Get relevant post-clustering analyses quickly

Samuel Granjeaud,

PhD, Research Engineer,

CRCM, Inserm, France

Meeting day

- French Society of Cytometry
- CyTOF, Spectral, Data analysis
- [Web](#)

CytoBatchNorm

O. Molendi-Coste

3

Post-clustering analyses

- Clusters or populations => features
  - cell counts in every sample
  - MFI of selected markers in every sample
- Identify the clusters that have a feature that makes a difference according to the question
- Time consuming and error prone if not automated
- Few software to perform an automated analysis, e.g. R/diffcyt, or commercial software

Key points

- Data quality is important
  - How to control it?

- Results reporting is important
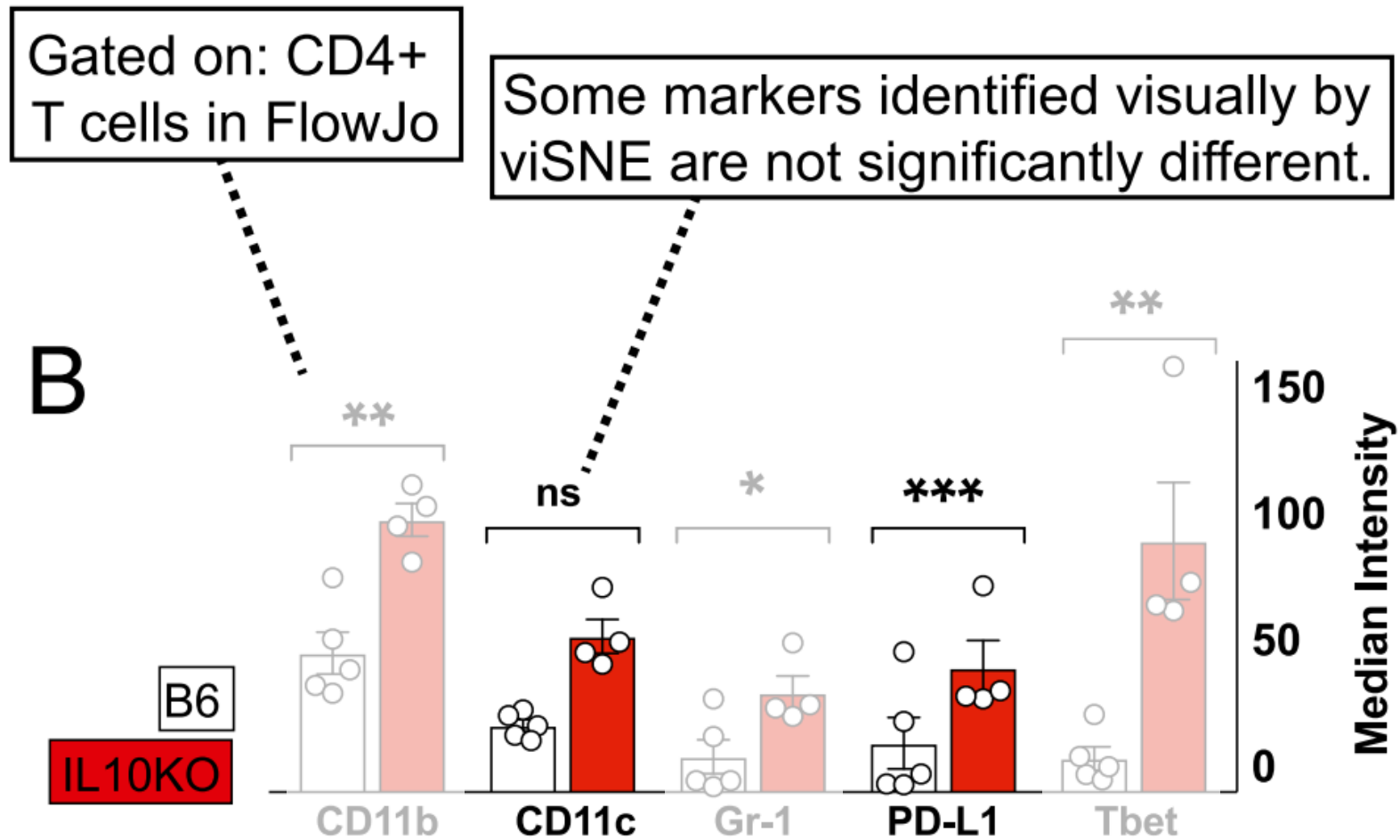 - Underestimated

The right position?

