Supporting Information for

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

Feifei Li[1,2], Shanshan Tang[3], Jitao Lv[1,2], Shiyang Yu[1,2], Xu Sun[1,4], Dong Cao[1], Yawei Wang[1,2,5], Guibin Jiang[1,2,5]

[1]State Key Laboratory of Environmental Chemistry and Eco-toxicology, Research Center for Eco-environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

[2]College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China

[3]Collaborative Center for Physics and Chemistry, Institute of International Innovation, Beihang University, Yuhang District, Hangzhou 311115, China

[4]Beijing Urban Ecosystem Research Station, State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

[5]School of Environment, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou 310024, China

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

List of supporting information

Texts

Tables

Figures

Text S1: Chemicals

Ultrapure water (18.2 MΩ·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, ≥ 99%, CAS 57981-02-9) obtained from Sigma-Aldrich and Sodium borohydride ($NaBH_4$, ≥ 98%, CAS 16940-66-2) obtained from Sinopharm Chemical Reagent Co. LTD.

Dithiothreitol (DTT, ≥ 99.4%, Biotechnology grade, Amresco), Diethylenetriaminepentaacetic acid (DTPA, ≥ 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 $PM_{2.5}$ samples.

Text S2: Solid phase extraction procedure

The water-soluble fraction of $PM_{2.5}$ 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, and 20 mL were passed through the cartridges by gravity at a flow rate of approximately 2 mL min^{-1}. Cartridges were rinsed with three volumes of 0.01 M HCl for removal of salts, dried with a stream of N_2, 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_2 to 1 mL and then stored at -4 °C in the dark.

Text S3: ESI-FT-ICR MS analysis

Samples for ESI FT-ICR-MS analysis were continuously infused into the ESI unit by syringe infusion at a flow rate of 120 μL h^{-1}. 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 hexapol 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 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:

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

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

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

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

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:

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

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)

Text S4: 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 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 (Figure 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.

Text S5: Supplement statistical analysis

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 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 (Figure S4).

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 (Figure S5). Further multivariate OPLS-DA (1 + 1 + 0) models explained the seasonal differences in the abundance of pro-oxidative carbonyls (Figure S17), and the model was validated by permutation tests (n=200) (Figure S6).

Table S1: Detailed information on representative $PM_{2.5}$ samples.

Sampling site	Sampling number	Sampling season	Data		TOC (mg/L)	WSOM		WISOM	Number of carbonyl molecules	Number of non-carbonyl molecules	Number of all molecules
						DTT_m (pmol/min/ngC)	DTT_v (pmol/min/m^3)	DTT_v (pmol/min/m^3)			
Mountain site	M-1	Winter	19 Jan. 2022	20 Jan. 2022	116.40	66.33	150.26	110.51	2051	8308	10359
	M-2	Winter	23 Jan. 2022	24 Jan. 2022	323.20	37.81	178.37	104.45	1768	8633	10401
	M-3	Winter	31 Jan. 2022	1 Feb. 2022	42.94	55.41	17.36	16.58	1585	5784	7369
	M-4	Olympics	9 Feb. 2022	10 Feb. 2022	199.10	43.72	127.06	73.52	1795	8032	9827
	M-5	Olympics	6 Feb. 2022	14 Feb. 2022	61.17	57.89	25.84	14.13	1645	6781	8426
	M-6	Olympics	8 Feb. 2022	18 Feb. 2022	115.27	64.88	54.58	22.03	1764	8345	10109
	M-7	Summer	10 Jul. 2022	27 Jul. 2022	130.07	7.96	7.40	3.57	1230	7119	8349
	M-8	Summer	16 Jul. 2022	24 Jul. 2022	166.13	10.70	25.40	8.61	1396	8442	9838
	M-9	Summer	11 Jul. 2022	20 Jul. 2022	166.33	8.97	21.31	11.58	1612	7810	9422
Suburban site	S-1	Winter	19 Jan. 2022	20 Jan. 2022	189.30	57.98	160.19	29.81	1825	7572	9397
	S-2	Winter	23 Jan. 2022	24 Jan. 2022	488.70	40.18	286.58	159.05	1306	8839	10145
	S-3	Winter	31 Jan. 2022	1 Feb. 2022	41.30	84.16	50.73	14.40	1809	7083	8892
	S-4	Olympics	9 Feb. 2022	10 Feb. 2022	328.10	52.44	251.07	89.09	1579	7821	9400
	S-5	Olympics	6 Feb. 2022	14 Feb. 2022	82.57	58.75	35.40	17.16	1769	7501	9270
	S-6	Olympics	8 Feb. 2022	18 Feb. 2022	93.19	57.88	78.72	43.25	1714	7138	8852
	S-7	Summer	10 Jul. 2022	27 Jul. 2022	112.07	11.57	9.26	8.28	1185	7680	8865
	S-8	Summer	16 Jul. 2022	24 Jul. 2022	213.23	11.54	35.14	19.89	1676	8227	9903
	S-9	Summer	11 Jul. 2022	20 Jul. 2022	136.73	10.98	21.79	15.27	1433	7489	8922
Downtown site	D-1	Winter	19 Jan. 2022	20 Jan. 2022	162.20	52.87	166.87	101.65	1686	7104	8790
	D-2	Winter	23 Jan. 2022	24 Jan. 2022	391.30	37.80	215.90	116.35	1705	8218	9923
	D-3	Winter	31 Jan. 2022	1 Feb. 2022	85.70	45.42	56.81	33.85	1290	6222	7512
	D-4	Olympics	9 Feb. 2022	10 Feb. 2022	263.40	51.40	197.59	122.90	1640	7699	9339
	D-5	Olympics	6 Feb. 2022	14 Feb. 2022	106.87	36.76	28.67	16.19	1501	6518	8019
	D-6	Olympics	8 Feb. 2022	18 Feb. 2022	74.79	71.00	77.50	42.01	1379	7894	9273
	D-7	Summer	10 Jul. 2022	27 Jul. 2022	134.07	12.57	12.04	8.94	1354	7687	9041
	D-8	Summer	16 Jul. 2022	24 Jul. 2022	230.73	11.29	37.21	20.04	1770	8227	9997
	D-9	Summer	11 Jul. 2022	20 Jul. 2022	306.97	9.90	21.72	15.33	1092	8585	9677

