



*Supplement of*

## **Critical contribution of chemically diverse carbonyl molecules to the oxidative potential of atmospheric aerosols**

**Feifei Li et al.**

*Correspondence to:* Jitao Lv ([jtlv@rcees.ac.cn](mailto:jtlv@rcees.ac.cn)) and Yawei Wang ([ywwang@rcees.ac.cn](mailto:ywwang@rcees.ac.cn))

The copyright of individual parts of the supplement might differ from the article licence.

## List of supporting information

### Sections

S1: Chemicals.

S2: Solid phase extraction procedure.

5 S3: ESI-FT-ICR MS analysis.

S4: Automatic screening procedure for carbonyl molecules

S5: Reduction of the carbonyl group by NaBH<sub>4</sub>.

S6: Supplement statistical analysis.

### Tables

10 Table S1: Detailed information on representative PM<sub>2.5</sub> samples.

Table S2: The number of single, double and triple derivatized carbonyl and non-carbonyl molecules.

### Fig.s

Fig. S1: Mass spectra of typical carbonyls and their PFBHA derivatives in the sample derivatized by PFBHA.

15 Fig. S2: Algorithm flow for identifying carbonyl molecules based on library comparison and derivatization reaction scheme of PFBHA and carbonyl molecules.

Fig. S3: Reproducibility of DTT measurements of samples.

Fig. S4: UV-Vis spectra during SRNOM reduction by NaBH<sub>4</sub>.

Fig. S5: Validation plot of PLSR model for carbonyl molecules significantly positively correlated with DTT<sub>OC</sub>.

Fig. S6: PCA score plots for pro-oxidative carbonyls.

20 Fig. S7: Multivariate OPLS-DA (1 + 1 + 0) models for the pro-oxidative carbonyls.

Fig. S8: Daily PM<sub>2.5</sub> concentration at the downtown site, suburban site, and mountainous site in Beijing during each sampling season.

Fig. S9: The proportion of molecules in five categories was estimated based on the molecular numbers.

25 Fig. S10: Five categories differ in the percentage distribution of all organic, non-carbonyl, and carbonyl molecules based on the molecular numbers.

Fig. S11: The proportion of each group in all organic, non-carbonyl, and carbonyl molecules was estimated based on the molecular numbers.

Fig. S12: Differences in the percentage distribution of elemental composition based on the molecular numbers in all organic, non-carbonyl, and carbonyl molecules.

30 Fig. S13: DTT consumption rates of organic in PM<sub>2.5</sub> over Beijing at the downtown site, suburban site, and mountainous site.

Fig. S14: Correlation analysis between the normalized molecular intensities of carbonyl and non-carbonyl molecules.

Fig. S15. The DTT activity of WSOM from PM<sub>2.5</sub> sample before and after NaBH<sub>4</sub> treatment was measured.

Fig. S16. The magnitude-weighted average parameters of carbonyl molecules.

Fig. S17: The difference in the percentage distribution based on the molecular numbers of carbonyl molecules in winter, winter Olympics, and summer samples for the five categories.

Fig. S18: The van Krevelen diagram of carbonyl molecules from different elemental compositions.

Fig. S19: Spearman correlation matrix of carbonyl molecule characteristic parameters and DTT<sub>OC</sub>.

Fig. S20: OPLS-DA score plots for normalized intensity of individual pro-oxidative carbonyls.

## 40 S1: Chemicals

Ultrapure water (18.2 M $\Omega$ ·cm) was supplied by a Milli-Q water 100 purification system. Acetonitrile and methanol (HPLC grade) were obtained from Fisher. *O*-(2,3,4,5,6-Pentafluorobenzyl) hydroxylamine hydrochloride (PFBHA,  $\geq$  99%, CAS 57981-02-9) obtained from Sigma-Aldrich and Sodium borohydride (NaBH<sub>4</sub>,  $\geq$  98%, CAS 16940-66-2) obtained from Sinopharm Chemical Reagent Co. LTD.

45 Dithiothreitol (DTT,  $\geq$  99.4%, Biotechnology grade, Amresco), Diethylenetriaminepentaacetic acid (DTPA,  $\geq$  99%, CAS 67-43-6, Sigma-Aldrich), 5,5-Dithiobis (2-nitrobenzoic acid) (DTNB,  $>$  99%, CAS 69-78-3, J&K Scientific) and phosphate buffer saline (PBS, 0.1M, pH 7.2-7.4, Solarbio) were used to measure the OP of PM<sub>2.5</sub> samples.

9,10-anthraquinone (CAS 84-65-1, C<sub>14</sub>O<sub>2</sub>H<sub>8</sub>), 5-Hydroxy-1,4-naphthoquinone (CAS 481-39-0, C<sub>10</sub>O<sub>3</sub>H<sub>6</sub>), benzoylformic acid (CAS 611-73-4, C<sub>8</sub>O<sub>3</sub>H<sub>6</sub>), trans-cinnamic acid (CAS 140-10-3, C<sub>9</sub>O<sub>2</sub>H<sub>8</sub>), syringaldehyde (CAS 134-96-3, C<sub>9</sub>H<sub>10</sub>O<sub>4</sub>), and  
50 acetovanillone (CAS 498-02-2, C<sub>9</sub>H<sub>10</sub>O<sub>3</sub>) were used to test the reaction efficiency of carbonyls with PFBHA.

## S2: Solid phase extraction procedure

The water-soluble fraction of PM<sub>2.5</sub> was desalted and purified using Varian Bond Elute PPL cartridges (200 mg per 3 mL). Briefly, the cartridges were rinsed with 3 mL of methanol (MS grade) and 9 ml of ultrapure water (pH 2) prior to use. The samples were acidified to pH 2 with HCl (32%, ultrapure) to increase the extraction efficiency for organic acids and phenols,  
55 and 20 mL were passed through the cartridges by gravity at a flow rate of approximately 2 mL min<sup>-1</sup>. Cartridges were rinsed with three volumes of 0.01 M HCl for removal of salts, dried with a stream of N<sub>2</sub>, and immediately extracted with two volumes of methanol (MS grade). The extracted WEOM samples were stored at -4 °C in the dark, and the WISOM was concentrated with N<sub>2</sub> to 1 mL and then stored at -4 °C in the dark.

## S3: ESI-FT-ICR MS analysis

60 Samples for ESI FT-ICR-MS analysis were continuously infused into the ESI unit by syringe infusion at a flow rate of 120  $\mu$ L h<sup>-1</sup>. The ESI needle voltage was set to -3.8 kV. All the samples were analyzed in negative ionization mode with broadband detection. Ions accumulated in a hexapole ion trap for 0.2 s before being introduced into the ICR cell. The lower mass limit was set to  $m/z = 120$  Da and the upper mass limit to  $m/z = 920$  Da. 300 mass spectra were averaged per sample. The spectra were externally calibrated with 10 mM of sodium formate solution in 50% isopropyl alcohol using a linear  
65 calibration and then internally recalibrated using an in-house reference mass list. After internal calibration, the mass error was  $< 0.5$  ppm over the entire mass range.

Based on previous studies, the following sample molecular parameters were calculated:

$$\text{DBE} = 1 + \frac{2C-H+N}{2} \text{ (Lv et al., 2016)}$$

$$\text{AImod} = (1 + C - \frac{O}{2} - \frac{H}{2} - S) / (C - \frac{O}{2} - N - S) \text{ (Wang et al., 2021)}$$

70

$$\text{NOSC} = 4 - [(4C + H - 3N - 2O - 2S) / C] \text{ (Zhang et al., 2023)}$$

$$\overline{\text{OS}}_C \approx 2 \frac{O}{2} - \frac{H}{2} \text{ (Kroll et al., 2011)}$$

Where C, H, N, O, and S represent the numbers of carbon, hydrogen, nitrogen, oxygen, and sulfur atoms in each formula.

The magnitude-weighted average parameters such as C, H, O, N, S, MW, O/C, H/C, NOSC, AImod, and DBE for each sample can be determined by the following formula:

75

$$(M)_W = [\sum_i I_i \times (M)_i] / \sum_i I_i$$

where  $I_i$  and  $(M)_i$  are the relative abundance and M value of peak  $i$ , respectively. The relative abundance is calculated as the abundance of the individual peak divided by the maximum abundances in a given spectrum (Lv et al., 2016).

#### S4: Automatic screening procedure for carbonyl molecules

Carbonyl molecules identification was automatically completed in R with the algorithm (Yu et al., 2024). Firstly, fluorine-  
 80 containing and non-fluorine-containing formula libraries were built separately and filtered using the following the rules:  $0 < H/C < 2.2$ ;  $0 < O/C < 1.2$ ;  $0 < DBE \leq 50$  (Zhang and Fu, 2022);  $DBE-O \leq 10$ ; nitrogen rule. Subsequently, the MS peaks with a signal-to-noise ratio  $(S/N) \geq 4$  in original and derivative DOM samples were automatically assigned formulas according to stringent criteria with elemental combinations of  $C_{5-60}H_{10-120}O_{1-40}N_{0-3}S_{0-1}$  and  $C_{5-60}H_{10-120}O_{1-40}N_{0-3}S_{0-1}F_{5or10or15}$  respectively by comparing the measured  $m/z$  values to the theoretical  $m/z$  values of the  
 85 filtered molecules (Fu et al., 2020). The optimum formula for an exact mass with multiple possible formulae was selected by following the rules with priority as follows: (1) the mass error is less than 0.5 ppm; (2) minimum number of heteroatoms (N+S); (3) lowest mass error (Koch et al., 2007). Once formula assignment of original and derivative samples was accomplished, the carbonyl molecules in  $PM_{2.5}$  were further screened according to the mass difference of oximation reaction pairs in Fig. S1.

