

Workshop:

Single Cell RNA-seq analysis using Cellenics



<https://www.biomage.net>

Oliver@biomage.net

Vicky@biomage.net

peter_kharchenko@hms.harvard.edu

GOAL

Q: How many of you have a single cell RNA-seq dataset that you are trying to / want to analyse?

-> Write 'yes' in the zoom chat

We will help you to get your analysis done!

Book a free 45-minute meeting with Vicky (biologist) and Oliver (bioinformatician) to discuss your analysis needs:

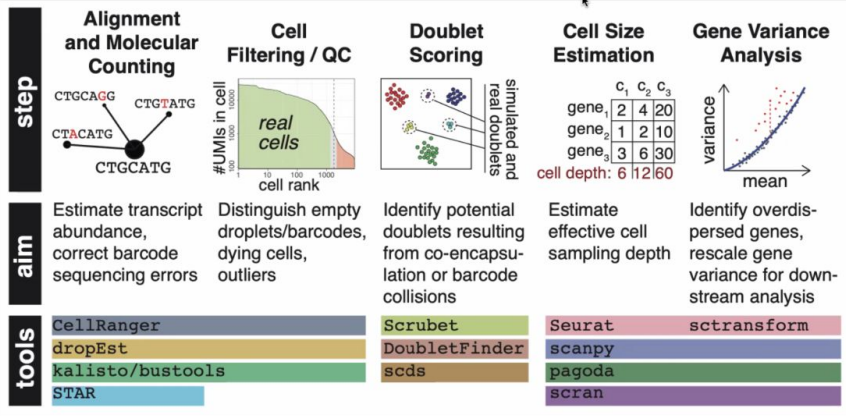
<https://calendly.com/vicky-morrison/cellenics-onboarding>

INTRODUCTION

By Peter Kharchenko, Associate Professor of Biomedical Informatics at Harvard Medical School

<https://www.nature.com/articles/s41592-021-01171-x>

scRNA-seq analysis steps



Review Article | [Published: 21 June 2021](#)

The triumphs and limitations of computational methods for scRNA-seq

[Peter V. Kharchenko](#)

[Nature Methods](#) **18**, 723–732 (2021) | [Cite this article](#)

15k Accesses | 1 Citations | 241 Altmetric | [Metrics](#)

Cellenics Overview

In this workshop you will be analysing a scRNA-seq dataset using Cellenics.

- Cellenics is cloud-based (AWS) - there is no need for you to install anything on your computer.
- Free for academic use
- Platform was released 16 weeks ago

We want to hear your feedback!

-> use the 'Feedback' button at the top

You may encounter issues/bugs

-> tell us and we will fix them!



If you need help now or at any point during the tutorial exercise, post a message in the zoom chat and one of the Biomage team will help.

Logging in

- We have created an account for each of you and we have already uploaded the example dataset in order to save time.

- Let's log in to Cellenics now using Chrome:

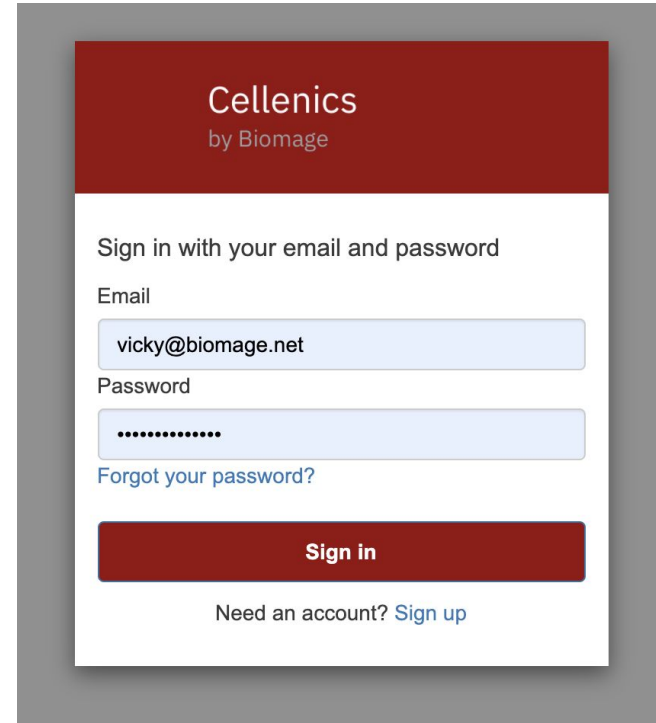
<https://scp.biomage.net/>

You should have received your login details by email ~1 hour ago.

Do check your junk folder.

If you can't find your login details or have issues, let us know in the zoom chat.

Once you have logged in, type "I'm in" in the zoom chat!



The screenshot shows the login interface for Cellenics by Biomage. It features a dark red header with the logo. Below the header, there is a white box containing the login form. The form includes a title, a sign-in instruction, an email input field with the address 'vicky@biomage.net', a password input field with masked characters, a 'Forgot your password?' link, a prominent red 'Sign in' button, and a 'Need an account? Sign up' link at the bottom.

Cellenics
by Biomage

Sign in with your email and password

Email

Password

[Forgot your password?](#)

Sign in

Need an account? [Sign up](#)

Data Management module

Cellenics
by Biomage

Data Management

Data Management

Feedback or issues? ↕ Invite a friend ↕ B

Data Management

No analysis

- Data Processing
- Data Exploration
- Plots and Tables

Team - Careers - Contact

© 2021 Biomage Ltd, affiliates & contributors.

Data Management

Projects

Create New Project

Filter by project name, project ID or analysis ID

GSE183716 - Covid19 MISC ✎ 🗑

Samples: 2

Created: [3 hours ago](#)

Modified: [2 hours ago](#)

Last analyzed: [2 hours ago](#)

Project Details

GSE183716 - Covid19 MISC

ID : 64b51570-3d8c-4fca-8aa9-3a2e8a116f9d

Description: [✎](#)

Add metadata Add samples Download [Go to Data Processing](#)

Sample	barcodes.tsv	genes.tsv	matrix.mtx	Species ?
☰ Convalescent ✎ 🗑	Uploaded	Uploaded	Uploaded	N.A. ⌵
☰ Acute ✎ 🗑	Uploaded	Uploaded	Uploaded	N.A. ⌵

Data Management module

Cellenics
by Biomage

Data Management

Data Management

Feedback or issues? ↕ Invite a friend ↕ B

Projects

Create New Project

Filter by project name, project ID or analysis ID

GSE183716 - Covid19 MISC ✎ 🗑️

Samples: 2

Created: [3 hours ago](#)

Modified: [2 hours ago](#)

Last analyzed: [2 hours ago](#)

Project Details

GSE183716 - Covid19 MISC

Add metadata Add samples Download [Go to Data Processing](#)

ID : 64b51570-3d8c-4fca-8aa9-3a2e8a116f9d

Description: [✎](#)

Sample	barcodes.tsv	genes.tsv	matrix.mtx	Species ?
☰ Convalescent ✎ 🗑️	Uploaded	Uploaded	Uploaded	N.A. ▼
☰ Acute ✎ 🗑️	Uploaded	Uploaded	Uploaded	N.A. ▼

List of your projects.
Select the Covid project.

Team - Careers - Contact

© 2021 Biomage Ltd,
affiliates & contributors.

Data Management module

The screenshot displays the Cellenics Data Management interface. On the left is a dark sidebar with navigation options: 'Data Management', 'No analysis', 'Data Processing', 'Data Exploration', and 'Plots and Tables'. The main content area is titled 'Data Management' and features a 'Projects' section with a 'Create New Project' button and a search filter. A project card for 'GSE183716 - Covid19 MISC' is highlighted, showing it has 2 samples and was created, modified, and last analyzed 2-3 hours ago. The 'Project Details' view for this project is shown, including its ID, a description, and a table of samples. The table lists two samples: 'Convalescent' and 'Acute', both with 'Uploaded' status for barcodes, genes, and matrix files. A red box highlights the project details section, and a red arrow points from a text box below to the sample table.