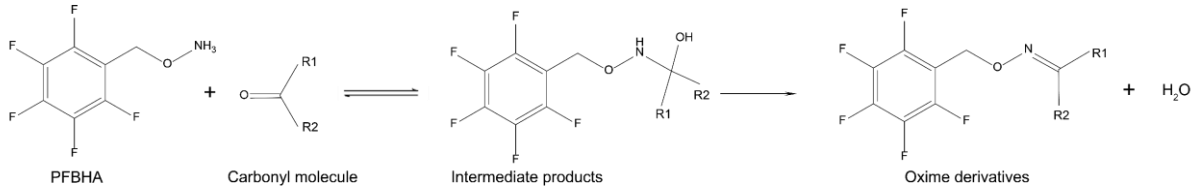

Figure S1. Derivatization reaction scheme of PFBHA and carbonyl molecules.

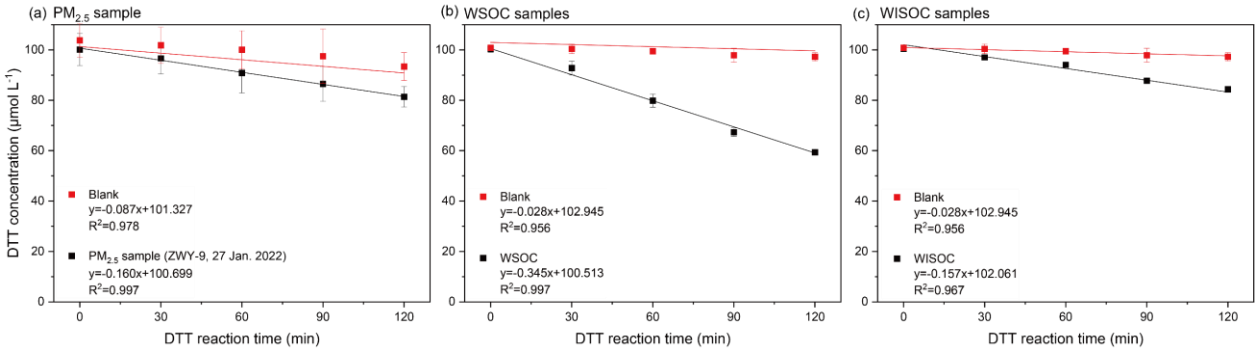

Figure S2. Reproducibility of DTT measurements of samples. (a) DTT measurements of $PM_{2.5}$ samples and backgrounds, using the downtown site on January 27, 2022, as an example. (b) DTT measurements of a water-soluble organic carbon (WSOM) quality control sample, which was prepared containing a mixture 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.

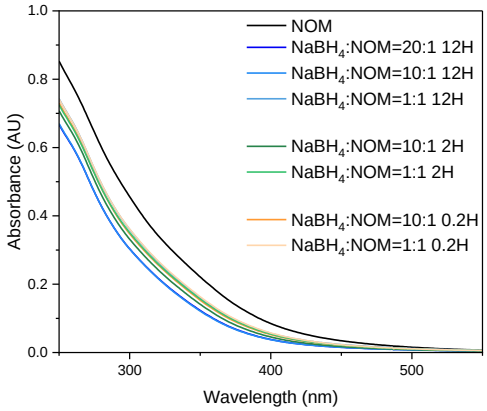

Figure S3. UV-Vis spectra during SRNOM reduction by $NaBH_4$.

