Rapid chemical screening of microplastics and nanoplastics by thermal desorption and pyrolysis mass spectrometry with unsupervised fuzzy clustering

Thomas P. Forbes*, John M. Pettibone, Eric Windsor, Joseph M. Conny, and Robert A. Fletcher

National Institute of Standards and Technology, Materials Measurement Science Division, Gaithersburg, MD 20899, USA

* Corresponding author: TPF: *e-mail: thomas.forbes@nist.gov, Tel: 1-301-975-2111*

Abstract

The transport and chemical identification of microplastics and nanoplastics (MNPs) are critical to the concerns over plastic accumulation in the environment. Chemically and physically transient MNP species present unique challenges for isolation and analysis due to many factors such as their size, color, surface properties, morphology, and potential for chemical change. These factors contribute to the eventual environmental and toxicological impact of MNPs. As analytical methods and instrumentation continue to be developed for this application, analytical test materials will play an important role. Here, a direct mass spectrometry screening method was developed to rapidly characterize manufactured and weathered MNPs, complementing lengthy pyrolysis-gas chromatography mass spectrometry analyses. The chromatography-free measurements took advantage of Kendrick mass defect analysis, in-source collision induced dissociation, and advancements in machine learning approaches for data analysis of the complex mass spectra. In this study, we applied Gaussian mixture models and fuzzy c-means clustering for the unsupervised analysis of MNP sample spectra, incorporating clustering stability and information criterion measurements to determine latent dimensionality. These models provided insight into the composition of mixed and weathered MNP samples. The multiparametric data acquisition and machine learning approach presented improved confidence in polymer identification and differentiation.

Keywords: Microplastics; Nanoplastics; Environment; Mass Spectrometry; Chemical Characterization; Machine Learning;

Introduction

The accumulation of plastic in natural environments, especially aquatic environments, continues to garner societal focus.[1, 2] These larger plastic pieces successively degrade and breakdown into ever smaller debris. In recent years, attention has concentrated on the detection and characterization of micro- and nano-sized plastic particles, commonly referred to as microplastics and nanoplastics (MNPs), in the environment, food, water, and human tissue.[3-6] The complexity in detection and characterization of these materials comes from the variety and transient nature (*i.e.,* the continuous physicochemical degradation) of plastic materials, colors, sizes, shapes, polymer compositions, environmental degradation, surface properties, and more. These MNP aspects will ultimately play a role in their environmental or toxicological impact. To adequately evaluate and ultimately mitigate potential health risks and environmental impact posed by MNPs, a cohesive suite of analytical techniques, documented methods, and test or reference materials is necessary for MNP measurement and quantification. The development of MNP test and reference materials is vital to advancing methods and instrumentation, as well as characterizing measurement performance and uncertainty. Transferrable MNP control materials (especially down into the nanoplastic regime) that appropriately mimic real-world samples will support the measurements employed by regulators in assessing health risks.

The detection, chemical identification, and quantification of MNPs has been demonstrated through numerous mass-based and particle-based analytical techniques.[4, 7] Particle-based identifications often employ Raman spectroscopy or Fourier transform infrared (FTIR) spectroscopy to measure polymer type and optical microscopy for morphological characteristics (size and shape).[7-9] Recent open source spectra processing and matching software aids in chemical identification.[8] As the sizes of plastic particles decrease, particle-based techniques become increasingly difficult. Polymer identification can also be challenging in the presence of pigments, dyes, and other additives, which may dominate the chemical spectra.[10] Alternatively,

mass-based chemical identifications are most frequently obtained by mass spectrometry (MS) coupled with a thermal desorption or pyrolysis step and gas chromatography (GC).[7, 11-14] Substantial mass spectral libraries aid in chemical identification of a wide-array of encountered compounds.[15] In addition, incorporating chromatography enables more robust analysis of complex mixtures, separation of interfering pigments, leaching compounds, and additives, as well as quantification.[16]

Though chromatography provides many benefits for separating complex mixtures, the technique significantly limits throughput, with separations often taking 30 min to 60 min for each sample. Pyrolysis GC-MS and related techniques are generally coupled with electron ionization and quadrupole mass analyzers. High energy electron ionization (EI) yields fragmented molecules and quadrupole mass spectrometry provides unit resolution spectra, both of which can complicate mass spectral matching for polymers and additive identification. Recently, chemical analysis with ambient ionization and high resolution mass spectrometry has demonstrated the power of such complimentary techniques.[17] Here, no GC separation is conducted, drastically reducing analysis time with a trade-off of more complicated spectra. In addition, high resolution mass analyzers and soft ionization provide more intact molecules and more accurate elemental compositions relative to EI-quadrupole MS. Ambient ionization techniques such as direct analysis in real time (DART)[18-22] and arc plasma-based dissociation (APD)[23] have demonstrated chemical analysis of polymers, with more recent works on DART[24, 25] and atmospheric solids analysis probe (ASAP)[26, 27] moving into microplastics analysis.

Rapidly screening MNP samples for characteristic plastic fingerprints, additives, and leachable compounds with semi-quantitative capabilities and spectral feature comparison provides powerful complimentary information to traditional spectroscopic and GC-MS techniques. In this article, we investigate the high throughput screening of pristine and weathered microplastics and nanoplastics by thermal desorption and pyrolysis mass spectrometry. The

reduction in analysis time by removing chromatography comes with an increase in spectral complexity. Advancements in multivariate statistics and machine learning have presented approaches for processing complex data. We employ the use of multiparametric instrumental methods (*e.g.,* in-source collision induced dissociation settings for extent of fragmentation), data processing techniques (*e.g.*, Kendrick mass defect analysis), and unsupervised and semi-supervised machine learning (*e.g.,* principal component analysis, gaussian mixture models, and fuzzy c-means clustering) to demonstrate the characterization and classification of MNP materials. Ultimately, capturing size dependent chemical signatures, stability, and degradation pathways during an MNP test material's life cycle, as well as how these data inform environmental samples, are priorities for regulators and industry.

Methods

Materials. Polystyrene (PS), polypropylene (PP), high density polyethylene (HDPE), poly(ethylene terephthalate) (PET), and nylon 6 nurdles were purchased from Sigma Aldrich (St. Louis, MO, USA) or Goodfellow (Pittsburg, PA, USA). Duke Standard (#4225A) polystyrene microspheres of (24.61 ± 0.22) µm diameter were also purchased from Sigma Aldrich and poly(methyl methacrylate) (PMMA) particles (approximately 50 µm to 150 µm diameter) were purchased from Scientific Polymer (Ontario, NY, USA). Weathered materials were acquired from the plastic repository at the Center for Marine Debris Research.[28] The criteria for visually characterizing the extent of weathering was outlined in the literature.[28] Due to the inherent difficulties differentiating pyrolysis products of polyolefins (more discussion on this in the Results and Discussion), we chose to focus this study on PS, PP, PET, PMMA, and nylon 6.

Sample Preparation. Pure plastic samples were cryogenically ground using a CryoMill system (Retsch, Newtown PA, USA) with 50 mL stainless steel grinding jar and 25 mm diameter grinding

ball. Additional details of the cryomilling can be found in the supporting information. MNP suspensions for analysis were pipetted into cleaned and tempered sample cups (copper pans, 6 mm diameter × 1.5 mm high, Shimadzu U.S. Webstore, Columbia, MD, USA). These MNP samples were allowed to dry prior to thermal desorption and pyrolysis. Analyzed MNP samples were generally on the order of 10's of micrograms, with a range from 525 ng to 105 µg considered. Physical mixtures of MNPs were artificially produced, including PS-PP-PMMA (21 µg - 30 µg - 30 µg) and PP-PMMA-PET (20 µg - 30 µg - 25 µg).

