S1 L2MS spectra

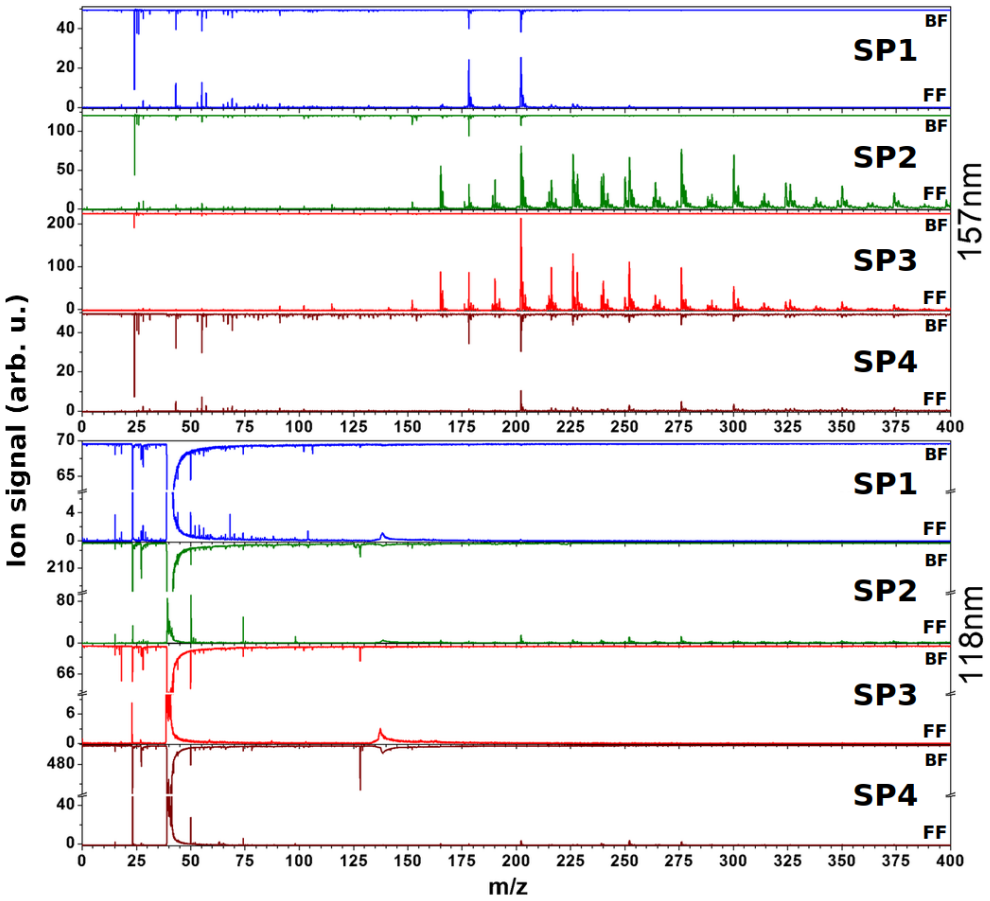

Figure S1. Comparison between mass spectra for SP1, SP2, SP3, and SP4 samples recorded with λ_i = 157 nm and λ_i = 118 nm for Front Filters (lower spectra) and Back Filters (upper spectra).

S2 Details of PCA statistical method

Before PCA is performed, mass spectra are baseline corrected and all the peaks that are coming from the substrate or are present in the analysis chamber (residual gas) are identified and removed. All the spectra were later normalized to the partial ion count (PIC) corresponding to the total remaining signal. The areas of the peaks, rather than their intensities, were used in order to take into account the full contribution of a certain molecule at their specific *m/z*, regardless of the mass resolution.

Scree plots representing the cumulative contribution of principal components are derived from the PCA. These are presented for each wavelength in Figure S2. These data allow us to choose the number of components to retain for further analysis while still preserving most of the information about the dataset. In this case, for the data obtained with 266 nm ionization, the first three principal components account for more than 91 % of the variation in the dataset (PC1 ~59 %, PC2 ~20 %, PC3 ~12 %) , therefore in the analyzes presented here, the data can be treated as being only three-dimensional. However, the information regarding the difference between the Front and Back Filters can be derived from the first two components alone (almost 80 %), therefore from now on only the first two components will be considered. The same can be said about the other two ionization wavelengths: in the case of 157 nm ionization, the first two components account for ~76 % of the variation (PC1 ~55 %, PC2 ~21 %) and in the case of 118 nm ionization, ~72 % (PC1 ~53 %, PC2 ~19 %).