Fig. 4B, Kimball et al., 2018

Errors in figures

- automate p-values drawing
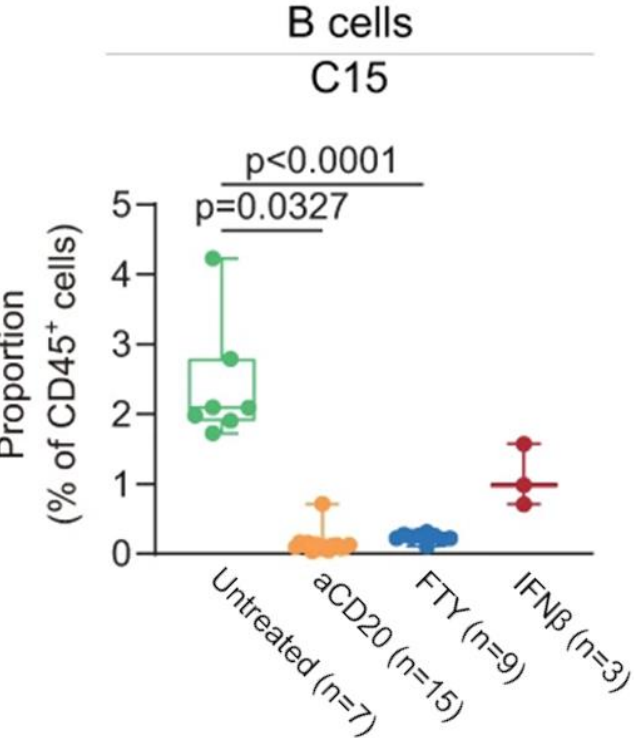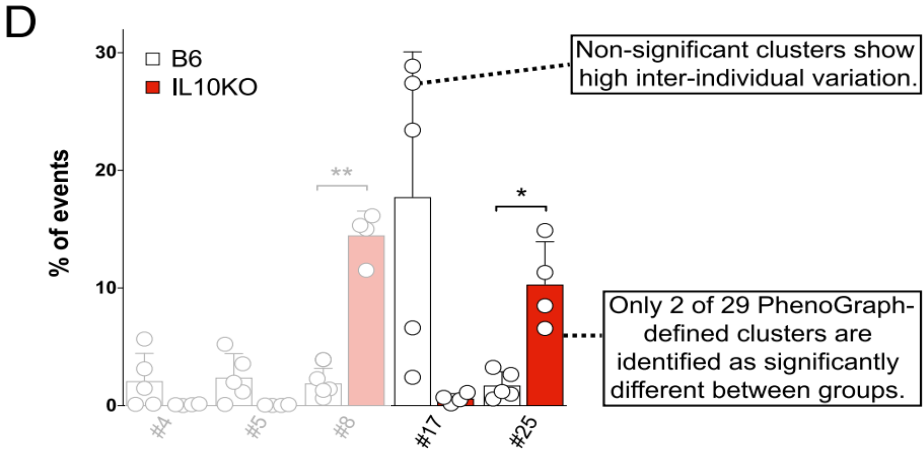- automate points labelling

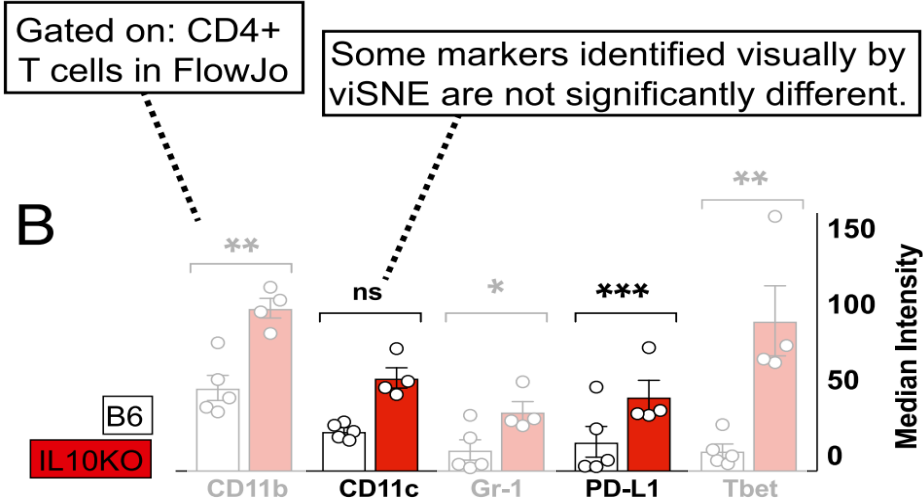Fig. 1D, Wang, 2023

Fig.3D, Kimball, 2018

Fig. 4B, Kimball, 2018

Objectives

- Control quality of features
- Reliable graphics
- Differential Abundance (percentage)
- Differential State (MFI)

- Import
  - cluster features and FCS annotations
  - from various software
- Graphical User Interface

analycyte

- Control quality of features
- Differential Abundance (percentage)
- Differential State (MFI)
- Reliable graphics

- Import
  - cluster features and FCS annotations
  - from **OMIQ, FlowJo, text files**
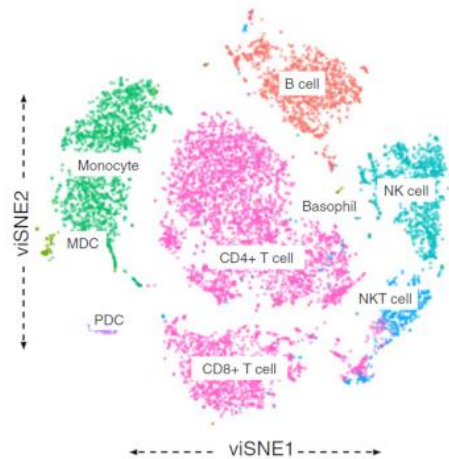- Graphical User Interface (**R/Shiny**)

ready-to-use reports

E. Lohmann

A-S. Chrétien

Dataset

Comprehensive innate immune profiling of chikungunya virus infection in pediatric cases

Daniela Michlmayr[1,†], Theodore R Pak[2,†], Adeeb H Rahman[2,3], El-Ad David Amir[2,3], Eun-Young Kim[4], Seunghee Kim-Schulze[2,3], Maria Suprun[5], Michael G Stewart[4], Guajira P Thomas[4], Angel Balmaseda[6], Li Wang[2], Jun Zhu[2], Mayte Suaréz-Fariñas[2,5], Steven M Wolinsky[4], Andrew Kasarskis[2] & Eva Harris[1,*]

We performed whole-blood RNA-seq, 37-plex **mass cytometry** of peripheral blood mononuclear cells (**PBMCs**), and serum cytokine measurements **of acute- and convalescent-phase** samples obtained from **42 children** naturally infected with CHIKV. Semi-supervised classification and clustering of single-cell events into **57 sub-communities** of canonical leukocyte phenotypes revealed a monocyte-driven response to acute infection, with the greatest expansions in "intermediate" **CD14++CD16+ monocytes** and an **activated subpopulation of CD14+ monocytes**. Increases in acute-phase CHIKV envelope protein E2 expression were highest for monocytes and dendritic cells.

