

# Step 1: Upload files

Files will be demultiplexed/quantified against the specified genome assembly.

Project name that could help to organize and identify your files

Sample name to quantify all files as the same sample

## Upload New Files

### 1) Sample type

Reference genome assembly ?

Human, Homo sapiens (GRCh38)

Project name ?

filtered\_feature\_bc\_matrixh5

Sample name ?

Write a sample name

File type

- FASTQ
- HDF5

Select the file format of files. FASTQ and HDF5 files are allowed. Alternatively, ENA project ID can be also specified on the "Import ENA project" menu.

Demultiplex ?

- 10x Chromium
- 10x Chromium v3
- Drop-seq

Demultiplex if this option is selected, samples will be demultiplexed with Alevin software. Demultiplexing is the process by which FASTQ reads are assigned to their cell of origin based on the sequence of their corresponding molecular barcode. This process will result in the quantification of gene expression for each input cell.


Sequencing type ?

- Single-end
- Paired-end

Sample name to quantify all files as the same sample

Advanced ▾

### 2) Select your FASTQ or HDF5 Files ?

Select or drag and drop one or more files (we recommend using google chrome 

No files selected.

Upload FASTQ or HDF5 files clicking on the 'browse...' button and selecting the desired files or doing a 'drag and drop' of your files on the square

# Step 2: Set Analysis parameters and launch analysis

File/Projects View Load Test Samples Search:

| <input type="checkbox"/> | Filename                        | Filename 2 | Sample name                  | Project name                 | Cells |
|--------------------------|---------------------------------|------------|------------------------------|------------------------------|-------|
| <input type="checkbox"/> | filtered_feature_bc_matrix.h5 ✓ | -          | filtered feature bc matrixh5 | filtered_feature_bc_matrixh5 | 1227  |

Show  entries Previous  Next

Number of cells in the HDF5 file

Shows demultiplexing and alignment statistics if FASTA files of ENA Project Identifier is provided

Descriptive name of the analysis. Results will be identified with this name. Parameters of each analysis can be queried on the "log" section.

Determines if genes/cells are filtered prior the analysis using the following methods:

- Gene expression mean/sd: filter genes with the lower expression mean or standard deviation.
- Seurat gene filtering: applies the FindVariableFeatures function of Seurat, setting mean.var.plot as selection method.
- DropletUtils Cell filtering: applies the emptyDrops function of the package dropletutils.
- Maximum cells: sets the maximum number of cells per sample which will be represented.

Cell expression is projected on two-dimensional plots. This projection is performed by the application of the following dimensional reduction methods:

- Principal Component Analysis (PCA).
- Classical Multi Dimensional Scaling (cMDS).
- Nonnegative Matrix Factorization (NMF).
- t-distributed stochastic neighbor embedding (t-SNE).
- Uniform Manifold Approximation and Projection (UMAP).
- Zero Inflated Factor Analysis (ZIFA).

Quality Control
Launch Analysis

### Analysis Parameters

Analysis name

#### Filtering

Filter  % of genes with lower expression

Filter  % of genes with lower sd

Seurat gene filter

DropletUtils Cell filtering

Maximum cells

#### Dimensionality reduction

PCA  cMDS

t-SNE  UMAP

NMF  ZIFA

### Clustering

Apply to components

Number of groups

Distance

Agglomeration

### Classification datasets

| Human:   | Mouse:                                 |
|--|--|
| <input type="checkbox"/> Human Primary Cell Atlas  | <input type="checkbox"/> ImmGen        |
| <input type="checkbox"/> Blueprint/ENCODE          | <input type="checkbox"/> Mouse RNA-seq |
| <input type="checkbox"/> Immune Cell Expression    |  |
| <input type="checkbox"/> Novershtern Hematopoietic |  |
| <input type="checkbox"/> Monaco Immune             |  |

SingleAnalyzer applies clustering methods. It applies hierarchical clustering (hclust), k-means, leiden and PAM methods. They obtain up to ten possible clustering configurations per method. These methods can be applied over the main components calculated with PCA. You can also configure the number of desired/expected groups. With the Auto option, SingleAnalyzer will determine the optimum number of clusters for your dataset. Moreover, you can also specify the distance metric and the agglomeration method for the hierarchical clustering.

SingleAnalyzer allows the automatic classification of Human and Mouse cells. It is performed with the SingleR package which performs a supervised classification of cells based on reference datasets. You can choose which one of the training datasets will be used for your cells.