Cellenics
by Biomage

Data Management

Data Management

Feedback or issues? ↕ Invite a friend ↕ B

Projects

Create New Project

Filter by project name, project ID or analysis ID

GSE183716 - Covid19 MISC ✎ 🗑

Samples: 2

Created: [3 hours ago](#)

Modified: [2 hours ago](#)

Last analyzed: [2 hours ago](#)

Project Details

GSE183716 - Covid19 MISC Add metadata Add samples Download Go to Data Processing

ID : 64b51570-3d8c-4fca-8aa9-3a2e8a116f9d

Description: [🔗](#)

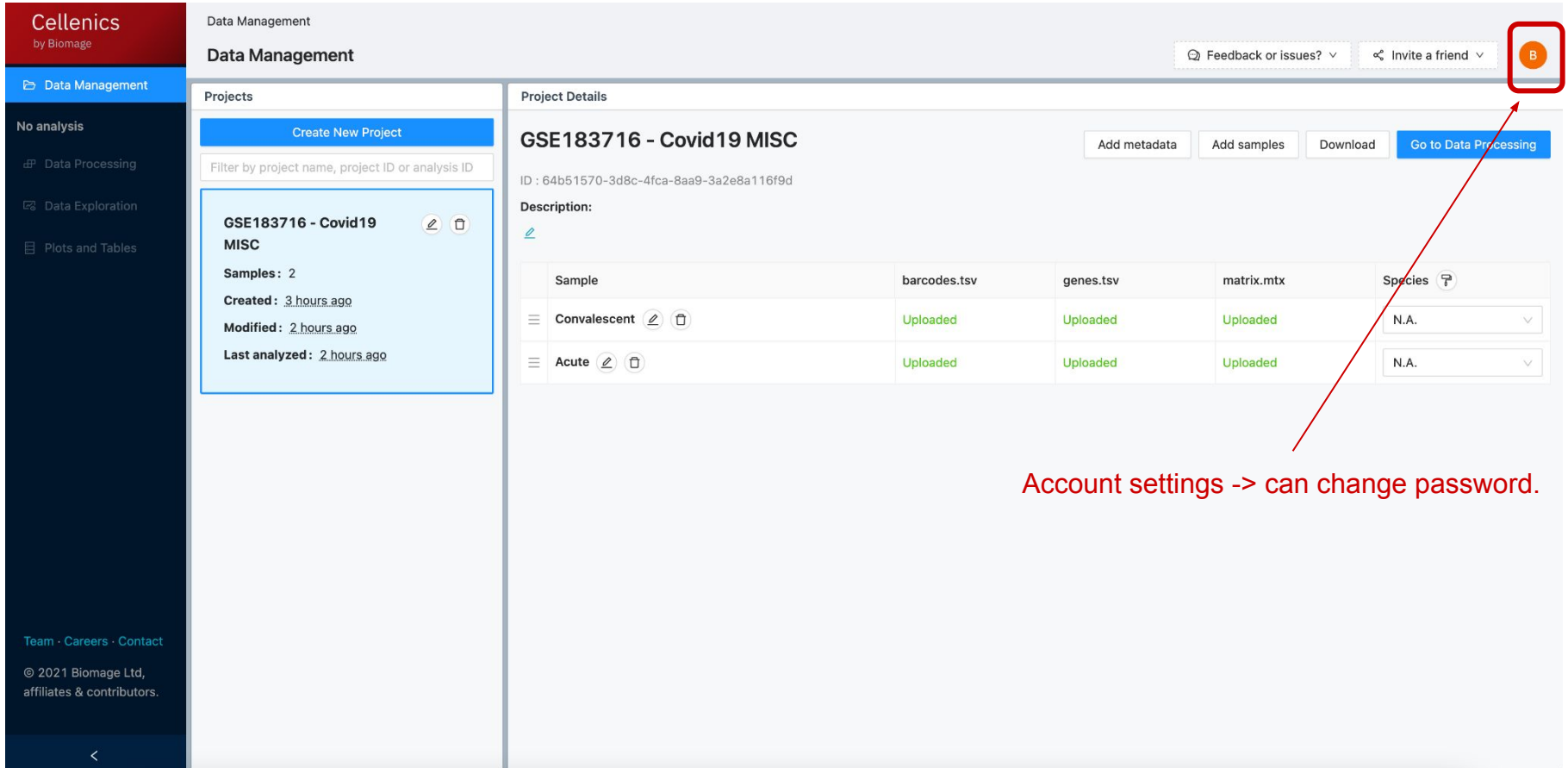
Sample	barcodes.tsv	genes.tsv	matrix.mtx	Species ?
☰ Convalescent ✎ 🗑	Uploaded	Uploaded	Uploaded	N.A. ▼
☰ Acute ✎ 🗑	Uploaded	Uploaded	Uploaded	N.A. ▼

List of samples within the selected project
You should see these 2 samples in the Covid project.

Team - Careers - Contact

© 2021 Biomage Ltd, affiliates & contributors.

Data Management module



Cellenics
by Biomage

Data Management

Data Management

Feedback or issues? ⌵ | Invite a friend ⌵ | **B**

Projects

Create New Project

Filter by project name, project ID or analysis ID

GSE183716 - Covid19 MISC ✎ 🗑️

Samples: 2

Created: [3 hours ago](#)

Modified: [2 hours ago](#)

Last analyzed: [2 hours ago](#)

Project Details

GSE183716 - Covid19 MISC ➤

ID : 64b51570-3d8c-4fca-8aa9-3a2e8a116f9d

Description: ✎

Add metadata | Add samples | Download | **Go to Data Processing**

Sample	barcodes.tsv	genes.tsv	matrix.mtx	Species ?
☰ Convalescent ✎ 🗑️	Uploaded	Uploaded	Uploaded	N.A. ⌵
☰ Acute ✎ 🗑️	Uploaded	Uploaded	Uploaded	N.A. ⌵

Team - Careers - Contact

© 2021 Biomage Ltd, affiliates & contributors.

<

Account settings -> can change password.

Tutorial Exercise Dataset

Reference:

The innate and adaptive immune landscape of SARS-CoV-2-associated multisystem inflammatory syndrome in children (MIS-C) from acute disease to recovery.

GEO accession GSE183716

<https://www.medrxiv.org/content/10.1101/2020.08.06.20164848v2.full.pdf>

About the dataset:

Publicly available dataset, published on GEO on Sep 10, 2021

Two PBMC samples isolated from humans with COVID-19 associated MIS-C:

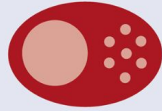
1. Acute patient
2. Convalescent post-treatment patient

PBMCs

- Critical components of the immune system
- Easily isolated from peripheral blood
- Widely used in areas relating to immunology, infectious disease, vaccine development, and others.



CD4⁺
T cells
25–60%



CD8⁺
T cells
5–30%



CD19⁺
B cells
5–10%



CD56⁺ CD3⁻
NK cells
10–30%



CD14⁺
monocytes
5–10%

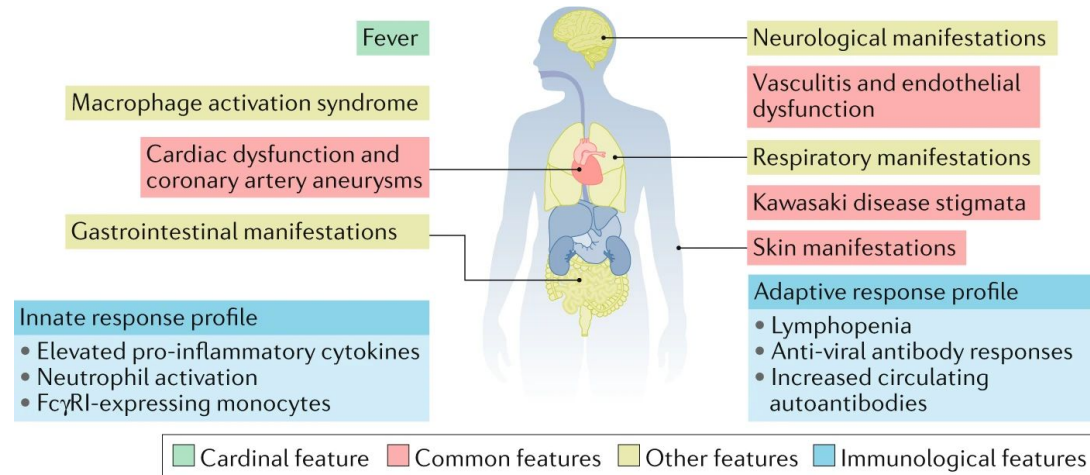


Dendritic
cells
1–2%

Common Cell Types in PBMCs: Lymphocytes, Monocytes and Dendritic cells.

Multisystem inflammatory syndrome in children (MIS-C)

- Associated with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Usually develops 4-6 weeks after SARS-CoV-2 infection.
- Characterized by a marked inflammatory state in various organs.
- Can be serious and potentially fatal: cardiac symptoms are common.
- Our understanding of this disease remains poor.



<https://www.nature.com/articles/s41584-020-00566-y>

Data Management module

Cellenics
by Biomage

Data Management

Data Management

Feedback or issues? | Invite a friend | B

Projects

Create New Project

Filter by project name, project ID or analysis ID

GSE183716 - Covid19 MISC

Samples: 2
Created: 3 hours ago
Modified: 2 hours ago
Last analyzed: 2 hours ago

Project Details

GSE183716 - Covid19 MISC

Add metadata | Add samples | Download | **Go to Data Processing**

ID : 64b51570-3d8c-4fca-8aa9-3a2e8a116f9d

Description:

Sample	barcodes.tsv	genes.tsv	matrix.mtx	Species
Convalescent	Uploaded	Uploaded	Uploaded	N.A.
Acute	Uploaded	Uploaded	Uploaded	N.A.