Instrumentation. Thermal desorption and pyrolysis of MNP samples were conducted by a custom infrared emitter-based platform designed for high temperature heating. Details of the infrared thermal desorber (IRTD) are available in the literature and supporting information (Figure S1 and Table S1).[29-31] Multiple rapid heating profiles were investigated, starting at approximately 150 °C and heating up to approximately 500 °C to 600°C (depending on duration) in 30 s to 120 s, where heating rates were altered by infrared emitter power level. Related high temperature heating platforms based on resistive heating are also commercially available.[21, 24] The infrared thermal desorber was coupled with a direct analysis in real time (DART, Ionsense, Saugus, MA, USA) ion source and time-of-flight mass spectrometer (AccuTOF, JEOL USA, Peabody, MA, USA) through a heated glass junction and aerodynamic-assist interface (Vapur, Ionsense). Details of the glass junction optimization,[32] heated junction configuration,[33] and DART-MS with N_2,[34] can be found in the literature with specific parameters in the supporting information (Table S1). Additional details and support instrumentation methods for pyrolysis-GC-MS, powder dispersion, aerodynamic particle sizing, and SEM imaging can be found in the supporting information.

Data Processing. Mass spectra were extracted from sample data for instrument parameters as specified in the text, using the Mass Center Main (JEOL) software. Kendrick mass defect analysis[35] was conducted using Mass Mountaineer software (Diablo Analytical, Antioch, CA,

USA). The remaining data processing and analysis used a custom MATLAB code (MATLAB 2022a, Mathworks, Inc., Natick, MA, USA).[36] A number of data reduction and unsupervised machine learning methods were employed in this study, including principal component analysis (PCA), Gaussian mixture models (GMM), and fuzzy c-mean (FCM) clustering. Data reduction was completed by PCA on the five pure MNP samples. Further samples, mixtures, and weathered samples were mapped onto the resulting principal components. We then employed GMM and FCM models for MNP sample and spectra differentiation, clustering, and mixture analysis. These unsupervised methods inherently require knowledge of the number of clusters or components. Therefore, an iterative method that measured the Silhouette statistic, Akaike information criterion (AIC), and Bayesian information criterion (BIC) for a GMM with 1 to 15 clusters determined the system's latent dimensionality (*i.e.,* maximizing the Silhouette statistic while minimizing the AIC and BIC). Additional method details are provided in the supporting information and Results and Discussion section.

Data and Code Availability. MNP mass spectra, extracted ion chronograms, Kenrick mass defect data, ion peaks areas as a function of mass, in-source collision induced dissociation (isCID) voltage, and IRTD parameters, IRTD temperature profiles, APS particle size distributions, SEM images, data analysis codes (MATLAB 2022a), GMM dimensionality measurements, and cluster assignments are available on the NIST Public Data Repository: https://doi.org/10.18434/mds2-2957.[36]

Results and Discussion

Microplastics and Nanoplastics Characterization. The screening of microplastics and nanoplastics by thermal desorption and pyrolysis mass spectrometry provided rapid polymer fingerprinting and temporal separation of additives and surfactant species. Figure 1 demonstrates

a few extracted ion chronograms along with mass spectra from multiple time points in the heating ramp for polypropylene and polystyrene MNPs. These samples were manipulated and deposited for analysis as suspensions. Multiple heating profiles were considered (examples in Figure S2(a) and Table S1) and peak areas demonstrated in Figure 1 insets. Additional details can be found in supporting information. Here, we aimed for rapid analysis and employed 50 % power for 50 s, as a trade-off between sensitivity and speed, in the remainder of the analyses. The displayed temperature profiles in Figure 1 represent the heating plate temperature (measured by thermocouple). Early in the heating ramp of PP MNPs a distribution of high mass peaks was observed in the range of *m/z* 585 to *m/z* 809 with *m/z* 28 spacing, representative of an alkene chain (Figure 1(a-i)). Intermixed with that distribution was a series of peaks indicative of the polymer antioxidant and process stabilizer tris(2,4-ditert-butylphenyl) phosphite ($C_{42}H_{63}O_3P$), commercialized as Irgafos 168 (peak assignments in Figure S2 caption). The soft ionization afforded by the DART ion source relative to the harsher electron ionization (EI) common to GC-MS instruments, enabled full intact organic molecules such as the Irgafos antioxidant to be detected. Monitoring the relative intensity of various forms of the compound may provide insight into the extent of oxidization experienced by the underlying plastic product.[37]

Elevated temperatures later in the heating ramp yielded the ion peak distribution for PP pyrolysis products (Figure 1(a-ii)). The dense peak distribution for the polyolefin polymer comprised many series of ions separated by C_3H_6 (*m/z* 42). Other polyolefins such as polyethylene (PE) are expected to generate similarly dense peak distributions.[24] In addition to temporally separating additives from the bulk MNP chemical signature, the rapid heating profile enabled removal of surfactant signatures included to improve sample handling and minimize agglomeration (*e.g.*, Triton X-100). Figure 1(b) displays the spectra for a PS MNPs suspension in deionized water and 0.05% Triton X-100. The ion distributions of these two components were readily isolated. Polystyrene produced a few series arranged predominantly around the monomer, dimer, trimer, and oxidized species; however, ion peaks for the tetramer, pentamer, and hexamer

were also observed (Figure 1(b-ii)). Notable and high intensity peaks included the styrene monomer (*m/z* 105 $[C_8H_8+H]^+$), styrene dimer (*m/z* 207 $[C_{16}H_{14}+H]^+$), oxidized dimer (*m/z* 221 $[C_{16}H_{12}O+H]^+$), styrene trimer (*m/z* 309 $[C_{24}H_{20}+H]^+$), and oxidized trimer (*m/z* 325 $[C_{24}H_{20}O+H]^+$).

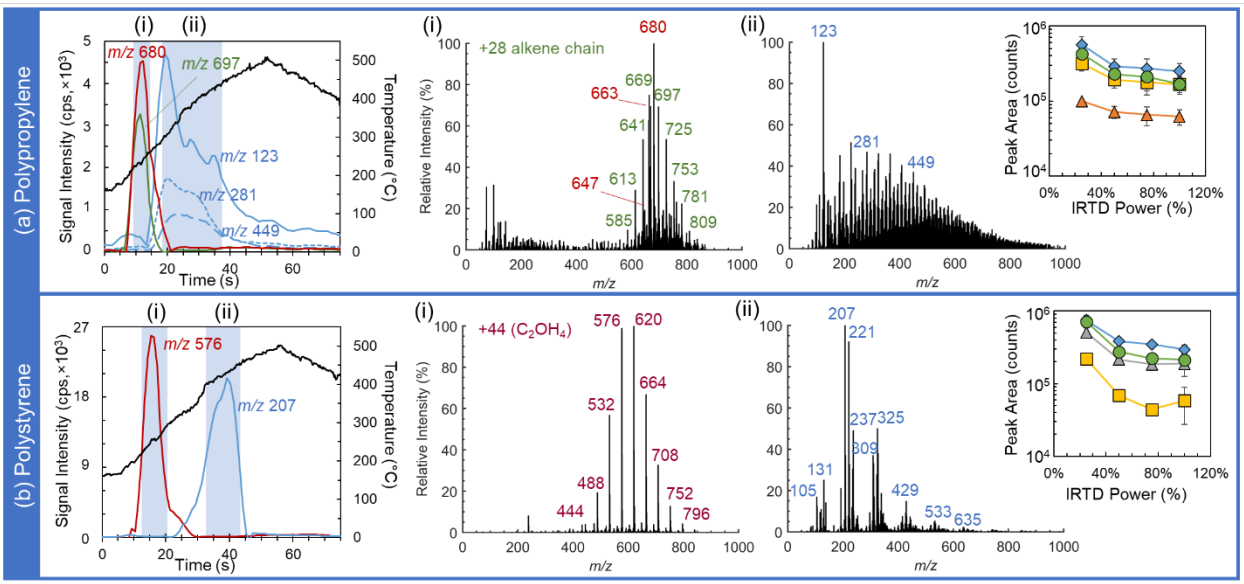

Figure 1. Illustrative extracted ion chronograms of select ions from (a) 53 µg PP and (b) 42 µg PS MNP samples. Mass spectra at (i) early and (ii) late times were extracted during the heating profile. Inset represents peak area of select ions as a function of increasing IRTD power and decreasing heating duration (*i.e.,* faster heating rates). Data points and uncertainty represent the average and standard deviation of triplicate measurements. Inset (a) PP: (▲) *m/z* 123 at 20 V isCID, (●) *m/z* 123 at 60 V isCID, (♦) *m/z* 109 at 60 V isCID, and (■) *m/z* 95 at 60 V isCID. Inset (b) PS: (■) *m/z* 327 at 20 V isCID, (▲) *m/z* 207 at 20 V isCID, (●) *m/z* 105 at 60 V isCID, and (♦) *m/z* 91 at 60 V isCID. Details of in-source collision induced dissociation are in the next sections.

