

User guide for ExNormalizeMets:

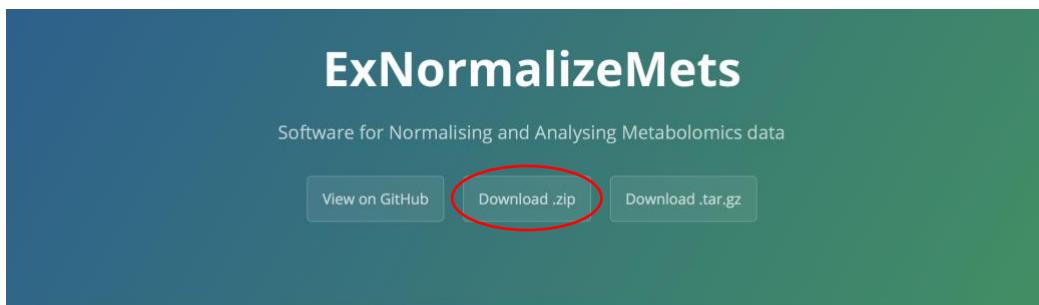
For this guide, a windows machine is being used. No prior software other than Microsoft Excel and a web browser are required for installation. The guide will explain how to install R to be able to use the NormalizeMets R package through its ExNormalizeMets Excel interface (no direct interaction with R or command line functions are needed).

A detailed guide with illustrations for using the software is also provided, showing its main functionalities in a ‘walk through’ guide of an example project.

For support contact: olshansky.g@unimelb.edu.au

Getting the files ready:

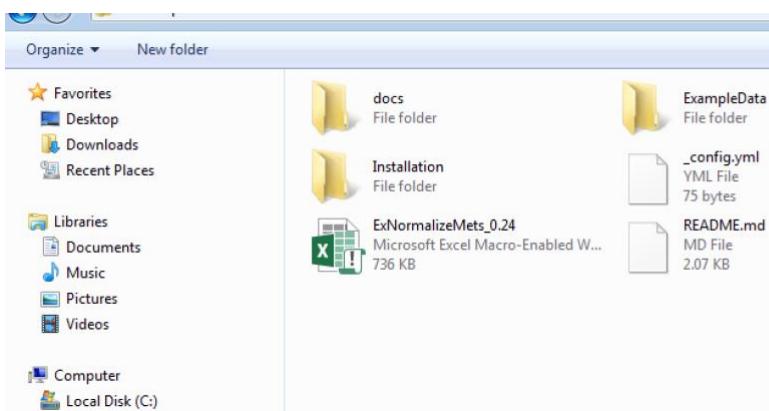
To get the required files, go to <https://metabolomicstats.github.io/ExNormalizeMets/> and download the latest version of the *ExNormalizeMet*:



ExNormalizeMets

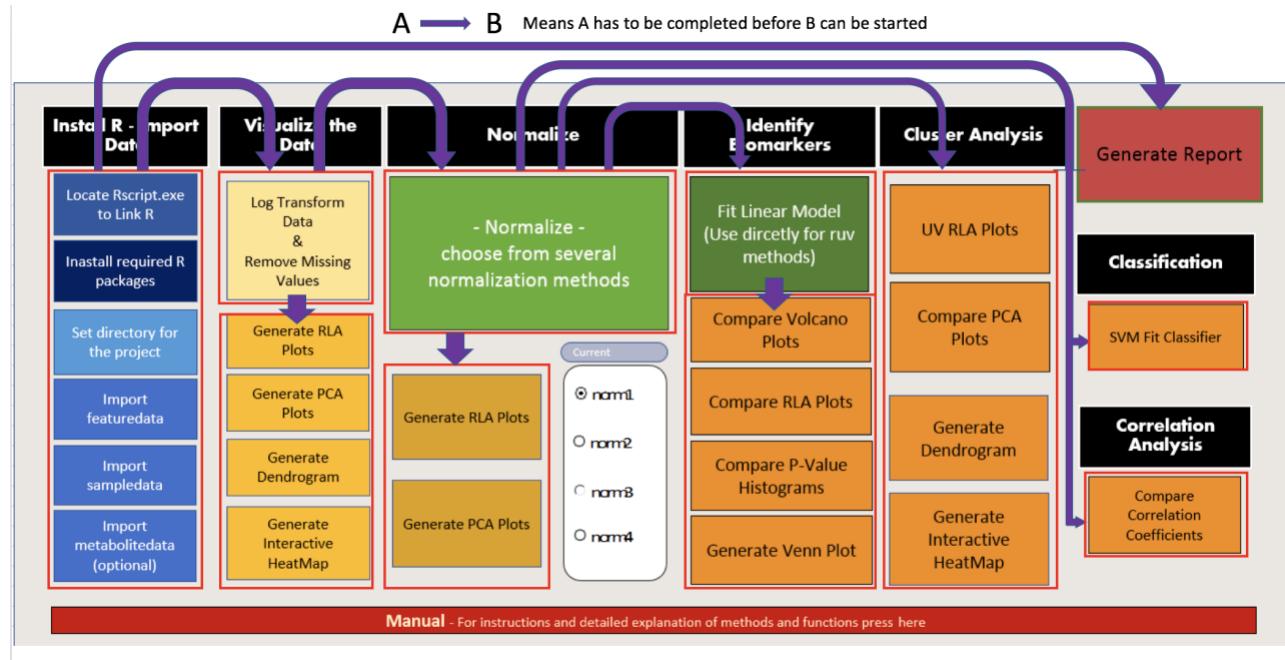
They are all the files needed for installing R and loading NormalizeMets through excel.

Once downloaded, open the folder:



Click on *ExNormalizeMets_0.24* to open the excel worksheet but otherwise don't modify the files in this folder (the downloaded folder can be copied to a different location on your computer for convenient access although it is only needed for first use).

General Workflow:



Installing R:

1. Install R from CRAN (The Comprehensive R Archive Network) by going to the following page:
<https://cran.r-project.org>

The Comprehensive R Archive Network

Download and Install R

Precompiled binary distributions of the base system and contributed packages, Windows and Mac users most likely want one of these versions of R:

- [Download R for Linux](#)
- [Download R for \(Mac\) OS X](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

Source Code for all Platforms

Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

- The latest release (2017-11-30, Kite-Eating Tree) [R-3.4.3.tar.gz](#), read [what's new](#) in the latest version.
- Sources of [R alpha and beta releases](#) (daily snapshots, created only in time periods before a planned release).
- Daily snapshots of current patched and development versions are [available here](#). Please read about [new features](#) and [bug fixes](#) before filing corresponding feature requests or bug reports.
- Source code of older versions of R is [available here](#).
- Contributed extension [packages](#)

Questions About R

- If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.

What are R and CRAN?

R is 'GNU S' a freely available language and environment for statistical computing and graphics which provides a wide variety of statistical and



R for Windows

Subdirectories:

[base](#)

Binaries for base distribution. This is what you want to [install R for the first time](#).

[contrib](#)

Binaries of contributed CRAN packages (for R >= 2.13.x; managed by Uwe Ligges). There is also information on [third party software](#) available for CRAN Windows services and corresponding environment and make variables.

[old contrib](#)

Binaries of contributed CRAN packages for outdated versions of R (for R < 2.13.x; managed by Uwe Ligges).

...

Tools to build R and R packages. This is what you want to build your own packages on Windows, or to build



R-3.4.3 for Windows (32/64 bit)

[Download R 3.4.3 for Windows](#) (62 megabytes, 32/64 bit)
[Installation and other instructions](#)
[New features in this version](#)

If you want to double-check that the package you have downloaded matches the package distributed by CRAN, you can compare the [md5sum](#) of the .exe to the [fingerprint](#) on the master server. You will need a version of md5sum for windows: both [graphical](#) and [command line versions](#) are available.

Frequently asked questions

- [Does R run under my version of Windows?](#)
- [How do I update packages in my previous version of R?](#)
- [Should I run 32-bit or 64-bit R?](#)

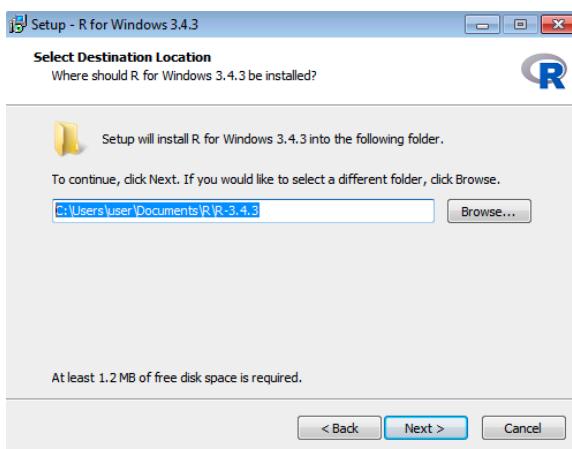
Please see the [R FAQ](#) for general information about R and the [Windows FAQ](#) for Windows-specific information.

Other builds

- Patches to this release are incorporated in the [r-patched snapshot build](#).
- A build of the development version (which will eventually become the next major release of R) is available in the [r-devel snapshot build](#).
- [Previous releases](#)

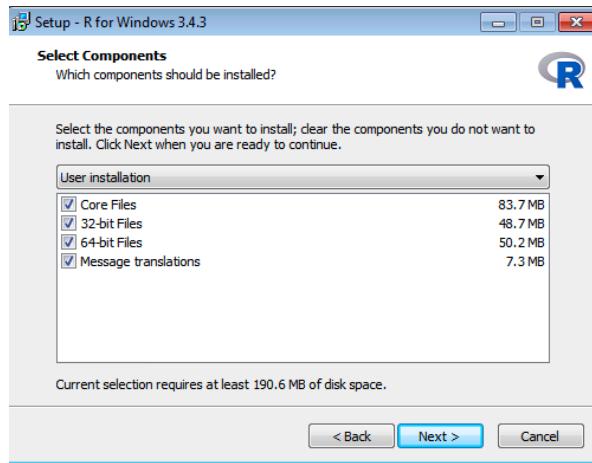
Note to webmasters: A stable link which will redirect to the current Windows binary release is
<CRAN MIRROR>/bin/windows/base/release.htm.

2. Once the installation is complete, run the downloaded file R-3.4.3-win.exe, follow the installation instructions and choose the location where R should be installed.

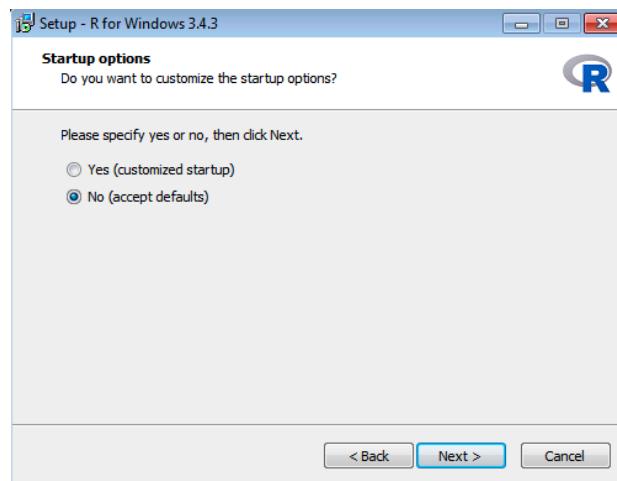


Click *Next* for default location (Recommended) or choose location manually. Make sure you **remember where R is installed** as you will need to locate this folder later to link R to Excel.

