Relational Graph Convolutional Network for Robust Mass Spectrum Classification

Raphaël La Rocca[1], Anthony Cioppa[2], Enrico Ferrarini[3], Monica Höfte[3], Marc Van Droogenbroeck[2], Edwin De Pauw[1], Gauthier Eppe[1] and Loïc Quinton[1*]

*E-mail: loic.quinton@uliege.be

[1] Mass Spectrometry Laboratory, MolSys Research Unit, University of Liège, B4000, Liège, Belgium

[2] TELIM, Montefiore Institute, University of Liège, B4000, Liège, Belgium

[3] Department of Plants and Crops, Faculty of Bioscience Engineering, Ghent University, B9000, Ghent, Belgium

Abstract

Supervised machine learning methods have shown impressive performance in interpreting mass signals and automatically segmenting spatially meaningful regions in Mass Spectrometry Imaging (MSI). Such segmentation generates maps that provide researchers with valuable insights into sample composition and serve as a foundation for downstream statistical analyses. However, these models often require dataset-specific pre-processing and do not fully exploit the rich mass features available in high-resolution mass spectrometry (HRMS). Unlike low-resolution mass spectrometry, HRMS reveals additional features such as mass defects and repeated nominal mass differences that carry important chemical information.

In this work, we propose a novel deep learning architecture based on a Relational Graph Convolutional Network (R-GCN) that captures and leverages those HRMS mass features. Our model explicitly encodes structural features such as mass defects and known mass differences

to represent each spectrum as a graph, enabling the learning of associations between chemically related ion families. To the best of our knowledge, no existing deep learning models for MSI classification incorporate this level of chemically informed mass structure. Most existing methods treat spectra as flat vectors or image-like inputs, thereby ignoring the underlying mass relationships.

We evaluate our R-GCN approach against several conventional machine learning and deep learning baselines across diverse MSI datasets, demonstrating its robustness to common signal variations (e.g., mass shift, ion loss). Finally, we integrate Class Activation Mapping (CAM) to enhance model interpretability, enabling the identification of ion families that are relevant to specific biological or spatial regions.

1. Introduction

Mass Spectrometry Imaging (MSI) has been recognized as a powerful tool for detecting biological regions of interest based on spectral signatures. It has been widely used to analyze a wide variety of biological samples such as biopsies[1,2], bacterial colonies[3], and even single-cells[4,5]. MSI is a specific application of mass spectrometry in which spectra are recorded, usually by Matrix-Assisted Laser Desorption Ionization (MALDI), at different spatial positions, called pixels, over a sample forming a two-dimensional image. Each pixel corresponds to a spectrum from which different amounts of ions are extracted and detected at specific m/z values (channels). The signals' resolution and mass accuracy are directly related to the instrumentation properties and tuning[6]. In a biological context, the high mass accuracy and the isotopic signature reached by high-resolution mass spectrometry (HRMS) instruments may be exploited to deduce molecular formulae from mass peaks, helping the interpretation of such signals[7].

The large number of signals gathered from MSI (HRMS) acquisitions requires specialized data mining tools to analyze, interpret, and exploit such images[8]. In this context, several unsupervised machine learning methods have been proposed for data exploration of MSI

signals. Those methods extract spatial regions with similar spectral signatures, thus delineating biologically meaningful regions[9]. More recently, supervised machine learning (ML) methods have been proposed to recognize spectra belonging to annotated classes automatically. These methods are particularly useful in a clinical diagnosis context for cell type classification and image segmentation, leading to automatic detection of tumor regions[10], single-cell classification[11], or automatic digital tissue staining[12]. Those methods are also useful for mass signal interpretation such as proposed by Tideman *et al.*[13] In their article, the authors combine ML classifier with Shapley additive explanations to identify mass signals linked to specific biological regions in MSI. Examples of commonly used algorithms for supervised machine learning on mass spectra are Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), and Random Forests (RF). These algorithms are included in most software packages dedicated to mass spectra or MSI analysis[4], such as MetaboAnalyst[14] or licensed software such as SCiLs (SCiLS, Bruker, Bremen, Germany). For all these ML algorithms, mass spectrum pre-processing is mandatory[15,16] and consists of selecting features corresponding to m/z signals' overall mass spectra within a certain tolerance. The model then typically uses these features to classify mass spectra into different predefined classes. Combined with interpretability measures, it is possible to highlight m/z values specific to biological regions, facilitating downstream analysis[11,13].

In MSI, signal variations are common due to sample preparation processes, intrinsic molecular heterogeneity from pixel to pixel, and sample to sample [17,18]. These variations affect the signal differently, for instance, by shifting the detected masses, adding noise to the measured intensities, or changing peak intensity ratios[19,20]. In HRMS, m/z variations such as mass shifts significantly impact the feature match between mass spectra, which skews data interpretation[21,22]. This can also have an impact on feature quality and, therefore, on classification performances. Those variations limit the applicability of the classical models to images only after applying specific pre-processing strategies, often guided by expert knowledge. For instance, recalibration of mass spectra necessitates a set of calibrant ions and a measure to evaluate the recalibration process[21]. Moreover, filtering out sample-unrelated

signals is possible by applying image filters with their own set of parameters[23]. To avoid extensive manual feature selection, Convolutional Neural Networks (CNN) have been proposed for spectra classification[10,24]. CNNs are successful deep learning architectures originally used to learn features from 2D RGB images. Convolution layers are applied to the images, and spatially invariant features are learned before being passed to the next layers[25]. CNN architectures can be adapted to spectral data from Raman spectroscopy[26] or mass spectrometry[10,24]. In a few words, convolution is usually applied to spectra with little pre-processing, consisting of dividing the mass range into bins of fixed size to form a unidimensional vector. The CNN network creates complex features (i.e. combination of bins) by exploiting local invariant intensity patterns. This achieves better classification performances compared to conventional ML methods without pre-processing or feature-engineering step[24]. While CNNs have shown great classification performances on low-resolution mass spectra, they are not adapted to HRMS data. In fact, HRMS requires a smaller bin size to maintain and to take advantage of the high-resolution information, notably for data interpretation. This results in an increase in model size, as the input and convolution kernel size need to be adapted to the larger input. The increase in the number of model parameters is such that CNN models become too large for common GPU (Graphics Processing Unit) and are therefore poorly adapted to HRMS.

Another data representation for mass spectrum is the Kendrick plot. It is a useful tool for visualizing mass patterns in HR mass spectra[27]. Kendrick plots (m/z = f(KMD), i.e. Kendrick Mass Defect) capitalize on the mass and the mass defect from the m/z scale of the ion's mass peaks to group them according to their atomic compositions. The more the peaks, represented by dots, are spaced in this plot, the less likely they have a similar atomic composition[28]. Here, we define mass patterns as peaks spaced by repetitive mass differences. Those differences correspond to the presence of isotopes, adducts, fragmentations, and chemical transformations[29,30] as depicted in Figure 1. The molecular families of peaks can be characterized according to their Kendrick positions, mass patterns, and intensity ratios. When

properly highlighted, molecular families can provide useful information for mass spectra classification and mass signal interpretability.

In this work, we propose to build a new supervised deep learning model from a Relational Graph Convolution Network (R-GCN) by modeling efficiently the following HRMS mass feature information: exact mass, intensity, mass defect, and mass patterns. In this study, we evaluate our new method over state-of-the-art methods for mass spectra classification. This evaluation is proposed in two steps: comparing the performances on the classification of background and foreground spectra from different MSI HRMS images and the classification of spectra from homeostatic and non-inflammatory Non-Alcoholic Fatty Liver Disease (NAFLD). In particular, we evaluate their robustness to different signal variations. Finally, we show how the model can extract signal produced by bacterial colonies grown on agar by using ion representation learned by the model and integrating class activation maps[31] to interpret class-specific m/z values and compare them to known molecular families in the dataset.