The MNP samples investigated here were predominantly a wide distribution of sized particles, cryomilled from commercial nurdles. Aerodynamic particle sizer measurements captured the distribution in the size range of 0.5 µm to 20 µm, with an additional bin to estimate events for particles < 0.5 µm. Generally, 20 % to 40 % of the MNP concentration were measured in the nanoplastic (sub-micron) regime (Figure S3). But, as is well documented for the nano-regime, these nanoplastics only accounted for around 0.2 % to 1.2 % of the total MNP mass across the measured size range (up to 2 % of the MNP mass for PS first filtered to 0.2 µm to 5

μm). MNPs are a small mass fraction of environmental samples and as future work focuses more on nanoplastic test material production and characterization, we aimed to minimize total masses investigated. Here, we focused on 10's of microgram MNP samples, 2 orders of magnitude less material that previous works.[24]

Figure 2 displays the spectra for select MNP samples, including nylon 6, PET, PMMA, PP, and PS. Polymers such as PS and nylon 6 yielded spectra that were easier to interpret and identify monomers, dimers, trimers, etc. (Figures 2(a) and 2(e)). The main peaks of PS were identified above. Similarly, the nylon 6 spectra exhibited a relatively neat spectra with major ions for the protonated monomer (*m/z* 114 $[C_6H_{11}NO+H]^+$), dimer, (*m/z* 227 $[(C_6H_{11}NO)_2+H]^+$), trimer (*m/z* 340 $[(C_6H_{11}NO)_3+H]^+$), and tetramer (*m/z* 453 $[(C_6H_{11}NO)_4+H]^+$). However, other polymers yielded large ion distributions that were complex and difficult to interpret. For example, PP and PMMA in Figure 2 produced large ion distributions with numerous series separated by the respective repeat unit. PMMA exhibited several presumptively identified ion series, including protonated molecules $[(C_5H_8O_2)_n+H]^+$, ammonium adducts $[(C_5H_8O_2)_n+NH_4]^+$, and related fragments $[(C_5H_8O_2)_n-CH_2+NH_4]^+$. Derivatizing thermochemolytic agents such as tetramethylammonium hydroxide (TMAH) typically used for pyrolysis-GC-MS of PMMA were not required here.[14, 38] Poly(ethylene terephthalate) (PET) produced major peaks for the protonated monomer (*m/z* 193 $[C_{10}H_8O_4+H]^+$) and fragments (*m/z* 149 $[C_9H_8O_2+H]^+$ and *m/z* 105 $[C_7H_5O]^+$), as well as numerous series with the PET $(C_{10}H_8O_4)_n$ repeat unit (Figure 2(b)). All MNP sample compositions were confirmed by pyrolysis-GC-MS (CDS Analytical Pyroprobe 6150 with Autosampler coupled to a Thermo Trace 1310 Gas Chromatograph / TSQ 8000evo triple quadrupole mass spectrometer). Details can be found in the supporting information. Compound gas chromatography peaks were identified by library matching with NIST MS Search 2.0 (Figure S4). Indicator compounds from the literature were used for MNP identification, including caprolactam (*m/z* 113), vinyl benzoate (*m/z* 105), methyl methacrylate (*m/z* 100), 2, 4-dimethyl-1-heptene (*m/z* 126), and styrene trimer (*m/z* 91)

for nylon, PET, PMMA, PP, and PS.[38-42] In addition, the potential for quantification was briefly considered (Figure 2(f)), with details found in the supporting information.

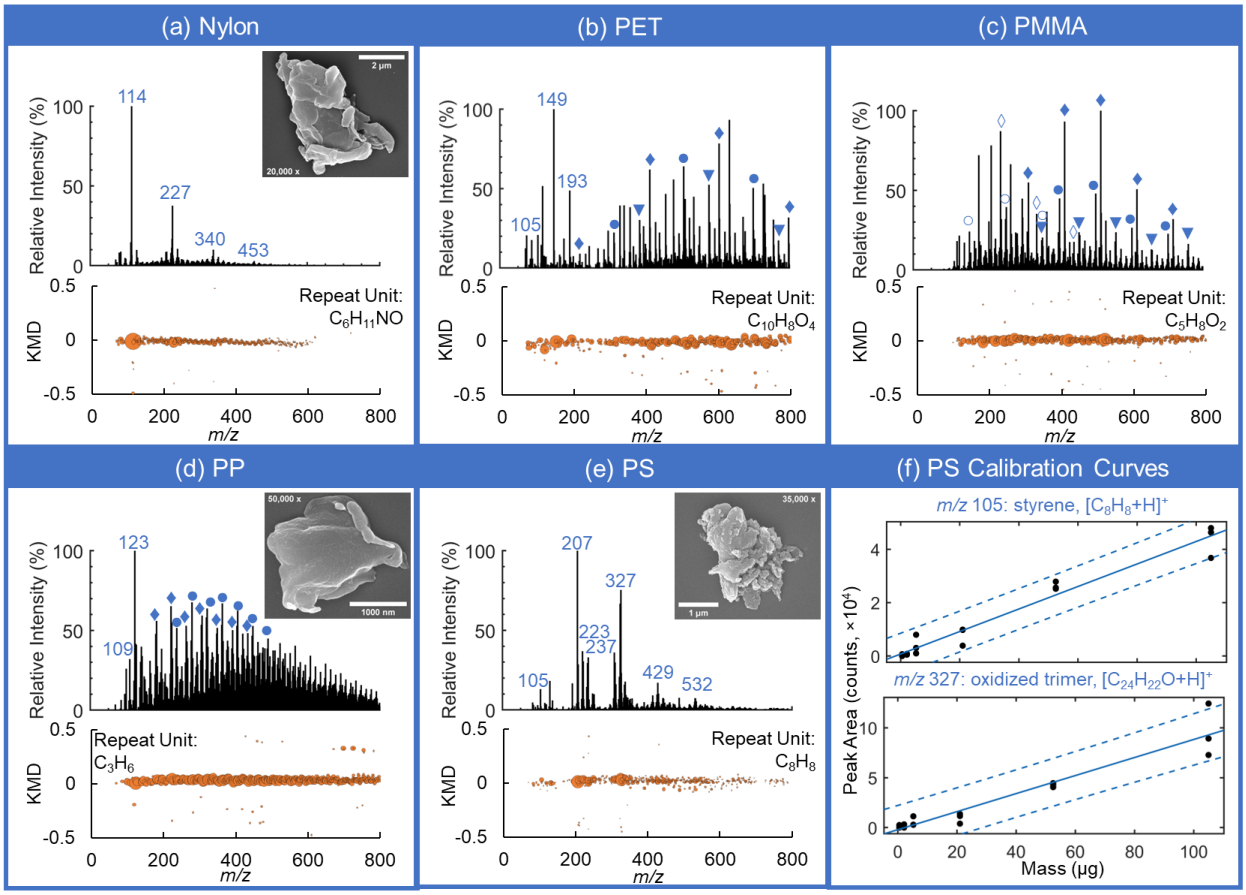

Figure 2. Representative mass spectra and Kendrick mass defect plots for (a) 40 µg nylon 6, (b) 40 µg PET (c) 60 µg PMMA, (d) 40 µg PP and (e) 52 µg PS. (f) Calibration curves for the styrene monomer (R^2: 0.96) and oxidized trimer (R^2: 0.92) for triplicate measurements of 525 ng to 105 µg. Solid line and dashed lines represent linear fit and 95 % confidence interval.

