H_{10}O_3$	0.13	3	(Smith et al., 2020b)
41	15.4	274.1077	isoliquiritigenin	$C_{15}H_{12}O_4$	0.31	3	
42	15.6	257.0998	1,1'-biisoquinoline	$C_{18}H_{12}N_2$	0.22	3	
43	15.6	245.1362	4-oxododecanedioic acid	$C_{12}H_{20}O_5$	1.80	3	
44	16.4	223.1419	Mexacarbate	$C_{12}H_{18}N_2O_2$	2.09	3	
45	17.0	275.1467	ascr#7	$C_{13}H_{22}O_6$	2.20	3	
46	17.3	349.1837	hydroquinidine	$C_{20}H_{26}N_2O_2$	6.31	3	

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
47	17.4	208.1333	Phenylalanine betaine	$C_{12}H_{17}NO_2$	0.11	3	
48	17.9	333.0584	2'-Deoxyinosine-5'-monophosphate	$C_{10}H_{13}N_4O_7P$	1.56	3	
49	17.9	390.0503	Sanguinarium chloride	$C_{20}H_{14}ClNO_4$	0.30	3	
50	18.05	253.0685	Harmaline	$C_{13}H_{14}N_2O$	5.26	2	
51	18.4	185.0573	Isosafrole	$C_{10}H_{10}O_2$	0	3	
52	18.4	319.1732	Exemestane	$C_{20}H_{24}O_2$	6.29	3	
53	18.5	303.1418	NCGC00180705-02	$C_{19}H_{20}O_2$	5.83	3	
54	18.9	368.1499	Corynoline	$C_{21}H_{21}NO_5$	0.09	3	
55	19.0	201.0524	Camalexin	$C_{11}H_8N_2S$	4.30	3	
56	20.0	316.1548	cephalotaxine	$C_{18}H_{21}NO_4$	5.21	3	
57	20.2	213.1123	Dehydro-piliformic-acid	$C_{11}H_{16}O_4$	0.16	3	
58	20.6	155.0703	terrein	$C_8H_{10}O_3$	0.03	3	
59	20.7	387.1332	Harpagide	$C_{15}H24O_{10}$	8.21	3	
60	20.8	202.1227	1-Cinnamoylpyrrolidine	$C_{13}H_{15}NO$	2.70	3	
61	21.0	295.0493	butein	$C_{15}H_{12}O_5$	0.70	3	
62	21.0	211.0966	Methylxanthoxylin	$C_{11}H_{14}O_4$	1.00	3	
63	21.3	377.2152	Fluorometholone	$C_{22}H_{29}FO_4$	0.21	3	
64	21.7	381.1526	methyl 2-(4,5-dihydroxy-6-methoxy-3,6-dimethyl-2-oxocyclohex-3-en-1-yl)oxy-4-hydroxy-3,6-dimethylbenzoate	$C_{19}H_{24}O_8$	0.82	3	
65	21.8	351.1420	(2S,3R,4S,5S,6R)-2-[4-(3-hydroxybutyl)phenoxy]-6-(hydroxymethyl)oxane-3,4,5-triol	$C_{16}H_{24}O_7$	0.18	3	
66	23.1	263.0529	maclurin	$C_{13}H_{10}O_6$	7.11	3	
67	23.5	321.1313	[2-[(E)-3,4-dihydroxypent-1-enyl]-6-oxooxan-3-yl] (E)-2-methylbut-2-enoate	$C_{15}H_{22}O_6$	0.67	3	
68	23.6	351.2260	andrographolide	$C_{20}H_{30}O_5$	6.01	3	
69	24.1	171.1017	Azelaic acid	$C_9H_{16}O_4$	0.70	3	(Sengupta et al., 2020)
70	24.1	252.1084	Cordycepin	$C_{10}H_{13}N_5O_3$	1.60	3	
71	24.1	415.1648	Garcinone C	$C_{23}H_{26}O_7$	5.22	3	
72	24.5	331.2097	ascr#17	$C_{17}H_{30}O_6$	1.80	3	
73	24.5	465.1742	Lusitanicoside	$C_{21}H_{30}O_{10}$	1.09	3	
74	25.0	272.2012	Dextromethorphan	$C_{18}H_{25}NO$	0.19	3	
75	25.4	274.1805	Tilidine	$C_{17}H_{23}NO_2$	0.30	3	

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
76	25.4	256.0970	Amfenac	$C_{15}H_{13}NO_3$	0.49	3	
77	25.6	315.1784	Lotusine	$C_{19}H_{24}NO_3$	1.58	3	
78	26.0	238.1416	N,N-Diethyl-3,4-dimethoxybenzamide	$C_{13}H_{19}NO_3$	2.17	3	
79	27.5	385.1630	Eupalinolide K	$C_{20}H_{26}O_6$	2.99	3	
80	27.8	285.1678	Dihydroqinghaosu, Dihydroartemisinin.	$C_{15}H_{24}O_5$	1.86	3	
81	27.9	319.1156	5-(3-furan-2-yl-acryloyl)-2,2-dimethyl-4,6-dioxo-cyclohexanecarboxylic acid methyl ester	$C_{17}H_{18}O_6$	4.40	3	
82	28.6	289.1051	(E)-5-(4-methoxy-5-methyl-6-oxopyran-2-yl)-3-methylhex-4-enoic acid	$C_{14}H_{18}O_5$	0.46	3	
83	29.4	333.1314	Duartin	$C_{18}H_{20}O_6$	0.79	3	
84	29.8	383.1472	Eupalinilide B	$C_{20}H_{24}O_6$	7.18	3	
85	30.3	299.1469	T-2 tetraol	$C_{15}H_{22}O_6$	1.50	3	
86	30.8	199.0366	Hymecromone	$C_{10}H_8O_3$	0.05	3	
87	31.3	373.2202	7-hydroxy-1,4a-dimethyl-9-oxo-7-propan-2-yl-2,3,4,4b,5,6,10,10a-octahydrophenanthrene-1-carboxylic acid	$C_{20}H_{30}O_4$	1.26	3	
88	31.7	379.2211	Neotame	$C_{20}H_{30}N_2O_5$	1.59	3	
89	34.1	391.1373	Resveratroloside	$C_{20}H_{22}O_8$	2.71	3	
90	36.4	282.1018	Erythrinine	$C_{18}H_{19}NO_4$	1.58	3	
91	37.6	325.1053	methyl 2-ethyl-4-[(3R,4R,5S)-5-hydroxy-4,5-dimethyl-2-oxooxolan-3-yl]-2-methyl-3-oxobutanoate	$C_{14}H_{22}O_6$	0.49	3	
92	37.6	347.2312	Calycanthine	$C_{22}H_{26}N_4$	9.22	3	
93	38.1	373.1059	Difucol Hexamethyl Ether	$C_{18}H_{22}O_6$	1.10	3	
94	38.1	357.1317	Pioglitazone	$C_{19}H_{20}N_2O_3S$	5.00	3	
95	38.1	244.1912	2-(7-hydroxy-6-methyloctyl)-2H-furan-5-one	$C_{13}H_{22}O_3$	0.20	3	
96	38.4	327.1596	Licarin A	$C_{20}H_{22}O_4$	0.39	3	
97	38.5	417.2469	N-(3-(dimethylamino)propyl)-2-((3,4,8,8-tetramethyl-2-oxo-2,8,9,10-tetrahydropyrano[2,3-f]chromen-5-	$C_{23}H_{32}N_2O_5$	6.90	3	

