ElemCor Tutorial

1 Introduction

ElemCor is a software tool to correct LC-MS data in isotope labeling experiments. Due to the contribution of naturally occurring isotopes and isotopic impurity from the tracer, the fractional abundances of measured isotopologues (FAM) collected from the instrument must be corrected to obtain mass distribution vectors (MDV) that correspond to the contribution from isotopic labeling. ElemCor is based on the correction matrix[1] and uses two different approaches to account for resolution effect, mass different theory[2], and unlabeled samples[3]. ElemCor corrects and improves the numerical schemes of both methods.

2 Building Software (Optional)

ElemCor is built using MATLAB Compiler. Note that building the software is not required as compiled files are also included in the package. It is recommended to directly install the software using compiled files (Section 3) unless changes need to be made to the source code. Before building the software, one needs to be sure MATLAB Compiler is properly installed. This can be done by typing "ver" at the MATLAB prompt. To build the software, simply click on the project file "elemcor.prj". All other files required to run the software are automatically identified if they are placed under the same directory as the project file (Fig. 1). After building the software, one should be able to obtain two files "MyAppInstaller_web.exe" and "ElemCor.exe". The former is the installation file that installs MATLAB Compiler Runtime, and the latter is the binary file to run the software.
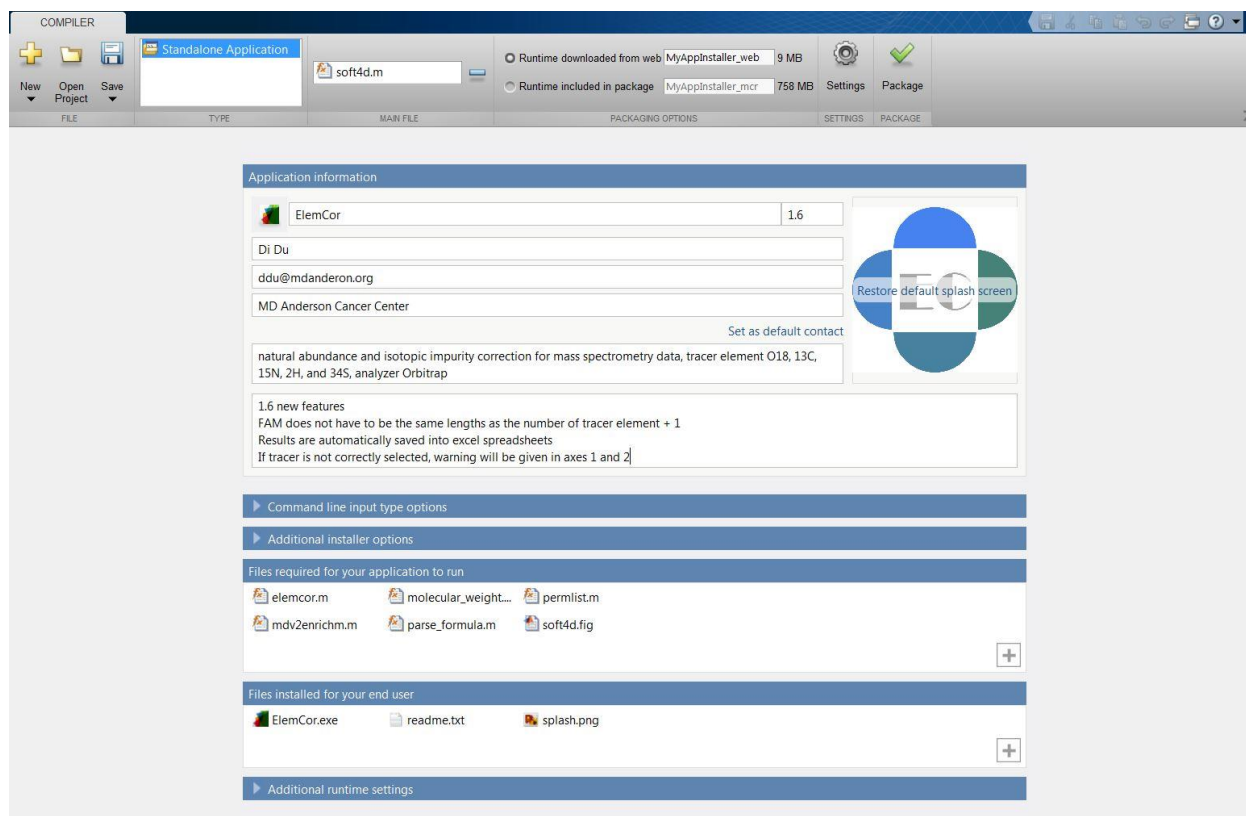
Figure 1. Building ElemCor using MATLAB Compiler