10

GUI

QC report

QC report

1 Features summary

| Samples informations | Metacluster informations | Markers informations |

Sample table with information on conditions (**cond**), batch (**acqdate**), ncells (**Number of Cells**)

| Copy | CSV | Excel | Print |

Show [ ] entries

Search: [ ]

percent.cluster	Number.of.Cells	fcs_id	filename
All	All	All	All
82.76	7944	1	160406_EHA001_1758_1_Patients_Acute.fcs.astrola
89.66	13727	2	160406_EHA001_1758_1_Patients_Conv.fcs.astrola
89.66	12080	3	160406_EHA001_1760_1_Patients_Acute.fcs.astrola
93.1	11500	4	160406_EHA001_1760_1_Patients_Conv.fcs.astrola
86.21	6981	5	160406_EHA001_1773_1_Patients_Acute.fcs.astrola

Showing 1 to 5 of 86 entries

Previous 1 2 3 4 5 … 18 Next

- **percent.cluster** : Percentage of clusters with cells (> 10 cells) in this fcs. see 1

13

QC report

2 Presence / Absence data

With a minimum of 10 per cluster.
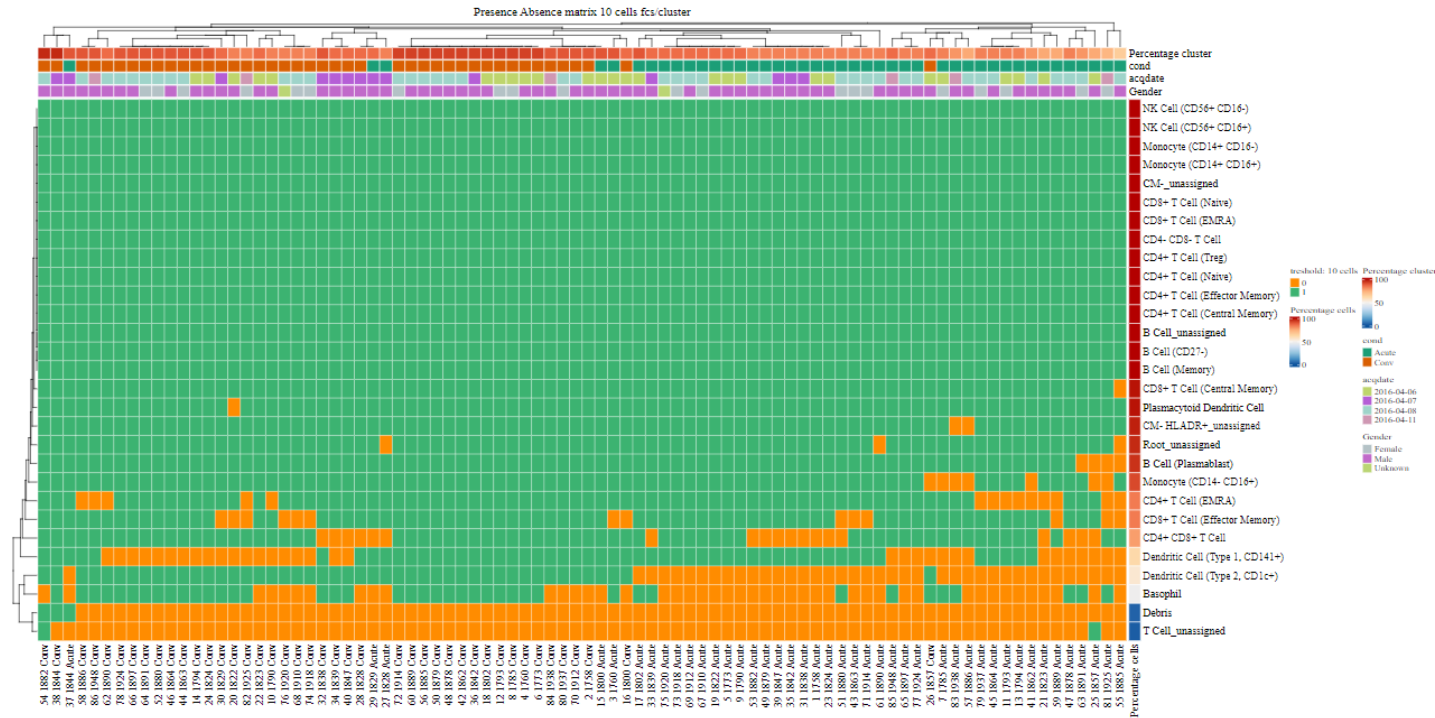
Figure 1: Heatmap Presence Absence

QC report

5 MFI's Heatmap

All Samples

Debris (0.11%)
B Cell (Memory) (2.25%)
B Cell (CD27-) (4.9%)
B Cell_unassigned (6.59%)
Monocyte (CD14+ CD16+) (1.81%)
Monocyte (CD14+ CD16-) (17.36%)
Plasmacytoid Dendritic Cell (0.66%)
Dendritic Cell (Type 2, CD1c+) (0.3%)
Dendritic Cell (Type 1, CD141+) (0.16%)
Monocyte (CD14- CD16+) (0.7%)
CM- HLADR+_unassigned (0.95%)
Root_unassigned (5.27%)
CD8+ T Cell (EMRA) (1.27%)
NK Cell (CD56+ CD16+) (6.81%)
NK Cell (CD56+ CD16-) (2.03%)
B Cell (Plasmablast) (0.45%)
Basophil (0.18%)
CM-_unassigned (2.21%)
CD4+ T Cell (Effector Memory) (0.82%)
CD4+ T Cell (EMRA) (0.23%)
CD4+ CD8+ T Cell (9.5%)
CD4+ T Cell (Naive) (9.8%)
CD4+ T Cell (Central Memory) (6.46%)
CD4+ T Cell (Treg) (2.14%)
CD8+ T Cell (Naive) (9.14%)
CD8+ T Cell (Central Memory) (1.43%)
CD8+ T Cell (Effector Memory) (0.24%)
T Cell_unassigned (0.01%)
CD4- CD8- T Cell (6.23%)