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
			yl)oxy)acetamide				
98	38.5	195.1018	Butylparaben	$C_{11}H_{14}O_3$	0.22	3	
99	38.7	271.1884	(E)-2-decylpent-2-enedioic acid	$C_{15}H_{26}O_4$	1.98	3	
100	38.8	431.2625	8-[2-(3,4-dihydroxy-2,5-dimethoxyoxolan-3-yl)ethyl]-4a,7,8-trimethylspiro[2,3,5,6,7,8a-hexahydro-1H-naphthalene-4,2'-oxirane]-2,3-diol	$C_{22}H_{38}O_8$	1.41	3	
101	38.8	228.1961	Andrachcinidine	$C_{13}H_{25}NO_2$	0.11	3	
102	38.8	288.2902	Sphinganine (C17 base)	$C_{17}H_{37}NO_2$	0.06	3	
103	39.1	302.3058	2,2'-(Tetradecylimino)diethanol	$C_{18}H_{39}NO_2$	0.40	3	
104	39.3	561.3626	Deferroxamine	$C_{25}H_{48}N_6O_8$	1.59	3	
105	39.3	393.3096	Deoxycholic acid	$C_{24}H_{40}O_4$	9.61	3	
106	39.4	304.3005	Tsitsikammamine A	$C_{18}H_{13}N_3O_2$	0.52	3	
107	39.5	311.1860	(5S)-10,13-dimethyldodecahydro-1H-cyclopenta[a]phenanthrene-3,17(2H,4H)-dione	$C_{19}H_{28}O_2$	4.00	3	
108	39.5	390.1600	5-Chloro-3-[(1E,3E)-3,5-dimethyl-1,3-heptadien-1-yl]-7-methyl-6,8-dioxo-2,6,7,8-tetrahydro-7-isoquinolinyl acetate	$C_{21}H_{24}ClNO_4$	2.99	3	
109	39.5	298.2747	Palmitoleoyl Ethanolamide	$C_{18}H_{35}NO_2$	0.58	3	
110	39.5	338.2672	Melophlin D/H/I/J	$C_{20}H_{35}NO_3$	0.21	3	
111	39.7	402.3586	Sorbitane Monopalmitate - Polysorbate 40 in-source fragment	$C_{22}H_{42}O_6$	1.38	3	
112	39.8	257.1154	5-butyl-3-methyl-7H-furo[3,2-g]chromen-7-one	$C_{16}H_{16}O_3$	4.61	3	
113	39.9	417.3462	tigogenin	$C_{27}H_{44}O_3$	6.19	3	
114	40.5	318.2409	Drofenine	$C_{20}H_{31}NO_2$	2.50	3	
115	40.6	249.1852	Matrine	$C_{15}H_{24}N_2O$	4.80	3	
116	40.6	235.1696	Isocurcumenol	$C_{15}H_{22}O_2$	0.40	3	
117	40.7	457.2784	NCGC00380119-01	$C_{12}H_{20}O_4$	2.60	3	
118	40.7	577.3939	NCGC00381059-01	$C_{30}H_{56}O_{10}$	0.06	3	
119	40.8	445.3146	Menaquinone-4	$C_{31}H_{40}O_2$	4.52	3	
120	40.8	357.2618	9,13-Epoxy-3,15,16,18-labdanetetrol	$C_{20}H_{36}O_5$	1.74	3	
121	40.8	401.2882	Murolic acid	$C_{21}H_{36}O_5$	1.80	3	
122	40.8	313.2356	(9Z,12E)-15,16-dihydroxyoctadeca-	$C_{18}H_{32}O_4$	0.10	3	

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
			9,12-dienoic acid				
123	40.9	411.0948	6,7-dimethoxy-8-oxo-8H-benzo[c]indolo[3,2,1-ij][1,5]naphthyridin-12-yl acetate	$C_{22}H_{16}N_2O_5$	5.22	3	
124	41.0	301.1416	NCGC00380262-01	$C_{18}H_{20}O_4$	1.41	3	
125	41.1	343.1490	8,9-epoxy-3,10-diisobutyryloxythymol	$C_{18}H_{24}O_5$	1.00	3	
126	41.2	413.2261	Tussilagone	$C_{23}H_{34}O_5$	3.91	3	
127	41.2	481.2137	anthothecol	$C_{28}H_{32}O_7$	6.29	3	
128	41.4	425.2155	HT-2 Toxin	$C_{22}H_{32}O_8$	0.48	3	
129	41.7	561.3990	Desferrioxamine B	$C_{25}H_{48}N_6O_8$	1.03	3	
130	41.7	517.3725	NCGC00380806-01	$C_{29}H_{50}N_4O_5$	2.50	3	
131	42.4	575.4145	h-14-19-norandrosterone	$C_{18}H_{28}O_2$	7.51	3	
132	42.4	531.3881	Arginine conjugated chenodeoxycholic acid	$C_{30}H_{52}N_4O_5$	1.89	3	
133	42.4	443.3353	"Cycloart-23-ene-3_, 25 diol"	$C_{30}H_{50}O_2$	1.29	3	
134	42.4	487.3618	(4R,6aS,8aS)-5',6a,8a,9-tetramethyldocosahydrospiro[naphtho[2',1':4,5]indeno[2,1-b]furan-10,2'-pyran]-4-yl butyrate	$C_{31}H_{50}O_4$	8.21	3	
135	42.7	425.2883	NCGC00385003-01	$C_{24}H_{42}O_7$	0.70	3	
136	44.2	284.2951	C18(Plasm)-18:1 PC	$C_{44}H_{86}NO_7P$	1.09	3	
137	45.2	425.3609	DINCH	$C_{26}H_{48}O_4$	1.59	3	
138	46.0	311.1269	Ovalitenin B	$C_{19}H_{18}O_4$	0.89	3	
Suggested compounds from MS-FINDER							
1	2.2	111.0917	2-ethyl-4-methylimidazole	$C_6H_{10}N_2$	0.02	3	
2	2.9	125.1075	2-butyl-1H-imidazole	$C_7H_{12}N_2$	0.18	3	
3	3.4	224.0446	Brassicanal C	$C_{10}H_9NO_3S$	7.00	3	
4	4.9	165.0183	4-hydroxy-2-benzofuran-1,3-dione	$C_8H_4O_4$	0.06	3	
5	8.3	223.0370	S-(2,5-Dimethyl-3-furanyl) 2-furancarbothioate	$C_{11}H_{10}O_3S$	5.34	3	
6	8.8	190.0554	Kynurenic acid	$C_{10}H_7NO_3$	5.53	3	
7	9.8	210.1490	bamethan	$C_{12}H_{19}NO_2$	0.14	3	
8	11.2	192.0713	5-Hydroxyindoleacetic acid	$C_{10}H_9NO_3$	5.78	3	
9	14.9	292.0796	Hawkinsin	$C_{11}H_{17}NO_6S$	5.33	3	
10	19.5	363.1995	Makomotine C	$C_{17}H_{30}O_8$	1.84	3	
11	19.5	249.0374	UNPD104949	$C_{12}H_8O_6$	1.96	3	

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
12	20.7	254.0704	Pyranonigrin C	$C_{11}H_{11}NO_6$	4.49	3	
13	21.2	475.3265	UNPD83280	$C_{26}H_{42}N_4O_4$	1.38	3	
14	21.8	198.0843	L-Dopa	$C_9H_{11}NO_4$	8.22	3	
15	23.6	180.0737	Hippuric acid	$C_9H_9NO_3$	8.18	3	
16	24.1	268.0860	ICM 0201;Osteoclast differentiation inhibitor F-1490;(3S,10aR)-3,4a-Dihydroxy-2,3,4,4a-tetrahydro-2H-pyrano[3,2-b]benzo[e] morpholine-9-carboxylic acid	$C_{12}H_{13}NO_6$	4.44	3	
17	26.4	302.2119	Phormidinine B	$C_{19}H_{27}NO_2$	0.44	3	
18	27.9	320.1192	Salbutamol 4-O-sulfate	$C_{13}H_{21}NO_6S$	2.97	3	
19	28.6	290.1084	3,5-Dimethyl-1-(3-Nitrophenyl)-1h-Pyrazole-4-Carboxylic Acid Ethyl Ester	$C_{14}H_{15}N_3O_4$	5.13	3	
20	29.1	519.2098	(-)-Gossypol	$C_{30}H_{30}O_8$	8.46	3	
21	29.1	389.2517	5,6-Dihydroxyprostaglandin F1a	$C_{20}H_{36}O_7$	1.68	3	
22	32.2	292.2275	Penbutolol	$C_{18}H_{29}NO_2$	0.39	3	
23	32.8	252.1576	N-desmethylmirtazapine	$C_{16}H_{17}N_3$	8.08	3	
24	37.7	447.2940	UNPD95242	$C_{19}H_{38}N_6O_6$	1.44	3	
25	37.8	491.3204	UNPD154329	$C_{25}H_{46}O_9$	1.06	3	
26	38.0	445.2785	Cyclipostin P2	$C_{23}H_{41}O_6P$	7.15	3	
27	38.5	384.3958	UNPD87915	$C_{23}H_{49}N_3O$	0.96	3	
28	38.7	290.2695	C16 phytosphingosine	$C_{16}H_{35}NO_3$	0.53	3	
29	38.8	503.3202	UNPD180468	$C_{26}H_{46}O_9$	1.26	3	
30	39.1	563.3784	PA(14:1(9Z)/12:0)	$C_{29}H_{55}O_8P$	7.67	3	
31	39.3	316.3216	2-Propanol, 1,1'-(tridecylimino)bis-	$C_{19}H_{41}NO_2$	0.59	3	
32	39.3	489.3048	2-([3,6-dihydroxy-4-(5-hydroxy-3-methylpent-3-en-1-yl)-8-(hydroxymethyl)-3,4a,8-trimethyl-decahydronaphthalen-1-yl]oxy)oxane-3,4,5-triol	$C_{25}H_{44}O_9$	1.00	3	
33	39.9	507.3782	Melyne C	$C_{34}H_{50}O_3$	5.07	3	
34	40.0	440.4107	Pentacosanoylglycine	$C_{27}H_{53}NO_3$	0.88	3	

