

1

2

3 **Rapid and Automatic Annotation of Multiple On-Tissue Chemical**

4 **Modifications in Mass Spectrometry Imaging with METASPACE**

5

6

7 Evan A. Larson^a, Trevor T. Forsman^a, Lachlan Stuart^b, Theodore Alexandrov^{b,c}, Young Jin Lee^a

8

9

10 a. Department of Chemistry, Iowa State University, IA 50011, USA

11 b. Structural and Computational Biology Unit, European Molecular Biology Laboratory

12 (EMBL), 69117, Heidelberg, Germany

13 c. Molecular Medicine Partnership Unit, EMBL, 69117, Heidelberg, Germany

14

15

16

17

18

19 Key Words: On-Tissue Derivatization, Metabolomics, Mass Spectrometry Imaging,

20 METASPACE, Metabolite Annotation

21

22 * To whom correspondence should be addressed. E-mail: yjlee@iastate.edu

23

Abstract

On-tissue chemical derivatization is a valuable tool for expanding compound coverage in untargeted metabolomics studies with matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI). Applying multiple derivatization agents in parallel increases metabolite coverage even further but results in large and more complex datasets that can be challenging to analyze. In this work, we present a pipeline to provide rigorous annotations for on-tissue derivatized MSI data using METASPACE. To test and validate the pipeline, maize roots were used as a model system to obtain MSI datasets after chemical derivatization with four different reagents, Girard's T and P for carbonyl groups, coniferyl aldehyde for primary amines and 2-picolyamine for carboxylic acids. Using this pipeline helped us annotate 631 unique metabolites from the CornCyc/BraChem database compared to 256 in the underivatized dataset yet at the same time, shortening the processing time compared to manual processing and providing robust and systematic scoring and annotation. We have also developed a method to remove false derivatized annotations, which can clean 5-25% of false derivatized annotations from the derivatized data, depending on the reagent. Taken together, our pipeline facilitates the use of broadly targeted spatial metabolomics using multiple derivatization reagents.

Introduction

Untargeted metabolomic analysis using mass spectrometry has become an important tool to better understand the mechanisms present in biological systems¹. However, these studies present a unique challenge in metabolite identification, or finding metabolites represented in mass spectrometry data, exacerbated by the large size of metabolome. The Human Metabolome Database contains over 110,000 metabolites^{2,3}. Plants have an even larger metabolome, estimated to be around 400,000 unique compounds which are responsible for numerous functions within a plant and within the plant's environment⁴. The chemical diversity of each metabolome contributes to the challenge of metabolite detection and identification. Peptides, amino acids, carbohydrates, acids, and lipids can comprise the metabolites present in a biological system, each with unique chemical properties and functionalities which makes comprehensive identification difficult⁵. Therefore, increasing the compound coverage of analytical techniques is paramount to better understanding the complex metabolome of biological species.

1 In addition to identifying compounds that are present, determining the spatial localization
2 of these compounds is necessary to better understand biological functions. Spatial metabolomics
3 using mass spectrometry imaging (MSI) has become a powerful tool to address this issue. Among
4 the MSI techniques, matrix-assisted laser desorption/ionization (MALDI) MSI is appealing due to
5 its high spatial resolution, a wide range of detectable molecules, and a variety of commercially
6 available instrumentation. One critical bottleneck in MSI is limited compound coverage due to the
7 low ionization efficiency of certain compounds and inherently small sampling size. The technique
8 of on-tissue chemical derivatization applied prior to MALDI imaging has shown to be a useful
9 strategy as it converts poorly ionizing compounds to positively charged or highly ionizable
10 compounds, therefore dramatically increasing their signals. Most applications of on-tissue
11 derivatization have focused on targeted MSI for specific chemical compounds. Girard's T (GT)
12 has been used to derivatize plant hormones in bean seeds⁶, triamcinolone acetonide in human
13 osteoarthritic tissue⁷, glycans and oligosaccharides^{8,9}, lactones in gram negative bacteria¹⁰, and
14 steroids in mice or rat tissues^{11,12}. Girard's P (GP) has been utilized to for detecting cholesterol
15 and derivatives in mouse brains¹³, sialylated oligosaccharides in human milk¹⁴, and N-glycans in
16 human cancer tissue¹⁵. Coniferyl aldehyde (CA) has been used for derivatizing primary amine
17 containing compounds such as amino acids, neurotransmitters, and short peptides¹⁶⁻¹⁹. 2-
18 Picolylamine (PA) has been applied to derivatize fatty acids in rat brain tissue²⁰.

19 METASPACE is a web-based platform for untargeted spatial metabolomics and able to
20 automatically and systematically perform metabolite and lipid annotation of high mass resolution
21 MSI data. The quality of the annotation is estimated through the metabolite score match score
22 (MSM), calculated as a product of several measures assessing the data quality as well as how it
23 matches to theoretically predicted properties including theoretical isotope patterns. The false
24 discovery rate (FDR) of the produced list of metabolite annotations is then estimated following the
25 target-decoy strategy widely used in other omics using a target database (including ions from a
26 metabolome of interest, considering plausible adducts e.g. H^+ , Na^+ , or K^+) and a decoy database
27 (including implausible ions calculated for the same metabolome e.g. B^+ , Db^+ , or Ag^+)²¹.
28 METASPACE additionally uses a machine learning algorithm to identify off-sample localized
29 features and removes them²². This platform has been applied to many applications such as studying
30 metabolite distribution in diabetic kidney tissue²³, microbial metabolites²⁴, lipid composition of
31 demyelinated mouse spinal cord²⁵, metabolite distribution in whole zebrafish²⁶, and N-glycans in

human kidney and mouse lung tissue²⁷. Additionally, METASPACE annotation quality has been used as an optimization metric for steps in the MALDI imaging workflow such as tissue storage²⁸, MALDI matrix spraying parameters²⁹, or post calibration of imaging datasets³⁰. Finally, METASPACE is designed to easily share MSI data and has become a popular public repository.

Recently, our group has proposed the use of multiple on-tissue derivatization reagents on serial tissue sections to dramatically increase the compound coverage in semi-targeted MALDI imaging^{31,32}. The challenge of these studies is the data analysis and interpretation, as it may take weeks to manually analyze multiple replicates of each derivatization reaction. Another challenge is the analyst bias and potential for human error which can result in missed or incorrect annotations. METASPACE was originally developed for the annotation and analysis of underivatized MSI data²¹. Here we present and make publicly available the novel functionality of METASPACE to support chemical derivatization so that chemically derivatized MS imaging datasets can be automatically annotated and scored with FDR. Previously, there has been no automatic annotation method for derivatized MSI data and any chemically derivatized MSI data had to be manually analyzed. Now, during the data upload or reprocessing, the submitter can enter the expected chemical modification which will be used for annotation. The focus of this work is to demonstrate this new functionality in METASPACE and, furthermore, propose a systematic workflow to identify and remove false positive annotations. Maize roots were used as a model system with the derivatization using four sets of chemical reagents.

Materials and Methods.

Brief Experimental Details.

Full details of the experimental section are found in the supporting information. Sample preparation was performed based on our previous work.^{31,32} In brief, B109 maize roots were grown in a damp paper towel to 10-11 cm, embedded in 10% (w/v) gelatin and cryosectioned at 2 cm away from the seed with 20 μ m thickness.³¹ Tissue sections were then dried down and derivatization reagents, additives and matrices were applied via TM sprayer (HTX Technologies, Chapel Hill, NC). Data was collected on a MALDI source (MALDI Injector; Spectrograph, Kennewick, WA) coupled to an Orbitrap mass spectrometer (QExactive HF; Thermo Fisher

Scientific, San Jose, CA). Three replicates of each condition were imaged at a 20 μm raster step at the mass resolution of 120,000 at m/z 200 for a scan range of m/z 100 – 1000. Once collected, data was converted to imzML and uploaded to METASPACE (<https://metaspace2020.eu/>). LC-MS(/MS) was collected to support annotations.

