Metabolomic CSI: High Resolution Lipidomic Profiling of Fingerprints using TD-SICRIT®-MS and Bioinformatics Pipeline

Introduction

In forensic science, identifying and differentiating fingerprints and the chemicals found on them are crucial for linking individuals to crime scenes, establishing timelines, and providing evidence in legal proceedings. Each individual's fingerprint carries a unique chemical signature derived from sweat, oils, and residues from substances touched, making it a valuable tool for forensic investigations. Discriminating between different fingerprints and their chemical compositions aids in identifying suspects, excluding innocent individuals, and reconstructing events. Rapid identification of these chemical constituents enhances the efficiency and accuracy of forensic investigations, expediting case resolution and ensuring justice is served swiftly.

In both metabolomics and lipidomics, distinguishing between different fingerprints and their chemical profiles is crucial for understanding an individual's physiological state, lifestyle, and potential health risks. Metabolites and lipid classes present in fingerprints reflect metabolic processes, dietary habits, exposure to environmental toxins, and underlying health conditions. Rapid identification and differentiation of these profiles enable researchers to uncover patterns, correlations, and biomarkers relevant to health, disease, and drug responses. This knowledge can inform personalized medicine, disease diagnosis, and therapeutic interventions, ultimately enhancing patient outcomes and the delivery of healthcare services. Additionally, lipidomics focuses on the analysis of lipid molecules, such as fatty acids, glycerides, and phospholipids, which are abundant on the skin's surface and contribute to the unique chemical composition of fingerprints. Investigating the role of different lipid classes in fingerprint analysis could provide valuable insights into individual variations and aid in forensic investigations and metabolomic research.

Desorption Electrospray Ionization Mass Spectrometry (DESI) is a powerful technique used for rapid fingerprint analysis in both forensic and metabolomic applications. DESI allows for direct ionization and detection of chemicals on a surface without the need for sample preparation. However, DESI may suffer from ionization limitations and disadvantages, including spatial resolution issues, ion suppression effects, and matrix interference. An additional limitation is that the main ionization mechanisms that drives DESI is an ESI-based mechanism. This means that those components that cannot be softly ionized through these typical methods, like the more non-polar components, are simply lost. These limitations can affect the accuracy and reliability of fingerprint analysis, particularly in complex samples with overlapping signals or low analyte concentrations.

Here, we propose an alternative method that expands upon the limitations and capabilities of DESI, along with a computational workflow to provide dual analysis for both targeted and untargeted forensic or metabolomic analysis. With the SICRIT® Ionization Technology, we are able to ionize a broader range of compounds than the gold standard method by employing not only an ESI-like ionization pathway, but also providing APCI-like and PI-like routes of ionization, which allows for some of the most non-polar compounds to be ionized. Furthermore, as shown in previous applications (Method Mimicry: Metabolomic Targeted Analysis of 79 pesticides to Study Matrix Effects and MS2 Library Analysis of the ESI and SICRIT®), we are able to minimize the matrix effect to the same, if not better, level than the traditional ESI-LC-MS set-up. This means that we also have the possibility of being less affected by matrix effects, while maintaining our high reproducibility and sensitivity, all of which are necessary for this type of analysis.

In this study, we aim to expand on the application presented in a previously published article (Conway, C. et al. (2023) 'Rapid desorption and analysis for illicit drugs and chemical profiling of fingerprints by SICRIT® ion source', Drug Testing and Analysis, pp. 1–8. doi:10.1002/dta.3623). In that work, an in-house thermal desorption device was coupled with the ion source (TD-SICRIT®-MS) to provide a proof-of-concept forensic experiment. Here, we further explore how this technique can also be applied to metabolomics analysis.

Am Mittleren Moos 48
D-86167 Augsburg
www.plasmion.com

T: +49 (0) 821 20713355
E: info@plasmion.de

Sample Preparation & Analysis Conditions

Polar and triglyceride lipid standards were diluted 1:10 in methanol, and 5 µL were pipetted onto glass slides, left to dry before measurement. The fingerprints were collected from eight individuals directly onto the glass slides and stored for up to 1 day at 5°C before analysis. Three replicate fingerprints were collected from each individual's left and right thumb. Collection of fingerprints was done on 2 days, 4 days apart, resulting in 6 to 12 samples of each individual.

Samples were desorbed with an in-house built thermal desorption device at a temperature of 300°C for 30 s, flushed with dry nitrogen at a flow rate of 3 L/min to ensure a stable and clean background. Slides were manually introduced with the sample facing the ion source. Ionization of the desorbed sample was carried out by a SICRIT® ionization source at an amplitude of 1600 V and a frequency of 15 kHz.