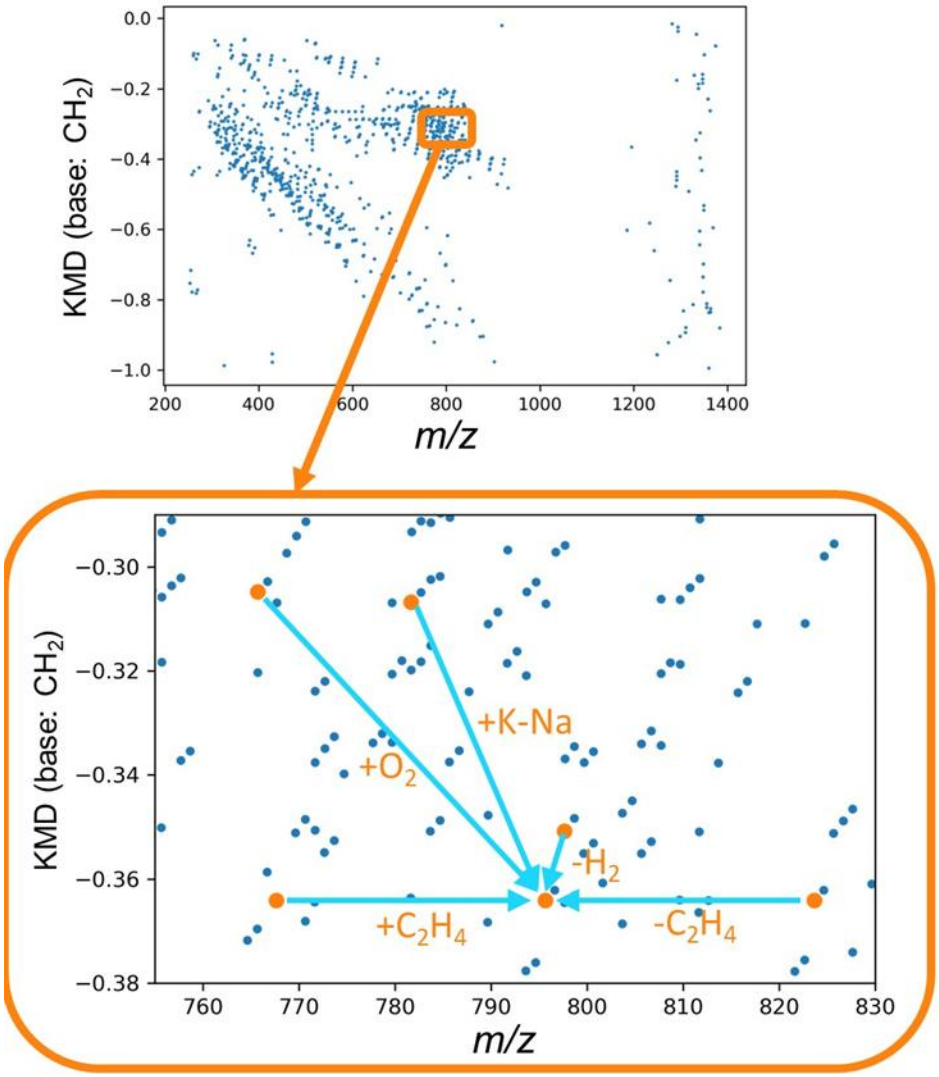


Figure 1: Representation of HRMS mass features by a Kendrick plot in CH_2 base of a MALDI MSI of a human lung biopsy. Blue points represent mass peaks. Light blue arrows represent specific mass differences corresponding to a particular difference in atomic composition.

2. EXPERIMENTAL SECTION

2.1 Material

In this study, we collect several MALDI MSI datasets from METASPACE (see Supporting Information, S1), a hepatocyte dataset including homeostatic state (Control) and non-inflammatory Non-Alcoholic Fatty Liver Disease (NAFLD) hepatocytes from Rappez et al.[5] and we provide an additional dataset from our laboratory, related to *Pseudomonas sessilinigenes*

CMR12a[32] bacterial colonies grown on agar. We include two additional MALDI-MSI datasets from Burguet et al.[33]. These datasets were designed to compare culture conditions that either activate or repress specialized metabolite production in Streptomyces scabiei. Sample preparation and spatial annotation are detailed in Supporting Information S6 for all datasets.

3. Method

3.1 General approach

In HRMS, the mass peak intensity values are not the only information available. Masses, mass defects, and mass differences between peaks contain important information from which a model can learn and extract essential patterns for data interpretation (Figure 1). Currently, those features are systematically overlooked by automatic state-of-the-art methods. However, this information is often used for manual analysis and provides a good estimation of the molecular composition. In this work, we develop a new deep learning algorithm for MSI's automatic mass spectrum classification to efficiently incorporate those extra mass features. To accomplish this, we select a graph representation for the spectrum described below.

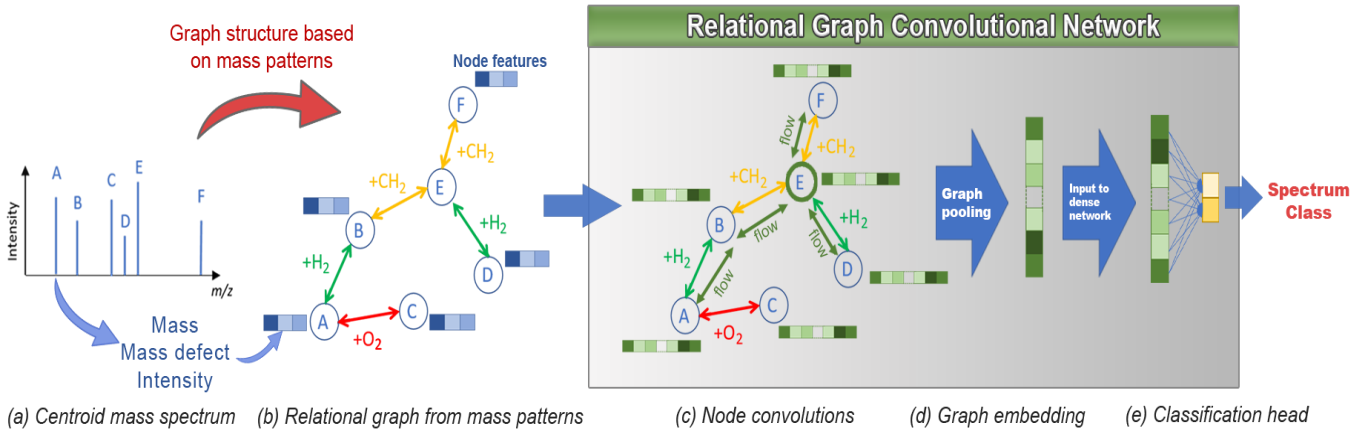


Figure 2: Pipeline of our R-GCN for HRMS classification. a) A centroid mass spectrum is transformed into a graph containing b) the HRMS features of each peak. c) graph convolution layers create a representation for each node, then d) aggregate it into a single vector representation by a pooling layer passed to e) a classification head.

3.2 Graph representation of the spectrum

As a first step, each spectrum of each pixel (see Figure 2 a) is converted into a graph. To this end, each centroid in the spectrum is represented as a node in the graph. For each node, three features are associated: the mass of the centroid, its intensity, and computed mass defect according to IUPAC as follows:

$$Mass\ Defect = Mass - \lceil Mass \rceil \quad (1)$$

Mass Defect is defined as the difference between the measured Mass (the ion mass in atomic mass units) and its ceiling Mass ($\lceil$Mass$\rceil$, the smallest integer greater than or equal to the mass). Note that the mass defect and mass values are scaled between −0.5 and 0.5 and the centroid intensities are log-transformed and then Z-score normalized. Then, two nodes are connected by a bi-directional edge if the absolute difference between their associated mass corresponds, within a certain mass tolerance (e.g., set at ±0.001 Da), to a predefined set of mass differences. For all experiments, we selected the same set composed of 102 mass differences. Those differences describe the most frequent mass patterns in MALDI-MSI experiments[29] (see Supporting Information, S3). For instance, an element in the set may represent the mass difference of a change of adducts, isotopic patterns, the masses of common chemical transformations such as the addition of H_2, CH_2, and O2, and mass differences from DHB and HCCA. As an extra feature, we provide this label to the bi-directional edge connecting the two nodes. Therefore, our spectrum representation is a bi-directional relational graph (see Figure 2 b)). Different methods for selecting mass differences have been explored (see Supporting Information, S8).

3.3 Relational graph convolutional network

Graph Convolutional Networks (GCN) is a particular network architecture that applies convolution operations on node features in a graph by aggregating the information of adjacent nodes. Performing multiple consecutive convolutions allows for the aggregation of node information further and further away in the graph. Relational-GCN (R-GCN) is a particular type of GCN where the graph edges contain labels. This allows us to model different relation types

between the nodes. In this case, the convolution aggregates adjacent node representations connected by similar edge types[34].

We used the PyTorch Geometric library[35] to build our own R-GCN network architecture. Our R-GCN takes as input a graph of mass differences and encodes the graph's node features with two graph convolution layers. The first convolution layer converts the initial three features of each node into eight features, and the second one transforms the eight features of each node into 16 features (see Figure 2 c)). Next, average pooling is applied to all nodes, feature by feature. This pooling operation allows the reduction of all the node's representation into a single vector of 16 features (see Figure 2 d)). This representation is then passed to a single fully connected layer that outputs two values. All layers are activated with the ReLU (rectified linear unit) function except the output layer, which is activated by a Softmax Activation function to produce the predictions for each class (see Figure 2 e)). A more detailed description of the network architecture can be found in Supporting Information S7). The impact of the method's parameters is also explored (see Supporting Information, S8).

3.4 Class activation maps for network interpretability

Classification activation maps (CAM)[31] is a method for interpreting, visualizing, and debugging CNN networks. We propose to adapt CAM to interpret learned mass features of our R-GCN. The CAM operation consists of applying the classification head of the network directly to each node representation before the pooling layer. This gives each node present in the graph two features corresponding to the contribution of the node to each of the two classes. This allows us to interpret which node (and therefore which peak in the spectrum) is characteristic to a class. We consider a peak to be characteristic of a particular class if its CAM feature, i.e., its contribution to that class, is greater than its contribution to the other class.