ID	RT (min)	m/z (Da)	Name	Formula	Mass diff. (mDa)	Confidence level	Ref.
35	40.0	326.3789	N-undecylundecan-1-amine	$C_{22}H_{47}N$	0.77	3	
36	40.1	399.3589	Episterol	$C_{28}H_{46}O$	3.24	3	
37	40.1	421.3409	Irisoquin C	$C_{26}H_{44}O_4$	9.66	3	
38	40.3	359.2410	UNPD229266	$C_{19}H_{34}O_6$	1.81	3	
39	40.6	354.4101	N,N-dioctyloctan-1-amine	$C_{24}H_{51}N$	0.67	3	
40	41.9	420.2155	Ancistrotanzanine B	$C_{26}H_{29}NO_4$	1.43	3	
41	42.0	341.2669	C00050790	$C_{20}H_{36}O_4$	1.74	3	
42	42.4	256.2637	Palmitic amide	$C_{16}H_{33}NO$	0.21	3	(Simoneit et al., 2003)
43	42.4	399.3090	(2-acetyloxy-3-hydroxypropyl) octadec-9-enoate	$C_{23}H_{42}O_5$	1.50	3	
44	43.3	560.1338	dTDP-alpha-D-desosamine	$C_{18}H_{31}N_3O_{13}P_2$	6.69	3	
45	43.6	369.2983	Docosanedioic acid	$C_{22}H_{40}O_4$	1.63	3	

Table S10. The contributions for the specific groups of compounds to BrC_{aq} derived from the LC-MS measurements using surrogate standards

Analysis methods	Carbohydrate	Hydrocarbon	Lipid	Peptide	Lignin	Tannin	Unclassified	Total
LC-MS (g/kg)	0.019	0.052	0.123	0.044	0.498	0.175	0.022	0.93 ± 0.47
TOC (g/kg)	-	-	-	-	-	-	-	0.95 ± 0.20
Gravimetry (g/kg)	-	-	-	-	-	-	-	1.00 ± 0.15

Table S11. The fifty highest concentration WSOCs detected with LC-MS

ID	Name	Concentration		Ref.	ID	Name	Concentration		Ref
		mg/kg	% (in total)				mg/kg	% (in total)	
1	1,2,4-	51.2	5.50	(Yee et al.,	26	2-Hydroxy-4-methyl-3H-	8.8	0.95	

ID	Name	Concentration		Ref.	ID	Name	Concentration		Ref
		mg/kg	% (in total)				mg/kg	% (in total)	
	Trihydroxybenzene			2013)		benzo[e][1,4]diazepin-5(4H)-one			
2	Methylxanthoxylin	28.1	3.02		27	Azelaic acid	8.4	0.90	(Sengupta et al., 2020)
3	Sorbic acid	27.2	2.92		28	Brassicanal C	8.2	0.88	
4	Columbianetin	23.0	2.46		29	Carbaryl	7.6	0.82	
5	4-hydroxybenzoic acid	21.4	2.30	(Divisekara, 2023)	30	4-Hydroxy-3-methoxycinnamaldehyde	7.2	0.77	
6	Coniferaldehyde	16.5	1.77	(Moschos et al., 2024; Fleming et al., 2020)	31	4-Ethoxy-4-oxobut-2-enoic acid	7.1	0.77	
7	p-Coumaric acid	15.1	1.62	(Oros et al., 2006)	32	1,2,3,4-tetrahydro-6-methoxy-1-oxo-beta-carboline	7.1	0.76	
8	Diethyl tartrate	14.0	1.50		33	Compound NP-008382	7.0	0.75	
9	Thymine	14.0	1.50		34	NCGC00180705-02	6.9	0.74	
10	Miltirone	13.0	1.39		35	Sulfadimethoxine	6.9	0.74	
11	2,3-Dihydroxybenzoic acid	12.7	1.37	(Graham et al., 2002b)	36	5-methoxybenzene-1,3-diol	6.8	0.73	
12	3',5'-Dimethoxy-4'-hydroxyacetophenone (acetosyringone)	12.6	1.35	(Hartner et al., 2024a)	37	Tryptophan	6.7	0.72	(Bianco et al., 2016)
13	2-(4-oxoquinazolin-3(4H)-yl)ethyl isobutyrate	12.5	1.34		38	[2-[(E)-3,4-dihydroxypent-1-enyl]-6-oxooxan-3-yl] (E)-2-methylbut-2-enoate (Compound NP-009265)	6.5	0.70	
14	Syringaldehyde	12.3	1.32	(Oros et al., 2006; Moschos et al., 2024)	39	5,7-Dihydroxy-4-methylcoumarin	6.4	0.69	
15	1-Isopropyl citrate	12.1	1.30		40	2'-Deoxyinosine-5'-monophosphate	6.3	0.68	
16	Succinic acid	11.3	1.21		41	2-ethyl-4-methylimidazole	6.3	0.67	
17	methyl 2-(2-amino-4-hydroxy-6-	11.2	1.20		42	3-Phenyllactic acid	6.2	0.67	

ID	Name	Concentration		Ref.	ID	Name	Concentration		Ref
		mg/kg	% (in total)				mg/kg	% (in total)	
	methylpyrimidin-5-yl)acetate								
18	isoliquiritigenin	10.7	1.14		43	3,5-Dimethoxycinnamic acid	6.2	0.66	(Moschos et al., 2024)
19	Ethionamide	10.5	1.12		44	Dihydro-O-methylsterigmatocystin	6.1	0.65	
20	(R)-(-)-mandelic acid	10.4	1.11	(Moschos et al., 2024; Oros and Simoneit, 2001b)	45	ICM 0201	6.0	0.65	
21	Norharman	10.3	1.10		46	Butein	6.0	0.65	
22	4-(3-Hydroxybutyl)phenyl beta-D-glucopyranoside	9.5	1.01		47	Orsellinic acid	6.0	0.65	
23	Umbelliferone	9.3	1.00	(Moschos et al., 2024; Fleming et al., 2020)	48	UNPD95242	6.0	0.64	
24	hydroquinidine	9.2	0.98		49	Glutaric acid	5.9	0.63	(Sengupta et al., 2020)
25	Prostaglandin G2 2-glyceryl Ester	8.8	0.95		50	Sinapoyl aldehyde	5.9	0.63	(Chan et al., 2020)

85

Table S12. Fifty WSOCs with the highest H scores tentatively identified with NTA

ID	Name	Emission (mg/kg)	*H* (M/atm) $\times 10^{-6}$	H_{score} $\times 10^{-6}$	ID	Name	Emission (mg/kg)	*H* (M/atm) $\times 10^{-6}$	H_{score} $\times 10^{-6}$
1	2'-Deoxyinosine-5'-monophosphate	6.3	95.6	605.8	26	1-O-E-Cinnamoyl-(6-arabinosylglucose)	0.8	67.4	51.5
2	Stipitatic acid	2.5	188.0	474.6	27	Shikimic Acid	4.1	10.6	43.3
3	Sulfadimethoxine	6.9	59.9	410.7	28	Drotaverine	2.8	14.9	41.2
4	benzene-1,3,5-tricarboxylic acid	5.6	71.6	404.5	29	Decaethylene glycol	2.6	14.9	38.6
5	Garcinone C	5.5	72.1	395.7	30	Isoliquiritigenin	10.7	3.6	38.3
6	Nonaethylene glycol	3.4	108.7	364.2	31	Tryptophan	6.7	5.3	35.8