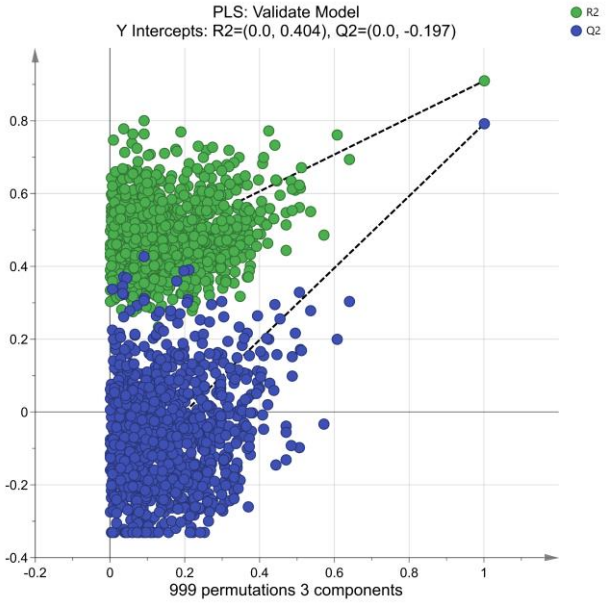

Figure S4. Validation plot of PLSR model for carbonyl molecules significantly positively correlated with DTT_{OC}, n = 999, R2 = (0.0, 0.404), Q2 = (0.0, -0.197).

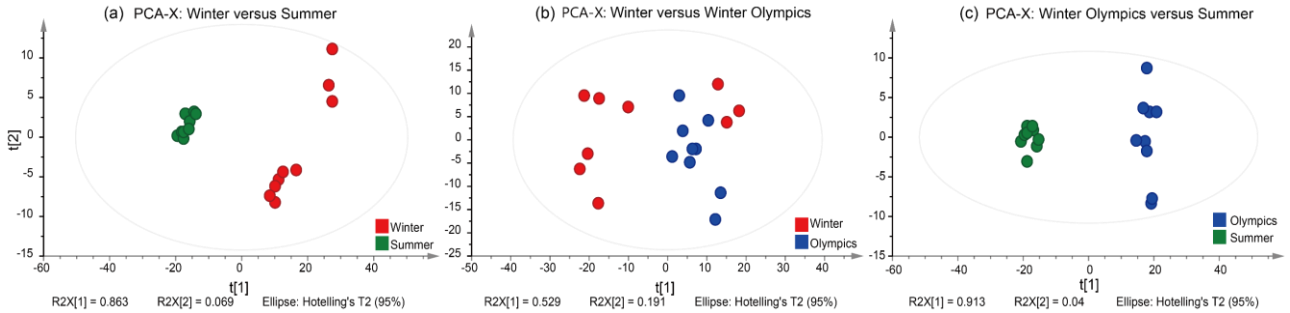

Figure S5. PCA score plots for pro-oxidative carbonyls in (a) winter versus summer, (b) winter versus winter Olympics, and (c) winter Olympics versus summer samples.

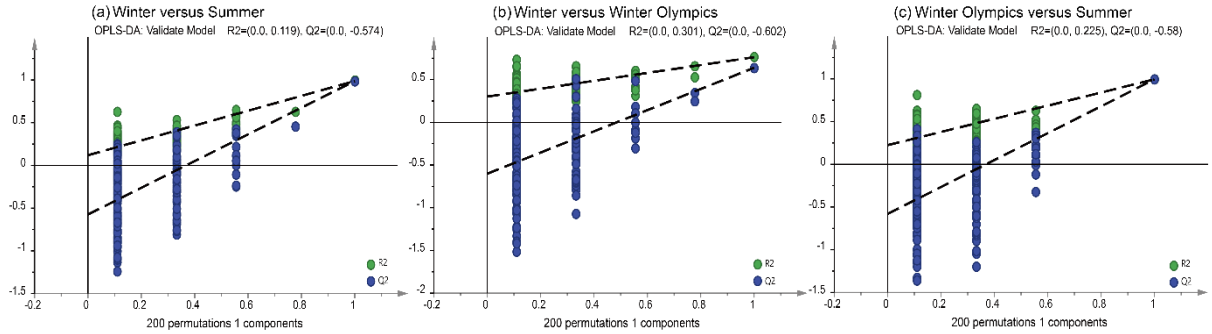

Figure S6. Multivariate OPLS-DA (1 + 1 + 0) models for the pro-oxidative carbonyls were validated by permutation test (n = 200). (a) Validation plot of OPLS-DA model obtained from carbonyl molecules in winter and summer samples, R2 = (0.0, 0.019), Q2 = (0.0, -0.574); (b) Validation plot of OPLS-DA model obtained from carbonyl molecules in winter and Winter Olympic samples, R2 = (0.0, 0.301), Q2 = (0.0, -0.602); (c) Validation plot of OPLS-DA model obtained from carbonyl molecules in winter Olympics and summer samples, R2 = (0.0, 0.225), Q2 = (0.0, -0.58).