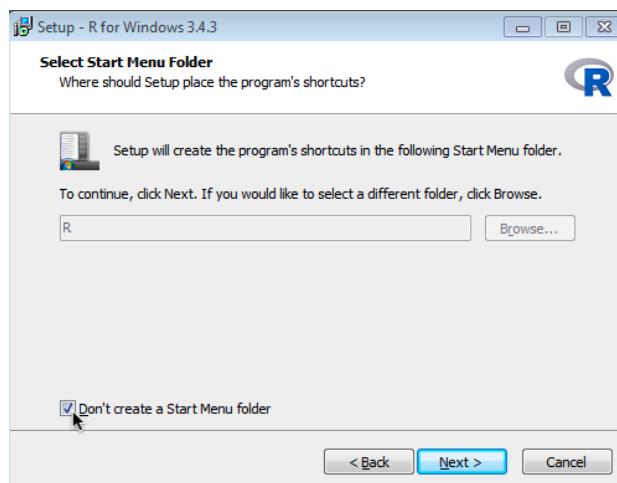
3. Click *Next* to install with the different required settings:

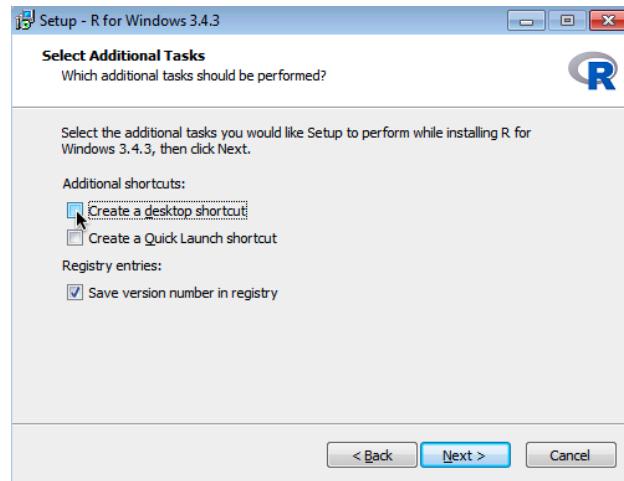


4. Click *Next* to install with default settings:

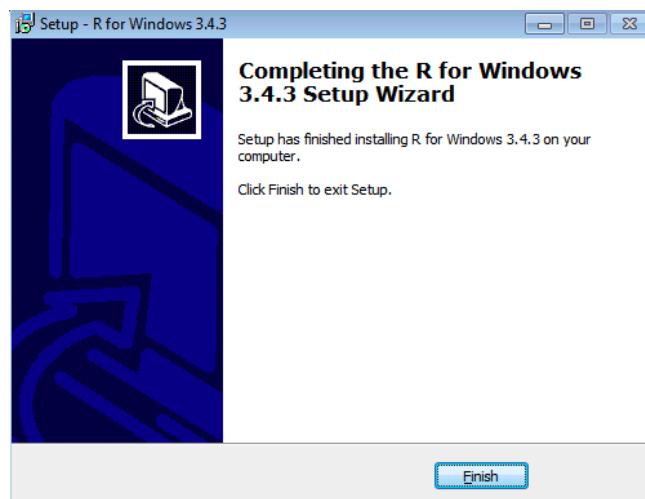


5. In the next screen check the *Don't create a Start Menu folder* icon if you don't intend to use R by itself and click next. Also uncheck the *Create a desktop shortcut* in the next screen:



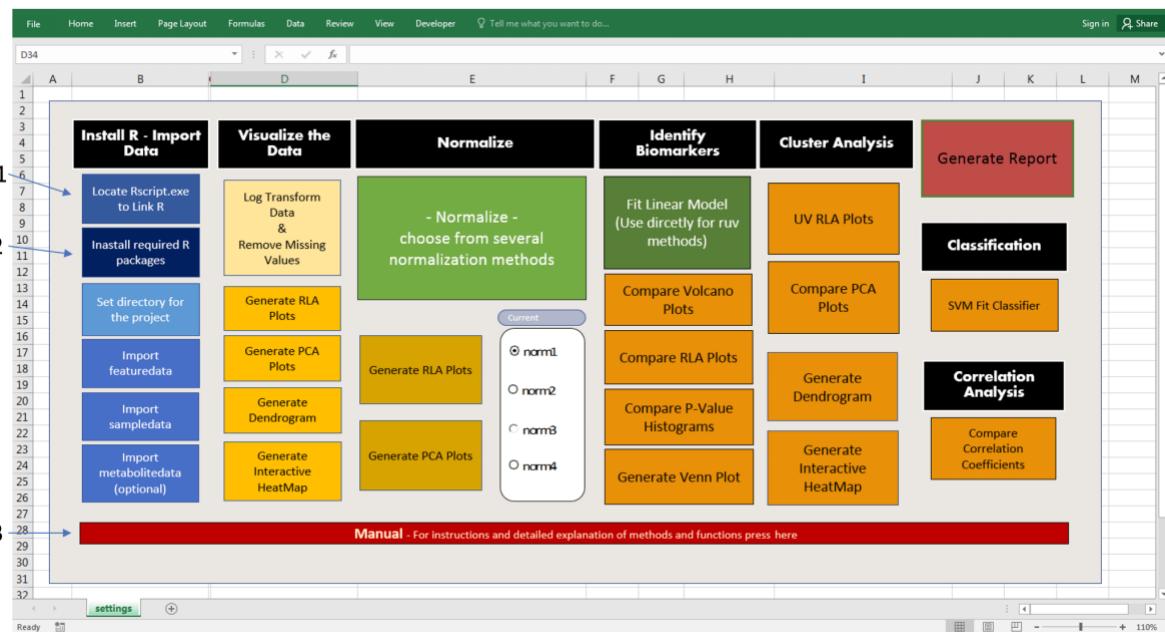


6. Wait until installation is done, the following screen should appear



Linking R to excel and installing the required packages (first use only):

After installing R open the excel file ExNormalizeMets.xlsx, this will open the Excel interface onto the settings sheet:

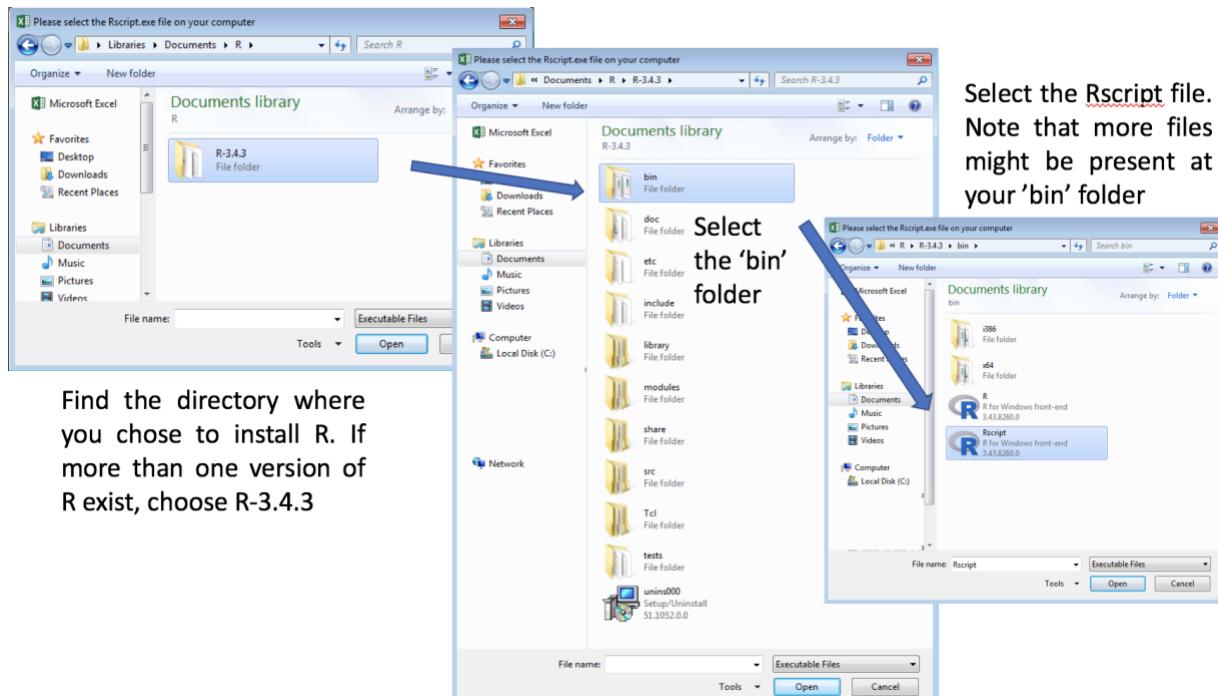


The *settings* sheet is your ‘**Control sheet**’, any function you want to run, from importing data, Normalizing, viewing results and opening the manual can be done from here.

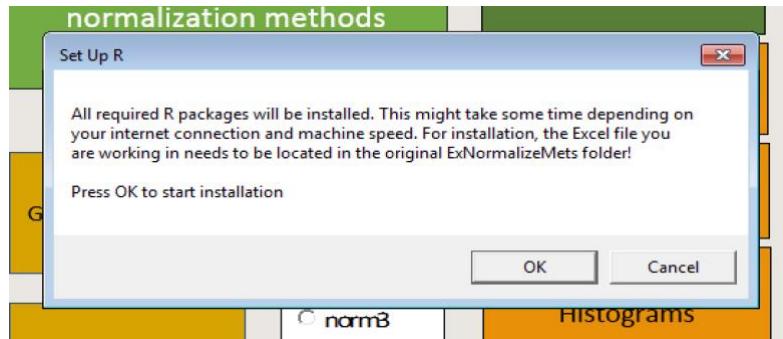
This manual can be accessed at any stage by clicking on (3) but more on this later. For now, first locate the *Rscripts.exe* file so that excel will know how to run R commands it generates.

Locating Rscript:

Press (1. ‘*Locate Rscript.exe to Link R*’) to locate the *Rscript* file in the window that opens up. Make sure to select the *Rscript* file in the *bin* folder of your R installation:



After selecting the file, to install all the required R packages and set up the needed dependencies, press (2. ‘*Install required R packages*’).



Pressing ok will start the installation, this might take a few minutes if you are using R for the first time as many of the base packages will need to be installed.

The installation window looks like this:

```

C:\Users\Oshinsky\Documents\R\win-3.4.3\bin\scripts\ex
downloaded 45 KB
trying URL 'http://cran.us.r-project.org/bin/windows/contrib/3.4/cran_0.0.20.zip'
Content type 'application/zip' length 246551 bytes <240 KB>
downloaded 240 KB
trying URL 'http://cran.us.r-project.org/bin/windows/contrib/3.4/knitr_1.18.zip'
Content type 'application/zip' length 913875 bytes <892 KB>
downloaded 892 KB
trying URL 'http://cran.us.r-project.org/bin/windows/contrib/3.4/rmarkdown_1.8.zip'
Content type 'application/zip' length 2291306 bytes <2.2 MB>
downloaded 2.2 MB
package 'GGally' successfully unpacked and MD5 sums checked
package 'plotly' successfully unpacked and MD5 sums checked
package 'ggplot2' successfully unpacked and MD5 sums checked
package 'htmlwidgets' successfully unpacked and MD5 sums checked

```

When the installation is done, a window with the message *Done!* will appear.

Now that the installation is complete, NormalizeMets is ready for use!

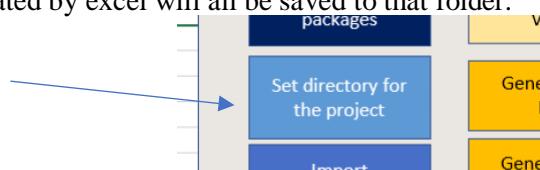
Using NormalizeMets

Data used in the following examples is provided with NormalizeMets (alldata_eg in R), it is located in the *ExampleData* folder in your downloaded *ExNormalizeMetsSetup* file. Future references in this guide refer to this data by default.

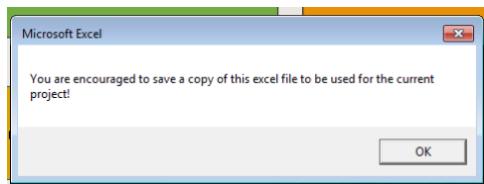
An example excel document containing the data used for the tutorial with all settings identical to those in the tutorial is provided (MyFirstNormalizeMetsProject_example.xlsx), if using this document, make sure to set the Rscript location and working directory for your machine.