Detection and quantification were performed using a high-resolution LTQ Orbitrap XL mass spectrometer (Thermo Scientific, San Jose, USA) with specific parameters: capillary voltage 2.6 V; tube lens voltage 70 V; capillary temperature 275°C; mass window 75 to 1200 m/z; micro scans 1; maximum injection time 250 ms. Automatic gain control (AGC) was applied. Measurements were performed in full scan mode with profile-mode acquisition and positive polarity, employing a resolution of 30,000 (FWHM at 400 m/z).

Lipidomic samples, standards, and fingerprints obtained from the Thermo-Orbitrap underwent processing through an R-script and a Python machine learning pipeline. Raw spectra were centroided, converted to mzML with MSConvert, and processed with PyOpenMS. Baseline correction, smoothing, peak detection, and background subtraction were applied to the extracted MS1 data averaged over a given retention time window. The cleaned MS1 peaks were exported as a feature table, with mz values as columns. To align the spectra across all samples, a virtual lock mass algorithm merged redundant features with a 0.01 mz-value window shift, resulting in a reduced feature table. Features missing from at least 90% of the samples were dismissed and features not consistently observed within two-thirds of the triplicate or sextuplicate samples were considered noise and discarded.

Following an initial PCA to observe sample separation, a machine-learning pipeline was constructed to find an optimal model for differentiation (Figure 1).

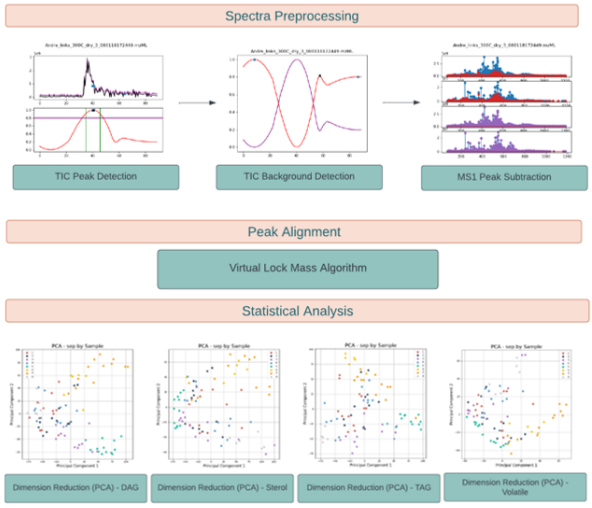


Figure 1: The basic spectra processing and analysis part of the computational pipeline

The data was split 70:30 into a training and test set, stratified and randomized for sample separation. Due to the multi-class nature, the data were binarized. The sklearn Pipeline class applied a series of dimension reduction methods (PCA, SVD, IsoMap, Locally Linear Embedding) and classifiers (Random Forest, K-Nearest Neighbor, Gaussian Process, Naïve Bayes, Decision Tree, Voting Ensemble method) to determine the best model combination. Constructed pipelines underwent Leave-One-Out cross-validation grid search to identify the optimal classifier model for each combination based on validation accuracy.

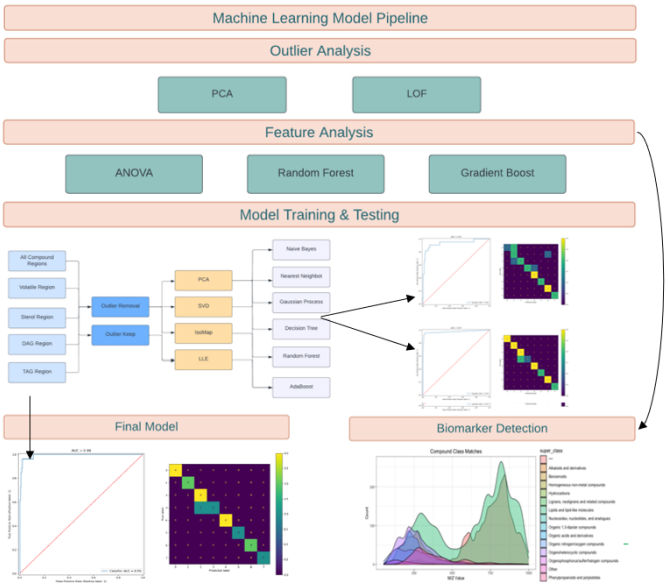


Figure 2: The machine learning model pipeline, coupled with the preliminary biomarker matching tool

Am Mittleren Moos 48
D-86167 Augsburg
www.plasmion.com

T: +49 (0) 821 20713355
E: info@plasmion.de

Additionally, these features collected were run through a preliminary biomarker detecting pipeline, where the mass range defect was set to 0.01 mz for a match.