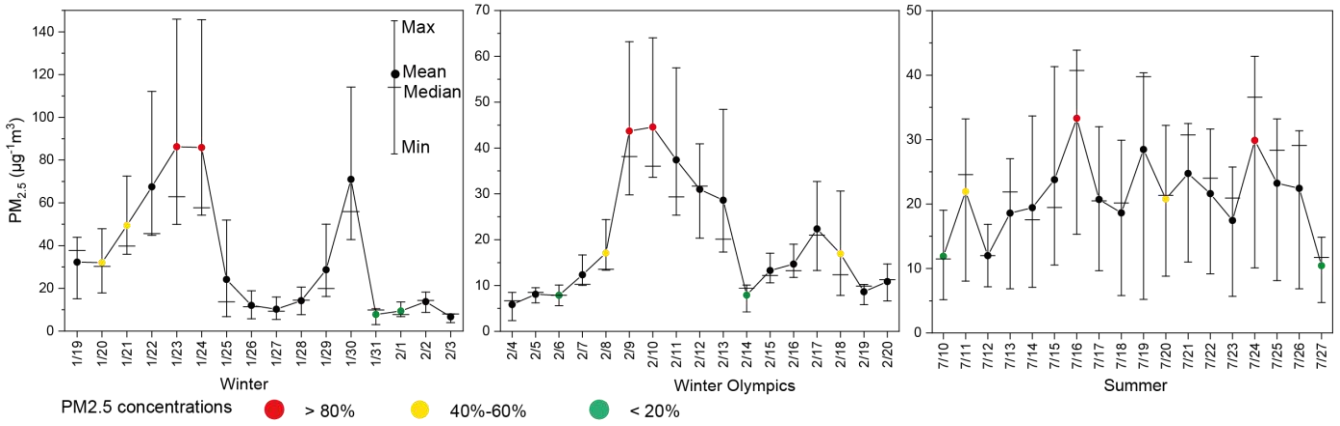

Figure S7. Daily $PM_{2.5}$ concentration at the downtown site, suburban site, and montane site in Beijing during each sampling season. The red dots represent the day on which $PM_{2.5}$ concentrations exceeded 80% of that sampling season. The yellow dots represent the day on which $PM_{2.5}$ concentrations are at 40-60% of that sampling season. The green dots represent the day on which $PM_{2.5}$ concentrations are below 20% of that sampling season.

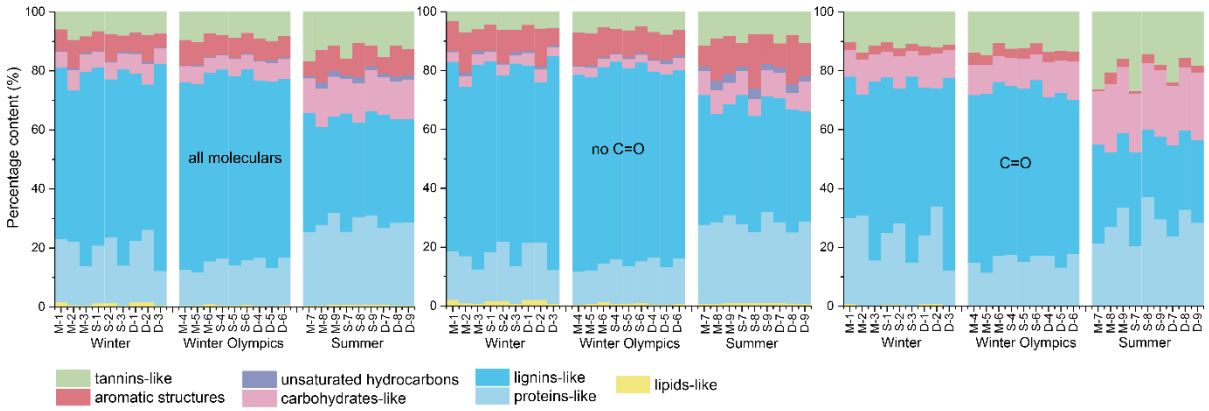

Figure S8. The proportion of molecules in seven compounds was estimated based on the normalized molecular intensities. Organic molecules detected in FT-ICP MS were distinguished into seven compounds based on H/C and O/C. (Group1: lipids-like, Group2: proteins-like, Group3: lignins-like, Group4: carbohydrates-like, Group5: unsaturated hydrocarbons, Group6: aromatic structures, Group7: tannins-like).

