

# Package ‘MAPFX’

May 20, 2024

**Title** MASSively Parallel Flow cytometry Xplorer (MAPFX): A Toolbox for Analysing Data from the Massively-Parallel Cytometry Experiments

**Version** 1.0.0

**Description** MAPFX is an end-to-end toolbox that pre-processes the raw data from MPC experiments (e.g., BioLegend's LEGENDScreen and BD Lyoplates assays), and further imputes the ‘missing’ infinity markers in the wells without those measurements. The pipeline starts by performing background correction on raw intensities to remove the noise from electronic baseline restoration and fluorescence compensation by adapting a normal-exponential convolution model. Unwanted technical variation, from sources such as well effects, is then removed using a log-normal model with plate, column, and row factors, after which infinity markers are imputed using the informative backbone markers as predictors. The completed dataset can then be used for clustering and other statistical analyses. Additionally, MAPFX can be used to normalise data from FFC assays as well.

**Depends** R (>= 4.4.0)

**License** GPL-2

**Encoding** UTF-8

**RoxygenNote** 7.3.1

**Roxygen** list(markdown = TRUE)

**biocViews** Software, FlowCytometry, CellBasedAssays, SingleCell, Proteomics, Clustering

**Imports** flowCore, Biobase, stringr, uwot, iCellR, igraph, ggplot2, RColorBrewer, Rfast, ComplexHeatmap, circlize, glmnetUtils, e1071, xgboost, parallel, pbapply, reshape2, gtools, utils, stats, cowplot, methods, grDevices, graphics

**Suggests** BiocStyle, knitr, rmarkdown, testthat

**VignetteBuilder** knitr

**NeedsCompilation** no

**URL** <https://github.com/HsiaoChiLiao/MAPFX>

**BugReports** <https://github.com/HsiaoChiLiao/MAPFX/issues>

**git\_url** <https://git.bioconductor.org/packages/MAPFX>

**git\_branch** RELEASE\_3\_19

**git\_last\_commit** 4bbfac7

**git\_last\_commit\_date** 2024-04-30

**Repository** Bioconductor 3.19

**Date/Publication** 2024-05-19

**Author** Hsiao-Chi Liao [aut, cre] (<<https://orcid.org/0000-0002-9586-1246>>),  
 Agus Salim [ctb],  
 infinityFlow [ctb]

**Maintainer** Hsiao-Chi Liao <[chelsea.acad@gmail.com](mailto:chelsea.acad@gmail.com)>

## Contents

|                                       |           |
|---------------------------------------|-----------|
| bkc_bkb . . . . .                     | 2         |
| bkc_pe . . . . .                      | 4         |
| cluster.analysis . . . . .            | 5         |
| cluster.analysis.bkbOnly . . . . .    | 6         |
| fcs_to_rds . . . . .                  | 6         |
| fcs_to_rds_bkb . . . . .              | 7         |
| imputation_bkb.predictors . . . . .   | 8         |
| initM . . . . .                       | 9         |
| MapfxFFC . . . . .                    | 10        |
| MapfxMPC . . . . .                    | 12        |
| mkImputeMT . . . . .                  | 16        |
| ord.fcs.raw.meta.df.out_ffc . . . . . | 17        |
| ord.fcs.raw.meta.df.out_mpc . . . . . | 17        |
| ord.fcs.raw.mt_ffc . . . . .          | 18        |
| ord.fcs.raw.mt_mpc . . . . .          | 18        |
| rmBatchEffect . . . . .               | 19        |
| rmWellEffect . . . . .                | 19        |
| <b>Index</b>                          | <b>21</b> |

---

bkc\_bkb

*Background correction for the backbone markers*

---

## Description

This function has been designed to do background correction for the backbone markers by using normal-exponential convolution model.

**Usage**

```
bkc_bkb(  
  paths,  
  bkb.v,  
  MPC,  
  bkb.upper.quantile = 0.9,  
  bkb.lower.quantile = 0.1,  
  bkb.min.quantile = 0.01,  
  plots = TRUE  
)
```