Tools and methods exist to aid in the interpretation of large ion distributions where peak series and repeating units are difficult to manually identify. For example, Kendrick mass defect (KMD) analysis supported the determination and visualization of the repeating unit across the potentially many peak series.[35, 43, 44] KMD plots corresponding to each mass spectrum are also displayed in Figure 2. The bubbles in these plots represented mass spectral peaks with bubble size increasing proportionally to peak intensity. In these examples, the KMD plots were generated for the designated polymer repeat unit (Table S2), creating horizontal lines of bubbles with zero

y-intercept. Variations in the slope corresponded to different repeat units, while variations in y-intercept represented alternative end groups. The KMD analysis clearly demonstrated the high-density peak distributions in each MNP sample were associated with individual polymers. The distributions of peaks seen around high intensity monomers, dimers, trimers, etc. for nylon 6 and PS were also associated with those polymers. In addition to traditional KMD analysis, "Reverse Kendrick Mass Defect Analysis" can be employed to determine the repeat units of unknown samples or polymers.[43] Reverse KMD analysis rotates the plot until the desired series is aligned horizontally, identifying the repeat unit and therefore polymer composition. However, reverse KMD analysis is most successful when some predetermined indication of what polymer or polymers might be present is known. Next, we employ machine learning approaches to aid MNP differentiation and preliminary mixture compositions for use with reverse KMD analysis.

MNP Differentiation and Mixture Analysis. The above chemical characterization of select MNP samples provided the foundation for exploring multivariate statistic and machine learning approaches for MNP differentiation and mixtures analysis. Investigating unknown samples and mixtures introduces several hurdles. As an alternative to a targeted approach to MNP composition determination introduced above, we also considered dimensionality reduction of the complex spectra by PCA. Details of the analysis and data preprocessing can be found in the Methods section and supporting information. PCA was completed for a dataset of 38 mass spectra, comprised of 5 to 11 spectra of pure MNP samples of nylon, PET, PMMA, PS, and PP. Figure 3(a) displays the percent explained variance and cumulative variance for increasing number of principal components. The first three principal components accounted for 78.3% of the data variance and enabled straight-forward 3-dimensional visualization. The Figure 3 inset shows the first three principal components for pure MNP samples, allowing for visual grouping of samples. The corresponding PCA loading plots can be found in the supporting information (Figure S5). A

preliminary characterization of the reduced data for pure MNP samples was completed with a probabilistic Gaussian mixture model (GMM).

Here, the GMM was employed to cluster the samples into a specified number of groups, where each cluster (or group or component) is comprised of a mean and covariance. Each MNP sample was assigned a proportion or fraction to each cluster. We determined the latent dimensionality (*i.e.,* number of clusters) by employing a variation, similar to the expansion of traditional non-negative matrix factorization (NMF) to NMFk,[45-47] which coupled GMM clustering with Silhouette statistics and information criteria – both the Akaike information criterion (AIC) and Bayesian information criterion (BIC). The AIC and BIC differ in the extent of penalizing model complexity (*e.g.,* AIC may overfit for small samples). The dimensionality reduced data in the Figure 3 inset was fit with GMMs of successively increasing number of clusters. Figure 3(b) displays the average Silhouette statistics, AICs, and BICs, from ten replicate fitted models with random initial conditions. The optimal number of clusters (*i.e.,* 5) combined a high Silhouette statistic (maximizing cluster stability) and low information criteria. The AIC and BIC yielded similar trends with the AIC tending toward a more complex model as expected. Figure 3(c) displays the 5-component GMM results for the pure single plastic MNP samples (circle datapoints). Though the datapoints in Figure 3(c) were colored with the maximum posterior probability, each point was a linear combination of all five spectral components to varying degrees (*i.e.,* membership score or partition values).

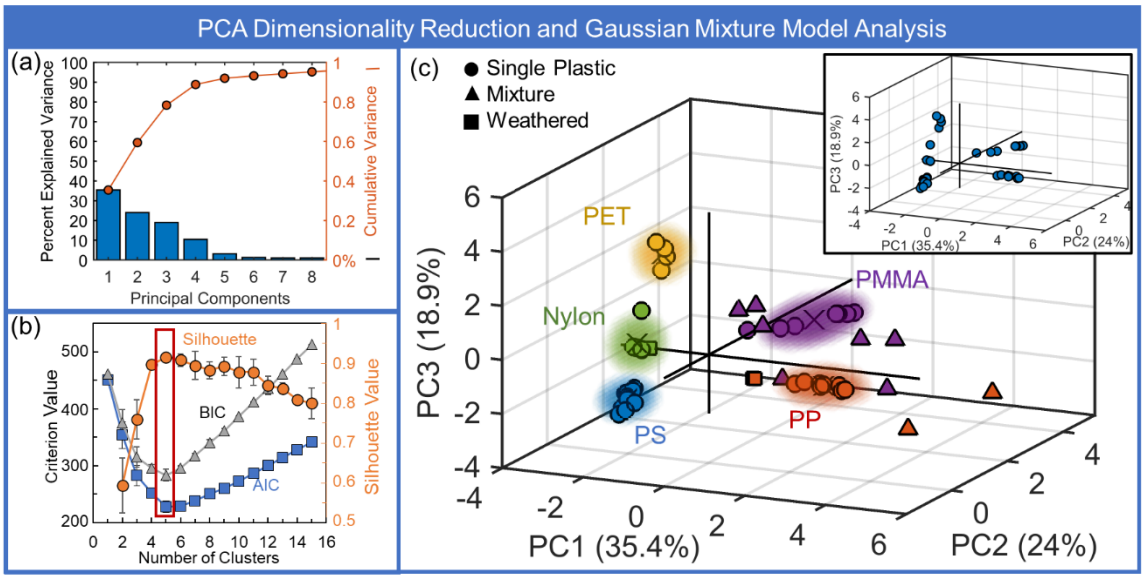

Figure 3. (a) Explained and cumulative variance of PCA as a function of increasing principal components. (b) Silhouette statistic (●), BIC (▲), and AIC (■) values for increasing number of clusters. (c) Pure MNP samples (●) clustered by data point color based on a 5-component GMM. Shaded ellipsoids represent 99 % confidence regions. Inset displays the first three principal components of reduced data prior to GMM clustering. MNP mixtures (▲) and weathered samples (■) were labeled/colored based on the preliminary pure MNP GMM. MNP samples were a wide range of masses, but all on the order of 10's of µg.

We took an unsupervised approach to clustering these data and generating the GMM, however, in this simplified demonstration with relatively few clusters, the cluster identities were determined. Manual inspection of the reconstructed mass spectra for each cluster mean (Figure S6) and select raw spectra of MNP samples from each cluster were used for identification, though comparison to mass spectral libraries and cosine similarity scores could also be employed. These composition labels were included in Figure 3(c). The GMM was then applied to a series of mixed MNP and weathered samples. These included physical mixtures of PS-PP-PMMA (21 µg - 30 µg - 30 µg) and PP-PMMA-PET (20 µg - 30 µg - 25 µg), and weathered PP and nylon. Again, the weathered MNPs (squares) and mixed MNPs (triangles) were labeled (colored) based on the cluster with maximum posterior probability. The mildly weathered nylon samples fell within the 99 % confidence ellipsoid of the pristine nylon cluster, with spectra dominated by the monomer, dimer, and trimer. The severely weathered PP MNP samples fell just outside the 99 % confidence

region, but still yielded the highest posterior probability for the PP cluster. Upon further manual inspection, the weathered samples exhibited peaks at *m/z* 91 and *m/z* 150 that weren't observed in the pristine PP samples (Figure S7). Future work will focus on the analysis of MNP weathering; however, this demonstrated the utility of rapid screening samples to detect differences. The mixed MNP samples were spread across the principal component space and predominantly exhibited maximum posterior probabilities for PMMA or PP.