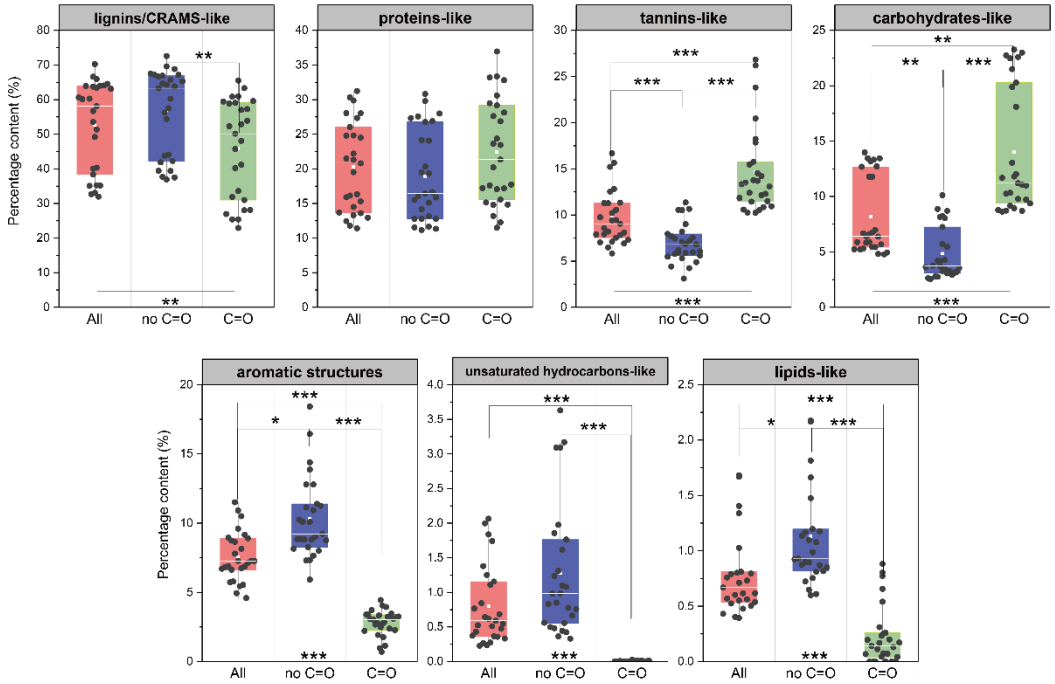

Figure S9. Seven compounds differ in the percentage distribution of all organic, non-carbonyl, and carbonyl molecules based on the normalized molecular intensities (Kruskal-Wallis tests, Differences between groups were considered statistically significant when $p < 0.05$, with $0.01 < p < 0.05$ marked by *; $p < 0.01$ marked by **; and $p < 0.001$ marked by *).**

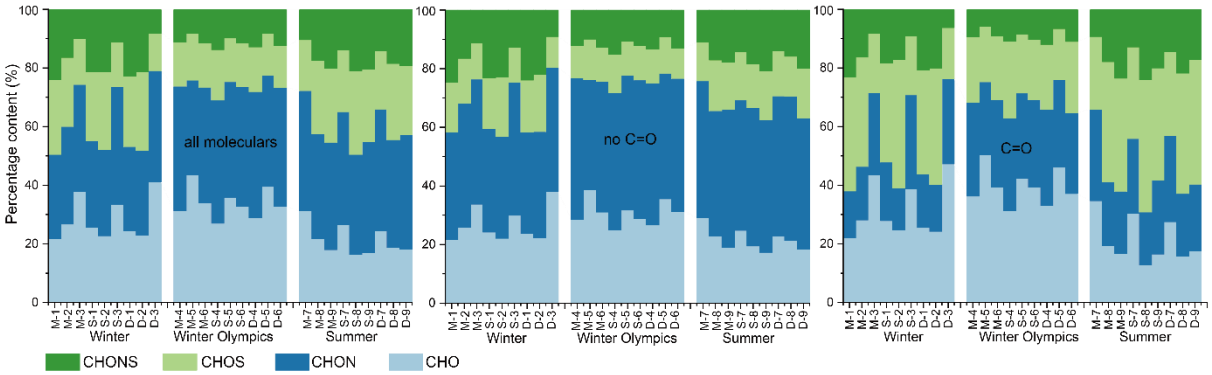

Figure S10. The proportion of each group in all organic, non-carbonyl, and carbonyl molecules was estimated based on the normalized molecular intensities. Organic molecules detected in FT-ICP MS were divided into four groups according to the elemental composition, namely CHO, CHON, CHOS, and CHONS.

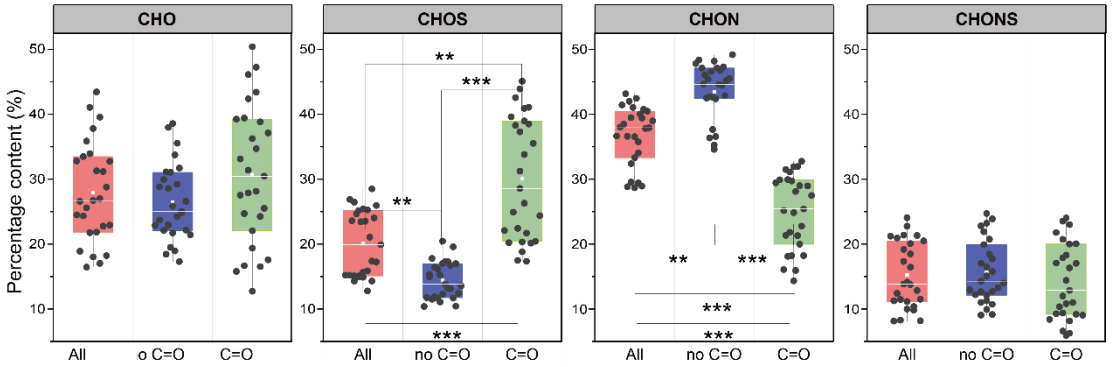

Figure S11. Differences in the percentage distribution of elemental composition based on the normalized molecular intensities in all organic, non-carbonyl, and carbonyl molecules (Kruskal-Wallis tests, Differences between groups were considered statistically significant when $p < 0.05$, with $0.01 < p < 0.05$ marked by *; $p < 0.01$ marked by **; and $p < 0.001$ marked by *).**

