MTXQCvX2 documentation

Christin Zasada

2018-10-26

Contents

Chapter 1

Welcome

This documentation introduced to you how to use MTXQCvX2 in order to assess the quality of your GC-MS derived data, perform the determination of calibration curves and absolute quantification. It furthermore provides you two normalisation strategies and the calculation of quantities in, e.g., pmol/1e+6 cells or pmol/mg tissue.

MTXQCvX2 does also enable the calculation of stable isotopic incorporation and the evaluation of the underlying data, the mass isotopomer distributions (MIDs).

The tool has been set up to support the in-lab developed workflow for quantitative metabolomics experiments using the in-house developed software Maui for the annotation of data. MTXQCvX2 bridges the gap between quality control and first data post-processing / analysis of GC-MS derived data (MTXQCvX2_part1, MTXQCvX2_part2).

Nevertheless MTXQCvX2 includes a module in order to integrate all kind of data provided in spreadsheet-format, e.g., derived from metmax, extracting required information and creating corresponding files (MTXQCvX2_part4).

Both workflows are introduced in the distinct chapters including their required input parameter (chapter **??**). Technical relevant information are summarised in chapter **??**.

Chapter 2

Introduction

Experimental and mathematical concepts have been introduced for the pulsed stable isotope resolved metabolomics (pSIRM) approach in (Pietzke et al., 2014).

Chapter 3

Workflow for Maui-annotation proejcts

3.1 Read this in case

- you have run a succesfull Maui project
- exported all required container (see **??**)
- you have a copy of sequence list and experimental conditions
- you know the extraction procedure

The following article describes briefly how to use MTXQCvX2 in case you used Maui for the annotation of your metabolomics project. It does not matter if you have performed an experiment including stable isotopes or if you just aim for the quantification of a few intermediates.

3.2 Quick view

1. Setup a new R-project and copy MTXQC-files
2. Knit with parameter: `MTXQC_init.Rmd`
3. Copy input files and rename ManualQuantTable.tsv (e18205cz.tsv)
4. Create annotation.csv and sample_extracts.csv files
5. Define the internal extraction standard[1]
6. Knit with parameter: `MTXQC_ExperimentalSetup.Rmd`
7. Knit with parameter: `MTXQC_part1.Rmd`
8. Knit with parameter: `MTXQC_part2.Rmd`
9. If required, proceed with `MTXQC_part3.Rmd` - ManualValidation

3.3 Input files

Input files for the MTXQCvX are generated by using the export functions implemented in Maui. Specific containers have to be exported and moved into their corresponding `input` folder. Follow the instructions which Maui container have to be exported described at 12.4.1.5.

[1] see below InternalStandard

Certain circumstances might require the combination of *multiple MAUI-projects* into one MTXQC-project. This might be the case when you run the same samples in split and splitless mode on the machine or your experimental setup has been measured in multiple batches in order to avoid derivatisation effects.

It is recommended to combine the input files derived from a different number of Maui projects beforehand. In that way you have to work with a single file `CalculationFileData.csv` containing all experimental data points

The herein described process provides a quick way how to combine the exported files from different Maui projects. The script `combine-sets.R` saves all combined files into the correct `input` folder. Just update the folder and subfolder names. All the rest has been taken care of for you.

1. Create in the MTXQC-project folder (e.g., `psirm_glucose/`) a new folder, e.g., `raw-data`
2. Create subfolder for each MAUI-run in `psirm_glucose/raw_data`
3. Copy the required input files into each subfolder
4. Update the parameter of template file `combine-sets.R` file[2]
5. Execute the script and check the generated files
6. Merged files have been generated and copied into the corresponding folder: `psirm_glucose/input-folder/gc/...` or `psirm_glucose/input-folder/inc/...`
7. Copy the renamed tsv-files separately into input/quant/…

3.4 Annotation-file

The annotation file relate file names with experimental conditions or specify quantification standards in your batch. Two columns - **File and Type** - are obligatory and have to be present in the annotation file. In the case of absence MTXQCvX_part1 stops processing and shows an error message.