The contribution of every mass peak in the covariance matrix to the principal components, in the form of loadings, is plotted in order to aid the interpretation of their physical meaning (Figure S3). Although all the detected peaks were used in this analysis, for visualization purposes, peaks with little to no contribution are not displayed in these plots.

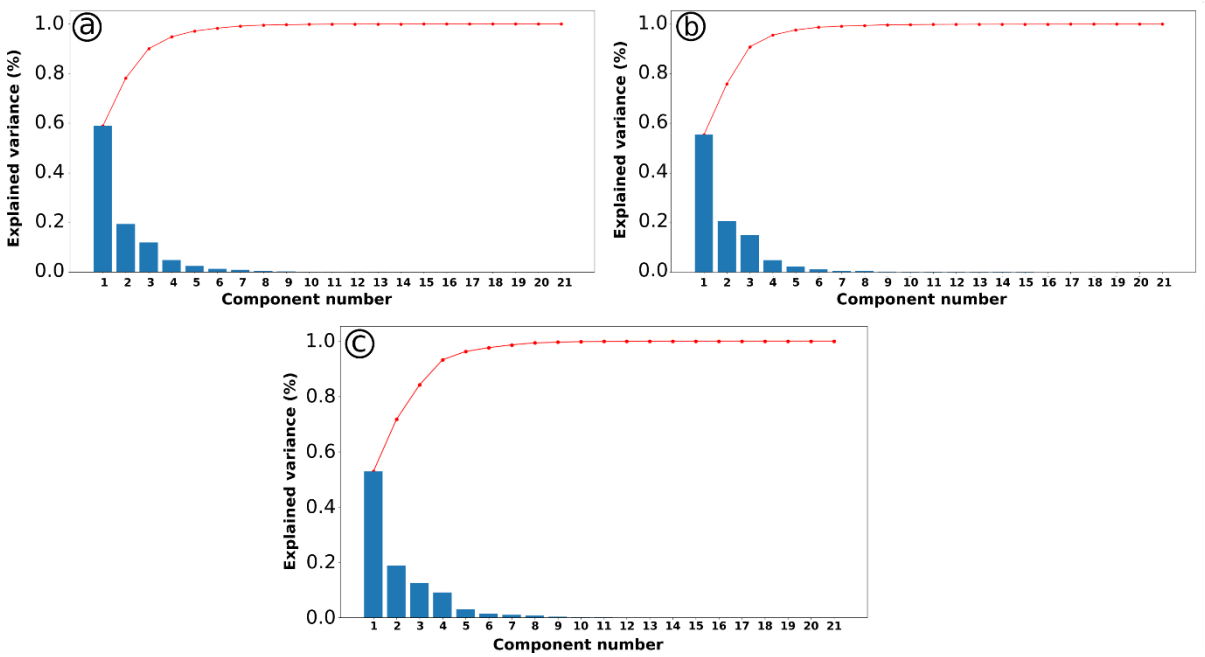

Figure S2. Scree plots representing the cumulative contribution of principal components for mass spectra obtained by L2MS at ionization wavelengths: (a) 266 nm, (b) 157 nm, and (c) 118 nm.