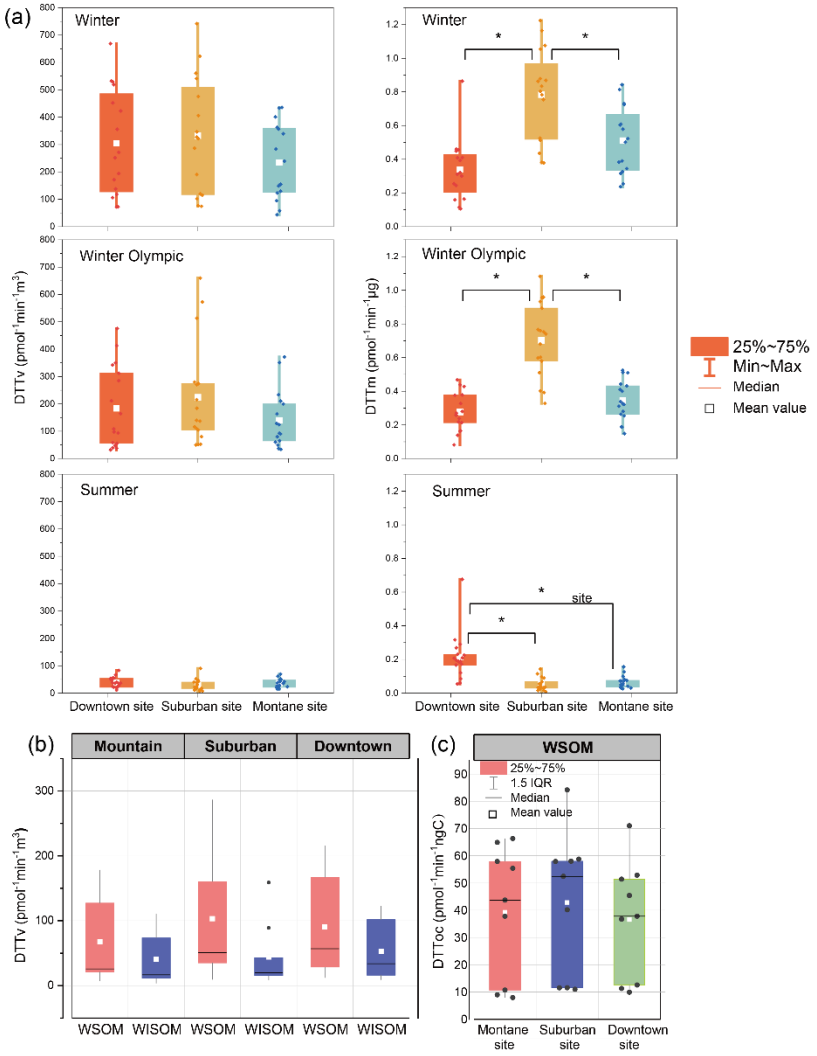

Figure S12. (a) Volume-normalized DTT consumption rate (DTT_v) and mass-normalized DTT consumption rate (DTT_m) of organic matter in $PM_{2.5}$ at the downtown site, suburban site, and montane site (Kruskal-Wallis tests, Differences between groups were considered statistically significant when $p < 0.05$, 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}).**

