

Supplemental Data

ImShot: An open-source software for probabilistic identification of proteins *in situ* and visualization of proteomics data

Wasim Aftab^{1,2}, Shibojyoti Lahiri^{1*} and Axel Imhof^{1*}

¹Biomedical Center, Protein Analysis Unit, Faculty of Medicine, Ludwig-Maximilians-Universität München, Großhaderner Strasse 9, 82152 Planegg-Martinsried, Germany.

²Graduate School for Quantitative Biosciences (QBM), Ludwig-Maximilians-Universität Munich, 81377 Munich, Germany.

* Corresponding authors

shibojyoti.lahiri@med.uni-muenchen.de

imhof@lmu.de

Table of Contents

1. Imputation from normal distribution (tiny)	3
2. Tolerance for deisotoping and MLP based filtering	3
3. ImShot log file.....	4
4. Guidelines for running ImShot on Windows 10 pc	4
4.1 Setting R environment	4
4.2 How to run ImShot desktop application	5
4.3 Testing ImShot desktop app.....	9
4.4 How to install ImShot R package.....	10
4.5 Testing ImShot R package.....	10
5. Usage guidelines for Linux desktop.....	11
5.1 Setting R environment	11
5.2 How to run ImShot desktop application	12
5.3 Testing ImShot desktop app.....	13
5.4 How to Install ImShot R package	13
5.5 Testing ImShot R package.....	13
6. Usage guidelines for MacOS.....	14
6.1 Setting R environment	14
6.2 How to run ImShot desktop application	14
6.3 Testing ImShot desktop app.....	15
6.4 How to Install ImShot R package	15
6.5 Testing ImShot R package.....	15
7. Imputation algorithms in ImShot	15
8. Description of output file headers after GO/Pathway Enrichment	17
9. Important Notes	18
Reference.....	19

1. Imputation from normal distribution (tiny)

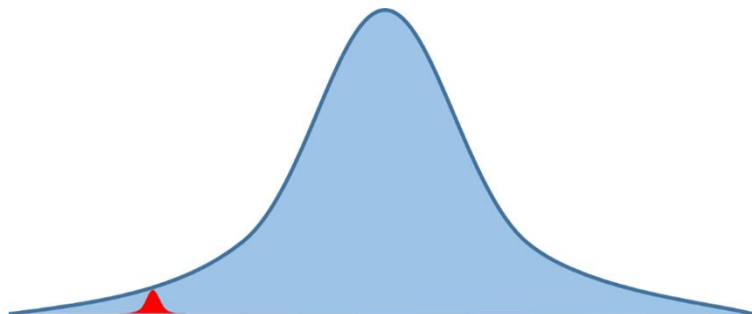
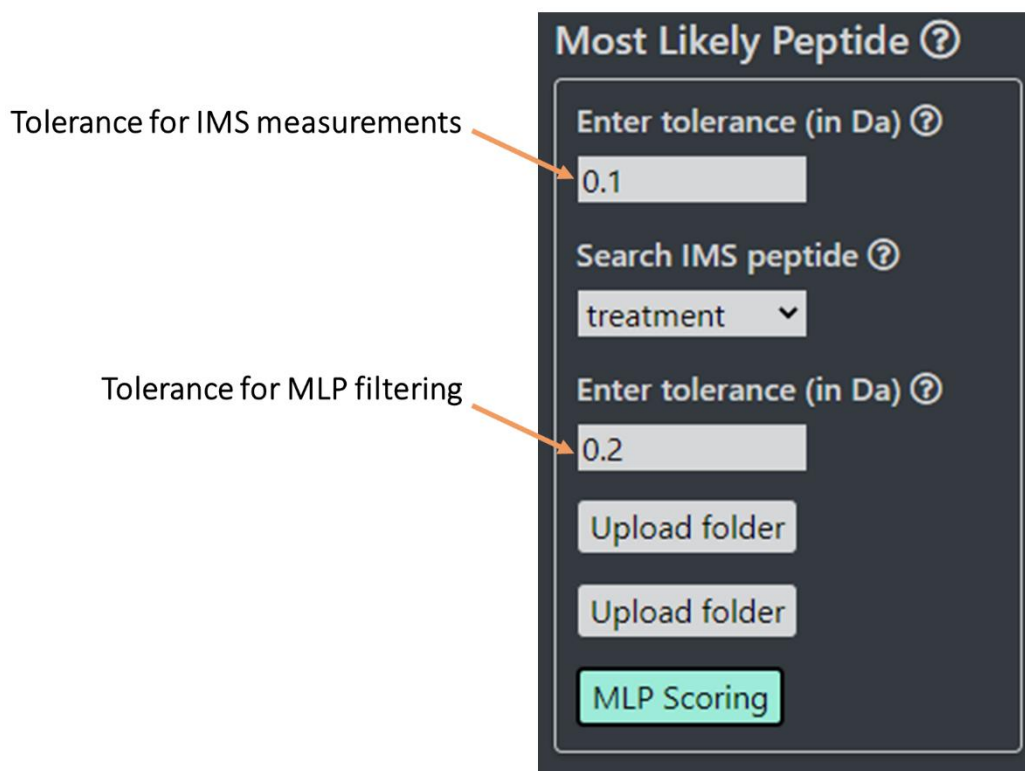