# Step 3: Explore Results

Select all   Select none   Delete   Search:

|                          | Date ⓘ     | Analysis Name ⓘ    | Quantification ⓘ | QC ⓘ      | Cluster ⓘ | Comparisons                      |
|--------------------------|------------|--------------------|------------------|-----------|-----------|----------------------------------|
| <input type="checkbox"/> | 10/03/2022 | cluster_1646898858 | waiting 🌀        | waiting 🌀 | waiting 🌀 | -                                |
| <input type="checkbox"/> | 09/03/2022 | cluster_1646826215 | done 📄           | done 📊    | done 📊    | -                                |
| <input type="checkbox"/> | 21/01/2022 | cluster_1642765741 | done 📄           | done 📊    | done 📊    | <a href="#">View Comparisons</a> |

Shows if the demultiplexing and quantification have finished. The icon gives access to statistics, gene counts and TPMs (Transcript Per Millions).

Shows the quality control graphs of the quantified samples.

Shows if some dimensional reduction and clustering processes have finished. It redirects to the central visualization which represents the dimensional reduction and clustering of cells. (next page)

Shows if the differential expression analysis (DE) between groups has finished. It will redirect to a table with all the comparisons done and with links to dynamic reports.

# Dimensional Reduction/Clustering Results

Groups selected by the user. They can be added automatically by clicking on a star of the clusters or cell classification results. A manual modification is also possible adding or removing groups with "My Groups" controls or changing the cell group with the use of "Add to..." and "Remove from groups" buttons.





Generates new scatter plots which represent the expression of two genes in each cell

Colors each point with a scale color which represents the expression of the selected gene.

Points can be selected, or color and shape attributes can be changed based on clustering results. Ten possible clustering results were calculated with k-means, PAM and hclust methods and can be explored clicking on the number.

Points can be selected, or color and shape attributes can be changed based on the cell population prediction performed by SingleR

cluster\_1645796176  

My Groups    

- B cells (199)
- Monocytes (432)
- NK cells (161)
- T cells, CD4+ (347)
- T cells, CD8+ (88)

New Gene Expression Plot

Write Gene Name...

Write Gene Name...

Color by Gene Expression

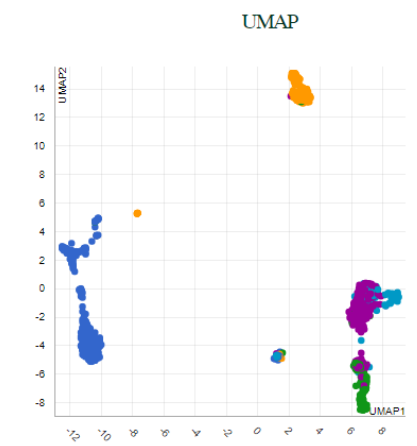
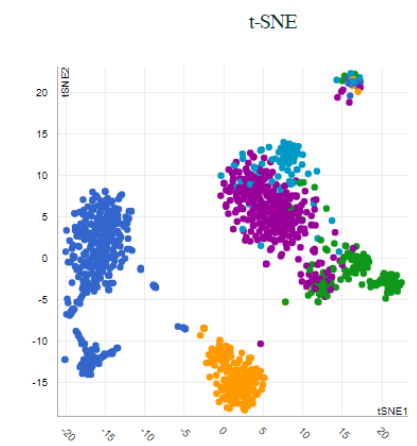
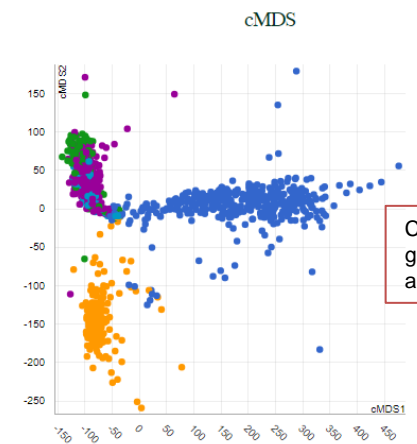
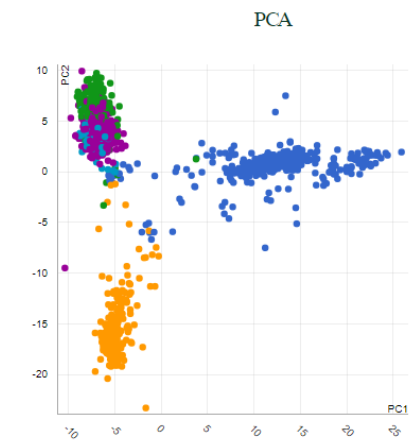
Write Gene Name...