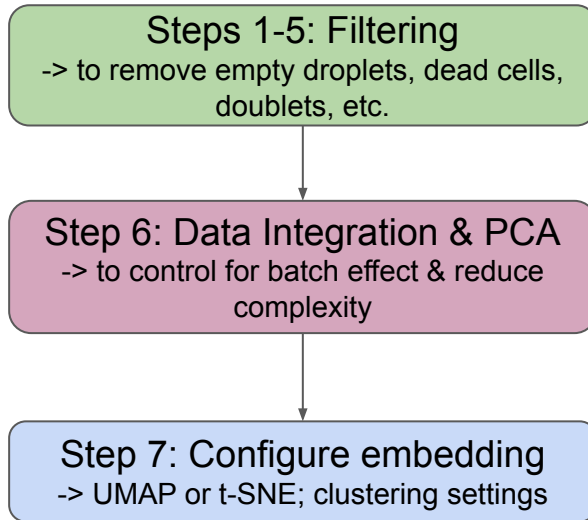
ACTION
Click 'Go to Data Processing'

If the button says 'Process project' click it and let us know in the zoom chat.

Team - Careers - Contact
© 2021 Biomage Ltd, affiliates & contributors.

Data Processing module

Data processing pipeline is already applied to the dataset:



Manual control over the settings:

The screenshot shows the 'Filtering Settings' panel. The 'Automatic' radio button is selected and highlighted with a red box. The 'Manual' radio button is unselected. Below the radio buttons, there are two sliders: 'Max percentage' set to 10 and 'Bin step' set to 0.05. A 'Copy to all samples' button is located below the sliders. At the bottom of the panel, there is a '> Plot styling' link.

Please DO NOT make and apply ('Run') changes to the Data Processing module right now

Data Processing module

Cellenics
by Biome

Analyses / GSE183716 - Covid19 MISC / Data Processing

Data Processing

Feedback or issues? ↕ Invite a friend ↕ A

X 1-Classifier filter ⌵ ⌚ Enable

- X 1-Classifier filter
- X 2-Cell size distribution filter
- ✓ 3. Mitochondrial content filter
- ✓ 4. Number of genes vs UMIs filter
- ✓ 5. Doublet filter
- ✓ 6. Data integration
- ✓ 7. Configure embedding

of the sample(s) is pre-filtered. Click 'Next' to continue processing your data.

Nothing to show yet
Results will appear here when they're available.

Filtering Settings

Automatic Manual

FDR: ⌚ 0.01

Copy to all samples

> Plot styling

Status: ▬▬▬▬▬▬ ✓ ⏪ ⏩

Pipeline status indicator

Exercise 1 - Assessment of sample quality [8mins]

Determine whether the two samples differ significantly in quality.

- Do they contain similar proportions of dead cells?
- Do they contain similar proportions of doublets?
- Are the 2 samples integrated well?

INSTRUCTIONS:

- Go to Data Processing module
- Use the 'next' button to navigate through the 7 steps
- Look in detail at the mitochondrial content filter (step 3), doublet filter (step 5) and integration (step 6)

**When you have completed it, write
“Done 1” in chat.**

The screenshot shows the Cellenics Data Processing interface. A sidebar on the left contains a navigation menu with 'Data Processing' highlighted in blue. A red box labeled '1.' points to this menu item. The main panel shows the 'Data Processing' workflow with a red box labeled '3.' pointing to the '3. Mitochondrial content filter' step. A 'Next' button (a right-pointing arrow) is circled in red and labeled '2.'. Below the workflow, a histogram shows the 'Percentage of cells' vs 'Percentage of mitochondrial reads' for 'Alive' (green) and 'Dead' (blue) cells. A vertical dashed red line is at approximately 10%. To the right, 'Filtering Settings' are shown with 'Automatic' selected, 'Max percentage' at 10, and 'Bin step' at 0.05. At the bottom, a 'Statistics' table is visible.

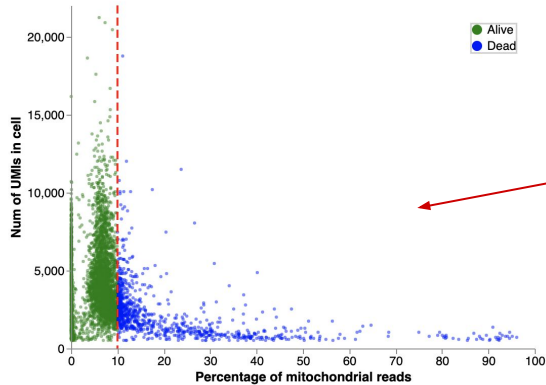
Statistics	# before	# after	% changed
Estimated number of cells	4386	3450	-21.341
Total number of genes	16175	16165	-0.062
Median number of genes per cell	1012	1085.5	+7.263

Answer 1 - Assessment of sample quality

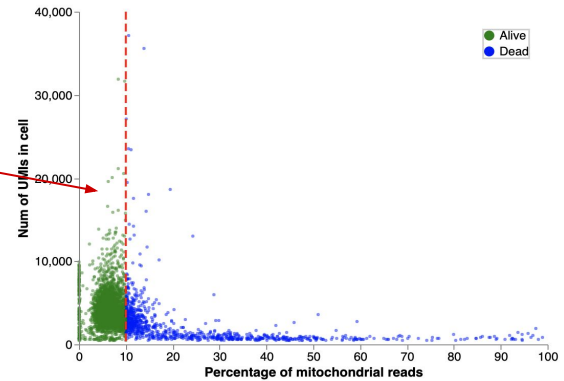
Determine whether the two samples differ significantly in quality.

a) Do they contain similar proportions of dead cells?

convalescent



acute



Data spread is similar

% of cells filtered out is similar

Statistics	# before	# after	% changed
Estimated number of cells	4479	3553	-20.674
Total number of genes	16031	16021	-0.062
Median number of genes per cell	1049	1094	+4.290
Median UMI counts per cell	3551	3877	+9.181

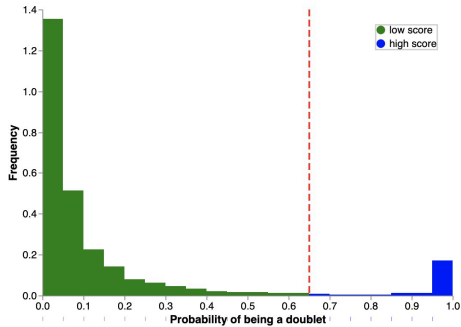
Statistics	# before	# after	% changed
Estimated number of cells	4386	3450	-21.341
Total number of genes	16175	16165	-0.062
Median number of genes per cell	1012	1085.5	+7.263
Median UMI counts per cell	3489	3954.5	+13.342

Answer 1 - Assessment of sample quality

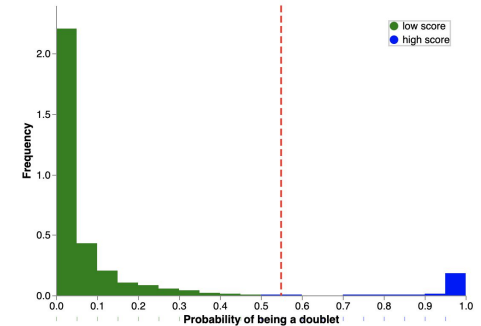
Determine whether the two samples differ significantly in quality.

b) Do they contain similar proportions of doublets?

convalescent



acute



Data spread is similar

% of cells filtered out is similar

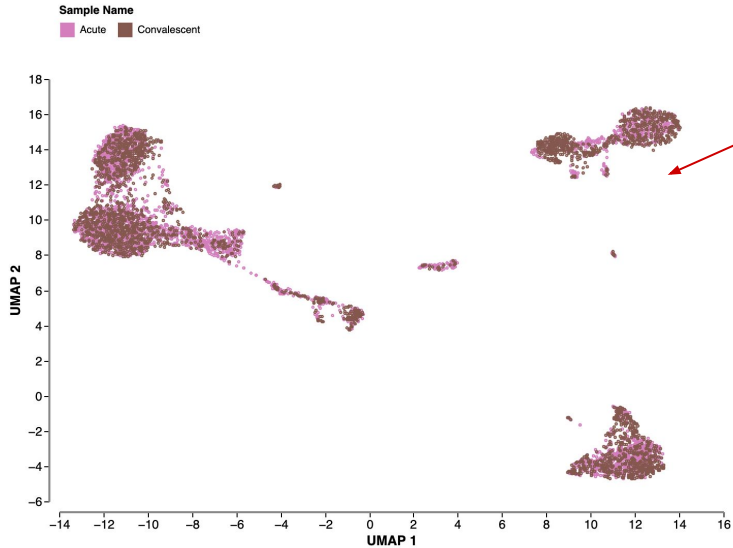
Statistics	# before	# after	% changed
Estimated number of cells	2754	2536	-7.916
Total number of genes	16149	16118	-0.192
Median number of genes per cell	1196.5	1173.5	-1.922
Median UMI counts per cell	3998	3897	-2.526

Statistics	# before	# after	% changed
Estimated number of cells	3406	3171	-6.900
Total number of genes	16009	15989	-0.125
Median number of genes per cell	1105	1089	-1.448
Median UMI counts per cell	3870.5	3747	-3.191

Answer 1 - Assessment of sample quality

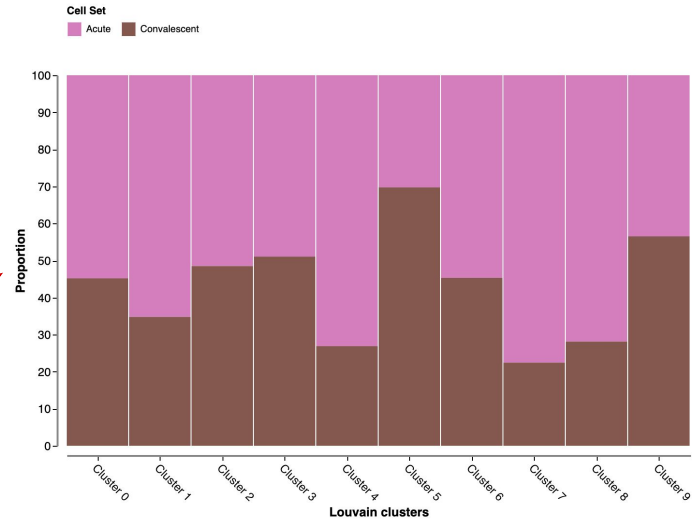
Determine whether the two samples differ significantly in quality.

c) Is the data well integrated?

