Journal of Visualized Experiments

Visualization and Quantification of High-Dimensional Cytometry Data Using Cytofast and the Upstream Clustering Methods FlowSOM and Cytosplore

--Manuscript Draft--

Article Type:	Invited Methods Article - JoVE Produced Video
Manuscript Number:	JoVE60525R2
Full Title:	Visualization and Quantification of High-Dimensional Cytometry Data Using Cytofast and the Upstream Clustering Methods FlowSOM and Cytosplore
Section/Category:	JoVE Cancer Research
Keywords:	flow and mass cytometry analysis, R package, high-dimensional analysis, single-cell analysis, visualization tool,
Corresponding Author:	Ramon Arens UNITED STATES
Corresponding Author's Institution:	
Corresponding Author E-Mail:	R.Arens@lumc.nl
Order of Authors:	Guillaume Beyrend
	Koen Stam
	Ferry Ossendorp
	Ramon Arens
Additional Information:	
Question	**Response**
Please indicate whether this article will be Standard Access or Open Access.	Open Access (US$4,200)
Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations.	Leiden, South-Holland, The Netherlands

TITLE:

Visualization and Quantification of High-Dimensional Cytometry Data Using Cytofast and the Upstream Clustering Methods FlowSOM and Cytosplore

AUTHORS AND AFFILIATIONS:

Guillaume Beyrend[1], Koen Stam[2], Ferry Ossendorp[1], Ramon Arens[1]

[1]Department of Immunohematology and Blood Transfusion, Leiden University Medical Center, Leiden, The Netherlands

[2]Department of Parasitology, Leiden University Medical Center, Leiden, The Netherlands

Corresponding Authors:

Guillaume Beyrend (G.Beyrend@lumc.nl)

Ramon Arens (R.Arens@lumc.nl)

E-mail Addresses of Co-authors:

Koen Stam (k.stam@lumc.nl)

Ferry Ossendorp (f.ossendorp@lumc.nl)

KEYWORDS:

flow and mass cytometry analysis, R package, high-dimensional analysis, single-cell analysis, visualization tool, data mining

SUMMARY:

Cytofast is a visualization tool used to analyze output from clustering. *Cytofast* can be used to compare two clustering methods: *FlowSOM* and *Cytosplore*. *Cytofast* can rapidly generate a quantitative and qualitative overview of mass cytometry data and highlight the main differences between different clustering algorithms.

ABSTRACT:

The complexity of data generated by mass cytometry has necessitated new tools to rapidly visualize analytic outcomes. Clustering methods like *Cytosplore* or *FlowSOM* are used for the visualization and identification of cell clusters. For downstream analysis, a newly developed R package, *Cytofast*, can generate a rapid visualization of results from clustering methods. *Cytofast* takes into account the phenotypic characterization of cell clusters, calculates the cell cluster abundance, then quantitatively compares groups. This protocol explains the applications of *Cytofast* to the use of mass cytometry data based on modulation of the immune system in the tumor microenvironment (i.e., the natural killer [NK] cell response) upon tumor challenge followed by immunotherapy (PD-L1 blockade). Demonstration of the usefulness of *Cytofast* with *FlowSOM* and *Cytosplore* is shown. *Cytofast* rapidly generates visual representations of group-related immune cell clusters and correlations with immune system composition. Differences are observed in the clustering analysis, but separation between groups are visible with both clustering methods. *Cytofast* visually shows the patterns induced by PD-L1 treatment that include a higher abundance of activated NK cell subsets, expressing a higher intensity of activation markers (i.e., CD54 or CD11c).