Figure S1: Missing value imputation algorithm (*tiny*) - Original data distribution (after filtering and \log_2 transformation) is shown in blue and the tiny normal distribution (obtained by shifting and shrinking the original distribution) from where missing values are drawn randomly is shown in red.

2. Tolerance for deisotoping and MLP based filtering



Most Likely Peptide ?

Enter tolerance (in Da) ?

0.1

Search IMS peptide ?

treatment ▼

Enter tolerance (in Da) ?

0.2

Upload folder

Upload folder

MLP Scoring

Figure S2: ImShot offers flexibility in choosing search tolerance – An excerpt from ImShot GUI to highlight the flexibility to search LC-MS peptides within user specified tolerance of an IMS peptide. Tolerance for MLP search is increased slightly to ensure the boundary is achieved. This feature is also provided in ImShot R package.

3. ImShot log file

```
[2021-03-21T15:24:06.532] [INFO] fileLog - ##### Rscript path setup #####  
[2021-03-21T15:24:06.622] [FATAL] fileLog - Rscript_path.json is NOT found, you must set Rscript path before continuing  
[2021-03-21T15:24:06.623] [INFO] fileLog - Running in production  
[2021-03-21T15:24:06.624] [INFO] fileLog - user_data_dir = C:\Users\W.Aftab\AppData\Roaming\ImShot_Electron_App  
[2021-03-21T15:24:42.073] [INFO] fileLog - RscriptPath = C:/Program Files/R/R-4.0.2/bin/x64  
[2021-03-21T15:24:42.087] [INFO] fileLog - Absolute path to RSCRIPT = C:/Program Files/R/R-4.0.2/bin/x64/Rscript  
[2021-03-21T15:25:20.882] [INFO] fileLog - ##### LIMMA analysis #####  
[2021-03-21T15:25:20.987] [INFO] fileLog - treatment_rep_names=Intensity AROM_Rep1,Intensity AROM_Rep2,Intensity AROM_Rep3  
[2021-03-21T15:25:20.987] [INFO] fileLog - control_rep_names=Intensity WT_Rep1,Intensity WT_Rep2,Intensity WT_Rep3  
[2021-03-21T15:25:20.987] [INFO] fileLog - remove_exclusive=YES  
[2021-03-21T15:25:20.988] [INFO] fileLog - fc_cutoff=1.5  
[2021-03-21T15:25:20.989] [INFO] fileLog - median_norm=NO  
[2021-03-21T15:25:20.989] [INFO] fileLog - k_out_N=1  
[2021-03-21T15:25:20.990] [INFO] fileLog - seq_cov=1  
[2021-03-21T15:25:20.991] [INFO] fileLog - uniq_pep=1  
[2021-03-21T15:25:20.991] [INFO] fileLog - fasta_db_val=Uniprot  
[2021-03-21T15:25:20.992] [INFO] fileLog - regex_gene_val=  
[2021-03-21T15:25:20.992] [INFO] fileLog - protein_or_peptide=peptides  
[2021-03-21T15:25:33.830] [INFO] fileLog - R_version = 4.0.2
```

Figure S3: ImShot log file – ImShot log captures the user inputs (in grey rectangle), error events (in red rectangle) and R version information (in green rectangle) which aid in debugging and reproducibility of the data analysis. This feature is only supported in the GUI version of ImShot.

4. Guidelines for running ImShot on Windows 10 pc

4.1 Setting R environment

Existing R users may skip the remainder of this section and proceed directly to Section 4.2, as the R environment may have already been configured. Nevertheless, it is important to note that we tested this guideline on R version 4.0.2 on 64-bit Windows 10 and recommend using R version $\geq 4.0.2$.

