A Protocol to Compare Methods for Untargeted Metabolomics

Lingjue Wang[1], **Fuad J. Naser**[1], **Jonathan L. Spalding**[1,2], and **Gary J. Patti**[1,2]

[1.]Department of Chemistry, Washington University, St. Louis, USA

[2.]Department of Medicine, Washington University School of Medicine, St. Louis, USA

Abstract

There are thousands of published methods for profiling metabolites with liquid chromatography/ mass spectrometry (LC/MS). While many have been evaluated and optimized for a small number of select metabolites, very few have been assessed on the basis of global metabolite coverage. Thus, when performing untargeted metabolomics, researchers often question which combination of extraction techniques, chromatographic separations, and mass spectrometers is best for global profiling. Method comparisons are complicated because thousands of LC/MS signals (so-called features) in a typical untargeted metabolomic experiment cannot be readily identified with current resources. It is therefore challenging to distinguish methods that increase signal number due to improved metabolite coverage from methods that increase signal number due to contamination and artifacts. Here, we present the *credentialing* protocol to remove the latter from untargeted metabolomic datasets without having to identify metabolite structures. This protocol can be used to compare or optimize methods pertaining to any step of the untargeted metabolomic workflow (e.g., extraction, chromatography, mass spectrometer, informatic software, etc.).

Keywords

Untargeted metabolomics; Metabolite profiling; Metabolism; *Credentialing*; Liquid chromatography; Mass spectrometry

1 Introduction

Metabolite profiling, or metabolomics, can be performed with either a targeted or an untargeted approach. In targeted metabolomics, a defined set of metabolites is analyzed. As such, these methods are relatively straightforward to optimize by using commercial standards [1]. In contrast to targeted metabolomics, which is effective at testing specific hypotheses, the objective of untargeted metabolomics is to measure as many metabolites in the sample as possible [2]. This systems-level assessment of metabolism is highly attractive because it can potentially reveal altered pathways that had not been previously anticipated [3]. Since the set of metabolites being profiled is not well defined and may include "unknown" compounds that have not yet been characterized, however, simple optimization of untargeted metabolomic methods with commercial standards is impractical.

In principle, untargeted metabolomic methods can be optimized by maximizing the number of metabolites detected in an experiment. In practice, this approach is complicated because it is typical to detect thousands of signals (or features) in an untargeted metabolomic

experiment that cannot be identified with current informatic workflows [4]. While some of these unidentified signals arise from bona fide metabolites in the biological sample, others do not and therefore the total number of detected signals does not reliably correlate with metabolome coverage [5]. Non-biological signals arise due to contamination and artifacts. Contaminants are chemical impurities introduced during sample handling and LC/MS analysis (e.g., solvent impurities, plastic leechables in the extraction process, carry over from previous experiments, etc.). Artifacts are signals that result from informatic errors. In some untargeted metabolomic experiments, contaminants and artifacts may represent a major fraction of the total LC/MS signals detected [6].

Considering the number of extraction protocols, separation methods, mass spectrometers, and informatic software packages, there are thousands of potential workflows for performing untargeted metabolomics [7–11]. A fundamental question is which combination of techniques, instrumentation, and equipment settings is best for achieving comprehensive metabolome coverage. New laboratories trying to establish an untargeted metabolomic platform, for example, may wish to compare the performance of different mass spectrometers prior to committing financial resources.

Historically, metabolite coverage in untargeted metabolomics has often been benchmarked by the total number of signals detected. It is important to emphasize that such an approach is highly unreliable because the total number of signals and the total number of metabolites poorly correlate [5]. By way of illustration, a method using dirty vials may increase the number of signals (as a result of contamination) without increasing the actual number of metabolites assayed. As an alternative approach to better assess metabolome coverage in LC/MS experiments without having to identify each signal, we present the *credentialing* protocol using *E. coli* as a representative complex biological sample. In brief, *credentialing* mixes uniformly labeled *E. coli* extracts with natural-abundance *E. coli* extracts at different ratios. Signals in the LC/MS data that correspond to bona fide metabolites will have an isotopic dance partner at the appropriate ratio, whereas signals corresponding to contaminants and artifacts will not. After removing contaminants and artifacts, the number of remaining signals can then be used as a better estimate of metabolite coverage between different untargeted metabolomic methods [12].

The following protocol consists of two main parts. First, we describe how to make *credentialed* samples (this part can be skipped when using a purchased *credentialing* kit). Then, we describe in detail how to process *credentialing* data. The output of this work-flow will be a list of signals or features from which non-*credentialed* contaminants and artifacts have been removed. We note that the approach we present for data analysis can be applied to any set of appropriately labeled biological materials (e.g., plants, mammalian cells, animal tissues, etc.). Given the costs and experimental challenges of generating such samples, however, we focus on using *E. coli* here as a representative complex biological matrix for method optimization. *E. coli* samples that have been labeled for *credentialing* analysis can be commercially obtained from Cambridge Isotope Laboratories.