A quick way to generate an annotation file is described below:

1. Copy the first row / header of `quantMassAreaMatrix.csv` file
2. Paste & transpose the content into a new Excel-File into column A
3. Change Metabolite to File
4. Remove the entry QuantMasses at the very end of the column A
5. Add the column Type and specify each file either as sample or addQ1_dilution[3]
6. Add more columns specifying your experimental conditions, e.g., Cellline and Treatment …[4]
7. Save the content as `csv-file` in the `psirm_glucose/input/...`

3.5 Sample_extracts-file

The `sample_extracts.csv` file helps to determine correct absolute quantities in the manner of pmol/1e+6 cells or pmol/mg tissue in the `CalculationFileData.csv`.

This file requires two obligatory columns and have to be included: **Extract_vol** and **Unit**[5]. Please specify for each experimental condition the amount of extracted cells (count), tissue (mg) or volume of blood/plasme (ul) in the unit shown in the brackets.
The names of the columns of the experimental conditions have to match up with the annotation file. Save the file in the folder `psirm_glucose/input/...`.

If the defined experimental conditions do not match up with the annotation MTXQCvX2_part1.Rmd exit data processing.

[2] inst/template_files/…
[3] see for further details additionalQuant
[4] optimal: two-three parameter, max: four parameter. Consider possible combinations, e.g., HCT116-control, HCT116-BPTES
[5] Define: count, mg or ul

A template file can be modified and reused: `inst/template_files/...`

3.6 Internal Standard

MTXQCvX2 allows the specification of project-specific internal extraction standards. The only thing you need to do is to define the corresponding compounds as an internal standard in the `conversion_metabolite.csv` file. To do so add `InternalStandard` in last column `Standard`.

For an classical pSIRM experiment in the Kempa lab we are using cinnamic acid. The evaluation of this compound has been integrated into maui and peak areas are exported from a distinct container called `cinAcid`. The exported file has to be renamed to `InternalStandard.csv` and moved to `psirm_glucose/input/gc/...`.

If you have used a different compound as an internal extraction standard you might need to extract the peak areas of this compound from the file `quantPeakAreasMatrix.csv` file and save it in the folder `psirm_glucose/input/gc/InternalStandard.csv`, respectively. Prerequisite - you annotated the compound in Maui.

The report of `MTXQCvX2_part1.Rmd` includes the detected internal standard for each project.

Chapter 4

Workflow for Metmax-extracted projects

4.1 You want to follow this ...

- in case you have measured samples and quantification standards by GC-MS
- performed the annotation of intermediates in ChromaToF / vendor software
- exported all files into a txt-format containing information about metabolite names, peak area

4.2 Introduction

This document describes how to use MTXQCvX2 in combination with metmax[1].

Historically, MTXQCvX2 has been developed and optimized for Maui-derived input files. The `MTXQCvX2-part4.Rmd` functions as a converter of metmax-derived files in order to create suitable input formats for MTXQCvX-part1. This module could also be used to convert tables derived from other programs as long as they are confirm with the herein described tables. Mandatory columns are referenced in the text.

The general workflow of the NMTXQCvX2 project is briefly shown below in **quick view**. More detailed instructions are summarised in the following paragraphs. For more detailed explanations about the individual input parameter for each module of MTXQCvX2 please proceed to read the documentation about the individual modules and their knitting parameter. The relation of knitting parameter, input and output files are described in each section.

4.3 Quick view

1. Generate input files: run `MTXQC_part4.Rmd`[2]
2. Setup R-project and copy MTXQC-files
3. Knit with parameter: `MTXQC_init.Rmd`
4. Copy input files into corresponding folders
5. Create annotation.csv and sample_extracts.csv files[3]

[1] http://gmd.mpimp-golm.mpg.de/apps/metmax/default.htm
[2] read here the instructions
[3] Details further down this document

6. Update metabolite names in `conversion_metabolite.csv`[4]
7. Define the internal standard and/or alkanes[5]
8. Knit with parameter: `MTXQC_ExperimentalSetup.Rmd`
9. Knit with parameter: `MTXQC_part1.Rmd`
10. Knit with parameter: `MTXQC_part2.Rmd`
11. If required - proceed with `MTXQC_part3.Rmd` - ManualValidation

4.4 Required input files derived from a ChromaToF/Metmax

If you need an introduction about how to use metmax - have a look at the separate documentation `Metmax_intro`.

