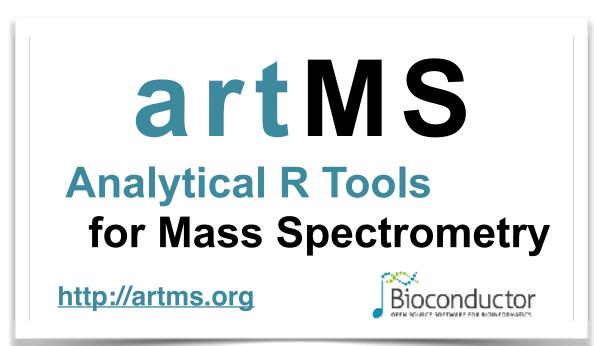


http://artms.org



David Jimenez-Morales, PhD

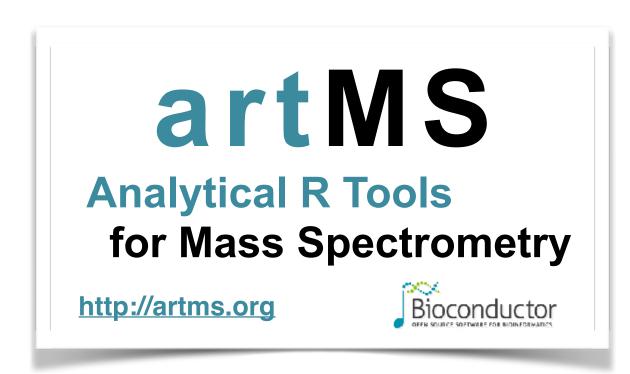




Goals

Facilitate the analysis of Mass-Spectrometry based Proteomics data using the programming language

Provide different levels of analysis to ensure data quality, quantification, and frame results in a biological context.



Quantification Methods

Label Free (+/- Fractionation) SILAC

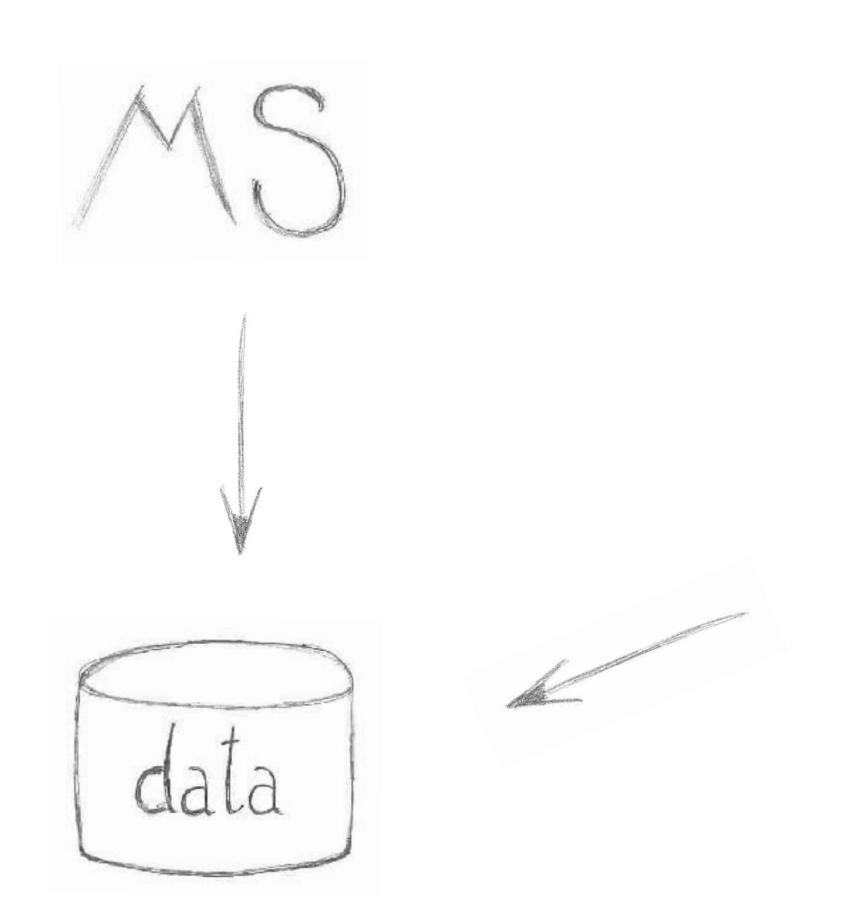
Analysis of

+Global Proteome +PostTranslational Modifications

Phosphorylation
Ubiquitination
Acetylation (coming soon)



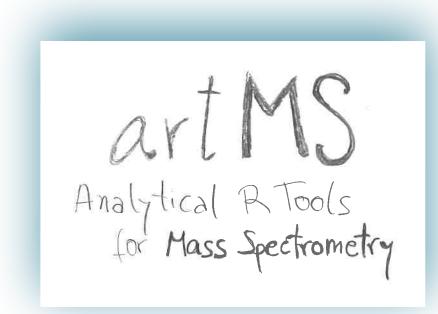
MS-Proteomics experts



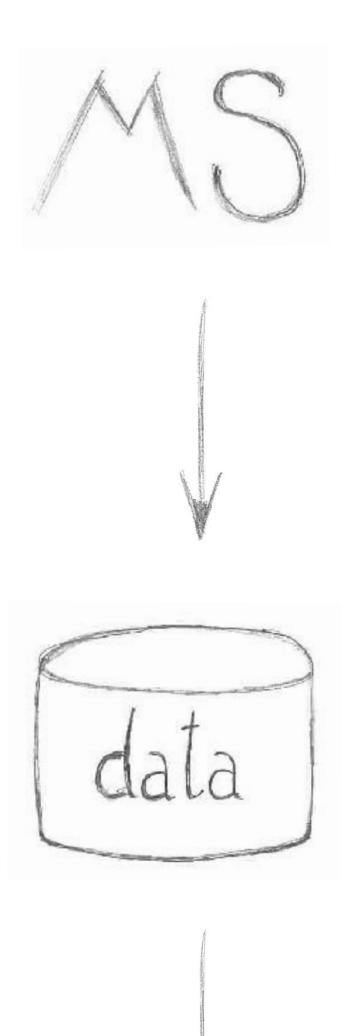
System biologists

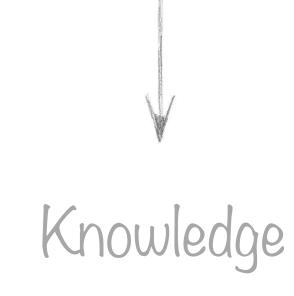
Data scientists

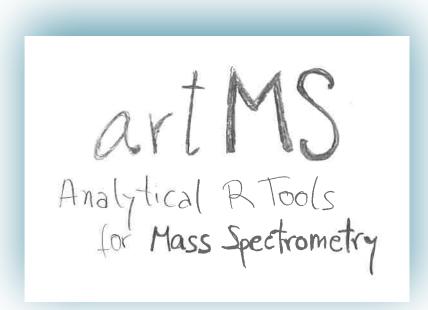




MS-Proteomics experts



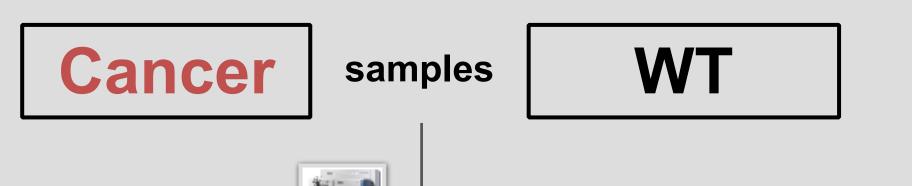






System biologists

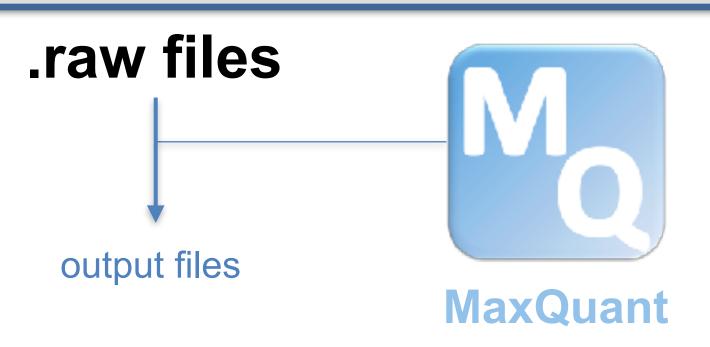
Data scientists



Proteomics pipeline PH, UB, AB, APMS...