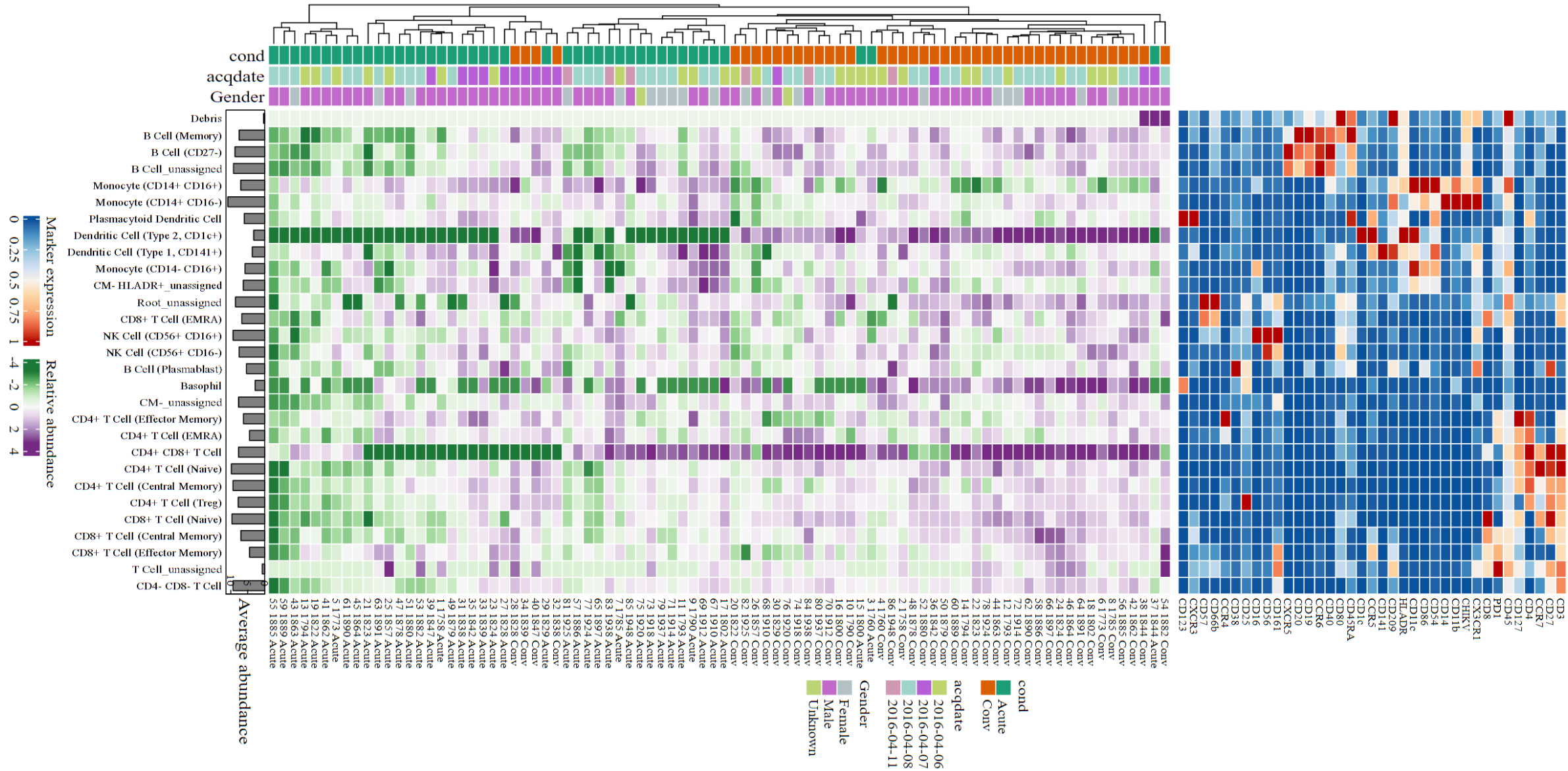Marker expression

0   0.25 0.5 0.75 1

QC report

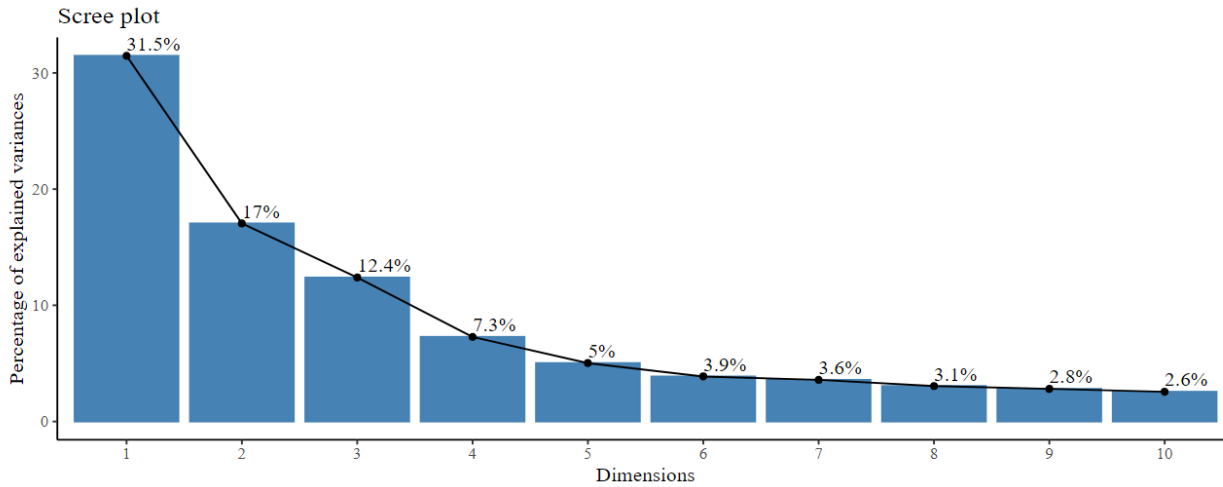# QC report

QC report

8 PCA of samples using percentages per cluster