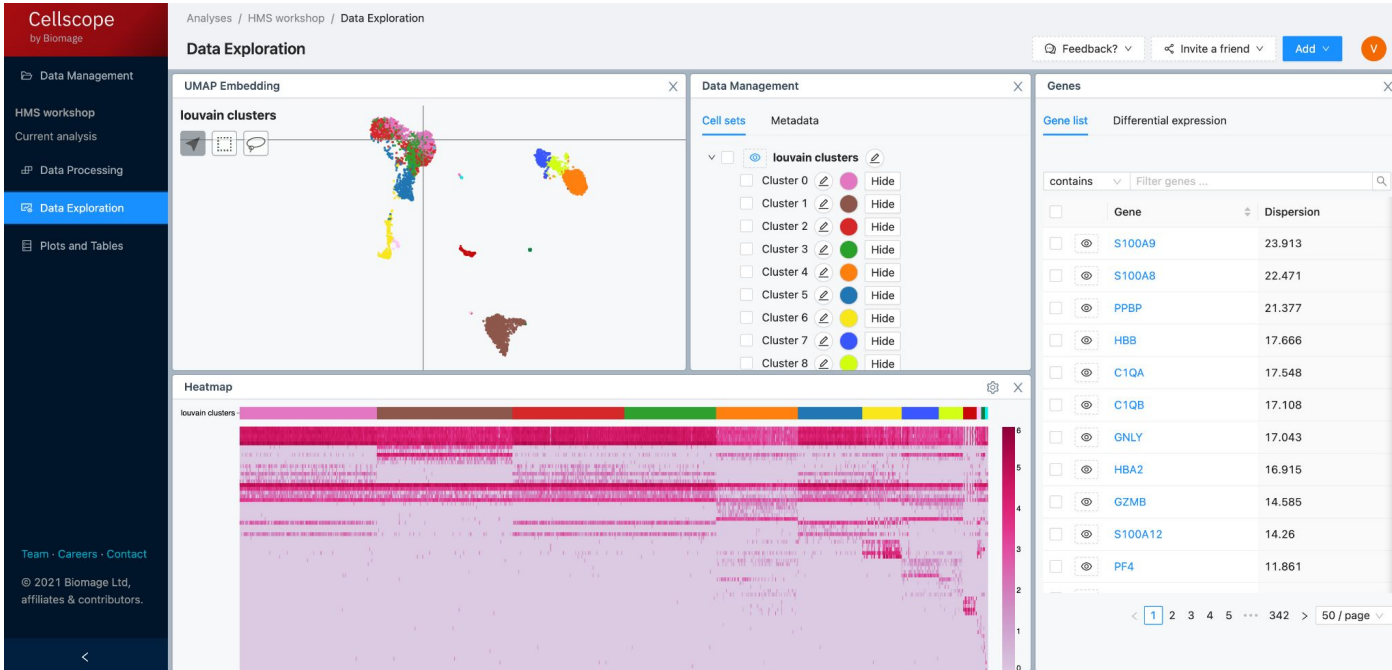


The two samples show similar distribution on the UMAP embedding.

The frequency plot shows that all clusters contain cells from both samples.



Data Exploration module



- fast, interactive data visualisation

- cluster annotation

- view gene expression in just a few clicks

- calculate differential expression between clusters or groups

-> Define what cell types are present -> annotate clusters

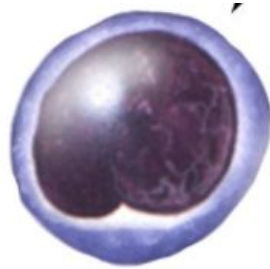
-> Find interesting things from the dataset