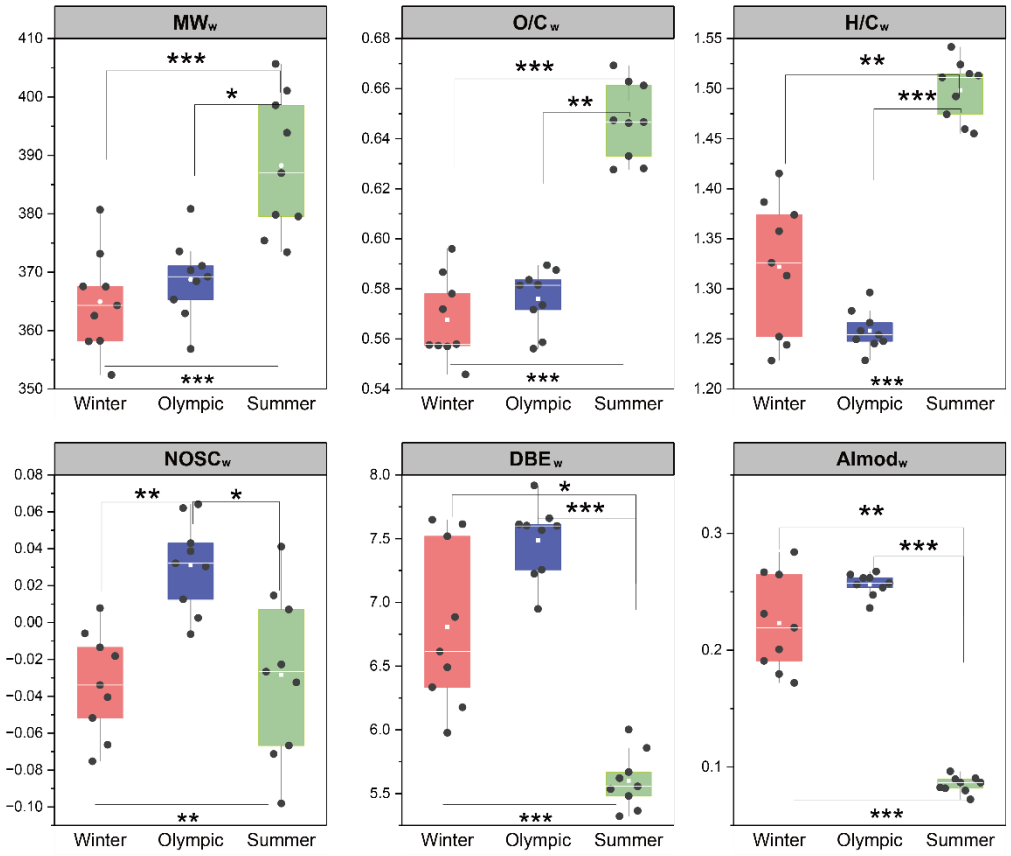

Figure S13. The magnitude-weighted average values of molecular weight (MWw), O/C, H/C, the nominal carbon oxidation states (NOSC), double bond equivalence (DBE), and modified aromaticity index (AImod) of carbonyl molecules (Kruskal-Wallis tests, Differences between groups were considered statistically significant when $p < 0.05$, with $0.01 < p < 0.05$ marked by *; $p < 0.01$ marked by **; and $p < 0.001$ marked by *).**

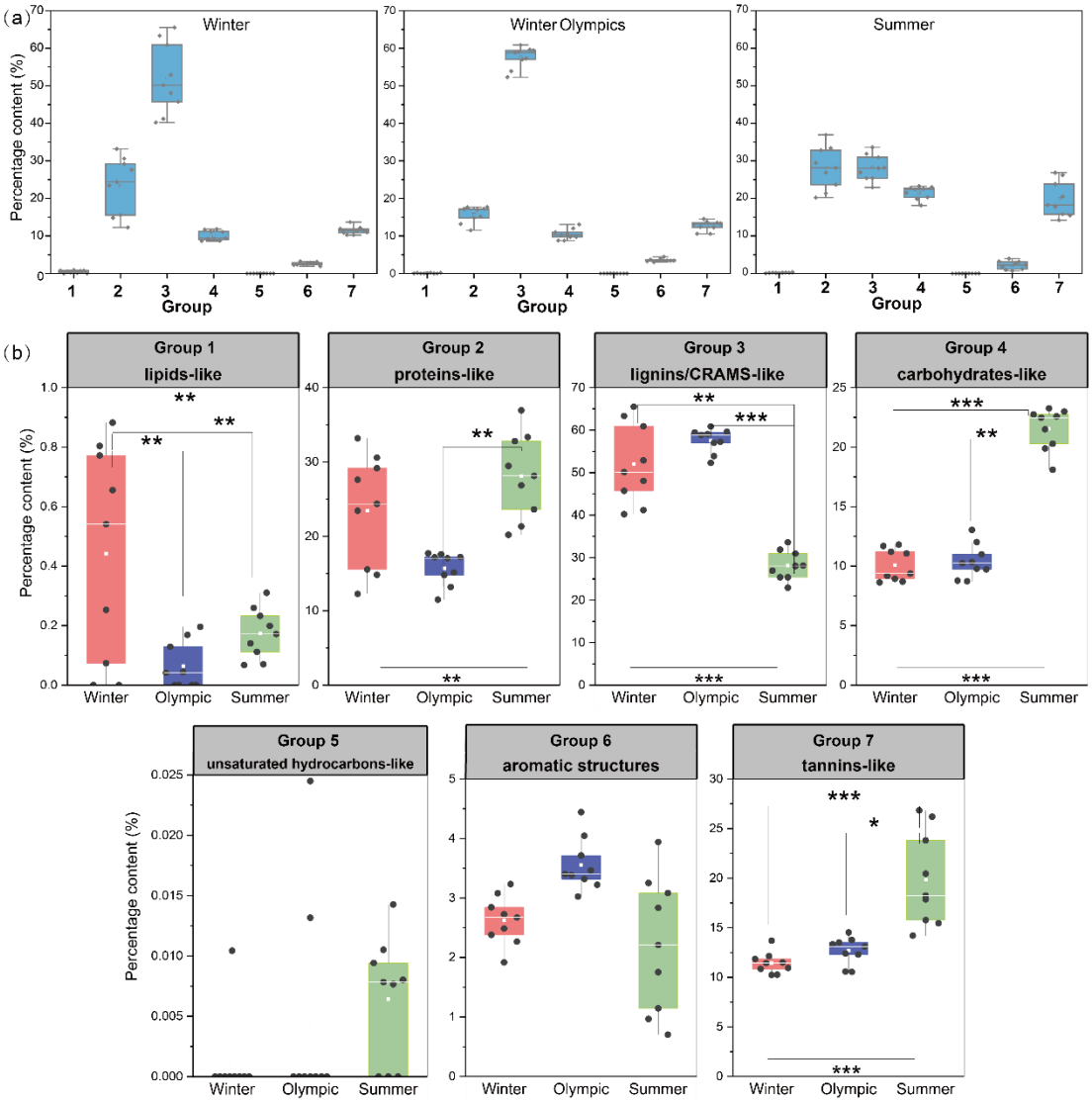

Figure S14. The difference in the percentage distribution based on the normalized molecular intensities of carbonyl molecules in winter, winter Olympics, and summer samples for the seven compound groups (Group1: lipids-like, Group2: proteins-like, Group3: lignins/CRAMs-like, Group4: carbohydrates-like, Group5: unsaturated hydrocarbons, Group6: aromatic structures, Group7: tannins-like).

