

## LipidIN Installation and Launch Guide

After downloading the **LipidIN** project from GitHub, navigate to the LipidIN GUI folder. Open the file ‘code for launch UI.R’ using Rstudio (we recommend using R version 4.1 and above.).

### 1. Check Required Packages

Lines 1–10 of the script are used to check whether the required R packages are installed. Select these lines, then click **Run** in the top-right corner of your R console. Wait for the necessary packages to be installed. Besides, we have packaged and uploaded all necessary dependency packages to Zenodo. you can simply download and copy these packages into their R library directory, ensuring consistency in the required R packages.

### 2. Modify File Path

Once all the dependencies are installed, update the file path on **line 15** of the script. Replace the placeholder file path with the path to the downloaded ‘LipidIN\_2.0.0.1.tar.gz’ file. Note that here you need to write the full address of LipidIN\_2.0.0.1.tar.gz, the code is only an example, please modify the address according to the actual situation.

### 3. Install LipidIN

After updating the file path, run **lines 12–19** to install LipidIN. During the installation process, a pop-up window will prompt you to confirm the installation of private packages. Click **Yes** to proceed.

### 4. Launch the LipidIN UI

Finally, run **lines 22–23** of the script. This will launch the LipidIN user interface.

## Parameter input and precautions

### 1. Step 1: Input mzML File Path

In the first input field of the UI, enter the file path of the mzML file you wish to annotate. If necessary, you can use the **MSConvert** software to convert your files into the mzML format. The default parameters of MSConvert are sufficient for this conversion, and additional operations such as centralization are not required. Please ensure that there are no spaces or special characters in

the file path. Please note that the UI does not like long addresses, so please keep your address as accurate, concise, and free of special characters as possible.

## 2. Step 2: Input Spectral Library Path

In the second input field, provide the file path to the spectral library to be used for annotation. Note that the spectral library must be in **RDA format**. For instructions on converting an MSP-format spectral library to RDA format, refer to the guide: [How to Convert Your MSP Format Spectral Library to RDA Format](#)

(<https://github.com/LinShuhaiLAB/LipidIN/tree/main/How%20to%20Convert%20Your%20MSP%20Format%20Spectral%20Library%20to%20RDA%20Format>). Alternatively, you can use the pre-compiled RDA spectral library available on our GitHub repository: [LipidIN/LipidIN 4-level hierarchical library/pos\\_ALL.rda](#). As with the mzML file path, ensure that the spectral library path as accurate, concise, and free of special characters as possible.

## 3. Step 3: Select Filtering Threshold

Choose a filtering threshold to exclude peaks with intensities lower than a certain percentage of the maximum intensity in the spectrum. Recommended values for this threshold are **0.01**, **0.05**, or **0.1**.

## 4. Step 4: Select Ionization Mode

Select the ionization mode for your analysis. The available options are:

- Positive ionization mode
- Negative ionization mode ([M+HCOO]-)
- Negative ionization mode ([M+OAc]-)

## 5. Step 5: Set Mass Tolerance for Precursor Ions

Specify the mass-to-charge ratio (m/z) tolerance for precursor ions during spectral library matching. Input the desired tolerance value in this step.

## 6. Step 6: Set Mass Tolerance for Fragment Ions

Specify the mass-to-charge ratio (m/z) tolerance for fragment ions during spectral library matching. Input the desired tolerance value in this step.

## 7. Step 7: Select MS2 scan time or MS1 Rt to use for the LCI module.

MS1 retention time (RT) provides higher accuracy as it reflects the precise chromatographic peak maxima. However, it requires additional computational time for peak extraction and isotope removal. In contrast, MS2 scan time, while less accurate than MS1 RT, significantly reduces processing steps and improves analysis speed, making it a practical choice for high-throughput lipid annotation workflows.

## 8. Step 8: Fill in the XCMS peak picking parameters

For detailed explanations of the parameters and recommendations on their usage, please refer to the comprehensive documentation available at DOI: 10.18129/B9.bioc.xcms.

## 9. Step 9: Start the Annotation Process

After verifying the inputs for all the steps above, click the **Start** button to initiate the lipid annotation process.

The screenshot shows the 'Compound Annotations' section of the LipidIN Shiny application. The interface is divided into a left sidebar and a main content area. The sidebar has a 'Compound Annotation' section with a sub-section 'Annotation for EAD'. The main content area contains several input fields and a 'Start' button. The inputs are: 'Folder Path' (D:\bio\_inf\LipidINDesk\XCMS\test), 'Enter the Full Path of the File' (D:\bio\_inf\LipidINDesk\XCMS\library\_pos.rda), 'Abundance Filter (0-1)' (0), 'Ionization Mode' (Positive Ion Mode), 'MS1 Tolerance (ppm)' (10), 'MS2 Tolerance (ppm)' (40), 'Use MS1 Rt or MS2 scan time for LCI' (MS2 scan time), 'MS1 Rt' (unchecked), 'MS2 scan time' (checked), 'Peak width left' (10), 'Peak width right' (50), and 'Noise Level' (5). A 'Start' button is at the bottom. The steps are labeled on the right: Step1 (Folder Path), Step2 (Enter the Full Path of the File), Step3 (Abundance Filter), Step4 (Ionization Mode), Step5 (MS1 Tolerance), Step6 (MS2 Tolerance), Step7 (Use MS1 Rt or MS2 scan time for LCI), Step8 (Peak width left, Peak width right, Noise Level), and Step9 (Start button).