2 Materials

Prepare all solutions and perform cell culture with ultrapure water. All glassware referred to in this protocol should be sterile unless noted otherwise. Sterilization can be achieved prior to *E. coli* growth by thoroughly rinsing glassware with ultrapure water, covering with aluminum foil, and baking in an oven at 250 °C for 3–5 h.

2.1 Purchasing Metabolic Extracts of Credentialed E. coli

Metabolic extracts of *credentialed E. coli* (strain K12, MG1655) may be purchased from Cambridge Isotope Laboratories (Tewks-bury, MA, USA). Cambridge offers two *Credentialed E. coli* Cell Extract products: an extract kit in solution (Item #: MSK-CREDKIT) and a dried down extract kit (Item #: MSK-CRED-DD-KIT). Both kits contain separate vials of unlabeled extract and uniformly ^{13}C-labeled extract. Products in solution have been suspended in acetonitrile:water (1:1). *E. coli* have been extracted as previously detailed [13].

2.2 Media Preparation for Credentialed E. coli Growth

1. 5× M9 minimal salts. Final concentration: 33 g/L disodium phosphate, 15 g/L monopotassium phosphate, 2.5 g/L sodium chloride, and 5 g/L ammonium chloride.

2. Lennox B broth powder (10 g/L enzymatic digest of casein, 5 g/L yeast extract, and 5 g/L sodium chloride).

3. 1 M magnesium sulfate.

4. 0.1 M calcium chloride.

5. Ultrapure water.

6. Two 500 mL Erlenmeyer flasks.

7. 500 mL media filter with a 0.22 μm pore size.

8. 250 mL media filter with a 0.22 μm pore size.

9. Parafilm.

10. Vacuum suction pump.

2.3 Initial E. coli Overnight Growth

1. *E. coli* stock strain K12, MG1655. Store at −80 °C and do not let the stock thaw.

2. Styrofoam box.

3. Ice.

4. Two 250 mL Erlenmeyer flasks.

5. 50 mL LB media.

6. Sterile Pasteur pipets.

7.   Rotary shaker.

2.4   Credentialed E. coli Growth

1.   Two 10 mL beakers.

2.   Natural abundance D-glucose (200 mg).

3.   Uniformly labeled ^{13}C-D-glucose (207 mg).

4.   Ultrapure water.

5.   Charged auto-pipetter.

6.   Two 25 mL disposal plastic pipettes.

7.   M9 media (220 mL).

8.   Two 0.22 μm syringe filters.

9.   Two 3 mL plastic syringes.

10.   Two 1 L Erlenmeyer flasks.

11.   Rotary shaker.

12.   Pipette and pipette tips (1 mL).

13.   UV-Vis spectrophotometer.

14.   Spectrophotometer cuvettes.

2.5   Credentialed E. coli Harvest

1.   Two 50 mL conical tubes.

2.   Centrifuge for 50 mL conical tubes.

3.   Charged auto-pipetter.

4.   Disposable plastic pipettes (25 mL or 50 mL).

5.   Liquid nitrogen.

6.   Ice.

7.   Lyophilizer.

8.   Eppendorf tubes.

2.6   LC/MS

1.   LC/MS-grade solvents (e.g., water, acetonitrile, methanol, and isopropanol). We recommend solvents from Honeywell Burdick & Jackson (Muskegon, MI, USA).

2.   LC/MS-grade eluent additives such as ammonium acetate. We recommend purchasing mobile phase additives from Sigma Aldrich (St. Louis, MO, USA).

3.   LC/MS method of your choice.

2.7 Data Processing and Analysis

1. R and RStudio. *See* Subheading 4.1 for more details about R installation.

2. The following R packages: *Credential3.1, data.table, utils,* and *xcms* (or an equivalent software package for generating a features table).

3. ProteoWizard MSConvert (if using the XCMS workflow).

3 Methods

3.1 E. coli Media Preparation

1. Prepare M9 media and LB media 1 day before *E. coli* harvest.

2. Combine 500 mL of ultrapure water and 5.5 g of 5× M9 minimal salts in a sterile 500 mL Erlenmeyer flask.

3. Cover with parafilm and mix by inversion. Ensure that salts are fully dissolved before filtration step (about 5 min).

4. Add via pipette 500 μL of 1 M $MgSO_4$ and 250 μL of 0.1 M $CaCl_2$.

5. Cover with parafilm and mix by inversion.

6. Filter via vacuum suction through a 500 mL media filter bottle.

7. In a separate 500 mL Erlenmeyer flask, combine 250 mL ultra-pure water and 5 g of LB powder.

8. Cover with parafilm and mix by inversion. Ensure that powder is fully dissolved before filtration step (about 5 min).

9. Filter via vacuum suction through a 250 mL media filter.

3.2 Initial E. coli Overnight Growth

1. We recommend performing this portion of the experiment by 5:00 pm the day before harvest to give adequate time for initial overnight growth and subsequent *credentialed E. coli* cultures.