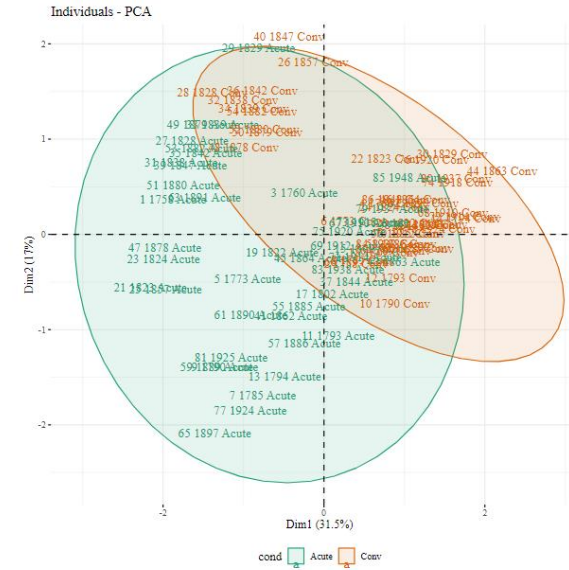### 8.1 Barplot of variance of each PC
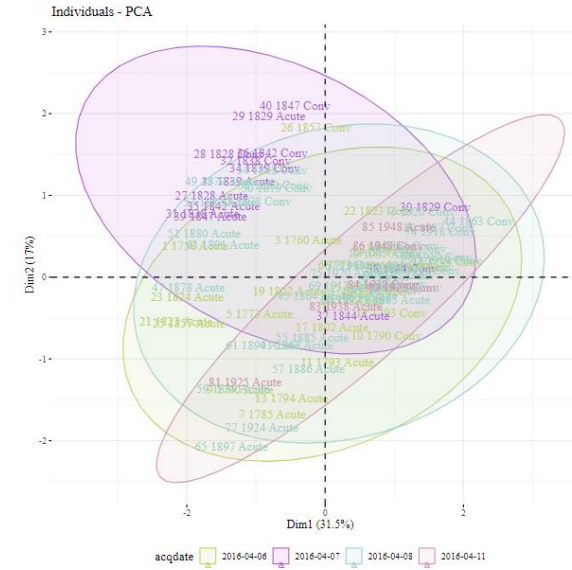
Figure 12: Dim.1 vs Dim.2 by cond
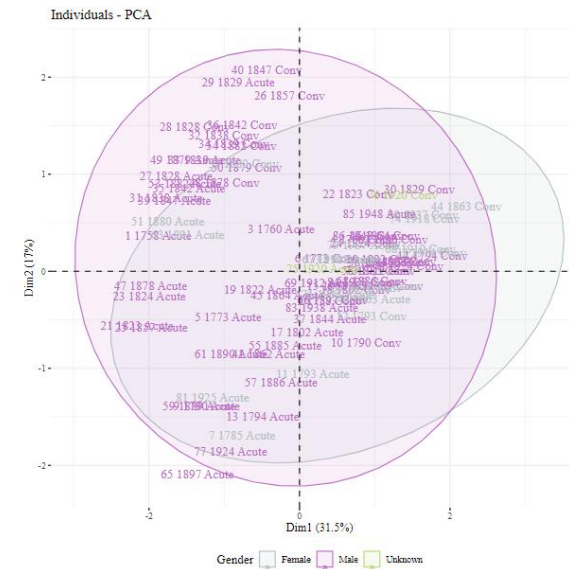