Experimental

Computational



Mass Spectrometry

Peptide Identification

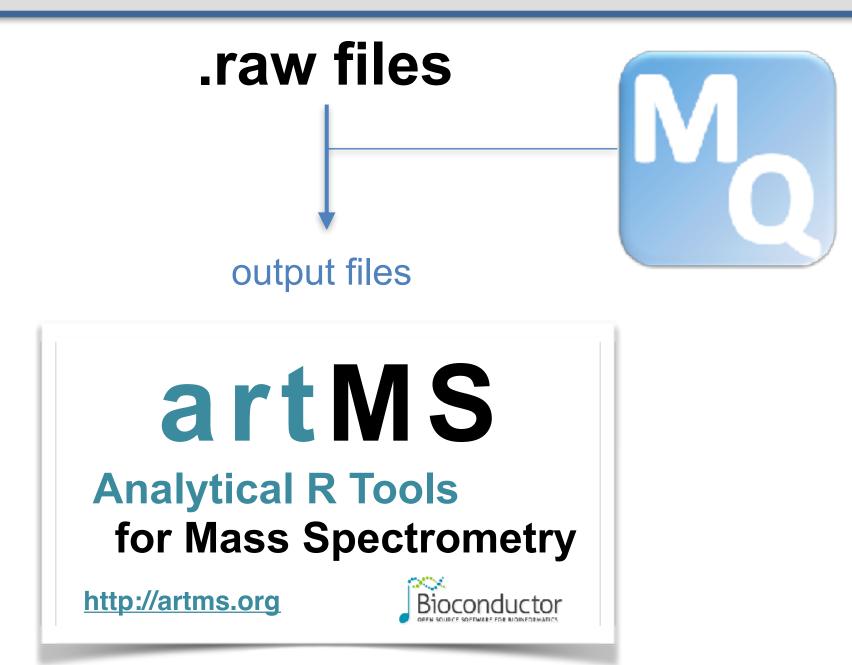
- Tool widely used
- Algorithms yield high mass accuracy and precision
- It covers a wide variety of proteomics experiments
- Simplified user interface
- It can now run on **Linux** (improved performance)



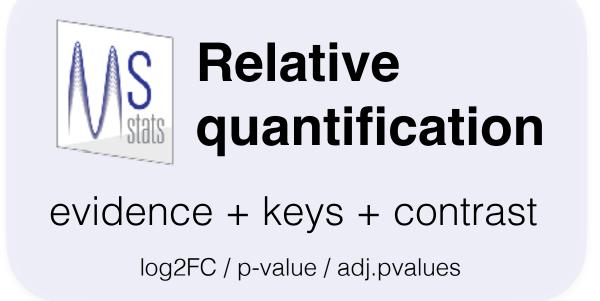
Proteomics pipeline PH, UB, AB, APMS...

Experimental

Computational



Quality control Only Proteins in BioReplicates Unique Proteins in BioRep



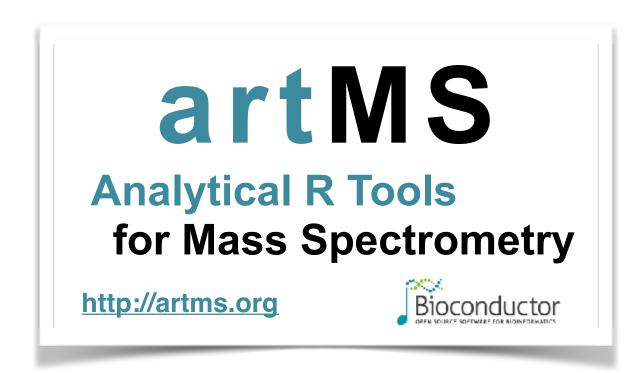
Functional analysis

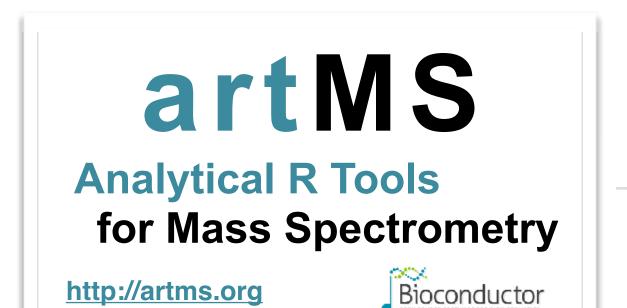
Data integration
Imputation
Enrichment analysis
Pathway analysis
Network generation

Miscellaneous

SAINTq SAINTexpress PHOTON Phosfate

How does it work?





Input Files

Configuration

config.yaml

M evidence.txt

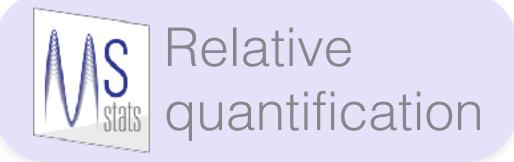
+

keys.txt

contrast.txt

Quality control

artmsQualityControlEvidenceBasic() artmsQualityControlEvidenceExtended() artmsQualityControlSummaryExtended()



artmsQuantification()

Functional analysis

artmsAnalysisQuantifications()

Miscellaneous

artmsEvidenceToSAINTq() artmsEvidenceToSaintExpress() artmsPhosfateOutput() artmsPhotonOutput() etc