3.5 Code Availability.

The R-GCN method implementation and the implementation of various classification methods are available on GitHub at *https://github.com/LaRoccaRaphael/Massconvnet*. Training time and

model complexity for each method on the background/foreground classification task are provided in the supplementary information (S11.1).

4. Results and discussion

As explained in the introduction, the current work proposes a new method for mass spectrum classification that takes advantage of the features provided by HRMS (exact masses, mass defects, mass patterns). We characterize our new R-GCN method for mass spectrum classification into two sections: (1) Its robustness toward signal variation and (2) how it can be used for m/z signal interpretation (R-GCN signal interpretability). To support our experiments, we compare our method against different spectrum classification methods: Random Forests (RF), Linear discriminant analysis (LDA), Support vector machine (SVM), eXtreme Gradient Boosting (XGBoost), and 1D Convolutional Neural Networks (1D-CNNs). Each method of pre-processing is described in Supporting Information S9.

4.1 Robustness toward signal variation

4.1.1 Experimental setup

For this first experiment, we evaluate the method's robustness toward signal variation on two different classification tasks: classification of foreground (FG) and the background (BG) spectra from 4 different MSI and classification of human HepaRG hepatocytes cells into control and non-inflammatory NAFLD states. For this purpose, we create:

- A training set and a validation set from the four METASPACE MSI datasets were gathered by randomly sampling 64% and 16% of the spectra, respectively (6,400 and 1,600). The remaining 20% of spectra (2,000) are kept as a variation-free test set.
- A training set and a validation set from 2 images of each class in the hepatocyte's dataset were gathered by randomly sampling 80% and 20% of the spectra, respectively (4,000 and 1,000). The remaining four images (2 for each class) are kept as a variation-free test set (5,000).

From each dataset test set, we generate four variation test sets by simulating diverse signal variations on the data, which are constant and random mass shifts, a decrease in ion detection intensity, and a mix of signal classes (see below). All six methods (RF, SVM, LDA, XGBoost, 1D-CNN, and R-GCN) are trained on the same training set without any mass variation, except for signal mixing variation. Then, the models are evaluated on the four variation test sets to study their behavior when faced with these variations. For the evaluation metric, we choose balanced accuracy since the background/foreground datasets are highly imbalanced towards the background (see Supporting Information S11.1) . Balanced accuracy is defined by:

$$balanced\ accuracy = \frac{1}{2}\left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP}\right) \quad (2)$$

Where: TP – true positive, TN – true negative, FP – false positive, FN – false negative.

The four signal variations simulated in the test sets are the following.

Test 1. A *mass shift* was simulated by adding a constant value to every signal in each spectrum.

Test 2. A *random mass shift* was simulated by adding a random value drawn from a normal distribution N(0,σ) to every signal in each spectrum. An increase in the standard deviation of this distribution corresponds to a decrease in mass precision.

Test 3. A *decrease in ion detection sensitivity* was imposed by removing a proportion of the lowest intensity peaks in each spectrum.

Test 4. We simulate signal mixing by modifying the test and train sets. For each set, we randomly take 50% of the spectra and add to each spectrum a spectrum from the other class. The added spectrum is multiplied by a value between 0 and 1, where a value of 1 indicates an equal mix of the two class signals.

Below, we define the key concepts of mass spectrum signal characterization to provide clearer context for the types of signal variation studied in this work.

- Mass accuracy refers to how close the measured m/z value of an ion is to its theoretical value. It is crucial for accurate compound identification.
- *Mass resolution* describes the instrument's ability to distinguish between ions with similar m/z values; high resolution enables detection of fine isotopic or adduct patterns.
- *Mass precision* refers to the consistency of mass measurements across spectra, low precision results in variability in peak positions across the dataset.
- *Mass shift* is a systematic or random deviation in m/z values across a dataset, often due to calibration drift or instrument instability.
- *Ion detection sensitivity* indicates the instrument's ability to detect low-abundance ions; lower sensitivity can lead to missing weak but potentially relevant peaks.
- *Spatial resolution* in MSI defines the physical size of each sampled area (pixel); lower spatial resolution may lead to signal mixing from adjacent tissue regions.

Note that each of the signal variations described above often co-occurs in real cases; simulation allows us to isolate them, which is essential to study their impact on the various classification methods.

Balanced Accuracy	R-GCN *(this work)*	1D-CNN	RF	SVM	LDA	XGBoost
Background/Foreground	0.874 ± 0.001	0.860 ± 0.002	0.871 ± 0.001	0.873 ± 0.0	0.861 ± 0.0	0.863 ± 0.0
Control and non-inflammatory NAFLD human HepaRG hepatocytes	0.991 ± 0.002	0.996 ± 0.001	0.985 ± 0.001	0.999 ± 0.0	0.999 ± 0.0	0.996 ± 0.0

Table 1: Performance comparison between our R-GCN method against different methods using balanced accuracies *on the BG/FG classification task of the METASPACE MSI datasets and the classification of control and non-inflammatory NAFLD human HepaRG hepatocytes.*

As a benchmark, we evaluate all methods on the original test set without any signal variation. To account for the reproducibility of models with random weight initialization, we train 15 models for each method by varying the order in which the training samples are presented to the model during the training phase (called random seeds in the following). We provide the mean and standard deviation of the balanced accuracy across those 15 repetitions. Since RF, R-GCN, and 1D-CNN have random initialization and the order in which the training samples are presented matters in their optimization, the standard deviation may be greater than 0. As can be seen, all models achieve similar performances (see Table 1).

4.1.2. Classification results with mass variations

The performances over different signal variation ranges (x-axis of the plots) for each type of variation are shown in Figure 3. The models are trained on the same training set without signal variations for all six methods. To account for the variability that can arise during training, we train 15 models for each algorithm by shuffling each time the order of the samples in the training phase and the network initialization parameters when relevant. As explained above in the benchmark, only RF, R-GCN, and 1D-CNN can exhibit non-zero standard variations. Our R-GCN method shows excellent robustness to mass shift variation compared to other methods (Figure 3). Note that simulations with exaggerated shifts (mass shift >0.01 Da) are designed to examine the model's reliance on mass features and are not used to demonstrate typical instrument mass deviations. Accordingly, we also present mass shift values on a more realistic scale in the second row of Figure 3.

In fact, it outperforms the other methods over the entire mass shift range except the 1D-CNN on the hepatocyte's dataset. The R-GCN and 1D-CNN's performances vary much more with mass shift compared to the other methods. This non-monotonic behavior is primarily due to the fixed binning strategy used in the preprocessing of input spectra for the 1D-CNN. When mass shifts cause ions to misalign with their expected bins, the input representation leads to a

drop in classification accuracy. Interestingly, as the shift approaches 1 Da, performance begins to recover. This is because the shifted ion starts to align with the first isotopic peak (all ions are singly charged due to the MALDI process). This alignment effect partly compensates for the misalignment caused by the shift, resulting in the performance rebound observed in Figure 3 after a shift of 0.8 Da.

In contrast, the RF, SVM, LDA, and XGBoost models fail completely when they reach a mass shift of 0.001 Da. In fact, those methods are not robust to mass shift compared to R-GCN and 1D-CNN as they do not incorporate mass information. For the R-GCN, when we shift all masses by a constant value, the graph structure remains unchanged and only the node features (m/z and mass defect) are impacted. Indeed, extremely large changes in the values of these features (*e.g.,* 0.8 Da) impact the model performance with a bigger impact on the hepatocyte's dataset, which means that mass defect is an important feature for spectra classification in this dataset. This perfectly shows the robustness of our model towards this type of variation in the signal; moreover, the robustness toward implausible mass shifts (0.8 Da) demonstrates how the model depends on mass defect values for classification.