Figure 13: Dim.1 vs Dim.2 by acqdate

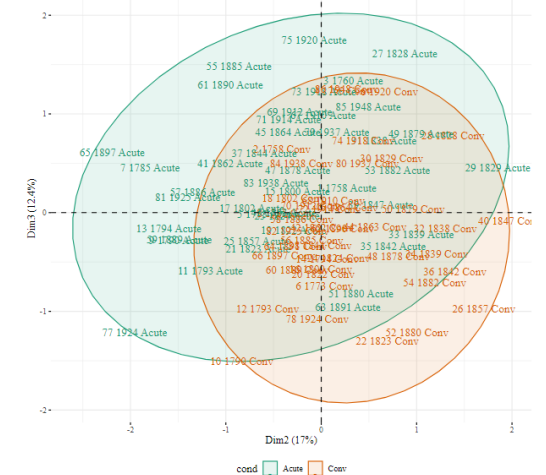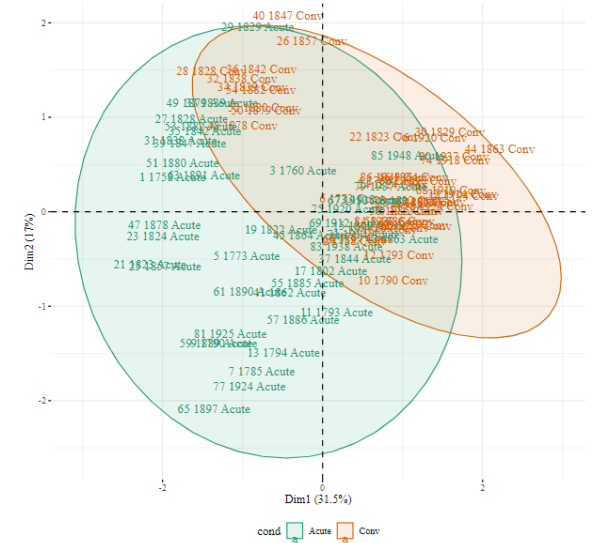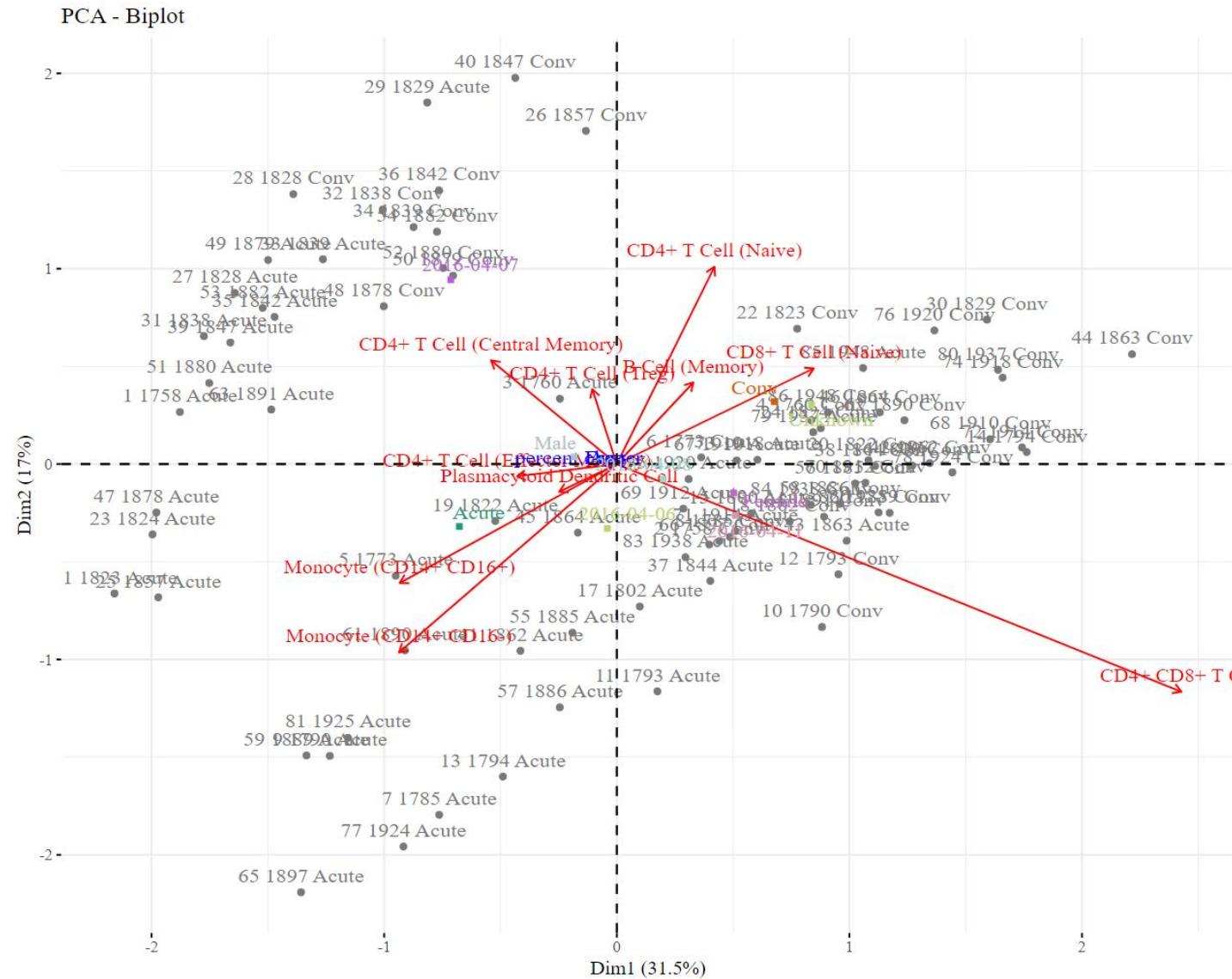Figure 14: Dim.1 vs Dim.2 by Gender

QC report

Figure 15: Dim.1 vs Dim.2

QC report

9 PCA of samples using MFI
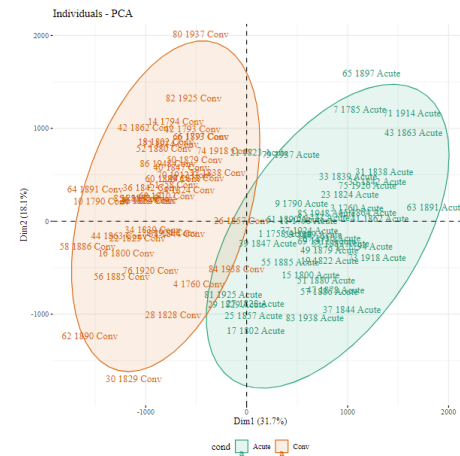
9.1 Dim.1 vs Dim.2
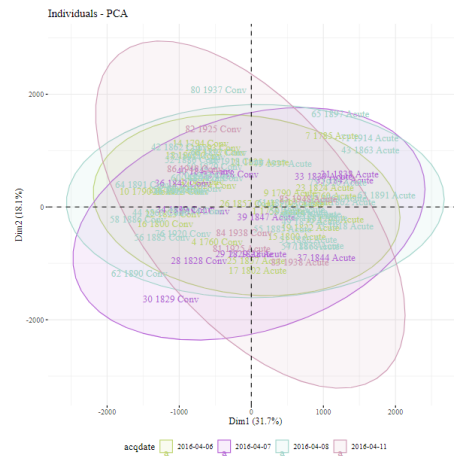