Basic Functions



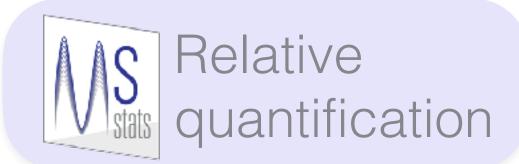
for Mass Spectrometry

http://artms.org



evidence.txt

Quality control



Functional analysis

Miscellaneous

Input Files



+

+

contrast.txt

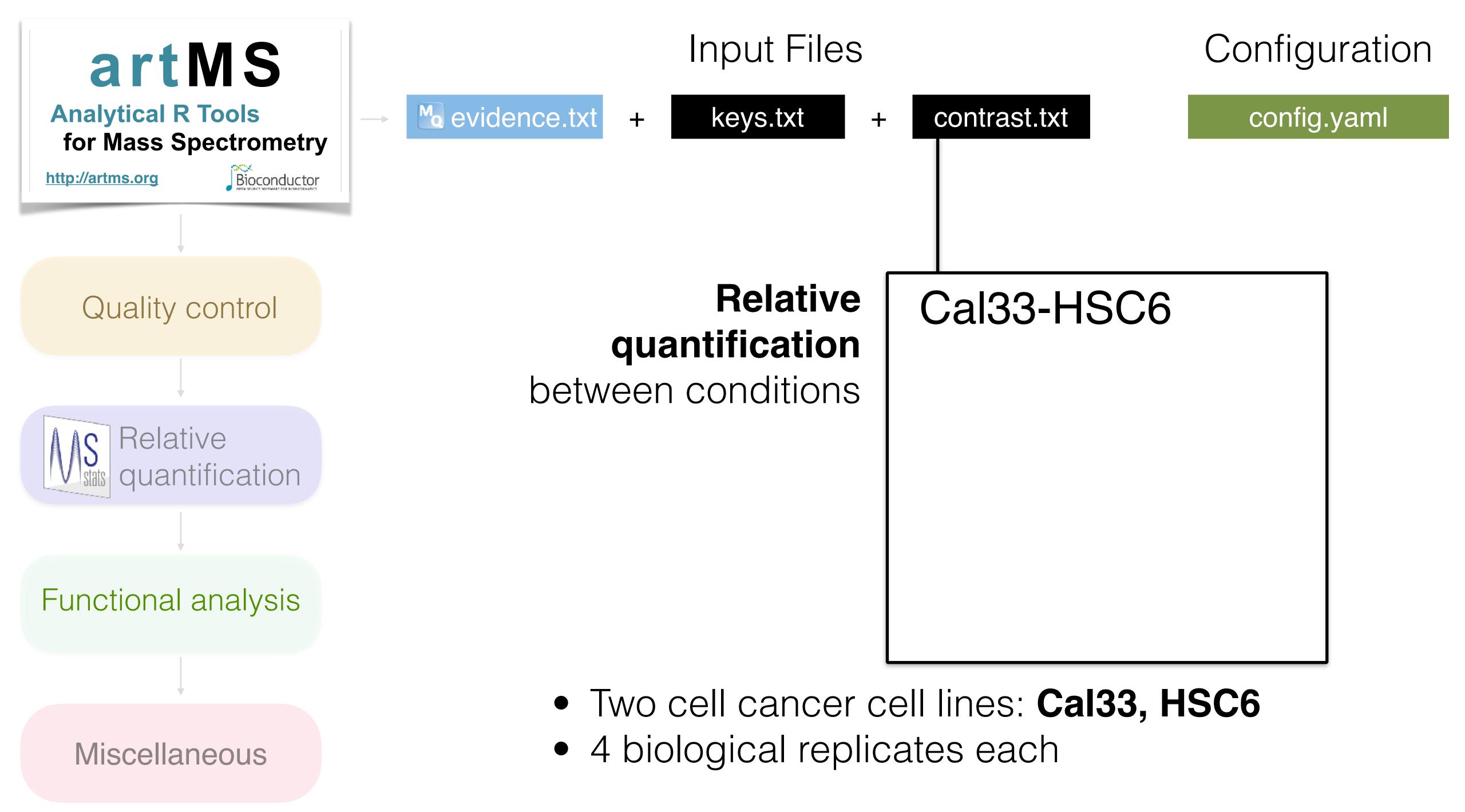
Configuration

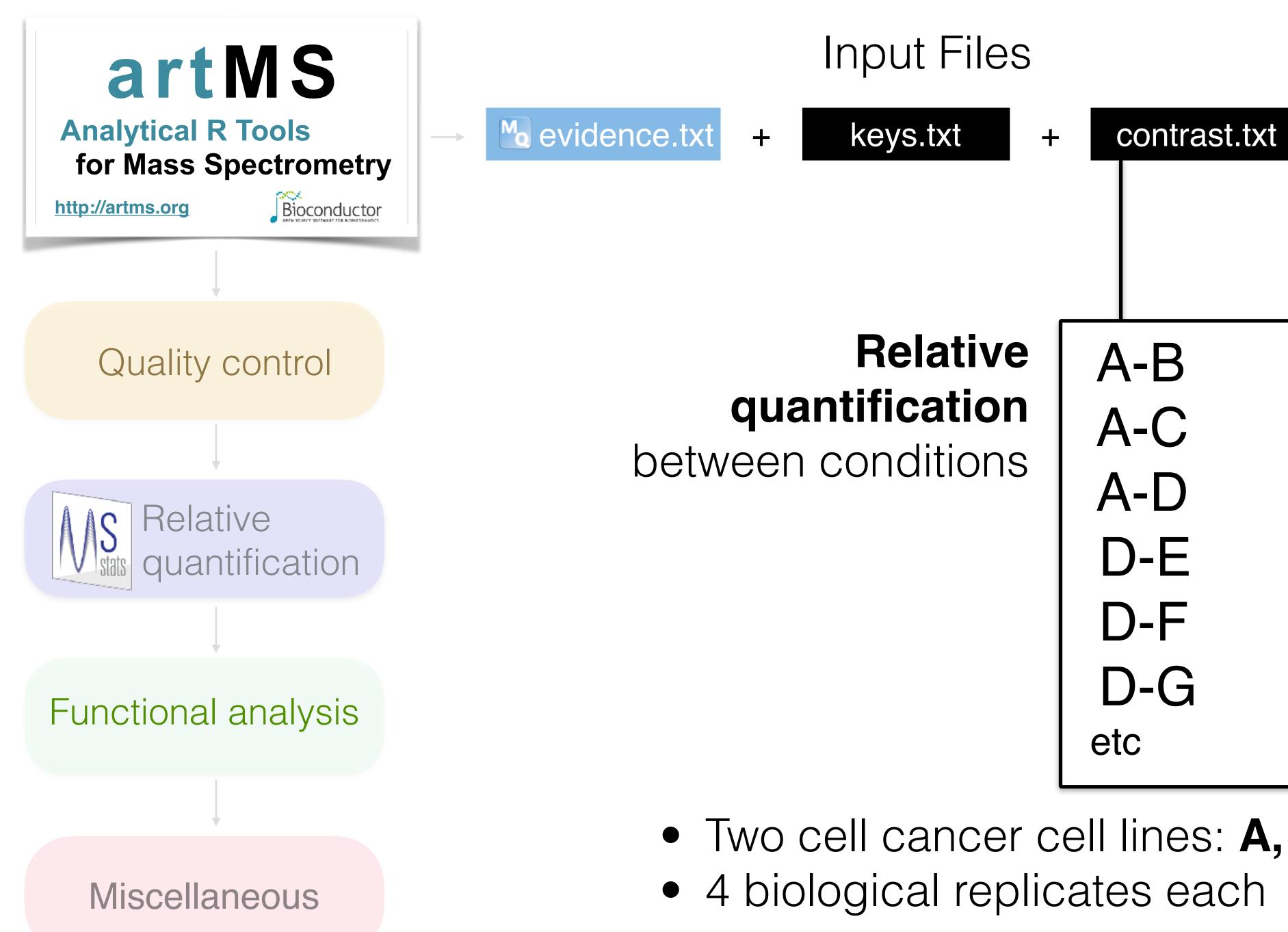
config.yaml

Experimental design

	Α	В	С	D	Е
1	RawFile	IsotopeLabelType	Condition	BioReplicate	Run
2	qx006145	L	Cal33	Cal33-1	1
3	qx006146	L	Cal33	Cal33-2	2
4	qx006147	L	Cal33	Cal33-3	3
5	qx006148	L	Cal33	Cal33-4	4
6	qx006154	L	HSC6	HSC6-1	5
7	qx006151	L	HSC6	HSC6-2	6
8	qx006152	L	HSC6	HSC6-3	7
9	qx006153	L	HSC6	HSC6-4	8
10					

- Two cell cancer cell lines: Cal33, HSC6
- 4 biological replicates each





Configuration config.yaml

- Two cell cancer cell lines: A, B, C, D, E, F, G
- 4 biological replicates each

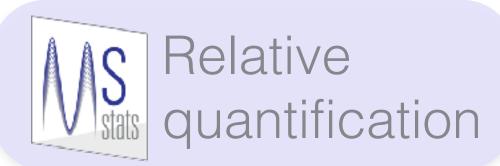
Analytical R Tools

for Mass Spectrometry

http://artms.org



Quality control



Functional analysis

Miscellaneous

```
config.yaml
                                                          Files
        evidence : '~/ph_ms/evidence.txt'
 3
             : '~/ph_ms/keys.txt'
        contrasts : '~/ph_ms/contrast.txt'
        output : '~/ph_ms/phglobal/phglobal-results.txt'
        basic: 1
        extended: 1
       enabled: 1
fractions:
  enabled: 0
10
11
12
13
          enabled: 0
14
15
        filters:
          enabled: 1
16
          contaminants: 1
18
          protein_groups: remove
19
          modifications: PH
20
        sample_plots: 1
      msstats:
   enabled: 1
   msstats_input:
21
23
        profilePlots: none
24
        normalization_method: equalizeMedians
25
26
        normalization_reference:
        summaryMethod: TMP
28
        censoredInt: NA
29
        cutoffCensored: minFeature
     MBimpute: 1
feature_subset: all
output_extras:
enabled: 1
30
31
33
34
         enabled: 1
35
36
          specie: HUMAN
        plots:
          volcano: 1
38
39
          heatmap: 1
40
          LFC: -1.5 1.5
          FDR: 0.05
41
          heatmap_cluster_cols: 0
42
          heatmap_display: log2FC
43
```

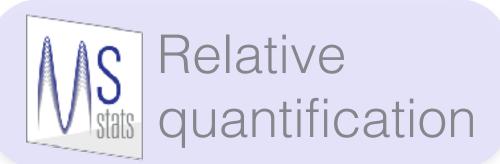
Configuration

Analytical R Tools for Mass Spectrometry

http://artms.org



Quality control



Functional analysis

Miscellaneous

```
config.yaml
      files :
        evidence : '~/ph_ms/evidence.txt'
        keys : '~/ph_ms/keys.txt'
        contrasts : '~/ph_ms/contrast.txt'
        output : '~/ph_ms/phglobal/phglobal-results.txt'
        basic: 1
                                                           QC
       extended: 1
       enabled: 1
fractions:
  enabled: 0
10
11
12
13
          enabled: 0
14
15
        filters:
         enabled: 1
16
          contaminants: 1
18
         protein_groups: remove
19
         modifications: PH
        sample_plots: 1
     msstats:
   enabled: 1
   msstats_input:
21
23
        profilePlots: none
24
        normalization_method: equalizeMedians
25
26
        normalization_reference:
        summaryMethod: TMP
28
        censoredInt: NA
        cutoffCensored: minFeature
29
     MBimpute: 1
feature_subset: all
output_extras:
enabled: 1
30
31
33
34
         enabled: 1
35
36
          specie: HUMAN
        plots:
          volcano: 1
38
39
          heatmap: 1
40
          LFC: -1.5 1.5
          FDR: 0.05
41
          heatmap_cluster_cols: 0
42
          heatmap_display: log2FC
43
```

Configuration

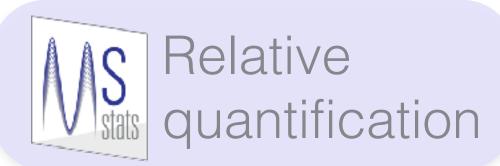
Analytical R Tools

for Mass Spectrometry

http://artms.org



Quality control



Functional analysis

Miscellaneous

```
config.yaml
     files :
       evidence : '~/ph_ms/evidence.txt'
       keys : '~/ph_ms/keys.txt'
       contrasts : '~/ph_ms/contrast.txt'
       output : '~/ph_ms/phglobal/phglobal-results.txt'
       basic: 1
       extended: 1
       ata:
       enabled: 1
fractions:
   enabled: 0
10
                                              Proteomics
11
12
                                                     details
13
         enabled: 0
       filters:
         enabled: 1
17
         contaminants: 1
18
         protein_groups: remove
         modifications: PH
19
20
        sample_plots: 1
21
      nsstats:
       enabled: 1
msstats_input:
23
       profilePlots: none
24
       normalization_method: equalizeMedians
25
26
       normalization_reference:
       summaryMethod: TMP
28
       censoredInt: NA
       cutoffCensored: minFeature
29
     MBimpute: 1
feature_subset: all
output_extras:
enabled: 1
30
31
33
34
         enabled: 1
35
36
         specie: HUMAN
       plots:
         volcano: 1
38
39
         heatmap: 1
40
         LFC: -1.5 1.5
41
         FDR: 0.05
         heatmap_cluster_cols: 0
42
         heatmap_display: log2FC
43
```