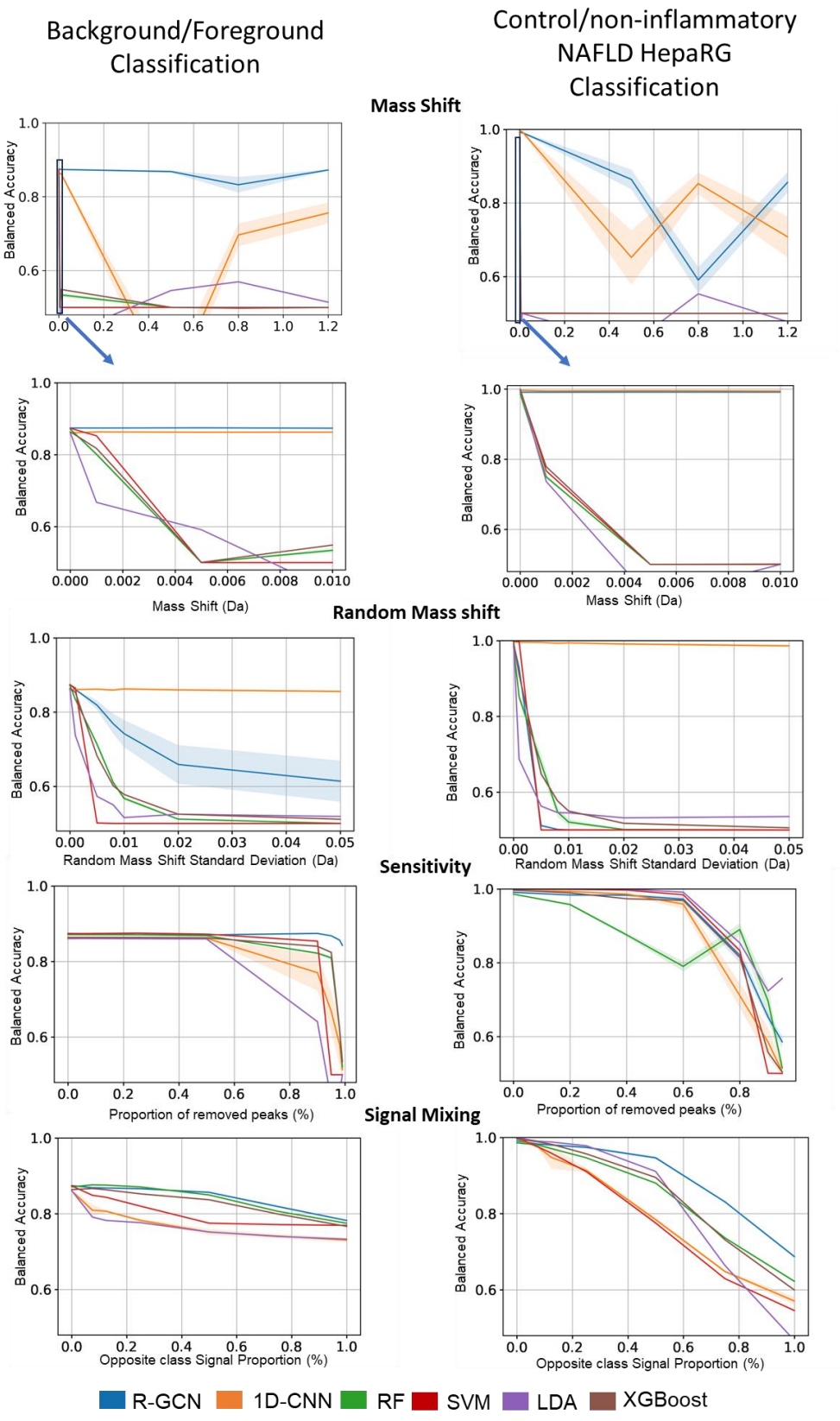

Background/Foreground Classification
Control/non-inflammatory NAFLD HepaRG Classification
Mass Shift
Balanced Accuracy
Mass Shift (Da)
Random Mass shift
Random Mass Shift Standard Deviation (Da)
Sensitivity
Proportion of removed peaks (%)
Signal Mixing
Opposite class Signal Proportion (%)
R-GCN
1D-CNN
RF
SVM
LDA
XGBoost

Figure *3: Performance of the binary classification Background/Foreground (BG/FG) and control/non-inflammatory NAFLD human HepaRG hepatocytes for RF, SVM, LDA, XGBoost, 1D-CNN, and R-GCN on four simulated signal variations: Mass shift, Random mass shift, Decrease in ion sensitivity (Lowest peak removal), Signal mixing. The error bars correspond to the standard deviation of the prediction evaluation for 15 models trained with random seeds for the RF, R-GCN, and 1D-CNN methods .*

A random mass shift, which corresponds to a decrease in precision, has a different impact on the models, as seen in Figure 3. Here, the 1D-CNN's bin size (0.1 Da) prevents the (limited) mass shift from affecting the model, as the mass shift range is much smaller than the bin size. In the case of our method, the graph structure is subjected to changes, and, as a result, the performance drops slowly in the BG/FG dataset and more rapidly in the hepatocyte dataset. The changes are mainly caused by the graph's edge removal as the distance between the mass centroids becomes larger than the method's predefined mass tolerance. In fact, our R-GCN is sensitive to a change in the mass distance between the centroids (see Supporting Information, S8 Table S3), which means that the mass precision of the MSI in the train set restricts its application to MSI with related mass precision. This problem can be avoided by correctly setting mass tolerance according to the original MSI precision of the test set.

For the signal variation of a decrease in ion detection sensitivity, the different methods demonstrate similar performances for 0-50% removal with lower performances for RF and LDA. For higher proportions, our R-GCN is consistently better than the other two for the BG/FG dataset.However, the differences in performance between datasets are visible. Those differences can be attributed to the underlying complexity of the dataset. The background/foreground dataset includes MSI from four diverse biological sources (e.g., tissue, bacteria), leading to stronger, more generalizable class distinctions. Background spectra often contain abundant matrix-related peaks (e.g., HCCA), while foreground spectra feature biologically relevant ions like endogenous lipids and metabolites. These signals remain even after significant peak removal, allowing the R-GCN to preserve classification accuracy by leveraging residual graph structure.

In contrast, the hepatocyte dataset involves more subtle metabolic differences between two similar cell states. Most discriminative peaks belong to chemically similar lipid species that are easily lost under high peak dropout, which weakens graph connectivity and model performance. This explains why the R-GCN shows greater resilience in the BG/FG task but converges toward other methods in the hepatocyte setting.

Signal mixing is a common phenomenon in MSI, as mass spectra often span over two different biological regions due to an inadequate spatial resolution. Therefore, the resulting mass spectra contain mass information from the two regions. The lower performances of the 1D-CNN demonstrate the importance of conserving high spectral resolution for classifying mass spectra with mixed biological origins (Figure 3). The 1D-CNN's bin size (0.1 Da) compared to the mass tolerance of the other methods (0.001 Da) greatly impacts the model performances.

The R-GCN is consistently better than the other methods for high signal mixing, especially in the hepatocyte's dataset, where the two classes have similar signals. Attempting to increase the number of bins in the 1D-CNN to match the resolution used by the R-GCN (e.g., 0.001 Da tolerance) would dramatically increase the CNN's complexity, from approximately 4.8 million to 480 million parameters. This scale of dimensionality imposes significant computational and memory demands and substantially increases the risk of overfitting, especially given the relatively small training sizes typical of MSI datasets.

In contrast, the R-GCN architecture is far more scalable and robust. It requires only 15,746 trainable parameters, and its complexity remains constant regardless of the number of bins or peaks. Instead, it depends on the structure of the graph defined by mass differences, making it inherently better suited for HRMS data where rich mass patterns and high resolution are key to molecular interpretation.

4.3 R-GCN signal interpretability

As shown in the previous sections, R-GCN exhibits improved robustness to variations in the signal and to predictions on unseen datasets with distinct signals (see Supporting Information, S10 Figure S3). These improvements can be attributed to better representation of HRMS

features, such as mass patterns, masses, and mass defects. These new features allow the method to use mass information more flexibly than 1D-CNN. Building on the R-GCN's ability to capture chemically meaningful mass relationships, we now explore whether its learned feature representations can also aid in interpreting MSI data. Specifically, we investigate whether R-GCN can help identify ions of interest from a list of m/z values by highlighting those that the model associates with spatial regions of interest in various case studies. Our approach consists of training an R-GCN model for binary classification of MSI spectra and then applying the trained model to a representative spectrum from each dataset. Rather than classifying spectra, we repurpose the model to assign class relevance scores to individual ions, thereby identifying which m/z values contribute most to distinguishing between the two classes.

To achieve this, we integrate a Class Activation Map (CAM) strategy into the R-GCN architecture and apply it across multiple datasets to extract m/z values associated with different spatial regions or experimental conditions—such as foreground (FG) vs. background (BG), or control vs. treated samples.

The resulting CAM scores allow us to refine the mean spectrum of each MSI dataset by filtering for class-relevant ions, highlighting signals corresponding to molecules produced under specific biological conditions. This not only improves interpretability but also simplifies downstream analysis by focusing on biologically meaningful signals.

4.3.1 Experimental setup

In contrast to the experiments in Section 4.1, here we modify the R-GCN input by setting all node intensities to 1 during both training and inference. This design ensures that the model learns exclusively from mass-based features, avoiding redundancy with intensity-based methods and enabling a more straightforward interpretation of mass-driven contributions. To improve robustness to small deviations in m/z values, the tolerance for graph generation is set to 0.003 Da across all datasets. For each dataset, a representative spectrum is constructed by computing the mean spectrum across all individual spectra of the dataset. Once the model is trained on each dataset, we apply the CAM strategy to the representative spectrum and its associated relational graph. The CAM produces two activation values per node (i.e., per m/z

value), corresponding to the contribution of each ion to the two respective classes. This allows us to identify m/z values that are most characteristic of each class and effectively classify or associate ions with specific biological or experimental conditions (see Methods for details). We evaluate this interpretability strategy across three distinct case studies, each representing a different biological context and classification task.