#### 90 S5: Reduction of the carbonyl group by $NaBH_4$

The aqueous solution of SRNOM was de-oxygenated by nitrogen purging for 15 min, and an excess of  $NaBH_4$  was added under nitrogen protection and reacted in the dark. (Phillips and Smith, 2014, 2015)

Changes in light absorption during reduction were monitored using UV 3600 (Shimadzu, Japan) to determine the mass and time of  $NaBH_4$  required for the carbonyl group to be completely reduced. The mass ratios of  $NaBH_4$  to SRNOM were  
 95 set to 1:1, 10:1, and 20:1, and the UV-vis absorbances after the reactions of 0.2H, 2H, and 12H were recorded, respectively (Fig. S3). It can be found that the absorbances of the mass ratios of 10:1 and 20:1 were the same after the reaction 12H, so  $NaBH_4$  at 10 times the sample mass completely consumed the carbonyl group in the sample after the reaction 12H.

## S6: Supplement statistical analysis

100 The pro-oxidative carbonyls are screened by the following steps. Carbonyl molecules with more than 50% detection rate in all samples were first filtered as common molecules (1314), and the correlation between the normalized molecular intensity of common carbonyl molecules and  $DTT_{OC}$  was analyzed by Spearman rank correlation. Due to the large dataset, the use of the FDR method can reduce the false positive results during multiple hypothesis testing in large datasets. The p-values in Spearman's rank correlation results were corrected by FDR and labeled as  $p_{fdr}$ . A total of 705 carbonyl molecules were screened for significant positive correlation with  $DTT_{OC}$  based on  $p_{fdr} < 0.05$  and Spearman's  $r > 0$ . Next, the PLSR model  
105 was used to select the key variables among oxidation potential-related carbonyl molecules to avert the influence of collinearity in multivariate correlation analysis. The criterion for screening important variables was set as a  $VIP > 1$ , which screened a total of 380 pro-oxidative carbonyls. The results of 999-time permutation tests indicated that the PLSR model was valid and non-overfit (Fig. S4).

110 The samples were categorized into winter, Winter Olympics, and summer groups according to the sampling time, and no outliers were detected by the quality assessment of the sample variants by the PCA model (Fig. S5). Further multivariate OPLS-DA (1 + 1 + 0) models explained the seasonal differences in the abundance of pro-oxidative carbonyls (Fig. S17), and the model was validated by permutation tests ( $n=200$ ) (Fig. S6).

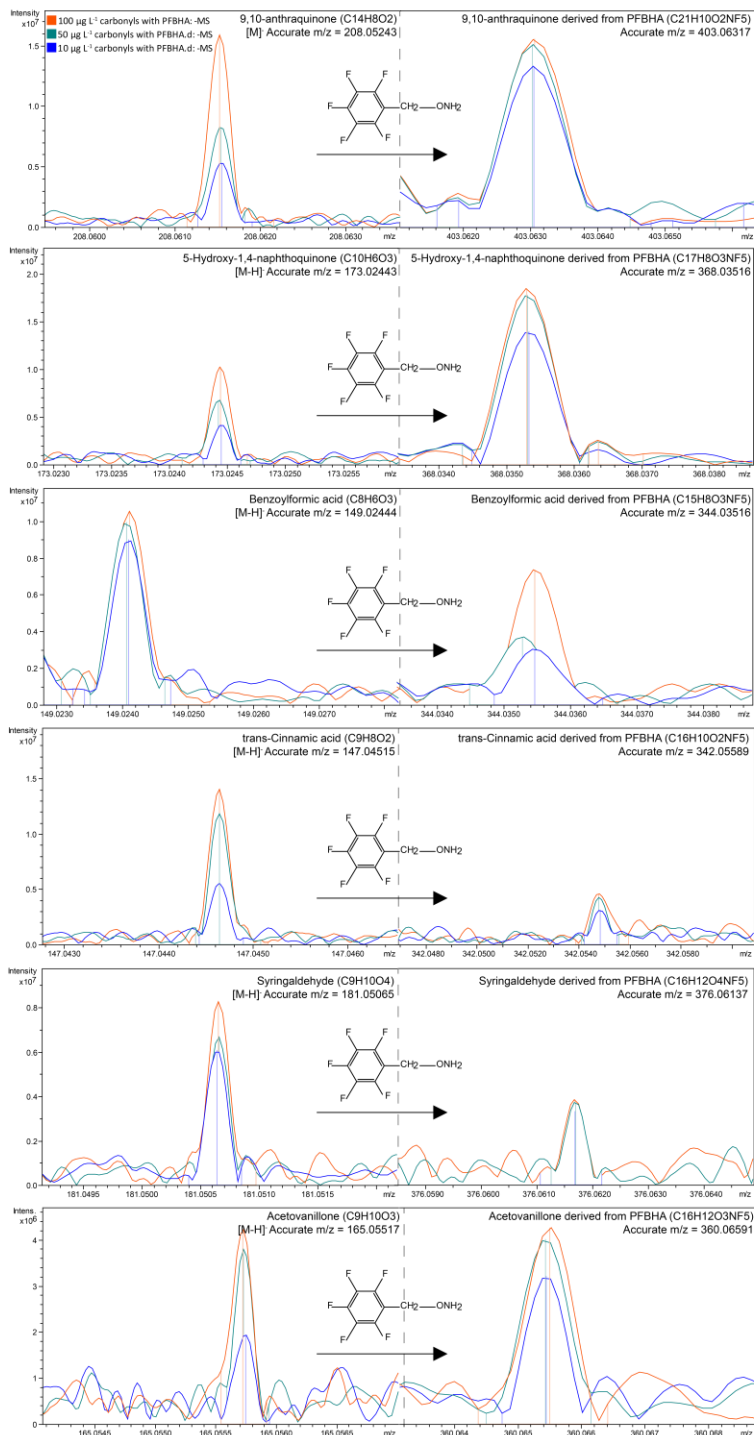
113 **Table S1: Detailed information on representative PM<sub>2.5</sub> samples.**

Sampling site	Sampling number	Sampling season	PM <sub>2.5</sub> concentrations				WSOM			WISOM
			Date	µg m <sup>-3</sup>	Date	µg m <sup>-3</sup>	TOC (mg L <sup>-1</sup> )	DTToc (pmol min <sup>-1</sup> ngC <sup>-1</sup> )	DTTv (pmol min <sup>-1</sup> m <sup>-3</sup> )	DTTv (pmol min <sup>-1</sup> m <sup>-3</sup> )
Mountainous site	M-1	Winter	19 Jan. 2022	37.73	20 Jan. 2022	30.26	116.40	66.33	150.26	110.51
	M-2	Winter	23 Jan. 2022	49.95	24 Jan. 2022	57.66	323.20	37.81	178.37	104.45
	M-3	Winter	31 Jan. 2022	9.86	1 Feb. 2022	7.71	42.94	55.41	17.36	16.58
	M-4	Olympics	9 Feb. 2022	29.78	10 Feb. 2022	36.03	199.10	43.72	127.06	73.52
	M-5	Olympics	6 Feb. 2022	7.86	14 Feb. 2022	10.11	61.17	57.89	25.84	14.13
	M-6	Olympics	8 Feb. 2022	13.56	18 Feb. 2022	12.35	115.27	64.88	54.58	22.03
	M-7	Summer	10 Jul. 2022	19.02	27 Jul. 2022	11.75	130.07	7.96	7.40	3.57
	M-8	Summer	16 Jul. 2022	43.88	24 Jul. 2022	36.60	166.13	10.70	25.40	8.61
	M-9	Summer	11 Jul. 2022	24.57	20 Jul. 2022	32.20	166.33	8.97	21.31	11.58
Suburban site	S-1	Winter	19 Jan. 2022	15.15	20 Jan. 2022	17.86	189.30	57.98	160.19	29.81
	S-2	Winter	23 Jan. 2022	62.86	24 Jan. 2022	54.23	488.70	40.18	286.58	159.05
	S-3	Winter	31 Jan. 2022	3.00	1 Feb. 2022	6.78	41.30	84.16	50.73	14.40
	S-4	Olympics	9 Feb. 2022	38.13	10 Feb. 2022	33.61	328.10	52.44	251.07	89.09
	S-5	Olympics	6 Feb. 2022	5.62	14 Feb. 2022	4.23	82.57	58.75	35.40	17.16
	S-6	Olympics	8 Feb. 2022	13.34	18 Feb. 2022	7.85	93.19	57.88	78.72	43.25
	S-7	Summer	10 Jul. 2022	11.49	27 Jul. 2022	14.85	112.07	11.57	9.26	8.28
	S-8	Summer	16 Jul. 2022	40.73	24 Jul. 2022	42.94	213.23	11.54	35.14	19.89
	S-9	Summer	11 Jul. 2022	33.22	20 Jul. 2022	21.35	136.73	10.98	21.79	15.27
Downtown site	D-1	Winter	19 Jan. 2022	43.86	20 Jan. 2022	47.92	162.20	52.87	166.87	101.65
	D-2	Winter	23 Jan. 2022	145.92	24 Jan. 2022	145.70	391.30	37.80	215.90	116.35
	D-3	Winter	31 Jan. 2022	10.49	1 Feb. 2022	13.65	85.70	45.42	56.81	33.85
	D-4	Olympics	9 Feb. 2022	63.18	10 Feb. 2022	64.05	263.40	51.40	197.59	122.90
	D-5	Olympics	6 Feb. 2022	10.11	14 Feb. 2022	9.39	106.87	36.76	28.67	16.19
	D-6	Olympics	8 Feb. 2022	24.38	18 Feb. 2022	30.63	74.79	71.00	77.50	42.01
	D-7	Summer	10 Jul. 2022	5.18	27 Jul. 2022	4.72	134.07	12.57	12.04	8.94
	D-8	Summer	16 Jul. 2022	15.30	24 Jul. 2022	10.11	230.73	11.29	37.21	20.04
	D-9	Summer	11 Jul. 2022	8.05	20 Jul. 2022	8.81	306.97	9.90	21.72	15.33