2. Remove *E. coli* stock from −80 °C, making sure not to let the stock thaw. Place it directly on ice in a Styrofoam box.

3. Label one 250 mL Erlenmeyer flask as "control" and one as "*E. coli.*"

4. Add 25 mL of LB media to each flask by pouring.

5. Swirl a Pasteur pipette in the "control" flask.

6. Inoculate the "*E. coli*" flask with *E. coli* stock by pressing a sterile Pasteur pipette into the frozen stock. Verify that some of the stock slush is attached to the pipette tip. Swirl the pipette tip in the "*E. coli*" flask, using a twisting of the finger at the top of the Pasteur pipette to evacuate the pipette of any stock liquid.

7. Place both flasks into the rotary shaker overnight at 300 rpm and 37 °C.

8. From this point forward, all manipulations are quantitative. Great care should be taken to ensure each of the *credentialed E. coli* cultures are treated identically.

3.3 Glucose Solution and E. coli Culture Preparation

1. Verify that the control overnight culture has not been contaminated.

2. The morning following overnight inoculation, label two 25 mL Erlenmeyer flasks as "^{12}C glucose" and "^{13}C glucose."

3. Add 2 mL of ultrapure water to each of the 25 mL flasks.

4. To the "^{12}C glucose" flask, add 200 mg of natural abundance D-glucose.

5. To the "^{13}C glucose" flask, add 207 mg of uniformly labeled ^{13}C-D-glucose.

6. Cover both flasks with parafilm.

7. Swirl gently and let it sit at room temperature until solutes are fully dissolved in water. This step may take up to 10 min. Do not let the glucose solutions come into contact with anything. Do not let any solution spill out of the flasks.

8. Label one sterile 1 L Erlenmeyer flask as "^{12}C" and another as "^{13}C."

9. Via auto-pipetter and a disposable 25 mL pipette, add 100 mL of M9 media to each flask.

10. Via syringe, remove all of the natural abundance glucose solution, filling the remainder of the syringe with air.

11. Attach a syringe filter to the syringe.

12. Slowly dispense the solution into the "^{12}C" flask while holding the filter in place. Verify final dispensing is air to ensure complete transfer of glucose.

13. Add 1 mL of ultrapure water to the empty 25 mL flask and swirl.

14. Remove the syringe filter and place it in its package. As above, remove the 1 mL of water and dispense through the same filter to ensure complete transfer.

15. Repeat **steps 10–14** for the uniformly labeled ^{13}C-D-glucose solution, using a fresh syringe and filter.

3.4 Credentialed E. coli Growth

1. Take the OD600 of the overnight culture by adding approximately 0.5 mL to a clear, plastic cuvette and taking an absorbance reading with a UV-VIS spectrophotometer. Be sure to blank the absorbance reading with leftover M9 media.

2. If the absorbance is greater than 1.0, dilute a small volume of the overnight culture $10\times$ in water until the OD600 is below 1.0. Calculate $OD600_{Effective}$ by multiplying the final absorbance reading by the dilution factor.

3. Calculate the desired volume of overnight culture to be added to each 1 L flask by the following equation: $1.5 \text{ mL}/(\text{OD600}_{\text{Effective}}/1000) = \mu\text{L}$ of overnight culture to add.

    The volume should be less than 1 mL but greater than 50 μL. If less than 100 μL is calculated, continue as written, but realize that growth times may be extended.

4. Add the calculated volume of overnight culture to each of the 1 L flasks via pipette. *This step is critical to ensuring your cultures grow at the same rate. Any error here is multiplied exponentially. Be sure to swirl the overnight culture before pipetting. Pipette and dispense twice into the overnight culture before pipetting into the 1 L culture flasks. Pipette directly into the culture, not down the side of the flask. Be mindful of extraneous drops.*

5. Place both flasks into the rotary shaker at 300 rpm and 37 °C.

6. Monitor growth with OD600 measurements until OD600 = 0.7, which is ready for harvest. Cultures should remain within 0.05 OD600 of each other at harvest time. If necessary, harvest at different times such that OD600 values are equal.

    a. Blank the UV-VIS spectrophotometer with leftover M9 media.

    b. We suggest taking readings every 2 h until OD600 = 0.5, every 1 h until OD600 = 0.6, and every 15 min until OD600 = 0.7.

3.5 Credentialed E. coli Harvest

1. Minimize the time between removal of the cultures from incubation and freezing of the pellets in liquid nitrogen. Keep this time constant between batches.