Exercise 2 - Dataset annotation

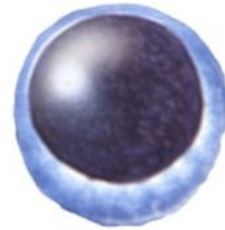


Monocytes

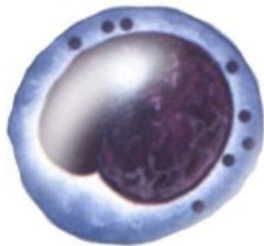
CD14+
CD16+



B cells



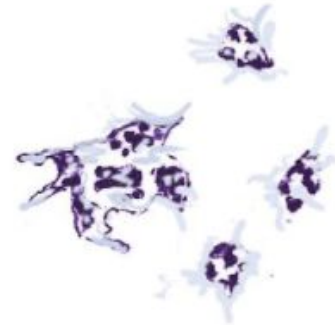
T cells



NK cells

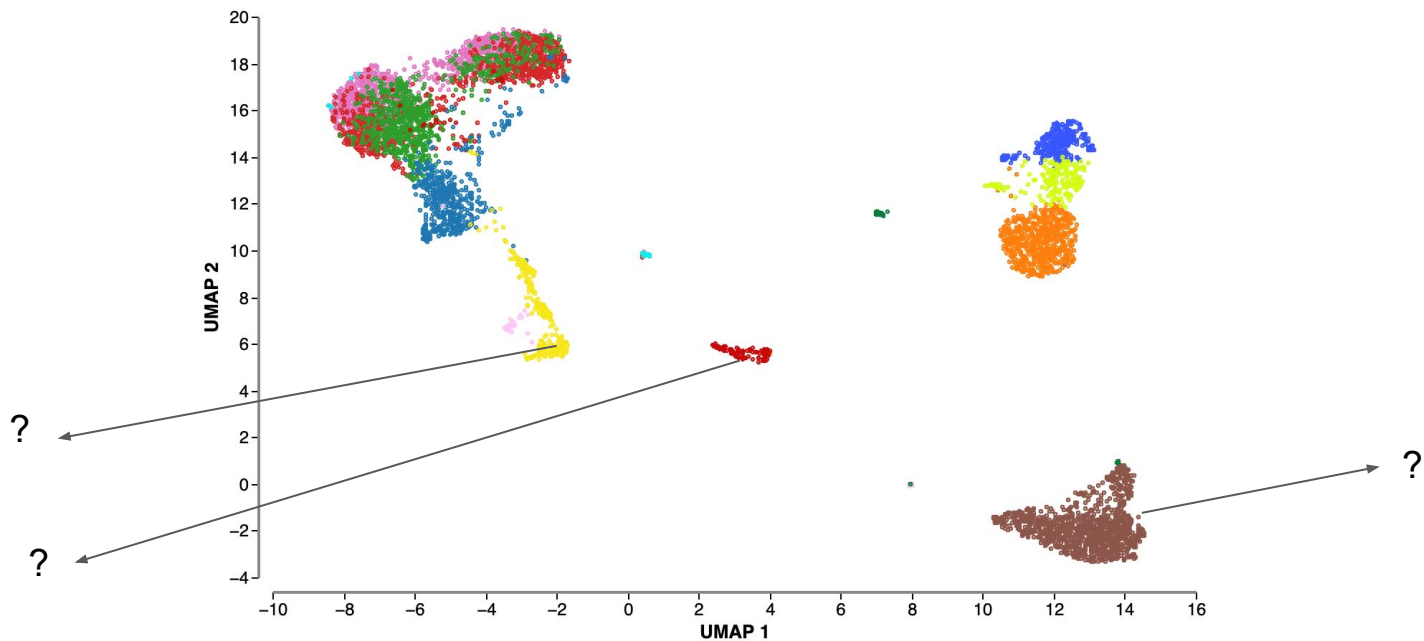


Erythrocytes

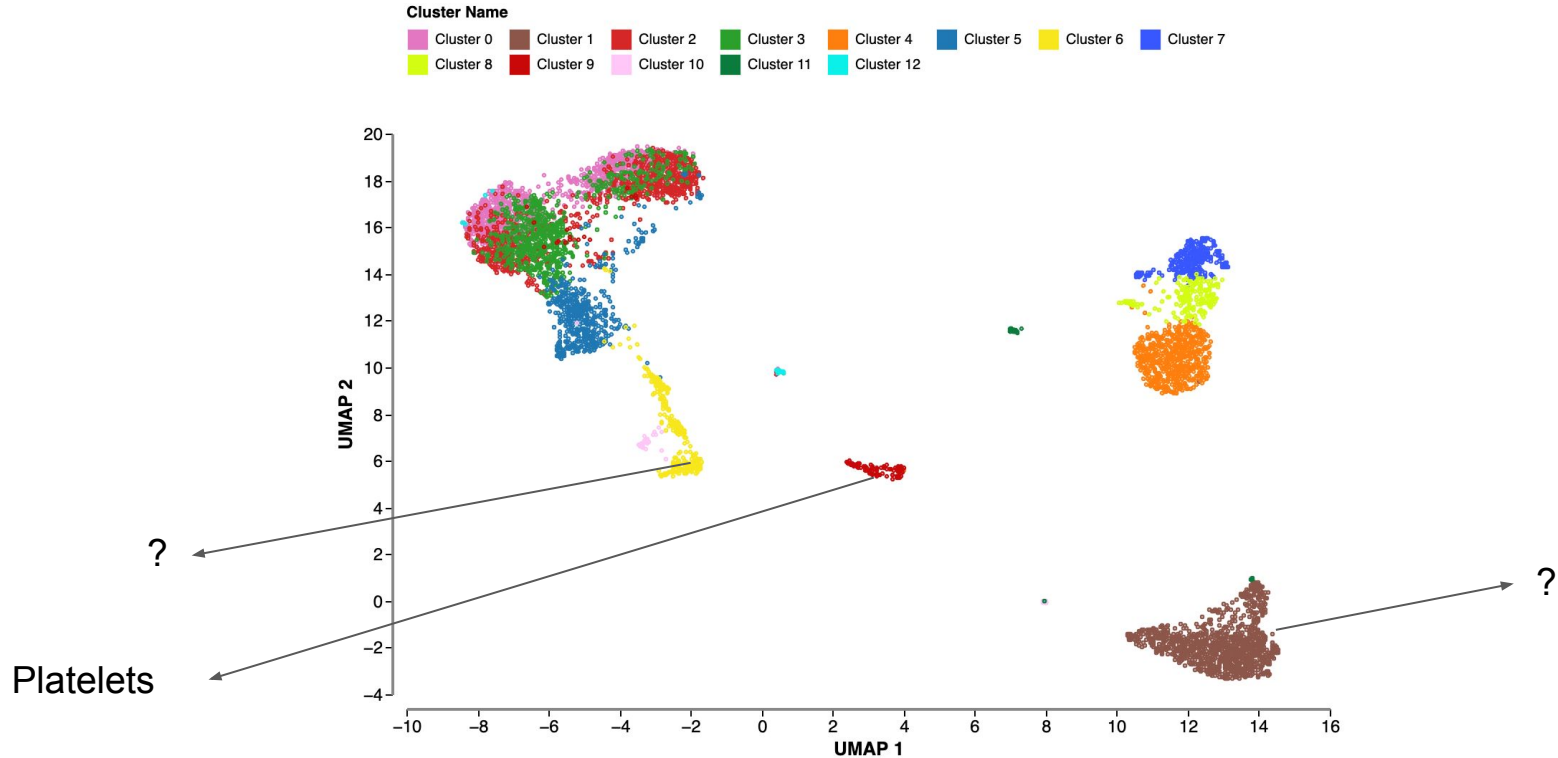


Platelets

Exercise 2 - Dataset annotation



Exercise 2 - Annotate these two unknown clusters [10mins]



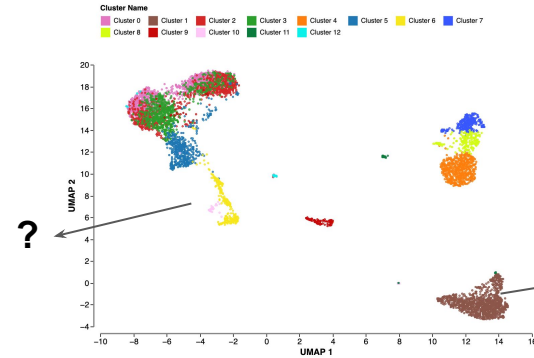
Exercise 2 - Annotate these 2 unknown clusters [10mins]

INSTRUCTIONS:

1. Go to Data Exploration module
2. For the **yellow cluster***: hover over the marker genes of this cluster in the heatmap.
3. Search for the relevant genes in the gene list.
4. For the **brown cluster***: Select the Differential Expression tab from the 'Genes' tile.
5. Select the relevant cluster; compare to all other cells; using all samples; click 'Compute'.
Look at the results (next slide).

When you have completed it, write
“Done 2” in chat.

If you have time left, annotate another cluster of
your choice.



**The cluster colors might be different but you will be able to identify the shapes*

The screenshot shows the Cellenics Data Exploration interface. The main window displays a UMAP plot of 'louvain clusters' and a heatmap below it. A red box labeled '1.' highlights the 'Data Exploration' option in the left sidebar. The heatmap shows a red box labeled '2.' over a specific gene. On the right, the 'Genes' panel is open, showing 'Differential expression' selected (red box '3.'). Below that, 'Compare cell sets within a sample/group' is selected (red box '4.'). At the bottom of the 'Genes' panel, the 'Compute' button is highlighted with a red box labeled '5.'

Exercise 2 - Annotate these 2 unknown clusters [10mins]

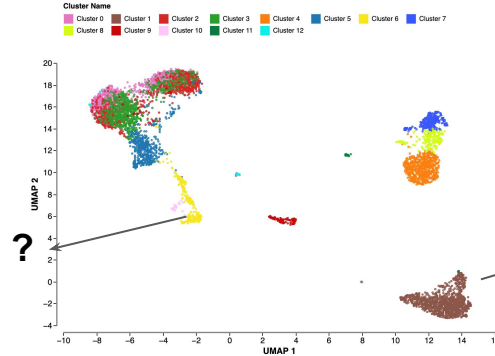
INSTRUCTIONS:

1. Go to Data Exploration module
2. For the **yellow cluster***: hover over the marker genes of this cluster in the heatmap.
3. Search for the relevant genes in the gene list.
4. For the **brown cluster***: Select the Differential Expression tab from the 'Genes' tile.
5. Select the relevant cluster; compare to all other cells; using all samples; click 'Compute'.
Look at the results (next slide).

When you have completed it, write

“Done 2” in chat.

If you have time left, annotate another cluster of your choice.



Click twice to sort descending by logFC

Click twice to visualize expression in the UMAP

Click to add genes to the heatmap

Genes

Gene list [Differential expression](#)

< Go back

Show settings

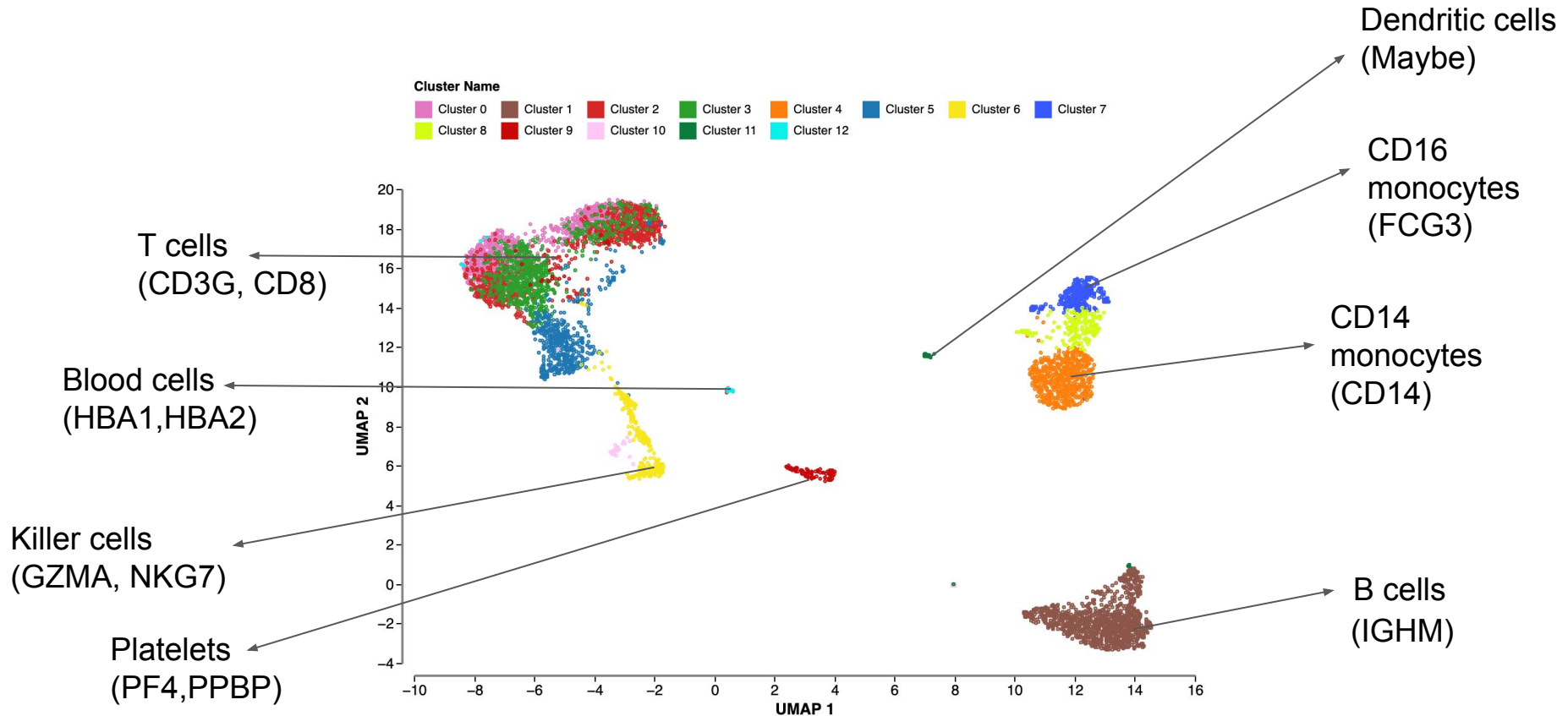
Export as CSV ...