Configuration

Analytical R Tools

for Mass Spectrometry

http://artms.org



Quality control



Functional analysis

Miscellaneous

```
config.yaml
      files :
       evidence : '~/ph_ms/evidence.txt'
       keys : '~/ph_ms/keys.txt'
       contrasts : '~/ph_ms/contrast.txt'
       output : '~/ph_ms/phglobal/phglobal-results.txt'
       basic: 1
       extended: 1
       enabled: 1
fractions:
  enabled: 0
10
11
12
13
         enabled: 0
15
        filters:
         enabled: 1
16
17
         contaminants: 1
18
         protein_groups: remove
19
         modifications: PH
        sample plots:
      enabled: 1
msstats_input:
profilePlots: none
21
                            Relative Quantification
22
23
24
                                                       (MSstats)
       normalization_method: equalizeMedians
25
26
       normalization_reference:
27
       summaryMethod: TMP
28
       censoredInt: NA
29
       cutoffCensored: minFeature
30
       MBimpute: 1
31
        feature_subset: all
32
     output_extras:
enabled: 1
33
34
         enabled: 1
35
36
         specie: HUMAN
       plots:
         volcano: 1
38
39
         heatmap: 1
40
         LFC: -1.5 1.5
41
         FDR: 0.05
         heatmap_cluster_cols: 0
42
         heatmap_display: log2FC
43
```

Configuration

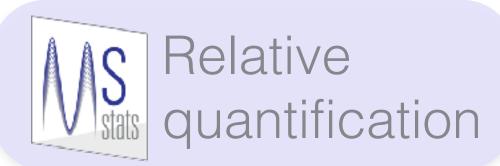
Analytical R Tools

for Mass Spectrometry

http://artms.org



Quality control

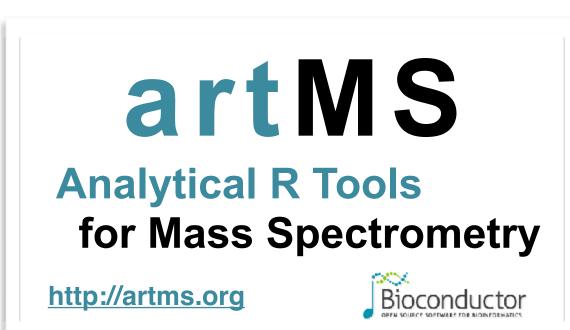


Functional analysis

Miscellaneous

```
config.yaml
      files :
        evidence : '~/ph_ms/evidence.txt'
        keys : '~/ph_ms/keys.txt'
        contrasts : '~/ph_ms/contrast.txt'
        output : '~/ph_ms/phglobal/phglobal-results.txt'
        basic: 1
        extended: 1
       enabled: 1
fractions:
  enabled: 0
10
11
12
13
          enabled: 0
14
15
        filters:
         enabled: 1
16
          contaminants: 1
18
         protein_groups: remove
19
         modifications: PH
        sample_plots: 1
21
      msstats:
   enabled: 1
   msstats_input:
23
        profilePlots: none
24
        normalization_method: equalizeMedians
25
26
        normalization_reference:
        summaryMethod: TMP
28
        censoredInt: NA
        cutoffCensored: minFeature
29
30
        MBimpute: 1
         <u>feature subset: all</u>
31
        itput_extras:
enabled: 1
33
                                        Outputs / extras
34
         enabled: 1
35
          specie: HUMAN
        plots:
38
          volcano: 1
39
          heatmap: 1
40
          LFC: -1.5 1.5
41
          FDR: 0.05
          heatmap_cluster_cols: 0 heatmap_display: log2FC
42
43
```

Configuration



Input Files

Me evidence.txt

keys.txt

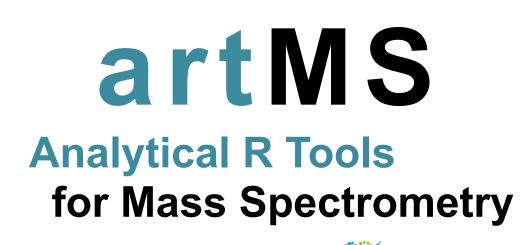
Quality control

artmsQualityControlEvidenceBasic()
artmsQualityControlEvidenceExtended()
artmsQualityControlSummaryExtended()



Functional analysis

Miscellaneous



Input Files

Me evidence.txt

keys.txt

http://artms.org



Quality control

artmsQualityControlEvidenceBasic()

artmsQualityControlEvidenceExtended()
artmsQualityControlSummaryExtended()

Relative quantification

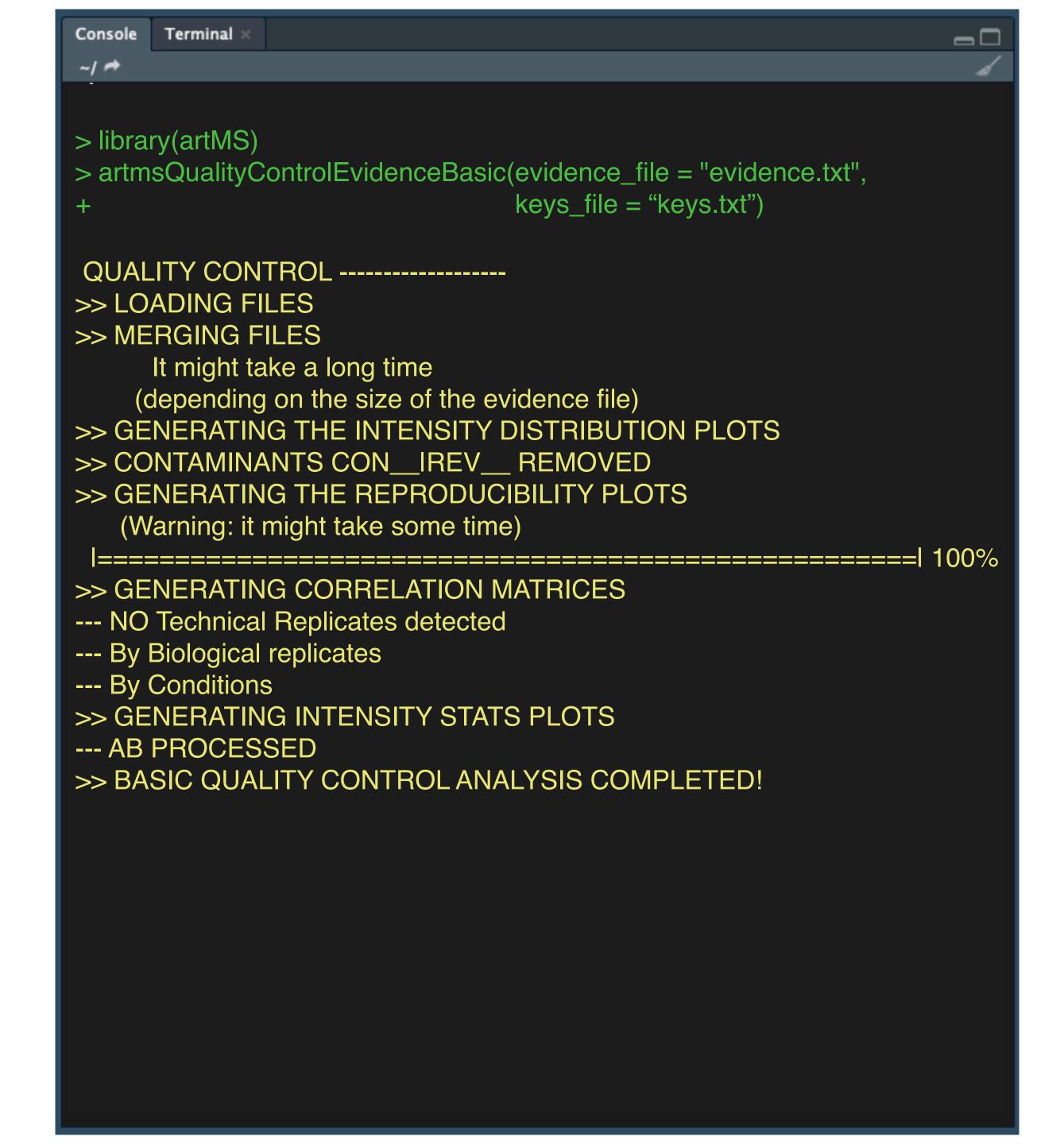
rtmsQuantification()