The chapter Workflows for Metmax-extracted data explains in detail how to use this module to generate suitable input files. Depending on your experiment you can select the kind of data that is extracted and transformed by the `MTXQCvX_part4`.

4.5 Annotation-file

The annotation file relate file names with experimental conditions or specify quantification standards in your batch. Two columns - **File and Type** - are obligatory and have to be present in the annotation file. In the case of absence MTXQCvX_part1 stops processing and shows an error message.

A quick way to generate an annotation file is described below:

1. Copy all file names from a file of your choice
2. Paste & transpose the content into a new Excel-File into column A
3. Give it the column name: File
4. Optional: Remove the entry QuantMasses at the very end of the column A
5. Add the column Type and specify each file either as sample or addQ1_dilution[6]
6. Add more columns specifying your experimental conditions, e.g., Cellline and Treatment …[7]
7. Save the content as `csv-file` in the `psirm_glucose/input/...`

4.6 Sample_extracts-file

The sample_extracts.csv file helps to determine correct absolute quantities in the manner of pmol/1e+6 cells or pmol/mg tissue in the `CalculationFileData.csv`.

This file requires two obligatory columns and have to be included: **Extract_vol** and **Unit**[8]. Please specify for each experimental condition the amount of extracted cells (count), tissue (mg) or volume of blood/plasme (ul) in the unit shown in the brackets.
The names of the columns of the experimental conditions have to match up with the annotation file. Save the file in the folder `psirm_glucose/input/...`.

If the defined experimental conditions do not match up with the annotation MTXQCvX2_part1.Rmd exit data processing.

A template file can be modified and reused: `inst/template_files/...`

[4] Column: Metabolite_manual
[5] Also in conversion_metabolite.csv; see below paragraph Standards
[6] see for further details additionalQuant
[7] optimal: two-three parameter, max: four parameter. Consider possible combinations, e.g., HCT116-control, HCT116-BPTES
[8] Define: count, mg or ul

4.7 Conversion_metabolite.csv

This file, saved in `config_mtx/`serves as a translational table. It defines alternative version of metabolite library names for the use in plotting your data. It is also used to define settings and standard classifications.

Detailed information for each column of the file are shown here: REF

4.7.1 Match your annotation with library names

Prior the analysis you need to match the names of your intermediates with the conversion_metabolite.csv file. You need to add the corresponding name for each intermediate in the column **Metabolite_manual**.

General suggestion for naming conventions in ChromaToF: Metabolite_Derivate, optional in case of main- (MP) and byproducts (BP) Metabolite_Derivate_MP/BP. An example for both: Lactic acid_(2TMS) or Glucose_(1MEOX)(5TMS)_MP.

If you have annotated intermediates that are not included so far in this table please follow the instructions how to extend `conversion_metabolite.csv`.

4.7.2 Define your internal standards and alkanes

MTXQCvX2 allows the specification of project-specific internal standards. Corresponding compounds have to be marked as an internal standard in `conversion_metabolite.csv` by adding the tag InternalStandard in the column Standard.

If you check the box - InternalStandard the module searches in your input file containing peak areas for the defined standard and extracts the information. It also generated the file `InternalStandard.csv` and stores it at `psirm_glucose/input/gc/`.

In the same way alkanes are be defined in conversion_metabolite.csv. Each alkane has to be flag tagged with **Alk** in the column Standard. This gives you the opportunity to implement customized mixtures of alkanes in order to determine the retention index. MTXQCvX_part4 recognises the flag tag and generates based on peakarea-file `Alcane_intensities.csv` and saves it at `psirm_glucose/input/gc/`[9].

The in-lab approach considers nine alkanes from c10 to c36. Standard annotation includes an hashtag, e.g., #c10. If you use this annotation even Metmax would be able to determine the retention index.