Clusters

- k-means    **10**
- PAM    **8**
- hclust    **7**
- leiden    **umap**

Cell Classification

- Samples   
- Human Primary Cell Atlas   



Locate Samples

Write a sample...

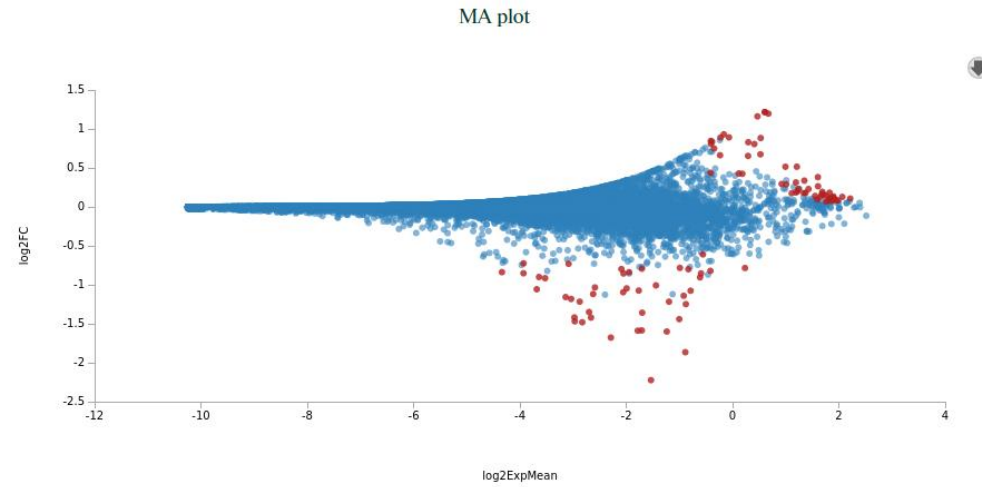
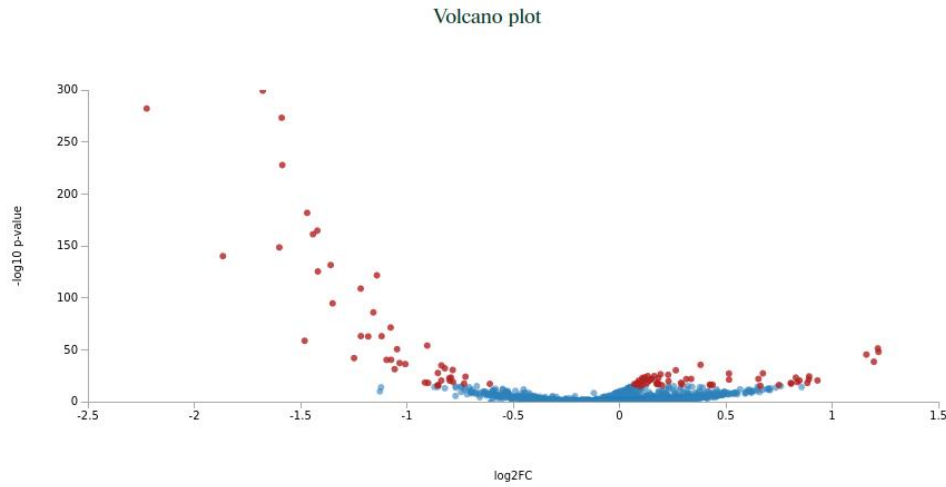
Write a sample pattern...

Selected Samples

- <empty>

Cells can be selected on the graphs writing their identifier or a text pattern.

# Differential Expression Results



685 significant genes of 16290

Select all Select none Show 100 entries

Search:

| ID   | Symbol | ExpMean | Exp k-means_6 | Exp k-means_1,k-means_2,k-means_3,k-means_4,k-means_5 | log2FC | lfcSE | Stat    | p-value  | padj     |
|--|--------|---------|---------------|---|--------|-------|---------|----------|----------|
| <input checked="" type="checkbox"/> ENSG00000134539.16 | KLRD1  | 0.205   | 2.429         | 0.072   | -1.677 | 0.076 | -37.021 | 5.3e-300 | 8.6e-296 |
| <input checked="" type="checkbox"/> ENSG00000115523.16 | GNLY   | 0.345   | 4.207         | 0.115   | -2.223 | 0.143 | -35.942 | 6.8e-283 | 5.5e-279 |
| <input checked="" type="checkbox"/> ENSG00000180644.7  | PRF1   | 0.29    | 2.488         | 0.16  | -1.589 | 0.079 | -35.367 | 5.5e-274 | 3.0e-270 |
| <input checked="" type="checkbox"/> ENSG00000077984.5  | CST7   | 0.306   | 2.524         | 0.174   | -1.586 | 0.079 | -32.271 | 1.8e-228 | 7.2e-225 |
| <input checked="" type="checkbox"/> ENSG00000137441.7  | FGFBP2 | 0.128   | 1.839         | 0.026   | -1.469 | 0.086 | -28.807 | 1.8e-182 | 5.8e-179 |
| <input checked="" type="checkbox"/> ENSG00000159674.11 | SPON2  | 0.158   | 1.833         | 0.058   | -1.421 | 0.087 | -27.407 | 2.3e-165 | 6.2e-162 |
| <input checked="" type="checkbox"/> ENSG00000172543.7  | CTSW   | 0.499   | 2.713         | 0.367   | -1.441 | 0.094 | -27.114 | 6.7e-162 | 1.6e-158 |
| <input checked="" type="checkbox"/> ENSG00000145649.7  | GZMA   | 0.424   | 2.874         | 0.278   | -1.6   | 0.11  | -26.017 | 3.2e-149 | 6.5e-146 |
| <input checked="" type="checkbox"/> ENSG00000105374.9  | NKG7   | 0.54    | 3.885         | 0.341   | -1.865 | 0.153 | -25.265 | 7.8e-141 | 1.4e-137 |

Ensembl Gene Identifier of each gene

Official Gene Symbol of each gene

Expression mean of each gene in the input samples

Mean expression in the first group samples

Mean expression in the first group samples

log2 scaled fold change of each gene

Standard error of the log2 scaled fold change

Statistic calculated by the selected analysis method

Unadjusted p-value

Adjusted p-value

# Functional Analysis Results

GO Process ▼

Select all Select none Show 10 entries

Search:

| <input type="checkbox"/> | GO_ID      | GO_TERM   | Category | NumDEInCat | NumDEAnot | DEPercent | NumInCat | NumGenesAnot | AnotPercent | p-value |
|--------------------------|------------|---|----------|------------|-----------|-----------|----------|--------------|-------------|---------|
| <input type="checkbox"/> | GO:0006614 | SRP-dependent cotranslational protein targeting to membrane         | Process  | 78         | 655       | 11.908    | 89       | 12490        | 0.713       | 1.7e-80 |
| <input type="checkbox"/> | GO:0006413 | translational initiation  | Process  | 91         | 655       | 13.893    | 131      | 12490        | 1.049       | 1.2e-78 |
| <input type="checkbox"/> | GO:0000184 | nuclear-transcribed mRNA catabolic process, nonsense-mediated decay | Process  | 81         | 655       | 12.366    | 116      | 12490        | 0.929       | 9.1e-70 |
| <input type="checkbox"/> | GO:0006412 | translation   | Process  | 81         | 655       | 12.366    | 171      | 12490        | 1.369       | 2.3e-52 |
| <input type="checkbox"/> | GO:0002181 | cytoplasmic translation   | Process  | 38         | 655       | 5.802     | 50       | 12490        | 0.4         | 4.1e-35 |
| <input type="checkbox"/> | GO:0043312 | neutrophil degranulation  | Process  | 89         | 655       | 13.588    | 432      | 12490        | 3.459       | 1.3e-28 |
| <input type="checkbox"/> | GO:0006955 | immune response   | Process  | 63         | 655       | 9.618     | 273      | 12490        | 2.186       | 1.2e-24 |
| <input type="checkbox"/> | GO:0050776 | regulation of immune response                                       | Process  | 35         | 655       | 5.344     | 124      | 12490        | 0.993       | 1.7e-16 |
| <input type="checkbox"/> | GO:0060333 | interferon-gamma-mediated signaling pathway                         | Process  | 22         | 655       | 3.359     | 62       | 12490        | 0.496       | 2.1e-13 |
| <input type="checkbox"/> | GO:0019221 | cytokine-mediated signaling pathway                                 | Process  | 35         | 655       | 5.344     | 200      | 12490        | 1.601       | 2.2e-10 |

Copy CSV Excel

Previous 1 2 3 4 5 ... 118 Next

Gene  
Ontology  
identifier

Gene  
Ontology  
term

Type of  
term in the  
Gene  
Ontology

N<sup>a</sup> of DE  
genes  
annotated to  
each term

Genes  
annotated  
to GO  
terms

Percentag  
e of DEGs  
annotated

Number of  
genes  
annotated  
to each  
GO term

Number of  
genes  
annotated  
to Gene  
Ontology

Percentag  
e of genes  
annotated  
to each  
term

P-value of  
a possible  
functional  
enrichmen  
t