Starting a new project:

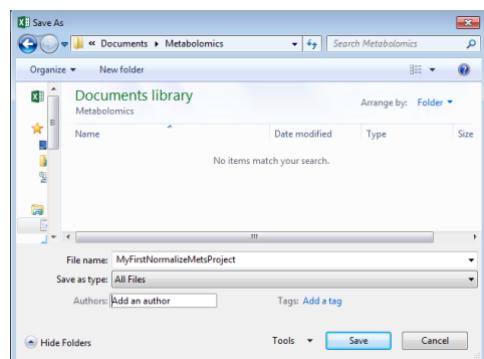
To start a new project, you will need to set up or use an existing directory where the project will ‘live’, data and plots generated by excel will all be saved to that folder.



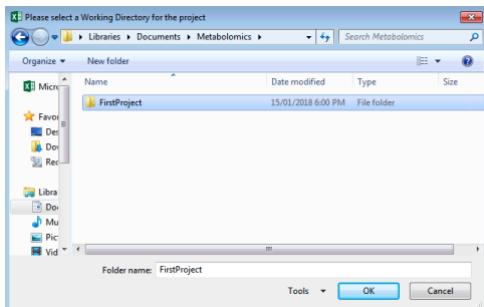
Before selecting the folder, you will be prompted to save a copy of the current version of the excel file, it is recommended to save it with a new name to make sure a ‘clean’ version always stays in your ExNormalizeMetsSetup folder.



After saving the workbook under the name of your choice:



Select the working directory where all files should be generated. We recommend making a new folder for each project.



Loading data:

To load data for the project, in turn click on the following to load the relevant data:

Import featuredata	Import the Data	Normalize	Identify Biomarkers	Cluster Analysis	Generate Report
Import featuredata	Locate Rscript.exe to Link R Install required R packages Set directory for the project	Log Transform Data & Remove Missing Values Generate RLA Plots Generate PCA Plot Generate Dendrogram Generate Interactive HeatMap	- Normalize - choose from several normalization methods radio buttons: ① norm1 ② norm2 ③ normR ④ normA	Fit Linear Model (Use directly for ruv methods) Compare Volcano Plots Compare RLA Plots Compare P-Value Histograms Generate Venn Plot	UV RLA Plots Compare PCA Plots Generate Dendrogram Generate Interactive HeatMap
Import sampledata				Classification SVM Fit Classifier	Correlation Analysis Compare Correlation Coefficients
Import metabolitedata (optional)					

Manual - For instructions and detailed explanation of methods and functions press here

Loaded data needs to be in .csv format. After clicking on the required file, it will open a new sheet, showing the imported data. Select the setting sheet import more data and get back to the options.

For *featuredata*, set metabolites in columns and samples in rows. Unique sample names should be provided as row names.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9	m_10	m_11	m_12	
2	10485.86719	33220.562	1112.979492	1408.639649	455.7529297	1100.402344	122.3631592	3855.804688	1637.482422	5194.292969	28793.65625	5200.71	
3	8960.46875	29995.5156	926.2890625	529.0449219	873.609375	2201.246094	173.0802002	5090.636719	2011.509977	8868.859375	31683.09375	6183.91	
4	10160.44531	28559.5406	1230.333008	1306.320313	1027.507813	2066.591797	269.7998047	4483.90625	1644.607422	7776.335938	44494.59375	10505.1	
5	8794.476563	27593.75	901.762707	1800.083008	675.039508	287.9992676	4949.261719	1508.802734	8405.171875	37030.40625	10726.1		
6	8956.921875	28161.70563	979.1723633	818.0366211	904.0253906	1245.991211	167.4133301	4302.033394	1460.012695	6976.429688	29579.53125	5421.91	
7	9092.257813	31685.3125	634.6899414	478.3103103	980.4231398	1259.563477	97.5625	4406.3475	2349.7891	7960.136719	39835.71875	5657.81	
8	2271.100000	7692.176000	42.166332	109.839900	78.045471875	138.791992	140.096600	100.40025	96.729400	2356.000000	20212.000000	4483.1	
9	7692.402344	28098.76000	822.091900	11.11500	50.580300	10.054500	127.770600	4000.000000	1000.274441	6300.000000	8013.000000	8013.000000	
10	10960.6625	21605.14063	408.0004883	1499.479492	452.2375488	914.5383203	234.4975806	465.7.535006	1847.687	6606.15625	34395.96875	7142.31	
11..10	1743.932017	7647.645431	210.4199219	127.4194434	408.2912598	999.5717773	74.82690533	1153.711719	895.0953516	2407.103516	18250.149063	2358.11	
12..11	1284.412109	4819.99609	61.62763451	15.51902008	394.8364258	1112.433555	1163.742188	620.701660	2640.0707013	10544.94531	1728.1		
13..12	2272.923828	6813.67968	243.9420166	280.7983394	434.3505859	1167.060938	86.25360107	1665.397852	1009.748047	3927.849609	21808.45313	6546.91	
14..13	9938.0625	28498.1875	577.8486318	1302.7236313	350.9599609	1200.636719	134.2883307	4006.34375	1777.508834	8467.085938	30871.17188	6412.01	
15..14	6795.752006	24264.4063	745.7470703	465.0424805	990.1313477	1704.889438	119.0528376	4312.5.390063	1480.686532	6865.367188	23930.875	8029.31	
16..15	1836.59375	6009.232881	142.5959473	37.65142822	491.6550293	70.08390381	1526.198242	941.4536133	3346.365234	20822.03125	2188.41		
17..16	1699.289063	6458.503906	96.57415771	200.641665	310.729248	774.0219727	183.0802246	1600.1981625	900.3241888	1143.683594	14266.88281	1877.81	
18..17	1497.520508	5609.605469	77.26855469	169.138916	226.6845703	567.8823242	79.01464844	1362.202148	830.0599609	2024.548828	9113.1325	1730.21	
19..18	9033.570313	21372.0625	1250.067383	1564.307617	892.3813477	1589.5886867	252.1553955	4251.226563	1510.931641	6424.839844	38623.5625	7712.21	

sampedata should have sample information matching featuredata (samples in rows).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	s_1	Batch 2 code_1	58.7	22.2													
2	s_1	Batch 2 code_0	67.6	23.7													
3	s_2	Batch 2 code_0	56.2	28.2													
4	s_3	Batch 2 code_0	56.2	28.2													
5	s_4	Batch 2 code_0	77.2	26.2													
6	s_5	Batch 2 code_0	74.3	26.4													
7	s_6	Batch 2 code_1	66.8	26.4													
8	s_7	Batch 1 code_0	65	29													
9	s_8	Batch 2 code_1	66.5	26.1													
10	s_9	Batch 2 code_0	70.2	27.3													
11	s_10	Batch 1 code_1	55	25													
12	s_11	Batch 1 code_0	51	28													
13	s_12	Batch 1 code_1	55	29													
14	s_13	Batch 2 code_0	79.2	30.7													
15	s_14	Batch 2 code_1	52.7	25.4													
16..15	s_15	Batch 1 code_0	42	32													

Optional *metabolitedata* should have metabolite information matching featuredata with metabolite names in rows.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
1	names	I5	neg_controls	pos_controls	gender											
2	m_1	0	1	0	0											
3	m_2	0	1	1	0											
4	m_3	0	1	0	0											
5	m_4	0	0	0	0											
6	m_5	0	0	0	0											
7	m_6	0	0	0	0											
8	m_7	0	0	0	0											
9	m_8	0	1	0	0											
10	m_9	0	1	0	0											
11	m_10	0	1	0	0											
12	m_11	0	0	0	0											
13	m_12	0	0	0	0											
14	m_13	1	1	0	0											
15	m_14	0	0	0	0											

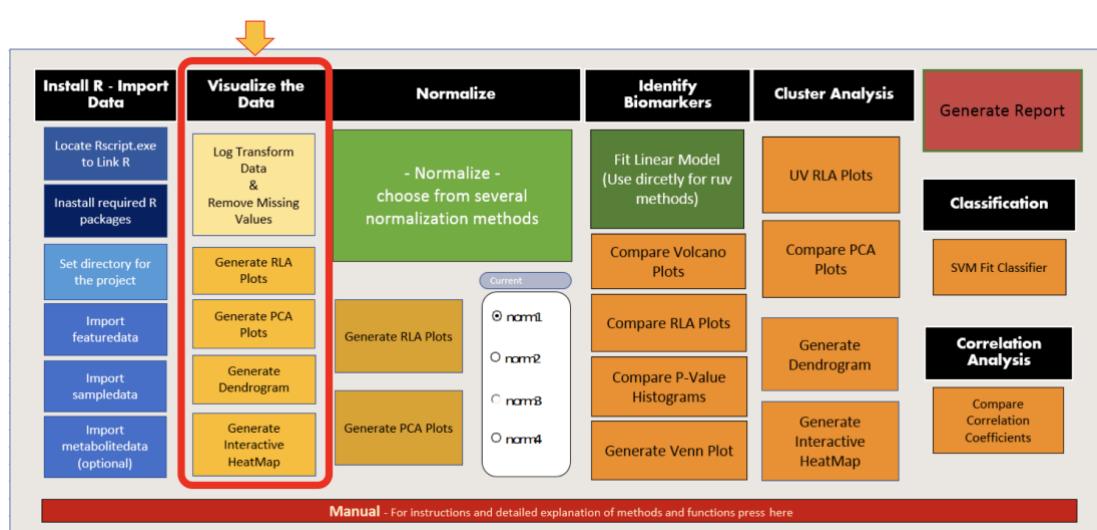
metabolitedata can include any metabolite information such as grouping structures, internal standard metabolites, negative control and positive control metabolites.

After the data is loaded, you are ready to proceed to analyse the data!

NormalizeMets Workflow:

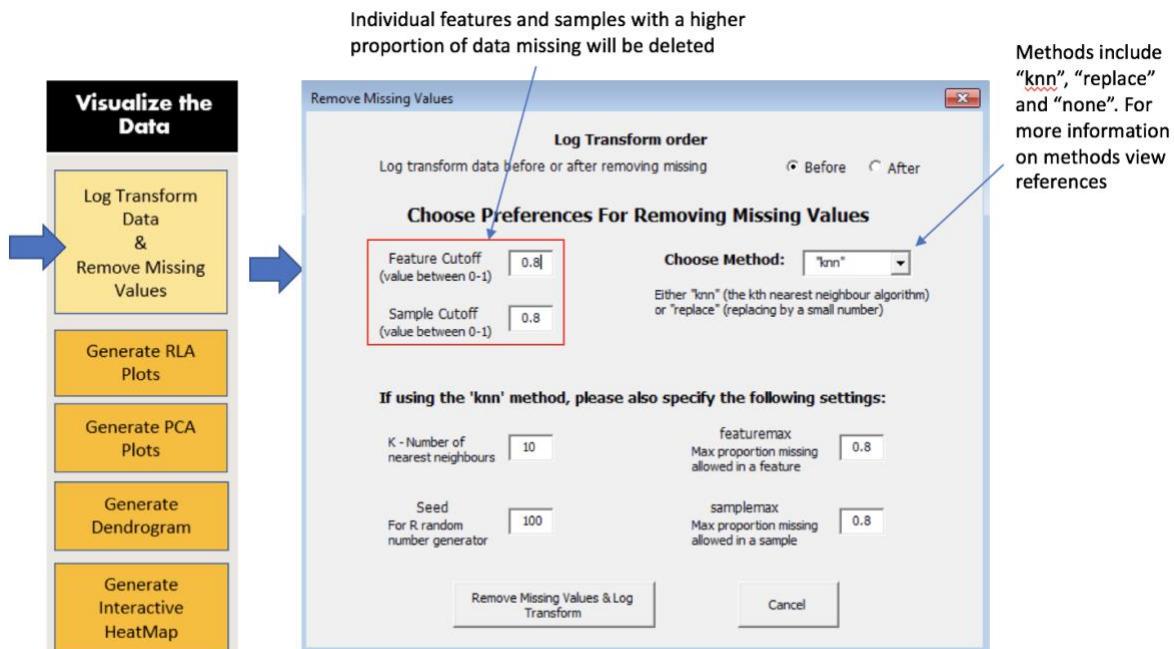
Visualize the Data

The following section refers to the visualize part:



Log Transforming the data and removing missing value (mandatory):

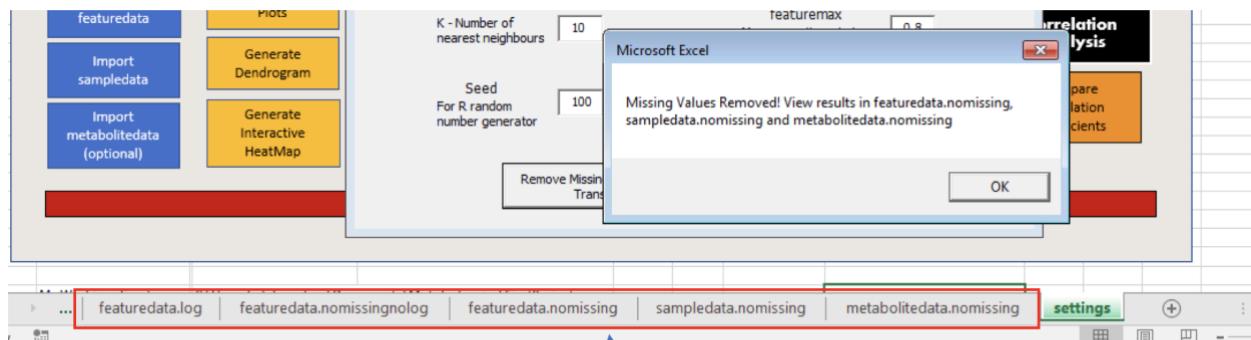
A frequent issue in metabolomics data sets is the occurrence of missing values. It is important to reduce the number of missing values as much as possible by using an effective pre-processing procedure. For example, a secondary peak picking method can be used for LC-MS data to fill in missing peaks which are not detected and aligned.



“knn” – use k nearest neighbours method to replace missing values.

“replace” – replaces missing values by half the minimum value in featuredata.

Clicking ‘Remove Missing Values & Log Transforming’ the following appears:



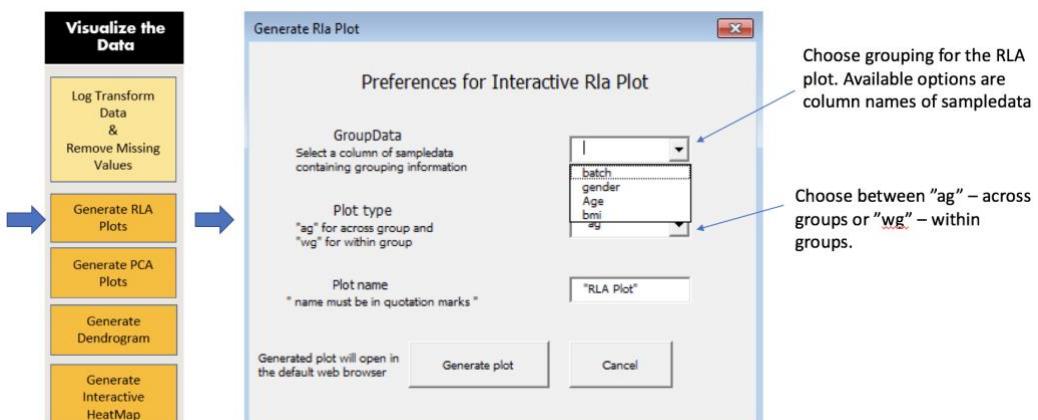
Note the new sheets that appeared, they have respectively the regular log transformed featuredata, the log transformed featuredata with missing values removed, featuredata with missing values removed and without log transformation, sampledata with rows removed corresponding to featuredata.nomissing, metabolitedata with rows removed corresponding to featuredata.nomissing.

Unless you are interested to view or copy any of this data, those sheets are only going to be used for further internal functions.

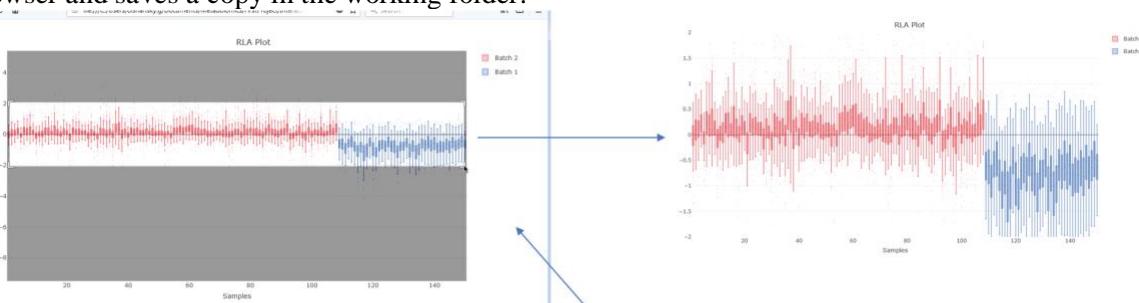
Now the plots in *Visualize the Data* can be generated! The data generated is also going to be used for the Normalization section.

RLA plots

One way of visualising the log transformed metabolomics data is the use of *across group* or *within group* relative log abundance (RLA) plots (De Livera et al. 2012 De Livera et al. (2015)).

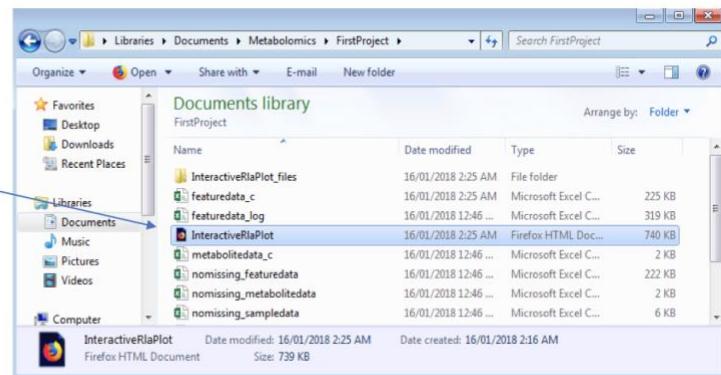


Setting groupdata to *batch* and selecting *Generate plot* opens the interactive plot in the default web browser and saves a copy in the working folder.



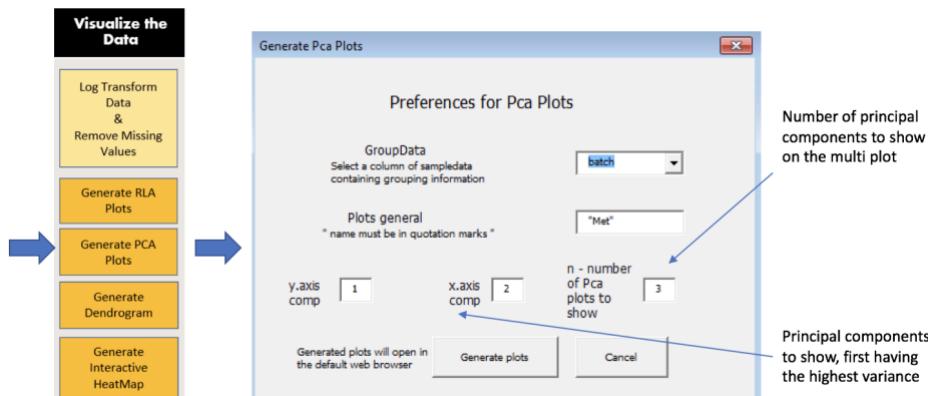
Interactive plot opens in the default browser,
to zoom in, simple select the required part

Copy of all plots together with all other generated files are stored in the working directory



PCA plots

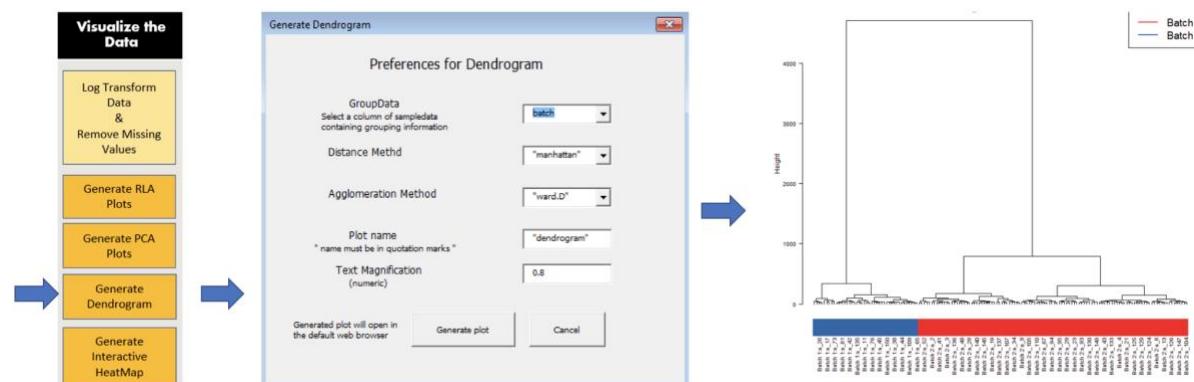
The following function can be used to obtain multiple plots for exploration of the principal components of the *featuredata* matrix: a bar plot indicating the variance explained by each principal component, scores and loading plots with specified axes (interactive and non-interactive), and a pairs plot of the first n principal components. These plots are useful in identifying any outlying samples and getting a preliminary understanding of the structure of the data.



All produced plots are stored in the working directory, with interactive plots opened in the browser and static plots located in the new *plots* sheet.

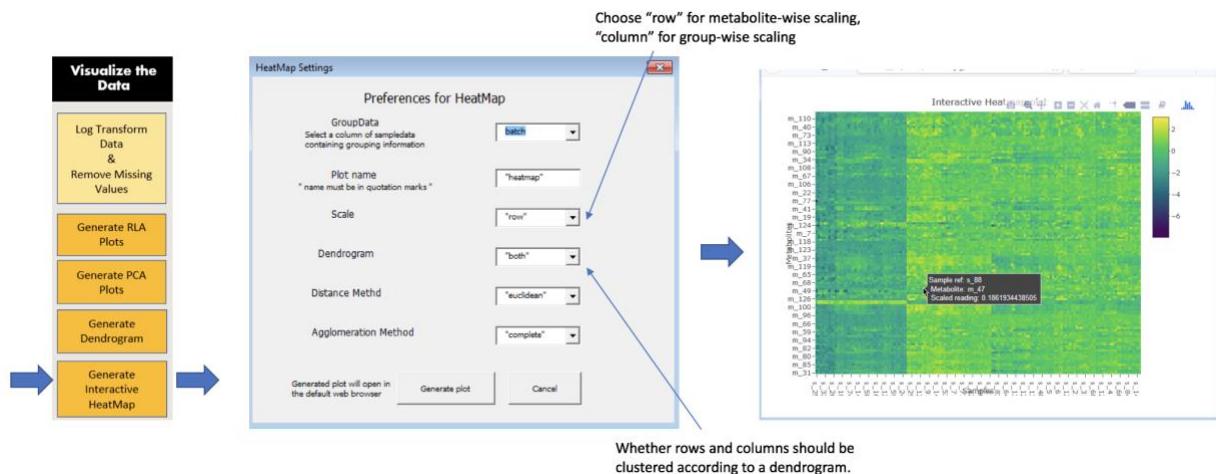
Dendrogram

Generates a dendrogram to visualise clustering structures in the data, many different methods are available.