Figure 25: Dim.1 vs Dim.2 by cond
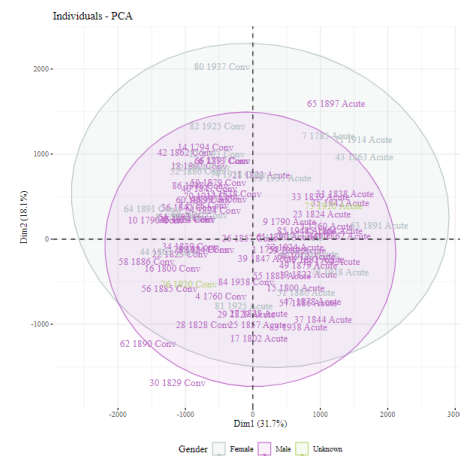

Figure 26: Dim.1 vs Dim.2 by acqdate

Figure 27: Dim.1 vs Dim.2 by Gender

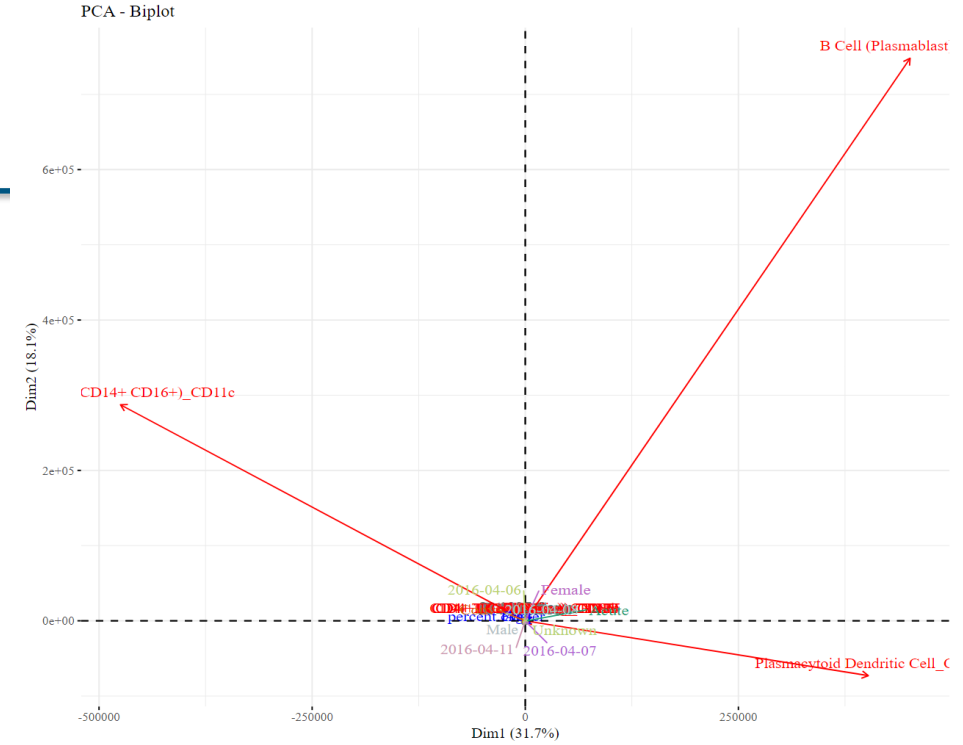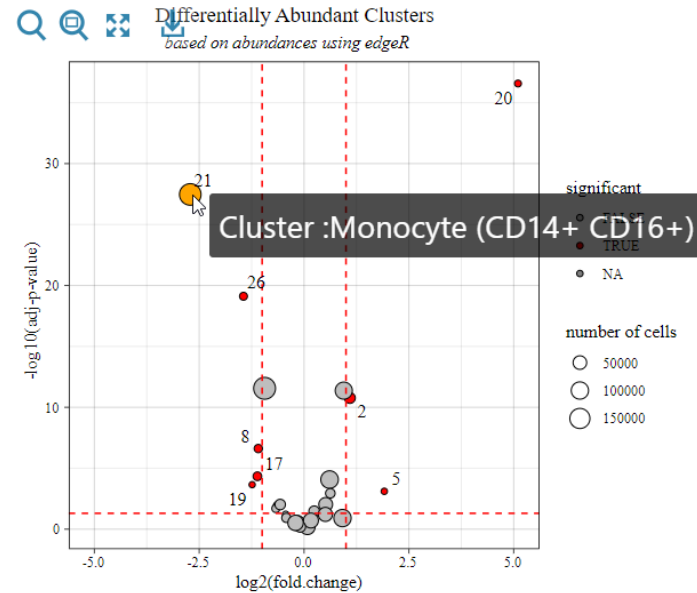

Figure 28: Dim.1 vs Dim.2
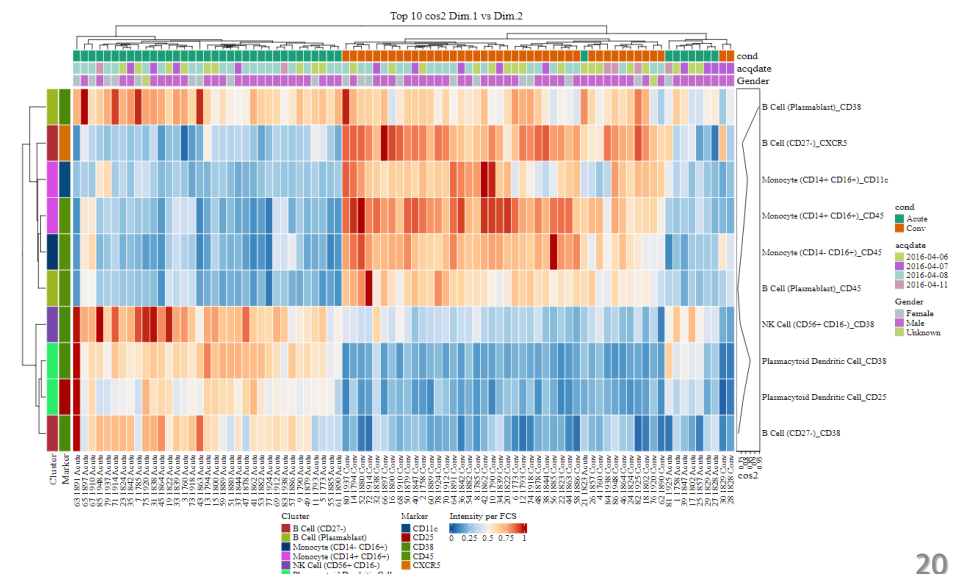
Figure 29: Heatmap Top 10 cos2 for :Dim.1 vs Dim.2