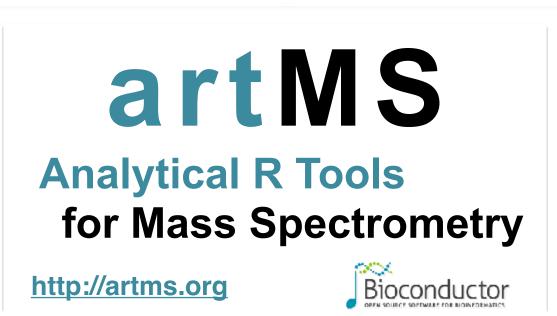
Functional analysis

artmsAnalysisQuantifications()

Miscellaneous

artmsEvidenceToSAINTq()
artmsEvidenceToSaintExpress()
artmsPhosfateOutput()
artmsPhotonOutput()





Input Files

evidence.txt

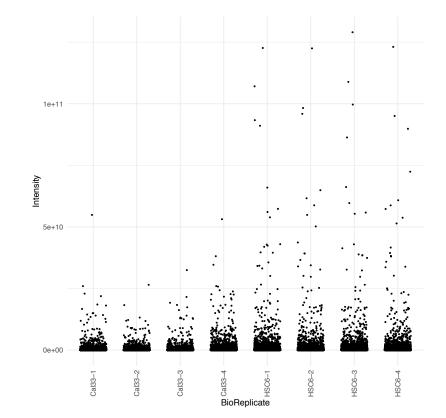
keys.txt

Quality control

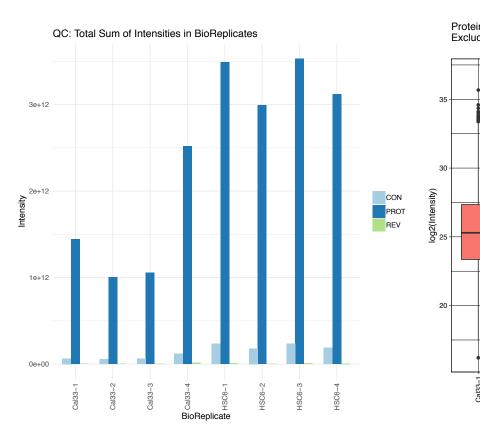
Relative quantification

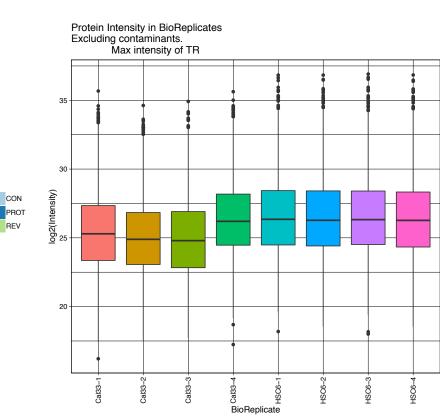
artmsQualityControlEvidenceBasic()

- Reproducibly
- MS Intensity
- Spectral Counts
- Contaminants
- PTMs



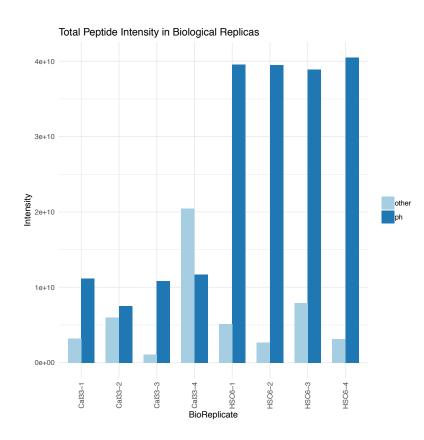
Peptide Reproducibility between Bioreplicas (condition: Cal33) Cal33–1 vs Cal33–2

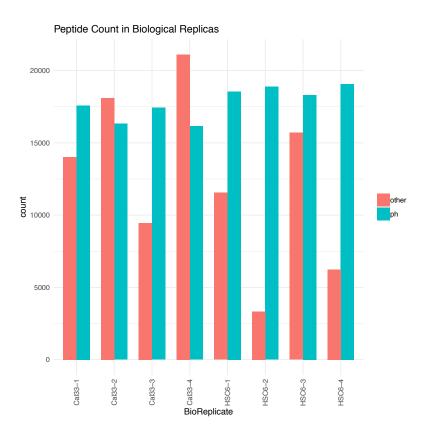




Functional analysis

Miscellaneous







evidence.txt

keys.txt

Input Files

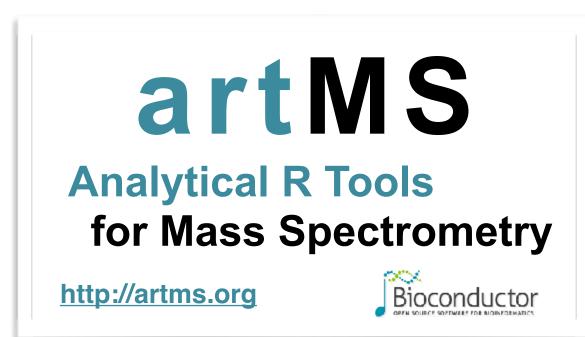
http://artms.org



Quality control

artmsQualityControlEvidenceExtended()

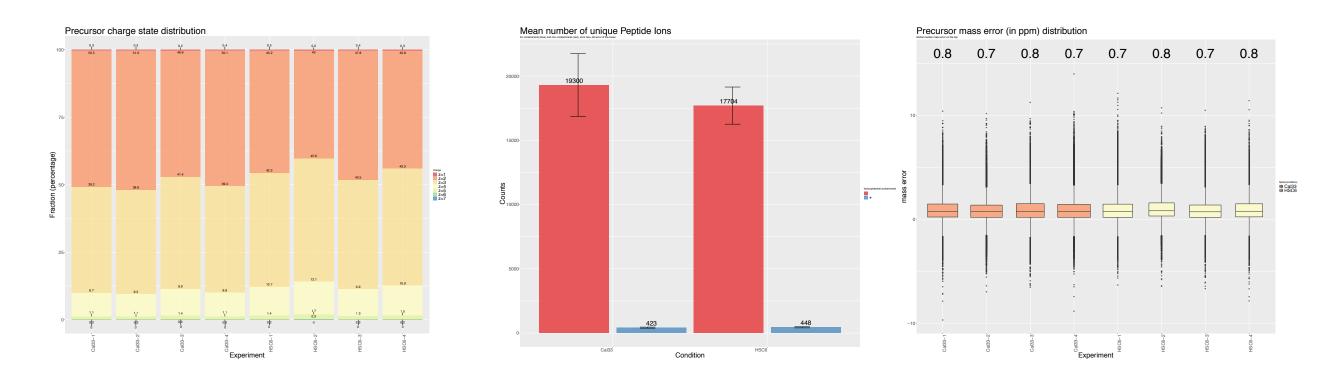
```
Terminal
Console
~/ #
> artmsQualityControlEvidenceExtended(evidence_file = "evidence.txt",
                                       keys_file = "keys.txt")
       >>EXTENDED QUALITY CONTROL ANALYSIS
       (evidence.txt based)-----
       >> MERGING FILES
             It might take a long time
            (depending on the size of the evidence file)
       >> GENERATING QC PLOTS
       --- Plot PSM done
       --- Plot IONS done
       --- Plot TYPE done
       --- Plot PEPTIDES done
       --- Plot PROTEINS done
       --- Plot Plot Ion Oversampling done
       --- Plot Charge State done
       --- Plot Mass Error done
       --- Plot Mass-over-Charge distribution done
       --- Plot Peptide Intensity CV done
       --- Plot Peptide Detection (using modified.sequence) done
       --- Plot Protein Intensity CV done
       --- Plot Protein Detection done
       --- Plot ID overlap done
       --- Plot Inter-Correlation done
       --- Plot Sample Preparation done
```