**Table S2: The number of single, double and triple derivatized carbonyl and non-carbonyl molecules.**

Sampling site	Sample number	Singly derivatized carbonyls (1 C=O)		Doubly derivatized carbonyls (2 C=O)		Triply derivatized carbonyls (3 C=O)		Number of carbonyl molecules	Number of non-carbonyl molecules	Number of all molecules
		Number	Proportion	Number	Proportion	Number	Proportion			
Mountainous site	M-1	1853	90%	197	10%	1	0%	2051	8308	10359
	M-2	1524	74%	242	12%	2	0%	1768	8633	10401
	M-3	1256	61%	327	16%	2	0%	1585	5784	7369
	M-4	1560	76%	233	11%	2	0%	1795	8032	9827
	M-5	1279	62%	360	18%	6	0%	1645	6781	8426
	M-6	1555	76%	205	10%	4	0%	1764	8345	10109
	M-7	1037	51%	182	9%	11	1%	1230	7119	8349
	M-8	1196	58%	180	9%	20	1%	1396	8442	9838
	M-9	1429	70%	168	8%	15	1%	1612	7810	9422
Suburban site	S-1	1655	81%	169	8%	1	0%	1825	7572	9397
	S-2	1185	58%	115	6%	6	0%	1306	8839	10145
	S-3	1449	71%	354	17%	6	0%	1809	7083	8892
	S-4	1470	72%	108	5%	1	0%	1579	7821	9400
	S-5	1502	73%	264	13%	3	0%	1769	7501	9270
	S-6	1519	74%	195	10%	0	0%	1714	7138	8852
	S-7	1042	51%	126	6%	17	1%	1185	7680	8865
	S-8	1387	68%	279	14%	10	0%	1676	8227	9903
	S-9	1282	63%	145	7%	6	0%	1433	7489	8922
Downtown site	D-1	1544	75%	140	7%	2	0%	1686	7104	8790
	D-2	1475	72%	228	11%	2	0%	1705	8218	9923
	D-3	1127	55%	156	8%	7	0%	1290	6222	7512
	D-4	1467	72%	167	8%	6	0%	1640	7699	9339
	D-5	1274	62%	226	11%	1	0%	1501	6518	8019
	D-6	1260	61%	117	6%	2	0%	1379	7894	9273
	D-7	1188	58%	153	7%	13	1%	1354	7687	9041
	D-8	1509	74%	253	12%	8	0%	1770	8227	9997
	D-9	977	48%	101	5%	14	1%	1092	8585	9677





**Fig. S1. Mass spectra of 9,10-anthraquinone, 5-hydroxy-1,4-naphthoquinone, benzoylformic acid, trans-Cinnamic acid, syringaldehyde, acetovanillone and their PFBHA derivatives in the sample derivatized by PFBHA.**

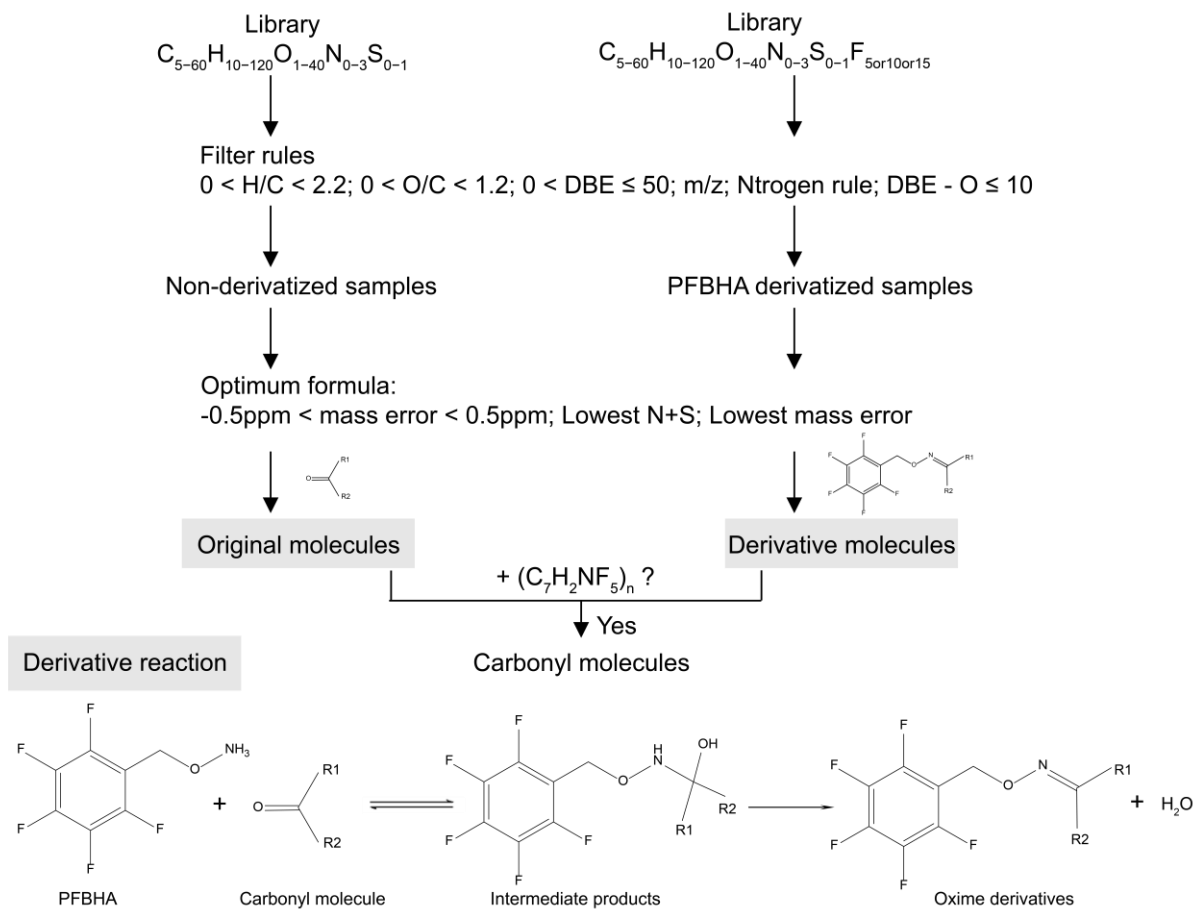
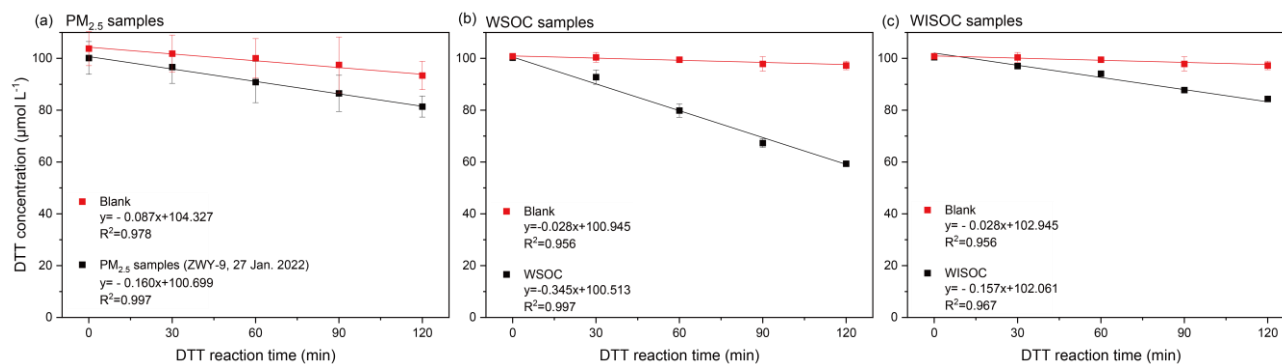


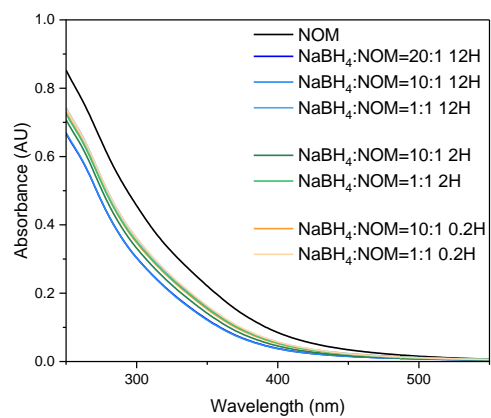
Fig. S2. Algorithm flow for identifying carbonyl molecules based on library comparison.