Step 1 - Download 64-bit R, RStudio and install

Step 2 - Install Rtools from <https://cran.r-project.org/bin/windows/Rtools/>

Step 3 - Run the following command in RStudio console

```
writeLines('PATH="%${RTOOLS40_HOME}\\usr\\bin;%${PATH}"', con = "~/Renvirom")
```

Step 4 - Restart RStudio and type the following command in the console

```
Sys.which("make")
```

Press enter

You should see output like the following highlighted lines

make

"C:\\rtools40\\usr\\bin\\make.exe"

Step 5 - If step 4 works, install an R package by typing the following command in R console and pressing enter

```
install.packages("jsonlite", type = "source")
```

If this succeeds, you're good to go for the next steps!

4.2 How to run ImShot desktop application

Step 1 – Download/clone the repository containing ImShot software (both R package and desktop app), datasets, and video tutorials from the following link,

<https://github.com/wasimaftab/ImShot>

In addition to the README file which contains information for the user about the software, the link contains eight folders and one file as shown in figure S4.

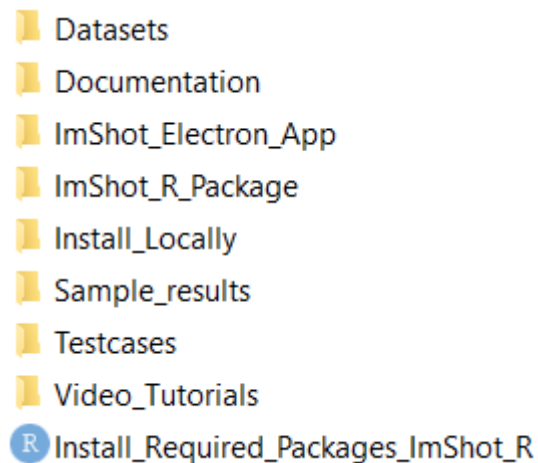


Figure S4: ImShot software shared folder structure

The **Install_Required_Packages_ImShot_R.R** file contains code to install the R packages required for successful execution of ImShot software. This script will install packages from CRAN/Bioconductor repositories as well as local packages.

The **ImShot_R_Package** folder contains the ImShot R package and because it is not part of CRAN/Bioconductor repositories therefore will be installed locally.

The **Install_Locally** folder contains the two supporting R packages that are also not part of CRAN/Bioconductor repositories and hence will be installed locally.

The **ImShot_Electron_App** folder contains three zipped files corresponding to the ImShot desktop applications for Windows, Linux, and Mac respectively.

The **Datasets** folder contains test datasets to test ImShot desktop app and R package.

The **Documentation** folder contains the present document in pdf format.

The **Testcases** folder contains various testcases to test the ImShot software.

The **Sample_results** folder contains the results after running the testcases. To give users an overview of the output data structures.

The **Video_Tutorials** folder contains videos to inform the users about the input, output and functioning of ImShot desktop app and the R package.

Step 2 - Open **Install_Required_Packages_ImShot_R.R** (downloaded earlier) in RStudio and click the **Source** button (near the top right corner of the Editor pane).

```
These packages have more recent versions available.  
It is recommended to update all of them.  
Which would you like to update?  
  
1: All  
2: CRAN packages only  
3: None  
4: TTR (0.24.2 -> 0.24.3) [CRAN]  
  
Enter one or more numbers, or an empty line to skip updates: 1
```

(A)

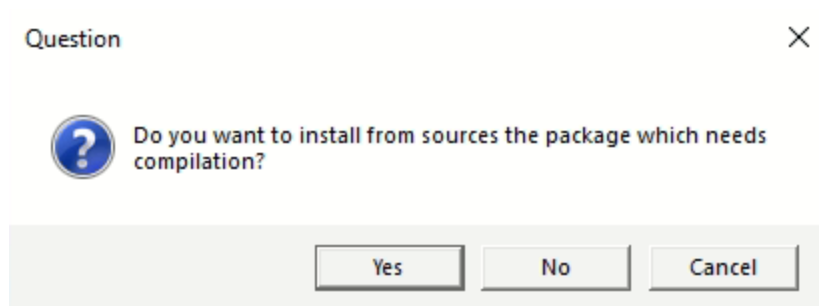


Figure S5: Installing helper packages for ImShot- (A) Update (if needed) all packages while installing local helper packages (B) Install from source and compile (Click Yes)

If the question as displayed in figure S5(A) appears during installation, enter 1 (as shown in the figure); if the question as depicted in figure S5(B) appears, click Yes. Proceed to the next steps only if you see the following output in RStudio console:

All the required packages are installed successfully

This step will install the required R packages and set up the backend. It is important to note that this will also install ImShot R package in the default R library path.