Results from Analysis of Standard Samples

The initial results that we wanted to obtain from such measurements pertained to the behavior of the desorption device and how well it handled both polar and nonpolar lipid standards, especially since this mass range is at the upper end of our detection limit and is a region that includes potential insight for finger differentiation and lipidomic studies. The list of included compounds and those that were found can be viewed in a table below, where only three could not be conclusively identified from the EIC in positive mode (Table 1). However, with further optimization and the use of both positive and negative ion mode, potentially these compounds should be visible with our source using the desorption device and have been well documented through LC-SICRIT® -MS (Lipidomics Decoded: Targeted Assignment of Polar Lipids using the SICRIT® LC-Module). With these standards we were able to obtain clean and reproducible MS1 spectra of the standards, even with the manual nature of the preparation and doping of the glass slides (Figure 3 & 4).

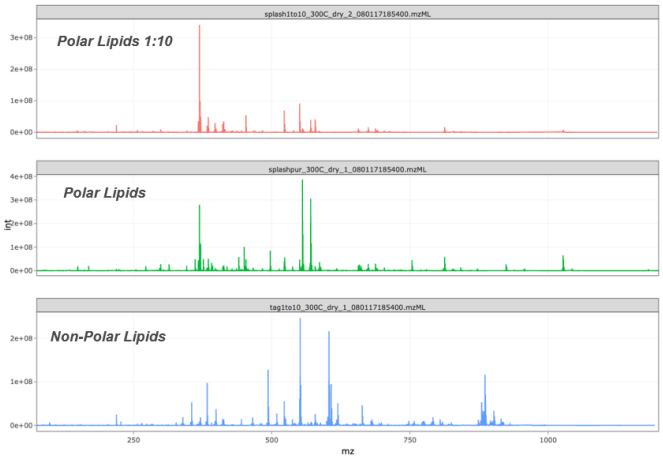


Figure 3: MS1 of one triplicate from each standard run, not background subtracted, including the additional Polar Lipid standard, diluted 1:10.

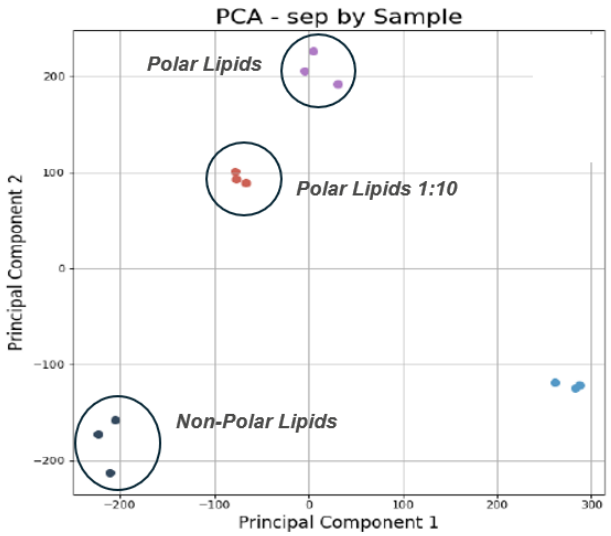
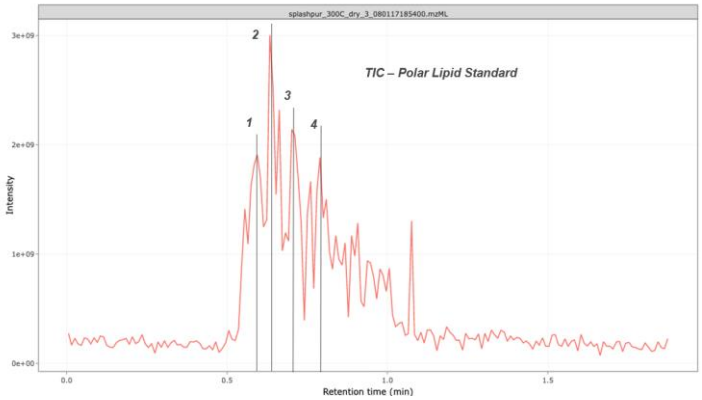


Figure 4: PCA of the replicates for the Non-Polar Lipids, Polar Lipids, and Polar Lipids 1:10. Additionally, in the blue, are the blanks of the glass slide itself to ensure that the reproducibility wasn't inherently due to reproducible blanks for any of the standards.

Upon further inspection of the individual compounds and their ionizations, we found that these compounds ionize with the identical ionizations found in a previous LC-SICRIT® -MS application note with these exact compounds (Lipidomics Decoded: Targeted Assignment of Polar Lipids using the SICRIT® LC-Module). Additionally, we found that the device itself provides deviations in retention times depending on how polar or non-polar the compound is (Figure 5).