**Arguments**

|                                 |                                                                                                                                                                   |
|---------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>paths</code>              | a vector of characters of paths to store input, intermediary results, outputs...                                                                                  |
| <code>bkb.v</code>              | a vector of the names of the backbone markers (MUST match to the names in the FCS file).                                                                          |
| <code>MPC</code>                | if the data is from MPC experiments, set MPC = TRUE. Setting FALSE represents data from the fluorescence flow cytometry (FFC) assay.                              |
| <code>bkb.upper.quantile</code> | the cut-off (default = 0.9) for selecting cells used for estimating the parameter of signal.                                                                      |
| <code>bkb.lower.quantile</code> | the cut-off (default = 0.1) for selecting cells used for estimating the parameters of noise.                                                                      |
| <code>bkb.min.quantile</code>   | the cut-off (default = 0.01) for omitting the cells with the smallest values to minimise the impact of outliers.                                                  |
| <code>plots</code>              | logical; if TRUE (default), produce scatter plots for pre- and post- background adjusted backbone markers (calibrated values on y-axis and raw values on x-axis). |

**Details**

Generating the calibrated measurements and save to `medpara_bkc.bkb_no.bkcPhy_mt.rds` file, and visualising the result with the scatter plots in the output directory.

**Value**

Background noise corrected backbone markers and graphs if specified

**Author(s)**

Hsiao-Chi Liao and Agus Salim

---

`bkc_pe`*Background correction for the well-specific markers (PE)*

---

**Description**

This function has been designed to do background correction for the well-specific markers (PE) by using normal-exponential convolution model.

**Usage**

```
bkc_pe(paths, pe.lower.quantile = 0.1, pe.min.quantile = 0.01, plots = TRUE)
```

**Arguments**

`paths` a vector of characters of paths to store input, intermediary results, outputs...

`pe.lower.quantile` the cut-off (default = 0.1) for selecting cells used for estimating the parameters of noise for infinity markers.

`pe.min.quantile` the cut-off (default = 0.01) for omitting the cells with the smallest values to minimise the impact of outliers for infinity markers.

`plots` logical; if TRUE (default), produce scatter plots for pre- and post- background adjusted infinity markers (calibrated values on y-axis and raw values on x-axis).

**Details**

Generating the calibrated measurements and save to `bkc.pe_mt.rds` file, and visualising the result with the scatter plots in the output directory.

**Value**

Background noise corrected infinity markers and graphs if specified

**Author(s)**

Hsiao-Chi Liao and Agus Salim

---

|                  |                                                                                        |
|------------------|----------------------------------------------------------------------------------------|
| cluster.analysis | <i>Cluster analysis with normalised backbone measurements and the complete dataset</i> |
|------------------|----------------------------------------------------------------------------------------|

---

## Description

This function has been designed to perform cluster analysis for the normalised backbone measurements and the complete dataset which includes the normalised backbone measurements and the imputed well-specific markers.

## Usage

```
cluster.analysis(paths, bkb.v, yvar = "Legend", control.wells, plots = TRUE)
```

## Arguments

|               |                                                                                                                                                                                         |
|---------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| paths         | a vector of characters of paths to store input, intermediary results, outputs...                                                                                                        |
| bkb.v         | a vector of the names of the backbone markers (MUST match to the names in the FCS file).                                                                                                |
| yvar          | the name of the well-specific marker in the FCS files (e.g., "Legend").                                                                                                                 |
| control.wells | the well label of the control wells, including the autofluorescence and the isotype controls (format: plate_well, e.g., P1_A01)                                                         |
| plots         | logical; if TRUE (default), produce an UMAP embedding plot from the normalised backbone markers and the imputed infinity markers to visualise the structure of the biological clusters. |

## Details

Updating the metadata for cells in the fcs\_metadata\_df.rds file, adding the information of the biological clusters from the clean and complete dataset, and visualising the result with the scatter plots in the output directory.

## Value

Metadata for cells with group labels from the cluster analysis

## Author(s)

Hsiao-Chi Liao

---

`cluster.analysis.bkbOnly`*Cluster analysis with normalised backbone measurements*

---

**Description**

This function has been designed to perform cluster analysis for the normalised backbone measurements.

**Usage**

```
cluster.analysis.bkbOnly(paths, bkb.v, plots = TRUE)
```

**Arguments**

|                    |                                                                                                                                                        |
|--------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>paths</code> | a vector of characters of paths to store input, intermediary results, outputs...                                                                       |
| <code>bkb.v</code> | a vector of the names of the backbone markers (MUST match to the names in the FCS file).                                                               |
| <code>plots</code> | logical; if TRUE (default), produce an UMAP embedding plot from the normalised backbone markers to visualise the structure of the biological clusters. |

**Details**

Updating the metadata for cells in the `fcs_metadata_df.rds` file, adding the information of the biological clusters from the clean and complete dataset, and visualising the result with the scatter plots in the output directory.