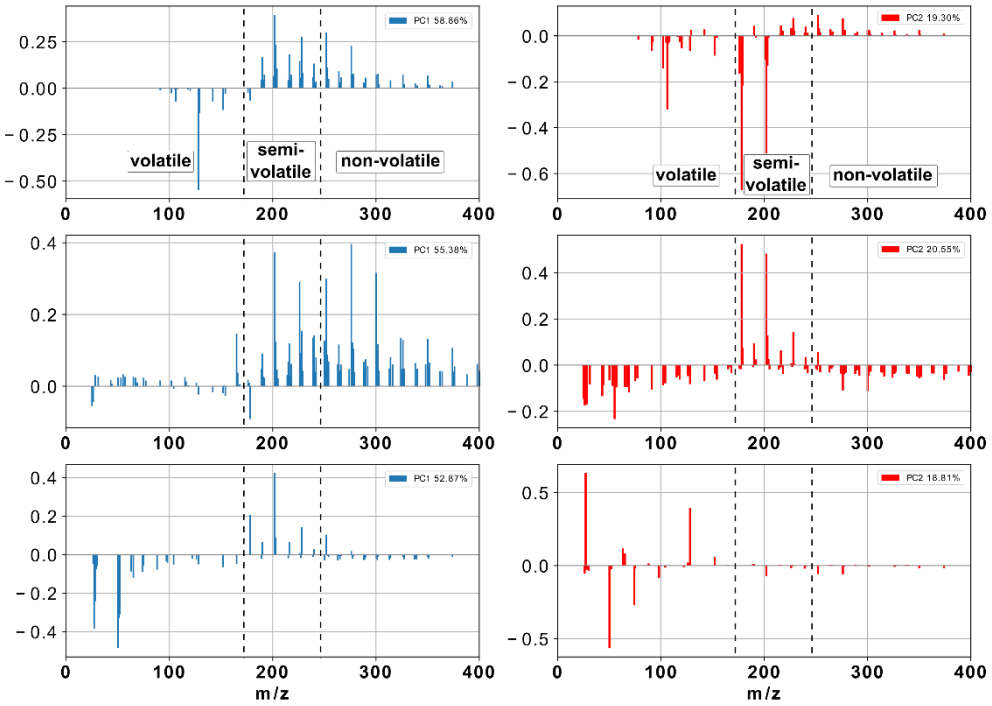

Figure S3. Loadings corresponding to the contribution of various species to PC1 (blue) and PC2 (red) derived from L2MS spectra at ionization wavelengths: (a) 266 nm (upper row), (b) 157 nm (middle row), and (c) 118 nm (lower row). Dashed lines indicate the limits of the three aromatic categories defined in the main text: non-volatile ($m/z \geq 252$), semi-volatile (m/z 176–242), and volatile (m/z 78–166) fractions.

S3 SIMS spectra and details of the PCA

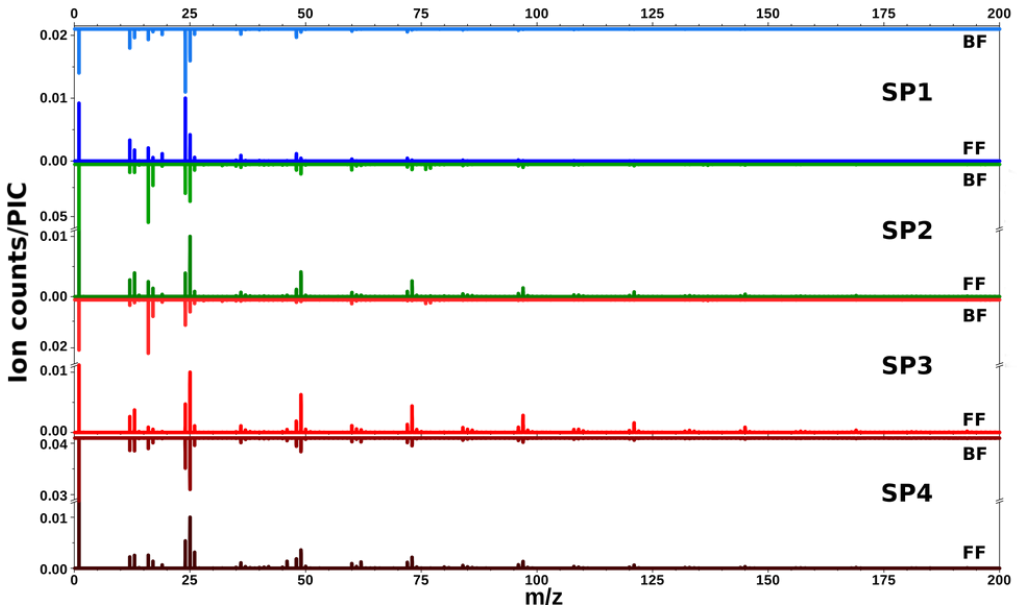

Figure S4. ToF-SIMS mass spectra of samples SP1, SP2, SP3, and SP4 obtained in negative polarity for Front Filters (lower spectra) and Back Filters (upper spectra).

When applying PCA, 89 mass peaks including all PAHs and their fragments up to *m/z* 300 were selected in the positive polarity mass spectra. As the ToF-SIMS analysis procedure was performed under the same conditions for all samples, the SIMS mass spectra were not normalized before using the PCA method in order to avoid losing information. The first two principal components represent almost 81 % of the variance on the samples.