Am Mittleren Moos 48
D-86167 Augsburg
www.plasmion.com

T: +49 (0) 821 20713355
E: info@plasmion.de

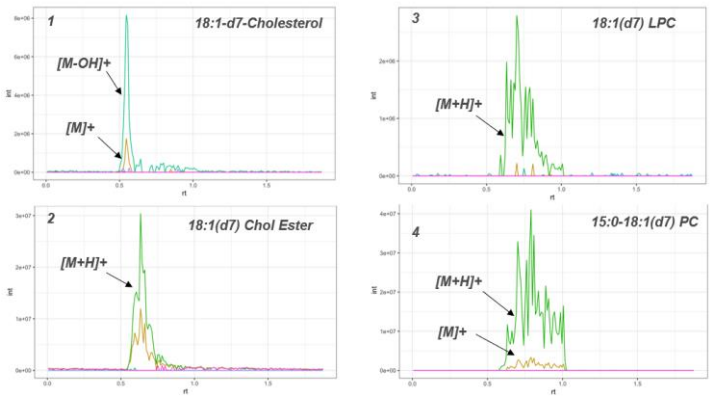


Figure 5: TIC of the Polar Lipid Standard with several "peaks" labelled to indicate the four compounds and their differentiating desorption times. The corresponding compounds and EICs can be seen below, along with the ionization provided.

Once we had concluded that the possibility of lipid detection, with a wide range of lipid classes, on the glass slide was possible in a targeted manner, we moved on to a more untargeted, more complex matrix, fingerprint residues. In general, we were able to produce a nice array of fingerprint profiles (Figure 6). Furthermore, we were able to determine which area of the spectra seemed to conserve the most information about the unique features for each person and which part of the spectra is the least conserved.

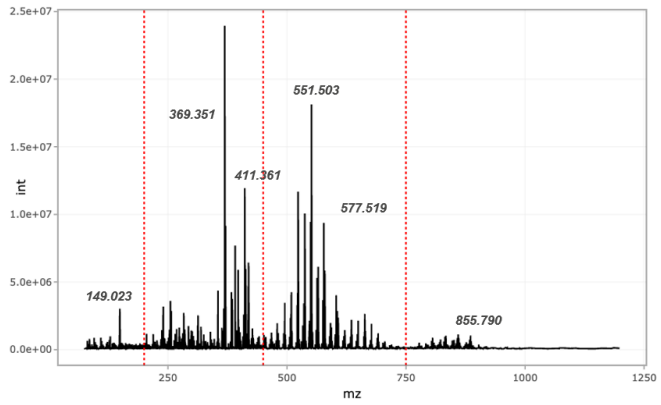


Figure 6: MS1 of a single fingerprint desorption

Similar to the outputs of the machine learning models that are fully described in the paper, we found that overall, the fingerprints are unique, however, the DAG region appears to be the most conserved between individuals, independent of which day or which hand was sampled. This can be seen in the discussion of the forensic paper. Whereas the volatile region had the most variety, which makes sense, since this area is clearly dependent on what you have been interacting with on a daily basis and can even depend on which hand is dominant. This relationship can be seen in using rather simplistic metrics, such as the spectra cosine similarity. Here, we were able to determine, essentially what the more complex algorithms had alluded to (Figure 7).