INTRODUCTION:

Mass cytometry (cytometry by time-of-flight, or CyTOF) allows detection of a wide range of intracellular or extracellular biomarkers in millions of single cells. The high-dimensional nature of mass cytometry data necessitates certain analysis tools, such as cell clustering techniques like *SPADE*[1], *FlowMaps*[2], *FlowSOM*[3], *Phenograph*[4], *VorteX*[5], and *scaffold maps*[6]. In addition, various dimensionality reduction-based techniques have been developed (i.e., principal component analysis [PCA][7], t-distributed stochastic neighbor embedding [t-SNE][8], hierarchical stochastic neighbor embedding [HSNE][9], uniform manifold approximation and projection [UMAP][10], and diffusion maps[11]) to improve the speed, interpretation and visualization of high-dimensional datasets.

Downstream analysis of high-dimensional flow and mass cytometric data often lacks automatic processes to perform statistical tests on cluster frequency and links with clinical outcomes. Previously, we developed an R-based workflow known as *Cytofast*[12], which allows for visual and quantitative downstream analyses of clustering techniques by *Cytosplore* or *FlowSOM*.

The protocol described here clarifies the use of *Cytofast* in R and shows how to generate quantitative and qualitative heatmaps and graphs. Furthermore, it facilitates the determination of connections between observed immune phenotypes and clinical outcomes. This report also describes the analysis of a specific mass cytometry dataset using two different clustering procedures: *FlowSOM* and *Cytosplore*. By using *Cytofast* with both clustering methods, it is correspondingly shown that the activation phenotype of NK cells is influenced by PD-L1 immune checkpoint blockade.

PROTOCOL:

All animal experiments were approved by the Animal Experiments Committee of LUMC and were executed according to the animal experimentation guidelines of the LUMC in compliance with the guidelines of Dutch and European committees.

NOTE: For experimental set-up, C57BL/6 mice were subcutaneously inoculated on the right flank with the murine colon tumor MC38 at a concentration of 0.3 x 10^6 cells/200 µL of phosphate-buffered saline (PBS). After 10 days, when tumors were palpable, the mice were treated with PD-L1 blocking antibodies (clone MIH-5, 200 µg/mouse, intra-peritoneal injection) or were mock-treated. The tumors were resected 3 days later after PD-L1 injection, processed ex vivo, and analyzed by CyTOF mass cytometry using 38 markers[13].

1. Equipment and software for data analysis

NOTE: Use a computer (Windows 7 or newer) and processor I5 at 2.4 GHz or equivalent, installed memory RAM 6 GB, and 10 GB of free hard drive space. The R package *Cytofast* uses existing functions: mainly *flowCore*, *pheatmap*, and *ggplot*. The command lines to be executed in R are included in the protocol. The resource for R instructions can be found at <https://www.rstudio.com/online-learning/>.

1.1. To install the *Cytofast* package, start R (version "3.6") and install *Bioconductor* version 3.9 by entering the following code:

```
if (!requireNamespace("BiocManager", quietly = TRUE))
 install.packages("BiocManager")
BiocManager::install("cytofast")
```

1.2. Ensure that the package is loaded in the desired environment by running the following:

```
library(cytofast)
```

2. Creating clusters

NOTE: To showcase the two clustering methods *Cytosplore* and *FlowSOM* with *Cytofast*, the NK cells ($CD161^+$) in the tumor micro-environment 3 days after PD-L1 treatment are analyzed.

2.1. Clustering performed by *Cytosplore*

2.1.1. After downloading and installing *Cytosplore*, hosted at <www.*Cytosplore*.org>, upload the .fcs files (**Supplementary Files 1.1–1.8** [Cytosplore input files]) by clicking **File | Open FCS File(s)** in *Cytosplore*. Add a unique sample tag as channel by clicking **Add unique sample tag as channel** when prompted and select a cofactor for hyperbolic arcsinh transformation (default is 5).

2.1.2. Select **Run H-SNE**, run a HSNE level of 3, and wait for the map to be generated.

NOTE: This step may require some time, depending on the number of cells analyzed and level of HSNE chosen.

2.1.3. On the first HSNE level, check the cells that are positive for CD161. Select the $CD161^+$ cells and right-click **Zoom into selection**. At the second level, repeat the procedure to reach the third level with only $CD161^+$ events.

