DecoID User Interface Help Page

Overview:

DecoID offers deconvolution and searching of MS/MS spectra. In its simplist form, it averages all spectra acquired for a particular feature of interest and then searches those spectra against a metabolomic database (mzCloud, MoNA, HMDB) based on accurate mass and similarity to the MS/MS spectrum in the database. In addition, "chimeric" MS/MS spectra can be deconvolved where instead of searching against each database spectrum individually, all possible linear combinations of database spectra are considered and the combination that results in the most similar spectrum to the acquired spectrum is used to improve the identification. For details on how DecoID works, please consult the publication. In this help page, I will detail the parameters available in the user interface and walk through an example of how to use DecoID on a typical DDA dataset.

Search Parameters:

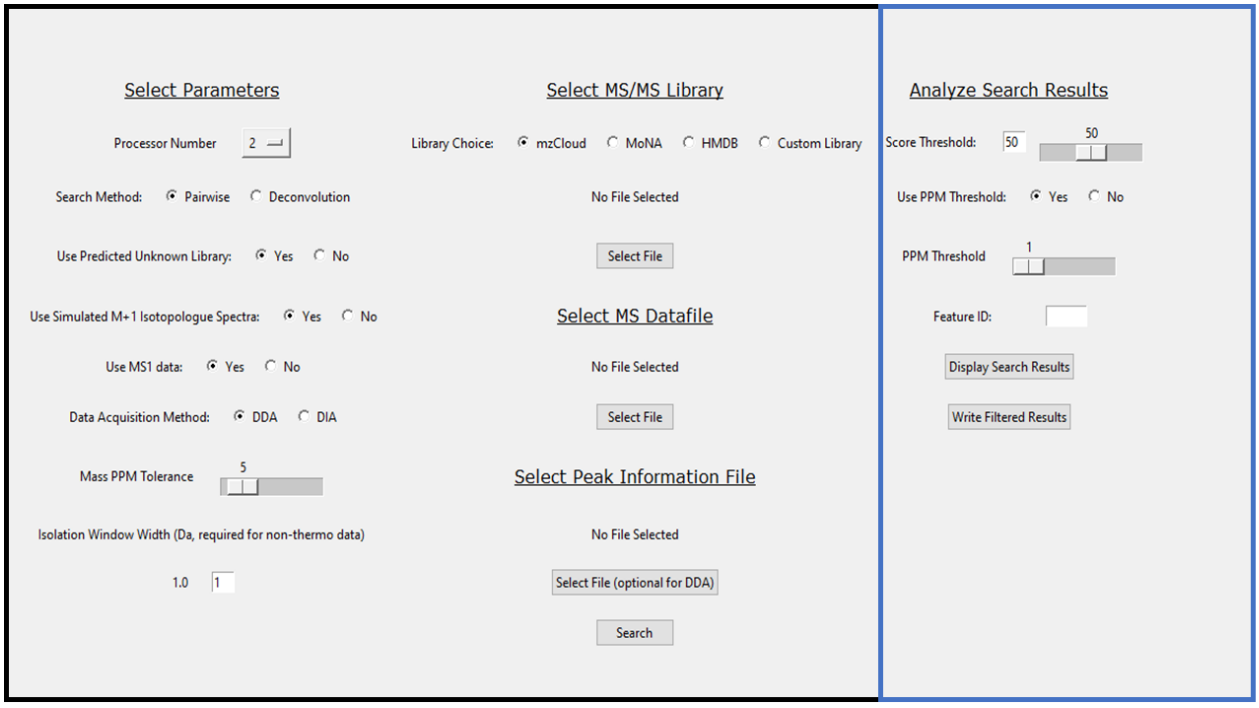
Below the search parameters are defined:

- **Processor Number**- # of concurrent processes to run at once. In other words, how many spectra do you want to search and deconvolve in parallel. This number should be set as high as you can afford based on the hardware. On most laptop 2-3 is a good choice. This will still leave you with enough computing power for other uses. On a normal desktop, 5-6 would likely work well. On a workstation computer, 10-15 will give fast performance. As a general rule, I try to set it to the # of processors in the computer – 1.
- **Search Method**- Pairwise: without deconvolution, simply performs spectral averaging and direct MS/MS search. Deconvolution: try to find an optimal linear combination of database spectra. If you suspect your data is chimeric (large isolation window, complex sample matrix), this will improve performance dramatically, but is slower than direct searching.
- **Use Predicted Unknown Library**- If toggled to "Yes", spectra with no match in the selected database, but are not chimeric, are used to deconvolve other spectra that were acquired. Usage of these spectra is restricted by the retention time of the "unknown" feature and the accurate mass. More details and examples can be found in the publication as the "on-the-fly unknown library". This can only be used in DDA workflows where MS1 full scan data is acquired along with the MS/MS spectra.
- **Use simulated M+1 isotopologue spectra**- If toggled "Yes", MS/MS fragments coming from orphan isotopologues are able to be deconvolved from a chimeric spectrum. This sort of contamination occurs when the parent M+0 is excluded from the isolation window, but the M+1 is not. The publication gives further details. In short, the M+1 isotopologue's MS/MS spectrum is modeled as a mixture of isotopomer spectra. Probable spectra for these isotopomers are predicted by shifting the top 5 most intense fragments in the M+0 spectrum by +1.003 m/z. This can only be used if MS1 full scan data is acquired along with the MS/MS as is common in DDA and DIA workflows.
- **Use MS1 data**- If toggled "Yes" and MS1 data was collected in addition to MS/MS, the nearest MS1 spectrum to each acquired MS/MS spectrum is used to filter which compounds to use with deconvolution by requiring that each compound used to reconstruct the chimeric MS/MS

spectrum have a peak at the corresponding precursors' m/z. This parameter is meaningless if "pairwise" is selected.

- **Data acquisition method**- Data dependent acquisition (DDA) for datasets where narrow isolation windows are used to target specific m/z values. In the DDA analysis workflow, the goal will be to identify the compounds targeted for MS/MS. In data independent acquisition (DIA). The goal is to identify all compounds in each spectrum which is assumed to be multiplexed. A common DIA workflow would be sequential acquisition of the entire m/z axis with 20 m/z isolation windows. A peak list is required for DIA (see below). Select the method that most closely matches your desired output and acquired data.
- **Mass PPM Tolerance-** This parameter determines the MS1 mass accuracy tolerance. This is used to assign spectra to features as well as filter database matches. This should be set to the mass accuracy of the instrument.
- **Isolation Window Width-** This parameter only needs to be entered for non-Thermo data. The value should be the width of the isolation window around each DDA spectrum (i.e. if the isolation window is 112-115 m/z, the entry should be 3.0).

Data Input:



search parameters

Visualization/analysis parameters