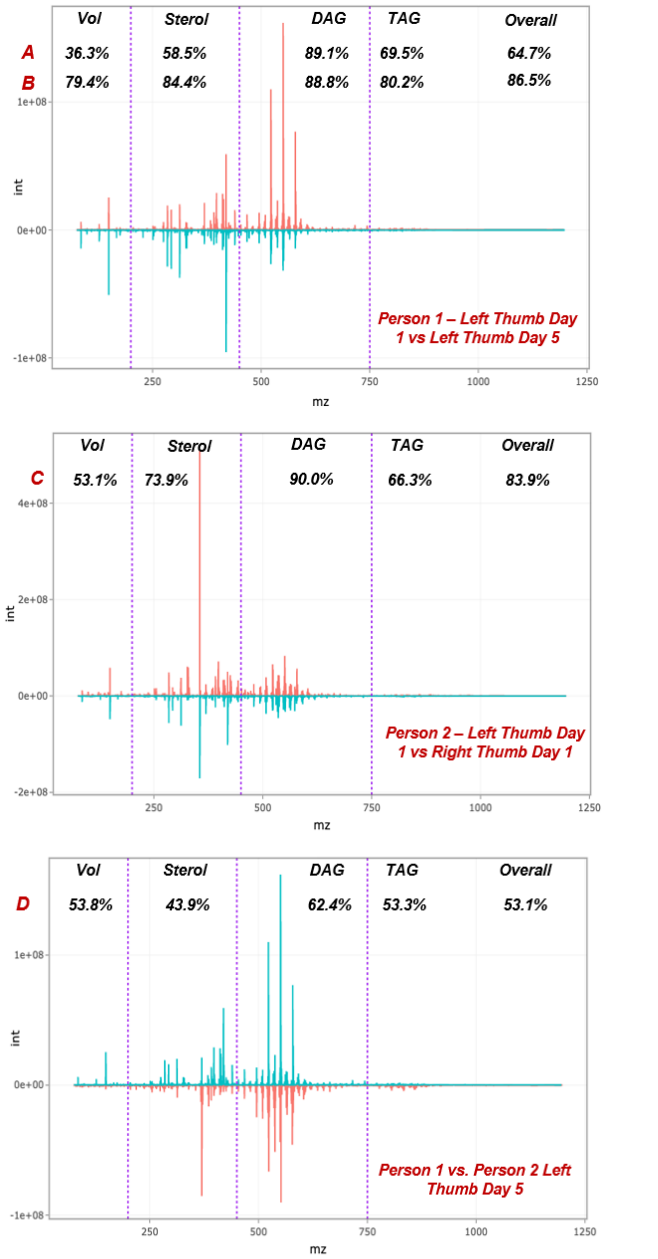


Figure 7: Visual comparisons of the fingerprint MS1. Additionally, several cosine similarity metrics are shown, broken down by each region and then with the overall similarity score: A. Person 1 – Left Day 1 vs Left Day 5, B. Person 1 – Left Day 1 vs Left Day 1, C. Person 2 – Left Day 1 vs Right Day 1, D. Person 1 Left vs Person 2 Left

Am Mittleren Moos 48
D-86167 Augsburg
www.plasmion.com

T: +49 (0) 821 20713355
E: info@plasmion.de

In this example, the overall similarity metric shows that the replicate similarity is close to 90%, while the similarity between the right and left replicates remains above 80%. The overall similarity does seem to fall between replicates taken on different days, in this case to about 63% similarity, which makes sense when looking at the volatile region, which is the most dissimilar. Finally, there is a reasonable difference between two different individuals for the same hand on the same day, even without a machine learning pipeline, of about 50% similarity. What is the most interesting about these metrics is for every comparison, the DAG region is the best at differentiating between hands and people, but conserves the information between replicates, even if several days apart (Figure 7).

Once we had established that we can preliminarily differentiate between fingerprints and look to the lipid-region for further information, it was time to take a deeper dive into the MS1 and look at the preliminary biomarker assignment.

What we generate from this biomarker database search are the particular features that could be the most interesting to focus on, filtering out those masses that simply don't have the proper mass defect to be a metabolite, and the search provides an array of potential compound candidates. For instance, when only looking at 5 main ionizations that come from the search, we are able to reduce the 57,000 features down to 330 unique features of interest, with 27,000 potential compound matches. The list of 27,000 compounds is then matched in ClassyFire to the molecular taxonomy, allowing us to understand the types of compounds that may be present and if these make sense. This simple search provides a significant reduction in feature analysis and the algorithm can be applied to any database or hand-curated list.

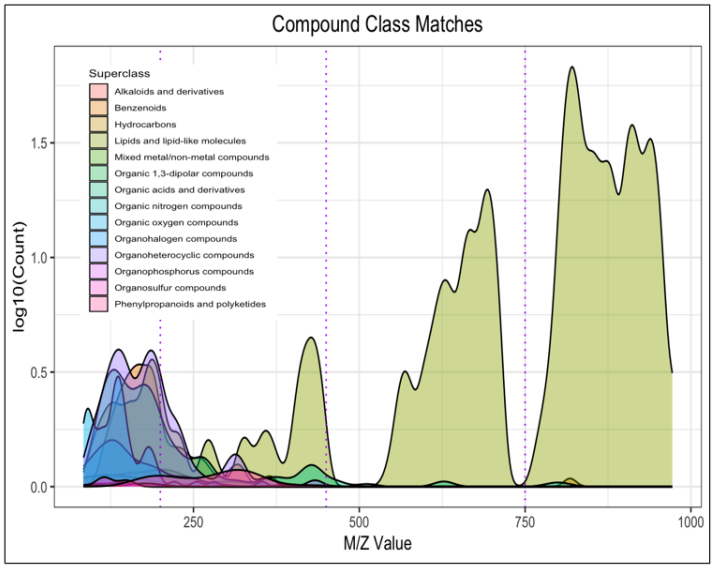


Figure 8: Density plot for every single unique mass across all samples that matched to a biomarker with the [M+H]+ ionization displayed by their Superclass given through ClassyFire