To take a closer look at the mixed MNP samples, we also considered a more constrained clustering approach, fuzzy c-means (FCM), which is like a limited GMM employing shared diagonal covariances. In FCM, each sample belonged to all clusters to varying extent. Fuzzy c-means enabled straightforward manipulation of the cluster overlap through a single parameter (*i.e.,* the fuzzy partition matrix exponent). We used a 5-cluster system as determined in the above analysis based on the Silhouette statistic, AIC, and BIC (Figure 3(b)). Figure 4(a) displays the MNP dataset mapped onto the same principal components from above (visualizing just the first two) with cluster assignments based on maximum FCM cluster membership score. The mixtures and weathered samples in Figure 4(a) were colored (and assigned) with the maximum membership score. Though, each sample (and especially the mixtures) exhibited numerous membership partitions with significant values. This was visualized as pie graphs of a couple select mixture samples as presented in the Figure 4 inset. The flexibility in cluster membership for these techniques was exploited in the identification of samples that were potential mixtures. We also adjusted membership fraction limits to identify samples that exhibited large contributions from multiple components/clusters. For example, the sample datapoints circled in Figure 4(a) represent MNP samples that exhibited a maximum membership less than 50%. More or fewer samples were identified based varying this parameter – a 50 % threshold identified 4 samples, a 60% threshold identified 7 samples, and a 70% threshold identified 11 samples.

We then employed the largest membership partitions as a potential mixture composition to inform the rotation of Kendrick mass defect graphs for identifying repeat units. Figure 4(b)

displays the mass spectrum from a MNP mixture of PP, PMMA, and PET. At first glance, the diverse peak distributions of each polymer created a convoluted spectrum, making manual interpretation challenging. The FCM membership partitions revealed large fractions for PP and PMMA. Associated repeat units were then readily identified in the KMD graph rotated for each (Figure 4(c)). For the mixtures investigated here, the PS or PET components were difficult to identify among the diverse PP and PMMA ion peak distributions (Figure S8). This was reflected in the mixed MNP sample locations in the principal component space and dominant cluster labeling with both the GMM and FCM approaches.

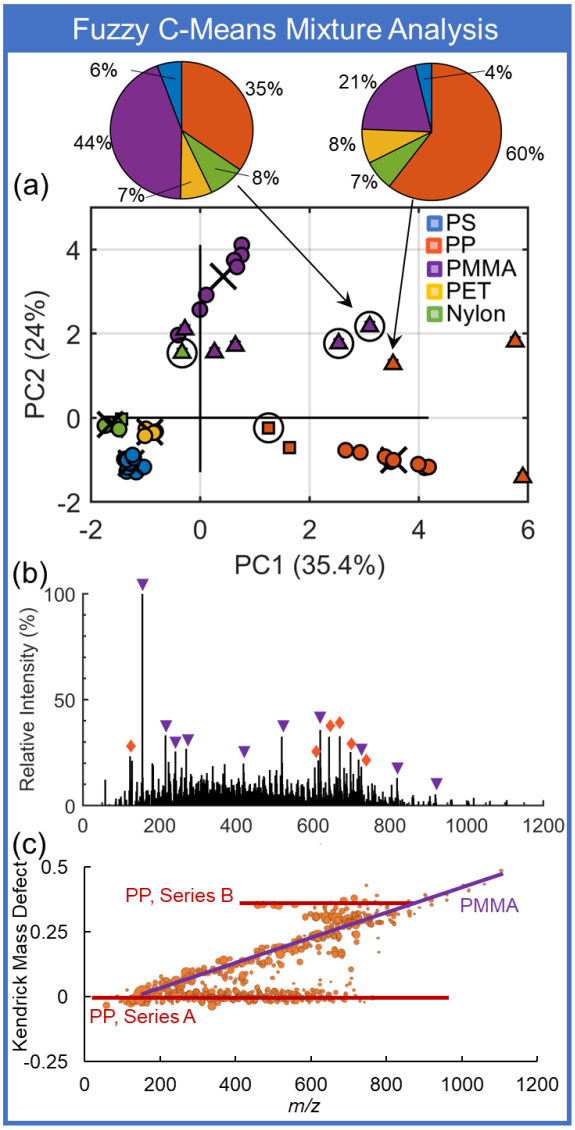

Figure 4. (a) Pure MNP samples (●) clustered by data point color based on a 5-component FCM model with fuzzy partition matrix exponent of 1.7. MNP mixtures (▲) and weathered samples (■) were labeled/colored based on the maximum membership partition. Datapoints circled represent samples with a maximum membership partition less than 50 %. MNP samples were a wide range of masses, but all on the order of 10's of µg. Mass spectra (b) and Kendrick mass defect plot rotated for PP repeat unit (c) of a mixed MNP sample. Inset pie charts represent membership partitions as percentages (%) for two MNP mixture samples.

In-source Collision Induced Dissociation. The mass spectrometer used in this study did not have tandem mass spectrometry capabilities, however, much work has been done employing in-source collision induced dissociation (isCID) for declustering, compound fragmentation, inorganic oxidizer detection, and gaining additional chemical information for library searching algoritms.[48-52] Here, isCID was conducted by varying voltages (orifice 1) in the differentially pumped region to characterize fragmentation and enhance detection of low(er) mass species. Figure 5 displays the integrated peak areas of select ions from PS, PP, PET, and PMMA as a function of increasing isCID. The spectra used in the investigations described above were from low(er) isCID values, generally 20 V to 30 V. Larger oligomers, clusters, and adducts were observed in this range (20 V to 30 V) and exhibited in Figure 2. Increasing isCID led to fragmentation of these larger species, corresponding with an increase in monomers and select fragments. For example, PS exhibited oxidized trimer and dimer ions that had maximum integrated peaks areas around 20 V to 30 V (Figure 5(a)). The oxidized trimer and dimer ions fragmented, decreasing in intensity beyond these isCID values, which corresponded to an increase in the PS monomer (*m/z* 105 $[C_8H_8+H]^+$), toluene (*m/z* 91 $C_7H_7^+$), and toluene fragment (*m/z* 77 $C_6H_5^+$ and *m/z* 65 $C_5H_5^+$) ions (Figure 5(e)). Similarly, Figures 5(c) and 5(f) demonstrate the fragmentation of the dense PET ion distribution of larger oligomers and adducts toward predominantly fragments for vinyl benzoate (*m/z* 149 $[C_9H_8O_2+H]^+$) and the benzoyl radical (*m/z* 105 $C_7H_5O^+$). In general, the MNPs investigated here demonstrated maxima in the select fragmentation ions peak areas in the range of 60 V to 90 V range (Figure 5). High isCID mass spectra for the remaining MNPs can be found in the Supporting Information (Figure S9).

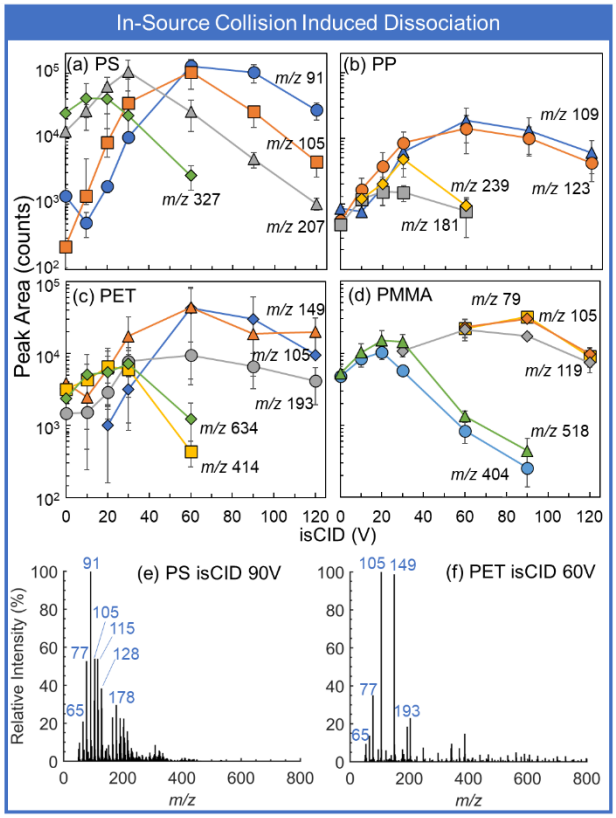

Figure 5. Integrated peak areas for select ions as a function of in-source collision induced dissociation (isCID) voltage for (a) 42 µg PS, (b) 42 µg PP, (c) 40 µg PET, and (d) 48 µg PMMA. Data points and uncertainty represent the average peak area and standard deviation from 3 to 5 replicate measurements. Representative fragmentation spectra at high isCID for (e) PS at 90 V and (f) PET at 60 V.