Step 3 - Unzip **release-builds-win-v1.2.1.7z** and launch the app by double clicking its icon (figure S6)

Name	Date modified	Type	Size
locales	09-Sep-21 11:10 AM	File folder	
resources	09-Sep-21 11:10 AM	File folder	
swiftshader	09-Sep-21 11:10 AM	File folder	
chrome_100_percent.pak	09-Sep-21 11:08 AM	PAK File	177 KB
chrome_200_percent.pak	09-Sep-21 11:08 AM	PAK File	314 KB
d3dcompiler_47.dll	09-Sep-21 11:08 AM	Application extension	3,628 KB
ffmpeg.dll	09-Sep-21 11:08 AM	Application extension	2,547 KB
icudtl	09-Sep-21 11:08 AM	DAT File	10,272 KB
ImShot_Electron_App	09-Sep-21 11:09 AM	Application	104,554 KB
libEGL.dll	09-Sep-21 11:08 AM	Application extension	348 KB
libGLSv2.dll	09-Sep-21 11:08 AM	Application extension	8,152 KB
LICENSE	09-Sep-21 11:08 AM	File	2 KB
LICENSES.chromium	09-Sep-21 11:08 AM	Chrome HTML Docume...	4,503 KB
resources.pak	09-Sep-21 11:08 AM	PAK File	4,229 KB
snapshot_blob.bin	09-Sep-21 11:08 AM	BIN File	50 KB
v8_context_snapshot.bin	09-Sep-21 11:08 AM	BIN File	168 KB
version	09-Sep-21 11:08 AM	File	1 KB
vk_swiftshader.dll	09-Sep-21 11:08 AM	Application extension	3,668 KB
vk_swiftshader_icd	09-Sep-21 11:08 AM	JSON File	1 KB
vulkan-1.dll	09-Sep-21 11:08 AM	Application extension	605 KB

Figure S6: Double click ImShot icon (enclosed in a black dashed rectangle) to launch the app

Do not delete any files/folders present in **release-builds-win-v1.2.1\ImShot_Electron_App-win32-ia32** directory. To ensure safety, once the program is launched for the first time, users may wish to pin it to the taskbar; this will save them from having to navigate to the **release-builds-win-v1.2.1** directory and clicking on the icon each time.

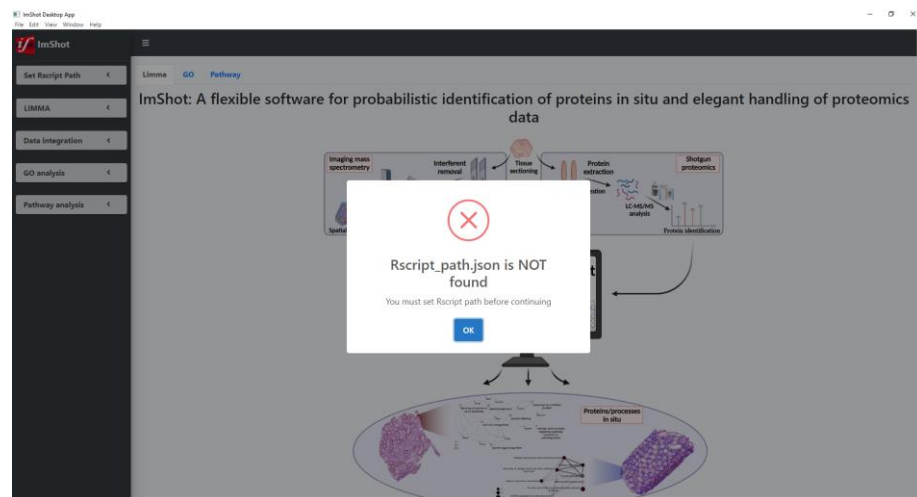


Figure S7: Pop-up message when ImShot is launched for the first time

ImShot will launch (first time) with a pop-up message (see figure S7) asking users to set Rscript path before attempting any analysis. Users must not ignore this message and must set the path as described in the next step.

Step 4 - In RStudio console run the following command and set the Rscript path as shown in the video file named **Set_Rscript_Path** (inside Video_Tutorials\For_ImShot_Desktop_App)

```
normalizePath(R.home(component = "bin"), winslash = '/')
```

Correct execution of this step is required for running ImShot desktop application successfully!