Input Files

evidence.txt

keys.txt



Quality control

Relative quantification

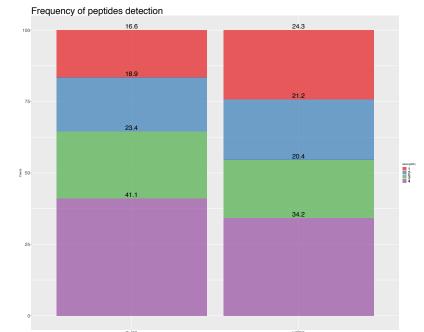
Functional analysis

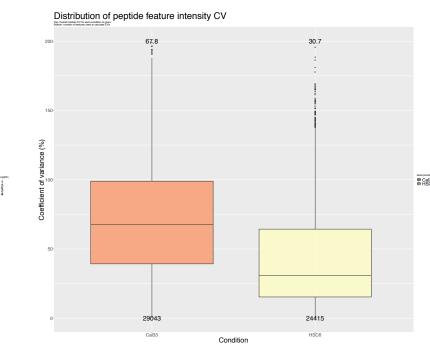
Miscellaneous

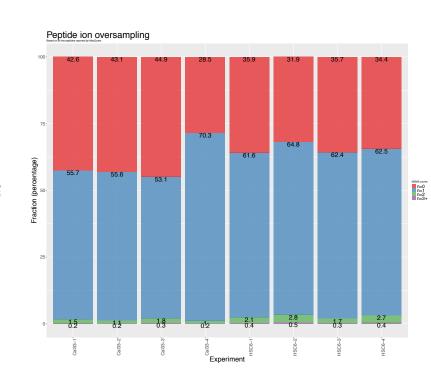
artmsQualityControlEvidenceBasic(

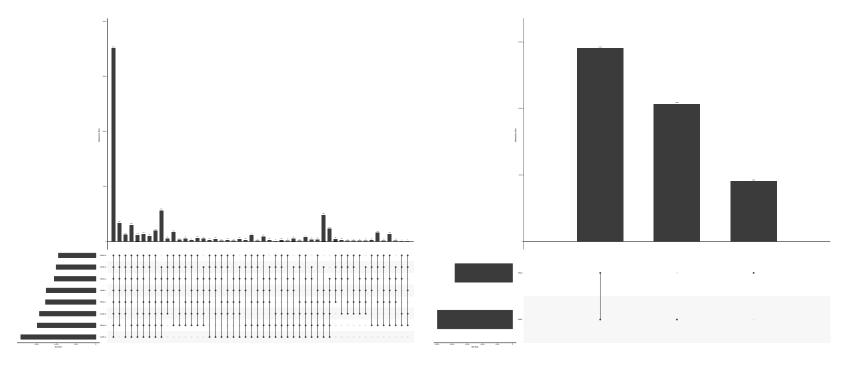
artmsQualityControlEvidenceExtended()

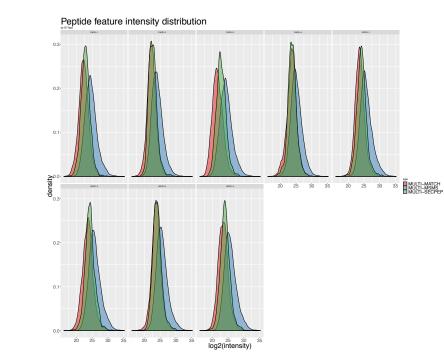
- Charge State Distribution
- ID overlap
- Peptide Ion Statistics
- Precursor Mass Error
- Precursor m/z error
- Frequency Peptide Detection
- Peptide Intensity CV
- Peptide Statistics
- Peptide Ion oversampling
- Protein detection frequency
- Protein Statistics / CV
- Peptide-spectrum matches
- MaxQuant Type statistics





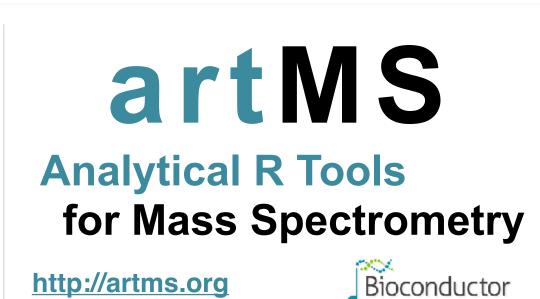






artmsPhotonOutput(

ot/



Input Files

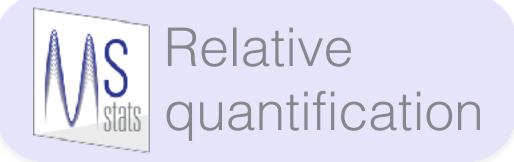
evidence.txt

keys.txt



Quality control

artmsQualityControlEvidenceBasic() artmsQualityControlEvidenceExtended() artmsQualityControlSummaryExtended()



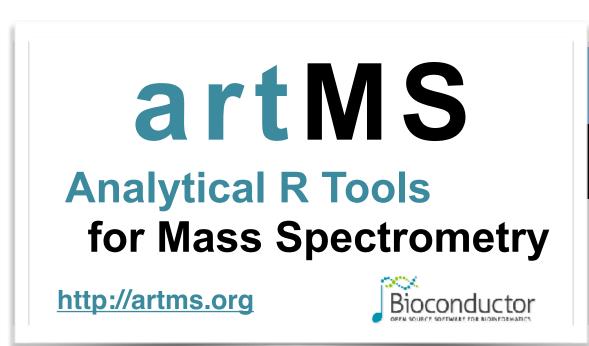
artmsQuantification()

Functional analysis

artmsAnalysisQuantifications()

Miscellaneous

artmsEvidenceToSAINTq() artmsEvidenceToSaintExpress() artmsPhosfateOutput() artmsPhotonOutput() etc



Input Files

evidence.txt

keys.txt

Quality control

artmsQualityControlEvidenceBasic()
artmsQualityControlEvidenceExtended()
artmsQualityControlSummaryExtended()



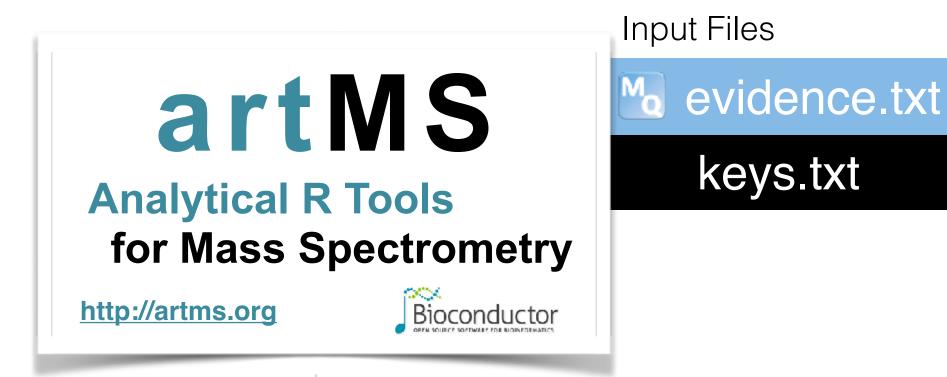
artmsQuantification()

Functional analysis

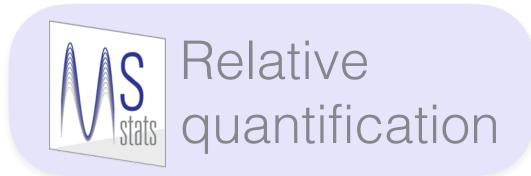
artmsAnalysisQuantifications()

Miscellaneous

artmsEvidenceToSAINTq() artmsEvidenceToSaintExpress() artmsPhosfateOutput() artmsPhotonOutput()



Quality control



artmsQuantification()

Functional analysis

Miscellaneous



MSstats is an open-source R package for statistical relative quantification of proteins and peptides in

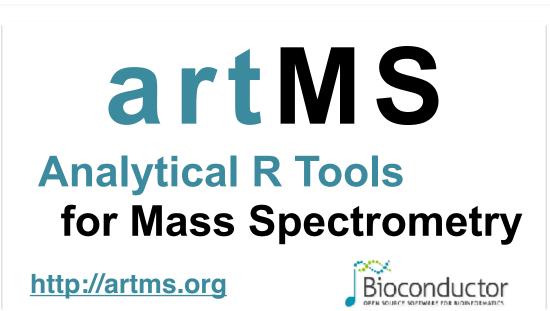
It uses a family of linear mixed models that attempts to:

global, targeted, and data-independent proteomics.