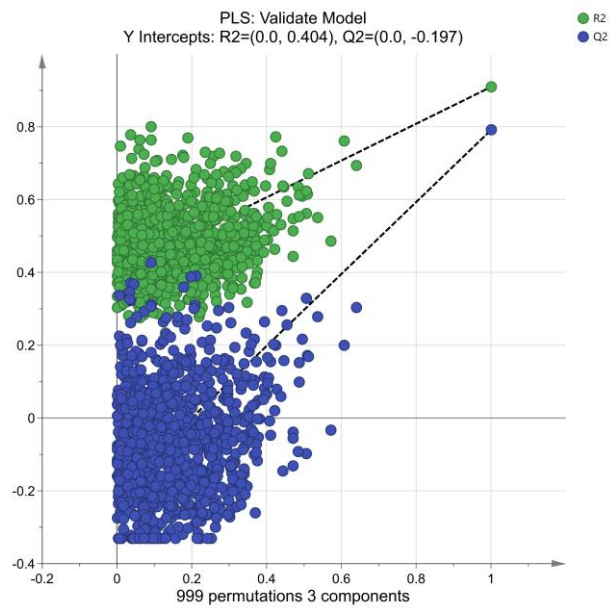
120

**Fig. S3. Reproducibility of DTT measurements of samples. (a) DTT measurements of  $PM_{2.5}$  samples and backgrounds, using the downtown site on January 27, 2022, as an example. (b) DTT measurements of a water-soluble organic carbon (WSOM) quality control sample, which was prepared containing a mixture of all WSOM samples. (c) DTT measurements of a water insoluble organic carbon (WISOM) quality control sample, which was prepared containing a mixture of all WISOM samples.**

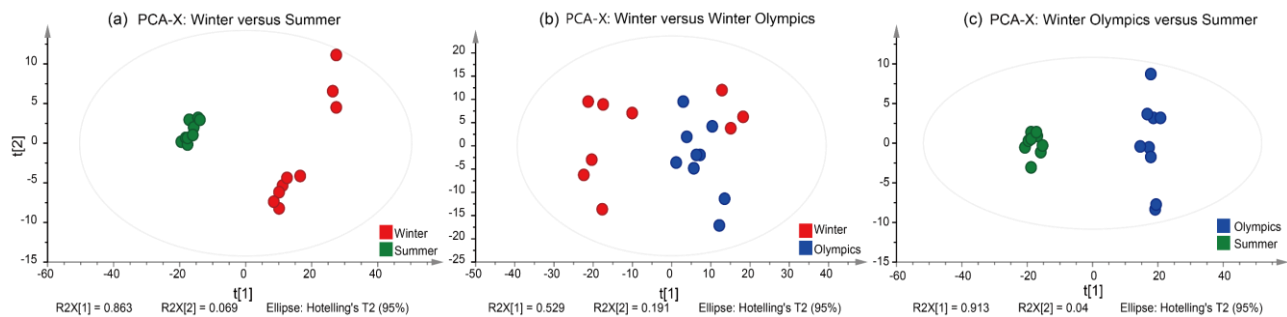
125



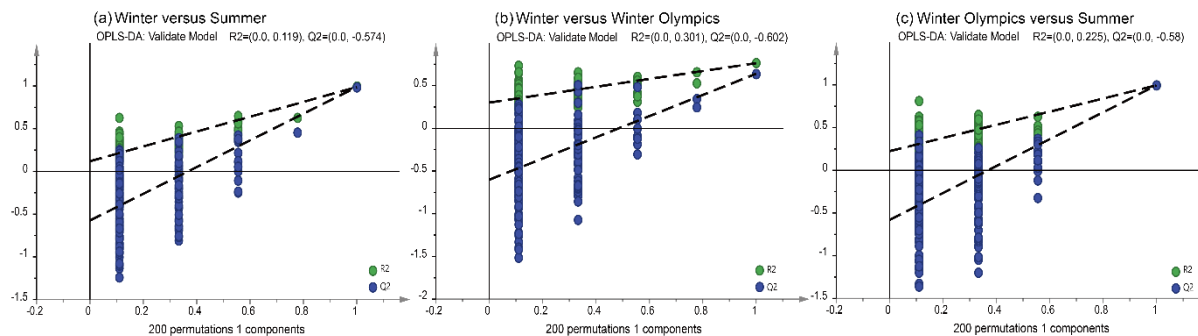
**Fig. S4. UV-Vis spectra during SRNOM reduction by NaBH<sub>4</sub>.**



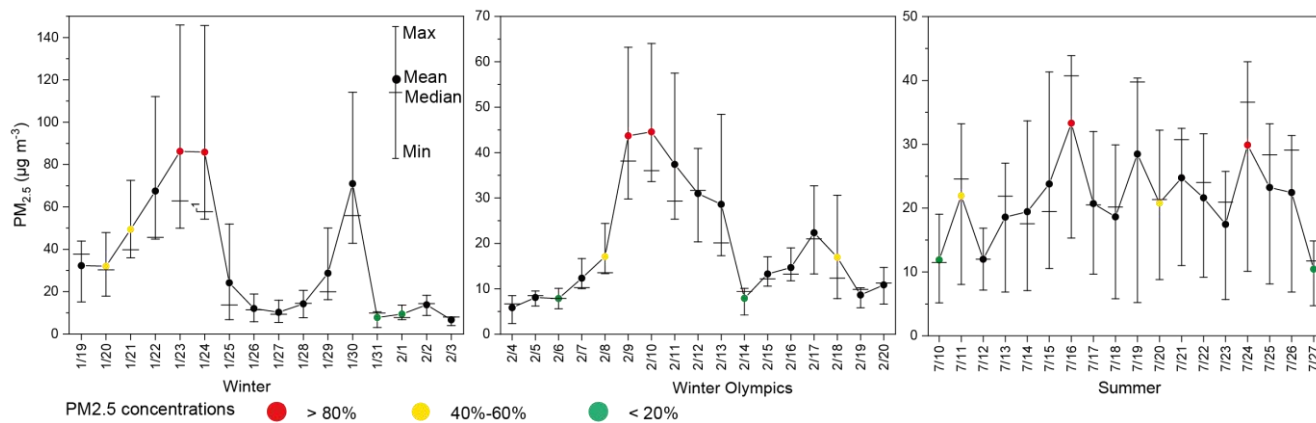
**Fig. S5.** Validation plot of PLSR model for carbonyl molecules significantly positively correlated with  $DTT_{OC}$ ,  $n = 999$ ,  $R^2 = (0.0, 0.404)$ ,  $Q^2 = (0.0, -0.197)$ .



**Fig. S6. PCA score plots for pro-oxidative carbonyls in (a) winter versus summer, (b) winter versus winter Olympics, and (c) winter Olympics versus summer samples.**



135 **Fig. S7. Multivariate OPLS-DA (1 + 1 + 0) models for the pro-oxidative carbonyls were validated by permutation test (n = 200). (a) Validation plot of OPLS-DA model obtained from carbonyl molecules in winter and summer samples,  $R^2 = (0.0, 0.019)$ ,  $Q^2 = (0.0, -0.574)$ ; (b) Validation plot of OPLS-DA model obtained from carbonyl molecules in winter and Winter Olympic samples,  $R^2 = (0.0, 0.301)$ ,  $Q^2 = (0.0, -0.602)$ ; (c) Validation plot of OPLS-DA model obtained from carbonyl molecules in winter Olympics and summer samples,  $R^2 = (0.0, 0.225)$ ,  $Q^2 = (0.0, -0.58)$ .**