METASPACE Data Analysis.

A total of 18 MSI datasets were uploaded to METASPACE: three replicates of four derivatizations (GP, GT, CA, PA) and two controls (positive and negative mode). We annotated data against three metabolite databases on METASPACE: BraChem/CornCyc, LipidMaps (2017-12-12) and ChEBI (2018-01). BraChem/CornCyc database is a custom combined database of BraChem (2018-01) and CornCyc (v9, 2018-07). For the control datasets, we considered $[\text{M}+\text{H}]^+$ and $[\text{M}+\text{K}]^+$ adducts in positive mode and deprotonated molecule, $[\text{M}-\text{H}]^-$, in negative mode. For the derivatized datasets, the adducts were selected dependent on the derivatization reaction as follows. GP and GT have permanent positive charges we considered the $[\text{M}]^+$ adduct and entered the chemical modification as ‘+C₇H₁₀N₃O-H₂O’ (or ‘+C₇H₈N₃’ or + 134.0718 Da) for GP and ‘+C₅H₁₄N₃O-H₂O’ (or ‘+C₅H₁₂N₃’ or + 114.1031 Da) for GT. CA and PA have no permanent charge so we considered $[\text{M}+\text{H}]^+$ and $[\text{M}+\text{K}]^+$ adducts with the chemical modifications entered as ‘+C₁₀H₁₀O₃-H₂O’ (or ‘+C₁₀H₈O₂’ or + 160.0524 Da) for CA and ‘+C₆H₈N₂-H₂O’ (or + 90.0581 Da) for PA. All METASPACE processed data is publicly available in the devoted project: https://metaspace2020.eu/project/MaizeB109_Roots_Deriv_2022. There are more annotated datasets in the project folder than the original 18 imzML datasets, a total of 48, because multiple METASPACE analyses were performed on each dataset, including different database searches or false derivatized annotation searches. A description of each dataset filename is found in **Supplementary Table 1**.

Results and Discussion.

Annotating Chemically-Derivatized Metabolites using METASPACE.

The workflows of two approaches to annotate chemically-derivatized MSI data are compared in **Figure 1**: the manual conventional approach and METASPACE-based semi-

automatic approach proposed in this work. Without an automatic annotation tool, the typical untargeted derivatized MSI workflow (**Figure 1a**) requires four major steps: 1) manual extractions of region-of-interest (ROI) features, 2) identification of ROI features unique to the derivatized sample by comparing to a control, 3) identification of tissue-localized features by generating images and manual filtering, and 4) manual database searching after subtracting the derivatization mass from each feature. Performing these steps is time-consuming and is subject to analyst biases which can often lead to missed or incorrect annotations. Depending on the size of data sets and the number of replicates, it may take at least a few days if not weeks to analyze the entire dataset. METASPACE provides an automated, rigorous, and efficient method to analyze derivatized MS imaging data using the chemical modification tool (**Figure 1b**). Steps 1, 3, and 4 above are greatly simplified as all are performed automatically by METASPACE, systematically providing annotations for the derivatized and underivatized data in minutes per dataset. The chemical modification tool accounts for the change in molecular formula corresponding to an applied derivatization reagent (e.g., the formula of chemical reagent) and annotates the “derivatized” molecular formulas (formulas where the chemical composition is updated accounting for the derivatization agent). A screenshot of METASPACE’s data submission page with the chemical modification tool highlighted is shown in **Figure S1**. After selecting signals for such derivatized molecular formulas, the rest of the process follows regular METASPACE annotation, including calculations of the MSM scores and FDR. However, step 2 in manual analysis is currently not implemented in METASPACE, which may result in false derivatized annotations by annotating underivatized compounds as if they were derivatized. These unintended false positives, however, can be removed by comparing the same analysis on MSI data of underivatized tissue as discussed in the next subsection.

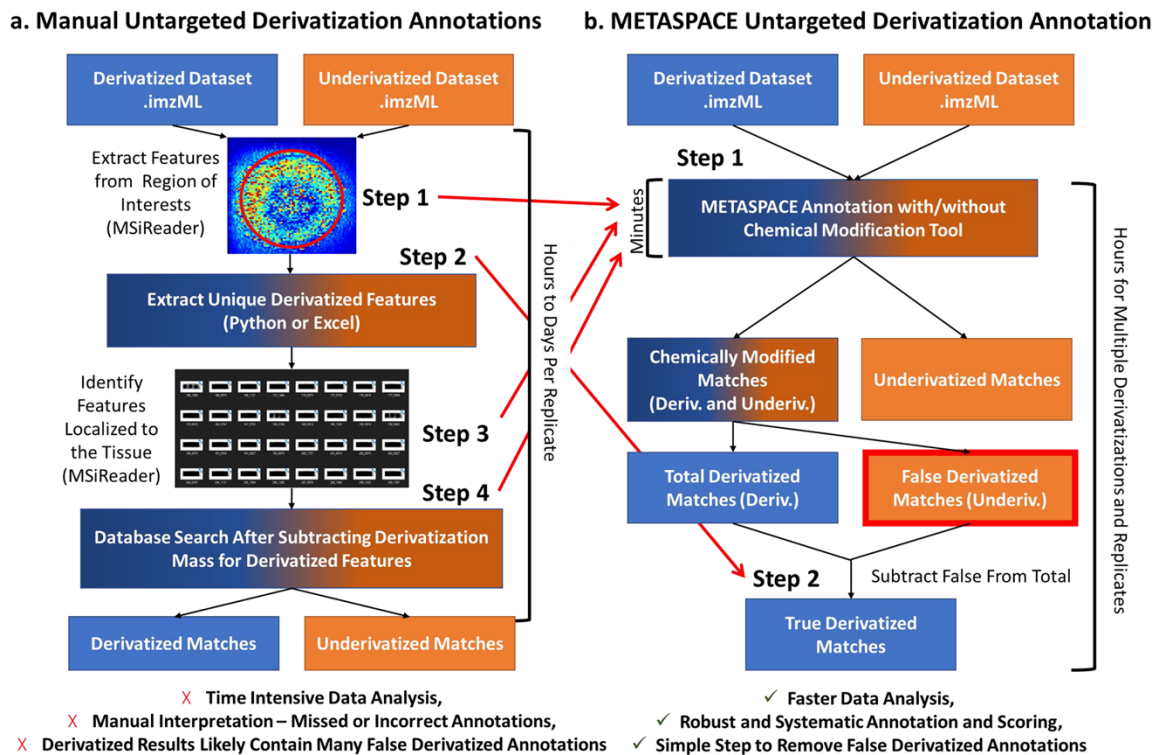


Figure 1. Workflow for (a) manual untargeted analysis of chemically derivatized MSI data and (b) automatic annotation proposed in this work using METASPACE including the removal of false derivatized annotations. Blue and orange color indicate the derivatization status (orange for underivatized, blue for derivatized) for the respective dataset, analysis step, or result.