ID	Name	Emission (mg/kg)	*H* (M/atm) $\times 10^{-6}$	H_{score} $\times 10^{-6}$	ID	Name	Emission (mg/kg)	*H* (M/atm) $\times 10^{-6}$	H_{score} $\times 10^{-6}$
7	2'-Deoxyinosine-5'-monophosphate sodium salt	3.4	95.6	324.3	32	columbianetin	23.0	1.5	34.7
8	Pyranonigrin C	2.2	127.4	276.8	33	3-Phosphoglyceric acid	2.8	12.2	34.0
9	Gaultherin	4.6	59.5	271.9	34	Methyl vanillate	4.5	6.9	31.3
10	Thyrotropin releasing hormone	3.9	59.6	231.3	35	Luteolin 7-O-glucuronide	0.4	64.7	29.0
11	Mortivinacin B	3.6	53.4	190.2	36	Diethyl tartrate	14.0	1.5	20.4
12	Australin A	2.4	75.3	183.7	37	Methylxanthoxylin	28.1	0.6	17.9
13	Morin	3.0	56.0	170.0	38	Umbelliferone	9.3	1.9	17.4
14	[2-[(E)-3,4-dihydroxypent-1-enyl]-6-oxooxan-3-yl] (E)-2-methylbut-2-enoate	6.5	24.0	155.9	39	Quinic acid	4.6	3.7	17.1
15	4,5-dihydroxy-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid	2.0	59.0	118.6	40	Syringaldehyde	12.3	1.4	16.8
16	1-Isopropyl citrate	12.1	9.6	116.3	41	3',5'-Dimethoxy-4'-hydroxyacetophenone	12.6	1.3	16.5
17	Licoflavone B	2.6	45.0	115.4	42	NCGC00380324-01	1.0	15.4	15.8
18	Glabrol	5.2	19.5	101.9	43	Orsellinic acid	6.0	2.6	15.4
19	Makomotine C	1.5	59.2	90.1	44	Olivanic acid	1.2	11.6	14.0
20	5,7-Dihydroxy-4-methylcoumarin	6.4	11.9	76.1	45	Cetirizine	2.9	4.6	13.4
21	Inosine-5'-monophosphate	0.8	94.0	74.2	46	Azelaic acid	8.4	1.6	13.3
22	Coatline A	1.1	52.7	58.0	47	ascr#7	4.0	3.1	12.3
23	Vanilloside	1.1	49.4	55.4	48	3,4-Hihydroxyphenylacetate	1.9	5.4	10.5
24	Glutaric acid	5.9	8.8	52.1	49	bamethan	2.6	3.6	9.4
25	Harpagide	1.0	54.6	52.0	50	NCGC00380119-01	2.5	3.5	8.7

90

Table S13. Fifty WSOCs with the highest LD_{50} scores tentatively identified with NTA

ID	Name	Emission (mg/kg)	LD_{50} (mg/kg)	LD_{50} Score $\times 10^2$	ID	Name	Emission (mg/kg)	LD_{50} (mg/kg)	LD_{50} Score $\times 10^2$
1	Propoxur	4.3	40.81	10.49	26	2-(4-oxoquinazolin-3(4H)-yl)ethyl isobutyrate	12.5	1808.18	0.69
2	1h-indole-3-butanoic acid	5.7	56.16	10.11	27	alternariol monomethyl ether	1.8	261.85	0.68
3	Prostaglandin G2 2-glyceryl Ester	8.8	137.53	6.42	28	Methylxanthoxylin	28.1	4153.86	0.68
4	1,2,4-Trihydroxybenzene	51.2	823.47	6.22	29	Norharman	10.3	1636.39	0.63
5	Coatline A	1.1	25.95	4.25	30	Drotaverine	2.8	446.60	0.62
6	Harmalol hydrochloride	5.4	153.14	3.54	31	Everninic Acid	3.6	590.30	0.62
7	1-[bis(2,4-dinitrophenyl)methyl]-2,4-dinitrobenzene	0.6	17.70	3.40	32	5-methoxybenzene-1,3-diol	6.8	1110.72	0.62
8	Carbaryl	7.6	231.07	3.30	33	Licoflavone B	2.6	421.64	0.61
9	hydroquinidine	9.2	334.11	2.75	34	Harmalol	0.9	153.14	0.58
10	Dehydrosalsolidine	2.3	85.54	2.71	35	columbianetin	23.0	4343.52	0.53
11	Dextromethorphan	4.8	187.81	2.55	36	3',5'-Dimethoxy-4'-hydroxyacetophenone	12.6	2391.23	0.53
12	Dihydro-O-methylsterigmatocystin	6.1	365.34	1.67	37	Isobergapten	2.7	540.14	0.50
13	Catechol	3.8	258.17	1.46	38	bamethan	2.6	528.65	0.49
14	Salicyclic acid	2.3	169.95	1.34	39	Coniferaldehyde	16.5	3852.02	0.43
15	1,2,3,4-tetrahydro-6-methoxy-1-oxo-beta-carboline	7.1	627.50	1.13	40	4-Hydroxybenzoic acid	21.4	5015.61	0.43
16	Miltirone	13.0	1210.97	1.07	41	Coumaric acid	15.1	3825.40	0.39
17	Orsellinic acid	6.0	582.39	1.04	42	Eupalinolide K	1.3	345.44	0.39
18	Glabrol	5.2	516.41	1.01	43	2-Mercaptobenzimidazole	1.1	299.75	0.38
19	Syringaldehyde	12.3	1388.33	0.89	44	Orcinol	3.2	839.43	0.38
20	Sorbic acid	27.2	3234.42	0.84	45	Succinic acid	11.3	3036.31	0.37
21	2-ethyl-4-methylimidazole	6.3	769.79	0.81	46	(Laudanosine	3.2	884.51	0.36
22	2,3-Dihydroxybenzoic acid	12.7	1577.33	0.81	47	5-Hydroxyindoleacetic acid	1.1	318.04	0.35

ID	Name	Emission (mg/kg)	LD_{50} (mg/kg)	LD_{50} Score $\times 10^2$	ID	Name	Emission (mg/kg)	LD_{50} (mg/kg)	LD_{50} Score $\times 10^2$
23	Ethionamide	10.5	1320.84	0.79	48	4-Ethoxy-4-oxobut-2-enoic acid	7.1	2077.15	0.34
24	5,8-dimethoxyquinoline-2-carbaldehyde	1.8	257.18	0.71	49	Isoliquiritigenin	10.7	3141.33	0.34
25	Riluzole	0.3	44.63	0.70	50	Butein	6.0	1839.39	0.33

Table S14. Correlations between LD_{50} and chemical properties of detected compounds from fine BrC, upper values are Pearson correlation coefficients, and p-values are given in brackets below. For p-values <0.05, the correlation is statistically significant; these results are highlighted in green

	LD_{50}	H (M/atm)	M.W. (Da)	nO	nN	nC	nS	DBE	Osc	O/C	H/C	N/O	Group
LD_{50}													
H (M/atm)	-0.24 (p=0.00)												
m/z	0.200 (p=0.00)	0.340 (p=0.00)											
nO	0.394 (p=0.00)	0.561 (p=0.00)	0.567 (p=0.00)										
nN	-0.07 (p=0.211)	0.073 (p=0.149)	0.131 (p=0.010)	-0.145 (p=0.004)									
nC	0.047 (p=0.406)	0.170 (p=0.001)	0.868 (p=0.00)	0.35 (p=0.00)	-0.001 (p=0.981)								
nS	-0.014 (p=0.799)	-0.033 (p=0.514)	0.04 (p=0.428)	-0.065 (p=0.204)	0.197 (p=0.00)	-0.062 (p=0.222)							
DBE	0.349 (p=0.00)	0.166 (p=0.001)	0.315 (p=0.00)	0.16 (p=0.002)	0.145 (p=0.004)	0.335 (p=0.00)	-0.001 (p=0.987)						
Osc	-0.033 (p=0.554)	0.215 (p=0.00)	-0.303 (p=0.00)	0.161 (p=0.002)	0.395 (p=0.00)	-0.517 (p=0.00)	0.108 (p=0.033)	0.308 (p=0.00)					
O/C	0.278 (p=0.00)	0.310 (p=0.00)	-0.11 (p=0.030)	0.583 (p=0.00)	-0.164 (p=0.001)	-0.353 (p=0.00)	-0.064 (p=0.205)	-0.184 (p=0.00)	0.505 (p=0.00)				
H/C	0.338 (p=0.000)	-0.016 (p=0.760)	0.3 (p=0.00)	0.102 (p=0.045)	0.046 (p=0.366)	0.279 (p=0.00)	-0.037 (p=0.463)	-0.693 (p=0.00)	-0.533 (p=0.00)	0.045 (p=0.378)			
N/O	-0.128 (p=0.025)	-0.059 (p=0.255)	0.005 (p=0.924)	-0.318 (p=0.00)	0.751 (p=0.00)	-0.034 (p=0.517)	0.182 (p=0.00)	0.078 (p=0.135)	0.148 (p=0.004)	-0.26 (p=0.00)	0.019 (p=0.721)		
Group	0.222 (p=0.000)	-0.091 (p=0.090)	0.126 (p=0.018)	-0.057 (p=0.283)	0.01 (p=0.855)	0.159 (p=0.003)	-0.001 (p=0.992)	-0.356 (p=0.00)	-0.335 (p=0.00)	-0.095 (p=0.075)	0.466 (p=0.00)	0.065 (p=0.231)	

Table S15. Data used to estimate the atmospheric lifetimes (see Eq. 8 in the main text) due to reaction with the OH for the potential $_{aq}$SOA precursors (Table S12 and Fig. 10)