[9] It should be al_k_ane, I know, but Maui doesn't…

Chapter 5

MTXQCvX2_init

MTXQCvX2_init.Rmd - why and how to use it. Advantages of the project folder.

Chapter 6

MTXQCvX_experimentalSetup.Rmd

Chapter 7

MTXQCvX_part1.Rmd

Chapter 8

MTXQCvX_part2.Rmd

Chapter 9

MTXQCvX_part3.Rmd

Chapter 10

MTXQCvX_part4.Rmd - Metmax parser

10.1 This section explains ...

- what MTXQCvX_part4.Rmd does
- how do input files need to look like
- which files are generated
- what the distinct checkboxes mean

This module provides the generation of suitable input files for MTXQCvX2 based on spreadsheet exported information by tools like metmax.

M

10.2 Input files

10.2.1 Quantification - PeakAreas.csv[1]

In order to perform absolute quantification of

You need a file containing all extracted peak areas for each metabolite and file[2]. The header of metmax-extracted files looks like shown below (see table 1). Please, remember to delete the second header row, representing the column loads for each file before saving as csv-file. Otherwise you end up with weird imported dataframes in R. Quantification masses have to be updated while processing in ChromaToF prior the export of the data e.g., with a reference search[3] or using statistical compare. pSIRM experiments require the definition of pTop5 masses[4] instead of top5 masses in the reference in order to take into account the shift of intensities induced by the application of stable isotopes[5]

name	mass	ri	row.load	file_1	file_2	file_x
Lac	219	1051	0.76	15423	135444	465486
Pyr	174	1042	0.65	56978	46888	4354544
Cit	273	1805	0.99	1326	23321	132121

[1]Required for: all parameter, just not `calculation stable isotope incorporation`
[2]Tools/Options/Retention analysis, Parameter: Area
[3]See `vignette/ReferenceSearch`
[4]Extended list of quant masses considering isotope incorporation
[5]Mandatory columns: name, mass, files

MTXQCvX_part4 takes care of the formatting and correct column names of the peak areas file and saves it[6]. MTXQCvX_part4 generates also the file PeakDensities-Chroma.csv[7], in case you have selected the option to include sum of area normalisation while knitting this module.

10.2.2 Isotope incorporation - MIDs.csv[8]

In order to determine the incorporation of stable isotopes MTXQCvX2 requires as an input the mass isotopomer distributions (MIDs) for each intermediate and measurement[9]. Fragments for each intermediate have to be pre-defined in metmax at Tools/Options/metabolite masses. They can be imported[10] or manually specified each by each. An example of the metmax output is shown in table 2. The output has to be saved as csv-file, including the deletion of the partial row `column.load`, respectively[11].

name	mass	ri	row.load	file_1	file_2	file_x
Lac	219	1051	0.85	31026	5165829	5829
Lac	220	1051	0.85	3607	662277	277
Lac	221	1051	0.85	1222	111481	81
Lac	222	1051	0.85	188	1003494	10023
Lac	223	1051	0.85	0	33542	342

MTXQCvX_part4 calculates the stable isotope incorporation and exports DataMatrix.csv as well as pSIRM_SpectraData.csv[12]. The mathematics behind are outlined in (Pietzke et al., 2014)

Important: Extracted MIDs have to match with defined mass couples for each metabolite in MTXQCvX2[13]. Please refer for more details to `vignettes/config_mtx-files`.

10.2.3 Derivatisation efficiency - mz73.csv[14]

The extraction of intensities for the ion m/z 73 works analogous to the extraction of MIDs[15]. Mass ranges have to be defined for each intermediate for the mass 73 by defining starting and end mass with 73. MTXQCvX_part4 generates the file MassSum-73.csv[16]. Check `inst\template_files\` for reference. Hopefully soon a new metmax button extracting specific intensities across the batch.