The above multivariate analyses (*i.e.,* PCA dimensionality reduction, Gaussian mixture modeling, and fuzzy c-means clustering) were completed on the corresponding high isCID (60 V to 90 V) spectra extracted from each MNP sample (Figures S10). The modeling and clustering largely yielded similar results to the low isCID data (additional details in the supporting information). To directly compare the cluster membership, Figures 6(a) and 6(b) represent the membership scores of the fuzzy c-means clustering for low and high isCID spectra as a function of sample in radar charts. Here, the pure MNP samples were ordered by cluster from Samples 1 to 38; Samples 39 through 47 represent the MNP mixture samples; and Samples 48 through 51

represent the handful of weathered MNPs. Generally, the composition of pure MNP samples was clear based on these analyses. However, several pure MNP samples that appeared to be mixtures when clustering at low isCID (*e.g.,* Samples 11 and 14 in Figure 6(a)), were now more definitively labeled when also considering the high isCID results.

Figure 6(c) displays the membership partitions as a percentage for the MNP mixtures and weathered MNP samples, as stacked column plots for a magnified view (relative to the full data radar charts). The distinct and relatively clean mass spectra of the weathered nylon samples exhibited the dominant monomer, dimer, and trimer species, all making differentiation straightforward (Samples 48-49). The low isCID of the weathered PP yielded a potential mixture, though PP made up the largest fraction of the membership scores (Samples 50-51). Fragmenting the weathered PP MNP samples resulted in definitively characteristic high isCID PP spectra (Figure 6(c-ii)). Both the low and high isCID spectra of the two MNP mixtures (PS-PP-PMMA [Samples 39-42] and PP-PMMA-PET [Samples 43-47]) were dominated by the PP and PMMA ions (Figure 6(c)). This was also observed in the principal component space in Figures 3 and 4. However, the high isCID results extracted more significant membership ratios for the PS and PET components of several of these samples (*e.g.,* Samples 40, 43, and 47). The combination of multiple spectra and analysis techniques improved confidence in polymer identification and differentiation.

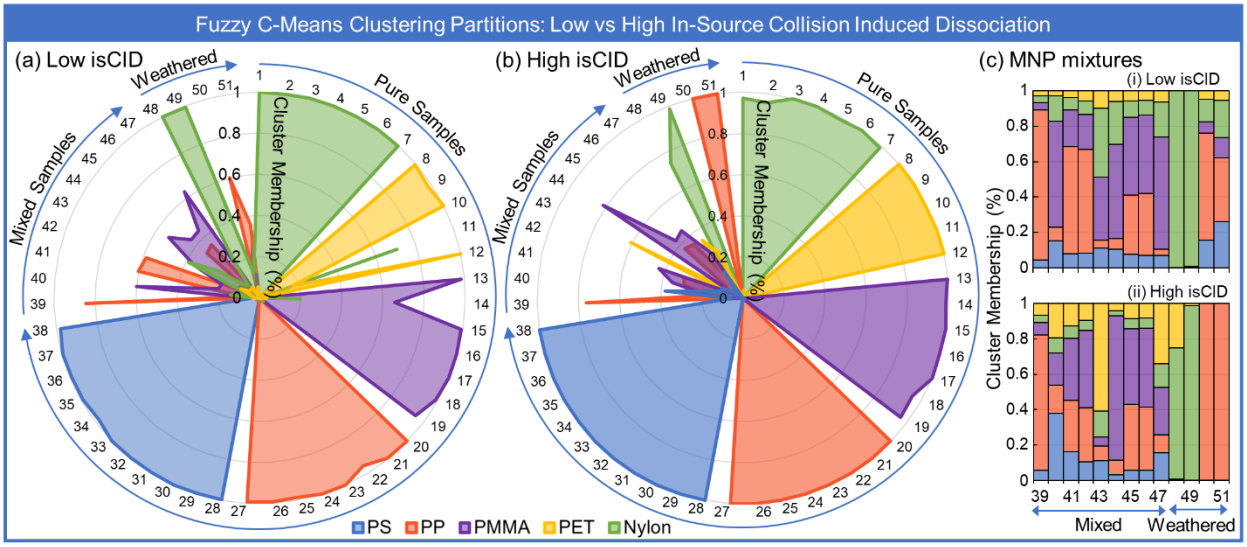

Figure 6. Radar charts (or spider charts) representing the FCM cluster membership partitions for pure (#1 - #38), mixed (#39 - #47), and weathered (#38 - #51) MNP samples at (a) low and (b) high in-source collision induced dissociation (isCID). (c) Isolated cluster membership partitions plotted as stacked column charts for the MNP mixtures only at (i) low and (ii) high isCID.

The analytical and data processing methods presented here provided an initial demonstration with select subset of rather simple pure, mixed, and weathered MNP samples. To truly apply such a process to environmental MNP samples, application to increasingly complex systems will be required. Chromatography-free analysis also introduces hurdles for differentiation of polyolefins from mass spectra and Kendrick mass defect graphs. Including HDPE in the present analysis yielded clustering with PP in the 5-component GMM (Figures S11). Methods for polyolefin differentiation is an aim of ongoing work. In addition to the increase in complexity from commercially available plastic compositions with polymer blends and additives, we must also consider sample preparation and data processing methods that can account for the added background diversity of natural samples (*e.g.,* natural organic material, dissolved organic material, and black carbon). This may also take the form of alternative methods that have a two-stage heating profile, one for thermal desorption of more labile species, and one for plastic pyrolysis. As with any chromatography-free ambient or atmospheric pressure mass spectrometry

techniques, competitive ionization and matrix effects may hinder detection of low concentration species in diverse environmental mixtures. Employed as a screening or triage method, these techniques can determine next steps and inform use of targeted pyrolysis-GC-MS methods and appropriate internal standards.

Conclusions

Direct thermal desorption and pyrolysis mass spectrometry supported by unsupervised machine learning enabled rapid chemical characterization of cryomilled and commercial microplastic and nanoplastic materials. Multiparametric (*i.e.,* mass spectra at multiple isCID voltages) datasets were collected simultaneously, providing robust chemical characterization (*e.g.,* chemical signatures with different fragmentation extents). These avenues complement traditional pyrolysis-GC-MS measurements in support of analytical test material production and examination of environmental or weathered MNPs. Gaussian mixture models and fuzzy c-means clustering in conjunction with Silhouette statistics and information criteria directly learned the number of constituent materials (*i.e.,* latent dimensions) from the dataset and provided insight into MNP composition through posterior probabilities and partition membership fractions. The proposed unsupervised learning framework is being scaled up to include more MNP plastics, blended plastics, plastics with a range of additives, and environmental samples. The framework will also be formulated into a semi-supervised model with ground truth MNP samples and a rigorous characterization of classification performance. The development of models specific to applications or sample types may be needed, but once a model is developed, further analyses will be simple data processing.

Pyrolysis GC-MS remains the gold standard for quantitative (and qualitative) analysis of MNPs beyond the detection limits and/or spatial resolution of spectroscopic techniques. The preliminary work here demonstrated the potential for rapid screening and semi-quantitative measurements.