Name	$k_{OH_{gas}}$ (ml^3 $molecules^{-1}$ s^{-1}) × 10^{11}	$k_{OH_{aq}}$ ($M^{-1}s^{-1}$) × 10^{-9}	H^{cc} × 10^{-9}
Stipitatic acid	3.58	1.9	4.6
Trimesic acid	0.87	4.3	1.7
Methyl vanillate	0.69	4.6	0.2
Methylxanthoxylin	6.00	7.6	0.0
Gaultherin	0.22	8.8	1.5
isoliquiritigenin	1.22	6.2	0.1
3-Hexenyl tiglate	6.66	3.5	0.6
Vanilloside	2.52	7.5	1.2
Orsellinic acid	2.52	5.0	0.1
Quinic acid	0.80	0.6	0.1
Morin	3.23	1.9	1.4
Licoflavone B	3.23	5.1	1.1
Glabrol	103.3	8.6	0.5
Luteolin 7-O-glucuronide	2.16	1.6	1.6
Rhein	1.59	8.9	1.4
Australin A	1.12	5.4	1.8
5,7-Dihydroxy-4-methylcoumarin	5.12	3.5	0.3
Umbelliferone	9.06	10.3	0.0
Pyranonigrin C	9.06	4.1	3.1
Mortivinacin B	2.04	6.8	1.3
1-Isopropyl citrate	0.18	0.3	0.2
Makomotine C	1.52	9.5	1.4
Tryptophan	5.49	13.7	0.1

References

GNPS tool, last

Amarandei, C., Olariu, R. I., and Arsene, C.: Offline analysis of secondary formation markers in ambient organic aerosols by liquid chromatography coupled with time-of-flight mass spectrometry, Journal of Chromatography A, 1702, 464092, https://doi.org/10.1016/j.chroma.2023.464092, 2023.

Arndt, D., Wachsmuth, C., Buchholz, C., and Bentley, M.: A complex matrix characterization approach, applied to cigarette smoke, that integrates multiple analytical methods and compound identification strategies for

non-targeted liquid chromatography with high-resolution mass spectrometry, Rapid Commun Mass Spectrom, 34, e8571, 2020.
Bhatia, M. P., Das, S. B., Longnecker, K., Charette, M. A., and Kujawinski, E. B.: Molecular characterization of dissolved organic matter associated with the Greenland ice sheet, Geoch. Cosm. Act., 74, 3768-3784, 2010.
Bianco, A., Passananti, M., Deguillaume, L., Mailhot, G., and Brigante, M.: Tryptophan and tryptophan-like substances in cloud water: Occurrence and photochemical fate, Atmospheric Environment, 137, 53-61, https://doi.org/10.1016/j.atmosenv.2016.04.034, 2016.
Burgay, F., Salionov, D., Huber, C. J., Singer, T., Eichler, A., Ungeheuer, F., Vogel, A., Schwikowski, M., and Bjelic, S.: Hybrid Targeted/Untargeted Screening Method for the Determination of Wildfire and Water-Soluble Organic Tracers in Ice Cores and Snow, Anal Chem, 95, 11456-11466, 10.1021/acs.analchem.3c01852, 2023.
Cao, F., Zhang, S.-C., Kawamura, K., Liu, X., Yang, C., Xu, Z., Fan, M., Zhang, W., Bao, M., Chang, Y., Song, W., Liu, S., Lee, X., Li, J., Zhang, G., and Zhang, Y.-L.: Chemical characteristics of dicarboxylic acids and related organic compounds in PM2.5 during biomass-burning and non-biomass-burning seasons at a rural site of Northeast China, Environmental Pollution, 231, 654-662, https://doi.org/10.1016/j.envpol.2017.08.045, 2017.
Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T. A., Brusniak, M.-Y., Paulse, C., Creasy, D., Flashner, L., Kani, K., Moulding, C., Seymour, S. L., Nuwaysir, L. M., Lefebvre, B., Kuhlmann, F., Roark, J., Rainer, P., Detlev, S., Hemenway, T., Huhmer, A., Langridge, J., Connolly, B., Chadick, T., Holly, K., Eckels, J., Deutsch, E. W., Moritz, R. L., Katz, J. E., Agus, D. B., MacCoss, M., Tabb, D. L., and Mallick, P.: A cross-platform toolkit for mass spectrometry and proteomics, Nat. Biotechnol, 30, 918-920, 10.1038/nbt.2377, 2012.
Chan, L. K., Nguyen, K. Q., Karim, N., Yang, Y., Rice, R. H., He, G., Denison, M. S., and Nguyen, T. B.: Relationship between the molecular composition, visible light absorption, and health-related properties of smoldering woodsmoke aerosols, Atmos Chem Phys, 20, 539-559, 2020.
Choudhary, V., Roson, M. L., Guo, X., Gautam, T., Gupta, T., and Zhao, R.: Aqueous-phase photochemical oxidation of water-soluble brown carbon aerosols arising from solid biomass fuel burning, Environmental Science: Atmospheres, 3, 816-829, 10.1039/d2ea00151a, 2023.
Daneshvar, N., Behnajady, M. A., and Zorriyeh Asghar, Y.: Photooxidative degradation of 4-nitrophenol (4-NP) in UV/H2O2 process: Influence of operational parameters and reaction mechanism, Journal of Hazardous Materials, 139, 275-279, https://doi.org/10.1016/j.jhazmat.2006.06.045, 2007.
https://cfmid.wishartlab.com/statistics, last access: 26/11/2024.
Divisekara, D. T. D. K.: UHPLC/FT-MS Non-Targeted Screening Approach for Biomass Burning Organic Aerosol and Liquid Smoke as Biomass Burning Organic Aerosol Surrogate, Michigan Technological University, 2023.
Divisekara, T., Schum, S., and Mazzoleni, L.: Ultrahigh performance LC/FT-MS non-targeted screening for biomass burning organic aerosol with MZmine2 and MFAssignR, Chemosphere, 338, 139403, https://doi.org/10.1016/j.chemosphere.2023.139403, 2023.
Evans, R. L., Bryant, D. J., Voliotis, A., Hu, D., Wu, H., Syafira, S. A., Oghama, O. E., McFiggans, G., Hamilton, J. F., and Rickard, A. R.: The importance of burning conditions on the composition of domestic biomass burning organic aerosol and the impact of atmospheric aging, EGUsphere, 2024, 1-38, 10.5194/egusphere-2024-2642, 2024a.
Evans, R. L., Bryant, D. J., Voliotis, A., Hu, D., Wu, H., Syafira, S. A., Oghama, O. E., McFiggans, G., Hamilton, J. F., and Rickard, A. R.: A Semi-Quantitative Approach to Nontarget Compositional Analysis of Complex Samples, Analytical Chemistry, 96, 18349-18358, 2024b.

Fan, X., Wei, S., Zhu, M., Song, J., and Peng, P.: Comprehensive characterization of humic-like substances in smoke PM2.5 emitted from the combustion of biomass materials and fossil fuels, Atmos. Chem. Phys., 16, 13321-13340, 10.5194/acp-16-13321-2016, 2016.

Fine, P. M., Cass, G. R., and Simoneit, B. R. T.: Chemical Characterization of Fine Particle Emissions from the Wood Stove Combustion of Prevalent United States Tree Species, Environmental Engineering Science, 21, 705-721, 10.1089/ees.2004.21.705, 2004.

Fleming, L. T., Lin, P., Roberts, J. M., Selimovic, V., Yokelson, R., Laskin, J., Laskin, A., and Nizkorodov, S. A.: Molecular composition and photochemical lifetimes of brown carbon chromophores in biomass burning organic aerosol, Atmos. Chem. Phys., 20, 1105-1129, 10.5194/acp-20-1105-2020, 2020.

Fleming, L. T., Lin, P., Laskin, A., Laskin, J., Weltman, R., Edwards, R. D., Arora, N. K., Yadav, A., Meinardi, S., Blake, D. R., Pillarisetti, A., Smith, K. R., and Nizkorodov, S. A.: Molecular composition of particulate matter emissions from dung and brushwood burning household cookstoves in Haryana, India, Atmos. Chem. Phys., 18, 2461-2480, 10.5194/acp-18-2461-2018, 2018.

https://external.gnps2.org/gnpslibrary last access: 26/11/2024.

Graham, B., Mayol-Bracero, O. L., Guyon, P., Roberts, G. C., Decesari, S., Facchini, M. C., Artaxo, P., Maenhaut, W., Köll, P., and Andreae, M. O.: Water-soluble organic compounds in biomass burning aerosols over Amazonia 1. Characterization by NMR and GC-MS, Journal of Geophysical Research: Atmospheres, 107, LBA 14-11-LBA 14-16, https://doi.org/10.1029/2001JD000336, 2002a.