2. Place 50 mL conical tubes on ice in preparation for harvest.

3. Remove the natural abundance and ^{13}C-labeled *E. coli* culture flasks from rotary shaker and place on ice.

4. Swirl and pipette 50 mL of the cultures into each tube using separate 50 mL disposable pipettes, noting the actual volume in the final tube.

5. Centrifuge at 0 °C and 3200 rcf for 10 min.

6. Decant the supernatant, taking care not to disturb the pellets.

7. Rinse the top of the pellet with 0.5 mL of ultrapure water by pipetting gently down the side of the tube at an incline. Decant the supernatant.

8. Place conical tubes upright in liquid nitrogen, completely freezing the pellets.

9. Transfer conical tubes to the lyophilizer, taking care to maintain liquid nitrogen temperatures. Remove lids and cover the tops of each tube with Kimwipes, secured with rubber bands.

10. Dry on the lyophilizer for 24 h, or until completely dry.

11. Weigh powder into separate 1.5 mL Eppendorf tubes and transfer to a −80 °C freezer.

3.6 Credentialed E. coli Extraction and LC/MS

1. Perform the metabolite extraction protocol of choice (for *credentialed E. coli* growth samples only). The amount of dry mass extracted may vary depending on the protocol, signal-to-noise of the mass spectrometer being used, number of replicates, etc.

2. Mix the *E. coli* extracts to create two *credentialed* samples. The first sample should have natural-abundance *E. coli* and uniformly labelled *E. coli* in a 1:1 ratio by volume. The second sample should have natural-abundance *E. coli* and uniformly labeled *E. coli* in a 1:2 ratio by volume.

3. To concentrate the *credentialed* samples, dry and reconstitute the *credentialed* extracts in solution by using your method of choice.

    If *E. coli* standard extracts were purchased in solution from Cambridge Isotopes, mix the extracts in ratios as detailed in **step 2**. If *E. coli* standard extracts were purchased from Cambridge Isotopes as a powder, reconstitute the extracts as indicated in **step 3** and mix the reconstituted extracts in the ratios as indicated in **step 2**.

4. Transfer reconstituted extracts to LC/MS vials.

5. Perform LC/MS using your choice of chromatography, mass spectrometer, data acquisition settings, etc.

3.7 LC/MS Data Analysis

Credentialing is compatible with most data-processing software, as illustrated in Fig. 1. In this protocol, we detail a method using XCMS for feature detection, since it is commonly used in untargeted metabolomics [14]. We also include details on how to use other data-processing pipelines with *credentialing*.

1. Data processing will likely take less than 2 h from this point forward (depending on computer-processing speed).

2. Data analysis will be performed with RStudio. *See* Subheading 4.1 to install R and R packages on your computer.

3. A detailed template is available online at the link below to assist you in data processing with XCMS and *credentialing*: https://github.com/pattilabwu/Credential3.1/blob/master/SampleScript.R.

3.7.1 Loading R Packages and Setting Up a Working Directory

1. In the RStudio software, run the following to load the required R packages:

```
library(xcms)

library(Credential3.1)

library(data.table)
```

```
library(utils)
```

1. Use one of the following to set up a working directory, depending on your operating system:

```
setwd("/Users/Lingjue/Credentialing Demo") # MacOS format

setwd("C:/Users/Mike/Desktop/Credentialing Demo") # Windows format
```

1. Note that every time a new R session is opened, all required packages should be loaded again.

2. If you are using a data-processing pipeline that does not rely upon XCMS for peak detection, you may skip Subheading 3.7.2.

3.7.2 Data Processing with XCMS

1. Open MSConvert.

2. On the file loading panel, click "Browse" to select all data files.

MS Raw Data Conversion

1. Select output directory.

2. In the "Filters" panel, select "Peak Picking" and click "Add" to add the filter.

3. Remove all default filters, such as "titleMaker."

4. Click "Start" to begin converting raw data into mzXML format.

5. Figure 2 shows the Graphical User Interface of MSConvert and highlights the relevant entries related to each step above.