- minimize bias and inefficiencies in spectrometry-based proteomics,
- distinguish the systematic variation from random artifacts,
- maximize the reproducibility of the results

Developed by Meena Choi and Olga Vitek

Used in Skyline



™ e\

Input Files

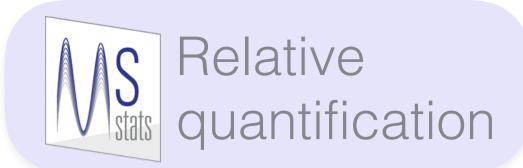
evidence.txt

keys.txt

contrast.txt

config.yaml

Quality control



artmsQuantification("config.yaml") -----

TXT FILES

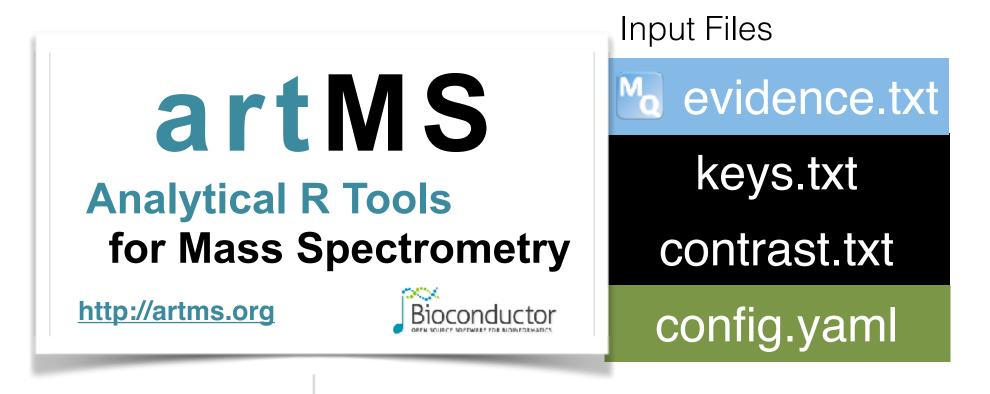
- quantification-results.txt
- quantification-results-annotated.txt
- Normalized abundance
 - * results_ModelQC.txt
 - * results_RunlevelData.txt
 - * results-mss-groupQuant.txt
 - * results-mss-normalized.txt
 - * results-mss-sampleQuant.txt
- results_sampleSize.txt
- results_experimentPower.txt

-unctional analysis

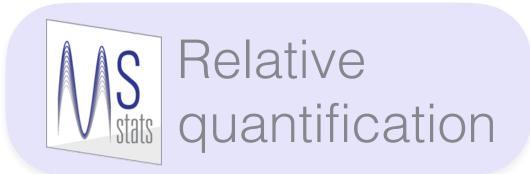
PDF FILES

- results-heatmap.pdf
- results-peptidecounts-perBait.pdf
- results-peptidecounts.pdf
- results-sign.pdf
- results-volcano.pdf

Miscellaneous



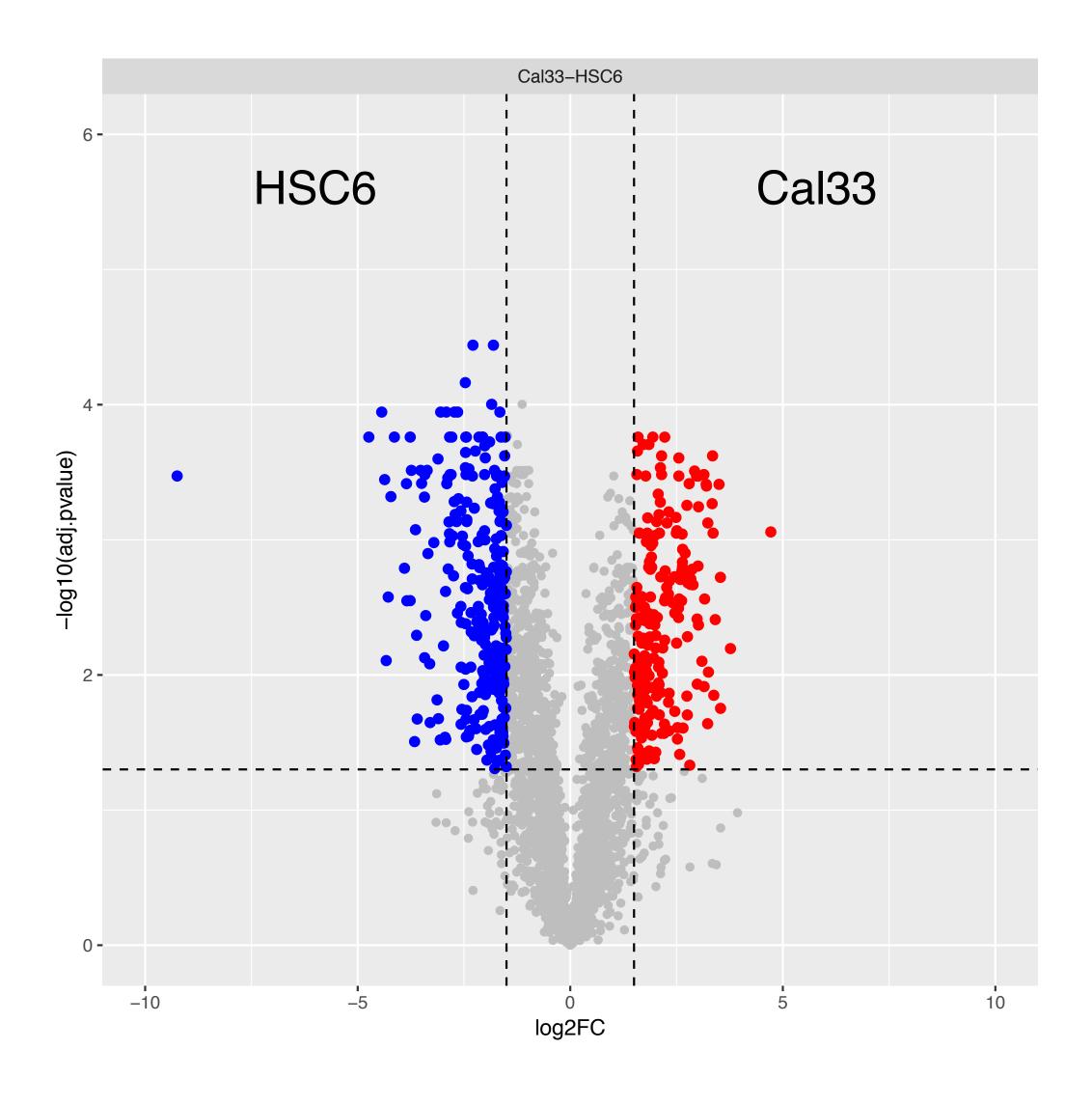
Quality control

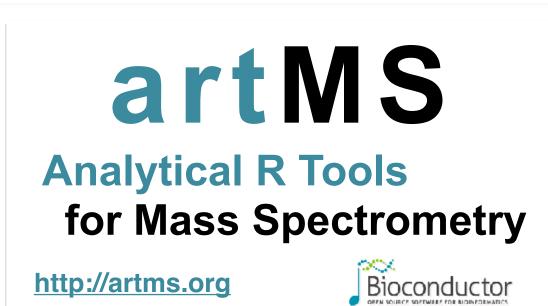


artmsQuantification("config.yaml")

Functional analysis

Miscellaneous





evidence.txt

keys.txt

contrast.txt

config.yaml

- Annotations
- Summary files in different format (xls, txt) and shapes
- Numerous summary plots
- Enrichment analysis using Gprofiler
- PCA of protein abundance
- PCA of quantifications
- Clustering analysis

Functional analysis

artmsAnalysisQuantifications()

Miscellaneous

artmsEvidenceToSAINTq() artmsEvidenceToSaintExpress() artmsPhosfateOutput() artmsPhotonOutput() etc