Graham, B., Mayol-Bracero, O. L., Guyon, P., Roberts, G. C., Decesari, S., Facchini, M. C., Artaxo, P., Maenhaut, W., Köll, P., and Andreae, M. O.: Water-soluble organic compounds in biomass burning aerosols over Amazonia 1. Characterization by NMR and GC-MS, Journal of Geophysical Research: Atmospheres, 107, LBA 14-11-LBA 14-16, 2002b.

Harrison, M. A. J., Barra, S., Borghesi, D., Vione, D., Arsene, C., and Iulian Olariu, R.: Nitrated phenols in the atmosphere: a review, Atmospheric Environment, 39, 231-248, https://doi.org/10.1016/j.atmosenv.2004.09.044, 2005.

Hartikainen, A., Tiitta, P., Ihalainen, M., Yli-Pirilä, P., Orasche, J., Czech, H., Kortelainen, M., Lamberg, H., Suhonen, H., Koponen, H., Hao, L., Zimmermann, R., Jokiniemi, J., Tissari, J., and Sippula, O.: Photochemical transformation of residential wood combustion emissions: dependence of organic aerosol composition on OH exposure, Atmos. Chem. Phys., 20, 6357-6378, 10.5194/acp-20-6357-2020, 2020.

Hartner, E., Gawlitta, N., Gröger, T., Orasche, J. r., Czech, H., Geldenhuys, G.-L., Jakobi, G., Tiitta, P., Yli-Pirila, P., and Kortelainen, M.: Chemical Fingerprinting of Biomass Burning Organic Aerosols from Sugar Cane Combustion: Complementary Findings from Field and Laboratory Studies, ACS Earth and Space Chemistry, 8, 533-546, 2024a.

Hartner, E., Gawlitta, N., Gröger, T., Orasche, J., Czech, H., Geldenhuys, G.-L., Jakobi, G., Tiitta, P., Yli-Pirilä, P., Kortelainen, M., Sippula, O., Forbes, P., and Zimmermann, R.: Chemical Fingerprinting of Biomass Burning Organic Aerosols from Sugar Cane Combustion: Complementary Findings from Field and Laboratory Studies, ACS Earth and Space Chemistry, 8, 533-546, 10.1021/acsearthspacechem.3c00301, 2024b.

Hatch, L. E., Luo, W., Pankow, J. F., Yokelson, R. J., Stockwell, C. E., and Barsanti, K. C.: Identification and quantification of gaseous organic compounds emitted from biomass burning using two-dimensional gas chromatography–time-of-flight mass spectrometry, Atmos. Chem. Phys., 15, 1865-1899, 10.5194/acp-15-1865-2015, 2015.

Hems, R. F. and Abbatt, J. P. D.: Aqueous Phase Photo-oxidation of Brown Carbon Nitrophenols: Reaction Kinetics, Mechanism, and Evolution of Light Absorption, ACS Earth and Space Chemistry, 2, 225-234, 10.1021/acsearthspacechem.7b00123, 2018.
Hems, R. F., Schnitzler, E. G., Liu-Kang, C., Cappa, C. D., and Abbatt, J. P. D.: Aging of Atmospheric Brown Carbon Aerosol, ACS Earth and Space Chemistry, 5, 722-748, 10.1021/acsearthspacechem.0c00346, 2021.
Hems, R. F., Schnitzler, E. G., Bastawrous, M., Soong, R., Simpson, A. J., and Abbatt, J. P. D.: Aqueous Photoreactions of Wood Smoke Brown Carbon, ACS Earth and Space Chemistry, 4, 1149-1160, 10.1021/acsearthspacechem.0c00117, 2020.
Heuckeroth, S., Damiani, T., Smirnov, A., Mokshyna, O., Brungs, C., Korf, A., Smith, J. D., Stincone, P., Dreolin, N., Nothias, L.-F., Hyötyläinen, T., Orešič, M., Karst, U., Dorrestein, P. C., Petras, D., Du, X., van der Hooft, J. J. J., Schmid, R., and Pluskal, T.: Reproducible mass spectrometry data processing and compound annotation in MZmine 3, Nature Protocols, 19, 2597-2641, 10.1038/s41596-024-00996-y, 2024.
Holmes, B. J.: Oligomerization of Levoglucosan in Proxies of Biomass Burning Aerosols, UVM ScholarWorks UVM ScholarWorks, University of Vermont, 2008.
Hughey, C. A., Hendrickson, C. L., Rodgers, R. P., Marshall, A. G., and Qian, K.: Kendrick Mass Defect Spectrum: A Compact Visual Analysis for Ultrahigh-Resolution Broadband Mass Spectra, Anal. Chem., 73, 4676-4681, 10.1021/ac010560w, 2001.
Ibáñez, A. B. and Bauer, S.: Analytical method for the determination of organic acids in dilute acid pretreated biomass hydrolysate by liquid chromatography-time-of-flight mass spectrometry, Biotechnology for Biofuels, 7, 145, 10.1186/s13068-014-0145-3, 2014.
Iinuma, Y., Brüggemann, E., Gnauk, T., Müller, K., Andreae, M. O., Helas, G., Parmar, R., and Herrmann, H.: Source characterization of biomass burning particles: The combustion of selected European conifers, African hardwood, savanna grass, and German and Indonesian peat, Journal of Geophysical Research: Atmospheres, 112, doi:10.1029/2006JD007120, 2007.
Jen, C. N., Hatch, L. E., Selimovic, V., Yokelson, R. J., Weber, R., Fernandez, A. E., Kreisberg, N. M., Barsanti, K. C., and Goldstein, A. H.: Speciated and total emission factors of particulate organics from burning western US wildland fuels and their dependence on combustion efficiency, Atmos. Chem. Phys., 19, 1013-1026, 2019.
Kaivosoja, T., Virén, A., Tissari, J., Ruuskanen, J., Tarhanen, J., Sippula, O., and Jokiniemi, J.: Effects of a catalytic converter on PCDD/F, chlorophenol and PAH emissions in residential wood combustion, Chemosphere, 88, 278-285, https://doi.org/10.1016/j.chemosphere.2012.02.027, 2012.
Kawamura, K., Izawa, Y., Mochida, M., and Shiraiwa, T.: Ice core records of biomass burning tracers (levoglucosan and dehydroabietic, vanillic and p-hydroxybenzoic acids) and total organic carbon for past 300years in the Kamchatka Peninsula, Northeast Asia, Geochimica et Cosmochimica Acta, 99, 317-329, https://doi.org/10.1016/j.gca.2012.08.006, 2012.
Koch, B. P. and Dittmar, T.: From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter, Rapid Communications in Mass Spectrometry, 20, 926-932, 10.1002/rcm.2386, 2006.
Koss, A. R., Sekimoto, K., Gilman, J. B., Selimovic, V., Coggon, M. M., Zarzana, K. J., Yuan, B., Lerner, B. M., Brown, S. S., Jimenez, J. L., Krechmer, J., Roberts, J. M., Warneke, C., Yokelson, R. J., and de Gouw, J.: Non-methane organic gas emissions from biomass burning: identification, quantification, and emission factors from PTR-ToF during the FIREX 2016 laboratory experiment, Atmos. Chem. Phys., 18, 3299-3319, 10.5194/acp-18-3299-2018, 2018.