Feature Detection

1. Sort mzXML files for the *credentialed* samples into two folders based on natural abundance and uniformly labeled mixing ratios (e.g., folders are labeled as "**1T1**" for 1:1 mixing ratio and "**1T2**" for 1:2 mixing ratio). The folders should be in the working directory you set up in Subheading 3.7.1. XCMS processing includes peak detection, feature grouping, retention time correction, and missing peak filling. Run the following to process the mzXML files in the 1:1 mixing ratio folder ("**1T1**"):

```
# 1. peak detection with centWave algorithm

xs_1 = xcmsSet("./1T1", method="centWave", ppm=20, peak-width = c(10,30),
snthresh=5, prefilter=c(3,100))

# 2. initial peak grouping, 1st round

xs_1= group(xs_1, bw=5, mzwid=.015, minfrac=0.5)
```

```
# 3. retention time correction

xs_1 = retcor(xs_1, method="obiwarp",profStep=1)

# 4. feature grouping, 2nd round

xs_1 = group(xs_1, bw=5, mzwid=.015, minfrac=0.5)

# 5. filling missing peaks

xs_1 = fillPeaks(xs_1)
```

1. Repeat **step 1** for the 1:2 mixing ratio files (in "**1T2**" folder). That is, copy and paste the above code and change *xs_1* to *xs_2* (or any other name) and *"./1T1"* to *"./1T2"*.

Please note that users need to adjust XCMS parameters based on their instrumentation and methods. For further information, see Subheading 4.2 below. For detailed documentation about XCMS parameters, see: https://bioconductor.org/packages/release/bioc/manuals/xcms/man/xcms.pdf.

3.7.3 Input Format for *Credentialing*—The format of input data for *credentialing* is a *data.table* object with only four columns: "**cc**"—feature index number, "**rt**"—retention time, "**mz**"—m/z ratio, and "**i**"—signal intensity. The steps below describe how to create such an object.

Automated Processing of XCMS Files by *Credentialing*: *xsGroupExtract()* is a function in the *Credential3.1 R* package. It extracts features from the *xcmsSet* object and adjusts the format for *credentialing*. Run the following to extract features in *xs_1* and *xs_2* from Subheading 3.7.2 "Feature Detection":

```
featureGroup1 = xsGroupExtract(xs_1,intchoice = "into", sampling = 1)

featureGroup2 = xsGroupExtract(xs_2,intchoice = "into", sampling = 1)

features1T1 = data.table(featureGroup1$credTable)

features1T2 = data.table(featureGroup2$credTable)
```

*Use the help(xsGroupExtract) **command to see the description of parameters and output of xsGroupExtract().**

Manual Processing of Files from Other Data Processing Pipelines: If data processing was performed by another software platform:

*Important: when using other data processing pipelines, do not apply any function to group or remove isotope patterns.

1. Export the resulting features from each mixing ratio condition ("**1T1**" and "**1T2**") to two CSV tables. The tables should include only four columns: **"cc", "mz", "rt," and "i"**. An example is shown in Fig. 3.

2. Place the CSV files in the working directory and import them with the following:

```
features1T1 = data.table(read.csv("features1T1.csv"))

features1T2 = data.table(read.csv("features1T2.csv"))
```

3.7.4 Credentialing—Run the following to perform *credentialing*:

```
credential = credentialing (features1T1,features1T2,ppm = 20, rtwin =
2,rtcom = 5, ratio1 = 1/1, ratio2 = 1/2, ratio_tol =0.1, ratios_tol = 0.2)
```

*Parameters of the credentialing () function:

- **ppm**—mass error tolerance for isotope pair searching and grouping.

- **rtwin**—retention time window in the first round of credentialed feature selection.

- **rtcom**—retention time window in the second round of credentialed feature selection.

- **ratio1**—mixing ratio of unlabeled to labeled extract for the first credentialed sample (Default: 1/1).

- **ratio2**—mixing ratio of unlabeled to labeled extract for the second credentialed sample (Default: 1/2).

- **ratio_tol**—a factor between [0, 1] that sets the acceptable range for the intensity ratio relative to the actual mixing ratio (Default:0.1).

- **ratios_tol**—a factor between [0, 1] that sets the acceptable range for the intensity ratio relative to the combined mixing ratio (ratio1/ratio2, Default: 0.2).

*Generally, users should set these parameters based on their LC/MS instrumentation and methods. For more information regarding parameters, see: https://github.com/pattilabwu/Credential3.1/blob/master/README.md.

3.7.5 Data Output and Interpretation—The output of the *credentialing* function in R is a list object that includes *credentialed* and *non-credentialed* features. By typing in *credential$* and hitting the "tab" button, a list of tables will appear. Press the enter key to display the highlighted table. Some of the tables are listed in Table 1.

For more information regarding the output of *credentialing*, see: https://github.com/pattilabwu/Credential3.1/blob/master/README.md.

4 Notes

4.1 R Initiation, Package Installation, and Computer Configuration

1.   To install R, *see*: https://cran.r-project.org/.

2.   RStudio is the recommended working environment. To install RStudio, see: https://rstudio.com.

3.   For a one-time installation of R packages, use the following:

```
source("https://bioconductor.org/biocLite.R")

biocLite("xcms") # XCMS

install.packages("data.table") # data.table

install.packages("devtools") # devtools

install.packages("utils") # utils

devtools::install_github("pattilabwu/Credential3.1") # Credentialing
```

1.   There are several versions of the *credentialing* software available. This protocol describes the Credential3.1 package, which is a development of the original *credentialing* software designed to simplify analysis for less experienced users.