**Value**

Metadata for cells with group labels from the cluster analysis

**Author(s)**

Hsiao-Chi Liao

---

`fcs_to_rds`*Converting FCS files to RDS files*

---

**Description**

This function has been designed to convert the raw FCS files to data matrix and export to RDS files.

**Usage**

```
fcs_to_rds(paths, file_meta, yvar)
```

**Arguments**

|           |                                                                                                                                                                 |
|-----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| paths     | a vector of characters of paths to store input, intermediary results, outputs...                                                                                |
| file_meta | if the file names of the FCS files are in the specified format, set file_meta="auto"; otherwise set file_meta="usr" and provide "filename_meta.csv" in FCSpath. |
| yvar      | the name of the well-specific marker in the FCS files (e.g., "Legend").                                                                                         |

**Details**

Generating fcs\_metadata\_df.rds and fcs\_rawInten\_mt.rds files in the output directory.

**Value**

Raw protein intensities and the corresponding metadata from MPC experiments

**Author(s)**

Hsiao-Chi Liao

---

|                |                                                                                     |
|----------------|-------------------------------------------------------------------------------------|
| fcs_to_rds_bkb | <i>Converting FCS files to RDS files (for the case without exploratory markers)</i> |
|----------------|-------------------------------------------------------------------------------------|

---

**Description**

This function has been designed to convert the raw FCS files to data matrix and export to RDS files.

**Usage**

```
fcs_to_rds_bkb(paths, file_meta, MPC)
```

**Arguments**

|           |                                                                                                                                                                 |
|-----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| paths     | a vector of characters of paths to store input, intermediary results, outputs...                                                                                |
| file_meta | if the file names of the FCS files are in the specified format, set file_meta="auto"; otherwise set file_meta="usr" and provide "filename_meta.csv" in FCSpath. |
| MPC       | if the data is from MPC experiments, set MPC = TRUE. Setting FALSE represents data from the fluorescence flow cytometry (FFC) assay.                            |

**Details**

Generating fcs\_metadata\_df.rds and fcs\_rawInten\_mt.rds files in the output directory.

**Value**

Raw protein intensities and the corresponding metadata from FFC experiments

**Author(s)**

Hsiao-Chi Liao

---

```
imputation_bkb.predictors
```

*Imputing the unmeasured well-specific markers with regression models*

---

## Description

This function has been designed to impute/predict the unmeasured well-specific markers with regression models.

## Usage

```
imputation_bkb.predictors(
  paths,
  chans,
  yvar = "Legend",
  cores = 4L,
  models.use,
  extra_args_regression_params,
  prediction_events_downsampling = NULL,
  impu.training,
  plots = TRUE
)
```

## Arguments

|                                             |                                                                                                                                                                                                       |
|---------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>paths</code>                          | a vector of characters of paths to store input, intermediary results, outputs...                                                                                                                      |
| <code>chans</code>                          | a vector of the names of the backbone markers (MUST match to the names in the FCS file).                                                                                                              |
| <code>yvar</code>                           | the name of the well-specific marker in the FCS files (has been changed to "Legend" in the first function).                                                                                           |
| <code>cores</code>                          | the number of cores used to perform parallel computation (default = 8L).                                                                                                                              |
| <code>models.use</code>                     | a vector of the names of the models used for imputation (an example: <code>c("LM", "LASSO3", "SVM", "XGBoost")</code> ). The length of the vector is arbitrary.                                       |
| <code>extra_args_regression_params</code>   | a list of the lists of the parameters for running regression models.                                                                                                                                  |
| <code>prediction_events_downsampling</code> | default = NULL (not doing subsampling). How many cells per well you want to have the imputation? (must be less than or equal to a half as we won't get the prediction for cells in the training set). |
| <code>impu.training</code>                  | logical; if FALSE (default), not impute the training set (the dataset used to train the imputation models).                                                                                           |
| <code>plots</code>                          | logical; if TRUE (default), visualise the distribution of R-sq from each infinity marker.                                                                                                             |



**Details**

This function returns the object of imputation of the unmeasured well-specific markers. In the output directory, the imputations are saved to predictions.Rds file. Visualisation of the imputation accuracy will be provided if specified.