Kroflič, A., Huš, M., Grilc, M., and Grgić, I.: Underappreciated and Complex Role of Nitrous Acid in Aromatic Nitration under Mild Environmental Conditions: The Case of Activated Methoxyphenols, Environmental Science & Technology, 52, 13756-13765, 10.1021/acs.est.8b01903, 2018.
Kundu, S., Kawamura, K., Andreae, T. W., Hoffer, A., and Andreae, M. O.: Molecular distributions of dicarboxylic acids, ketocarboxylic acids and α-dicarbonyls in biomass burning aerosols: implications for photochemical production and degradation in smoke layers, Atmos. Chem. Phys., 10, 2209-2225, 10.5194/acp-10-2209-2010, 2010.
Lamnoi, S., Boonupara, T., Sumitsawan, S., Vongruang, P., Prapamontol, T., Udomkun, P., and Kaewlom, P.: Residues of atrazine and diuron in rice straw, soils, and air post herbicide-contaminated straw biomass burning, Sci. Rep., 14, 13327, 10.1038/s41598-024-64291-2, 2024.
Laskin, A., West, C. P., and Hettiyadura, A. P. S.: Molecular insights into the composition, sources, and aging of atmospheric brown carbon, Chemical Society Reviews, 10.1039/D3CS00609C, 2025.
Laszakovits, J. R. and MacKay, A. A.: Data-based chemical class regions for Van Krevelen diagrams, Journal of the American Society for Mass Spectrometry, 33, 198-202, 2021.
Li, W., Ge, P., Chen, M., Tang, J., Cao, M., Cui, Y., Hu, K., and Nie, D.: Tracers from Biomass Burning Emissions and Identification of Biomass Burning, Atmosphere, 12, 10.3390/atmos12111401, 2021.
Li, Y. J., Huang, D. D., Cheung, H. Y., Lee, A. K. Y., and Chan, C. K.: Aqueous-phase photochemical oxidation and direct photolysis of vanillin – a model compound of methoxy phenols from biomass burning, Atmos. Chem. Phys., 14, 2871-2885, 10.5194/acp-14-2871-2014, 2014.
Lin, P., Fleming, L. T., Nizkorodov, S. A., Laskin, J., and Laskin, A.: Comprehensive Molecular Characterization of Atmospheric Brown Carbon by High Resolution Mass Spectrometry with Electrospray and Atmospheric Pressure Photoionization, Analytical Chemistry, 90, 12493-12502, 10.1021/acs.analchem.8b02177, 2018.
Lin, P., Bluvshtein, N., Rudich, Y., Nizkorodov, S. A., Laskin, J., and Laskin, A.: Molecular Chemistry of Atmospheric Brown Carbon Inferred from a Nationwide Biomass Burning Event, Environmental Science & Technology, 51, 11561-11570, 10.1021/acs.est.7b02276, 2017.
López-Caravaca, A., Vicente, E. D., Figueiredo, D., Evtyugina, M., Nicolás, J. F., Yubero, E., Galindo, N., Ryšavý, J., and Alves, C. A.: Gaseous and aerosol emissions from open burning of tree pruning and hedge trimming residues: Detailed composition and toxicity, Atmos. Environ., 338, 120849, https://doi.org/10.1016/j.atmosenv.2024.120849, 2024.
Mazzoleni, L. R., Zielinska, B., and Moosmüller, H.: Emissions of Levoglucosan, Methoxy Phenols, and Organic Acids from Prescribed Burns, Laboratory Combustion of Wildland Fuels, and Residential Wood Combustion, Environmental Science & Technology, 41, 2115-2122, 10.1021/es061702c, 2007.
McCord, J., Newton, S., and Strynar, M.: Validation of quantitative measurements and semi-quantitative estimates of emerging perfluoroethercarboxylic acids (PFECAs) and hexfluoroprolyene oxide acids (HFPOAs), Journal of Chromatography A, 1551, 52-58, 2018.
https://www.metaboanalyst.ca/docs/Databases.xhtml, last access: 26/11/2024.
https://www.metaboanalyst.ca/MetaboAnalyst/docs/About.xhtml, last access: last access 01/10/2024.
https://mona.fiehnlab.ucdavis.edu/downloads, last access: last access date 08/082024.
Moschos, V., Christensen, C., Mouton, M., Fiddler, M. N., Isolabella, T., Mazzei, F., Massabò, D., Turpin, B. J., Bililign, S., and Surratt, J. D.: Quantifying the Light-Absorption Properties and Molecular Composition of Brown Carbon Aerosol from Sub-Saharan African Biomass Combustion, Environmental Science & Technology, 58, 4268-4280, 2024.

https://systemsomicslab.github.io/compms/msdial/main.html#MSP, last access: last access date 08/08/2024.
Nothias, L.-F., Petras, D., Schmid, R., Dührkop, K., Rainer, J., Sarvepalli, A., Protsyuk, I., Ernst, M., Tsugawa, H., Fleischauer, M., Aicheler, F., Aksenov, A. A., Alka, O., Allard, P.-M., Barsch, A., Cachet, X., Caraballo-Rodriguez, A. M., Da Silva, R. R., Dang, T., Garg, N., Gauglitz, J. M., Gurevich, A., Isaac, G., Jarmusch, A. K., Kameník, Z., Kang, K. B., Kessler, N., Koester, I., Korf, A., Le Gouellec, A., Ludwig, M., Martin H, C., McCall, L.-I., McSayles, J., Meyer, S. W., Mohimani, H., Morsy, M., Moyne, O., Neumann, S., Neuweger, H., Nguyen, N. H., Nothias-Esposito, M., Paolini, J., Phelan, V. V., Pluskal, T., Quinn, R. A., Rogers, S., Shrestha, B., Tripathi, A., van der Hooft, J. J. J., Vargas, F., Weldon, K. C., Witting, M., Yang, H., Zhang, Z., Zubeil, F., Kohlbacher, O., Böcker, S., Alexandrov, T., Bandeira, N., Wang, M., and Dorrestein, P. C.: Feature-based molecular networking in the GNPS analysis environment, Nat. Methods, 17, 905-908, 10.1038/s41592-020-0933-6, 2020.
Oros, D. R. and Simoneit, B. R.: Identification and emission factors of molecular tracers in organic aerosols from biomass burning Part 2. Deciduous trees, Appl. Geochem., 16, 1545-1565, 2001a.
Oros, D. R. and Simoneit, B. R. T.: Identification and emission factors of molecular tracers in organic aerosols from biomass burning Part 2. Deciduous trees, Applied Geochemistry, 16, 1545-1565, https://doi.org/10.1016/S0883-2927(01)00022-1, 2001b.
Oros, D. R., Abas, M. R. b., Omar, N. Y. M. J., Rahman, N. A., and Simoneit, B. R. T.: Identification and emission factors of molecular tracers in organic aerosols from biomass burning: Part 3. Grasses, Applied Geochemistry, 21, 919-940, 10.1016/j.apgeochem.2006.01.008, 2006.
Pang, Z., Lu, Y., Zhou, G., Hui, F., Xu, L., Viau, C., Spigelman, Aliya F., MacDonald, Patrick E., Wishart, David S., Li, S., and Xia, J.: MetaboAnalyst 6.0: towards a unified platform for metabolomics data processing, analysis and interpretation, Nucleic Acids Research, 52, W398-W406, 10.1093/nar/gkae253, 2024.
Pietrogrande, M. C., Bacco, D., and Rossi, M.: Chemical characterization of polar organic markers in aerosols in a local area around Bologna, Italy, Atmos. Environ., 75, 279-286, https://doi.org/10.1016/j.atmosenv.2013.04.023, 2013.
Pihlström, T., Fernández-Alba, A. R., Amate, C. F., Poulsen, M. E., Hardebusch, B., Anastassiades, M., Lippold, R., Cabrera, L. C., de Kok, A., and ORegan, F.: Analytical quality control and method validation procedures for pesticide residues analysis in food and feed SANTE 11312/2021, Sante, 11312, 2021.
Pratt, K. A., Murphy, S. M., Subramanian, R., DeMott, P. J., Kok, G. L., Campos, T., Rogers, D. C., Prenni, A. J., Heymsfield, A. J., Seinfeld, J. H., and Prather, K. A.: Flight-based chemical characterization of biomass burning aerosols within two prescribed burn smoke plumes, Atmos. Chem. Phys., 11, 12549-12565, 10.5194/acp-11-12549-2011, 2011.
Schauer, J. J., Kleeman, M. J., Cass, G. R., and Simoneit, B. R. T.: Measurement of Emissions from Air Pollution Sources. 3. C1−C29 Organic Compounds from Fireplace Combustion of Wood, Environmental Science & Technology, 35, 1716-1728, 10.1021/es001331e, 2001.
Schmid, R., Heuckeroth, S., Korf, A., Smirnov, A., Myers, O., Dyrlund, T. S., Bushuiev, R., Murray, K. J., Hoffmann, N., Lu, M., Sarvepalli, A., Zhang, Z., Fleischauer, M., Dührkop, K., Wesner, M., Hoogstra, S. J., Rudt, E., Mokshyna, O., Brungs, C., Ponomarov, K., Mutabdžija, L., Damiani, T., Pudney, C. J., Earll, M., Helmer, P. O., Fallon, T. R., Schulze, T., Rivas-Ubach, A., Bilbao, A., Richter, H., Nothias, L.-F., Wang, M., Orešič, M., Weng, J.-K., Böcker, S., Jeibmann, A., Hayen, H., Karst, U., Dorrestein, P. C., Petras, D., Du, X., and Pluskal, T.: Integrative analysis of multimodal mass spectrometry data in MZmine 3, Nature Biotechnology, 41, 447-449, 10.1038/s41587-023-01690-2, 2023.