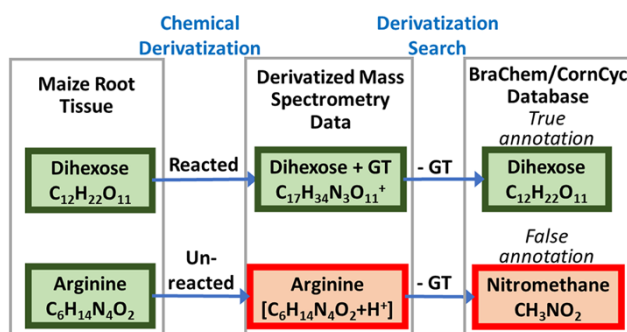
To evaluate the proposed METASPACE workflow, four different derivatization reagents were applied to maize root cross-sections: CA for primary amines, PA for carboxylic acids, GP and GT for carbonyls (**Figure S2**). Three replicates of each on-tissue modification and underivatized control were searched in three databases, BraChem/CornCyc³³, LipidMaps³⁴ and ChEBI³⁵. BraChem (Brassica Napus Database) is a database from an LC-MS/MS study of rapeseed totaling roughly 11,000 metabolites with ~5,000 unique formulas. To supplement the BraChem database, it was combined with the CornCyc database which is a compendium of maize specific metabolites and metabolic pathways with ~2,500 metabolites and ~1,750 unique formulas. This combination was uploaded as a publicly available custom database on METASPACE. LipidMaps is a comprehensive lipid database containing both experimentally identified lipids as well as computationally predicted lipids with the version on METASPACE totaling 42,022 lipids with 7,354 unique molecular formulas. ChEBI contains a broad range of natural metabolites from

various organisms as well as synthetic compounds that are biologically relevant, with the version on METASPACE totaling 34,748 metabolites with 13,505 unique molecular formulas. Only annotations from these databases which are present in two or three replicates are included in the following discussion. The results of the chemical modification search using METASPACE are discussed below and used to validate the tool as well as determine the improvements to the results.

False Annotations in Chemical Derivatization.

As with any untargeted metabolomics analysis, false positive matching is unavoidable in metabolite annotation. METASPACE minimizes this effect by calculating FDR by comparing MSM scores between annotations against plausible ions (target database) and implausible ions (decoy database). However, chemical derivatization analysis represents a specific challenge as it produces another type of false positives which we call “false derivatized annotations”. This can happen due to the fact that on-tissue chemical derivatization reactions do not produce 100% yield even in the most effective scenarios, because 1) the reaction time is limited as the solvent evaporates quickly, and 2) only metabolites extracted by solvent may react with the reagent. Thus, a “false derivatized annotation” occurs when an unmodified metabolite, which is present in the sample and detected by mass spectrometry, is isomeric to a molecular formula constructed by applying the derivatization reaction to another molecule from the target database. Additionally, any compounds which are not targeted by the derivatization reagent may also not react. **Figure 2** shows an example of true and false derivatized annotations. Chemical derivatization is properly considered in derivatization search for dihexose resulting in true annotation; however, unreacted protonated arginine ($C_6H_{14}N_4O_2$) is isomeric to nitromethane (CH_3NO_2) that is theoretically derivatized by GT, $[CH_3NO_2 + C_5H_{14}N_3O - H_2O]^+ = [C_6H_{14}N_4O_2 + H]^+$, resulting in false annotation. As each derivatization agent used in this work contains common functional groups made of biologically common atoms ($C_xH_yO_zN_w$), a molecular formula of a potentially derivatized metabolite is often isomeric to some other molecule from the target database. Additionally, besides isomeric matches between underivatized and derivatized molecular formulas, there can be also isobaric matches within the 3 ppm tolerance used in METASPACE. In manual untargeted derivatization annotations (**Figure 1a**), step 2 extracts only unique derivatized features by

1 comparing derivatized and underivatized datasets, typically done manually using an Excel sheet
 2 or semi-automatically using an in-house Python code.



3
 4 **Figure 2. Illustration of derivatized annotations.** Schematic for annotation of true and false
 5 derivatized annotations. Dihexose is present in maize root tissue, derivatized and annotated
 6 correctly. Arginine is not derivatized but results in a false derivatized annotation after matching to
 7 a database compound, nitromethane, after the derivatized search.

8 The chemical modification tool of METASPACE cannot solve this issue by analyzing the
 9 derivatized dataset alone; however, the same unreacted metabolites are also present in the control
 10 and the corresponding false annotation can be removed by analyzing the underivatized control
 11 samples as if they are derivatized. **Figure S3** further illustrates this process. If there is no matching
 12 in derivatization search of non-reacted compounds, they will result in only true derivatized
 13 annotations (**Figure S3a**). However, false derivatized annotation can occur when there happen to
 14 be a matching such as in arginine in search for GT derivatization (**Figure S3b**). To remove the
 15 false derivatized annotations, underivatized control data can be analyzed as if it is derivatized
 16 (**Figure S3c**). Any formula matches thus detected are all false derivatized annotations as there is
 17 no derivatization and can be removed from the derivatized dataset. In this work, this was
 18 accomplished by downloading annotations from METASPACE and comparing them in Excel and
 19 removing any overlap between the derivatized dataset and the derivatized search in the
 20 underivatized dataset.

21 This approach is expected to remove almost all of the false derivatized annotations. **Figure**
 22 **3** shows the percentage of peaks filtered out as false derivatized annotations for each derivatization
 23 reagent when searched against three metabolite databases. FDR 20% was used for all annotations.
 24 PA has a lower false derivatized annotation rate than the other three derivatization reagents when

searched against the largest database (ChEBI), likely due to the unique structure (i.e., pyridine ring). However, there are more false annotations (16.3%) in plant database (BraChem/CornCyc) than in a more general database such as ChEBI (5.3%). It may be due to plants have many more metabolites containing similar chemical formula with PA. CA has a very high false derivatized annotation rate when searched against LipidMaps (25%) than other databases because of the lack of tentative positives, only 18. This analysis gives an insight on how to improve true annotations in untargeted analysis using chemical derivatization; the use of derivatization reagents with unique formulas or functionality would alleviate false derivatized annotations from the start.

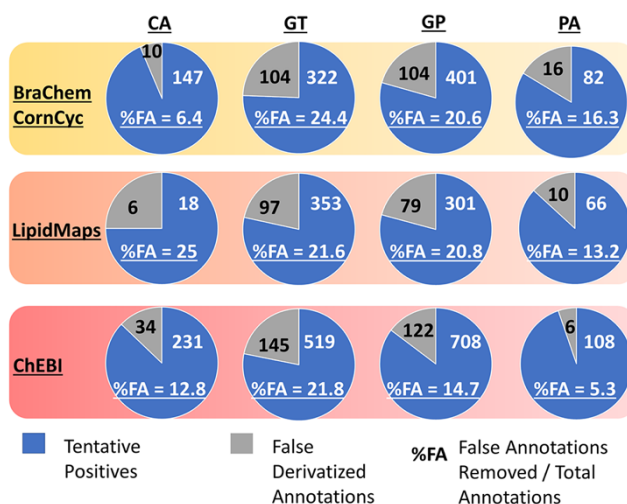


Figure 3. Pie chart for the number of true and false derivatized annotations in METASPACE analysis of derivatized maize root sections at 20% FDR.

The number of matches from the BraChem/CornCyc database after false derivatized annotation removal are shown in **Figure 4a**. The matches for underivatized control samples in positive and negative mode with DHB and DAN matrix, respectively, are compared with the derivatized sample sets. GP and GT have much higher numbers of matching at low FDR than the control, CA or PA modification. When using BraChem/CornCyc database with FDR 10% (**Figure 4a**), for example, the number of matches is 248 and 302 for GP and GT, respectively, but it is only 39, 84, 43, and 24 for positive control, negative control, CA, and PA, respectively. It is because GP and GT dramatically improved derivatized ion signals compared to CA and PA, which not only allowed for annotation but also improved quality of molecular signals (and thus higher numbers of annotations at the same FDR) due to high-quality isotope images. An example is $C_{28}H_{46}O$ and $C_{29}H_{46}O$, matched to campest-4-en-3-one and avenastenone in the CornCyc database,

in which they have a ~1,000 times signal increase for GT and ~200 times signal increase for GP compared to the control. Low matching with the CA modification at a low FDR is attributed to fewer compounds with primary amines in maize roots and less ion signal improvement with CA. Fewer high-quality matches for CA and PA can also be attributed to the lack of a permanent positive charge limiting the signal improvement compared to GP and GT. **Figure S4** shows matches for all three databases before and after false derivatization removal. As expected, ChEBI has the most matches at 20% FDR as it is the largest database (**Figure S4**); however, BraChem/CornCyc, the only plant specific database, has high-quality matching with comparable or higher number of matches at 5 or 10% FDR, especially for GP and GT.