4.2 XCMS Parameter Settings

Depending on user instrumentation and methods, optimized parameters for XCMS may vary [15]. Table 2 provides some suggestions for XCMS settings as previously published [16]. More information can be found at the link below:https://bioconductor.org/packages/release/bioc/manuals/xcms/man/xcms.pdf

Acknowledgments

References

1. Roberts LD, Souza AL, Gerszten RE, Clish CB (2012) Targeted metabolomics. Curr Protoc Mol Biol 30:Unit 30 32 31–Unit 30 32 24. 10.1002/0471142727.mb3002s98

2. Nikolskiy I, Mahieu NG, Chen YJ et al. (2013) An untargeted metabolomic workflow to improve structural characterization of metabolites. Anal Chem 85(16):7713–7719. 10.1021/ac400751j [PubMed: 23829391]

3. Milne SB, Mathews TP, Myers DS et al. (2013) Sum of the parts: mass spectrometry-based metabolomics. Biochemistry 52(22):3829–3840. 10.1021/bi400060e [PubMed: 23442130]

4. Benton HP, Ivanisevic J, Mahieu NG et al. (2015) Autonomous metabolomics for rapid metabolite identification in global profiling. Anal Chem 87(2):884–891. 10.1021/ac5025649 [PubMed: 25496351]

5. Mahieu NG, Huang X, Chen YJ, Patti GJ (2014) Credentialing features: a platform to benchmark and optimize untargeted metabolomic methods. Anal Chem 86(19):9583–9589. 10.1021/ac503092d [PubMed: 25160088]

6. Mahieu NG, Patti GJ (2017) Systems-level annotation of a metabolomics data set reduces 25000 features to fewer than 1000 unique metabolites. Anal Chem 89(19):10397–10406. 10.1021/acs.analchem.7b02380 [PubMed: 28914531]

7. Lindahl A, Saaf S, Lehtio J, Nordstrom A (2017) Tuning Metabolome coverage in reversed phase LC-MS metabolomics of MeOH extracted samples using the reconstitution solvent composition. Anal Chem 89(14):7356–7364. 10.1021/acs.analchem.7b00475 [PubMed: 28613827]

8. Vinayavekhin N, Saghatelian A (2010) Untargeted metabolomics. Curr Protoc Mol Biol Chapter 30:Unit 30 1.1–Unit 30 124. 10.1002/0471142727.mb3001s90

9. De Vos RC, Moco S, Lommen A et al. (2007) Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. Nat Protoc 2(4):778–791. 10.1038/nprot.2007.95 [PubMed: 17446877]

10. Weber RJM, Lawson TN, Salek RM et al. (2017) Computational tools and workflows in metabolomics: an international survey highlights the opportunity for harmonisation through galaxy. Metabolomics 13(2):12 10.1007/s11306-016-1147-x [PubMed: 28090198]

11. Patti GJ (2011) Separation strategies for untargeted metabolomics. J Sep Sci 34(24):3460–3469. 10.1002/jssc.201100532 [PubMed: 21972197]

12. Naser FJ, Mahieu NG, Wang L et al. (2018) Two complementary reversed-phase separations for comprehensive coverage of the semi-polar and nonpolar metabolome. Anal Bioanal Chem 410(4):1287–1297. 10.1007/s00216-017-0768-x [PubMed: 29256075]

13. Ivanisevic J, Zhu ZJ, Plate L et al. (2013) Toward 'omic scale metabolite profiling: a dual separation-mass spectrometry approach for coverage of lipid and central carbon metabolism. Anal Chem 85(14):6876–6884. 10.1021/ac401140h [PubMed: 23781873]

14. Mahieu NG, Genenbacher JL, Patti GJ (2016) A roadmap for the XCMS family of software solutions in metabolomics. Curr Opin Chem Biol 30:87–93. 10.1016/j.cbpa.2015.11.009 [PubMed: 26673825]

15. Libiseller G, Dvorzak M, Kleb U et al. (2015) IPO: a tool for automated optimization of XCMS parameters. BMC Bioinformatics 16:118 10.1186/s12859-015-0562-8 [PubMed: 25888443]