As a triage approach, the rapid screening may inform targeted methods for further pyrolysis-GC-MS analyses. Ongoing work is further investigating the potential for semi-quantitative capabilities by including deuterated internal standards and considering potential matrix effects from mixtures and environmental background. Future work will also necessarily consider alternative sample preparation methods (*e.g.,* liquid extraction). Here, direct MNP suspensions were measured, however, as we consider more dilute suspensions at decreasing mass, Poisson statistics and the MNP size distributions will begin to play a role. Finally, ongoing and future work will focus on controlled MNP aging investigations of test material chemical changes and how they compare to environmental samples.

Supporting Information

Additional experimental method details, discussion of results, system parameter and peak assignment tables, and figures as noted in the text can be found in the online supporting information.

Acknowledgments

The authors would like to thank Chip Cody at JEOL for his discussions on the use of Mass Mountaineer and reverse Kendrick mass defect analysis. We thank Jennifer Lynch at NIST for providing the weathered plastic samples, and Kim Rogers and Mark Surette at EPA for discussions and initial cryomilling of these samples.

The authors declare the following competing financial interest(s): Thomas P. Forbes is an inventor on a patent application (describing infrared thermal desorption (IRTD)) to the United States of America as represented by the Secretary of Commerce, The National Institute of Standards and Technology. The remaining authors declare no competing financial interests.

Notes

† Certain commercial equipment, instruments, or materials are identified in this article in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by NIST, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

‡ Official contribution of the National Institute of Standards and Technology; not subject to copyright in the United States.

References

(1) Geyer, R.; Jambeck, J. R.; Law, K. L. Production, use, and fate of all plastics ever made. *Sci. Adv.* **2017**, *3*, e1700782.
(2) da Costa, J. P.; Santos, P. S. M.; Duarte, A. C.; Rocha-Santos, T. (Nano)plastics in the environment – Sources, fates and effects. *Sci. Total Environ.* **2016**, *566-567*, 15-26.
(3) Sorensen, R. M.; Jovanović, B. From nanoplastic to microplastic: A bibliometric analysis on the presence of plastic particles in the environment. *Mar. Pollut. Bull.* **2021**, *163*, 111926.
(4) Vitali, C.; Peters, R. J. B.; Janssen, H.-G.; Nielen, M. W. F.; Ruggeri, F. S. Microplastics and nanoplastics in food, water, and beverages, part II. Methods. *TrAC, Trends Anal. Chem.* **2022**, *157*, 116819.
(5) Vitali, C.; Peters, R.; Janssen, H.-G.; W.F.Nielen, M. Microplastics and nanoplastics in food, water, and beverages; part I. Occurrence. *TrAC, Trends Anal. Chem.* **2022**, 116670.
(6) Leslie, H. A.; van Velzen, M. J. M.; Brandsma, S. H.; Vethaak, A. D.; Garcia-Vallejo, J. J.; Lamoree, M. H. Discovery and quantification of plastic particle pollution in human blood. *Environ. Int.* **2022**, *163*, 107199.
(7) Ivleva, N. P. Chemical Analysis of Microplastics and Nanoplastics: Challenges, Advanced Methods, and Perspectives. *Chem. Rev.* **2021**, *121*, 11886-11936.
(8) Cowger, W.; Steinmetz, Z.; Gray, A.; Munno, K.; Lynch, J.; Hapich, H.; Primpke, S.; De Frond, H.; Rochman, C.; Herodotou, O. Microplastic Spectral Classification Needs an Open Source Community: Open Specy to the Rescue! *Anal. Chem.* **2021**, *93*, 7543-7548.
(9) Gillibert, R.; Balakrishnan, G.; Deshoules, Q.; Tardivel, M.; Magazzù, A.; Donato, M. G.; Maragò, O. M.; Lamy de La Chapelle, M.; Colas, F.; Lagarde, F.; Gucciardi, P. G. Raman Tweezers for Small Microplastics and Nanoplastics Identification in Seawater. *Environ. Sci. Technol.* **2019**, *53*, 9003-9013.
(10) Xu, J.-L.; Thomas, K. V.; Luo, Z.; Gowen, A. A. FTIR and Raman imaging for microplastics analysis: State of the art, challenges and prospects. *TrAC, Trends Anal. Chem.* **2019**, *119*, 115629.
(11) Yakovenko, N.; Carvalho, A.; ter Halle, A. Emerging use thermo-analytical method coupled with mass spectrometry for the quantification of micro(nano)plastics in environmental samples. *TrAC, Trends Anal. Chem.* **2020**, *131*, 115979.
(12) Blancho, F.; Davranche, M.; Hadri, H. E.; Grassl, B.; Gigault, J. Nanoplastics Identification in Complex Environmental Matrices: Strategies for Polystyrene and Polypropylene. *Environ. Sci. Technol.* **2021**, *55*, 8753-8759.

(13) Sullivan, G. L.; Gallardo, J. D.; Jones, E. W.; Hollliman, P. J.; Watson, T. M.; Sarp, S. Detection of trace sub-micron (nano) plastics in water samples using pyrolysis-gas chromatography time of flight mass spectrometry (PY-GCToF). *Chemosphere* **2020**, *249*, 126179.
(14) Seeley, M. E.; Lynch, J. M. Previous successes and untapped potential of pyrolysis–GC/MS for the analysis of plastic pollution. *Anal. Bioanal. Chem.* **2023**.
(15) *NIST Mass Spectral Library*. National Institute of Standards and Technology, http://chemdata.nist.gov. (accessed 2022 -07-06).
(16) La Nasa, J.; Biale, G.; Fabbri, D.; Modugno, F. A review on challenges and developments of analytical pyrolysis and other thermoanalytical techniques for the quali-quantitative determination of microplastics. *J. Anal. Appl. Pyrolysis* **2020**, *149*, 104841.
(17) Feider, C. L.; Krieger, A.; DeHoog, R. J.; Eberlin, L. S. Ambient Ionization Mass Spectrometry: Recent Developments and Applications. *Anal. Chem.* **2019**, *91*, 4266-4290.
(18) Cody, R. B.; Fouquet, T. N. J.; Takei, C. Thermal desorption and pyrolysis direct analysis in real time mass spectrometry for qualitative characterization of polymers and polymer additives. *Rapid Commun. Mass Spectrom.* **2020**, *34*, e8687.
(19) Zughaibi, T. A.; Steiner, R. R. Differentiating Nylons Using Direct Analysis in Real Time Coupled to an AccuTOF Time-of-Flight Mass Spectrometer. *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 982-985.
(20) Yamane, S.; Nakamura, S.; Inoue, R.; Fouquet, T. N. J.; Satoh, T.; Kinoshita, K.; Sato, H. Determination of the Block Sequence of Linear Triblock Copolyethers Using Thermal Desorption/Pyrolysis Direct Analysis in Real-Time Mass Spectrometry. *Macromolecules* **2021**, *54*, 10388-10394.
(21) Liang, J.; Frazier, J.; Benefield, V.; Chong, N. S.; Zhang, M. Forensic Fiber Analysis by Thermal Desorption/Pyrolysis-Direct Analysis in Real Time-Mass Spectrometry. *Anal. Chem.* **2020**, *92*, 1925-1933.
(22) Sisco, E.; Forbes, T. P. Forensic applications of DART-MS: A review of recent literature. *Forensic Chem.* **2021**, *22*, 100294.
(23) Zhu, S.-z.; Zhou, B.-w.; Zhang, L.; Zhang, J.; Guo, Y.-l. Rapid Characterization of Polymer Materials Using Arc Plasma-Based Dissociation-Mass Spectrometry. *Anal. Chem.* **2021**, *93*, 12480-12486.
(24) Zhang, X.; Mell, A.; Li, F.; Thaysen, C.; Musselman, B.; Tice, J.; Vukovic, D.; Rochman, C.; Helm, P. A.; Jobst, K. J. Rapid fingerprinting of source and environmental microplastics using direct analysis in real time-high resolution mass spectrometry. *Anal. Chim. Acta* **2020**, *1100*, 107-117.
(25) Schirinzi, G. F.; Llorca, M.; Seró, R.; Moyano, E.; Barceló, D.; Abad, E.; Farré, M. Trace analysis of polystyrene microplastics in natural waters. *Chemosphere* **2019**, *236*, 124321.
(26) Smith, M. J. P.; Cameron, N. R.; Mosely, J. A. Evaluating Atmospheric pressure Solids Analysis Probe (ASAP) mass spectrometry for the analysis of low molecular weight synthetic polymers. *Analyst* **2012**, *137*, 4524-4530.
(27) Vitali, C.; Janssen, H.-G.; Ruggeri, F. S.; Nielen, M. W. F. Rapid Single Particle Atmospheric Solids Analysis Probe-Mass Spectrometry for Multimodal Analysis of Microplastics. *Anal. Chem.* **2022**.
(28) Brignac, K. C.; Jung, M. R.; King, C.; Royer, S.-J.; Blickley, L.; Lamson, M. R.; Potemra, J. T.; Lynch, J. M. Marine Debris Polymers on Main Hawaiian Island Beaches, Sea Surface, and Seafloor. *Environ. Sci. Technol.* **2019**, *53*, 12218-12226.
(29) Forbes, T. P.; Staymates, M. Infrared thermal desorber and performing infrared thermal desorption. U.S. Patent Application 17/144,232, 2021.
(30) Forbes, T. P.; Staymates, M.; Sisco, E. Broad spectrum infrared thermal desorption of wipe-based explosive and narcotic samples for trace mass spectrometric detection. *Analyst* **2017**, *142*, 3002-3010.
(31) Forbes, T. P.; Sisco, E.; Staymates, M. Detection of Nonvolatile Inorganic Oxidizer-Based Explosives from Wipe Collections by Infrared Thermal Desorption—Direct Analysis in Real Time Mass Spectrometry. *Anal. Chem.* **2018**, *90*, 6419-6425.
(32) Sisco, E.; Forbes, T. P.; Staymates, M. E.; Gillen, G. Rapid analysis of trace drugs and metabolites using a thermal desorption DART-MS configuration. *Anal. Methods* **2016**, *8*, 6494-6499.
(33) Forbes, T. P.; Krauss, S. T. Confined DART-MS for Rapid Chemical Analysis of Electronic Cigarette Aerosols and Spiked Drugs. *J. Am. Soc. Mass Spectrom.* **2021**, *32*, 2274-2280.
(34) Sisco, E.; Staymates, M. E.; Forbes, T. P. Optimization of confined direct analysis in real time mass spectrometry (DART-MS). *Analyst* **2020**, *145*, 2743-2750.
(35) Kendrick, E. A mass scale based on CH2 = 14.0000 for high resolution mass spectrometry of organic compounds. *Anal.Chem.* **1963**, *35*, 2146-2154.
(36) Forbes, T. P.; Pettibone, J. M.; Windsor, E.; Conny, J. M.; Fletcher, R. A. *Microplastics and nanoplatics chemical characterization by thermal desorption and pyrolysis mass spectrometry with*