In this case, we further narrowed the search space to only [M+H]+ ionizations, resulting in 8,000 non-unique compound matches from 214 unique m/z features. In the volatile region, we observed a wide range of Superclasses, while in the DAG and TAG regions, lipid and lipid-like compounds dominated, along with a few hydrocarbons. This distribution of compounds passed the "sanity check." The more fascinating aspect, in terms of lipidomic analysis, was the distribution of lipid and lipid-like compounds (Figure 9).

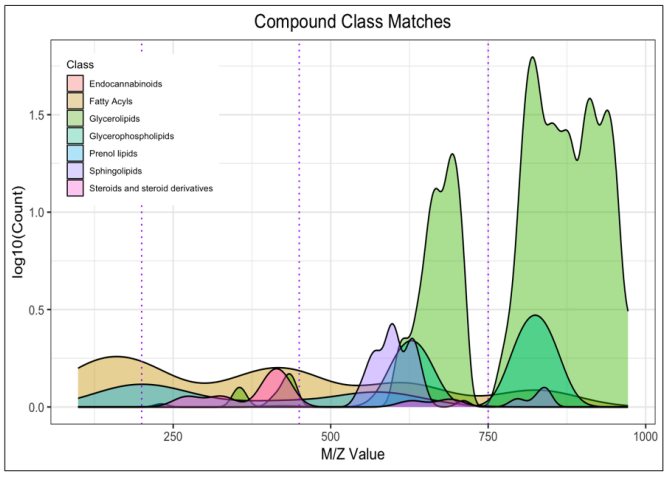


Figure 9: Density plot for every single unique mass across all samples that matched to a biomarker in the Lipid and Lipid-like Superclass with the [M+H]+ ionization displayed by their Lipid Class given through ClassyFire

Am Mittleren Moos 48
D-86167 Augsburg
www.plasmion.com

T: +49 (0) 821 20713355
E: info@plasmion.de

With a total of 6,800 matches to the lipid and lipid-like superclass out of the total 8,000 matches, this superclass was by far the majority of hits. This also makes sense, since most compounds that make-up the oil that coats the skin are lipids or lipid-like. From the literature there are several classes of lipids that are of interest when looking at the residues released from the skin, such as the Fatty Acyls, the Steroids or Hormone-like Compounds, Glycerolipids, and the Sphingolipids (including the ceramides). All these classes were identified within the features matched to the database, and their distribution across the MS1 mass range is clearly shown in Figure 9. There were 346 matches to the Fatty Acyls, 54 matches to steroids, 122 Sphingolipids, and a large amount of Glycerolipid matches of close to 6,000. This large number accounts for all the isomers that come with the DAG and TAG region, and this naturally inflates the number of matches per m/z value in that range.

With this information, we can begin to analyze the differences in class distribution between individuals that we had previously measured. In order to do this, the spectra for all replicates, for both hands, across both testing days were averaged for every subject tested, to obtain a lipid profile for each person. While many of classes were found in each individual, the differentiation lay in the magnitude of each compound with the class. This means that even if two individuals had the same compounds found, those values would not be the same, in fact, none of the distributions across any of the individuals are identical (Figure 10).