16. Patti GJ, Tautenhahn R, Siuzdak G (2012) Meta-analysis of untargeted metabolomic data from multiple profiling experiments. Nat Protoc 7(3):508–516. 10.1038/nprot.2011.454 [PubMed: 22343432]
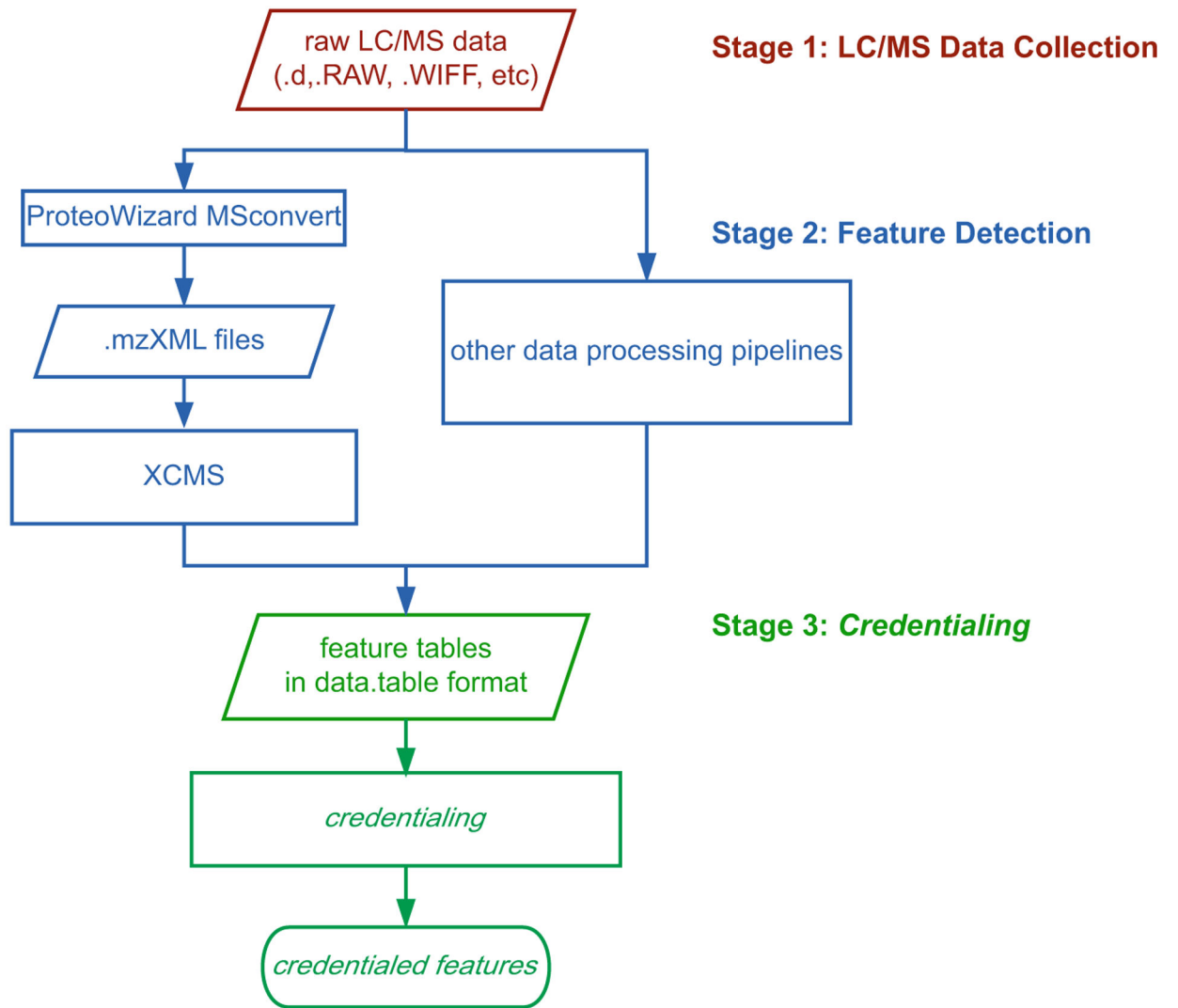
Fig. 1.
Three stages of the *credentialing* workflow
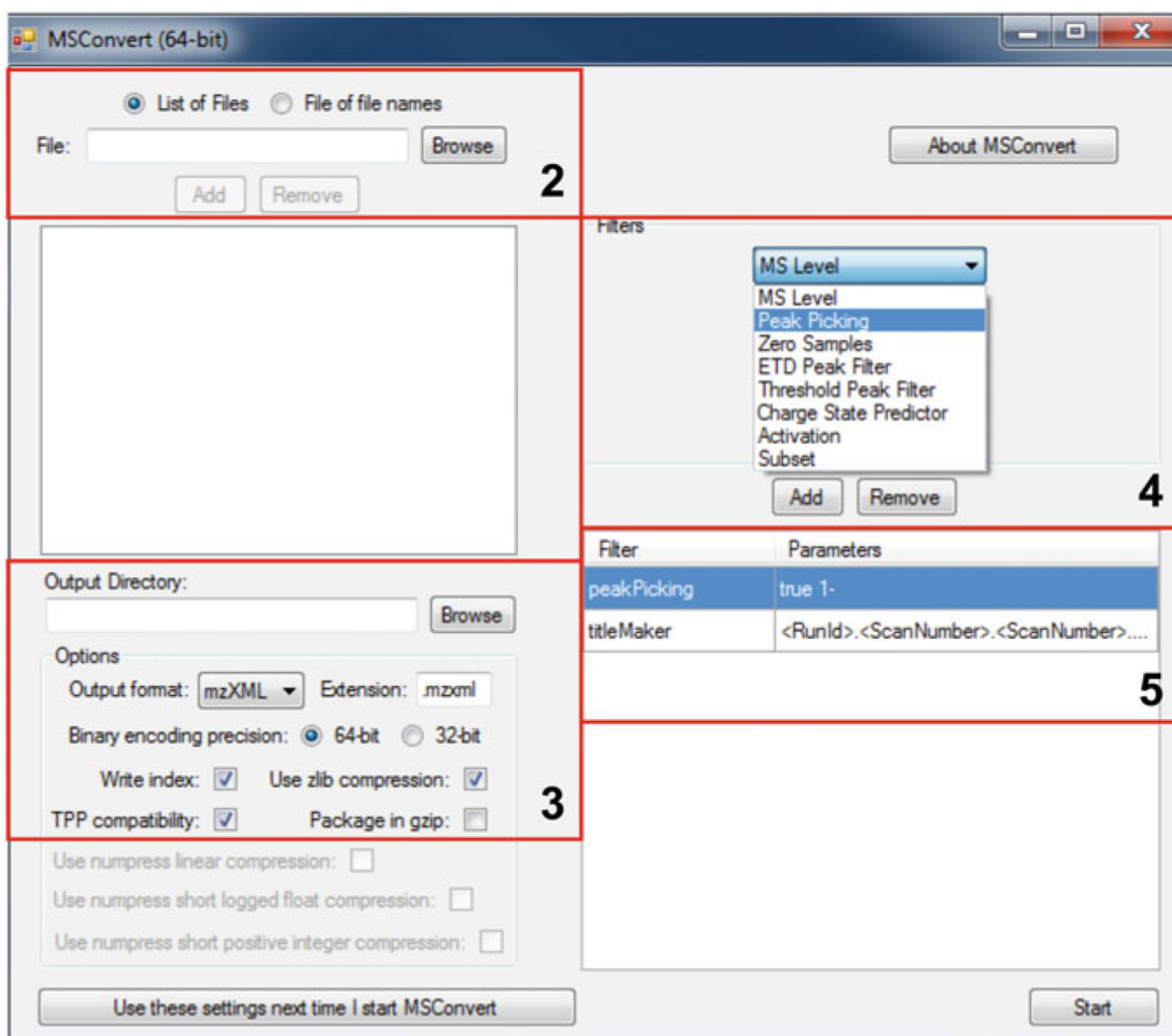
Fig. 2.
Screenshot illustrating the conversion of raw MS data with MSConvert