**Value**

A list of imputations

**Author(s)**

Hsiao-Chi Liao and InfinityFlow (Becht et. al, 2021)

---

|       |                                    |
|-------|------------------------------------|
| initM | <i>Initial biological clusters</i> |
|-------|------------------------------------|

---

**Description**

Generating the M matrix for removing well effects.

**Usage**

```
initM(paths, assay, bkb.v, plots = TRUE)
```

**Arguments**

|       |                                                                                                                                                                          |
|-------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| paths | a vector of characters of paths to store input, intermediary results, outputs...                                                                                         |
| assay | the type of the input data - MPC or FFC.                                                                                                                                 |
| bkb.v | a vector of the names of the backbone markers (MUST match to the names in the FCS file).                                                                                 |
| plots | logical; if TRUE (default), produce a UMAP embedding plot to visualise the structure of the biological clusters to form the initial M matrix for removal of well effect. |

**Details**

This function has been designed to find initial biological clusters with centred transformed data. Updating the metadata for cells in the fcs\_metadata\_df.rds file, adding the information of the initial biological clusters, and visualising the result with the scatter plots in the output directory.

**Value**

Metadata for cells with the initial biological clusters labels added

**Author(s)**

Hsiao-Chi Liao

---

 MapfxFFC

*Normalising data from the Fluorescence Flow Cytometry (FFC) Experiments with mapfx.norm*


---

## Description

This function is used to normalise, including background correction and removal of batch effects, protein intensity data from FFC assays. The input data is in FCS format. The functions include data normalisation and cluster analysis.

## Usage

```
MapfxFFC(
  runVignette = FALSE,
  runVignette_meta = NULL,
  runVignette_rawInten = NULL,
  FCspath = NULL,
  Outpath = NULL,
  protein.v = NULL,
  protein.upper.quantile = 0.9,
  protein.lower.quantile = 0.1,
  protein.min.quantile = 0.01,
  plots.bkc.protein = TRUE,
  plots.initM = TRUE,
  plots.rmBatchEffect = TRUE,
  cluster.analysis.protein = TRUE,
  plots.cluster.analysis.protein = TRUE
)
```

## Arguments

**runVignette** logical; if FALSE (default), specify a path to FCspath argument; TRUE for running Vignette using built-in data.

**runVignette\_meta** the argument for the built-in metadata when running Vignette; NULL (default).

**runVignette\_rawInten** the argument for the built-in raw intensities when running Vignette; NULL (default).

**FCspath** path to the input directory where filename\_meta.csv and FCS files are stored. filename\_meta.csv should be saved under FCspath/FCS/meta/ and FCS files should be saved under FCspath/FCS/fcs/. See Vignette for details.

**Outpath** path to the output directory where intermediate results and final results will be stored.

**protein.v** a vector of the names of the protein markers (MUST be the same as the names in the FCS files). For example, protein.v = c("FSC-H", "FSC-W", "SSC-H", "SSC-W", "CD3", "CD4", "CD8

|                                             |                                                                                                                                                                                  |
|---------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>protein.upper.quantile</code>         | the cut-off (default = 0.9) for selecting cells used for estimating the parameter of signal for protein markers.                                                                 |
| <code>protein.lower.quantile</code>         | the cut-off (default = 0.1) for selecting cells used for estimating the parameters of noise for protein markers.                                                                 |
| <code>protein.min.quantile</code>           | the cut-off (default = 0.01) for omitting the cells with the smallest values to minimise the impact of outliers during estimation.                                               |
| <code>plots.bkc.protein</code>              | logical; if TRUE (default), produce scatter plots for pre- and post- background adjusted protein markers (calibrated values on y-axis and raw values on x-axis).                 |
| <code>plots.initM</code>                    | logical; if TRUE (default), produce an UMAP embedding plot to visualise the structure of the biological clusters used to form the initial M matrix for removal of batch effects. |
| <code>plots.rmBatchEffect</code>            | logical; if TRUE (default), produce heatmaps to visualise the unwanted (batch) effects and biological effects in the pre- and post- adjusted datasets.                           |
| <code>cluster.analysis.protein</code>       | logical; if TRUE (default), perform cluster analysis using normalised protein markers.                                                                                           |
| <code>plots.cluster.analysis.protein</code> | logical; if TRUE (default), produce an UMAP embedding plot from the normalised protein markers to visualise the structure of the biological clusters.                            |

### Details

In the output directory, this function produces the normalised protein measurements, cell group labels from the cluster analysis using normalised proteins, and graphs will be provided if specified.