Regarding details of input, output, and other modules of the ImShot software, video tutorials have been created. Based on the user's requirement they can go through the videos present inside Video_Tutorials\For_ImShot_Desktop_App folder. All the operations, inputs along with R version are logged in a file named **ImShot.log** at %appdata%\ImShot_Electron_App folder, so that users can reproduce the analysis a later time point. Moreover, **ImShot.log** (figure S3) will also help users, developers to report and debug errors.

4.3 Testing ImShot desktop app

The folder **Video_Tutorials\For_ImShot_Desktop_App** contains all the videos that provide hands-on training on various ImShot desktop app features. For testing ImShot desktop app data from **Datasets** folder can be used.

The **Datasets** folder contains following 4 sub-folders,

- Enrichment_Data
- IMS_cluster_masslists
- IMS_Spectrum
- MQ_output_files

MQ_output_files: It contains a proteingroups.txt and a peptides.txt file. These are MaxQuant output files to be used for two group comparisons using ImShot. To test two group comparisons upload peptides.txt/proteingroups.txt file while doing Limma analysis using ImShot desktop app. See **limma_operation_and_output_files** video.

IMS_cluster_masslists: It contains excel files holding the IMS cluster mass-lists which need to be deisotoped. See **data_integration** video.

IMS_Spectrum: It contains IMS spectrum files for both diseased and healthy groups. The spectrum file is needed to correct IMS peptide peaks. See **data_integration** video.

Enrichment_Data: It contains excel files with gene lists that need to be used in GO/Pathway analyses and subfolder **Background** which in turn a list of uniprot ids that can be used as background gene list during the over-representation test. Both the background file (must be in .txt format) and the gene list file (must be in .xlsx format) must contain a column containing uniprot ids and the column names must be “uniprot” (See **GO_analysis** and **Pathway_analysis** videos). In the absence of a background file, the users can opt for the ‘default’ option in the background dropdown within the ImShot desktop widget.

4.4 How to install ImShot R package

If you have already executed the R code mentioned in step 2 of section 4.2 (How to run ImShot desktop application) then ImShot R package should be installed in your machine. Otherwise, the follow instructions given in that step.

4.5 Testing ImShot R package

The folder **Video_Tutorials\For_ImShot_R_package** contains all the videos that demonstrate how to test ImShot. The dataset for testing is same as described in section 4.3. There are five R scripts present inside **Testcases** folder (figure S8) to test the R package. These five testcases tests the all the functionalities offered by ImShot. The GUI can also be tested for these testcases.






Name	Date modified	Type
 test_deisotoping_analysis	08-Sep-21 6:28 PM	R File
 test_go_enrichment_analysis	08-Sep-21 6:38 PM	R File
 test_limma_analysis	08-Sep-21 6:26 PM	R File
 test_MLP_filtering_analysis	07-Sep-21 6:34 PM	R File
 test_pathway_enrichment_analysis	08-Sep-21 6:42 PM	R File

Figure S8: Testcases for testing ImShot

5. Usage guidelines for Linux desktop

5.1 Setting R environment

Step 1 - Download 64-bit RStudio and install

Step 2 - To install R (on Ubuntu 20.04), run following command in the terminal:

01. Install the dependencies necessary to add a new repository over HTTPS:

```
sudo apt install dirmngr gnupg apt-transport-https ca-certificates
software-properties-common
```

02. Add the CRAN repository to your system sources' list:

```
sudo apt-key adv --keyserver keyserver.ubuntu.com --recv-keys
E298A3A825C0D65DFD57CBB651716619E084DAB9
```

if the above command fails try the following one

```
gpg --keyserver hkp://keyserver.ubuntu.com:80 --recv-keys
E298A3A825C0D65DFD57CBB651716619E084DAB9
```

Then add the repository

```
sudo add-apt-repository 'deb https://cloud.r-
project.org/bin/linux/ubuntu focal-cran40/'
```

03. Install R by typing:

```
sudo apt install r-base
```

04. The installation may take a few minutes to complete. Once completed, verify it by printing the R version:

```
R -version
```

05. To be able to compile R packages, you need to install the build-essential package:

```
sudo apt install build-essential
```

```
sudo apt-get build-dep r-base-core
```

5.2 How to run ImShot desktop application

Step 1 – Follow step 1 of section 4.2.

Step 2 – Follow step 2 of section 4.2.