contains Filter genes ...

Gene	adj logFC	p-value	Pct 1	Pct 2
IGHM	3.693	0	96.79	5.28
HLA-DRA	3.219	0	99.61	20.9
CD74	3.171	0	99.80	37.1
IGKC	2.975	0	65.76	7.01
HLA-DRB1	2.899	0	98.93	20.7
CD79A	2.890	0	98.73	6.23
HLA-	2.406	^	96.79	17.3

< 1 2 3 4 5 ... 342 > 50 / page

Answer 2 - Annotate the unknown clusters



Plots & Tables module

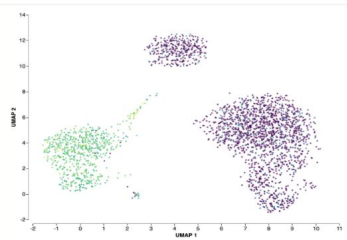
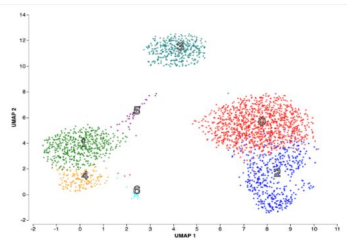
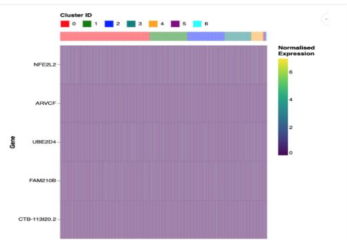
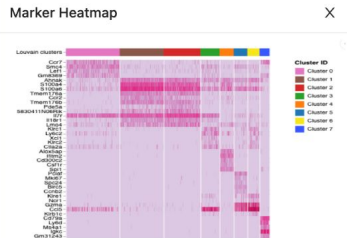
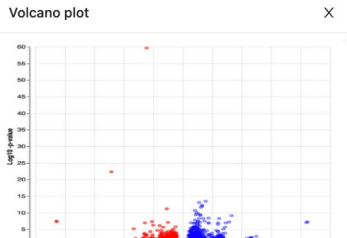

Cellscape
by Biome

Analyses / HMS workshop / Plots and Tables

Plots and Tables

Feedback? v Invite a friend v Open Existing v Create v

Recommended

- Continuous Embedding  Last updated: never
- Categorical Embedding  Last updated: never
- Custom Heatmap  Last updated: never
- Marker Heatmap  Last updated: never
- Volcano plot  Last updated: never
- Frequency Plot  Last updated: never

- range of plots pre-loaded with your data
- full customization of plot styling
- export differentially expressed gene lists
- export publication-quality figures

-> Create and customise figures for publication

Exercise 3 - Visualise differences between cell types [5mins]

Now that we have annotated one or more clusters, let's create some plots to sustain our findings.

Create and download:

- a Violin Plot, and
 - a Continuous Embedding plot
- ...to justify the annotation.

INSTRUCTIONS:

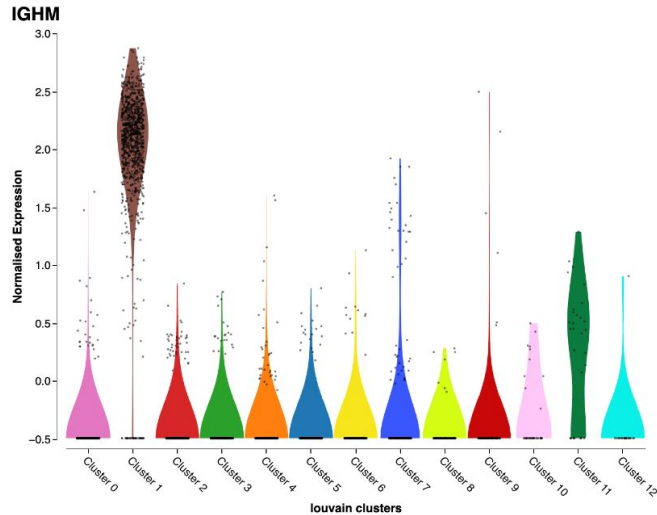
- Go to the Plots and Tables module
- Select the Violin plot (or continuous embedding)
- In the “Gene selection” menu, input one of the genes that you found in the Differential Expression results (e.g. IGHM)
- Download the plot using the “...” menu.

When you have completed it, write “Done 3” in chat.

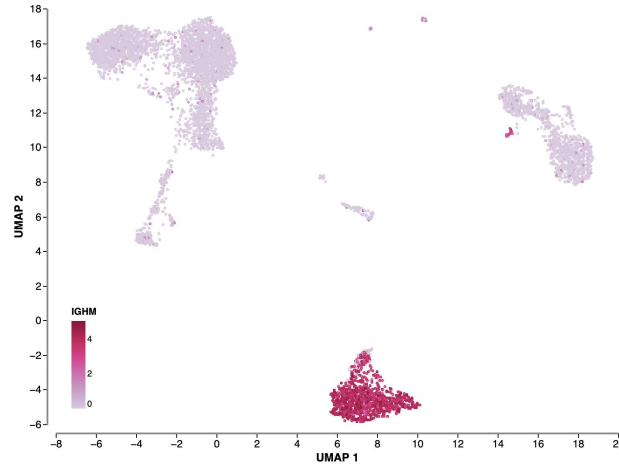
The screenshot displays the Cellenics web interface. On the left is a dark sidebar with a menu where 'Plots and Tables' is highlighted in blue, with a red box and the number '1.' next to it. The main area shows a grid of plots: 'Marker Heatmap', 'Volcano plot', 'Frequency Plot', and 'Violin Plot'. The 'Violin Plot' is selected and enlarged, showing a violin plot for the gene 'Lyz2' with a red box and the number '2.' next to it. Below this is another plot for 'IGHM' with a red box and the number '4.' next to it. On the right side, a 'Gene Selection' menu is open, showing a search bar with 'IGHM' entered and a 'Search' button, with a red box and the number '3.' next to it. The bottom of the interface shows a 'Preview' section with a violin plot for 'IGHM' across various clusters.

Answer 3 - Visualise differences between the cell types

A **violin plot** is a great idea to showcase the difference between clusters



A **continuous embedding plot** can also be used to visualise gene expression



FYI: In both plots, you can view all cells or specific samples using the “select data” menu.

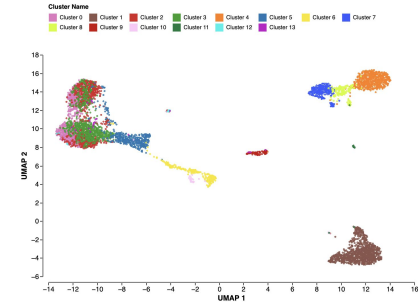
A screenshot of a software interface showing a menu for selecting data to view on an embedding. The menu is titled 'Select Data' and includes a dropdown menu with the following options: 'All', 'Samples', 'Convalescent', and 'Acute'. The 'All' option is currently selected. The interface also shows other menu items: 'Gene Selection', 'Axes and Margins', 'Colours', 'Markers', and 'Legend'.

Exercise 4 - Investigate the unknown population [20mins]

a) Investigate the population of cells that is excluded in filter 4.