DA report

5.1.1 volcano and abundance plot

5.1.2 violin plot
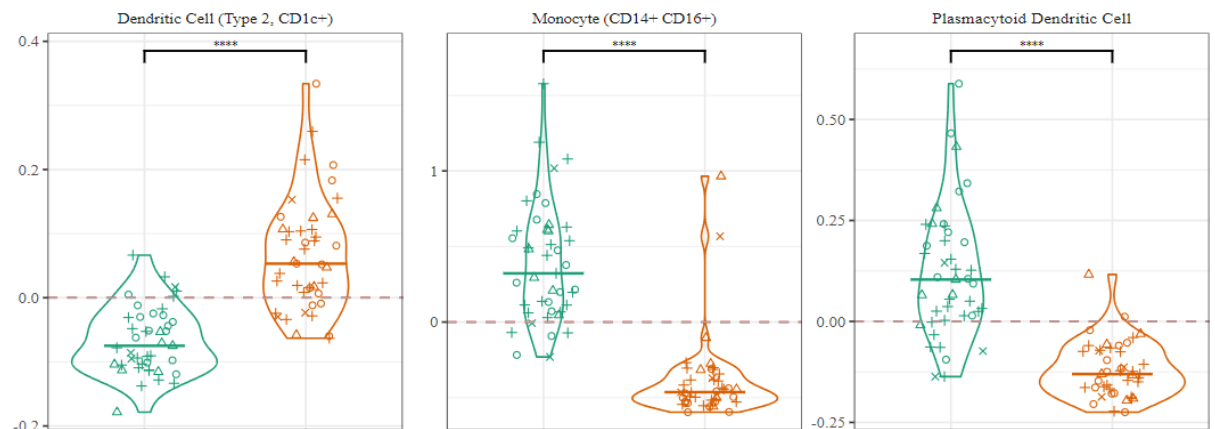
DS report

5 limma

Copy   CSV   Excel   Print

Show 10 entries

Search:

cluster_id	CellSubset	marker_id	reference	contrast	logFC	explic FC
All	All	All	All	All	All	All
22	Monocyte (CD14+ CD16-)	CD25	Acute	Conv	-2.452	-5.4?
22	Monocyte (CD14+ CD16-)	CCR7	Acute	Conv	-6.657	-100
8	CD4+ T Cell (Effector Memory)	CD45	Acute	Conv	91.224	2.8913
22	Monocyte (CD14+ CD16-)	CD209	Acute	Conv	-3.08	-8.4!
22	Monocyte (CD14+ CD16-)	CHIKV	Acute	Conv	-6.89	-118
22	Monocyte (CD14+ CD16-)	CD11c	Acute	Conv	235.168	6.2023
7	CD4+ T Cell (Central Memory)	CD45	Acute	Conv	73.722	1.5579
3	B Cell (Plasmablast)	CD45	Acute	Conv	45.216	4.0876
11	CD4+ T Cell (Treg)	CD45	Acute	Conv	67.598	2.234?
22	Monocyte (CD14+ CD16-)	CD123	Acute	Conv	-16.793	-1.1352

Showing 1 to 10 of 780 entries

Previous 1 2 3 4 5 … 78 Next

⚠ 10.34% NA results

22

analycyte

- Quality control
- Comparison 1-vs-1, all-vs-ref
- Automatically placed p-values
- Reproducible HTML reports
  - based on Quarto templates
- Interactive plots and SVG graphics

Monocyte (CD14- CD16+)_CD20

analycyte
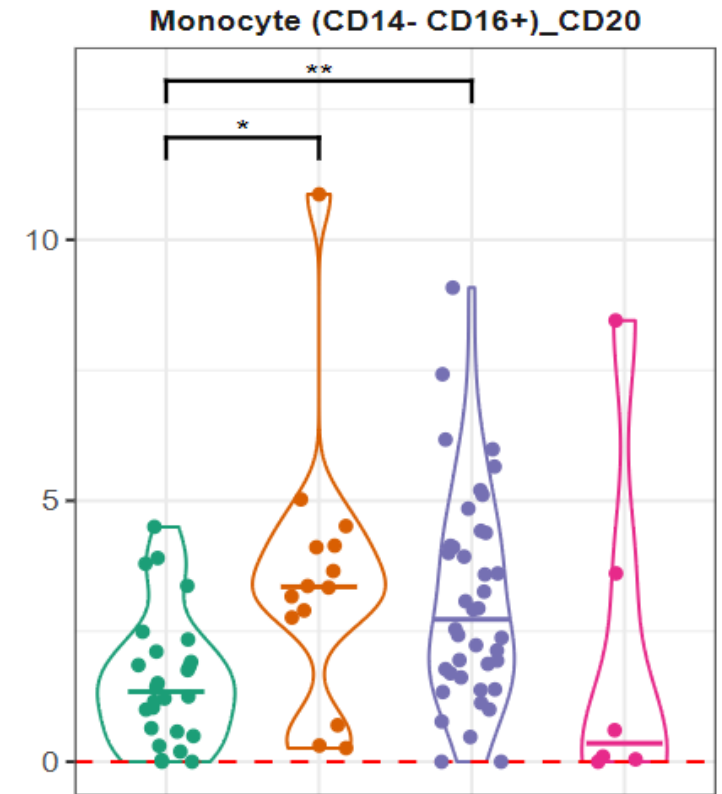
- https://i-cyto.github.io
- docker image & R packages
- demo reports
- current presentation

- documentation
- multi-variate analyses (ML)