Case Study 1: Spatially Resolved Ion Detection in *Pseudomonas* Colonies

The first case study focuses on identifying ions expressed in distinct spatial regions within a single MSI dataset. We apply this to a newly generated MSI of *Pseudomonas* colonies grown on agar (5th dataset in Supporting Information S6, Figure S1). An R-GCN model is trained to classify spectra from the colony regions (foreground) versus the surrounding medium (background). This task illustrates how the model can differentiate microbial metabolic signals from those of the media (Figure 4 left part).

Case Study 2: Natural Product Discovery in *Streptomyces scabiei*

The second case study explores the use of R-GCN for natural product discovery, using two MALDI-MSI datasets from Burguet et al [33]. These datasets were designed to compare culture conditions that either activate or repress specialized metabolite production in *Streptomyces scabiei*:

- Dataset 1: Cellobiose-Induced Pathogenicity

Here, we compare spectra from *S. scabiei* cultures exposed to cellobiose (which induces virulence-related metabolite production) against negative controls treated with water or sugar (Supporting Information S6). An R-GCN model is trained to classify spectra from the induced versus control conditions.

- Dataset 2: Iron-Regulated Siderophore Production

In this dataset, *S. scabiei* is grown under iron-depleted versus iron-supplemented conditions to study siderophore biosynthesis. The R-GCN model is again trained to distinguish between spectra from the iron-starved (experimental) and iron-rich (control) cultures.

Case Study 3: Lipid Remodeling in a Human Hepatocyte NAFLD Model

The third case study uses a dataset from the SpaceM study (Rappez et al.[5])(Supporting Information S6), also evaluated in earlier sections of this paper. It involves an MSI of human hepatocytes (dHepaRG) under homeostatic conditions (control) and under stimulation with oleic and palmitic fatty acids (+FA) to model non-inflammatory NAFLD. SpaceM analysis previously revealed metabolic remodeling in this model, with an increase in triglyceride and diglyceride expression in NAFLD hepatocytes and glycerophospholipid enrichment in controls. We train an R-GCN model to classify spectra from +FA (NAFLD) versus control hepatocytes, enabling interpretability in the context of lipid dysregulation.

Finally, we annotate selected m/z values in each dataset to evaluate biological relevance:

Pseudomonas dataset. The m/z values from the representative spectrum are matched to a curated list of exact masses corresponding to known molecular formulae, using a tolerance of ±3 ppm (see Supporting Information, S4 and S5) . Those molecular formulae correspond to HCCA ion masses and various lipopeptides (orfamides, sessilins) and phospholipids (PE, PC) known to be produced by this *Pseudomonas* strain[36]. Examples of annotated m/z values are shown in Figure 4, along with ion images and bacterial colony regions. To further support these assignments, Spearman correlations were computed between ion images and annotation masks to assess spatial co-localization.

S. scabiei cultures. We reference known ion identities from Burguet et al. For the cellobiose-induced dataset, Thaxtomin A and B are expected in experimental condition. For the iron-repressed dataset, siderophores such as Desferrioxamine B, Desferrioxamine E, and Scabichelin are reported (see Supplementary Information S4).

Hepatocyte (SpaceM) dataset. Annotations are retrieved from the original METASPACE results (as described in the SpaceM publication). The key lipid classes are reported onto our data and include diglycerides, triglycerides, and glycerophospholipids, which reflect the reported metabolic remodeling in the non-inflammatory NAFLD model (see Supplementary Information S4).

4.3.3 Signal filtering and identification in *Pseudomonas* colonies

In this section, we examine how our method filters m/z values in the Pseudomonas dataset using the CAM-derived scores. As shown in Figure 4, the Kendrick plot highlights distinct molecular families associated with each class.

Background-specific (BG) ions cluster tightly in a specific Kendrick region (labeled as the "Media" below 1200 m/z in Figure 4). While these ions remain unannotated, their high oxygen content (suggested by mass defect) indicates likely origin from agar components, particularly dextrose. Spatially, these ions are most abundant around—but not within—the bacterial colonies. In contrast, foreground-specific (FG) ions span several molecular families. These include:HCCA-derived ions from the MALDI matrix, Phospholipids, typically found within colony interiors, and

Lipopeptides, such as orfamides and sessilins, known to be produced by *Pseudomonas* strains. Interestingly, lipopeptide ions exhibit diverse spatial behaviors: some (e.g., orfamides) diffuse outward from the colonies, while others (e.g., sessilins) appear more localized within the colony area. Despite this spatial heterogeneity, all lipopeptide ions are still classified as FG-specific by the model. This highlights an important strength of our approach: it can associate ions with biological origin even when their spatial distribution is only weakly correlated with the class labels (e.g., Spearman correlation near zero or even negative).

Similarly, HCCA ions—present across the entire MSI surface—show varied localization depending on their adduct type and molecular composition. Some are more abundant in the colonies, while others dominate the background.

Overall, the R-GCN + CAM approach enables the filtering of background ions (e.g., media-related) and the classification of biologically relevant molecular families. Importantly, the method can identify class-associated ions even when their spatial correlation is weak, reflecting a better understanding of molecular similarity rather than relying solely on spatial co-localization.

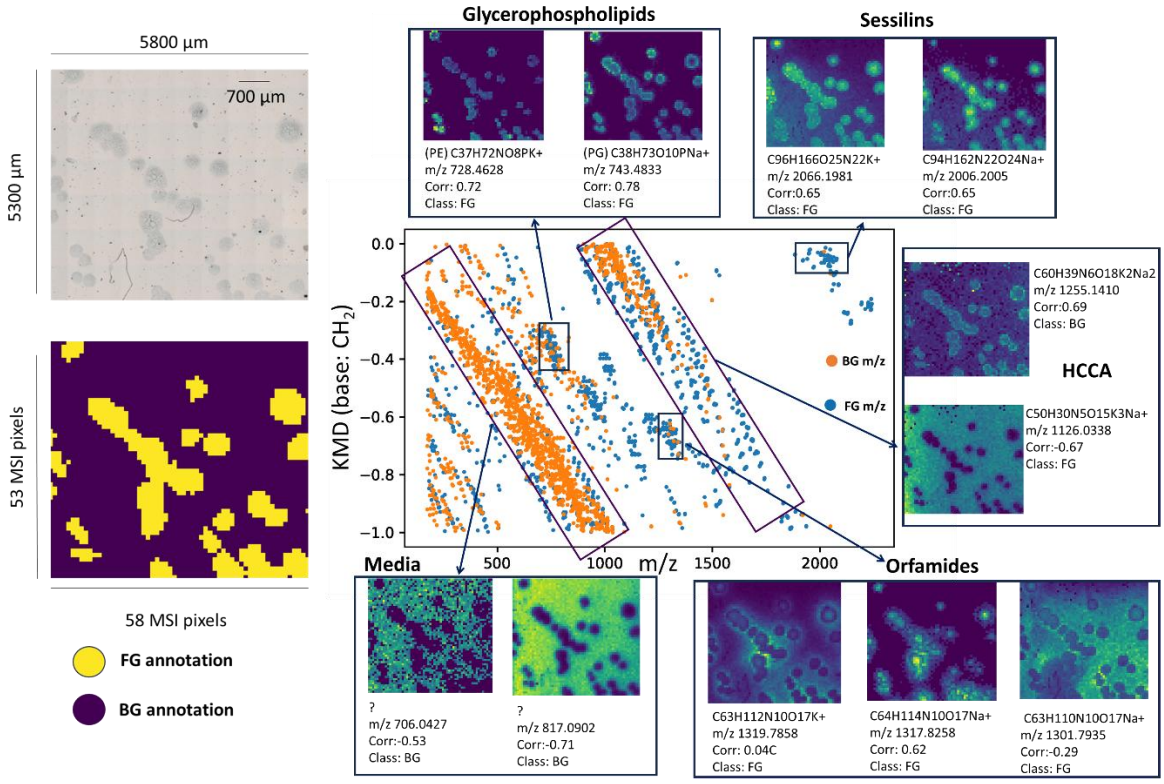


Figure 4: *On the left, a brightfield microscope image of the colonies (top) and the corresponding annotation of the BG and FG in the MSI (bottom). On the right (b), a Kendrick plot in CH2 base of the most specific signals selected by CAM for each class: FG (blue) and BG (orange) of the Pseudomonas colonies MSI. Different molecular families are highlighted with rectangles. Surrounding the Kendrick plot, we display various representative ion images as examples of the different molecular families.*

4.3.2 Analysis of m/z classification

To assess the consistency of m/z classification across different model initializations, we trained five R-GCN models with distinct random seeds and computed pairwise accuracies between their predicted class labels for each m/z value. For the Pseudomonas dataset, the models showed strong agreement, with an average accuracy of 0.98 (see Supporting Information S11. 3 Table S5). Similar levels of consistency were observed in the other datasets. These results confirm that our ion-level predictions are robust to random initialization. Based on this consistency, we selected a representative model per dataset for downstream interpretability analyses.