Step 3 – Unzip **release-builds-linux-v1.2.1.7z**. In the Linux terminal, change directory to the **release-builds-linux-v1.2.1\ImShot_Electron_App-linux-x64** folder and launch the app by running the following command (See figure S9).

```
./ImShot_Electron_App
```

```
-rwxr-xr-x 1 wasim wasim 7937624 Feb 2 13:41 libvulkan.so
-rw-r--r-- 1 wasim wasim 4349187 Feb 2 13:41 resources.pak
-rw-r--r-- 1 wasim wasim 107 Feb 2 13:41 vk_swiftshader_icd.json
-rw-r--r-- 1 wasim wasim 1060 Feb 2 13:41 LICENSE
-rw-r--r-- 1 wasim wasim 4458315 Feb 2 13:41 LICENSES.chromium.html
-rw-r--r-- 1 wasim wasim 6 Feb 2 13:41 version
-rwxr-xr-x 1 wasim wasim 132301272 Feb 2 13:41 ImShot_Electron_App ! ← ImShot binary
-rwxr-xr-x 1 wasim wasim 9393752 Feb 2 13:41 libGLSv2.so
-rw-r--r-- 1 wasim wasim 322069 Feb 2 13:41 chrome_200_percent.pak
-rw-r--r-- 1 wasim wasim 51262 Feb 2 13:41 snapshot_blob.bin
-rwxr-xr-x 1 wasim wasim 5811400 Feb 2 13:41 chrome-sandbox
-rwxr-xr-x 1 wasim wasim 253464 Feb 2 13:41 libEGL.so
-rw-r--r-- 1 wasim wasim 180782 Feb 2 13:41 chrome_100_percent.pak
-rwxr-xr-x 1 wasim wasim 3193952 Feb 2 13:41 libffmpeg.so
-rw-r--r-- 1 wasim wasim 10518160 Feb 2 13:41 icudtl.dat
drwxrwxr-x 2 wasim wasim 4096 Feb 2 13:41 swiftshader
-rw-r--r-- 1 wasim wasim 171504 Feb 2 13:41 v8_context_snapshot.bin
drwxrwxr-x 2 wasim wasim 4096 Feb 2 13:41 locales
-rwxr-xr-x 1 wasim wasim 3873656 Feb 2 13:41 libvk_swiftshader.so
drwxrwxr-x 3 wasim wasim 4096 Feb 2 13:41 resources
wasim@DNA2: [red box] ./ImShot_Electron_App ! ← Run it this way
```

Figure S9: Run the binary (in red rectangle) to launch the app

ImShot will launch (first time) with a pop-up message (See figure S7, section 4.2) asking users to set Rscript path before attempting any analysis. Users must not ignore this message and must set the path as described in the step 4 of section 4.2.

5.3 Testing ImShot desktop app

Testing the ImShot desktop application can be done as described in section 4.3. Although the videos referenced in Section 4.3 were recorded on Windows, Linux users can still follow along. For Linux users, the only difference in the videos is the output directory. In Linux the outputs are written in the following folder `/home/<username>/.config/ImShot_Electron_App`

5.4 How to Install ImShot R package

To install ImShot R package on Linux follow the instructions from section 4.4.

5.5 Testing ImShot R package

To test ImShot R package on Linux follow the instructions from section 4.5. However, as mentioned in section 5.3, Linux users need to look for output in a different directory than what is shown in the videos.

6. Usage guidelines for MacOS

6.1 Setting R environment

Step 1 - Download 64-bit R, RStudio and install

Step 2 (optional) - Download and install XCode from the Mac AppStore: <http://itunes.apple.com/us/app/xcode/id497799835?mt=12>. Within XCode go to Preferences : Downloads and install the Command Line Tools. Attempt this step only if you get compilation error for some packages in Step 2 of the next section.

6.2 How to run ImShot desktop application

Step 1 – Follow step 1 of section 4.2.

Step 2 – Follow step 2 of section 4.2.

Step 3 – Unzip **release-builds-mac-v1.2.1.zip**. Go to the **release-builds-mac-v1.2.1/ImShot_Electron_App-darwin-x64** folder and launch the app by double clicking its icon (See figure S10, ImShot icon is highlighted using a red rectangle)