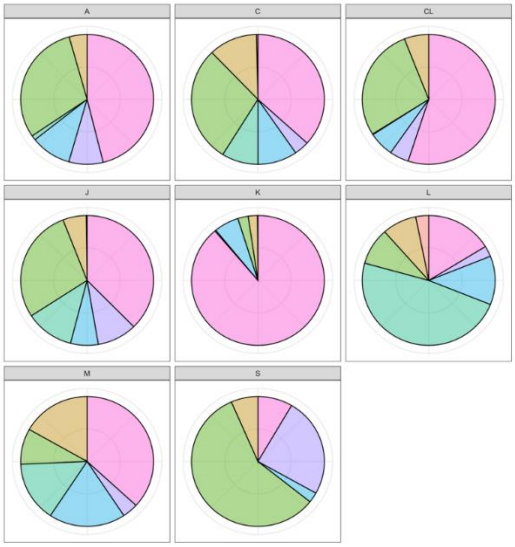
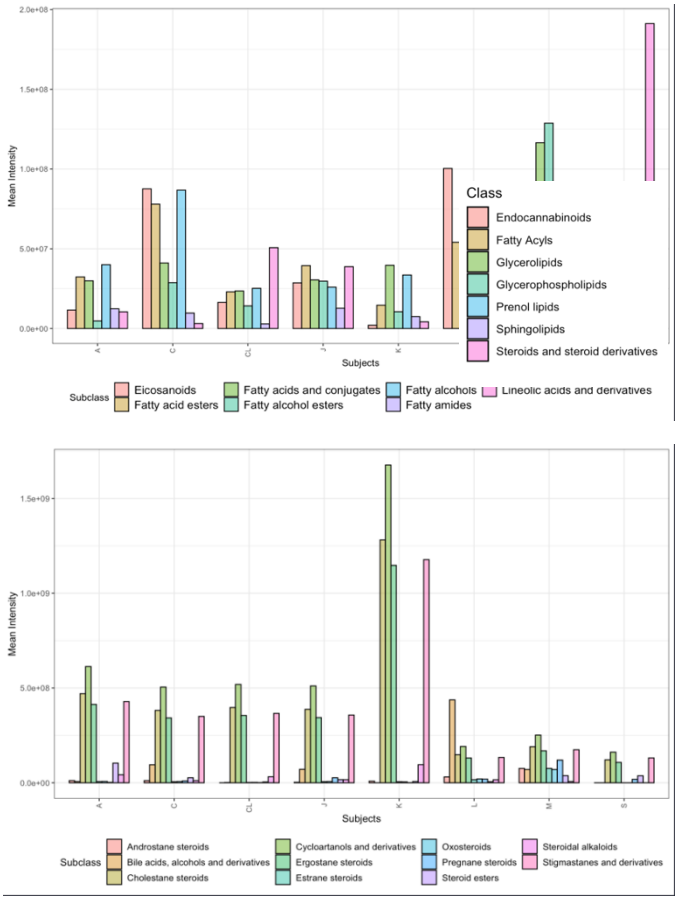


Figure 10: Distribution of each lipid class, broken out by each subject's averaged spectra across all days and both hands combined.

Furthermore, when analyzing the subclasses within each of these lipid classes, we are able to focus in on the types of compounds within the targeted lipid classes that are used to analyze the skin (Fatty Acyls, Glycerolipids, Sphingolipids, and Steroids). What we can show is that the ratios for each of these are unique and allow for unique profiles to be created for each individual (Figure 11).



Am Mittleren Moos 48
D-86167 Augsburg
www.plasmion.com

T: +49 (0) 821 20713355
E: info@plasmion.de

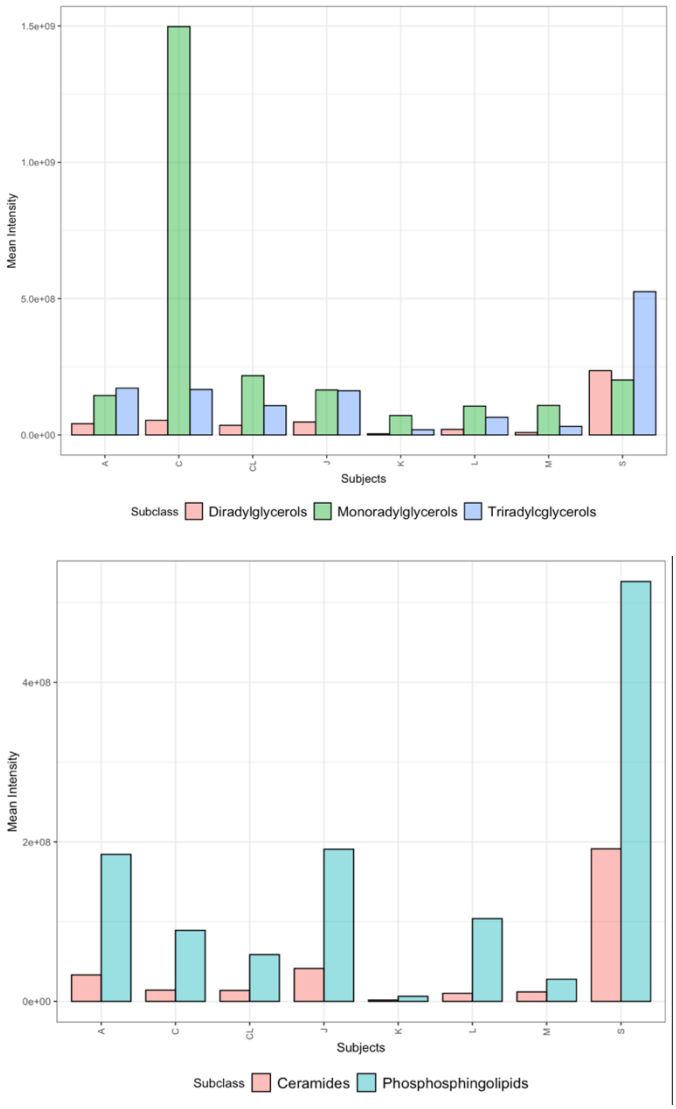


Figure 11: Ratios of each subject for each subclass pertaining to the following Lipid Classes (Top to Bottom): Steroids, Fatty Acyls, Glycerolipids, and Sphingolipids