2.1.4. Once the last tSNE map is generated, save the clusters defined by *Cytosplore* by right-clicking on the tSNE map and choose **Save Clusters**. Choose the directory of the output files as prompted by *Cytosplore* and note this location, because this directory will then be used to load the .fcs files into R.

NOTE: The number of subsets can be manually changed by changing the sigma value. The sigma value is set by default at 30; however, the actual number of subsets depends on the input. Here, *Cytosplore* detected 10 different subsets, with each file representing one subset.

2.1.5. Use a simple name (characters only) when renaming the output files, which will make identification and further handling easier. Save the output files by selecting **Save**.

NOTE: After saving, a folder is being created by *Cytosplore* with the .fcs files, with each file corresponding to the identified clusters in *Cytosplore*. The next step will be loading the files into R with the help of *Cytofast*. Here, the generated output files are provided in **Supplementary Files 2.1–2.10** (Cytosplore output files).

2.1.6. Load the output files generated by *Cytosplore* into R with the designated function: readCytosploreFCS.

```
dirFCS <- "C:\\Users\\username\\Desktop\\tostudy"
cfData <- readCytosploreFCS(dir = dirFCS, colNames = "description")
```

2.1.7. Clean the data by removing some parameters such as “Time” and “Background”. Check the position of the column related to its unnecessary parameters and remove it from the matrix.

```
colnames(cfData@expr)
cfData@expr <- cfData@expr[,-c(3,4,6,8:10,46:49,51:54)]
```

NOTE: The unnecessary columns can be seen by reading the column names of the generated matrix. By running colnames(cfData@expr) once more, ensure that only the desired parameters are obtained.

2.1.8. Re-order the markers so that lineage markers are displayed first, followed by functional markers.

```
cfData@expr <- cfData@expr[,c(1,2,3,35,36,31,9,10,18,8,37,20,
 29,40,5,30,33,11,34,14,19,
 32,28,6,7,4,12,13,17,16,15,
 21,22,24,25,26,27,38,39)]
```

NOTE: Step 2.1.8 is optional.

2.1.9. Link the meta datafile to the generated data from *Cytosplore* by uploading the spreadsheet metafile containing clinical information (**Supplementary File 3: sample_id Cytosplore.xlsx**).

```
library(readxl)
meta <- read_excel("C:\\Users\\username\\Desktop\\sample_id.xlsx")
cfData@samples <- data.frame(meta)
```

NOTE: The clustering being performed by *Cytosplore* is now finished. An alternative clustering option to *Cytosplore* is *FlowSOM* and is described in section 2.2. Once one of the two clustering steps is performed, proceed with the visualization step (section 3).

2.2. Clustering performed by *FlowSOM*

2.2.1. First install *FlowSOM* in R by running the following command:

```
if (!requireNamespace("BiocManager", quietly = TRUE))
install.packages("BiocManager")
BiocManager::install("FlowSOM")
```

```
library(FlowSOM)
```

2.2.2. Install the flowCore package using a similar method and load it in the environment by running the following:

```
if (!requireNamespace("BiocManager", quietly = TRUE))
install.packages("BiocManager")
BiocManager::install("flowCore")
library(FlowSOM)
```

2.2.3. Load the raw data provided in **Supplementary Files 4.1–4.8** (FCS FlowSOM input), which were previously gated on CD161+ events, in R with the read.flowSet function.

```
fcs_raw <- read.flowSet(path="C:\\Users\\username\\Desktop\\tocluster", pattern = ".fcs", transformation = FALSE, truncate_max_range = FALSE, seed=123)
```