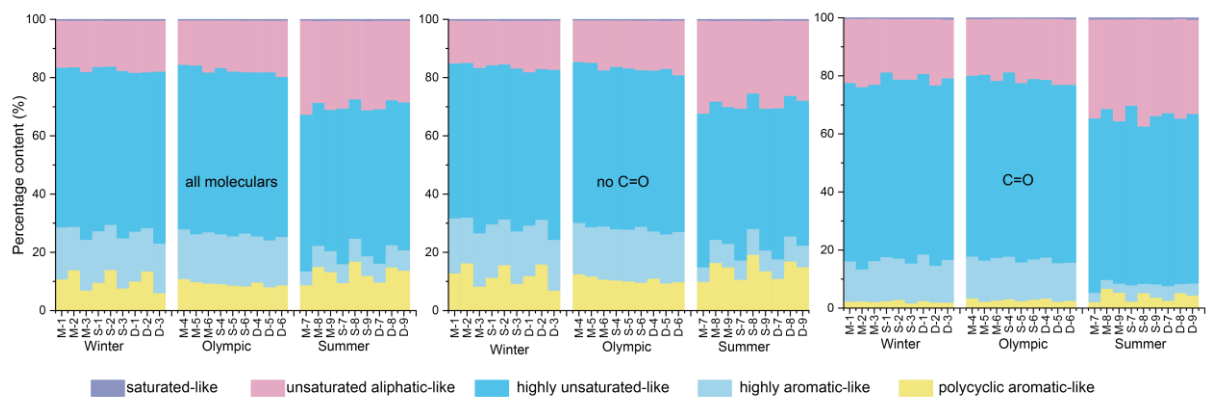


140

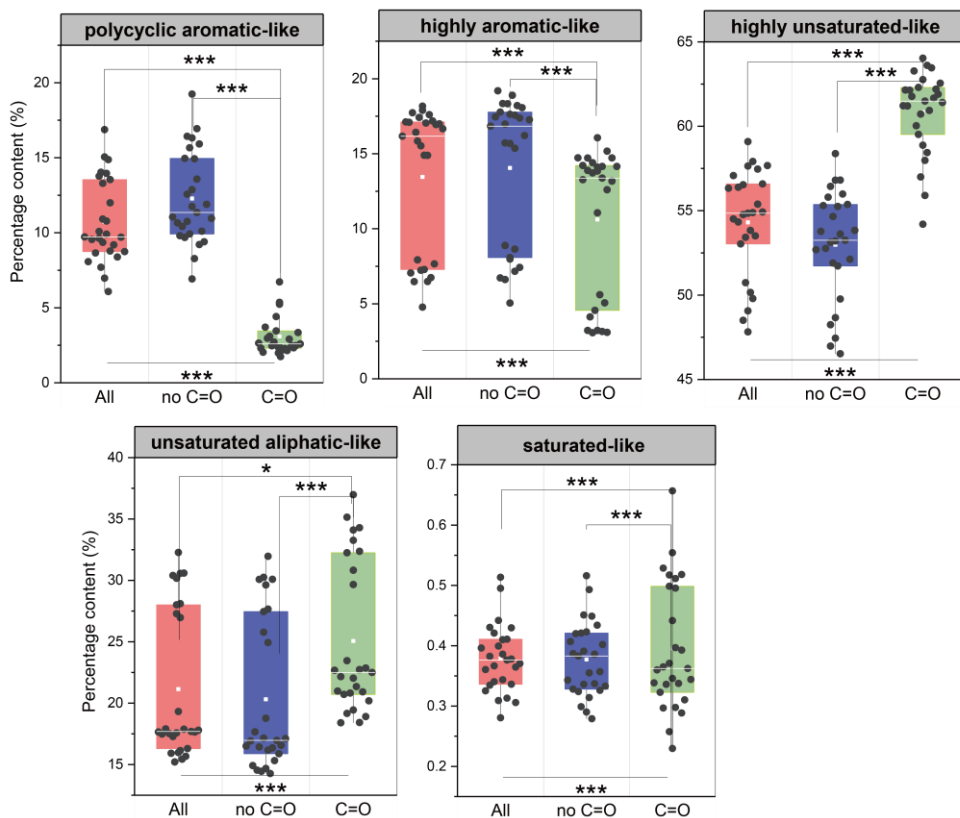
**Fig. S8. Daily PM<sub>2.5</sub> concentration at the downtown site, suburban site, and mountainous site in Beijing during each sampling season. The red dots represent the day on which PM<sub>2.5</sub> concentrations exceeded 80% of that sampling season. The yellow dots represent the day on which PM<sub>2.5</sub> concentrations are at 40-60% of that sampling season. The green dots represent the day on which PM<sub>2.5</sub> concentrations are below 20% of that sampling season.**

145

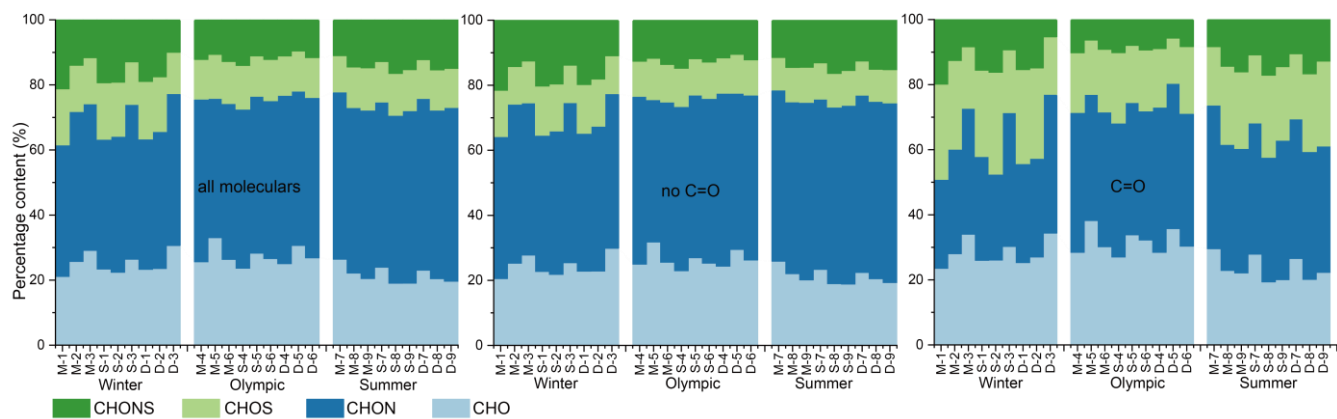




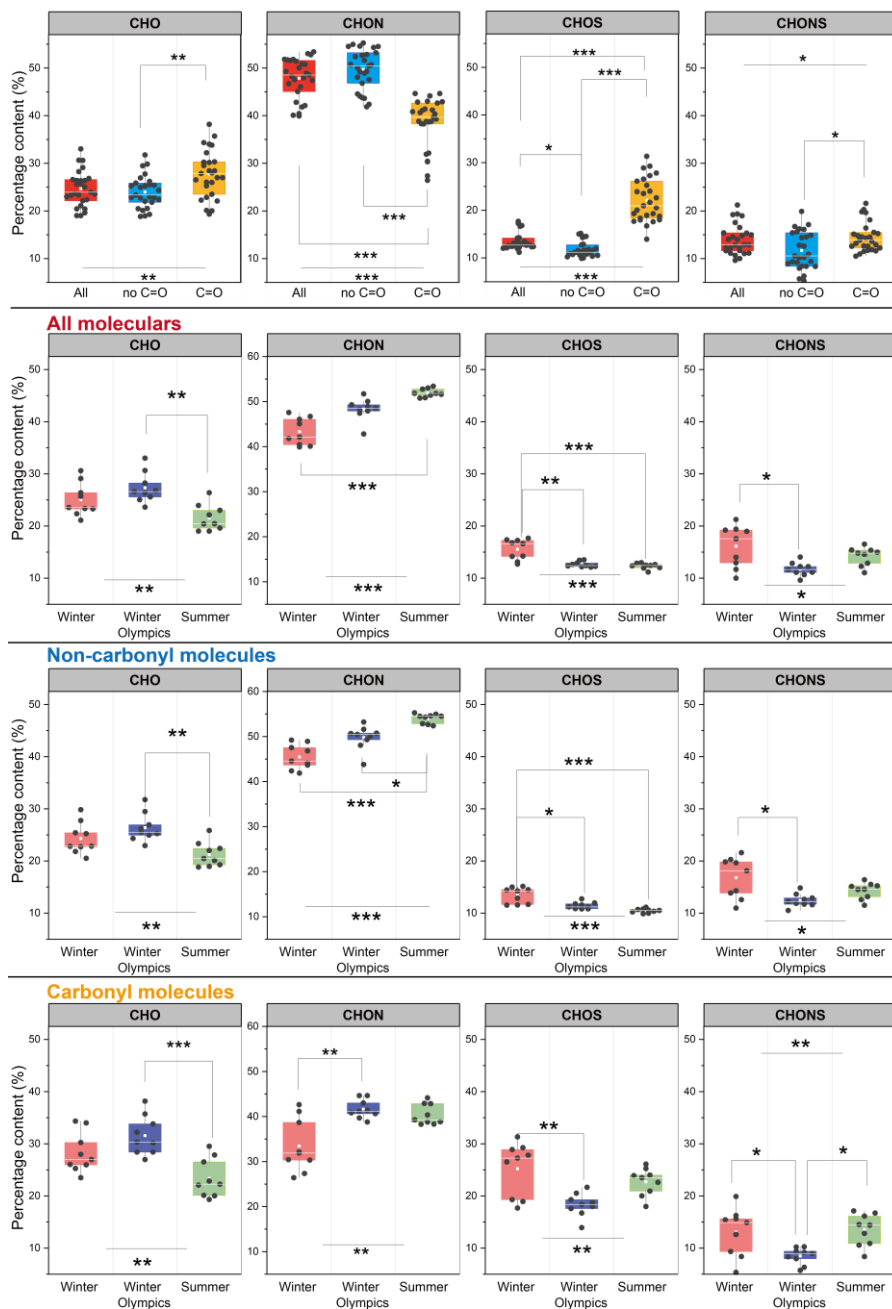
**Fig. S9. The proportion of molecules in five categories was estimated based on molecular numbers. Organic molecules detected in FT-ICP MS were distinguished into five compounds based on H/C, O/C, and AImod. (Group1: polycyclic aromatic-like molecules, Group2: highly aromatic-like molecules, Group3: highly unsaturated-like molecules, Group4: unsaturated aliphatic-like molecules, Group5: saturated-like molecules).**