### Value

Normalised protein markers on log scale and metadata for cells

### Author(s)

Hsiao-Chi Liao, Agus Salim

### Examples

```
# import built-in data
data(ord.fcs.raw.meta.df.out_ffc)
data(ord.fcs.raw.mt_ffc)

# create an Output directory for the MapfxFFC function
dir.create(file.path(tempdir(), "FFCnorm_Output"))

MapfxFFC_obj <- MapfxFFC(
  runVignette = TRUE, #set FALSE if not running this example
```

```

runVignette_meta = ord.fcs.raw.meta.df.out_ffc, #set NULL if not running this example
runVignette_rawInten = ord.fcs.raw.mt_ffc, #set NULL if not running this example
FCspath = NULL, # users specify their own input path if not running this example
Outpath = file.path(tempdir(), "FFCnorm_Output"),
protein.v = c("CD3", "CD4", "CD8", "CD45"),
protein.upper.quantile = 0.9,
protein.lower.quantile = 0.1,
protein.min.quantile = 0.01,
plots.bkc.protein = TRUE,
plots.initM = TRUE,
plots.rmBatchEffect = TRUE,
cluster.analysis.protein = TRUE, plots.cluster.analysis.protein = TRUE)

```

---

MapfxMPC

*MAssively Parallel Flow cytometry Xplorer (MAPFX)*


---

## Description

This function is an end-to-end toolbox for analysing single-cell protein intensity data from the Massively-Parallel Cytometry (MPC) Experiments in FCS format. The functions include data normalisation, imputation (using backbone markers), and cluster analysis.

## Usage

```

MapfxMPC(
  runVignette = FALSE,
  runVignette_meta = NULL,
  runVignette_rawInten = NULL,
  FCspath = NULL,
  Outpath = NULL,
  file_meta = "auto",
  bkb.v = NULL,
  yvar = "Legend",
  control.wells = NULL,
  bkb.upper.quantile = 0.9,
  bkb.lower.quantile = 0.1,
  bkb.min.quantile = 0.01,
  inf.lower.quantile = 0.1,
  inf.min.quantile = 0.01,
  plots.bkc.bkb = TRUE,
  plots.bkc.inf = TRUE,
  plots.initM = TRUE,
  plots.rmWelleffect = TRUE,
  impute = TRUE,
  models.use = c("XGBoost"),
  extra_args_regression_params = list(list(nrounds = 1500, eta = 0.03)),

```

```

prediction_events_downsampling = NULL,
impu.training = FALSE,
plots.imputation = TRUE,
cluster.analysis.bkb = TRUE,
plots.cluster.analysis.bkb = TRUE,
cluster.analysis.all = TRUE,
plots.cluster.analysis.all = TRUE,
cores = 4L
)

```

## Arguments

|                                   |                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|-----------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>runVignette</code>          | logical; if FALSE (default), specify a path to <code>FCspath</code> argument; TRUE for running Vignette using built-in data.                                                                                                                                                                                                                                                                                                                  |
| <code>runVignette_meta</code>     | the argument for the built-in metadata when running Vignette; NULL (default).                                                                                                                                                                                                                                                                                                                                                                 |
| <code>runVignette_rawInten</code> | the argument for the built-in raw intensities when running Vignette; NULL (default).                                                                                                                                                                                                                                                                                                                                                          |
| <code>FCspath</code>              | path to the input directory where <code>filename_meta.csv</code> and FCS files are stored (one file per well). <code>filename_meta.csv</code> should be saved under <code>FCspath/FCS/meta/</code> and FCS files should be saved under <code>FCspath/FCS/fcs/</code> (See Vignette for details.)                                                                                                                                              |
| <code>Outpath</code>              | path to the output directory where intermediate results and final results will be stored.                                                                                                                                                                                                                                                                                                                                                     |
| <code>file_meta</code>            | if the file names of the FCS files are in the specified format, set <code>file_meta = "auto"</code> ; otherwise set <code>file_meta = "usr"</code> and provide a <code>filename_meta.csv</code> file in <code>FCspath/FCS/meta/</code> .                                                                                                                                                                                                      |
| <code>bkb.v</code>                | a vector of the names of the backbone markers (MUST be the same as the names in the FCS files). For example, <code>bkb.v = c("FSC-H", "FSC-W", "SSC-H", "SSC-W", "CD69-CD301b", "MHCII", "CD4", "CD44", "CD8", "CD11c", "CD11b", "F480", "Ly6C", "Lineage", "CD45a488", "CD24", "CD103")</code> .                                                                                                                                             |
| <code>yvar</code>                 | the name of the well-specific exploratory marker in the FCS files (e.g., "Legend").                                                                                                                                                                                                                                                                                                                                                           |
| <code>control.wells</code>        | the well label of the control wells, including the autofluorescence and the isotype controls (format: <code>plate_well</code> , e.g., <code>P1_A01</code> ). Users need to provide this information when <code>cluster.analysis.all = TRUE</code> . For example, <code>control.wells = c("P1_A01", "P2_A01", "P3_A01", "P3_F04", "P3_F05", "P3_F06", "P3_F07", "P3_F08", "P3_F09", "P3_F10", "P3_F11", "P3_F12", "P3_G01", "P3_G02")</code> . |
| <code>bkb.upper.quantile</code>   | the cut-off (default = 0.9) for selecting cells used for estimating the parameter of signal for backbone markers.                                                                                                                                                                                                                                                                                                                             |
| <code>bkb.lower.quantile</code>   | the cut-off (default = 0.1) for selecting cells used for estimating the parameters of noise for backbone markers.                                                                                                                                                                                                                                                                                                                             |