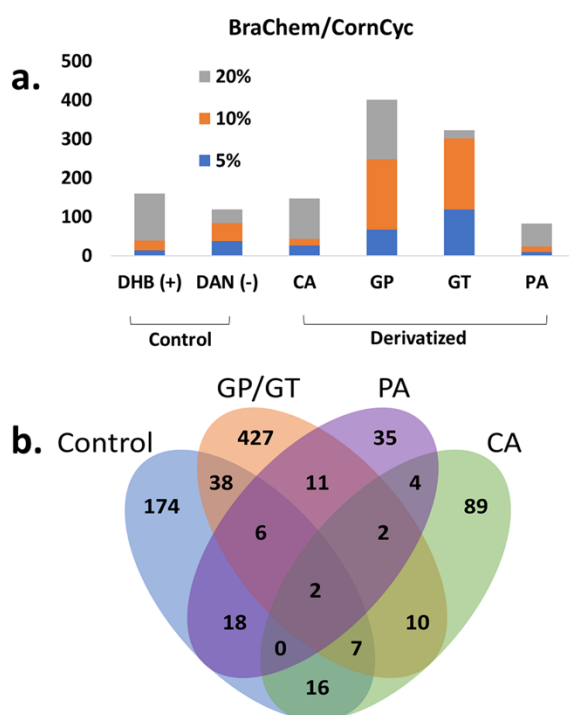


Figure 4. (A) Number of annotations after the removal of false derivatized annotation for the BraChem/CornCyc database at the FDR cutoffs of 5%, 10%, and 20%. **(B)** A Venn diagram showing the overlap in unique formulas annotated from the BraChem/CornCyc at 20% FDR. Girard's T and Girard's P were combined as they both target carbonyls.

Figure 4b shows a Venn diagram comparing the METASPACE annotated metabolite features between the data sets when searched against BraChem/CornCyc database at 20% FDR. There is only a little overlap in matches between underivatized and derivatized data highlighting the number of new features observed through derivatization. Combining both positive and negative

mode, a total of 261 annotations are made in control whereas 450, 52, and 105 additional annotations are made by the derivatization with GP/GT, PA, and CA, respectively, resulting in a total of 578 annotations uniquely found only in chemical derivatization. The overlap between each derivatization is relatively small, suggesting the benefits of targeting multiple chemical functionalities. For example, 27% and 11% of PA derivatized annotations are also detected in GP/GT and CA annotations, respectively, and 16% and 6.1% of CA derivatized annotations are also found in GP/GT and PA annotations, respectively. There are 21 compounds which are annotated in both the CA dataset and either the PA or GP/GT datasets. It is possible that some of them might be due to side reactions. This highlights a challenge with chemical derivatization that highly reactive derivatization reagents, such as CA, may partially react with untargeted functional groups. We are systematically investigating potential side reactions of CA (work in progress). Further study is necessary to identify side reactions for each derivatization reagent as we have previously done for on-tissue boronic acid modifications³⁶, in which the METASPACE chemical derivatization tool will be useful to annotate side reactions.

Amino Acids and Other Tentative Matches with METASPACE.

Amino acids, being amphiprotic, have low ionization efficiency in MALDI-MS; however, CA works well to derivatize and increase the signal of amino acids. In this work, nineteen derivatized amino acids including gamma-aminobutyric acid (GABA) have sufficient signals to be visualized (**Figure 5**) compared to three in the underivatized control. Of the 19 amino acids, sixteen were matched in all three replicates and three were detected in two replicates. Fourteen of them have a low FDR, eight 5% and six 10% FDR, respectively. Two have a medium FDR of 20%, and three have high FDR of 50%. This wide range of FDR is due to the abundance difference of amino acid ion signals (**Figure S5A**), making amino acids a good system to investigate the efficiency of derivatization, how it contributes to the detection, and to annotation by METASPACE. Tyrosine was annotated as protonated and potassiated ions; however, the protonated ion is removed by METASPACE's on-sample algorithm as it is largely delocalized and likely isomeric or isobaric by a background signal (not shown). The MSM score, a primary score to calculate FDR, is a product of three different scores: ρ Chaos quantifies the level of spatial chaos (versus structure) in the monoisotopic ion image, ρ Spatial quantifies the spatial co-localization

1 between isotopic ion images, and ρ Spectral quantifies the similarity of the relative isotopic
2 intensities to the theoretically predicted isotope intensities for a molecule of interest. **Figure S5B**
3 presents each of the scores composing MSM for Asn, an amino acid detected with 5% FDR, and
4 Phe, an amino acid detected with 50% FDR. Asn has a much higher MSM (0.908) than Phe
5 (0.2145), but their ρ Spectral and ρ Chaos scores are both above 0.95, indicating ρ Spatial as the
6 major determinant for the MSM score (0.914 vs 0.225). While ρ Spectral also measures the
7 likelihood of chemical formula and affected by ion signals, ρ Spatial is more affected by the
8 detected molecular intensities. This is explained by the fact that isotopic ion images for low-
9 intensity molecules have low signal-to-noise ratio thus leading to a decrease in spatial co-
10 localization between isotopic ion images (calculated across all pixels). Some isotopes may even
11 become below the limit of detection in some pixels and nullified altogether. Generally, lower
12 abundance metabolites have higher FDRs and lower MSMs, primarily due to the ρ Spatial score.
13 However, like Phe, some low-scoring metabolite annotations which are expected to be present may
14 require additional investigation as they may be missed with too strict of an FDR filter. For the rest
15 of this manuscript, we will restrict the discussion to 20% FDR as a cutoff searched against
16 BraChem/CornCyc database as summarized in **Figure 4**, providing a higher sensitivity in
17 annotation compared to 10% FDR yet excluding more false positives compared to 50% FDR.

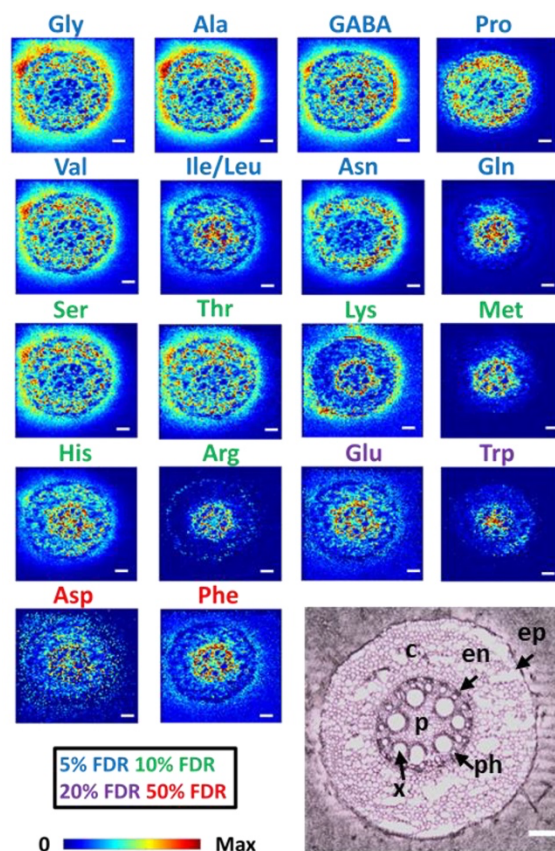


Figure 5. Heatmap images of 18 amino acids including GABA annotated with METASPACE in maize root sections. Tyrosine is significantly overlapping with a background ion signal and not shown. The names of amino acids are colored according to the FDR they were annotated with: blue for 5% FDR, green for 10% FDR, violet for 20% FDR, and red for 50% FDR. Maize root morphology is included for reference: ep, epidermis; c, cortex; p, pith; en, endodermis; ph, phloem; x, xylem. Scale bar is 200 μ m.