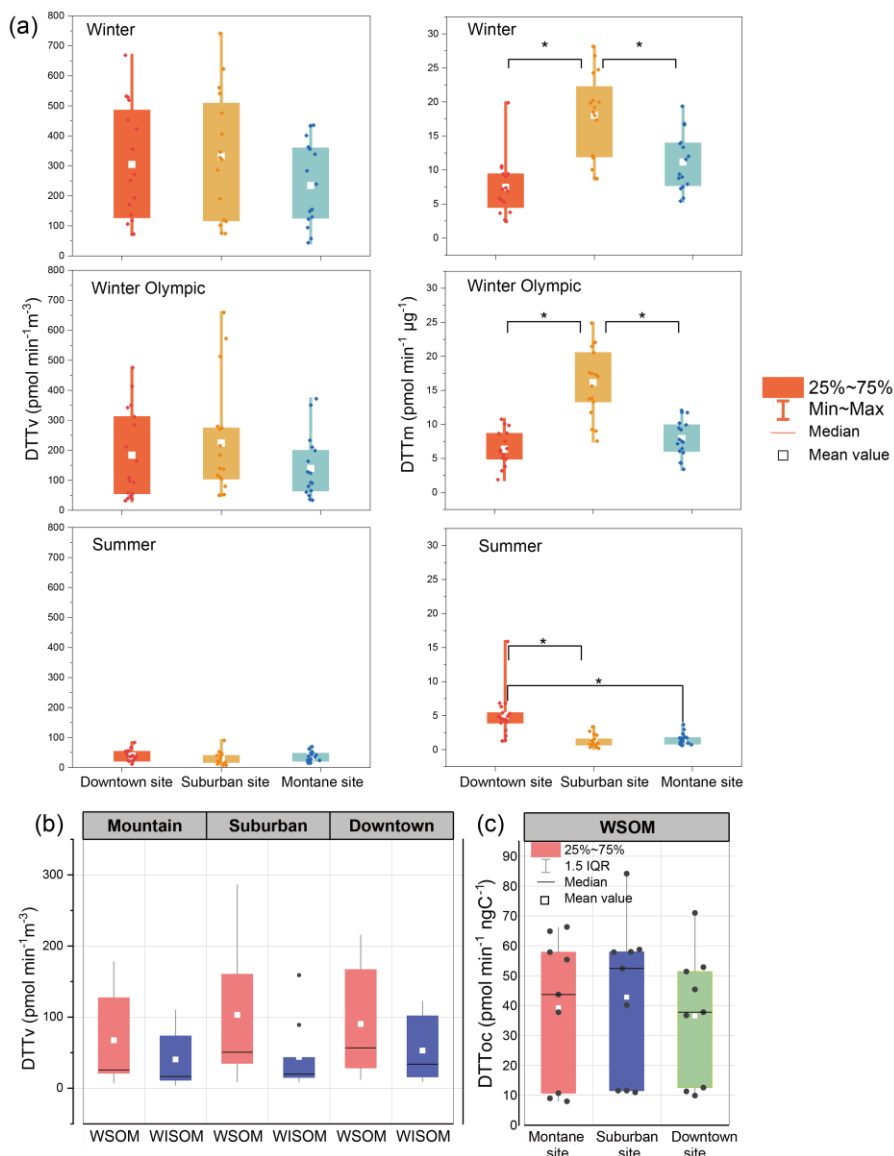
**Fig. S10.** Five categories differ in the percentage distribution of all organic, non-carbonyl, and carbonyl molecules based on the molecular numbers (Kruskal-Wallis tests, Differences between groups were considered statistically significant when  $p < 0.05$ , with  $0.01 < p < 0.05$  marked by \*;  $p < 0.01$  marked by \*\*; and  $p < 0.001$  marked by \*\*\*).



**Fig. S11.** The proportion of each group in all organic, non-carbonyl, and carbonyl molecules was estimated based on the molecular numbers. Organic molecules detected in FT-ICP MS were divided into four groups according to the elemental composition, namely CHO, CHON, CHOS, and CHONS.

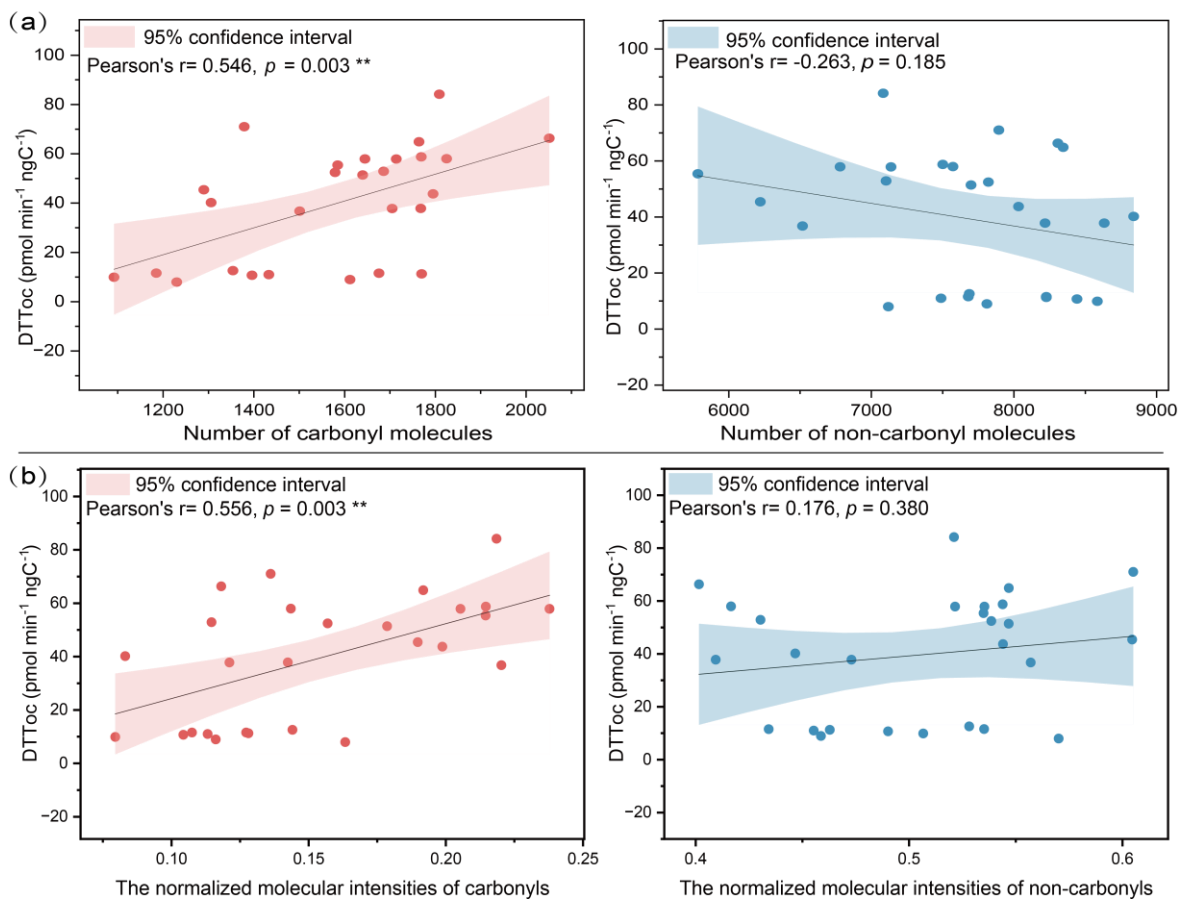


160 Fig. S12. Differences in the percentage distribution of elemental composition based on the molecular numbers of all organic, non-carbonyl, and carbonyl molecules in winter, summer and Winter Olympics (Kruskal-Wallis tests, Differences between groups were considered statistically significant when  $p < 0.05$ , with  $0.01 < p < 0.05$  marked by \*\*;  $p < 0.001$  marked by \*\*\*).

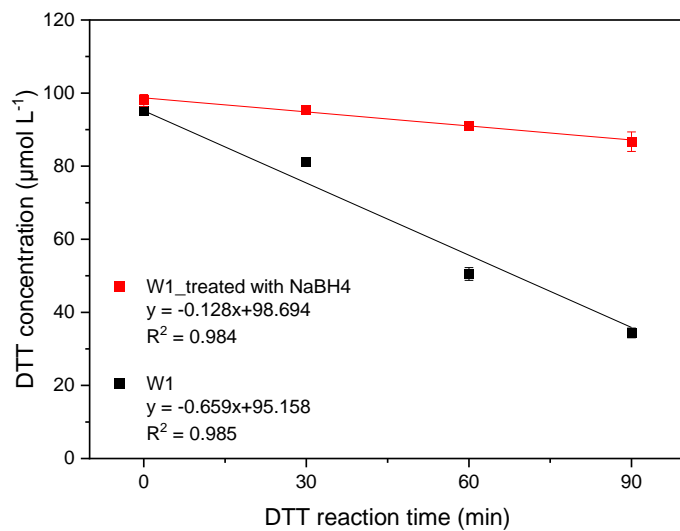


165 Fig. S13. (a) Volume-normalized DTT consumption rate ( $DTT_v$ ) and mass-normalized DTT consumption rate ( $DTT_m$ ) of organic matter in  $PM_{2.5}$  at the downtown site, suburban site, and mountainous site (Kruskal-Wallis tests, Differences between groups were considered statistically significant when  $p < 0.05$ , with  $0.01 < p < 0.05$  marked by \*;  $p < 0.01$  marked by \*\*; and  $p < 0.001$  marked by \*\*\*). (b)  $DTT_v$  of water-soluble organic carbon (WSOM) and organic carbon (WISOM) for representative samples. (c) DTT consumption rate of WSOM normalized by mass concentration of TOC ( $DTT_{oc}$ ).