	A	B	C	D
1	cc	mz	rt	i
2	4548	528.238213	849.141	13465.3301
3	4557	529.242157	849.139	3760.92787
4	4642	536.257912	848.891	2981.36
5	4655	537.268064	849.141	5050.95995
6	4671	538.272184	848.891	15082.402
7	4679	539.27359	849.137	2833.09017
8	4510	525.183383	78.16	118789.039
9	4522	526.186445	78.159	31650.8108
10	4751	545.249504	78.16	26642.374
11	4761	546.253073	78.16	79862.3325
12	4783	547.256382	77.91	109770.141
13	4798	548.251597	78.16	5604.67703

Fig. 3.
Screenshot of a CSV file with appropriate format for *credentialing*

Table 1

Select tables generated by the *credentialing() function*

Name	Content
CredentialedFeatureR2F	Final *credentialed* features with second ratio filter
CredentialedFeatureR2	Final *credentialed* features without second ratio filter
CredentialedFeature1N2	First *credentialed* features from "1T1" excluded by second *credentialing*
CredentialedFeature2N2	First *credentialed* features in "1T2" excluded by second *credentialing*
CredentialedFeature1R1	First *credentialed* features from "1T2" (1/1 ratio)
CredentialedFeature2R1	First *credentialed* features from "1T2" (1/2 ratio)

Table 2

XCMS parameters depend on LC/MS instrumentation [16]

Instrument	ppm	Peakwidth	bw	mzwid	Prefilter
HPLC/Orbitrap	2.5	c(10,60)	5	0.015	c(3,5000)
UPLC/Orbitrap	2.5	c(5,20)	2	0.015	c(3,5000)
HPLC/Q-TOF (high resolution)	15	c(10,60)	5	0.015	c(0,0)
UPLC/Q-TOF (high resolution)	15	c(5,20)	2	0.015	c(0,0)