HeatMap

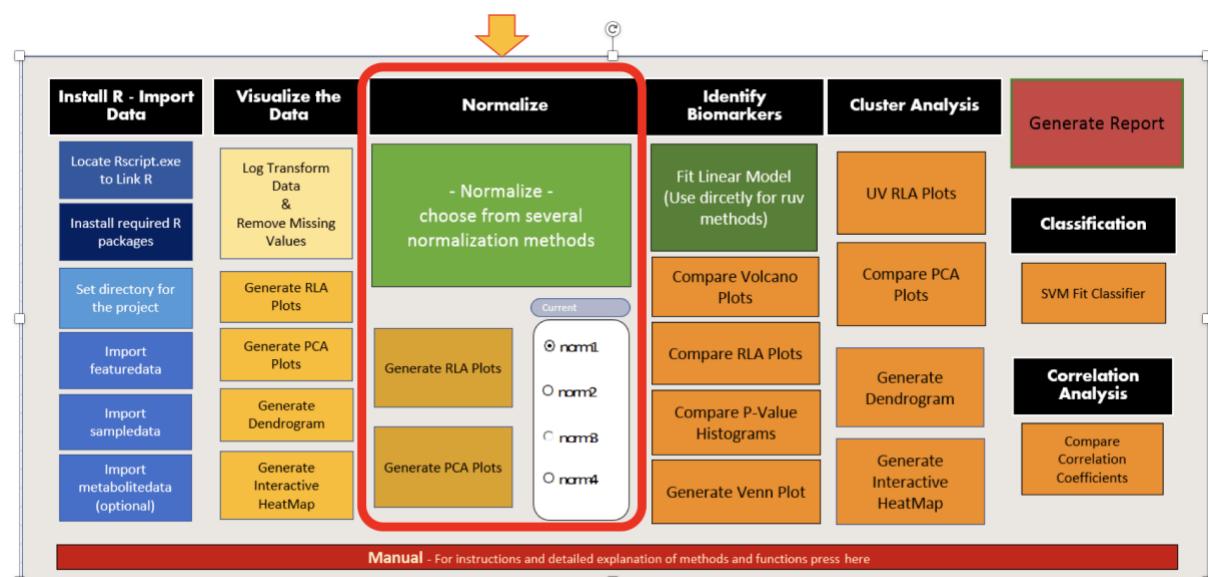
The HeatMap produced can reveal interesting structures in the data.



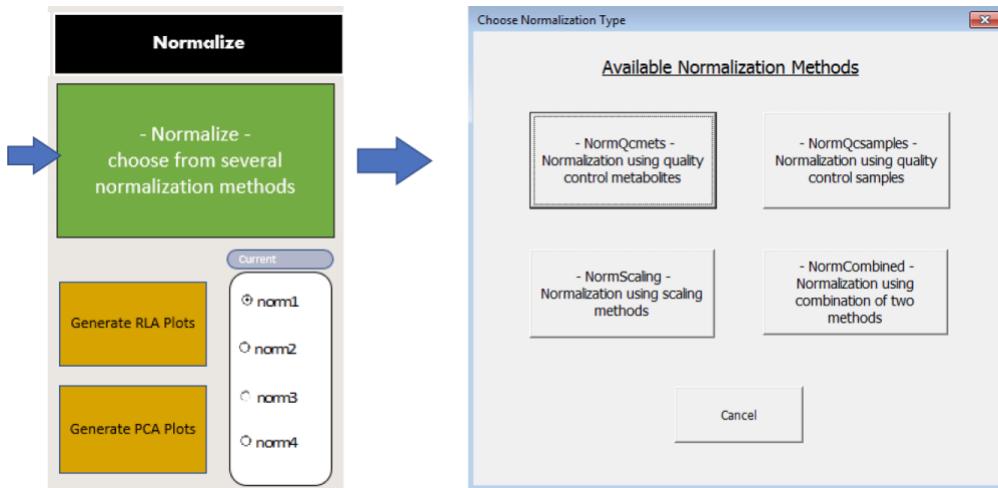
Normalization

Normalization methods presented in this package are divided into four categories, as those which use (i) internal, external standards and other quality control metabolites (*NormQcmets*) (Sysi-Aho et al. 2007, Redestig et al. (2009), De Livera et al. (2012), De Livera et al. (2015), Gullberg et al. (2004)) (ii) quality control samples (*NormQcsamples*) (Dunn et al. 2011), (iii) scaling methods (*NormScaling*) (Scholz et al. 2004, Wang et al. (2003)), and (iv) combined methods (*NormCombined*) (Kirwan and Broadhurst (2013)).

The normalization methods are accessible in the following section:



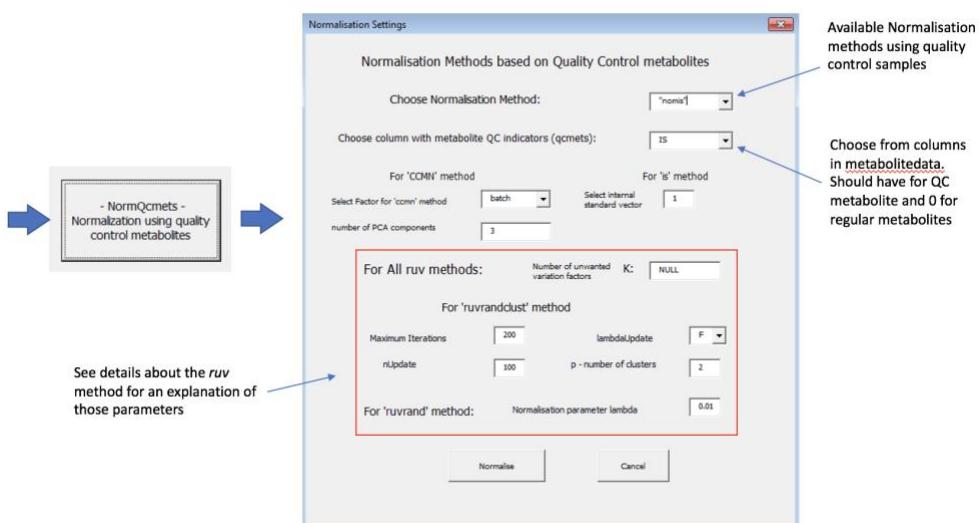
Clicking on the Normalize button opens the following menu enabling the choice of different normalization methods.



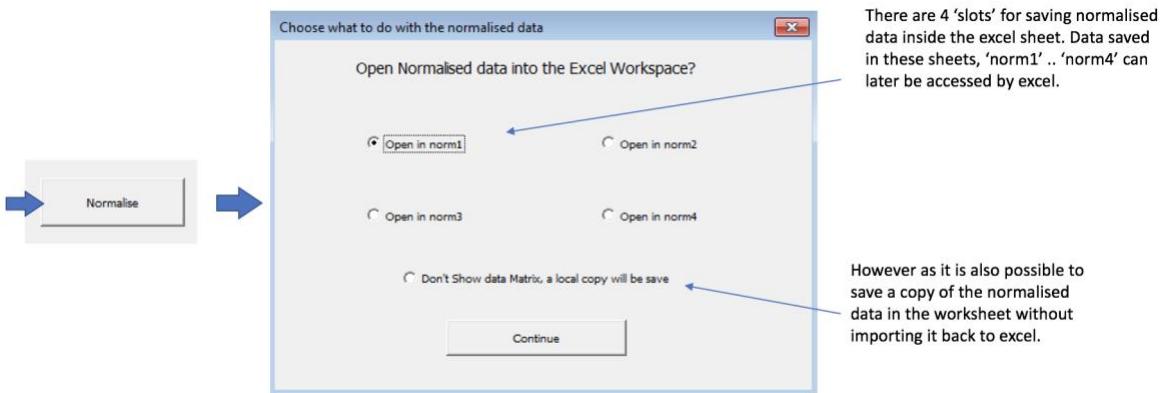
NormQcmets

The approaches in *NormQcmets* use internal, external standards and other quality control metabolites. These include the *is* method which uses a single standard (Gullberg et al. 2004), the *ccmn* (cross contribution compensating multiple internal standard) method (Redestig et al. 2009), the *nomis* (normalization using optimal selection of multiple internal standards) method (Sysi-Aho et al. 2007), and the remove unwanted variation methods (J. A. Gagnon-Bartsch, Jacob, and Speed 2014) as applied to metabolomics using “ruv2” (De Livera et al. 2012), “ruvrnd” and “ruvrndclust” (De Livera et al. 2015). Note that *ruv2* is an application specific method designed for identifying biomarkers using a linear model that adjusts for the unwanted variation component.

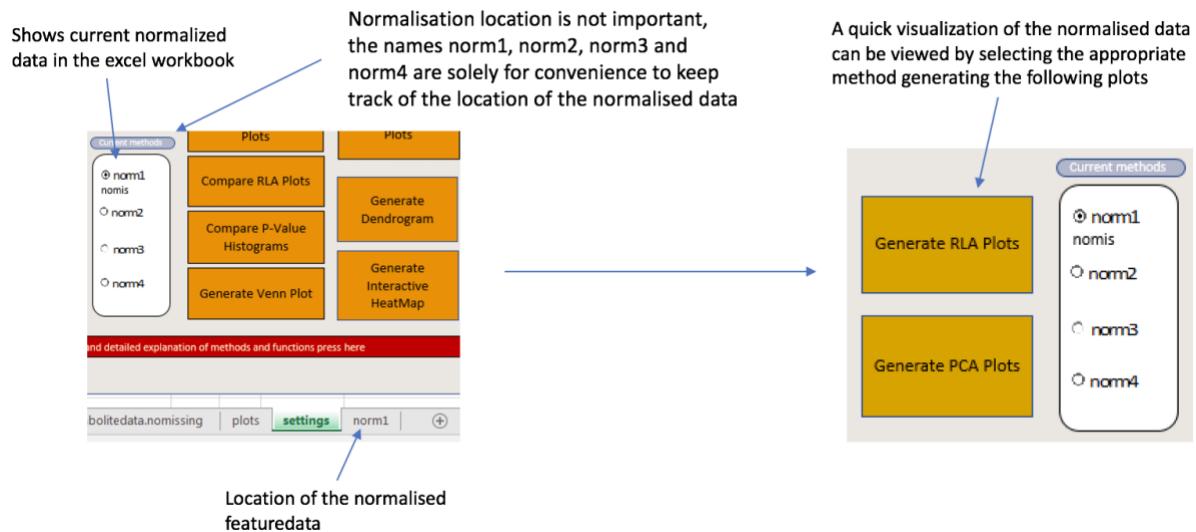
To Normalize:



After Clicking the Normalise button the screen asking you where the normalized data is to be saved appears.