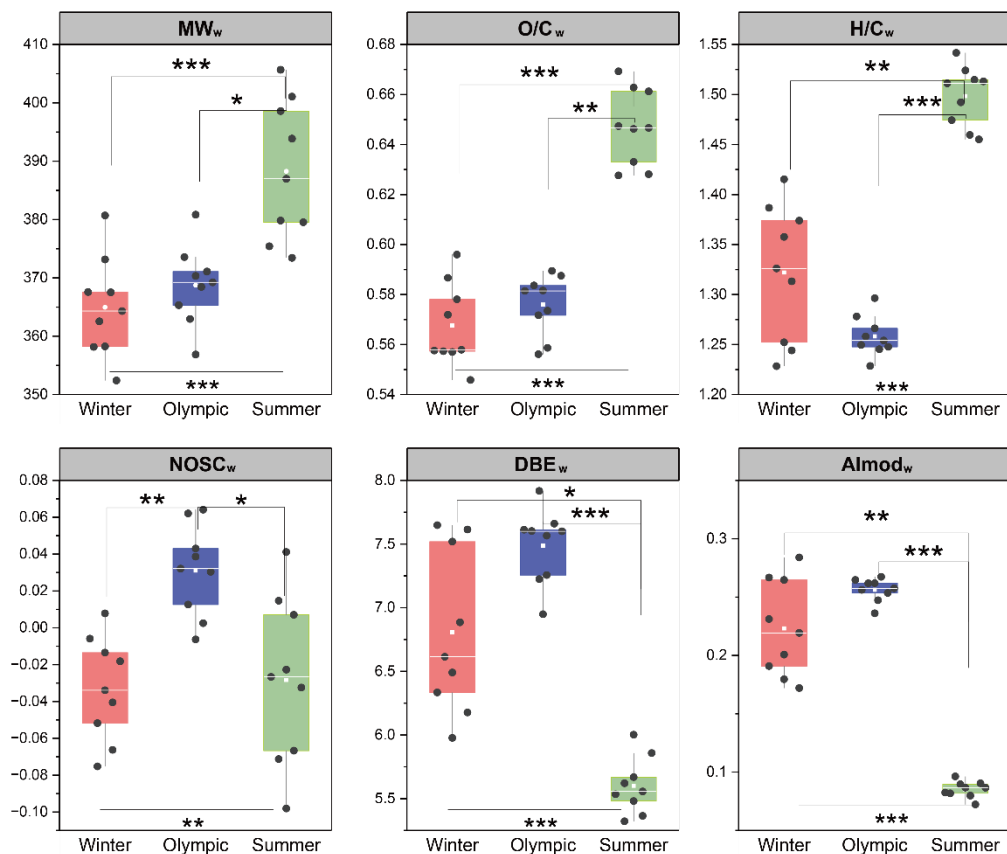
170



**Fig. S14. (a) Correlation analysis between the number of carbonyl and non-carbonyl molecules and  $\text{DTT}_{\text{OC}}$ . (b) Correlation analysis between the normalized molecular intensities of carbonyl and non-carbonyl molecules (sulfur-containing molecules were eliminated) and oxidative potential of WSOM. Sulfur-containing molecules were excluded because of their high ionization efficiency, which could potentially mask the effects of other molecules in intensity calculations (Xie et al., 2022).**

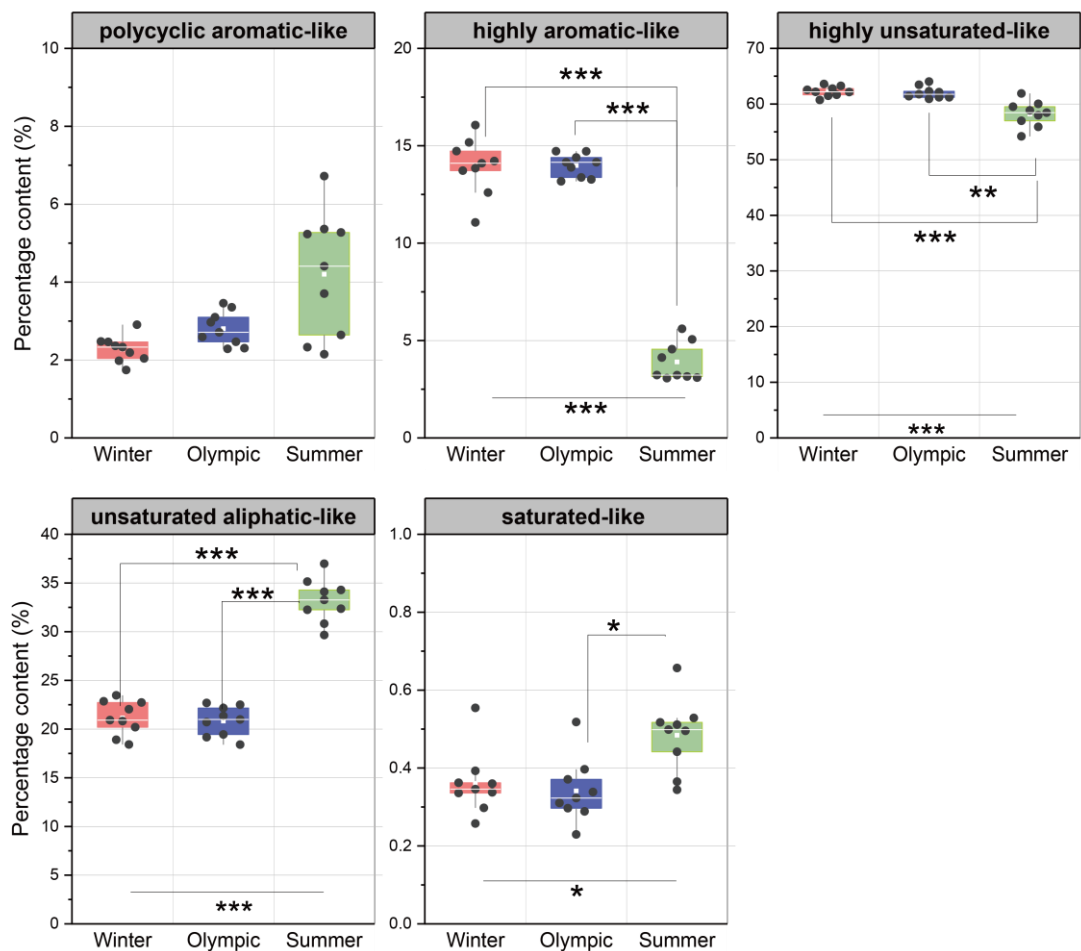


**Fig. S15.** The DTT activity of WSOM from  $\text{PM}_{2.5}$  sample before and after  $\text{NaBH}_4$  treatment was measured.



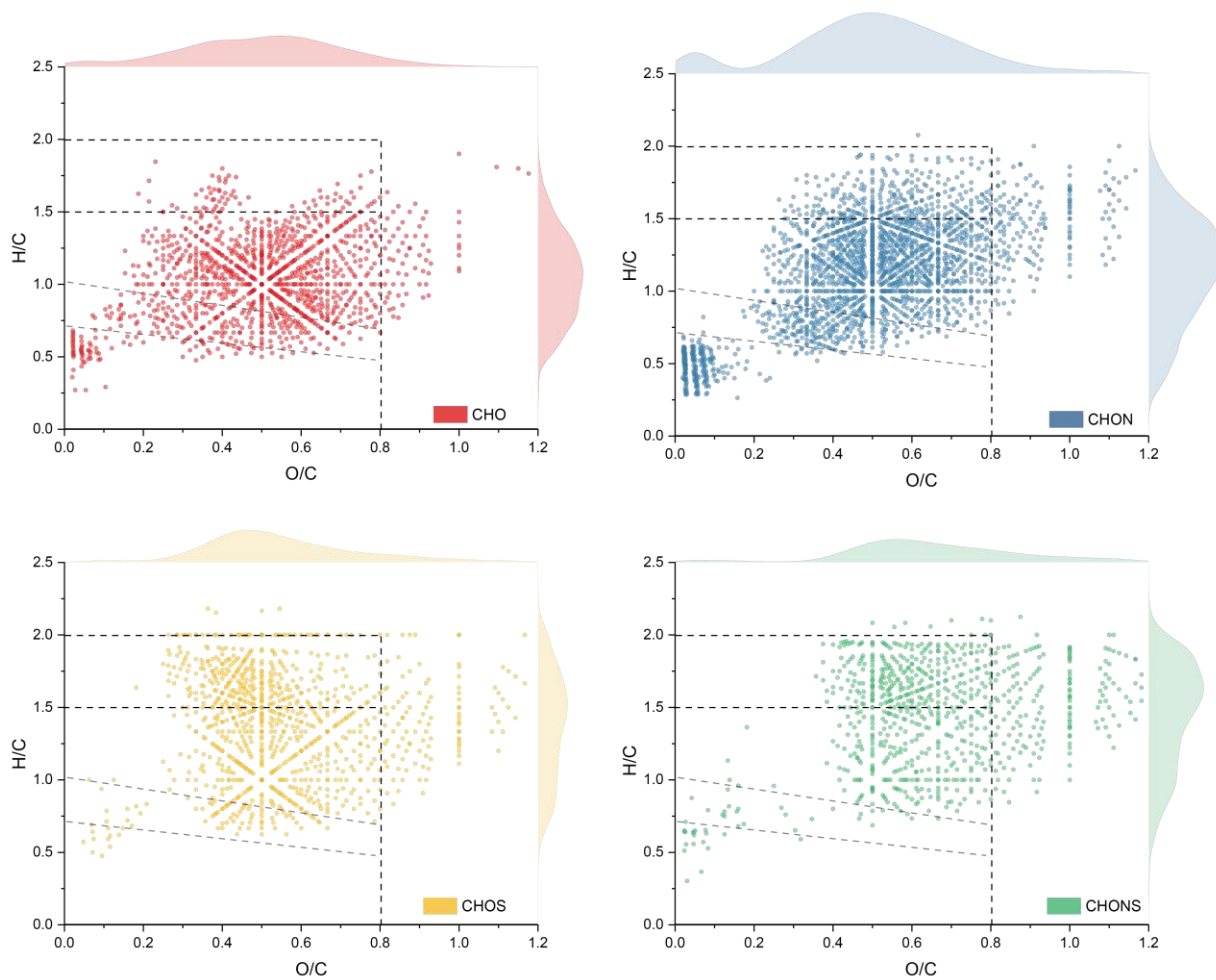
**Fig. S16.** The magnitude-weighted average values of molecular weight (MW<sub>w</sub>), O/C, H/C, the nominal carbon oxidation states (NOSC), double bond equivalence (DBE), and modified aromaticity index (Almod<sub>w</sub>) of carbonyl molecules (Kruskal-Wallis tests, Differences between groups were considered statistically significant when  $p < 0.05$ , with  $0.01 < p < 0.05$  marked by \*;  $p < 0.01$  marked by \*\*; and  $p < 0.001$  marked by \*\*\*).