The m/z values classified by R-GCN + CAM are visualized in Figure 5 (left column) using Kendrick plots. For comparison, we also applied a fold-change (FC) analysis, where FC of an m/z is defined as the ratio between the mean intensity of that m/z in the experimental condition and the control. An m/z is considered specific to the experimental condition if FC > 1, or to the control if FC < 1. These results are shown in Figure 5 (right column). We compare the outputs of R-GCN + CAM and FC across all three datasets (Supplementary information S4):

Pseudomonas dataset. In the Kendrick plot, notable differences are observed between the two methods. The orfamide region, a known lipopeptide family produced by Pseudomonas, is correctly associated with the foreground (colony) class by R-GCN, whereas FC misclassifies many of these ions (orange points in the orfamide region).

S. scabiei Datasets (Cellobiose and Iron-Depleted Conditions). In both datasets, R-GCN + CAM identifies a greater number of ions associated with the control condition than FC does, as indicated by the higher number of orange points. While it may appear that CAM removes too many experimental-condition ions (e.g., fewer blue points), it successfully retains key m/z values corresponding to Thaxtomin A/B and siderophores (Desferrioxamines, Scabichelin) identified in the original study (see Supporting Information S4).

Hepatocyte (NAFLD) Dataset. In this dataset, R-GCN and FC produce broadly similar results. Both methods associate triglycerides (TG, 800–1000 m/z, KMD −0.2 to 0) and diglycerides (DG, 550–700 m/z, KMD −0.2 to 0) with the experimental condition (+FA), consistent with findings from the original SpaceM paper. One difference is that R-GCN classifies a broader range of DG (19/20 DG) ions as experimental-specific, whereas FC shows more variability across the same region (13/20 DG).

4.3.4 Summary and Implications for MSI Interpretation

Overall, these results demonstrate that R-GCN + CAM not only provides interpretable and chemically informed ion-level classification but also effectively removes media-associated ions and misclassified ions that are often retained by fold-change-based methods. This

highlights the value of incorporating structural mass relationships into MSI interpretation beyond traditional statistical comparisons.

Fold-change analysis and the R-GCN + CAM approach rely on fundamentally different principles. Fold change considers only intensity differences between conditions, without regard to mass structure. In contrast, our R-GCN model—configured in this experiment to ignore intensities by setting all node values to 1—relies solely on mass-based features such as mass defect and nominal mass differences. This structural perspective enables the model to identify entire molecular families that share characteristic mass patterns, even if not all members are overexpressed in the experimental condition.

For example, in the Pseudomonas dataset, the model correctly classifies the orfamide lipopeptide family as foreground-specific, even though some orfamide ions are diffusely distributed and weakly spatially correlated with the colony region. Similarly, in the hepatocyte dataset, the model associates most diglycerides (DG) and triglycerides (TG) with the NAFLD condition, despite not all members of these lipid families being overexpressed individually.

One of the advantages of the R-GCN + CAM strategy lies in its ability to cluster and identify ions from the same molecular family, even when their spatial behavior differs. This stems from the graph representation: during graph convolutions, node features are propagated and aggregated from neighboring nodes based on their connectivity and edge types. As a result, CAM scores reflect not only a node's own characteristics but also those of chemically related ions nearby in the graph. This makes it possible for a weakly class-specific ion to be correctly classified if it is embedded within a chemically coherent neighborhood.

While these preliminary results suggest that R-GCN + CAM can aid manual interpretation and discovery of ions of interest, we acknowledge that the clear advantage of our approach was most apparent in the *Pseudomonas* dataset, where the model uniquely highlighted orfamides. In the other two case studies, both R-GCN and FC identified the major ions described by the original authors, showing comparable performance on known signals. However, the R-GCN occasionally highlights weakly spatially correlated ions not previously reported—molecules that could merit further investigation but for which we currently lack ground truth validation.

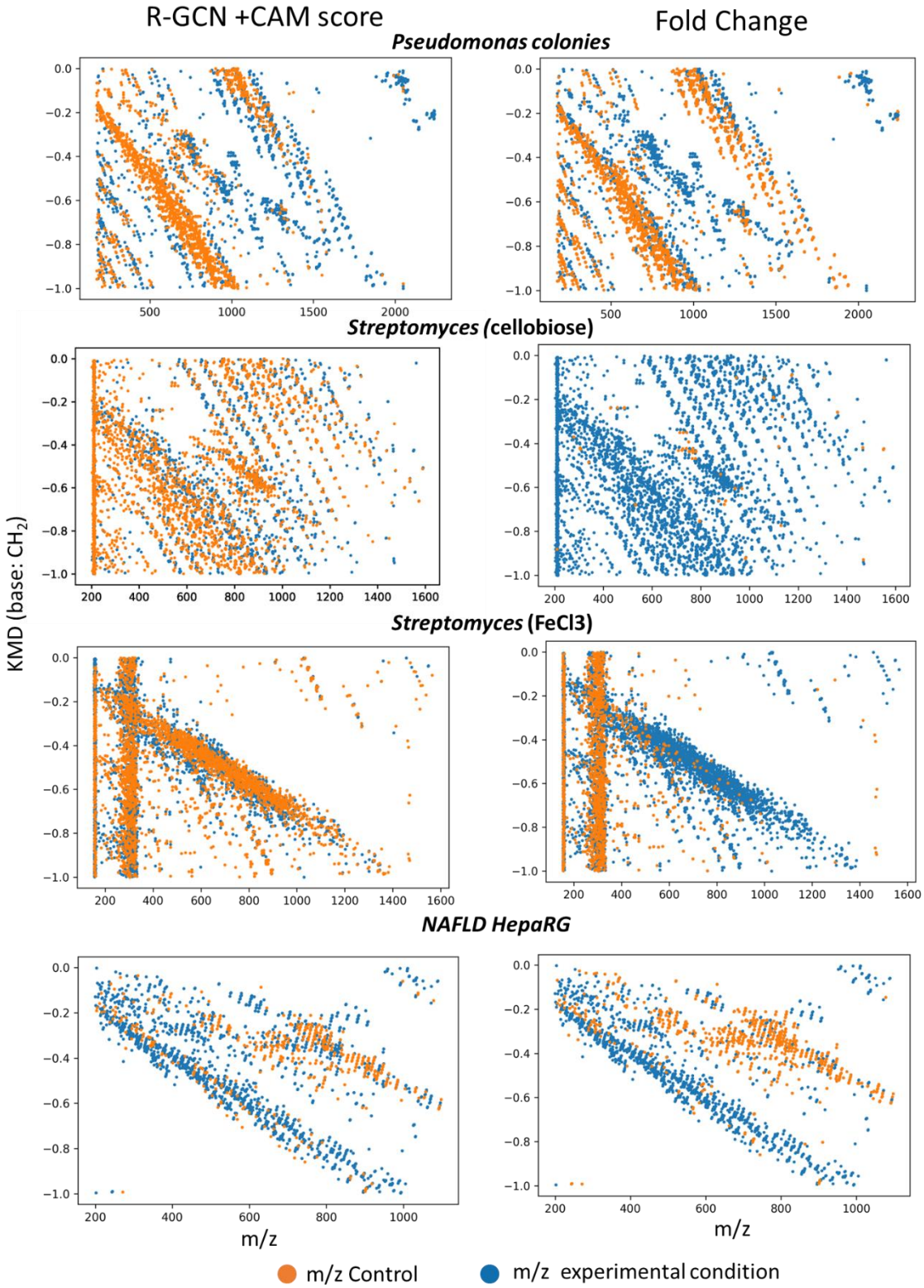

R-GCN +CAM score
Fold Change
Pseudomonas colonies
Streptomyces (cellobiose)
Streptomyces (FeCl3)
NAFLD HepaRG
KMD (base: CH2)
m/z
m/z Control
m/z experimental condition

Figure 5: This figure presents the comparison of m/z values extracted across four key datasets: Pseudomonas colonies, Streptomyces scabiei grown with cellobiose, Streptomyces scabiei treated with FeCl3, and human HepaRG hepatocytes in a NAFLD context. The data were analyzed using two distinct methods: the R-GCN+CAM and fold change calculations to identify discriminant m/z values across various conditions. In the Kendrick plot, m/z values are categorized based on their association with experimental conditions. Values related to the control condition are marked in orange, while those associated with experimental conditions are represented in blue. This visual representation enables the swift identification of m/z values related to specific biological conditions.

5. Conclusion and perspectives

This work is pioneering because it exploits HRMS mass features to classify spectra. To achieve this, we integrate HRMS mass features using a Relational Graph Convolution Network (R-GCN), representing a mass spectrum as a graph. We show that our method outperforms diverse classification methods when subjected to signal variations, especially in the case of mass shift and signal mixing, which is often the case in MSI experiments. Finally, we propose to integrate a CAM strategy into our model to interpret how each m/z value is used for predictions. We demonstrate that our model combines ions belonging to different known molecular families in a biologically relevant way. Based on these results discussed through concrete and meaningful examples, we recommend the scientific community explore using our R-GCN CAM as a new tool for extracting biologically meaningful m/z from MSI mass spectra.