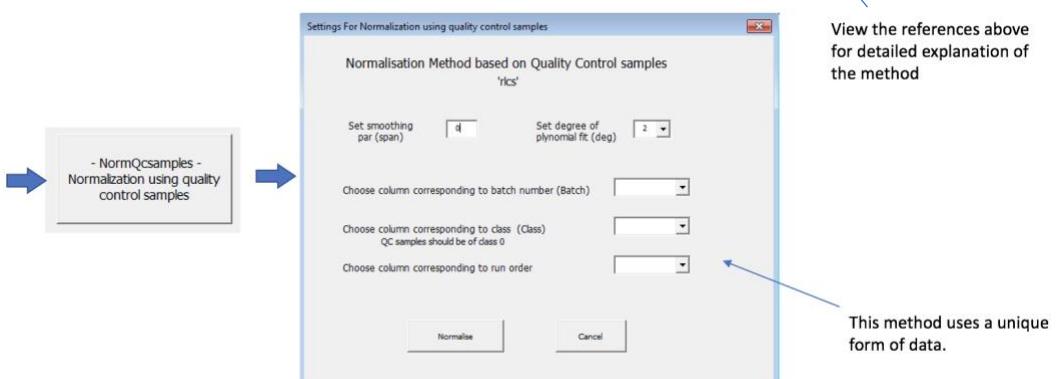
Upon clicking continue, you will return back to the settings sheet but you can notice some changes:



NormQcsamples

This function is based on the quality control sample based robust LOESS (locally estimated scatterplot smoothing) signal correction (QC-RLSC) method as described by Dunn et al. (2011) and implemented statTarget (Luan 2017)

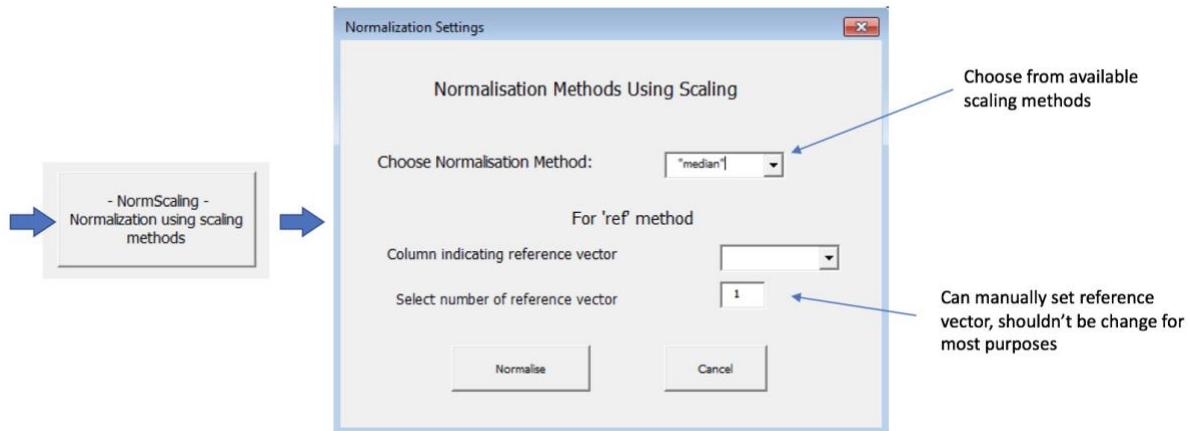
To Normalize:



NormScaling

The scaling normalization methods (Scholz et al. 2004, Wang et al. (2003)) included in the package are normalization to a total sum, normalisation by the median or mean of each sample, and are denoted by *sum*, *median*, and *mean* respectively. The method *ref* normalises the metabolite abundances to a specific reference vector such as the sample weight or volume.

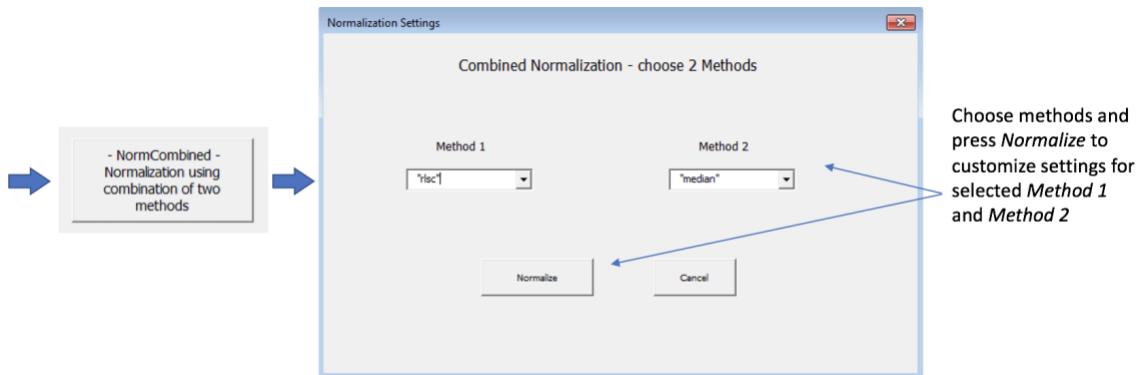
To Normalize:



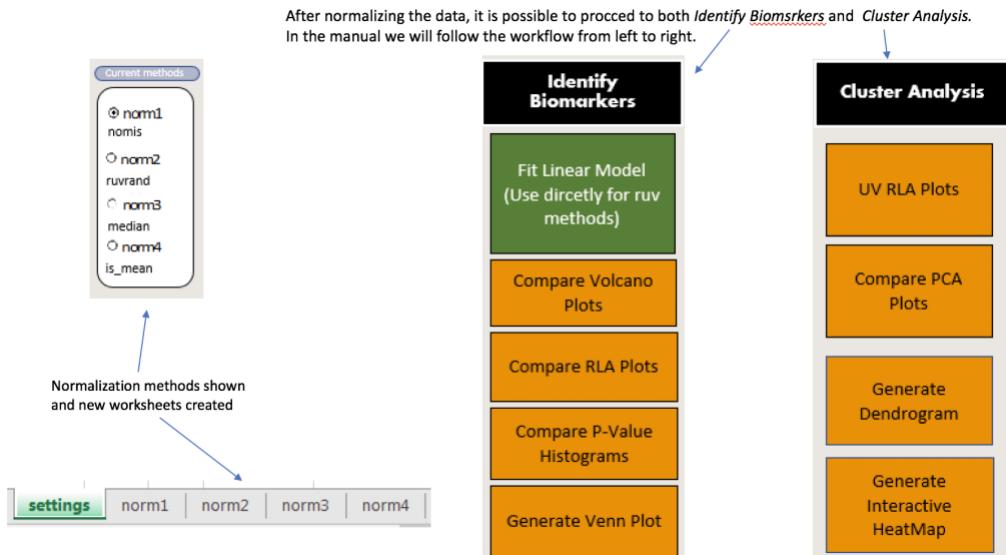
NormCombined

In some circumstances, researchers use a combination of the above normalizations (i.e., one method followed by another). This can be achieved using the *NormCombined* function. The function defaults to employing 'rlsc' approach followed by the 'median'.

To Normalize:



Note that normalizing the data is not necessary to proceed to fitting a linear model although it is highly recommended to try a few normalization methods when analysis data.

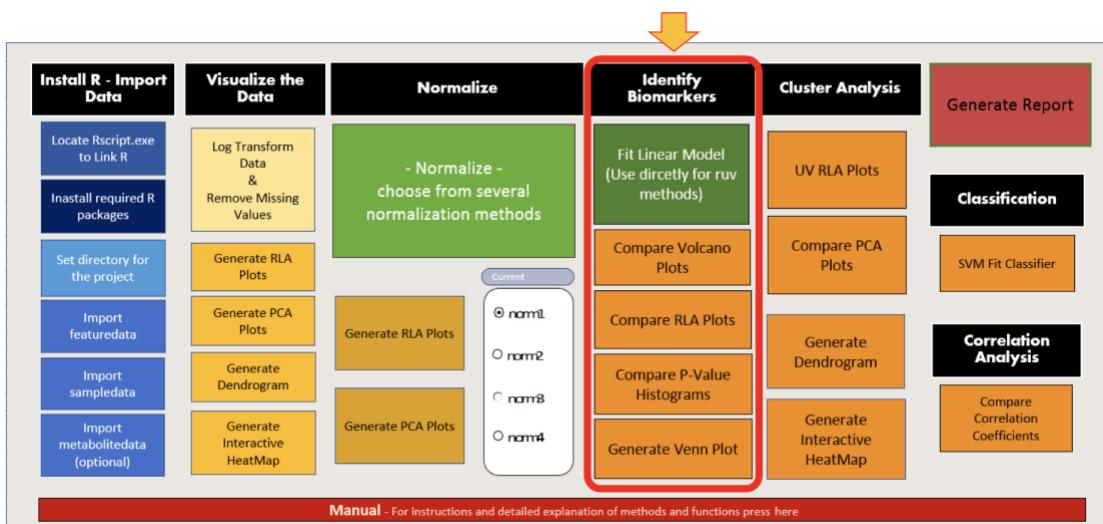


Assessing and choosing normalization methods

The criteria for assessing and choosing a normalization method implemented NormlizeMets have been described in detail by De Livera et al. (2012), De Livera et al. (2015) and J. A. Gagnon-Bartsch, Jacob, and Speed (2014).

Identifying Biomarkers

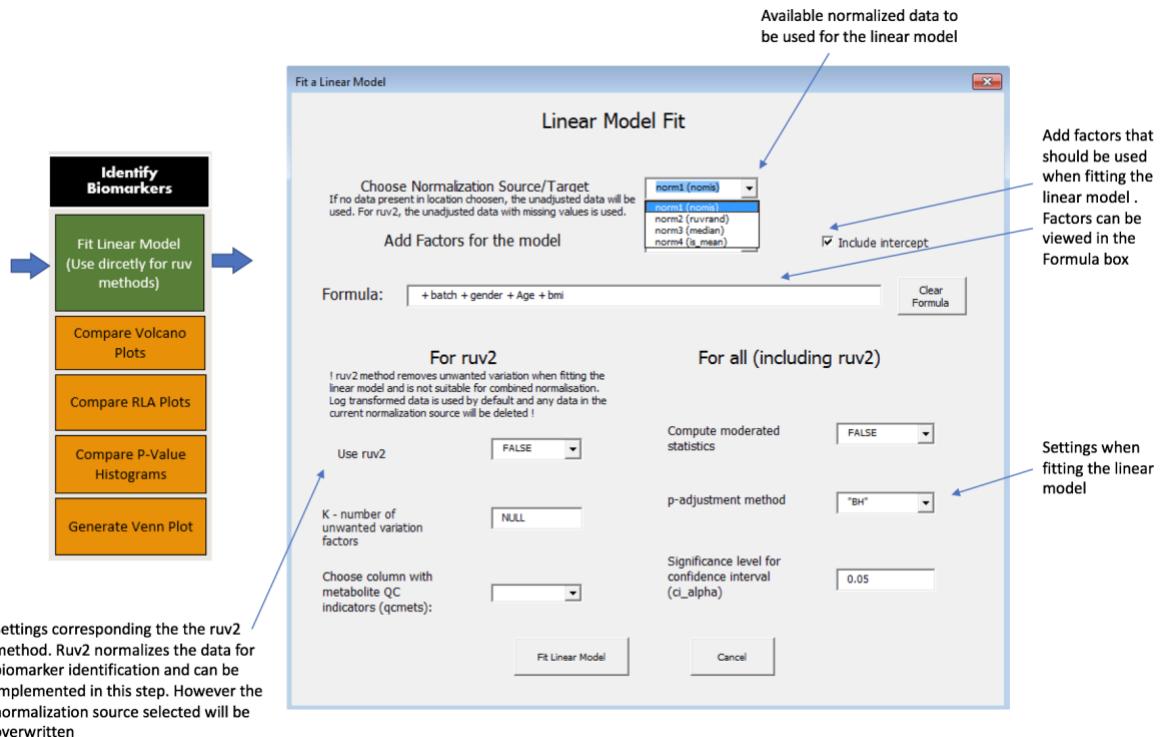
To view and compare the biomarkers identified, first a linear model has to be fitted to the data.



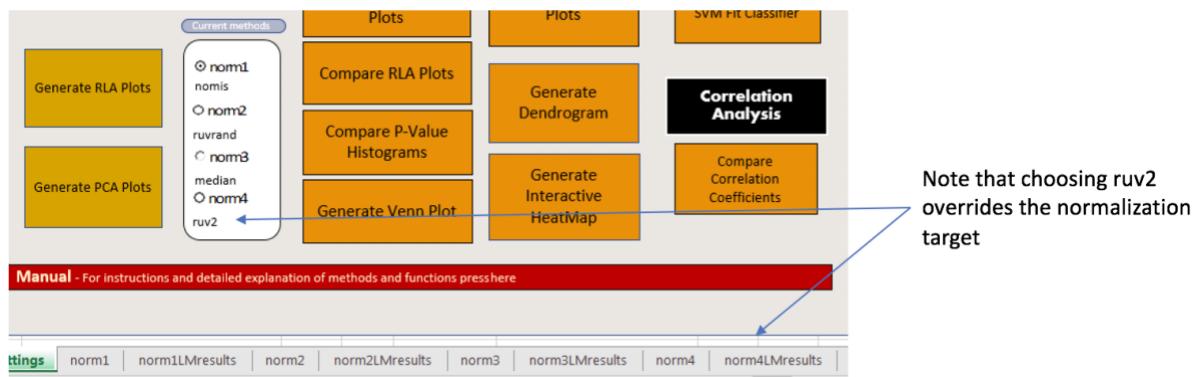
Fit Linear Model

A linear model has to be fitted for every Normalization method that is to be used down the line for Biomarker identification. Setting from one ‘run’ of the Fit Linear model will be saved for the next.