|                                             |                                                                                                                                                                                                                                                                                                                                                                                |
|---------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>bkb.min.quantile</code>               | the cut-off (default = 0.01) for omitting the cells with the smallest values to minimise the impact of outliers during estimation (backbone).                                                                                                                                                                                                                                  |
| <code>inf.lower.quantile</code>             | the cut-off (default = 0.1) for selecting cells used for estimating the parameters of noise for infinity markers.                                                                                                                                                                                                                                                              |
| <code>inf.min.quantile</code>               | the cut-off (default = 0.01) for omitting the cells with the smallest values to minimise the impact of outliers during estimation (infinity).                                                                                                                                                                                                                                  |
| <code>plots.bkc.bkb</code>                  | logical; if TRUE (default), produce scatter plots for pre- and post- background adjusted backbone markers (calibrated values on y-axis and raw values on x-axis).                                                                                                                                                                                                              |
| <code>plots.bkc.inf</code>                  | logical; if TRUE (default), produce scatter plots for pre- and post- background adjusted infinity markers (calibrated values on y-axis and raw values on x-axis).                                                                                                                                                                                                              |
| <code>plots.initM</code>                    | logical; if TRUE (default), produce an UMAP embedding plot to visualise the structure of the biological clusters used to form the initial M matrix for removal of well effects.                                                                                                                                                                                                |
| <code>plots.rmWellEffect</code>             | logical; if TRUE (default), produce heatmaps to visualise the unwanted (well) effects and biological effects in the pre- and post- adjusted datasets.                                                                                                                                                                                                                          |
| <code>impute</code>                         | logical; if TRUE (default), impute the missing infinity markers.                                                                                                                                                                                                                                                                                                               |
| <code>models.use</code>                     | a vector of the names of the models used for imputation. For example, <code>models.use = c("LM", "LASSO3", "SVM", "XGBoost")</code> .                                                                                                                                                                                                                                          |
| <code>extra_args_regression_params</code>   | a list of the lists of the parameters for running regression models. The order should be the same as the models specified in <code>models.use</code> . For example, <code>extra_args_regression_params = list(list(degree = 1), list(nfolds = 10, degree = 3), list(type = "nu-regression", cost = 8, nu = 0.5, kernel = "radial"), list(nrounds = 1500, eta = 0.03))</code> . |
| <code>prediction_events_downsampling</code> | integer (default = NULL); the number of samples used for the downsampling for the prediction.                                                                                                                                                                                                                                                                                  |
| <code>impu.training</code>                  | logical; if FALSE (default), not impute the training set (the dataset used to train the imputation models).                                                                                                                                                                                                                                                                    |
| <code>plots.imputation</code>               | logical; if TRUE (default), visualise the distribution of R-sq values of infinity markers.                                                                                                                                                                                                                                                                                     |
| <code>cluster.analysis.bkb</code>           | logical; if TRUE (default), perform cluster analysis using normalised backbone markers for all cells.                                                                                                                                                                                                                                                                          |
| <code>plots.cluster.analysis.bkb</code>     | logical; if TRUE (default), produce an UMAP embedding plot from the normalised backbone markers to visualise the structure of the biological clusters for all cells.                                                                                                                                                                                                           |