2.2.4. Select the relevant biological markers (remove “Background” or “Time”) by selecting the proper columns and transform the data in an arcsinh5 manner as shown in the code below (here, remove columns 1, 2, 4, 5, 6, 17, 21, 24, 25, 34, 35, 37, 38, 51, which do not correspond to any biological marker). Apply a cofactor of 5 as previously applied with *Cytosplore*, by choosing **cofactor=5** in the function below.

```
fcs_raw <- fsApply(fcs_raw, function(x, cofactor=5){
colnames(x) <- fcs_raw[[1]]@parameters@data$desc
expr <- exprs(x)
expr <- asinh(expr[,-c(1,2,4,5,6,17,21,24,25,34,35,37,38,51)]/ cofactor)
exprs(x) <- expr
return(x)})
```

2.2.5. Cluster the data using the *FlowSOM* function. To compare *FlowSOM* and *Cytosplore*, choose to cluster the data in ten subsets as the output previously yielded by *Cytosplore*.

```
fsom <- FlowSOM(fcs_raw, transformFunction = FALSE, scale = FALSE,
 scaled.center = FALSE, scaled.scale = FALSE, silent = FALSE, colsToUse = c(1:37),
 nClus = 10, maxMeta = 10, importance = NULL, seed = 123)
```

NOTE: This can be changed manually by the user.

2.2.6. Assign each cell to its identified subset and sample ID.

```
subset_id <- as.factor(fsom$FlowSOM$map$mapping[,1])
levels(subset_id) <- fsom$metaclustering
head(subset_id)
```

2.2.7. Load the metadata file in R (available in **Supplementary File 5: sample_id *FlowSOM*.xlsx**) containing the group assignment and link it to the .fcs files.

```
sampleid <- read_excel("C:\\Users\\username\\Desktop\\sample_id.xlsx")
sampleid <- na.omit(sampleid)

sampleid$sampleID <- as.factor(sampleid$sampleID)
sampleid$group <- as.factor(sampleid$group)
sampleid$CSPLR_ST <- as.factor(sampleid$CSPLR_ST)

sampleid <- as.data.frame(sampleid)
names(sampleid)[3] <- "sampleID"
sampleID <- lapply(fsom$FlowSOM$metaData, function(x){rep(x[1], each =
length(x[1]:x[2]))})
attr(sampleID, 'names') <- NULL
sampleID <- as.factor(unlist(sampleID))
sampleid <- data.frame(sampleid)
levels(sampleID) <- paste("ID", 1:dim(sampleid)[1], sep="_")

df <- data.frame(subset_id, sampleID, fsom$FlowSOM$data[, c(1:37)])
rename <- data.frame(colpar=fcs_raw[[1]]@parameters@data$desc)
colnames(df) <- c("clusterID", "sampleID", rename$colpar[c(1:37)])
df$ clusterID <- as.factor(df$ clusterID)
df$sampleID <- as.factor(df$sampleID)
```

2.2.8. Create a cfList based on the dataframe obtained from *FlowSOM* by running the following script:

```
cfData <- cfList(samples = sampleid,
expr = df)
```

2.2.9. Re-order the markers to appear similarly to the output from the *Cytosplore* analysis.

```
cfData@expr <- cfData@expr[,c(1,2,34,36,37,10,23,24,31,22,
 38,15,8,3,27,9,11,28,35,26,
 14,33,17,20,21,18,25,29,13,
 30,12,16,32,4,5,6,7,19,39)]
```

NOTE: The clustering by *FlowSOM* is now finished. Next, perform visualization of the clustering output.

3. Visualization: post-processing clustering analysis

NOTE: This step is a method that is common to both clustering methods. Therefore, it can be performed after clustering either with *FlowSOM* or *Cytosplore*.

3.1. Before creating the heatmaps, generate the count table per sample by using the function cellCounts as shown in the code below. Since some clusters contain fewer cells than others, scale the data per cluster by specifying “scale = TRUE” inside the function cellCounts, so that the dispersion among samples can be easily seen.

```
cfData <- cellCounts(cfData, frequency = TRUE, scale = TRUE)
```

NOTE: The data can now be visualized.