To Fit Linear Model:



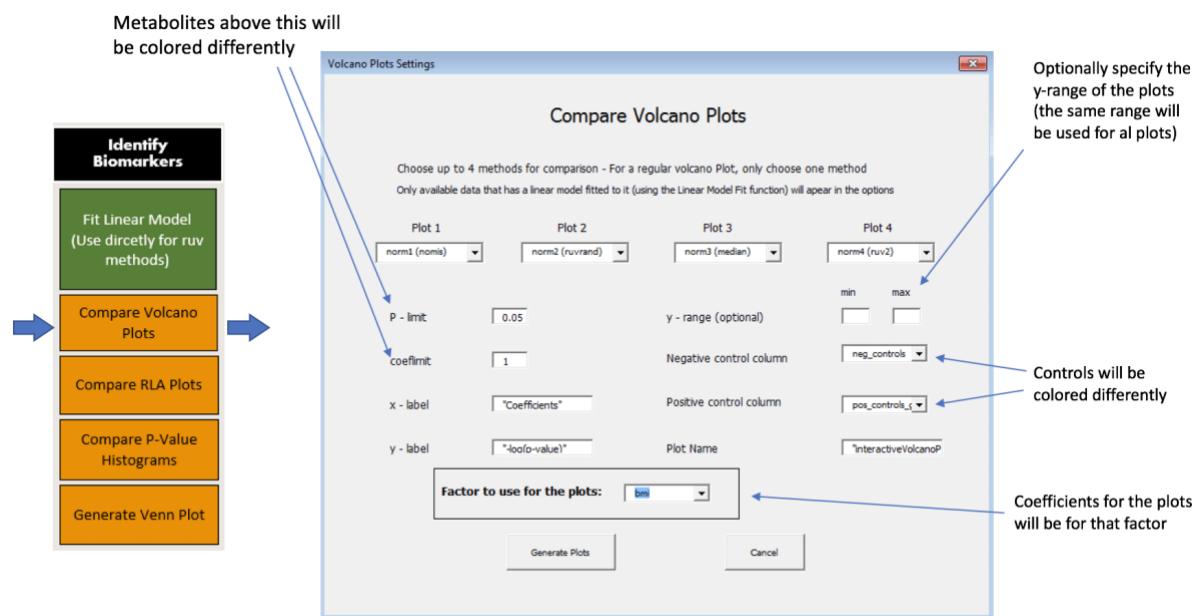
	A	B	C	D	E	F	G	H	I	J	K	L	M	
	F stat	P value	Adjusted F p value	cooff (intercept)	cooff batchbatch2	cooff gendergender1	cooff Age	cooff bmi	t stat (intercept)	t stat batchbatch2	t stat gendergender1	t stat Age	t stat bmi	
1	m_3	367.0640508	1.83e-40	-8.95e-01	8.34935337	0.000308303	0.000154971	0.00029044	42.8550360	0.599057272	0.0001749204	1.843111314	0.10733345	
2	m_1	1.000000000	0.99	4.23e-01	4.979238000	0.000100000	0.000000000	0.000000000	2.591000000	1.488000000	0.000000000	1.000000000	0.000000000	
3	m_3	39.818177081	1.29e-25	1.67e-25	6.793188000	0.020313445	0.300187081	-0.01077832	13.55311266	0.181800463	3.705513605	-0.672314342		
4	m_3	14.540055238	1.53e-11	1.76e-11	4.98533120	0.004521339	0.283799155	0.00104796	7.735175731	0.39048884	2.724893926	1.77603426		
5	m_4	74.81961081	2.05e-58	2.32e-58	6.6677930	0.124584715	0.065037851	0.01006986	18.91329620	1.320179023	1.147900604	-0.33111304		
6	m_3	15.47639232	1.23e-25	2.53e-25	2.059380000	0.000100000	-0.006372874	0.00021232	8.650757264	0.420692827	-0.705953196	0.454894642		
7	m_7	15.47639238	3.49e-12	4.08e-12	4.531485299	0.04880504	-0.063728784	0.013507399	85.50233518	0.494160694	-0.78447664	0.18047732		
8	m_8	1462.397870	4.02e-122	2.47e-120	8.10725599	0.054713137	-0.013513148	0.000184273	85.50233518	0.494160694	-0.78447664	0.18047732		
9	m_9	1226.58959	1.05e-16	3.23e-115	7.21770872	0.009903118	-0.004356857	-8.91e-05	78.30533970	0.48068160	-0.2919364	-0.10733345		
10	m_1	1.000000000	0.99	4.23e-01	4.979238000	0.000100000	0.000000000	0.000000000	2.591000000	1.488000000	0.000000000	1.000000000	0.000000000	
11	m_21	309.103156	1.93e-75	1.79e-75	10.0000000	0.000000000	0.000000000	0.000000000	39.15273064	1.251201934	1.843111314	-0.46717266		
12	m_21	309.103156	1.93e-75	1.79e-75	10.0000000	0.000000000	0.000000000	0.000000000	39.15273064	1.251201934	1.843111314	-0.46717266		
13	m_32	89.17162333	1.84e-42	3.23e-42	7.645185156	0.039796591	0.080723464	0.000568386	0.02113052	20.89799548	0.486640156	1.35377297	1.74875421	
14	m_34	88.20131119	2.13e-42	3.08e-42	8.13817702	0.035334734	0.126133374	0.000481462	-0.02077858	20.9552018	0.407048264	2.096520704	-1.320504835	
15	m_34	88.20131119	2.13e-42	3.08e-42	8.13817702	0.035334734	0.126133374	0.000481462	-0.02077858	20.9552018	0.407048264	2.096520704	-1.320504835	
16	m_19	45.54797620	4.31e-28	5.96e-28	5.25110527	0.09770374	0.040021189	0.03232399	0.01393338	14.40952391	1.17053969	0.714362020	0.004895076	
17	m_17	149.098252	2.60e-55	5.81e-55	7.47379830	0.022929971	0.174416195	0.00119121	0.00744416	27.01482123	0.472646389	8.317067487	-0.047635338	
18	m_18	32.90009500	2.29e-22	3.85e-22	4.9554007	0.008160398	0.044135203	0.002846402	0.03401139	12.64600705	0.913413199	0.493808483	-0.00550207	
19	m_20	18.10842953	4.09e-41	4.79e-41	5.9933842	0.114407096	0.004116894	-0.02088421	20.38594503	0.420692827	1.320504835	0.000000000		
20	m_20	18.10842953	4.09e-41	4.79e-41	5.9933842	0.114407096	0.004116894	-0.02088421	20.38594503	0.420692827	1.320504835	0.000000000		
21	m_21	13.1284853	1.51e-10	1.71e-10	2.81292127	0.174571041	0.051311434	0.002053416	0.05851318	6.11175203	1.71110084	-0.299121519	0.591837814	
22	m_22	261.150600	1.34e-70	1.44e-70	7.90517900	0.000179087	0.047239371	0.000180144	0.000121204	42.7720200	0.49517799	1.139411808	1.3075048	
23	m_23	1.000000000	0.99	4.23e-01	4.979238000	0.000100000	0.000000000	0.000000000	2.591000000	1.488000000	0.000000000	1.000000000	0.000000000	
24	m_24	190.9438527	6.23e-42	1.59e-41	6.416501638	0.051411949	0.20307016	0.019395348	29.7046369	1.064710241	7.445599934	2.41040648		
25	m_25	367.24408011	2.00e-40	1.02e-79	9.71953441	-0.009912337	0.104940607	-0.015994997	-0.004993575	42.4454536	-0.19504779	2.810682881	-0.75292248	
26	m_26	192.1110341	4.24e-42	1.11e-61	8.78784813	0.030552211	0.190005358	-0.000604004	-0.007441415	30.7026805	-0.553124672	4.091900082	-0.240493202	
27	m_27	26.18120479	2.05e-29	5.81e-29	5.81071000	0.000100000	0.000000000	0.000000000	2.591000000	1.488000000	0.000000000	1.330120808		
28	m_28	65.18120508	1.13e-33	1.66e-33	8.949487898	0.296430414	-0.238164862	-0.000503517	0.078147492	18.186695913	2.269644663	-2.4671454	-0.1330311376	



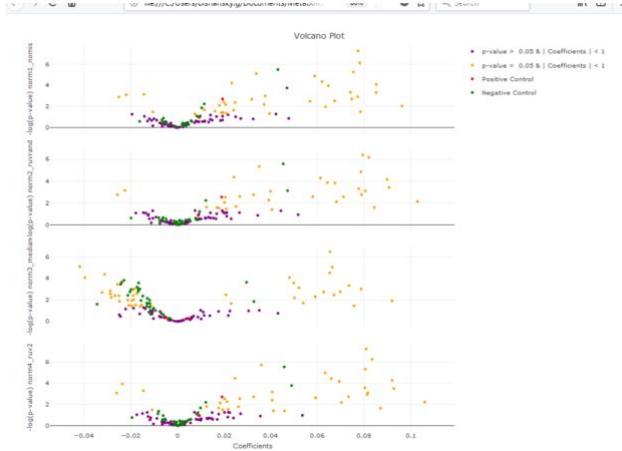
Volcano Plots

Volcano plots are useful in identifying biomarkers and generally assessing the normalization.

Compare Volcano Plots:



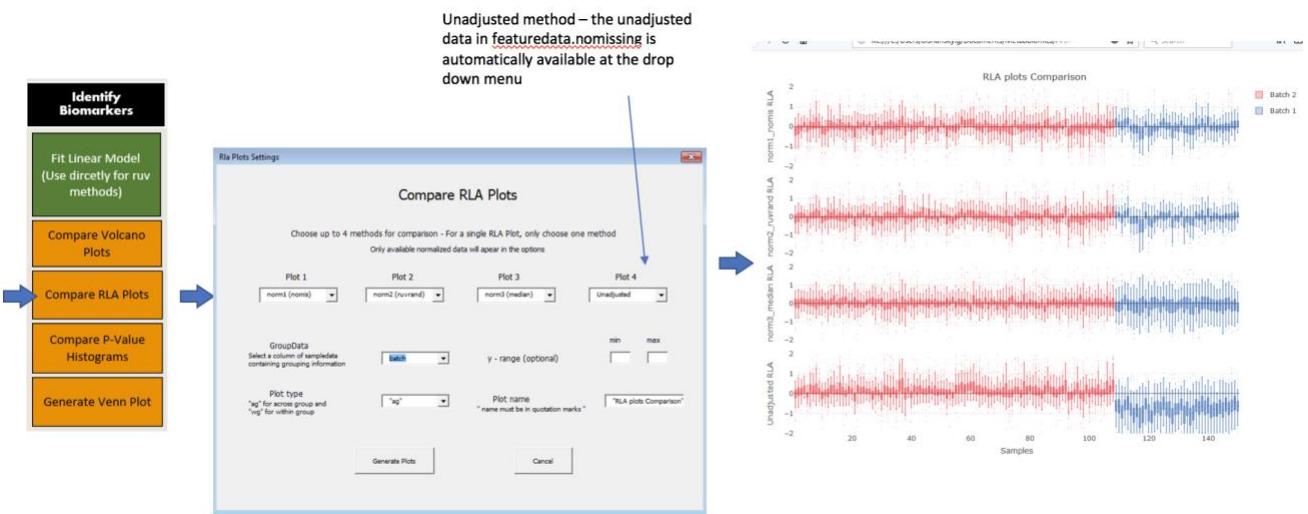
The plot is saved in the working directory and opens in the default browser.