Finally, we would like to put the novel approach proposed in this work into perspective. A key feature of our model is using a set of common mass differences. This set is composed of mass patterns frequently identified on MALDI MSI data. The mass differences between HCCA and DHB are already coded in this list. However, by adding the corresponding mass difference to the list, one can easily adapt the list to other mass patterns (mass differences), such as other MALDI matrices or other chemical transformations.

The flexibility of our approach also opens new opportunities for integration with other types of mass spectrometry data. For instance, information from electrophoretic or chromatographic techniques (e.g., retention time), as well as from ion mobility spectrometry (e.g., drift time or collision cross section), could be incorporated seamlessly as additional node features—alongside m/z, mass defect, and intensity. This would extend the applicability of R-GCNs beyond MSI, offering value to researchers working with LC-MS or ion mobility-MS for tasks such as sample classification or signal interpretation.

Furthermore, we demonstrate that our model, when combined with Class Activation Mapping (CAM), can extend beyond spectrum-level classification to interpret and annotate individual m/z features. While we currently use a model trained for spectrum classification and apply CAM post hoc to identify informative ions, the same architecture could be adapted for direct m/z-level classification. This could be achieved by removing the pooling layer and applying the classification head directly to the ion embeddings after the graph convolution layers.

Such a setup could, for example, be used to train a model to identify contaminant ions. By training on a small curated list of known contaminants, the R-GCN could potentially generalize to recognize other, previously unseen contaminant m/z values that share similar structural properties. This would provide a powerful tool for automated noise filtering and data cleaning in MSI and other MS-based workflows.

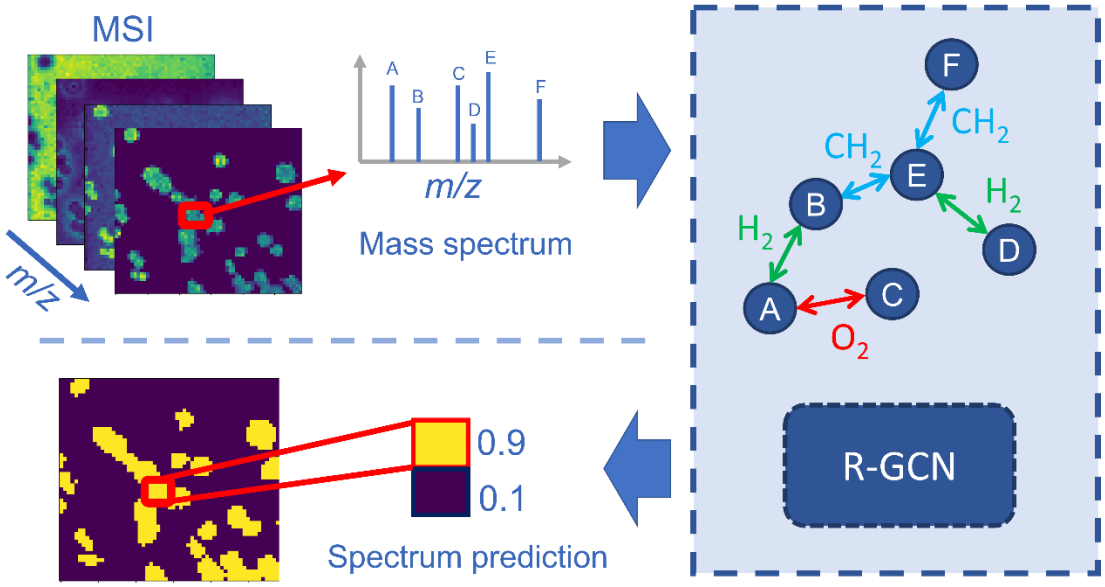
Overall, our results establish the foundation for a new generation of deep learning models in mass spectrometry—models that do not just classify but also interpret, generalize, and extend to new tasks in chemically meaningful ways.

Acknowledgement

The authors acknowledge financial support from the European Union's Horizon 2020 research and innovation program under grant agreement no. 731077 (EU FT-ICR MS project, INFRAIA-02-2017) from the European Union and Wallonia program FEDER BIOMED HUB Technology Support (no. 2.2.1/996) for the funding of the SolariX XR 9.4T, support from the Excellence Of Science Program of the FWO/FNRS F.R.S (Rhizoclip – EOS ID 30650620), and from the Interreg

EMR project: EURLIPIDS (R8598). Anthony Cioppa is funded by the FNRS (https://www.frs-fnrs.be/en/). We thank the METASPACE submitters who shared their data publicly and helped us create our dataset: Dusan Velickovic (PNNL), Jessica Lukowski (LungMAP), Don Nguyen (EMBL), Veronika Saharuka (EMBL).

TOC Graphic



Supporting information

S1: Description and id of the 4 METASPACE MSI data according to the initial uploader and an additional dataset from our laboratory.

S2: Background and Foreground annotation for each selected pixel of the 4 METASPACE MSI.

S3: Set of mass differences of the most frequent mass patterns in MALDI-MSI experiments from the work of Janda M. et al1.

S4: Annotated m/z of all dataset and their computed CAM values, Fold change and correlation.

S5: Ion images corresponding to the annotated m/z in S4, Pseudomonas.

S6 (pages S-2 - S-9): Material and Sample annotations.

S7 (page S-10): R-GCN Implementation Description

S8 (pages S-11 - S-12): Performance comparison of different R-GCN parameters.

S9 (page S-13): 1D-CNN, SVM, RF, LDA, XGBoost pre-processing.

S10 (pages S-14 - S-18): Evaluation of unseen data.

S11 (pages S-19 - S-21): Model evaluation details

References

(1) Seeley, E. H.; Caprioli, R. M. MALDI Imaging Mass Spectrometry of Human Tissue: Method Challenges and Clinical Perspectives. *Trends in Biotechnology* **2011**, *29* (3), 136–143. https://doi.org/10.1016/j.tibtech.2010.12.002.

(2) Holzlechner, M.; Eugenin, E.; Prideaux, B. Mass Spectrometry Imaging to Detect Lipid Biomarkers and Disease Signatures in Cancer. *Cancer Reports* **2019**, *2* (6). https://doi.org/10.1002/cnr2.1229.

(3) Watrous, J. D.; Dorrestein, P. C. Imaging Mass Spectrometry in Microbiology. *Nat. Rev. Microbiol.* **2011**, *9* (9), 683–694. https://doi.org/10.1038/nrmicro2634.

(4) Hu, H.; Laskin, J. Emerging Computational Methods in Mass Spectrometry Imaging. *Adv. Sci.* **2022**, 2203339. https://doi.org/10.1002/advs.202203339.

(5) Rappez, L.; Stadler, M.; Triana, S.; Gathungu, R. M.; Ovchinnikova, K.; Phapale, P.; Heikenwalder, M.; Alexandrov, T. SpaceM Reveals Metabolic States of Single Cells. *Nat. Methods* **2021**, *18* (7), 799–805. https://doi.org/10.1038/s41592-021-01198-0.

(6) Buchberger, A. R.; DeLaney, K.; Johnson, J.; Li, L. Mass Spectrometry Imaging: A Review of Emerging Advancements and Future Insights. *Anal. Chem.* **2018**, *90* (1), 240–265. https://doi.org/10.1021/acs.analchem.7b04733.

(7) 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.

(8) Alexandrov, T. MALDI Imaging Mass Spectrometry: Statistical Data Analysis and Current Computational Challenges. *BMC Bioinf.* **2012**, *13* (S16), S11. https://doi.org/10.1186/1471-2105-13-S16-S11.

(9) Verbeeck, N.; Caprioli, R. M.; Van de Plas, R. Unsupervised Machine Learning for Exploratory Data Analysis in Imaging Mass Spectrometry. *Mass Spectrom. Rev.* **2020**, *39* (3), 245–291. https://doi.org/10.1002/mas.21602.

(10) Behrmann, J.; Etmann, C.; Boskamp, T.; Casadonte, R.; Kriegsmann, J.; Maaβ, P. Deep Learning for Tumor Classification in Imaging Mass Spectrometry. *Bioinformatics* **2018**, *34* (7), 1215–1223. https://doi.org/10.1093/bioinformatics/btx724.

(11) Xie, Y. R.; Castro, D. C.; Bell, S. E.; Rubakhin, S. S.; Sweedler, J. V. Single-Cell Classification Using Mass Spectrometry through Interpretable Machine Learning. *Anal. Chem.* **2020**, *92* (13), 9338–9347. https://doi.org/10.1021/acs.analchem.0c01660.

(12) Hanselmann, M.; Köthe, U.; Kirchner, M.; Renard, B. Y.; Amstalden, E. R.; Glunde, K.; Heeren, R. M. A.; Hamprecht, F. A. Toward Digital Staining Using Imaging Mass Spectrometry and Random Forests. *J. Proteome Res.* **2009**, *8* (7), 3558–3567. https://doi.org/10.1021/pr900253y.