unsupervised machine learning. National Institute of Standards and Technology, 2023. https://doi.org/10.18434/mds2-2957 (accessed 2023 -03-24).
(37) Fouyer, K.; Lavastre, O.; Rondeau, D. Direct Monitoring of the Role Played by a Stabilizer in a Solid Sample of Polymer Using Direct Analysis in Real Time Mass Spectrometry: The Case of Irgafos 168 in Polyethylene. *Anal. Chem.* **2012**, *84*, 8642-8649.
(38) Fischer, M.; Scholz-Böttcher, B. M. Simultaneous Trace Identification and Quantification of Common Types of Microplastics in Environmental Samples by Pyrolysis-Gas Chromatography–Mass Spectrometry. *Environ. Sci. Technol.* **2017**, *51*, 5052-5060.
(39) Okoffo, E. D.; Ribeiro, F.; O'Brien, J. W.; O'Brien, S.; Tscharke, B. J.; Gallen, M.; Samanipour, S.; Mueller, J. F.; Thomas, K. V. Identification and quantification of selected plastics in biosolids by pressurized liquid extraction combined with double-shot pyrolysis gas chromatography–mass spectrometry. *Sci. Total Environ.* **2020**, *715*, 136924.
(40) Okoffo, E. D.; Tscharke, B. J.; O'Brien, J. W.; O'Brien, S.; Ribeiro, F.; Burrows, S. D.; Choi, P. M.; Wang, X.; Mueller, J. F.; Thomas, K. V. Release of Plastics to Australian Land from Biosolids End-Use. *Environ. Sci. Technol.* **2020**, *54*, 15132-15141.
(41) Toapanta, T.; Okoffo, E. D.; Ede, S.; O'Brien, S.; Burrows, S. D.; Ribeiro, F.; Gallen, M.; Colwell, J.; Whittaker, A. K.; Kaserzon, S.; Thomas, K. V. Influence of surface oxidation on the quantification of polypropylene microplastics by pyrolysis gas chromatography mass spectrometry. *Sci. Total Environ.* **2021**, *796*, 148835.
(42) Zhou, X.-x.; Hao, L.-t.; Wang, H.-y.-z.; Li, Y.-j.; Liu, J.-f. Cloud-Point Extraction Combined with Thermal Degradation for Nanoplastic Analysis Using Pyrolysis Gas Chromatography–Mass Spectrometry. *Anal. Chem.* **2019**, *91*, 1785-1790.
(43) Cody, R. B.; Fouquet, T. "Reverse Kendrick Mass Defect Analysis": Rotating Mass Defect Graphs to Determine Oligomer Compositions for Homopolymers. *Anal. Chem.* **2018**, *90*, 12854-12860.
(44) Fouquet, T.; Sato, H. Extension of the Kendrick Mass Defect Analysis of Homopolymers to Low Resolution and High Mass Range Mass Spectra Using Fractional Base Units. *Anal. Chem.* **2017**, *89*, 2682-2686.
(45) Alexandrov, Ludmil B.; Nik-Zainal, S.; Wedge, David C.; Campbell, Peter J.; Stratton, Michael R. Deciphering Signatures of Mutational Processes Operative in Human Cancer. *Cell Rep.* **2013**, *3*, 246-259.
(46) Stanev, V.; Vesselinov, V. V.; Kusne, A. G.; Antoszewski, G.; Takeuchi, I.; Alexandrov, B. S. Unsupervised phase mapping of X-ray diffraction data by nonnegative matrix factorization integrated with custom clustering. *npj Comput. Mater.* **2018**, *4*, 43.
(47) Forbes, T. P.; Gillen, J. G.; Souna, A. J.; Lawrence, J. Unsupervised Pharmaceutical Polymorph Identification and Multicomponent Particle Mapping of ToF-SIMS Data by Non-Negative Matrix Factorization. *Anal. Chem.* **2022**, *94*, 16443-16450.
(48) Weinmann, W.; Stoertzel, M.; Vogt, S.; Wendt, J. Tune compounds for electrospray ionisation/in-source collision-induced dissociation with mass spectral library searching. *J. Chromatogr. A* **2001**, *926*, 199-209.
(49) Forbes, T. P.; Sisco, E. In-source collision induced dissociation of inorganic explosives for mass spectrometric signature detection and chemical imaging. *Anal. Chim. Acta* **2015**, *892*, 1-9.
(50) Musah, R. A.; Cody, R. B.; Domin, M. A.; Lesiak, A. D.; Dane, A. J.; Shepard, J. R. E. DART–MS in-source collision induced dissociation and high mass accuracy for new psychoactive substance determinations. *Forensic Sci. Int.* **2014**, *244*, 42-49.
(51) Forbes, T. P.; Verkouteren, J. R. Forensic Analysis and Differentiation of Black Powder and Black Powder Substitute Chemical Signatures by Infrared Thermal Desorption–DART-MS. *Anal. Chem.* **2019**, *91*, 1089-1097.
(52) Moorthy, A. S.; Tennyson, S. S.; Sisco, E. Updates to the Inverted Library Search Algorithm for Mixture Analysis. *J. Am. Soc. Mass Spectrom.* **2022**, *33*, 1260-1266.

For Table of Contents Only