|                                         |                                                                                                                                                                                                                                                                  |
|-----------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>cluster.analysis.all</code>       | logical; must set FALSE if <code>impute = FALSE</code> ; if TRUE (default), perform cluster analysis using normalised backbone markers and imputed infinity markers for cells in testing set.                                                                    |
| <code>plots.cluster.analysis.all</code> | logical; must set FALSE if <code>impute = FALSE</code> ; if TRUE (default), produce an UMAP embedding plot from the normalised backbone markers and the imputed infinity markers to visualise the structure of the biological clusters for cells in testing set. |
| <code>cores</code>                      | the number of cores used to perform parallel computation during the imputation process (default = 4L).                                                                                                                                                           |

### Details

In the output directory, this function produces the normalised backbone measurements, the background corrected infinity measurements, and imputed infinity markers (if set `impute = TRUE`), cell group labels from the cluster analysis using both normalised backbones and the completed dataset (if `impute = TRUE`), and graphs will be provided if specified.

### Value

Normalised backbone markers on log scale, background noise corrected infinity markers, imputations, and metadata for cells

### Author(s)

Hsiao-Chi Liao, Agus Salim, and InfinityFlow (Becht et. al, 2021)

### Examples

```
# import built-in data
data(ord.fcs.raw.meta.df.out_mpc)
data(ord.fcs.raw.mt_mpc)

# create an Output directory for the MapfxMPC function
dir.create(file.path(tempdir(), "MPC_imp_u_Output"))

# When `impute = TRUE`, randomly selecting 50% of the cells in each well for model training
set.seed(123)
MapfxMPC_imp_u_obj <- MapfxMPC(
  runVignette = TRUE, #set FALSE if not running this example
  runVignette_meta = ord.fcs.raw.meta.df.out_mpc, #set NULL if not running this example
  runVignette_rawInten = ord.fcs.raw.mt_mpc, #set NULL if not running this example
  FCSpath = NULL, # users specify their own input path if not running this example
  Outputpath = file.path(tempdir(), "MPC_imp_u_Output"),
  file_meta = "auto",
  bkb.v = c(
    "FSC-H", "FSC-W", "SSC-H", "SSC-W", "CD69-CD301b", "MHCII",
    "CD4", "CD44", "CD8", "CD11c", "CD11b", "F480",
    "Ly6C", "Lineage", "CD45a488", "CD24", "CD103"),
  yvar = "Legend",
```

```

control.wells = c(
  "P1_A01", "P2_A01", "P3_A01",
  "P3_F04", "P3_F05", "P3_F06", "P3_F07", "P3_F08",
  "P3_F09", "P3_F10", "P3_F11", "P3_F12",
  "P3_G01", "P3_G02"),
bkb.upper.quantile = 0.9,
bkb.lower.quantile = 0.1,
bkb.min.quantile = 0.01,
inf.lower.quantile = 0.1,
inf.min.quantile = 0.01,
plots.bkc.bkb = TRUE, plots.bkc.inf = TRUE,
plots.initM = TRUE,
plots.rmWellEffect = TRUE,
impute = TRUE,
models.use = c("XGBoost"),
extra_args_regression_params = list(list(nrounds = 1500, eta = 0.03)),
prediction_events_downsampling = NULL,
impu.training = FALSE,
plots.imputation = TRUE,
cluster.analysis.bkb = TRUE, plots.cluster.analysis.bkb = TRUE,
cluster.analysis.all = TRUE, plots.cluster.analysis.all = TRUE,
cores = 2L)

```

---

mkImputeMT

*Making the input for imputation*


---

## Description

This function has been designed to combine the normalised backbone measurements and the normalised PE markers for later imputation.

## Usage

```
mkImputeMT(paths)
```

## Arguments

paths                    a vector of characters of paths to store input, intermediary results, outputs...

## Details

Generating the combined data and saving to `impu.input_log.mt.rds` (on log scale) in the output directory.