(13) Tideman, L. E. M.; Migas, L. G.; Djambazova, K. V.; Patterson, N. H.; Caprioli, R. M.; Spraggins, J. M.; Van de Plas, R. Automated Biomarker Candidate Discovery in Imaging Mass Spectrometry Data through Spatially Localized Shapley Additive Explanations. *Anal. Chim. Acta* **2021**, *1177*, 338522. https://doi.org/10.1016/j.aca.2021.338522.

(14) Xia, J.; Wishart, D. S. Using MetaboAnalyst 3.0 for Comprehensive Metabolomics Data Analysis. *Curr. Protoc. Bioinf.* **2016**, *55* (1), 14. https://doi.org/10.1002/cpbi.11.

(15) Yang, C.; He, Z.; Yu, W. Comparison of Public Peak Detection Algorithms for MALDI Mass Spectrometry Data Analysis. *BMC Bioinf.* **2009**, *10* (1), 4. https://doi.org/10.1186/1471-2105-10-4.

(16) Balluff, B.; Elsner, M.; Kowarsch, A.; Rauser, S.; Meding, S.; Schuhmacher, C.; Feith, M.; Herrmann, K.; Röcken, C.; Schmid, R. M.; Höfler, H.; Walch, A.; Ebert, M. P. Classification of HER2/Neu Status in Gastric Cancer Using a Breast-Cancer Derived Proteome Classifier. *J. Proteome Res.* **2010**, *9* (12), 6317–6322. https://doi.org/10.1021/pr100573s.

(17) Shimma, S.; Sugiura, Y. Effective Sample Preparations in Imaging Mass Spectrometry. *Mass Spectrom.* **2014**, *3* (Special_Issue), S0029–S0029. https://doi.org/10.5702/massspectrometry.S0029.

(18) Balluff, B.; Hopf, C.; Porta Siegel, T.; Grabsch, H. I.; Heeren, R. M. A. Batch Effects in MALDI Mass Spectrometry Imaging. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (3), 628–635. https://doi.org/10.1021/jasms.0c00393.

(19) Veselkov, K.; Sleeman, J.; Claude, E.; Vissers, J. P. C.; Galea, D.; Mroz, A.; Laponogov, I.; Towers, M.; Tonge, R.; Mirnezami, R.; Takats, Z.; Nicholson, J. K.; Langridge, J. I. BASIS: High-Performance Bioinformatics Platform for Processing of Large-Scale Mass Spectrometry Imaging Data in Chemically Augmented Histology. *Sci. Rep.* **2018**, *8* (1), 4053. https://doi.org/10.1038/s41598-018-22499-z.

(20) Deininger, S.-O.; Cornett, D. S.; Paape, R.; Becker, M.; Pineau, C.; Rauser, S.; Walch, A.; Wolski, E. Normalization in MALDI-TOF Imaging Datasets of Proteins: Practical Considerations. *Anal. Bioanal. Chem.* **2011**, *401* (1), 167–181. https://doi.org/10.1007/s00216-011-4929-z.

(21) 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.

(22) Tiquet, M.; La Rocca, R.; Kirnbauer, S.; Zoratto, S.; Van Kruining, D.; Quinton, L.; Eppe, G.; Martinez-Martinez, P.; Marchetti-Deschmann, M.; De Pauw, E.; Far, J. FT-ICR Mass Spectrometry Imaging at Extreme Mass Resolving Power Using a Dynamically Harmonized ICR Cell with 1ω or 2ω Detection. *Anal. Chem.* **2022**, *94* (26), 9316–9326. https://doi.org/10.1021/acs.analchem.2c00754.

(23) Inglese, P.; Correia, G.; Takats, Z.; Nicholson, J. K.; Glen, R. C. SPUTNIK: An R Package for Filtering of Spatially Related Peaks in Mass Spectrometry Imaging Data. *Bioinformatics* **2019**, *35* (1), 178–180. https://doi.org/10.1093/bioinformatics/bty622.

(24) Seddiki, K.; Saudemont, P.; Precioso, F.; Ogrinc, N.; Wisztorski, M.; Salzet, M.; Fournier, I.; Droit, A. Cumulative Learning Enables Convolutional Neural Network Representations for

Small Mass Spectrometry Data Classification. *Nat. Commun.* **2020**, *11* (1), 5595. https://doi.org/10.1038/s41467-020-19354-z.

(25) LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521* (7553), 436–444. https://doi.org/10.1038/nature14539.

(26) Liu, J.; Osadchy, M.; Ashton, L.; Foster, M.; Solomon, C. J.; Gibson, S. J. Deep Convolutional Neural Networks for Raman Spectrum Recognition: A Unified Solution. *Analyst.* **2017**, *142* (21), 4067–4074. https://doi.org/10.1039/C7AN01371J.

(27) Kune, C.; McCann, A.; Raphaël, L. R.; Arias, A. A.; Tiquet, M.; Van Kruining, D.; Martinez, P. M.; Ongena, M.; Eppe, G.; Quinton, L.; Far, J.; De Pauw, E. Rapid Visualization of Chemically Related Compounds Using Kendrick Mass Defect As a Filter in Mass Spectrometry Imaging. *Anal. Chem.* **2019**, *91* (20), 13112–13118. https://doi.org/10.1021/acs.analchem.9b03333.

(28) Richardson, L. T.; Neumann, E. K.; Caprioli, R. M.; Spraggins, J. M.; Solouki, T. Referenced Kendrick Mass Defect Annotation and Class-Based Filtering of Imaging MS Lipidomics Experiments. *Anal. Chem.* **2022**, *94* (14), 5504–5513. https://doi.org/10.1021/acs.analchem.1c03715.

(29) Janda, M.; Seah, B. K. B.; Jakob, D.; Beckmann, J.; Geier, B.; Liebeke, M. Determination of Abundant Metabolite Matrix Adducts Illuminates the Dark Metabolome of MALDI-Mass Spectrometry Imaging Datasets. *Anal. Chem.* **2021**, *93* (24), 8399–8407. https://doi.org/10.1021/acs.analchem.0c04720.

(30) Simon, C.; Dührkop, K.; Petras, D.; Roth, V.-N.; Böcker, S.; Dorrestein, P. C.; Gleixner, G. Mass Difference Matching Unfolds Hidden Molecular Structures of Dissolved Organic Matter. *Environ. Sci. Technol.* **2022**, *56* (15), 11027–11040. https://doi.org/10.1021/acs.est.2c01332.

(31) Zhou, B.; Khosla, A.; Lapedriza, A.; Oliva, A.; Torralba, A. Learning Deep Features for Discriminative Localization. In *IEEE Int. Conf. Comput. Vis. and Pattern Recogn. (CVPR)*; 2016; pp 2921–2929.

(32) Perneel, M.; Heyrman, J.; Adiobo, A.; De Maeyer, K.; Raaijmakers, J. M.; De Vos, P.; Höfte, M. Characterization of CMR5c and CMR12a, Novel Fluorescent *Pseudomonas* Strains from the Cocoyam Rhizosphere with Biocontrol Activity. *Journal of Applied Microbiology* **2007**, *103* (4), 1007–1020. https://doi.org/10.1111/j.1365-2672.2007.03345.x.

(33) Burguet, P.; La Rocca, R.; Kune, C.; Tellatin, D.; Stulanovic, N.; Rigolet, A.; Far, J.; Ongena, M.; Rigali, S.; Quinton, L. Exploiting Differential Signal Filtering (DSF) and Image Structure Filtering (ISF) Methods for Untargeted Mass Spectrometry Imaging of Bacterial Metabolites. *J. Am. Soc. Mass Spectrom.* **2024**, *35* (8), 1743–1755. https://doi.org/10.1021/jasms.4c00129.

(34) Schlichtkrull, M.; Kipf, T. N.; Bloem, P.; van den Berg, R.; Titov, I.; Welling, M. Modeling Relational Data with Graph Convolutional Networks. In *Extended Semantic Web Conference*; Gangemi, A., Navigli, R., Vidal, M.-E., Hitzler, P., Troncy, R., Hollink, L., Tordai, A., Alam, M., Eds.; Springer International Publishing: Cham, 2018; pp 593–607.

(35) Fey, M.; Lenssen, J. E. Fast Graph Representation Learning with PyTorch Geometric. **2019**. https://doi.org/10.48550/ARXIV.1903.02428.

(36) Andric, S.; Meyer, T.; Rigolet, A.; Prigent-Combaret, C.; Höfte, M.; Balleux, G.; Steels, S.; Hoff, G.; De Mot, R.; McCann, A.; De Pauw, E.; Arias, A. A.; Ongena, M. Lipopeptide Interplay Mediates Molecular Interactions between Soil Bacilli and Pseudomonads. *Microbiol. Spectrum* **2021**, *9* (3), 17. https://doi.org/10.1128/spectrum.02038-21.

TOC graphic