Interpreting the spatial localizations of the detected amino acids, we focus on the two morphological regions of the root: cortex and pith (see maize root morphology in **Figure 5**)¹⁹. Previously, we studied the genotypic difference of B73 and Mo17 species of maize as well as their hybrids. In this work we use B109 which is a cross of B73 and another Iowa strain, BS20(S)C1-73-1-1, and is phenotypically similar to B73³⁷. Comparing the amino acid localizations between B109 (this work) and B73 (previous study) leads to the following observations. It should be noted that the compounds in xylem or phloem are often difficult to differentiate from those in pith due to the delocalization during the sample prep or their low abundance. As they are all involved in

1 amino acid transportation or storage, we will combine them as pith in the following discussion. In
2 contrast, cortex is more involved in amino acid synthesis. The ion images for Gln and Leu/Ile are
3 similar between B73 and B109 in that they are mostly present in pith. Gly is also similar for being
4 present mostly in cortex. In contrast, Ala, Val, and Asn differ in their localization between B109
5 and B73. In B109, all three show localization in cortex. In B73, Ala and Val are localized mostly
6 in pith and Asn shows a homogeneous distribution. Further study would be necessary to interpret
7 these differences, but it may be explained by the hypothesis that B73 relies more on the transport
8 from seeds through pith. A few other amino acids have interesting localizations in B109 that we
9 could not detect in our previous work in B73 or Mo17, thanks mostly to higher ion signals in this
10 work achieved by using a better instrument and more optimized protocol; a total of 16 amino acids
11 were detected in the previous work compared to 19 detected in this work. Among the newly
12 detected amino acids, Lys is primarily localized in the epidermis and endodermis whereas Met,
13 Arg, and Trp are present in high abundance in the pith.

14 In addition to amino acids, numerous other metabolites were annotated by METASPACE.
15 **Figure S6** shows examples of ion images for each derivatization agent, obtained when using the
16 BraChem/CornCyc database with 20% FDR. A few of the unique localizations include m/z
17 281.150 in PA which is localized to the epidermis and endodermis. m/z 426.333 in GT and m/z
18 428.291 in GP are localized mostly to the endodermis. m/z 262.102, m/z 321.160, and m/z 481.101
19 in CA and m/z 596.442 in GP are all localized to the pith. m/z 385.139 and m/z 467.085 in CA, and
20 m/z 307.140 and m/z 443.155 in GT are all localized to the cortex. Finally, m/z 524.457 in GT and
21 m/z 544.426 in GP are localized to the phloem. Diverse localizations detected by several different
22 derivatization reactions further support the usefulness of adopting multiple chemical derivatization
23 strategy.

24 To cross-validate some of the annotated compounds, we performed an untargeted LC-MS
25 and LC-MS/MS analysis with a quadrupole-time of flight mass spectrometer for the solvent-based
26 extracts of roots grown in the same condition as the sectioned root. The accurate mass matching
27 was performed for LC-MS data using the MassProfiler software and searched against the
28 BraChem/CornCyc database. MassProfiler matches the accurate masses as well as isotope
29 abundances, similar to the ρ Spectral score used in METASPACE. LC-MS/MS data was analyzed
30 using MS-DIAL and matched against the MONA spectral library. A total of 257 annotations were

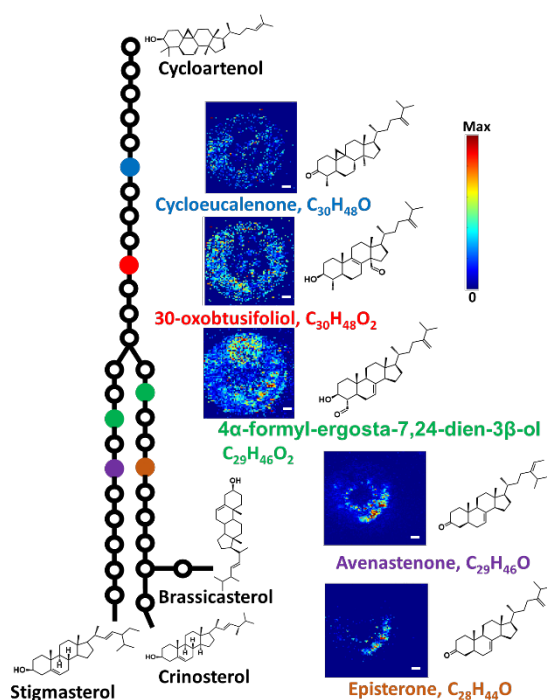
obtained for LC-MS/MS and 2078 annotations for LC-MS combining both polar and non-polar extracts in positive and negative mode at the cutoff score of 60 and 80, respectively. We expect the overlap in annotated compounds is limited between LC-MS(/MS) and MALDI imaging due to the differences in sample processing and ionization mechanisms, and more importantly the fact there is no derivatization used for LC-MS(/MS). Overall, coverage of 24-29% is obtained in LC-MS MassProfiler data analysis among the METASPACE matches with 20% FDR (**Figure S7A**); 94 out of 322 in GT, 100 out of 401 in GP, 27 out of 101 in CA, 18 out of 76 in PA. There are fewer matches in LC-MS/MS (**Figure S7B**), likely due to the difference in the database and insufficient MS/MS signals. In **Supplementary Table 2**, the LC-MS/MS matches are shown for each derivatization reagent.

CornCyc Annotations.

The METASPACE annotations when using the CornCyc database can give information about the compound coverage and biological value of our derivatized data. As expected, derivatization increased the number of tentatively identifiable metabolites. Within the CornCyc database, 75 molecular formulas were matched to known maize metabolites within the positive and negative mode control data. With the addition of each derivatization reagent an additional 178 compounds are matched in the database for a total of 253 compounds. This number is much smaller than a total of 578 when we used the combined database of CornCyc/Brachem, which is attributed to the smaller number of compounds in the CornCyc database (only ~1,750 unique formulas). The value of using this database is the pathway annotation available using the Plant Metabolic Network pathway tools.³⁸ In total, metabolites from 195 pathways were annotated in the control data while using derivatization allowed to detect metabolites from 343 pathways.

To determine the specific classes of metabolites and pathways where derivatization provided improvement, enrichment analysis was performed on each derivatization reagent and the results are shown in **Figure S8**. Enrichment provides an interesting interpretation of derivatized data that shows where improvement in compound coverage is gained due to derivatization. As the number of matches is limited (**Figure 4b**), PA did not give valuable results (not shown). Molecules detected when using CA, as expected, demonstrated a high enrichment in the amino acid biosynthesis ($p = 1.9\text{e-}15$) and other amino acid related pathways as well as ethene and ethylene

1 biosynthesis pathways. GP and GT have similar enrichment results, which is expected as they have
2 the same derivatization target. Many of the terpenoids and sterols have a carbonyl group and their
3 biosynthesis are highly enriched in both GP and GT. The phytosterol biosynthesis is a subclass
4 pathway of both terpenoid and sterol biosynthesis. As a result, it is most highly enriched among
5 the subclass pathways, also partially due to its importance in developing roots^{39,40}. As shown in
6 **Figure 6** for its simplified pathways, all five metabolites containing a ketone or aldehyde were
7 annotated in this pathway, with GT and GP derivatization covering 100% of compounds in the
8 pathway. Overall, they are abundant in cortex where their biosynthesis might be occurring but
9 avenasterone and episterone are also localized to the phloem, possibly indicating transportation
10 through the developing root.