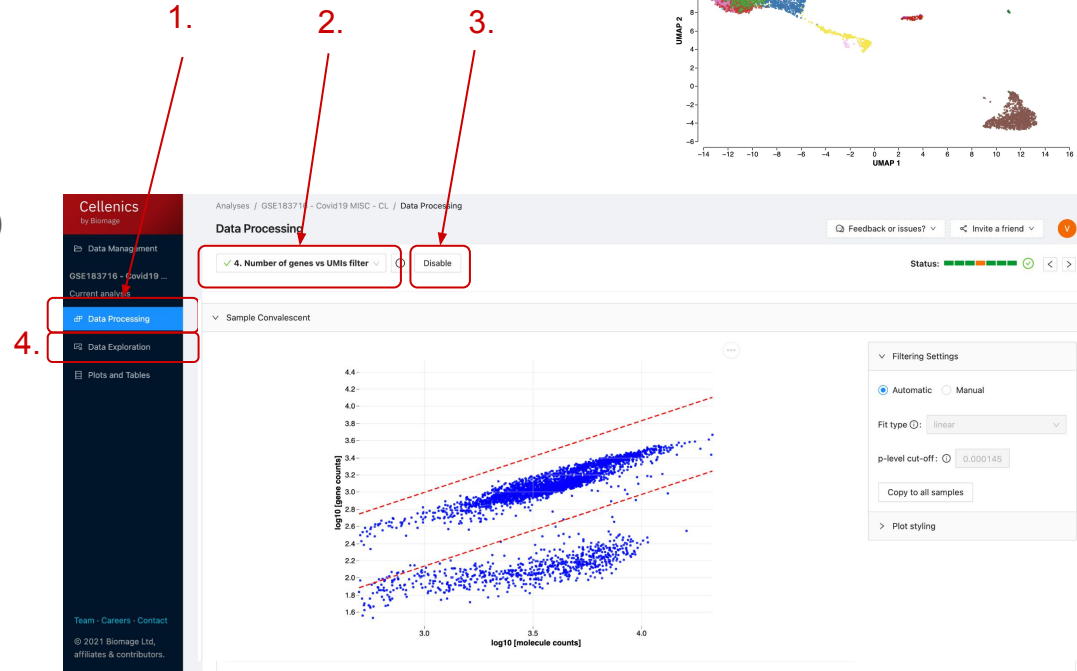
What are they? Create a plot to illustrate your conclusion.

The old embedding:



INSTRUCTIONS:

1. Go to Data Processing module
2. Navigate to step 4 (genes vs UMIs filter)
3. Disable this filter - you'll need to re-run the pipeline
4. Look in Data Exploration for a new cluster, use DE tool to identify them
5. Go to Plots & Tables - create a plot to illustrate your findings



When you have completed it, write

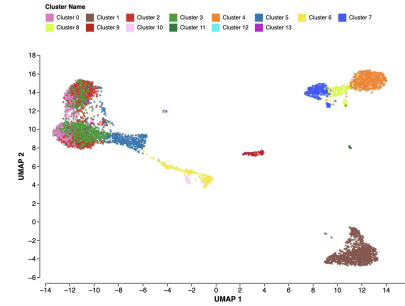
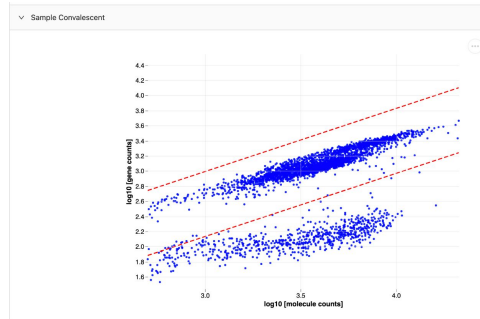
“Done 4” in chat.

Answer 4 - Investigate the unknown population

a) Investigate the population of cells that is excluded in filter 4.

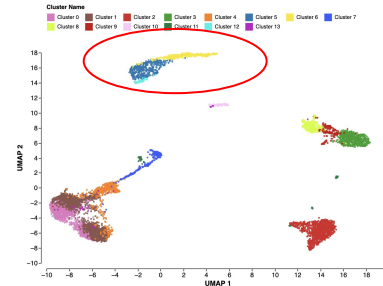
What are they? Create a plot to illustrate your conclusion.

Exclude the population:



Include the population:

This filter is disabled. You can still modify and save changes, but the filter will not be applied to your data.



Answer 4 - Investigate the unknown population

a) Investigate the population of cells that is excluded in filter 4.

What are they? Create a plot to illustrate your conclusion.



Genes

Gene list [Differential expression](#)

< Go back

[Hide settings](#)

new cluster vs. Background in All

[Export as CSV ...](#)

contains Filter genes ...

<input type="checkbox"/>	Gene	logFC	adj p-value	Pct 1	Pct 2
<input type="checkbox"/>	HBB	7.8875	0	100	45.111
<input type="checkbox"/>	HBA2	7.2446	0	100	17.605
<input type="checkbox"/>	HBA1	6.6572	0	100	11.795
<input type="checkbox"/>	ALAS2	2.5627	0	80.7646	1.2287
<input type="checkbox"/>	SLC25A37	2.2813	0	94.9821	17.447
<input type="checkbox"/>	HBM	2.2712	0	82.3178	0.6319
<input type="checkbox"/>	YBX3	1.9927	0	89.3668	27.031
<input type="checkbox"/>	SLC25A39	1.7693	0	83.3931	12.954

< 1 2 3 4 5 ... 342 > 50 / page

HBB = Hemoglobin subunit beta
HBA2 = Hemoglobin subunit alpha 2
HBA1 = Hemoglobin subunit alpha 1

Red Blood Cells

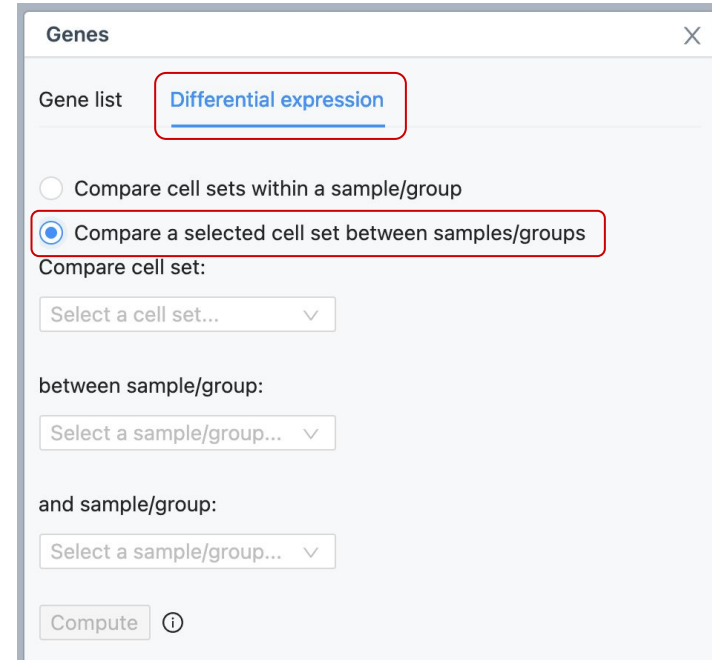
Exercise 5 - Find differences between the samples [10mins]

Determine if the two samples are from the same patient.

INSTRUCTIONS:

1. Perform Differential Expression, comparing 'All' cells between the two samples.
2. Sort the results by ascending / descending logFC.
3. Explore the top 5 genes.
4. Generate a plot (e.g. volcano plot) to illustrate your findings.

Write Done 5 in the chat



The screenshot shows a 'Genes' window with a close button (X) in the top right corner. Below the title bar, there is a 'Gene list' section with a dropdown menu currently set to 'Differential expression'. Below this, there are two radio button options: 'Compare cell sets within a sample/group' (unselected) and 'Compare a selected cell set between samples/groups' (selected). Below the selected option is a 'Compare cell set:' dropdown menu with the text 'Select a cell set...'. Further down, there is a 'between sample/group:' dropdown menu with the text 'Select a sample/group...'. Below that is an 'and sample/group:' dropdown menu with the text 'Select a sample/group...'. At the bottom left, there is a 'Compute' button, and at the bottom right, there is an information icon (i).

Answer 5 - Find differences between the samples

Determine if the two samples are from the same patient.

The Convalescent sample expresses various Y-linked genes (e.g RPS4Y1).

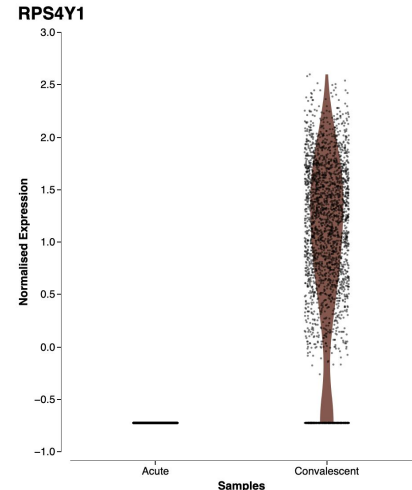
These Y-linked genes are only expressed in the Convalescent sample and not in the acute sample.

These findings lead us to believe that the two samples actually come from two different patients: the Acute sample is from a female and the Convalescent sample is from a male.

DE results, ordered by descending logFC.