Compare RLA plots

Used to assess normalization by comparing relative log abundance plots, similar input to the *Generate RLA* plots function

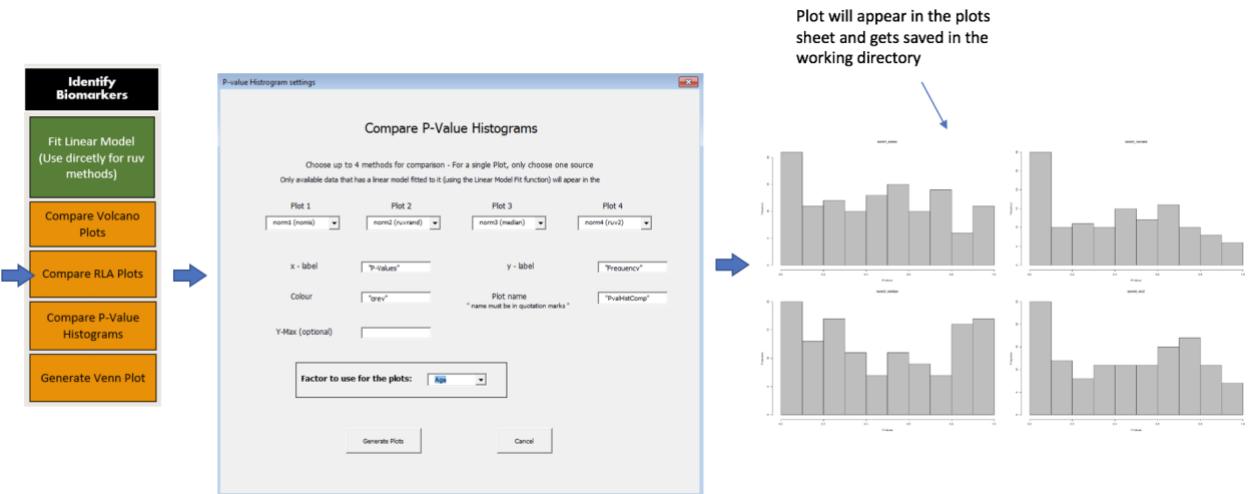
Compare Rla Plots:



Compare P-Value Histograms

Compare histograms of the coefficient's p-values. The distribution of the p-values should be used to assess the success of the normalization.

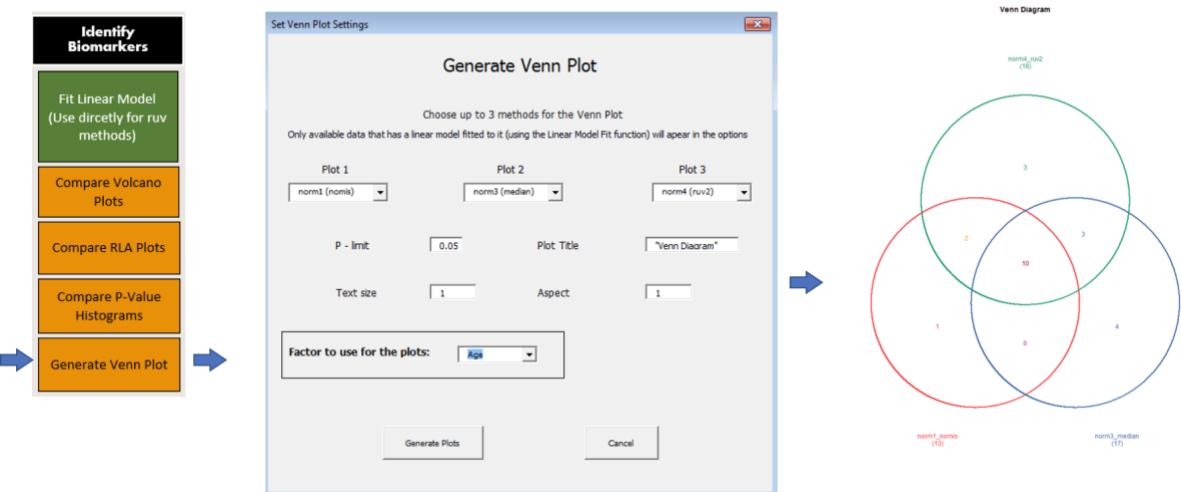
Compare P-Value histograms:



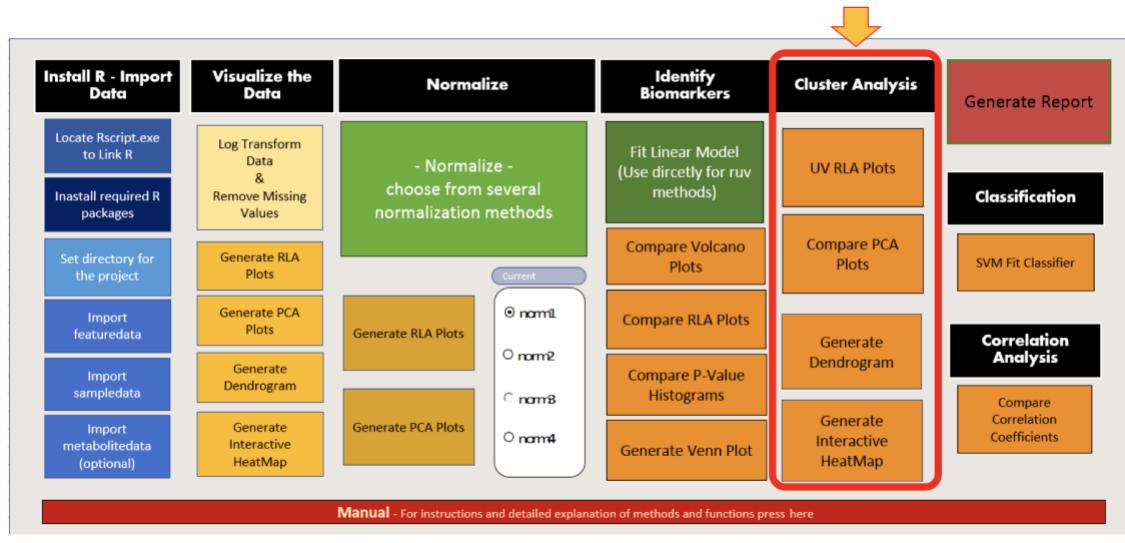
Generate Venn Plot

Generates a Venn plot that compares the biomarkers identified by the different normalization methods.

Generate Venn Plot:

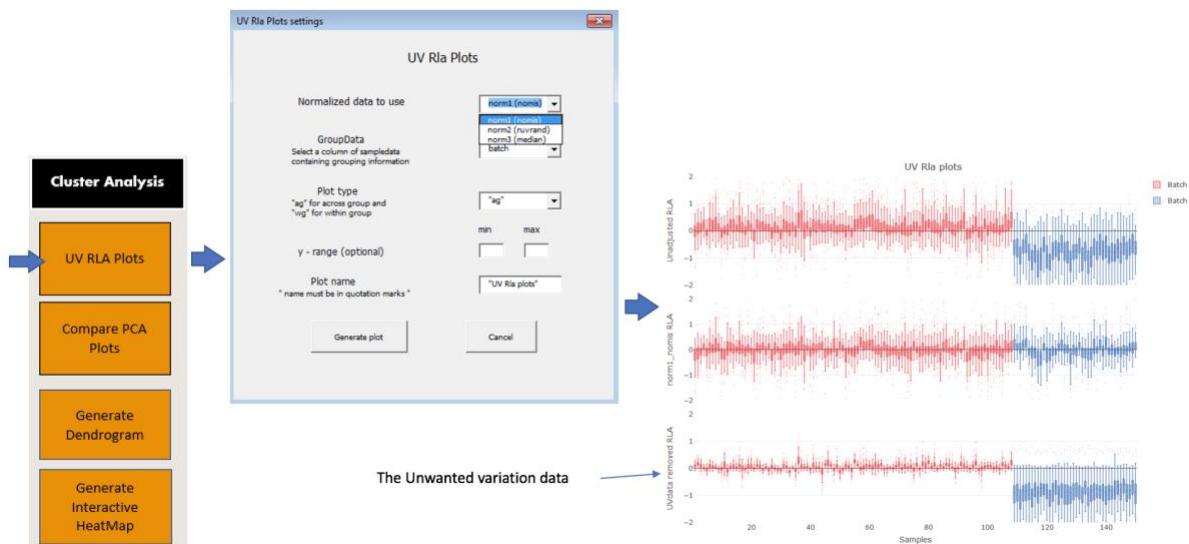


Cluster Analysis



UV RLA plots

Unwanted Variation relative log abundance plots enable visualisation of the unwanted variation removed by each normalization method.



Compare PCA Plots

Compare principal component multi-plots for differed normalization methods.



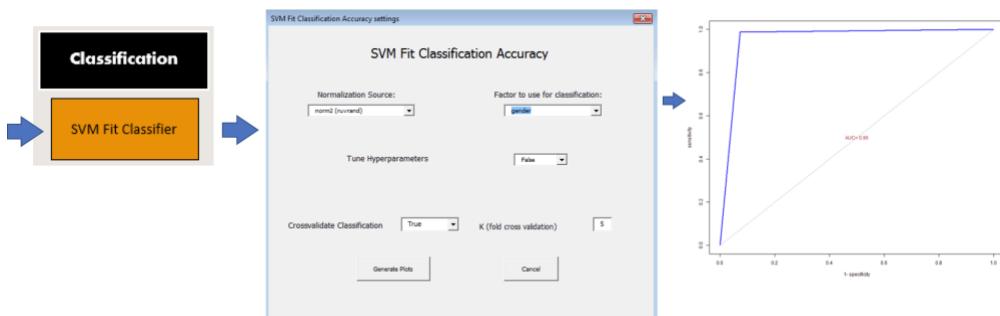
Generate Dendrogram and **Generate Interactive HeatMap** are identical to those discussed in the *Visualize the Data* section. The user has to choose the normalized data to be used.

Classification

Classification accuracy is a good way to assess the success of a given normalization method

SVM-Fit

The Support Vector Machines method is used to classify the data, the classification accuracy is then assessed based on the factor specified.



Correlation Analysis

It is important to look at correlation coefficients when normalizing data.

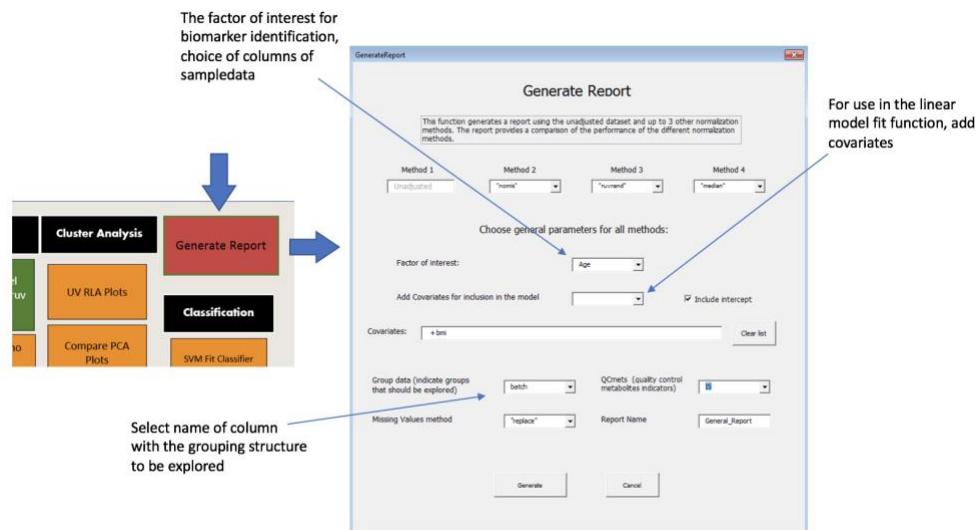
Compare Correlation coefficients



Generate Report

The *Generate Report* function generates an interactive report based on basic user input. There is a choice of up to 3 normalization methods to be included together with the unadjusted data. The report includes various plots and diagnostic to assess the normalization. Guidance on interpretation of the various plots, together with notes of what the user should look for when assessing the results is provided in the generated document.

Generate Report:



An example report is available in your downloaded ExNormalizeMets docs folder.

Full Package Vignette:

For package vignette with detailed explanations of methods and workflow, follow the link to:
https://cran.r-project.org/web/packages/NormalizeMets/vignettes/NormalizeMets_vignette.html