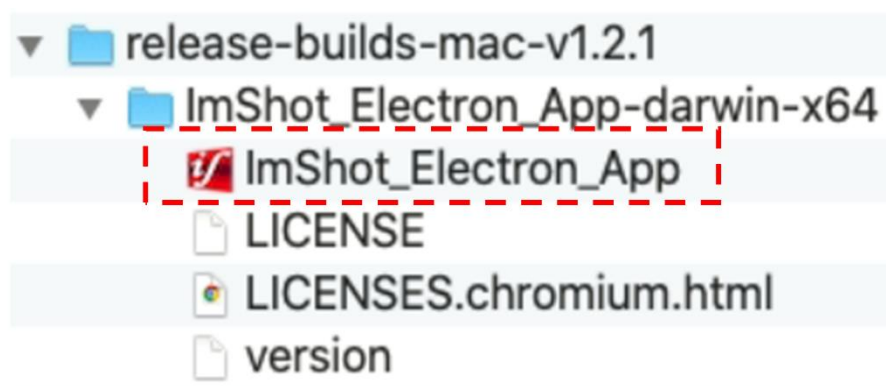


Figure S10: Double click ImShot icon (in dashed red rectangle) to launch the app

ImShot will launch (first time) with a pop-up message (See figure S7, section 4.2) asking users to set Rscript path before attempting any analysis. Users must not ignore this message and must set the path as described in the step 4 of section 4.2.

6.3 Testing ImShot desktop app

Testing the ImShot desktop application can be done as described in section 4.3. Although the videos referenced in Section 4.3 were recorded on Windows, Mac users can still follow along. For Mac users, the only difference in the videos is the output directory. In MacOS the outputs are written in the following folder

```
/Users/<username>/Library/Application Support/ImShot_Electron_App
```

6.4 How to Install ImShot R package

To install ImShot R package on Mac follow the instructions from section 4.4.

6.5 Testing ImShot R package

To test ImShot R package on MacOS follow the instructions from section 4.5. However, as mentioned in section 6.3, Mac users need to look for output in a different directory than what is shown in the videos.

7. Imputation algorithms in ImShot

ImShot provides ten widely used imputation procedures for dealing with missing values in LC-MS datasets. Each algorithm is briefly outlined below:

- **impSeq:** It begins with a complete subset M_c (i.e., with no missing values) from original dataset M and estimates the missing values in the incomplete part of data sequentially. For this method to work a set of proteins (in a given dataset) with no missing values is needed. Detailed algorithm is presented in this paper (1).
- **impSeqRob:** The impSeq method described earlier is not reliable when there are outliers present in a dataset because it utilizes mean and covariance of the data to calculate missing values. impSeqRob is a robust implementation of impSeq method, here the improved imputation is achieved by using robust estimators for mean and scatter. For in-depth explanation of this algorithm see this paper (2).
- **missForest:** It is non-parametric method, uses random forest algorithm to impute missing values. It handles the missingness in data using an iterative imputation scheme i.e., by training an RF on

observed values in a first phase, followed by predicting the missing values and then continuing iteratively. Detailed algorithm is presented in this paper (3).

- **imputePCA:** This method impute missing values in a proteomic dataset by employing EM method or regularised iterative PCA approach. See this article for more information (4)
- **ppca:** Probabilistic PCA is a PCA technique that combines an expectation maximization (EM) approach with a probabilistic model. The EM model is founded on the idea that both latent variables and noise are normally distributed. See this article (5) for more information.
- **MIPCA:** This method generates D imputed datasets applying a PCA model. The observed values are consistent across datasets; however, the imputed values vary. The missing values are predicted using the variation among the imputed values. See this article (6) for more information.
- **bpca:** This method imputes missing values based on Bayesian pca. By switching between updating the posterior distribution of the pca parameters and the posterior distribution of the missing values, bpca iteratively replaces missing values with estimates. See this article (7) for more information.
- **SVDImpute:** This algorithm estimates missing values as a linear combination of the k most significant eigenproteins. For more information see this article (8)
- **kNNImpute:** Here missing values in a protein is imputed by computing weighted average of the intensity values from K closest (in Euclidian space) proteins. For detailed explanation see this paper (8).
- **tiny:** See main text point (iv) under *Data processing (LC-MS)*

On the basis of our AROM+ dataset, we compared the performance of the aforementioned imputation techniques via two-group comparisons using Limma statistics. The number of statistically significant proteins are listed in Table S1, where the diagonal entries imply the number of statistically significant proteins obtained after imputing missing values using corresponding algorithms. And any off-diagonal entry corresponds to the number of statistically significant proteins shared by two algorithms. As shown in Table S1, all imputation algorithms perform similarly in terms of extracting statistically significant proteins from our dataset, with the exception of kNNImpute, which yields approximately 20%–23% fewer proteins.

Additionally, the results in Table S1 demonstrate that Limma is quite stable when the imputation method is changed; however, it also depicts that imputation has some effect on the statistical analyses. As a result, we propose experimenting with several imputation approaches and selecting the one that produces the best biologically meaningful results. However, one should have greater confidence in proteins that are found statistically significant by all/most of the approaches, rather than just one or two. Because ImShot features an intuitive GUI and facilitates identification of the biological themes associated with a set of proteins/genes (using GO and Pathway enrichment analyses modules) users can quickly experiment with different imputation algorithms on their datasets.