Acute vs. Convalescent in All					
Gene	logFC	adj p-value	Pct 1	Pct 2	AUC
RPS4Y1	-1.7745	0	0	84.9763	0.0751

Violin plot showing RPS4Y1 expression in the acute and convalescent samples.

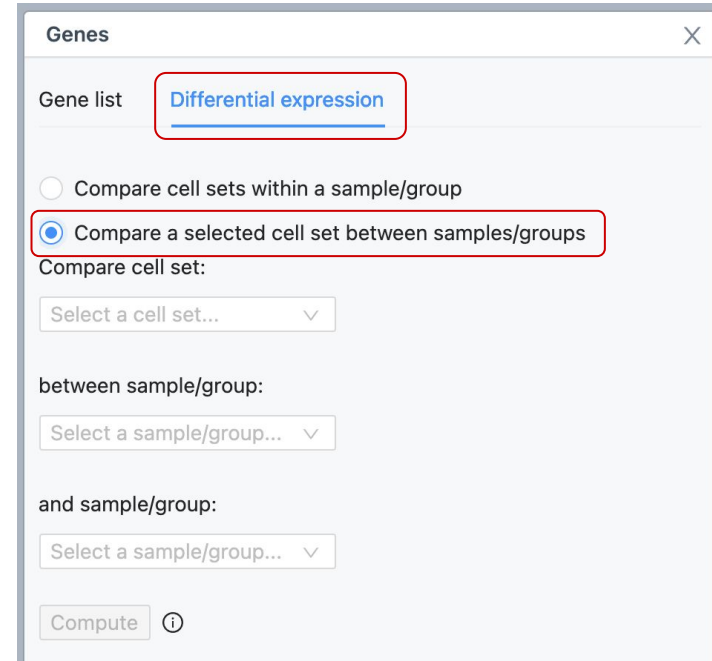


Homework - Find differences between the samples

Find differences in the Killer cell cluster (yellow cluster from exercise 2) between the acute and convalescent samples.

INSTRUCTIONS:

1. Perform Differential Expression, comparing the Killer cell cluster between the two samples.
2. Sort the results by ascending / descending logFC.
3. Explore the top 5 genes.
4. Generate a plot (e.g. volcano or violin plot) to illustrate your findings.



The screenshot shows a 'Genes' window with a 'Gene list' dropdown set to 'Differential expression'. Below this, there are two radio button options: 'Compare cell sets within a sample/group' (unselected) and 'Compare a selected cell set between samples/groups' (selected). Under the selected option, there are three dropdown menus: 'Compare cell set:' (set to 'Select a cell set...'), 'between sample/group:' (set to 'Select a sample/group...'), and 'and sample/group:' (set to 'Select a sample/group...'). At the bottom, there is a 'Compute' button and an information icon.

Data upload

To upload your own dataset

INSTRUCTIONS:

1. Go to Data Management module
2. Create and name a new Project
3. Click “Add samples” to upload the files.
4. Click “Add metadata” to allocate metadata to each sample

The screenshot shows the Cellscope Data Management interface. The left sidebar has a 'Data Management' button highlighted with a red box and labeled '1.'. The main content area shows a 'Projects' list with a 'Create New Project' button highlighted with a red box and labeled '2.'. The 'Project Details' section for 'Lung gamma delta T cells' shows 'Add samples' and 'Add metadata' buttons highlighted with red boxes and labeled '3.' and '4.' respectively. Below these buttons is a table of samples with columns for Sample, Barcodes.csv, Genes.csv, Matrix.mtx, Species, and Genotype.

Sample	Barcodes.csv	Genes.csv	Matrix.mtx	Species	Genotype
KO1	Uploaded	Uploaded	Uploaded	Mouse	KO
WT1	Uploaded	Uploaded	Uploaded	Mouse	WT
WT2	Uploaded	Uploaded	Uploaded	Mouse	WT

Supported file types

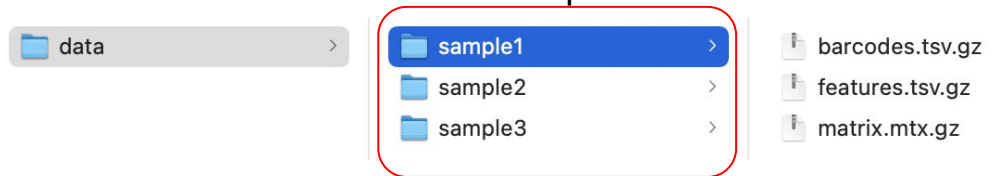
Datasets generated using 10X Chromium technology

Use the raw (unfiltered) count matrices that are output from Cell Ranger - These are usually stored in a folder named 'raw_feature_bc_matrix'.

For each sample you need 3 files:

1. `features.tsv` or `features.tsv.gz` or `genes.tsv` or `genes.tsv.gz`
2. `barcodes.tsv` or `barcodes.tsv.gz`
3. `matrix.mtx` or `matrix.mtx.gz`

Include each of these 3 files in a folder that is named with the sample ID:



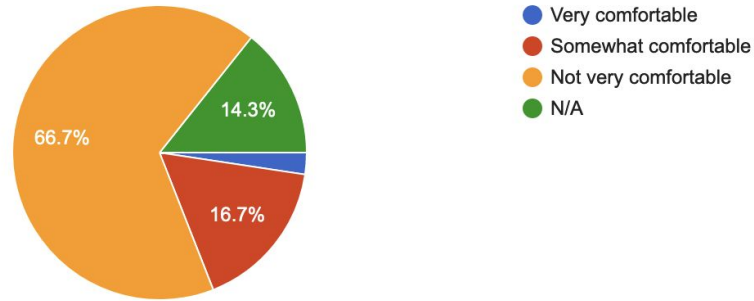
Upload the folders (named with the sample IDs) to Cellenics

Need help? Email: vicky@biomage.net

Summing up

How comfortable are you with single cell data analytics?

42 responses



What we hope to have achieved in this workshop:

- That you understand how Cellenics can support your single cell analysis
- That you feel comfortable exploring a dataset in Cellenics
- That you know how to get help with your analysis in Cellenics

OUR GOAL

We will help you to get your analysis done!

ACTIONS:

1. Upload your own data or any dataset that you are interested in exploring.
2. Book a free 45-minute meeting with Vicky (biologist) and Oliver (bioinformatician) to discuss your analysis needs:

<https://calendly.com/vicky-morrison/cellenics-onboarding>

vicky@biomage.net and oliver@biomage.net

Thank you!

References

- Centers for Disease Control and Prevention. Multisystem Inflammatory Syndrome (MIS) <https://www.cdc.gov/mis/index.html>. Page last reviewed February 24, 2021.
- Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002 Jan 1;30(1):207-10
- Giacalone, M., Scheier, E., & Shavit, I. (2021). Multisystem inflammatory syndrome in children (MIS-C): a mini-review. International Journal of Emergency Medicine, 14(1), 1-4.
- Henderson, L. A., & Yeung, R. S. (2021). MIS-C: early lessons from immune profiling. Nature Reviews Rheumatology, 17(2), 75-76.
- Kleiveland, C. R. (2015). Peripheral blood mononuclear cells. The impact of food bioactives on health, 161-167.
- National Cancer Institute - <https://www.cancer.gov/>
- Peripheral Blood | Handbook. <https://www.miltenyibiotec.com/US-en/resources/macs-handbook/human-cells-and-organs/human-cell-sources/blood-human.html>
- The innate and adaptive immune landscape of SARS-CoV-2-associated multisystem inflammatory syndrome in children (MIS-C) from acute disease to recovery. (GEO accession GSE183716)
- Weakley, S. M., Wang, H., Yao, Q., & Chen, C. (2011). Expression and function of a large non-coding RNA gene XIST in human cancer. World journal of surgery, 35(8), 1751-1756.

Homework answer - Find differences between the samples

b) Find differences in the Killer cell cluster (yellow cluster from exercise 2) between the acute and convalescent samples.

There are some really interesting immunological genes!

Granzyme is involved in the killing of virally-infected cells

Gene	logFC	adj p-value	Pct 1	Pct 2
GNLY	-2.074	5.8184e-12	46.0123	94.285
CCL4	-1.5134	4.9096e-21	24.5399	75.714
GZMB	-1.4853	5.8184e-12	43.5583	85.714
RPS4Y1	-1.4721	7.3392e-35	0	73.571
SPON2	-1.233	8.9661e-13	28.2209	72.142
PRF1	-1.2131	3.3308e-15	61.3497	91.428

Gene	logFC	adj p-value	Pct 1	Pct 2
LTB	1.2767	4.3857e-9	65.6442	35.714
CD52	0.9932	4.6798e-8	87.1166	62.142
CD3D	0.8539	0.0000098105	74.8466	40.714
NFKBIA	0.8537	2.0777e-11	59.5092	17.142
IL7R	0.8443	0	55.2147	27.857
RPS4X	0.8184	2.2481e-21	100	97.142

A negative regulator (inhibitor) of the NFkB pathway. NFkB is a transcription factor that drives production of inflammatory cytokines.