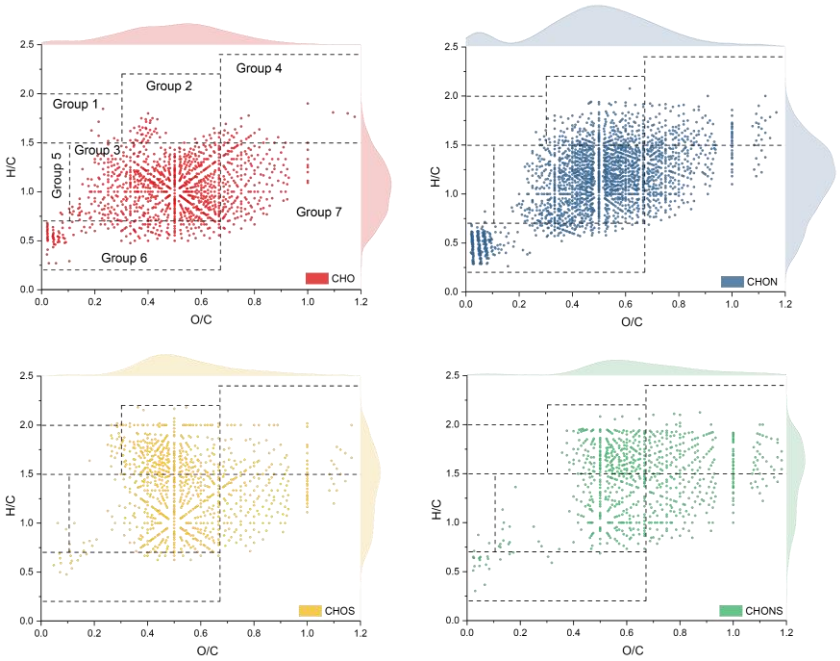

Figure S15. The van Krevelen diagram of carbonyl molecules from different elemental compositions (Group1: lipids-like, Group2: proteins-like, Group3: lignins/CRAMs-like, Group4: carbohydrates-like, Group5: unsaturated hydrocarbons, Group6: aromatic structures, Group7: tannins-like).

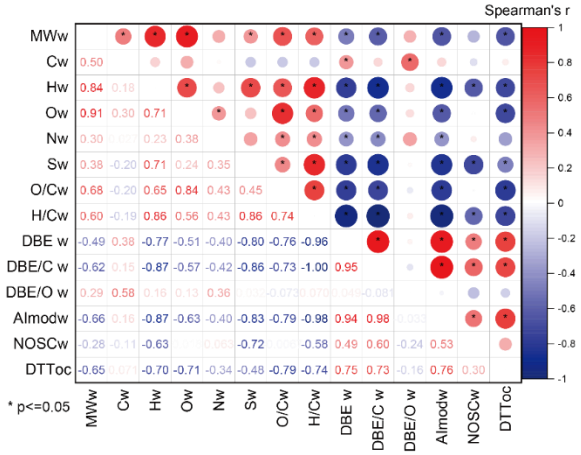

Figure S16. Spearman correlation matrix of carbonyl molecule characteristic parameters and DTT consumption rate of WSOM normalized by mass concentration of TOC (DTT_{OC}). Colors from blue to red represent changes in the Spearman correlation coefficient from −1 to 1.

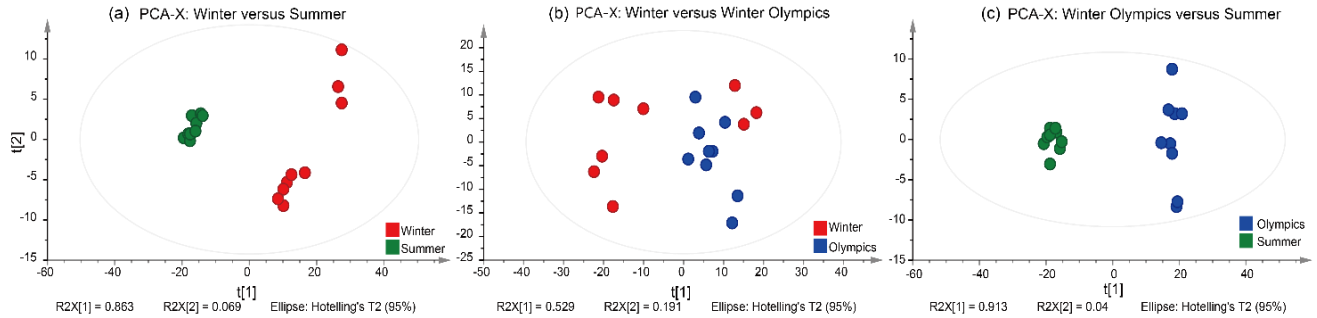

Figure S17. OPLS-DA score plots for normalized intensity of individual pro-oxidative carbonyls in (a) winter and winter Olympics, (b) winter and summer, and (c) winter Olympics and summer samples.

References

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.

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.

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.

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.