8. Description of output file headers after GO/Pathway Enrichment

It is highly recommended to follow the folder/sub-folder structure of **Datasets/Enrichment_Data** folder when doing GO/Pathway enrichment analysis. The header definitions of files containing GO/Pathway analyses results are provided below.

ID: ID of GO process (BP, MF, CC) or Pathway

Description: Description of GO process/Pathway

GeneRatio: It is the ratio between the number of genes in your input list associated with the given GO term or Pathway (e.g., 5) and the total number of input genes.

BgRatio: It is the ratio between the number of genes in your background list associated with the given GO term or Pathway (e.g., 115) and the total number of background genes.

pvalue: To determine whether any GO/Pathway term annotates a gene list at a frequency greater than that would be expected by chance, p-values are computed using hypergeometric distribution model.

p.adjust: Adjusted pvalue using one of the following correction methods: Holm's, Hochberg's, Hommel's, Bonferroni's, Benjamini and Hochberg, Benjamini and Yekutieli, FDR.

qvalue: Alternative approach to calculate FDR/control for multiple testing. The q-value of a test measures the proportion of false positives incurred (called the false discovery rate) when that particular test is called significant.

geneID: Genes from your input list that are found to be enriched in the given GO term/Pathway.

Count: Number of genes from your input list that are found to be enriched in the given GO term/Pathway.

9. Important Notes

1. The steps within **R environment Setup** and steps 2-3 in **How to run ImShot** are for setting the backend R environment and need to execute only once (when the ImShot is launched for the first time). If you want to use a different version of R in the backend, then you must set R-script path in ImShot afresh.
2. We tested the guidelines mentioned in this document on 64-bit Windows 10, Ubuntu 20.04, macOS Catalina (version 10.15.7) using 64-bit R (version 4.0.2) and recommend using R version $\geq 4.0.2$.
3. You might need to install 7 zip (<https://www.7-zip.org/>) to unzip the ImShot desktop applications.
4. For reproducibility, we recommend you use the parameter settings as shown in the videos while testing the software using the provided datasets, especially for data integration and GO/Pathway analyses.
5. Deisotoping expects a folder where mass lists from only one group (control /treatment) are kept and not mixed. Moreover, it also expects an appropriate IMS spectrum file, i.e., if you're deisotoping clusters from control then select/provide a spectrum file for control.
6. While doing Limma analysis using ImShot desktop app using the databases other than Uniprot and Flybase, a regular expression is needed to specify for extracting the gene names. For example: a row in *Fasta headers* column in MaxQuant output file (proteingroups.txt, peptides.txt) could be as follows `"Tb927.5.4450:mRNA-p1 | transcript=Tb927.5.4450:mRNA | gene=Tb927.5.4450 | organism=Trypanosoma_brucei_brucei_TREU927 | gene_product=hypothetical protein, conserved | transcript_product=hypothetical protein, conserved | location=Tb927_05_v5.1:1372470-137297"`. To extract the gene names (to appear in volcano plot) user must supply a regex via ImShot GUI and in this particular case it is `"gene=(.*) | organism"`. It is important to note here that the regex will vary depending on the patterns of the strings in *Fasta headers*.

Reference

1. Verboven, S., Branden, K. V., and Goos, P. (2007) Sequential imputation for missing values. *Computational Biology and Chemistry* 31, 320-327
2. Branden, K. V., and Verboven, S. (2009) Robust data imputation. *Computational Biology and Chemistry* 33, 7-13
3. Stekhoven, D. J., and Bühlmann, P. (2012) MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics* 28, 112-118
4. Josse, J., and Husson, F. (2012) Handling missing values in exploratory multivariate data analysis methods. *Journal de la Société Française de Statistique* 153, 79-99
5. Roweis, S. (1997) EM algorithms for PCA and SPCA. *Advances in neural information processing systems* 10
6. Josse, J., Pagès, J., and Husson, F. (2011) Multiple imputation in principal component analysis. *Advances in data analysis and classification* 5, 231-246
7. Oba, S., Sato, M.-a., Takemasa, I., Monden, M., Matsubara, K.-i., and Ishii, S. (2003) A Bayesian missing value estimation method for gene expression profile data. *Bioinformatics* 19, 2088-2096
8. Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., and Altman, R. B. (2001) Missing value estimation methods for DNA microarrays. *Bioinformatics* 17, 520-525