11
12 **Figure 6.** Plant phytosterol biosynthesis pathway annotated using METASPACE analysis of GT
13 data with the CornCyc database. All five compounds were also identified in the GP dataset in high
14 confidence with 5% FDR. The five compounds annotated are the only compounds containing a
15 ketone or aldehyde in the pathway. Scale bar is 200 micron.

16 The brassinosteroid pathway is another highly enriched subclass pathway, especially in
17 GP-derivatization. Brassinosteroids are important plant growth hormones that are known to be
18 biosynthesized in the roots, especially during root development. Eight unique formulas were

1 annotated with the GP data in this pathway, compared to seven in the GT data (**Figure S9**),
2 including a few previously identified in GC-MS studies⁴¹. Only 2 of these carbonyl containing
3 compounds are annotated in the underivatized datasets with 20% FDR in plant phytosterol and
4 brassinosteroid pathways, further confirming on-tissue chemical derivatization greatly increases
5 the compound coverage in an untargeted spatial metabolomics study. Recently, MALDI-2 is
6 demonstrated for its ability to enhance ion signals for native sterols in MS imaging⁴² without any
7 derivatization. The two techniques, MALDI-2 and derivatization with GP or GT, are expected to
8 be complimentary each other as MALDI-2 detects sterols mostly as a water-loss, $[M-H_2O+H]^+$,
9 while GP or GT can only derivatize ketone or aldehyde group.

11 **Conclusions**

12 Using chemical derivatization in spatial metabolomics provides enhanced sensitivity yet
13 previously required extensive manual data analysis. We present a publicly available workflow for
14 annotation of derivatized metabolites implemented in METASPACE. As compared to laborious
15 manual data analysis and interpretation as well as subjective annotation, METASPACE provides
16 a fast, automated platform that robustly annotates and scores derivatized annotations. Moreover,
17 using METASPACE provides confidence in those annotations in the form of their FDR scoring.
18 Using this tool, the analysis of large datasets including multiple derivatization reagents and
19 replicates could be performed in a systematic manner yet achievable in hours rather than days.
20 This makes derivatization far more amenable in large untargeted spatial metabolomics studies. As
21 demonstrated with the search against the BraChem/CornCyc database, the use of multiple
22 derivatization reagents in parallel dramatically increases the compound coverage and improves the
23 understanding of spatial metabolism and enriched metabolic pathways.

24 False derivatized annotations represent a challenge with any derivatization method in
25 spatial metabolomics because of potential isomeric or isobaric relationships between underivatized
26 metabolites and other derivatized metabolites. We addressed this issue by searching for each
27 chemical modification in the underivatized dataset to identify and remove potential false
28 annotations. Moreover, we learned the following lessons that can be helpful in designing novel
29 strategies for chemical derivatization in the future. First, the structure and formula of the
30 derivatization reagent is important; the more unique the formula, the less likely derivatized and

underivatized molecules will be isomeric. For example, a derivatization reagent with a rare atom in biological system (e.g., F or Cl) or isotope labels would be much less likely to result in a random isomeric match. Second, the use of a smaller, more targeted database leads to fewer randomly annotated compounds, which minimizes false positive annotations and improves annotation of real metabolites. Unexpected side reactions are a potential pitfall when using chemical derivatization in untargeted spatial metabolomics. Systematic studies should be performed to better understand possible side reactions and their reaction efficiencies to avoid such pitfalls. Further development of derivatization reagents focused on MS imaging will lead to the expansion of untargeted spatial metabolomics, especially improving reaction efficiencies, reducing false positives, and increasing molecular coverage. The METASPACE chemical derivatization tool as well as the methods described herein are applicable to any derivatization reaction, currently available or to-be newly developed, and would accelerate applications of MS imaging by robust high-throughput analysis.

Acknowledgements

This work was supported by National Science Foundation under Grant No. (1905335) and the European Union's Horizon 2020 grant under the agreement number 825184. We acknowledge the W.M. Keck Metabolomics Research Laboratory (Office of Biotechnology, Iowa State University, Ames IA) for LC-MS analysis and we thank Dr. Lucas J. Showman for his assistance and support.

Supporting Information

Full experimental details, screenshot of the chemical modification tool, derivatization reaction scheme, illustration and removal of false derivatized annotations, METASPACE annotations of derivatized and underivatized MSI datasets, the effect of MSM in amino acids, MSI images of selected tentative annotations, summary of LC-MS(/MS) matches, pathway enrichment analysis, the metabolite coverage in brassinosteroid biosynthesis pathway, filenames of the dataset in METASPACE, and BraChem/CornCyc matches also detected in LC-MS/MS.

Notes

The authors claim no conflict of interest.

Citations

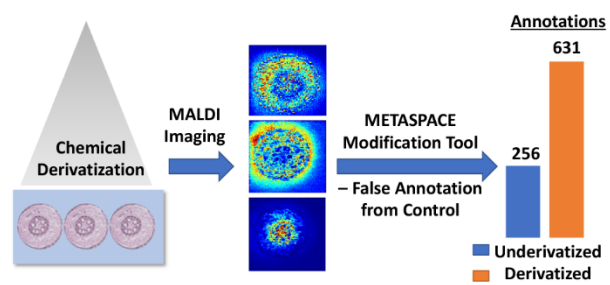
- (1) Doerr, A. Global Metabolomics. *Nat Methods* **2017**, *14* (1), 32–32. <https://doi.org/10.1038/nmeth.4112>.
- (2) Sajed, T.; Marcu, A.; Ramirez, M.; Pon, A.; Guo, A. C.; Knox, C.; Wilson, M.; Grant, J. R.; Djoumbou, Y.; Wishart, D. S. ECMDB 2.0: A Richer Resource for Understanding the Biochemistry of E. Coli. *Nucleic Acids Research* **2016**, *44* (D1), D495–D501. <https://doi.org/10.1093/nar/gkv1060>.
- (3) Wishart, D. S.; Feunang, Y. D.; Marcu, A.; Guo, A. C.; Liang, K.; Vázquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; Sayeeda, Z.; Lo, E.; Assempour, N.; Berjanskii, M.; Singhal, S.; Arndt, D.; Liang, Y.; Badran, H.; Grant, J.; Serra-Cayuela, A.; Liu, Y.; Mandal, R.; Neveu, V.; Pon, A.; Knox, C.; Wilson, M.; Manach, C.; Scalbert, A. HMDB 4.0: The Human Metabolome Database for 2018. *Nucleic Acids Research* **2018**, *46* (D1), D608–D617. <https://doi.org/10.1093/nar/gkx1089>.
- (4) Rai, A.; Saito, K.; Yamazaki, M. Integrated Omics Analysis of Specialized Metabolism in Medicinal Plants. *The Plant Journal* **2017**, *90* (4), 764–787. <https://doi.org/10.1111/tpj.13485>.
- (5) Alseekh, S.; Aharoni, A.; Brotman, Y.; Contrepois, K.; D’Auria, J.; Ewald, J.; C. Ewald, J.; Fraser, P. D.; Giavalisco, P.; Hall, R. D.; Heinemann, M.; Link, H.; Luo, J.; Neumann, S.; Nielsen, J.; Perez de Souza, L.; Saito, K.; Sauer, U.; Schroeder, F. C.; Schuster, S.; Siuzdak, G.; Skirycz, A.; Sumner, L. W.; Snyder, M. P.; Tang, H.; Tohge, T.; Wang, Y.; Wen, W.; Wu, S.; Xu, G.; Zamboni, N.; Fernie, A. R. Mass Spectrometry-Based Metabolomics: A Guide for Annotation, Quantification and Best Reporting Practices. *Nat Methods* **2021**, *18* (7), 747–756. <https://doi.org/10.1038/s41592-021-01197-1>.
- (6) Enomoto, H.; Sensu, T.; Yumoto, E.; Yokota, T.; Yamane, H. Derivatization for Detection of Abscissic Acid and 12-Oxo-Phytodienoic Acid Using Matrix-Assisted Laser Desorption/Ionization Imaging Mass Spectrometry. *Rapid Communications in Mass Spectrometry* **2018**, *32* (17), 1565–1572. <https://doi.org/10.1002/rcm.8200>.
- (7) Barré, F. P. Y.; Flinders, B.; Garcia, J. P.; Jansen, I.; Huizing, L. R. S.; Porta, T.; Creemers, L. B.; Heeren, R. M. A.; Cillero-Pastor, B. Derivatization Strategies for the Detection of Triamcinolone Acetonide in Cartilage by Using Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Imaging. *Anal. Chem.* **2016**, *88* (24), 12051–12059. <https://doi.org/10.1021/acs.analchem.6b02491>.
- (8) Zhang, Y.; Wang, B.; Jin, W.; Wen, Y.; Nan, L.; Yang, M.; Liu, R.; Zhu, Y.; Wang, C.; Huang, L.; Song, X.; Wang, Z. Sensitive and Robust MALDI-TOF-MS Glycomics Analysis Enabled by Girard’s Reagent T on-Target Derivatization (GTOD) of Reducing Glycans. *Analytica Chimica Acta* **2019**, *1048*, 105–114. <https://doi.org/10.1016/j.aca.2018.10.015>.
- (9) Gouw, J. W.; Burgers, P. C.; Trikoupi, M. A.; Terlouw, J. K. Derivatization of Small Oligosaccharides Prior to Analysis by Matrix-Assisted Laser Desorption/Ionization Using