## Value

Combined normalised backbone and infinity markers for imputation



**Author(s)**

Hsiao-Chi Liao

---

`ord.fcs.raw.meta.df.out_ffc`

*Metadata of the example MPC data*

---

**Description**

Metadata of the example MPC data

**Usage**

`data(ord.fcs.raw.meta.df.out_ffc)`

**Format**

a data.frame

---

`ord.fcs.raw.meta.df.out_mpc`

*Metadata of the example FFC data*

---

**Description**

Metadata of the example FFC data

**Usage**

`data(ord.fcs.raw.meta.df.out_mpc)`

**Format**

a data.frame

---

|                    |                                                                                          |
|--------------------|------------------------------------------------------------------------------------------|
| ord.fcs.raw.mt_ffc | <i>Subset of the single-cell murine lung data at steady state from an MPC experiment</i> |
|--------------------|------------------------------------------------------------------------------------------|

---

**Description**

Subset of the single-cell murine lung data at steady state from an MPC experiment

**Usage**

```
data(ord.fcs.raw.mt_ffc)
```

**Format**

a matrix containing cells from 266 wells (50 cells/well)

**Source**

<https://flowrepository.org/id/FR-FCM-Z2LP>

---

|                    |                                                                                                |
|--------------------|------------------------------------------------------------------------------------------------|
| ord.fcs.raw.mt_mpc | <i>Subset of the sorted CD4+ and CD8+ T cells from mice splenocytes from an FFC experiment</i> |
|--------------------|------------------------------------------------------------------------------------------------|

---

**Description**

Subset of the sorted CD4+ and CD8+ T cells from mice splenocytes from an FFC experiment

**Usage**

```
data(ord.fcs.raw.mt_mpc)
```

**Format**

a matrix containing cells from 5 batches (50 cells/batch)

**Source**

<http://flowrepository.org/id/FR-FCM-Z6UG>

---

|               |                                                  |
|---------------|--------------------------------------------------|
| rmBatchEffect | <i>Removing batch effect from the data (FFC)</i> |
|---------------|--------------------------------------------------|

---

**Description**

This function has been designed to remove the unwanted effects (batch effects) from the background corrected measurements.

**Usage**

```
rmBatchEffect(paths, plots = TRUE)
```

**Arguments**

|       |                                                                                                                                                        |
|-------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| paths | a vector of characters of paths to store input, intermediary results, outputs...                                                                       |
| plots | logical; if TRUE (default), produce heatmaps to visualise the unwanted (batch) effects and biological effects in the pre- and post- adjusted datasets. |

**Details**

Generating the calibrated measurements and saving to bkc.adj.bkb\_logScale\_mt.rds (on log scale) and bkc.adj.bkb\_linearScale\_mt.rds (on linear scale), and visualising the result with the heatmaps in the output directory.

**Value**

Normalised markers on log scale

**Author(s)**

Hsiao-Chi Liao

---

|              |                                           |
|--------------|-------------------------------------------|
| rmWellEffect | <i>Removing well effect from the data</i> |
|--------------|-------------------------------------------|

---

**Description**

This function has been designed to remove the unwanted effects (well effects) from the background corrected measurements.

**Usage**

```
rmWellEffect(paths, plots = TRUE)
```

**Arguments**

`paths` a vector of characters of paths to store input, intermediary results, outputs...  
`plots` logical; if TRUE (default), produce heatmaps to visualise the unwanted (well) effects and biological effects in the pre- and post- adjusted datasets.

**Details**

Generating the calibrated measurements and saving to `bkc.adj.bkb_logScale_mt.rds` (on log scale) and `bkc.adj.bkb_linearScale_mt.rds` (on linear scale), and visualising the result with the heatmaps in the output directory.

**Value**

Normalised backbone markers on log scale

**Author(s)**

Hsiao-Chi Liao

# Index

## \* datasets

ord.fcs.raw.meta.df.out\_ffc, [17](#)

ord.fcs.raw.meta.df.out\_mpc, [17](#)

ord.fcs.raw.mt\_ffc, [18](#)

ord.fcs.raw.mt\_mpc, [18](#)

bkc\_bkb, [2](#)

bkc\_pe, [4](#)

cluster.analysis, [5](#)

cluster.analysis.bkbOnly, [6](#)

fcs\_to\_rds, [6](#)

fcs\_to\_rds\_bkb, [7](#)

imputation\_bkb.predictors, [8](#)

initM, [9](#)

MapfxFFC, [10](#)

MapfxMPC, [12](#)

mkImputeMT, [16](#)

ord.fcs.raw.meta.df.out\_ffc, [17](#)

ord.fcs.raw.meta.df.out\_mpc, [17](#)

ord.fcs.raw.mt\_ffc, [18](#)

ord.fcs.raw.mt\_mpc, [18](#)

rmBatchEffect, [19](#)

rmWellEffect, [19](#)