3 Installing Software

To run the software, one needs to install MATLAB Compiler Runtime, which is similar to Java Runtime and included in the package. Double-clicking "MyAppInstaller_web.exe" will initialize the installation. After it is complete, one can use ElemCor as an executable by double-clicking "ElemCor.exe" from the **download folder** (not the installation directory). The installation requires network connection, but after that the software tool can be used offline. MATLAB installation is not required to run the software.

4 Instruction

The software tool takes data in a spreadsheet (XLSX) format. In Steps 1 and 2, labeled and unlabeled data are loaded. Step 2 is optional, and when it is not performed, ElemCor runs based on mass difference theory only. In Step 3, isotopic purity of nutrient and nominal instrument resolution are specified. In Step 4, the tracer element is selected. In addition to 13C, 2H, and 15N, ElemCor allows 18O and 34S as the tracer element for correction. **Please always make sure the correct tracer is selected**. In Step 5, the mass analyzer is selected. Then the loaded data are analyzed and isotopic enrichment is calculated for each compound in Step 6. Users can select a cell in the data table, and FAM and MDV for the corresponding compound and sample before and after correction are shown in the figure above.

5 Examples

The package includes example data files from labeled samples "test_labelN_sim.xlsx" and "test_labelS_sim.xlsx", and the associated files from unlabeled samples are "test_unlabelN_sim.xlsx" and "test_unlabelS_sim.xlsx". They are simulated from XCalibur with N and S as trace elements, and at nominal instrument resolutions of 140,000 and 280,000, respectively. Both are from Orbitrap analyzer. The files show standard format of input files. Each column vector correspond to FAM of all compounds in each sample.

The files from the N tracer are used as an example to illustrate the workflow (Fig. 2). First select the labeled and unlabeled samples, "test_labelN_sim.xlsx" and "test_unlabelN_sim.xlsx", in Steps 1 and 2. Since they are simulated data for which isotopic impurity is not considered, isotopic purity is specified as 1 in Step 3. The resolution for N tracer data is 140,000. In Step 4, select N as tracer element. **Please always make sure the correct tracer is selected**. In Step 5, select Orbitrap as analyzer. The loaded data are analyzed in Step 6. After correction, the MDVs of corresponding samples are shown in the figure window in ElemCor and also stored in a different sheet in the same input file. More specifically, MDVs calculated using mass difference theory are shown in the table on the left, and stored in "test_labelN_sim.xlsx". MDVs calculated using unlabeled samples are shown in the table on the right, and stored in "test_unlabelN_sim.xlsx". Other quantities such as enrichment and fold change of pool size are also saved in the input files accordingly.