Sengupta, D., Samburova, V., Bhattarai, C., Watts, A. C., Moosmüller, H., and Khlystov, A. Y.: Polar semivolatile organic compounds in biomass-burning emissions and their chemical transformations during aging in an oxidation flow reactor, Atmos. Chem. Phys., 20, 8227-8250, 10.5194/acp-20-8227-2020, 2020.
Shen, M., Ho, K. F., Dai, W., Liu, S., Zhang, T., Wang, Q., Meng, J., Chow, J. C., Watson, J. G., Cao, J., and Li, J.: Distribution and stable carbon isotopic composition of dicarboxylic acids, ketocarboxylic acids and α-dicarbonyls in fresh and aged biomass burning aerosols, Atmos. Chem. Phys., 22, 7489-7504, 10.5194/acp-22-7489-2022, 2022.
Simoneit, B. R., Rushdi, A., Bin Abas, M., and Didyk, B.: Alkyl amides and nitriles as novel tracers for biomass burning, Environ. Sci. Technol., 37, 16-21, 2003.
Simoneit, B. R. T.: Biomass burning — a review of organic tracers for smoke from incomplete combustion, Applied Geochemistry, 17, 129-162, https://doi.org/10.1016/S0883-2927(01)00061-0, 2002.
Smith, D. M., Cui, T., Fiddler, M. N., Pokhrel, R. P., Surratt, J. D., and Bililign, S.: Laboratory studies of fresh and aged biomass burning aerosol emitted from east African biomass fuels – Part 2: Chemical properties and characterization, Atmos. Chem. Phys., 20, 10169-10191, 10.5194/acp-20-10169-2020, 2020a.
Smith, D. M., Cui, T., Fiddler, M. N., Pokhrel, R. P., Surratt, J. D., and Bililign, S.: Laboratory studies of fresh and aged biomass burning aerosol emitted from east African biomass fuels–Part 2: Chemical properties and characterization, Atmos Chem Phys, 20, 10169-10191, 2020b.
Song, J., Li, M., Jiang, B., Wei, S., Fan, X., and Peng, P. a.: Molecular Characterization of Water-Soluble Humic like Substances in Smoke Particles Emitted from Combustion of Biomass Materials and Coal Using Ultrahigh-Resolution Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, Environmental Science & Technology, 52, 2575-2585, 10.1021/acs.est.7b06126, 2018.
Szeto, C., Lloyd, N., Nicolotti, L., Herderich, M. J., and Wilkinson, K. L.: Beyond Volatile Phenols: An Untargeted Metabolomic Approach to Revealing Additional Markers of Smoke Taint in Grapevines (Vitis vinifera L.) cv. Merlot, Journal of Agricultural and Food Chemistry, 72, 2018-2033, 10.1021/acs.jafc.2c09013, 2024.
Tomaz, S., Cui, T., Chen, Y., Sexton, K. G., Roberts, J. M., Warneke, C., Yokelson, R. J., Surratt, J. D., and Turpin, B. J.: Photochemical Cloud Processing of Primary Wildfire Emissions as a Potential Source of Secondary Organic Aerosol, Environmental Science & Technology, 52, 11027-11037, 10.1021/acs.est.8b03293, 2018.
Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., Ikeda, K., Kanazawa, M., VanderGheynst, J., Fiehn, O., and Arita, M.: MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis, Nat Methods, 12, 523-526, 10.1038/nmeth.3393, 2015.
Vaniya, A., Samra, S. N., Palazoglu, M., Tsugawa, H., and Fiehn, O.: Using MS-FINDER for identifying 19 natural products in the CASMI 2016 contest, Phytochem Lett, 21, 306-312, 10.1016/j.phytol.2016.12.008, 2017.
Vicente, E. D., Figueiredo, D., Gonçalves, C., Lopes, I., Oliveira, H., Kováts, N., Pinheiro, T., and Alves, C. A.: In vitro toxicity of indoor and outdoor PM10 from residential wood combustion, Science of The Total Environment, 782, 146820, https://doi.org/10.1016/j.scitotenv.2021.146820, 2021.
Wan, X., Kawamura, K., Ram, K., Kang, S., Loewen, M., Gao, S., Wu, G., Fu, P., Zhang, Y., Bhattarai, H., and Cong, Z.: Aromatic acids as biomass-burning tracers in atmospheric aerosols and ice cores: A review, Environmental Pollution, 247, 216-228, https://doi.org/10.1016/j.envpol.2019.01.028, 2019.
Wang, F., Liigand, J., Tian, S., Arndt, D., Greiner, R., and Wishart, D. S.: CFM-ID 4.0: More Accurate ESI-MS/MS Spectral Prediction and Compound Identification, Anal Chem, 93, 11692-11700, 10.1021/acs.analchem.1c01465, 2021.

Wang, F., Allen, D., Tian, S., Oler, E., Gautam, V., Greiner, R., Metz, T. O., and Wishart, D. S.: CFM-ID 4.0 - a web server for accurate MS-based metabolite identification, Nucleic Acids Res, 50, W165-w174, 10.1093/nar/gkac383, 2022.

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., Nguyen, D. D., Watrous, J., Kapono, C. A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A. V., Meehan, M. J., Liu, W.-T., Crüsemann, M., Boudreau, P. D., Esquenazi, E., Sandoval-Calderón, M., Kersten, R. D., Pace, L. A., Quinn, R. A., Duncan, K. R., Hsu, C.-C., Floros, D. J., Gavilan, R. G., Kleigrewe, K., Northen, T., Dutton, R. J., Parrot, D., Carlson, E. E., Aigle, B., Michelsen, C. F., Jelsbak, L., Sohlenkamp, C., Pevzner, P., Edlund, A., McLean, J., Piel, J., Murphy, B. T., Gerwick, L., Liaw, C.-C., Yang, Y.-L., Humpf, H.-U., Maansson, M., Keyzers, R. A., Sims, A. C., Johnson, A. R., Sidebottom, A. M., Sedio, B. E., Klitgaard, A., Larson, C. B., Boya P, C. A., Torres-Mendoza, D., Gonzalez, D. J., Silva, D. B., Marques, L. M., Demarque, D. P., Pociute, E., O'Neill, E. C., Briand, E., Helfrich, E. J. N., Granatosky, E. A., Glukhov, E., Ryffel, F., Houson, H., Mohimani, H., Kharbush, J. J., Zeng, Y., Vorholt, J. A., Kurita, K. L., Charusanti, P., McPhail, K. L., Nielsen, K. F., Vuong, L., Elfeki, M., Traxler, M. F., Engene, N., Koyama, N., Vining, O. B., Baric, R., Silva, R. R., Mascuch, S. J., Tomasi, S., Jenkins, S., Macherla, V., Hoffman, T., Agarwal, V., Williams, P. G., Dai, J., Neupane, R., Gurr, J., Rodríguez, A. M. C., Lamsa, A., Zhang, C., Dorrestein, K., Duggan, B. M., Almaliti, J., Allard, P.-M., Phapale, P., Nothias, L.-F., Alexandrov, T., Litaudon, M., Wolfender, J.-L., Kyle, J. E., Metz, T. O., Peryea, T., Nguyen, D.-T., VanLeer, D., Shinn, P., Jadhav, A., Müller, R., Waters, K. M., Shi, W., Liu, X., Zhang, L., Knight, R., Jensen, P. R., Palsson, B. Ø., Pogliano, K., Linington, R. G., Gutiérrez, M., Lopes, N. P., Gerwick, W. H., Moore, B. S., Dorrestein, P. C., and Bandeira, N.: Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking, Nature Nat. Biotechnol, 34, 828-837, 10.1038/nbt.3597, 2016.

Xie, M., Chen, X., Hays, M. D., and Holder, A. L.: Composition and light absorption of N-containing aromatic compounds in organic aerosols from laboratory biomass burning, Atmos. Chem. Phys., 19, 2899-2915, 10.5194/acp-19-2899-2019, 2019.

Yadav, S., Kapoor, T. S., Vernekar, P., and Phuleria, H. C.: Examining the Chemical and Optical Properties of Biomass-burning Aerosols and their Impact on Oxidative Potential, Aerosol and Air Quality Research, 23, 10.4209/aaqr.230102, 2023.

Yee, L., Kautzman, K., Loza, C., Schilling, K., Coggon, M., Chhabra, P., Chan, M., Chan, A., Hersey, S., and Crounse, J.: Secondary organic aerosol formation from biomass burning intermediates: phenol and methoxyphenols, Atmos Chem Phys, 13, 8019-8043, 2013.

Zhong, S., Chen, S., Deng, J., Fan, Y., Zhang, Q., Xie, Q., Qi, Y., Hu, W., Wu, L., Li, X., Pavuluri, C. M., Zhu, J., Wang, X., Liu, D., Pan, X., Sun, Y., Wang, Z., Xu, Y., Tong, H., Su, H., Cheng, Y., Kawamura, K., and Fu, P.: Impact of biogenic secondary organic aerosol (SOA) loading on the molecular composition of wintertime PM2.5 in urban Tianjin: an insight from Fourier transform ion cyclotron resonance mass spectrometry, Atmos. Chem. Phys., 23, 2061-2077, 10.5194/acp-23-2061-2023, 2023.