[6] `input/quant/quantMassAreasMatrix.csv`
[7] `input/gc/PeakDensities-Chroma.csv`
[8] Required for `calculation isotope incorporation`
[9] Tools/Options/Isotope concentrator; Parameter: IntensityOfMass
[10] `inst/template_files/MetMax_MIDs.txt`
[11] Mandatory columns: name, mass, files
[12] `input/inc/DataMatrix` & `pSIRM_SpectraData.csv`
[13] `config_mtx/incorpo_calc_masses.csv`
[14] Required for: `sum of area normalisation`
[15] Tools/Options/Isotope concentrator; Parameter: IntensityOfMass
[16] `input/gc/MassSum-73.csv`

Chapter 11

Configuration of MTXQCvX2

Herein explained are the customizable tables of the MTXQCvX2 universe.

11.1 conversion_metabolite.csv

```
data = read.csv("config_mtx/conversion_metabolite.csv", TRUE, sep = ";")
colnames(data)
```

```
##  [1] "Metabolite_manual" "Metabolite"        "Metabolite_short"
##  [4] "Lettercode"        "Q1_value"          "Mass_Pos"
##  [7] "SE_sel"            "Q_sel"             "nopsirm"
## [10] "Standards"
```

11.2 Metabolic profile

11.3 Defintion of intermediates

11.4 Declaration of standards

11.4.1 Internal standard

11.4.2 Alkanes

11.4.3 quant1-values.csv

11.4.4 incorporation_calc.csv

Chapter 12

Protocols

12.1 Sample extraction

12.1.1 Cell extracts

Materials:

- cell culture dishes (10 cm)
- 50 % MeOH plus 2 ug/ul cinnamic acid, ice-cold
- chloroform
- 15 ml falcon tubes
- cell lifter

Procedure:

- prepare cell culture dishes accordingly to your experimental conditions
- discard media / buffer
- add immediately 5 ml MeOH
- detach cells using cell lifter
- collect and transfer cell extract into 15 ml falcon
- add 1 ml chloroform
- incube for 30 - 60 min at cold temperature on rotary or thermo shaker
- centrifuge at max speed for 10 min
- collect polar and lipid phases
- dry under vacuum

In order to generate technical backups:

- resuspend dried extracts in 600 ul 20 % MeOH
- shake at cold temperature on thermo shaker for 30 min
- split volumes into equal parts in fresh eppendorf tubes
- dry under vacuum

Suggested cell density: $2 - 3e + 6$ cells / extract.

12.1.2 Tissue samples

12.1.3 Blood samples

12.2 Sample derivatisation

12.3 GC-MS measurement

12.4 Data processing

12.4.1 Processing for Maui annotation

12.4.1.1 Resampling

12.4.1.2 Combo-export

12.4.1.3 Maui quantification masses

12.4.1.4 MAUI pSIRM

12.4.1.5 Maui exports

1. Input-Folder: gc

- Alcane_intensities.csv (Diagnostics/Export Alcane intensities)
- InternalStandard.csv [1]
- MassSum-73.csv (Diagnostics/QC Mass Sum Export for mass 73)
- PeakDensities-Chroma.csv (Diagnostics/ExportPeakDensities)

2. Input-Folder: quant

- ManualQuantTable.tsv - **rename it** - e.g., e18125cz.tsv [2] [3]
- quantMassAreasMatrix.csv (Quantification export of samplesPeakGroups)

3. Input-Folder: inc

- DataMatrix.csv (Export % Label of pSIRM-samplesPeakGroups)
- pSIRM_SpectraData.csv (pSIRM Spectra Export)[4]

[1] see more details below
[2] Location: Maui-project/export/QM-AbsoluteQuantification/…
[3] C://Users/User-name/MauiProjects/…
[4] Requires the selection of Natural_MIDs.txt

12.4.2 Processing for Metmax data extraction

12.4.2.1 Resampling

12.4.2.2 1D-basic

12.4.2.3 Reference search

12.4.2.4 Export for Metmax

12.4.2.5 Data extraction

Chapter 13

Frequently Asked Questions

13.1 What are additional quantification standards

13.2 How do I extend conversion_metabolite.csv

Bibliography

Pietzke, M., Zasada, C., Mudrich, S., and Kempa, S. (2014). Decoding the dynamics of cellular metabolism and the action of 3-bromopyruvate and 2-deoxyglucose using pulsed stable isotope-resolved metabolomics. *Cancer & metabolism*, 2(1):9.