- Glycidyltrimethylammonium Chloride and Girard's Reagent T. *Rapid Communications in Mass Spectrometry* **2002**, 16 (10), 905–912. <https://doi.org/10.1002/rcm.654>.
- (10) Kim, Y.-W.; Sung, C.; Lee, S.; Kim, K.-J.; Yang, Y.-H.; Kim, B.-G.; Lee, Y. K.; Ryu, H. W.; Kim, Y.-G. MALDI-MS-Based Quantitative Analysis for Ketone Containing Homoserine Lactones in *Pseudomonas Aeruginosa*. *Anal. Chem.* **2015**, 87 (2), 858–863. <https://doi.org/10.1021/ac5039362>.
- (11) Cobice, D. F.; Mackay, C. L.; Goodwin, R. J. A.; McBride, A.; Langridge-Smith, P. R.; Webster, S. P.; Walker, B. R.; Andrew, R. Mass Spectrometry Imaging for Dissecting Steroid Intracrinology within Target Tissues. *Anal. Chem.* **2013**, 85 (23), 11576–11584. <https://doi.org/10.1021/ac402777k>.
- (12) Shimma, S.; Kumada, H.-O.; Taniguchi, H.; Konno, A.; Yao, I.; Furuta, K.; Matsuda, T.; Ito, S. Microscopic Visualization of Testosterone in Mouse Testis by Use of Imaging Mass Spectrometry. *Anal Bioanal Chem* **2016**, 408 (27), 7607–7615. <https://doi.org/10.1007/s00216-016-9594-9>.
- (13) Yutuc, E.; Angelini, R.; Baumert, M.; Mast, N.; Pikuleva, I.; Newton, J.; Clench, M. R.; Skibinski, D. O. F.; Howell, O. W.; Wang, Y.; Griffiths, W. J. Localization of Sterols and Oxysterols in Mouse Brain Reveals Distinct Spatial Cholesterol Metabolism. *PNAS* **2020**, 117 (11), 5749–5760. <https://doi.org/10.1073/pnas.1917421117>.
- (14) Gao, X.; Lu, Y.; Wei, M.; Yang, M.; Zheng, C.; Wang, C.; Zhang, Y.; Huang, L.; Wang, Z. Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry Analysis of Human Milk Neutral and Sialylated Free Oligosaccharides Using Girard's Reagent P On-Target Derivatization. *J. Agric. Food Chem.* **2019**, 67 (32), 8958–8966. <https://doi.org/10.1021/acs.jafc.9b02635>.
- (15) Zhang, H.; Shi, X.; Vu, N. Q.; Li, G.; Li, Z.; Shi, Y.; Li, M.; Wang, B.; Welham, N. V.; Patankar, M. S.; Weisman, P.; Li, L. On-Tissue Derivatization with Girard's Reagent P Enhances N-Glycan Signals for Formalin-Fixed Paraffin-Embedded Tissue Sections in MALDI Mass Spectrometry Imaging. *Anal. Chem.* **2020**, 92 (19), 13361–13368. <https://doi.org/10.1021/acs.analchem.0c02704>.
- (16) Manier, M. L.; Spraggins, J. M.; Reyzer, M. L.; Norris, J. L.; Caprioli, R. M. A Derivatization and Validation Strategy for Determining the Spatial Localization of Endogenous Amine Metabolites in Tissues Using MALDI Imaging Mass Spectrometry. *Journal of Mass Spectrometry* **2014**, 49 (8), 665–673. <https://doi.org/10.1002/jms.3411>.
- (17) Guo, S.; Tang, W.; Hu, Y.; Chen, Y.; Gordon, A.; Li, B.; Li, P. Enhancement of On-Tissue Chemical Derivatization by Laser-Assisted Tissue Transfer for MALDI MS Imaging. *Anal. Chem.* **2020**, 92 (1), 1431–1438. <https://doi.org/10.1021/acs.analchem.9b04618>.
- (18) Esteve, C.; Tolner, E. A.; Shyti, R.; van den Maagdenberg, A. M. J. M.; McDonnell, L. A. Mass Spectrometry Imaging of Amino Neurotransmitters: A Comparison of Derivatization Methods and Application in Mouse Brain Tissue. *Metabolomics* **2016**, 12 (2), 30. <https://doi.org/10.1007/s11306-015-0926-0>.
- (19) O'Neill, K. C.; Lee, Y. J. Visualizing Genotypic and Developmental Differences of Free Amino Acids in Maize Roots With Mass Spectrometry Imaging. *Frontiers in Plant Science* **2020**, 11, 639. <https://doi.org/10.3389/fpls.2020.00639>.
- (20) Wu, Q.; Comi, T. J.; Li, B.; Rubakhin, S. S.; Sweedler, J. V. On-Tissue Derivatization via Electrospray Deposition for Matrix-Assisted Laser Desorption/Ionization Mass