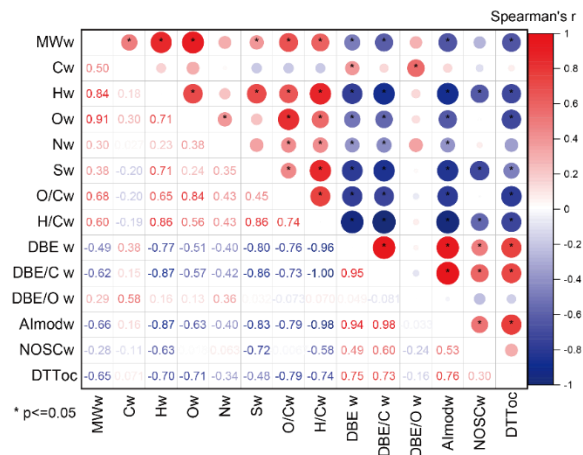


**Fig. S17.** The difference in the percentage distribution based on the molecular numbers of carbonyl molecules in winter, winter Olympics, and summer samples for the five categories (Group1: polycyclic aromatic-like molecules, Group2: highly aromatic-like molecules, Group3: highly unsaturated-like molecules, Group4: unsaturated aliphatic-like molecules, Group5: saturated-like molecules).

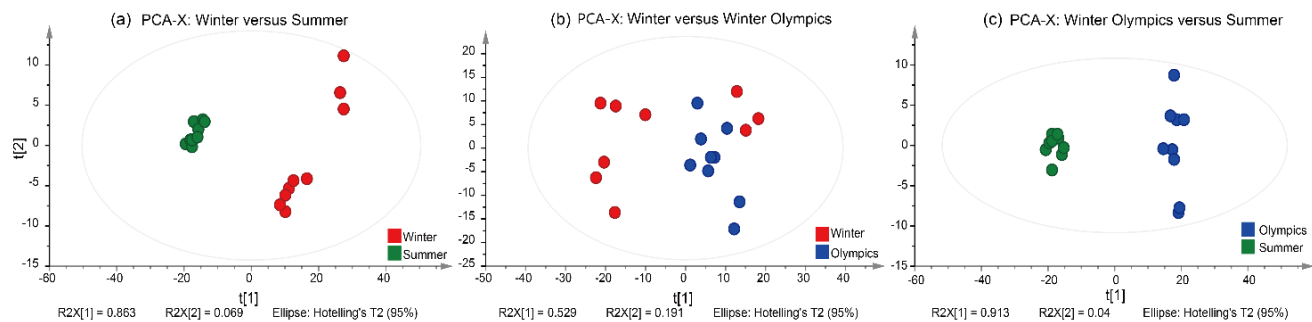
185



190 **Fig. S18.** The van Krevelen diagram of carbonyl molecules from different elemental compositions (Group1: polycyclic aromatic-like molecules, Group2: highly aromatic-like molecules, Group3: highly unsaturated-like molecules, Group4: unsaturated aliphatic-like molecules, Group5: saturated-like molecules).



**Fig. S19. Spearman correlation matrix of carbonyl molecule characteristic parameters and DTT consumption rate of WSOM normalized by mass concentration of TOC (DTT<sub>TOC</sub>). Colors from blue to red represent changes in the Spearman correlation coefficient from -1 to 1.**



**Fig. S20.** OPLS-DA score plots for normalized intensity of individual pro-oxidative carbonyls in (a) winter and winter Olympics, (b) winter and summer, and (c) winter Olympics and summer samples.

## References

- 200 Fu, Q.-L., Fujii, M., and Riedel, T.: Development and comparison of formula assignment algorithms for ultrahigh-resolution mass spectra of natural organic matter, *Analytica Chimica Acta*, 1125, 247–257, <https://doi.org/10.1016/j.aca.2020.05.048>, 2020.
- Koch, B. P., Dittmar, T., Witt, M., and Kattner, G.: Fundamentals of Molecular Formula Assignment to Ultrahigh Resolution Mass Data of Natural Organic Matter, *Anal. Chem.*, 79, 1758–1763, <https://doi.org/10.1021/ac061949s>, 2007.
- 205 Kroll, J. H., Donahue, N. M., Jimenez, J. L., Kessler, S. H., Canagaratna, M. R., Wilson, K. R., Altieri, K. E., Mazzoleni, L. R., Wozniak, A. S., Bluhm, H., Mysak, E. R., Smith, J. D., Kolb, C. E., and Worsnop, D. R.: Carbon oxidation state as a metric for describing the chemistry of atmospheric organic aerosol, *Nat. Chem.*, 3, 133–139, <https://doi.org/10.1038/nchem.948>, 2011.
- 210 Lv, J., Zhang, S., Wang, S., Luo, L., Cao, D., and Christie, P.: Molecular-Scale Investigation with ESI-FT-ICR-MS on Fractionation of Dissolved Organic Matter Induced by Adsorption on Iron Oxyhydroxides, *Environ. Sci. Technol.*, 50, 2328–2336, <https://doi.org/10.1021/acs.est.5b04996>, 2016.
- Phillips, S. M. and Smith, G. D.: Light Absorption by Charge Transfer Complexes in Brown Carbon Aerosols, *Environ. Sci. Technol. Lett.*, 1, 382–386, <https://doi.org/10.1021/ez500263j>, 2014.
- 215 Phillips, S. M. and Smith, G. D.: Further Evidence for Charge Transfer Complexes in Brown Carbon Aerosols from Excitation–Emission Matrix Fluorescence Spectroscopy, *J. Phys. Chem. A*, 119, 4545–4551, <https://doi.org/10.1021/jp510709e>, 2015.
- Wang, L., Lin, Y., Ye, L., Qian, Y., Shi, Y., Xu, K., Ren, H., and Geng, J.: Microbial Roles in Dissolved Organic Matter Transformation in Full-Scale Wastewater Treatment Processes Revealed by Reactomics and Comparative Genomics, *Environ. Sci. Technol.*, 55, 11294–11307, <https://doi.org/10.1021/acs.est.1c02584>, 2021.
- 220 Xie, Q., Su, S., Dai, Y., Hu, W., Yue, S., Cao, D., Jiang, G., and Fu, P.: Deciphering <sup>13</sup>C and <sup>34</sup>S Isotopes of Organosulfates in Urban Aerosols by FT-ICR Mass Spectrometry, *Environ. Sci. Technol. Lett.*, 9, 526–532, <https://doi.org/10.1021/acs.estlett.2c00255>, 2022.
- 225 Yu, S., Tang, S., Lv, J., Li, F., Huang, Z., Zhao, L., Cao, D., and Wang, Y.: High throughput identification of carbonyl compounds in natural organic matter by directional derivatization combined with ultra-high resolution mass spectrometry, *Water Research*, 258, 121769, <https://doi.org/10.1016/j.watres.2024.121769>, 2024.
- Zhang, H. and Fu, S.: Highly efficient catalytic transfer hydrogenolysis for the conversion of Kraft lignin into bio-oil over heteropoly acids, *Green Chemistry*, 24, 6619–6630, <https://doi.org/10.1039/D2GC01808J>, 2022.
- 230 Zhang, Q., Ma, H., Li, J., Jiang, H., Chen, W., Wan, C., Jiang, B., Dong, G., Zeng, X., Chen, D., Lu, S., You, J., Yu, Z., Wang, X., and Zhang, G.: Nitroaromatic Compounds from Secondary Nitrate Formation and Biomass Burning Are Major Proinflammatory Components in Organic Aerosols in Guangzhou: A Bioassay Combining High-Resolution Mass Spectrometry Analysis, *Environ. Sci. Technol.*, <https://doi.org/10.1021/acs.est.3c04983>, 2023.