With these tentative lipid assignments, we can begin to understand the fuller picture of what these MS1 spectra contain and how, in the future, these methods could be used to differentiate fingerprints through the MS1 profile, while determining which classes of compounds or compounds themselves contribute to these large differences, allowing us to elucidate metabolomic patterns of interest, going beyond the conventional forensic fingerprint identification.

Conclusion

Overall, this study shows the capabilities of the TD-SICRIT® -MS and how they can be utilized in the world of metabolomics and lipidomics, with the full potential of such a study being realized through the combination of chemistry and computation.

With such a device and pipeline, you can go beyond the conventional forensic uses, without imaging or sample prep or only one main mechanism of ionization or significant matrix effects. This allows you to extend the range of ionizable metabolites, and in turn, expanding the level of features that are possible to visualize, to derive a clear chemical profile of a biological matrix, like a fingerprint, without worrying about sample prep destroying valuable metabolomic information or that the matrix effect will fully suppress the desired outcome. Furthermore, you can take this expansive list of features and with a simple database pipeline for feature filtering, you can focus in on those features that have potential biological significance.

In the end, you have a device and computation pipeline that allows you to sequester all of the polar and non-polar compounds that allow for exposomic, lipidomic, forensic, or metabolomic studies, it is an all-in-one prototype that could allow us to change the way we view biological matrices and the information that lies within.

Am Mittleren Moos 48
D-86167 Augsburg
www.plasmion.com

T: +49 (0) 821 20713355
E: info@plasmion.de

Supplementary Information

Compound	Standard	Absolute amount / ng	Observed	DAG Observed
Trilinolenin	Non-Polar Standard	45	Yes	Yes
Trilinolein	Non-Polar Standard	105	Yes	Yes
Tripalmitin	Non-Polar Standard	20	Yes	Yes
Triolein	Non-Polar Standard	305	Yes	Yes
Tristearin	Non-Polar Standard	10	Yes	Yes
Trieicosenoin	Non-Polar Standard	5	Yes	Yes
Trierucin	Non-Polar Standard	5	Yes	Yes
Triarachidin	Non-Polar Standard	5	Yes	Yes
15:0-18:1(d7) PC	Polar Standard	80	Yes	N/A
15:0-18:1(d7) PE	Polar Standard	2.5	Yes	N/A
15:0-18:1(d7) PA	Polar Standard	3.5	Yes	N/A
18:1(d7) LPC	Polar Standard	12.5	Yes	N/A
18:1(d7) LPE	Polar Standard	2.5	Yes	N/A
18:1(d9) SM	Polar Standard	15	Yes	N/A
18:1(d7) Chol Ester	Polar Standard	175	Yes	N/A
18:1(d7) MG	Polar Standard	1	Yes	N/A
15:0-18:1(d7) DG	Polar Standard	5	Yes	N/A
15:0-18:1(d7)-15:0 TG	Polar Standard	27.5	Yes	Yes
15:0-18:1(d7) PS	Polar Standard	2.5	No	N/A
15:0-18:1(d7) PI	Polar Standard	5	No	N/A
15:0-18:1(d7) PG	Polar Standard	15	No	N/A

Table 1: Compounds in each standard with their absolute amounts and if they were found, along with any fragmentation that could occur in the form of a DAG.

References

1. Conway, C., Weber, M., Ferranti, A., Wolf, J., & Haisch, C. (2023). Rapid desorption and analysis for illicit drugs and chemical profiling of fingerprints by SICRIT® ion source. Drug Testing and Analysis. https://doi.org/10.1002/dta.3623

2. Djoumbou Feunang, Y., Eisner, R., Knox, C., Chepelev, L., Hastings, J., Owen, G., Fahy, E., Steinbeck, C., Subramanian, S., Bolton, E., Greiner, R., & Wishart, D. S. (2016). Classyfire: Automated Chemical Classification with a comprehensive, computable taxonomy. Journal of Cheminformatics, 8(1). https://doi.org/10.1186/s13321-016-0174-y

Am Mittleren Moos 48
D-86167 Augsburg
www.plasmion.com

T: +49 (0) 821 20713355
E: info@plasmion.de