- Spectrometry Imaging of Endogenous Fatty Acids in Rat Brain Tissues. *Anal. Chem.* **2016**, *88* (11), 5988–5995. <https://doi.org/10.1021/acs.analchem.6b01021>.
- (21) Palmer, A.; Phapale, P.; Chernyavsky, I.; Lavigne, R.; Fay, D.; Tarasov, A.; Kovalev, V.; Fuchser, J.; Nikolenko, S.; Pineau, C.; Becker, M.; Alexandrov, T. FDR-Controlled Metabolite Annotation for High-Resolution Imaging Mass Spectrometry. *Nat Methods* **2017**, *14* (1), 57–60. <https://doi.org/10.1038/nmeth.4072>.
- (22) Ovchinnikova, K.; Kovalev, V.; Stuart, L.; Alexandrov, T. OffsampleAI: Artificial Intelligence Approach to Recognize off-Sample Mass Spectrometry Images. *BMC Bioinformatics* **2020**, *21* (1), 129. <https://doi.org/10.1186/s12859-020-3425-x>.
- (23) Zhang, G.; Zhang, J.; DeHoog, R. J.; Pennathur, S.; Anderton, C. R.; Venkatachalam, M. A.; Alexandrov, T.; Eberlin, L. S.; Sharma, K. DESI-MSI and METASPACE Indicates Lipid Abnormalities and Altered Mitochondrial Membrane Components in Diabetic Renal Proximal Tubules. *Metabolomics* **2020**, *16* (1), 11. <https://doi.org/10.1007/s11306-020-1637-8>.
- (24) Nguyen, D. D.; Saharuka, V.; Kovalev, V.; Stuart, L.; Del Prete, M.; Lubowiecka, K.; De Mot, R.; Venturi, V.; Alexandrov, T. Facilitating Imaging Mass Spectrometry of Microbial Specialized Metabolites with METASPACE. *Metabolites* **2021**, *11* (8), 477. <https://doi.org/10.3390/metabo11080477>.
- (25) Sekera, E. R.; Saraswat, D.; Zemaitis, K. J.; Sim, F. J.; Wood, T. D. MALDI Mass Spectrometry Imaging in a Primary Demyelination Model of Murine Spinal Cord. *J. Am. Soc. Mass Spectrom.* **2020**, *31* (12), 2462–2468. <https://doi.org/10.1021/jasms.0c00187>.
- (26) Stutts, W. L.; Knuth, M. M.; Ekelöf, M.; Mahapatra, D.; Kullman, S. W.; Muddiman, D. C. Methods for Cryosectioning and Mass Spectrometry Imaging of Whole-Body Zebrafish. *J. Am. Soc. Mass Spectrom.* **2020**, *31* (4), 768–772. <https://doi.org/10.1021/jasms.9b00097>.
- (27) Veličković, D.; Bečejac, T.; Mamedov, S.; Sharma, K.; Ambalavanan, N.; Alexandrov, T.; Anderton, C. R. Rapid Automated Annotation and Analysis of N-Glycan Mass Spectrometry Imaging Data Sets Using NGlycDB in METASPACE. *Anal. Chem.* **2021**. <https://doi.org/10.1021/acs.analchem.1c02347>.
- (28) Lukowski, J. K.; Pamreddy, A.; Velickovic, D.; Zhang, G.; Pasa-Tolic, L.; Alexandrov, T.; Sharma, K.; Anderton, C. R. Storage Conditions of Human Kidney Tissue Sections Affect Spatial Lipidomics Analysis Reproducibility. *J. Am. Soc. Mass Spectrom.* **2020**, *31* (12), 2538–2546. <https://doi.org/10.1021/jasms.0c00256>.
- (29) Veličković, D.; Zhang, G.; Bezbradica, D.; Bhattacharjee, A.; Paša-Tolić, L.; Sharma, K.; Alexandrov, T.; Anderton, C. R. Response Surface Methodology As a New Approach for Finding Optimal MALDI Matrix Spraying Parameters for Mass Spectrometry Imaging. *J. Am. Soc. Mass Spectrom.* **2020**, *31* (3), 508–516. <https://doi.org/10.1021/jasms.9b00074>.
- (30) La Rocca, R.; Kune, C.; Tiquet, M.; Stuart, L.; Eppe, G.; Alexandrov, T.; De Pauw, E.; Quinton, L. Adaptive Pixel Mass Recalibration for Mass Spectrometry Imaging Based on Locally Endogenous Biological Signals. *Anal. Chem.* **2021**, *93* (8), 4066–4074. <https://doi.org/10.1021/acs.analchem.0c05071>.
- (31) Dueñas, M. E.; Larson, E. A.; Lee, Y. J. Toward Mass Spectrometry Imaging in the Metabolomics Scale: Increasing Metabolic Coverage Through Multiple On-Tissue Chemical Modifications. *Frontiers in Plant Science* **2019**, *10*, 860. <https://doi.org/10.3389/fpls.2019.00860>.

- (32) O'Neill, K. C.; Dueñas, M. E.; Larson, E.; Forsman, T. T.; Lee, Y.-J. Enhancing Metabolite Coverage for Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Imaging Through Multiple On-Tissue Chemical Derivatizations. In *Mass Spectrometry Imaging of Small Molecules: Methods and Protocols*; Lee, Y.-J., Ed.; Methods in Molecular Biology; Springer US: New York, NY, 2022; pp 197–213. https://doi.org/10.1007/978-1-0716-2030-4_14.
- (33) Xue, B.; Schlapfer, P.; Yon Rhee, S.; Zhang, P.; Xu, A.; Hawkins, C.; Fryer, E. *Summary of Zea mays mays, version 10.1.0*. <https://pmn.plantcyc.org/organism-summary?object=CORN> (accessed 2021-12-27).
- (34) Sud, M.; Fahy, E.; Cotter, D.; Brown, A.; Dennis, E. A.; Glass, C. K.; Merrill, A. H.; Murphy, R. C.; Raetz, C. R. H.; Russell, D. W.; Subramaniam, S. LMSD: LIPID MAPS Structure Database. *Nucleic Acids Research* **2007**, *35* (Database), D527–D532. <https://doi.org/10.1093/nar/gkl838>.
- (35) Hastings, J.; Owen, G.; Dekker, A.; Ennis, M.; Kale, N.; Muthukrishnan, V.; Turner, S.; Swainston, N.; Mendes, P.; Steinbeck, C. ChEBI in 2016: Improved Services and an Expanding Collection of Metabolites. *Nucleic Acids Res* **2016**, *44* (D1), D1214–9. <https://doi.org/10.1093/nar/gkv1031>.
- (36) Forsman, T. T.; Dueñas, M. E.; Lee, Y. J. On-Tissue Boronic Acid Derivatization for the Analysis of Vicinal Diol Metabolites in Maize with MALDI-MS Imaging. *J Mass Spectrom* **2021**, *56* (3), e4709. <https://doi.org/10.1002/jms.4709>.
- (37) Hallauer, A. R.; Lamkey, K. R.; White, P. R. Registration of B107, B108, and B109 Inbred Lines of Maize. *Crop Science* **1998**, *38* (6), cropscl1998.0011183X003800060076x. <https://doi.org/10.2135/cropscl1998.0011183X003800060076x>.
- (38) Karp, P.; Billington, R.; Holland, T.; Kothari, A.; Krummenacker, M.; Weaver, D.; Latendresse, M.; Paley, S. Computational Metabolomics Operations at BioCyc.Org. *Metabolites* **2015**, *5* (2), 291–310. <https://doi.org/10.3390/metabo5020291>.
- (39) Schrick, K.; Fujioka, S.; Takatsuto, S.; Stierhof, Y.-D.; Stransky, H.; Yoshida, S.; Jürgens, G. A Link between Sterol Biosynthesis, the Cell Wall, and Cellulose in Arabidopsis. *The Plant Journal* **2004**, *38* (2), 227–243. <https://doi.org/10.1111/j.1365-313X.2004.02039.x>.
- (40) Benveniste, P. Biosynthesis and Accumulation of Sterols. *Annual Review of Plant Biology* **2004**, *55* (1), 429–457. <https://doi.org/10.1146/annurev.arplant.55.031903.141616>.
- (41) Kim, Y.-S.; Kim, T.-W.; Kim, S.-K. Brassinosteroids Are Inherently Biosynthesized in the Primary Roots of Maize, *Zea Mays* L. *Phytochemistry* **2005**, *66* (9), 1000–1006. <https://doi.org/10.1016/j.phytochem.2005.03.007>.
- (42) Bien, T.; Hambleton, E. A.; Dreisewerd, K.; Soltwisch, J. Molecular Insights into Symbiosis—Mapping Sterols in a Marine Flatworm-Algae-System Using High Spatial Resolution MALDI-2-MS Imaging with Ion Mobility Separation. *Anal Bioanal Chem* **2021**, *413* (10), 2767–2777. <https://doi.org/10.1007/s00216-020-03070-0>.

1



2