If no data are loaded in Step 2, after submission, the table on the left will show corrected MDVs, whereas the table on the right will be empty. No change will be made to "test_unlabelN_sim.xlsx" as it is not loaded.
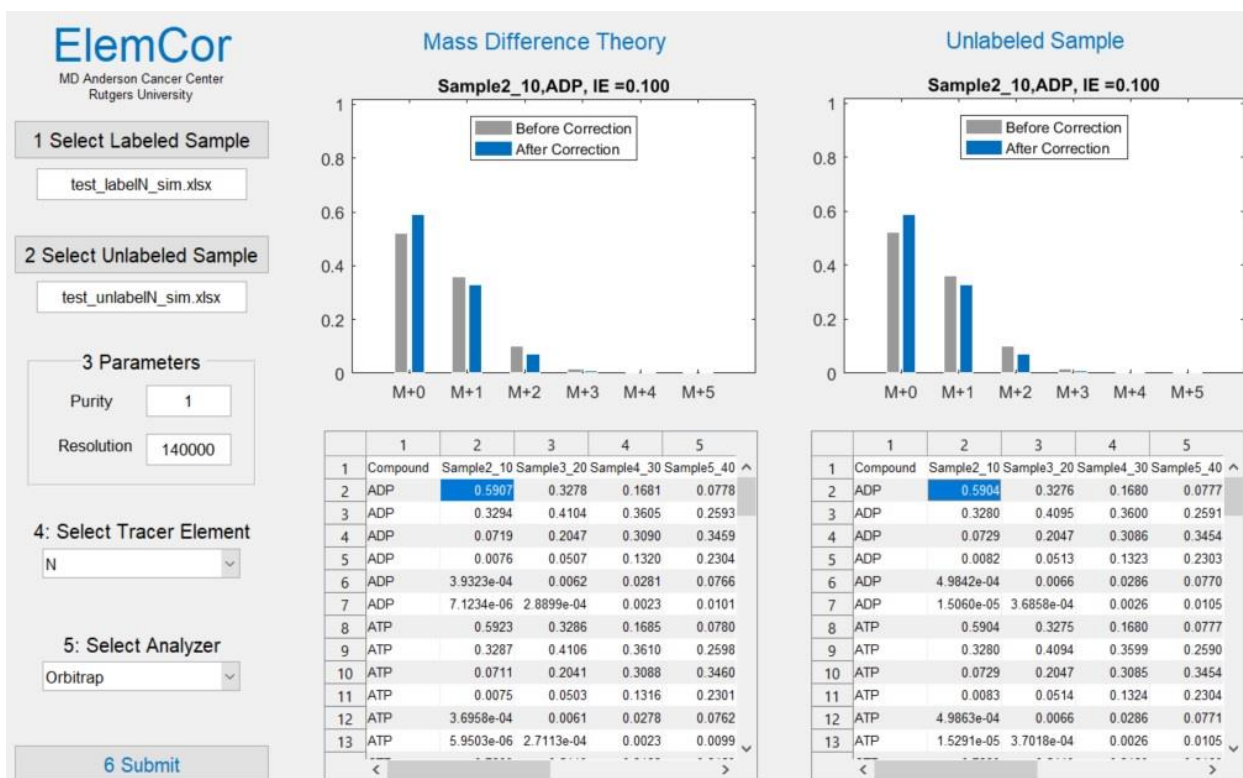

Figure 2. Example

6 Q&A

1. Where are the final results?

The final results are previewed in the graphic interface. They are automatically saved into additional sheets in the original XLSX file. The final results include MDV, enrichment, and fold change of pool size for different compounds and samples.

2. Can the FAM have a different length than one plus the number of tracer atoms?

FAM typically has a length of Nt + 1 to cover M+0, M+1, ..., M+Nt isotopologues. Here Nt stands for the number of tracer atoms. But ElemCor is able to handle FAM that has a length other than Nt + 1, and the answer is Yes.

3. What is the advantage of ElemCor over IsoCor?

IsoCor performs correction based on combinatorics without considering resolution effect. When instrument resolution is infinitely low, ElemCor is identical to IsoCor[1]. When instrument resolution is infinitely high, ElemCor performs no correction at all. Therefore, the amount of correction perform by ElemCor is monotonically dependent on instrument resolution.

4. What is the expected processing time?

The example files shown in Fig. 2 can be processed within 10 seconds on a 2.6GHz Core i7 processor, and 20 seconds on a 1.5GHz Core i5 processor. If the results do not come out within a minute, please check whether the correct tracer is selected, as the warning message suggests. If the 15N data are loaded, but C or S is selected as tracer, then the software tool would not run properly. This is because some compounds may not have C or S in their formula, and the algorithm will report error if the tracer element is not present in the compound.

7 References

1. Millard, P., et al. IsoCor: correcting MS data in isotope labeling experiments. *Bioinformatics* 2012; 28:1294-1296

2. Su, X., et al. Metabolite Spectral Accuracy on Orbitraps. *Analytical Chemistry* 2017; 89:5940–5948

3. Trefely, S., et al. FluxFix: automatic isotopologue normalization for metabolic tracer analysis. *BMC Bioinformatics* 2016; 17:485