EXCEL

PDF FILES

-results-summary.xlsx

-results-abundance-wide.txt

-results-abundance-long.txt

-results-enrich-MAC-allsignificants-corum.txt

-results-enrich-MAC-allsignificants.txt

-results-enrich-MAC-negatives-corum.txt

TXT FILES -results-enrich-MAC-negatives.txt

-results-enrich-MAC-positives-corum.txt

-results-enrich-MAC-positives.txt

-results-log2fc-long.txt

-results-log2fc-wide.txt

-results.log2fc-clusterheatmap-enriched.txt

-results.log2fc-clusterheatmap.txt

-results-enrich-MAC-allsignificants-corum.pdf

-results-enrich-MAC-negatives-corum.pdf

-results-enrich-MAC-positives-corum.pdf

-results-pca-correlations.pdf

-results-pca-pca01.pdf

-results-pca-pca02.pdf

-results-pca-pca03.pdf

-results.clustering.abundance.all-overview.pdf

-results.clustering.abundance.all-zoom.pdf

-results.clustering.log2fc.all-overview.pdf

-results.clustering.log2fc.all-zoom.pdf

-results.clustering.log2fcSign.all-overview.pdf

-results.clustering.log2fcSign.all-zoom.pdf

-results.distributions.pdf

-results.distributionsFil.pdf

-results.imputation.pdf

-results.log2fc-clusterheatmap.pdf

-results.log2fc-clusters.pdf

-results.log2fc-corr.pdf

-results.log2fc-dendro.pdf

-results.log2fc-individuals-pca.pdf

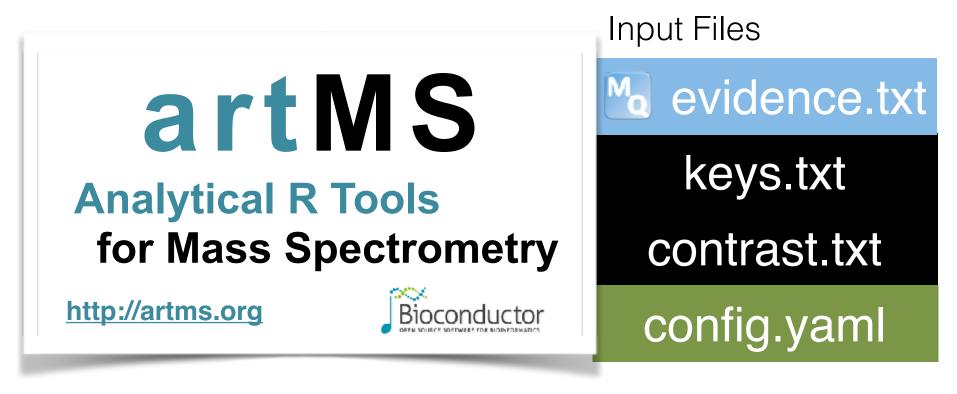
-results.TotalQuantifications.pdf

-results.correlationConditions.pdf

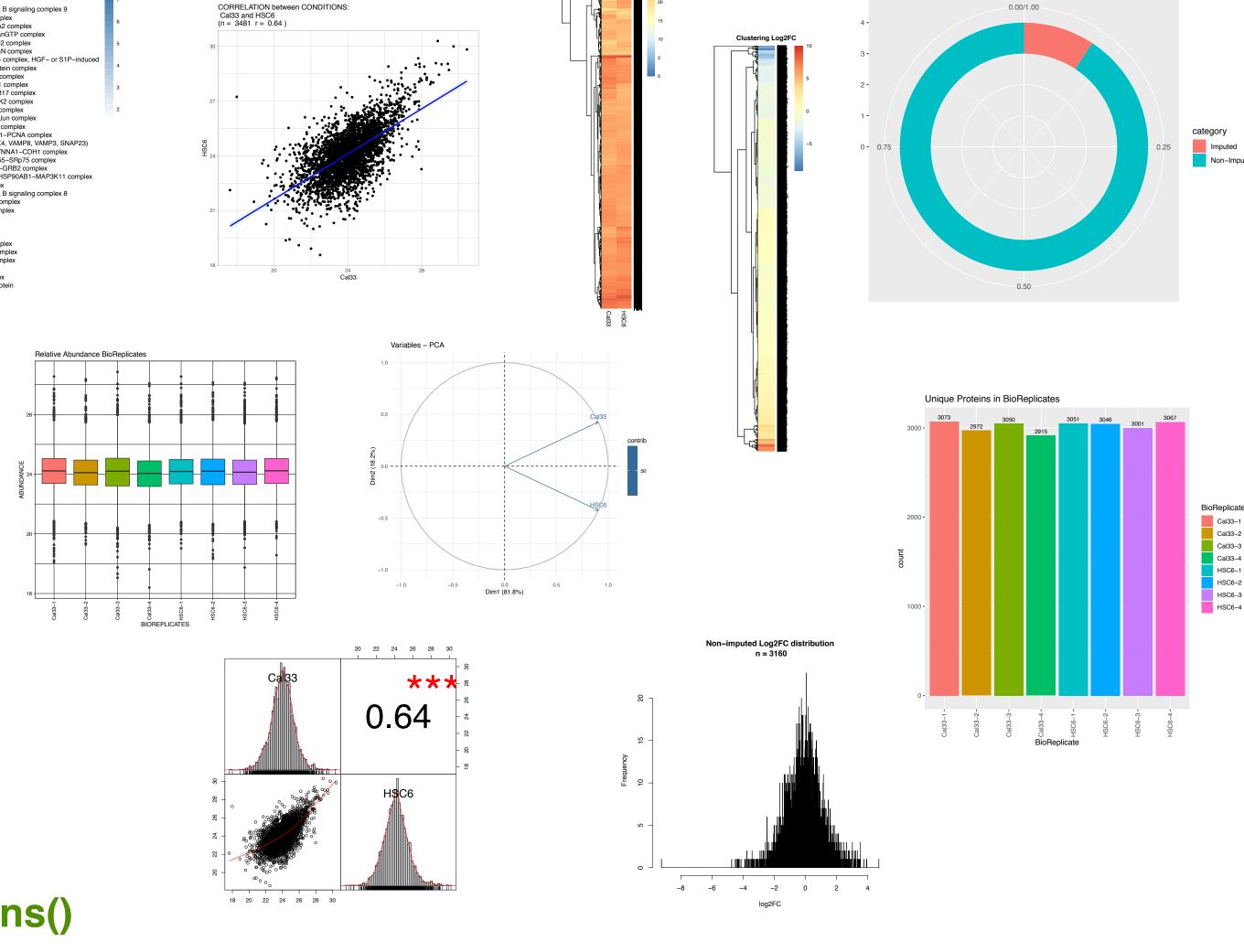
-results.correlationQuantifications.pdf

-results.relativeABUNDANCE.pdf

-results.reproducibilityAbundance.pdf



- Annotations
- Summary files in different format (xls, txt) and shapes
- Numerous summary plots
- Enrichment analysis using Gprofiler
- PCA of protein abundance
- PCA of quantifications
- Clustering analysis



Functional analysis

artmsAnalysisQuantifications()

log2fcImputed

log2fclmpExt

wide_iLog2fc

wide_iPvalue enrichALL

-1.643170638 0.020905156 0.063210769 no. -1.643170638 0.063210769 F2 Etranscription factor 8 4 HSC6 4 A0AVK6 E2F transcription factor 8 79733 0.692623032 0.740104187 0.809434278 no 0.692623032 0.809434278 human extended synaptotagmin 2 57488 ADFGR8_S676 -2.4675906 0.352698385 0.468321968 no -2.4675906 0.468321968 5 A0FGR8 Cal33-HSC6 A0FGR8 2 HSC6 human extended synaptotagmin 2 57488 5 A0FGR8 ADFGR8 5691 -0.198128315 0.405278123 0.518470215 no -0.198128315 0.518470215 3 HSC6 human extended synaptotagmin 2 57488 ADEGR8 S691;A0EGR8 S693 Cal33-HSC6 3.66610245 0.0198395 human 7 A0FGR8 0 Cal33 A0FGR8 extended synaptotagmin 2 57488 ADFGR8 5693 Cal33-HSC6 0.950304175 0.15494643 0.254530287 no 0.950304175 0.254530287 4 Cal33 A0FGR8 8 A0FGR8 human extended synaptotagmin 2 57488 ADFGR8 5736 Cal33-HSC6 -0.019383937 0.932856234 0.952326477 no -0.019383937 0.952326477 9 A0FGR8 3 HSC6 A0FGR8 human extended synaptotagmin 2 57488 ADFGR8 5743 Cal33-HSC6 4 HSC6 A0FGR8 10 A0FGR8 ESYT2 -7.793852072 0.0244517 human extended synaptotagmin 2 57488 11 A0FGR8 ADEGR8 \$758;A0EGR8 \$761 Cal33-HSC6 -0.388729606 0.906408486 0.934755075 no -0.388729606 0.934755075 3 HSC6 A0FGR8 human extended synaptotagmin 2 57488 -0.710786851 0.002631099 0.016113125 no 12 A0FGR8 ADFGR8 5761 Cal33-HSC6 -0.710786851 0.016113125 4 HSC6 A0FGR8 human extended synaptotagmin 2 57488 A0FGR8_T701;A0FGR8_T705 Cal33-HS06 -1.205399144 0.017277448 0.055442631 no -1.205399144 0.055442631 13 A0FGR8 4 HSC6 ADEGR8 human mediator complex subunit 1 219541 -0.628616166 0.213593795 0.32335184 no 14 A0JLT2 ADJLT2_S226 -0.628616166 0.32335184 4 HSC6 A0JLT2 human

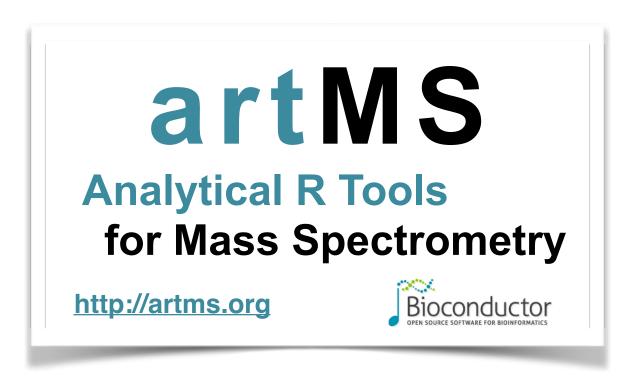
enrichMACpos

enMACallCorum

enMACpcsCorum

enMACnegCorum

Miscellaneous



available at:



https://bioconductor.org/packages/release/bioc/html/artMS.html



https://github.com/biodavidjm/artMS



http://artms.org



artMS: Analytical R Tools for Mass Spectrometry-based Proteomics

David Jimenez-Morales^{1*}, Alexandre Rosa Campos², John Von Dollen³, Matthew T Wheeler¹, Euan A Ashley¹, Nevan J. Krogan³, Danielle L. Swaney^{3*}

In preparation

1. Stanford

2. SBP3. UCSF

acknowledgements





Nevan Krogan
Danielle Swaney
Jeff Johnson
Ruth Huttenhain
John Von Dollen



Alexandre Rosa



Euan Ashley Matthew Wheeler



Alex Pico