## LipidIN Process Instructions

### 1. Package Loading and Installation

After clicking "Run," the first message (Message 1) will indicate that the system



errors at this step, please ensure the following:

- Your mzML files contain MS2 spectra.
- The input file path does not contain spaces or special characters.

### **3. Library Loading**

After file conversion, the system will load the spectral library, which will be indicated by the third message (Message 3). This step can take a significant amount of time, especially if your spectral library contains millions of entries.

### **4. Spectral Matching**

Once the library is loaded, spectral matching will begin, and the fourth message (Message 4) will appear. This process is typically very fast. If it becomes stuck or an error occurs, check the format and content of your spectral library for accuracy.

### **5. Lipid Retention Time Analysis**

If MS2 scan time is selected, the peak extraction step will be bypassed, and the LCI module will proceed directly. Following spectral matching, the system will execute lipid retention time prediction, as indicated by Message 5. This step typically requires approximately 1–2 minutes to complete.

### **6. XCMS quantitateion**

Upon completion of the LCI step, the XCMS and CAMERA packages will be utilized for quantitative analysis, peak alignment, isotope removal, and gap filling.

### **7. Completion**

Finally, when all annotation processes are complete, the system will display the sixth message (Message 7).

### **8. Extra Message for using MS1 Rt**

If MS1 retention time (RT) is selected, peak extraction will be performed using XCMS prior to the retention time prediction and evaluation module.

Analysis Complete

The analysis is complete. Processed files are in D:\bio\_info\LipidDesk\UCMS\test

Enter the Full Path of the File

D:\bio\_info\LipidDesk\UCMS\test

Abundance Filter (0-1)

0

Ionization Mode

Positive Ion Mode

MS1 Tolerance (ppm)

10

MS2 Tolerance (ppm)

40

Use MS1 RT or MS2 scan time for L2?

MS2 scan time

MS2 scan time

Selecting the MS2 scan time will skip the peak detection approach for false positive removal.

Selecting the MS1 RT will utilize XCMS for peak detection. However, this method is more time-consuming than we only provide an interface to XCMS, and so process. In addition, due to the peak picking process not be a result of supplementary parameter settings.

Peak width left

10

Peak width right

50

Start

Verify if all the required packages are installed.

Waiting for package loading.

Working directory set to:

D:\bio\_info\LipidDesk\UCMS\test

Successfully processed file:

D:\bio\_info\LipidDesk\UCMS\test\NIST\_1000.ms

Successfully processed file:

D:\bio\_info\LipidDesk\UCMS\test\NIST\_1000.ms

Loading spectral library...

Successfully loaded spectral library.

Starting MS2 searching...

Waiting for XCMS peak detection...

Starting peak picking...

Successfully processed file:

D:\bio\_info\LipidDesk\UCMS\test\NIST\_1000.ms

Starting peak picking...

Successfully processed file:

D:\bio\_info\LipidDesk\UCMS\test\NIST\_1000.ms

Starting L2 retention time prediction and judgment...

Successfully processed file:

D:\bio\_info\LipidDesk\UCMS\test\NIST\_1000.ms

Merge all annotations.

Quantitative\_NIST\_1000.ms

Quantitative\_NIST\_1000.ms

Message1

Message2

Message3

Message4

Message5

Message6

Message7

Analysis Complete

The analysis is complete. Processed files are in D:\bio\_info\LipidDesk\UCMS\test

Enter the Full Path of the File

D:\bio\_info\LipidDesk\UCMS\test

Abundance Filter (0-1)

0

Ionization Mode

Positive Ion Mode

MS1 Tolerance (ppm)

10

MS2 Tolerance (ppm)

40

Use MS1 RT or MS2 scan time for L2?

MS2 scan time

MS2 scan time

Selecting the MS2 scan time will skip the peak detection approach for false positive removal.

Selecting the MS1 RT will utilize XCMS for peak detection. However, this method is more time-consuming than we only provide an interface to XCMS, and so process. In addition, due to the peak picking process not be a result of supplementary parameter settings.

Peak width left

10

Peak width right

50

Start

Working directory set to:

D:\bio\_info\LipidDesk\UCMS\test

Successfully processed file:

D:\bio\_info\LipidDesk\UCMS\test\NIST\_1000.ms

Successfully processed file:

D:\bio\_info\LipidDesk\UCMS\test\NIST\_1000.ms

Loading spectral library...

Successfully loaded spectral library.

Starting MS2 searching...

Waiting for XCMS peak detection...

Starting peak picking...

Successfully processed file:

D:\bio\_info\LipidDesk\UCMS\test\NIST\_1000.ms

Starting peak picking...

Successfully processed file:

D:\bio\_info\LipidDesk\UCMS\test\NIST\_1000.ms

Starting L2 retention time prediction and judgment...

Successfully processed file:

D:\bio\_info\LipidDesk\UCMS\test\NIST\_1000.ms

Merge all annotations.

Quantitative\_NIST\_1000.ms

Quantitative\_NIST\_1000.ms

Extra Message