3.2. Visualize with heatmap.

NOTE: One of the main functions of this package is cytoHeatmaps, which is used to visualize the phenotype of the created clusters as well as their heterogeneity with respect to the samples.

```
cytoHeatmaps (cfData, group="group", legend=TRUE)
```

3.3. Visualization with boxplots

NOTE: Data can be represented in a quantitative way by calling the function cytoBoxplots. The output of this function represents the proportion of each sample for every cluster.

3.3.1. Generate the cell count as done in step 3.1, but do not scale the data to obtain the frequency of each cluster.

```
cfData <- cellCounts(cfData, frequency = TRUE, scale = FALSE)
cytoBoxplots(cfData, group = "group")
```

3.4. Visualization with median intensity signal histogram

NOTE: The data can also be visualized by representing the median intensity signal histogram.

3.4.1. Visualize the expression intensity of three markers: CD45, CD11c, and CD54.

3.4.2. Check the names of the desired markers by calling the following line. Note the marker names and include it in the msiPlot function.

```
names(cfData@expr)
```

NOTE: Here, focus on CD45, CD11c, and CD54. Check the exact spelling of the markers and adjust if needed:

```
msiPlot(cfData, markers = c("89Y_CD45", "167Er_CD11c" ,"164Dy_CD54"), byGroup='group')
```

REPRESENTATIVE RESULTS:

The workflow *Cytofast* **(Figure 1)** is meant to provide a quantitative and qualitative overview of the data originally clustered by analysis software (i.e., *FlowSOM* or *Cytosplore*). *Cytofast* runs several possible outputs, including the heatmap of all clusters identified in the analysis and based on marker expression **(Figure 2** and **Figure 3)**. The dendrogram on the top represents the hierarchical similarity between the identified clusters. The upper panel

displays another heatmap showing the relative quantity of corresponding subsets in each sample. The dendrogram on the left shows the similarity between samples and is based on hierarchical clustering performed on the Euclidean distances between samples. The combined heatmaps are shown for *FlowSOM* followed by *Cytofast* in **Figure 2** and for *Cytosplore* followed by *Cytofast* in **Figure 3**. *Cytofast* can also be used to present the data quantitatively and display the results in boxplots (by using cytoBoxplots function), as shown in **Figure 4** and **Figure 5**.

Similar clusters were found between the two different methods (e.g., cluster 8 from *Cytosplore* corresponds to cluster 10 from *FlowSOM*), and co-expression of some inhibitory markers like PD-1 and LAG-3 were still visible in both methods). Both clustering methods allowed discrimination between PD-L1 *vs*. PBS treated mice. In contrast, some differences between both methods can be highlighted. *FlowSOM* identifies 2 clusters (MHC-II^{+}), whereas *Cytosplore* shows only one cluster (MHC-II^{+dim}). This is due to the initial gating strategy in which NK cells were manually gated on CD161^{+} cells, then further processed by *FlowSOM*. However, *Cytosplore* automatically gated cells from the CD45^{+} population on the first HSNE level, which were then clustered in a higher hierarchical level. Thus, *Cytosplore* defined the NK cell subsets more precisely than how manual gating focused on CD161. Nevertheless, hierarchical clustering of the samples was preserved, as shown in the dendrogram on the right, indicating that segregation between the two groups (PD-L1 and PBS) was not dependent on the chosen clustering method.

The number of clusters can be manually defined using both methods. *Cytofast* enables the user to assess the heterogeneity of their data and can provide insight into how to choose the number of clusters into which the data should be divided. Other features are included in the *Cytofast* package, such as the msiPlot function (step 3.4.2), showing the median signal Intensity (MSI) plot of every marker per group (**Figure 6** and **Figure 7)**. This function allows detection of global changes, such as increases in the expression of CD54 or CD11c in NK cells of the PD-L1-treated group. Optional features can be incorporated in the *Cytofast* package, such as displaying data in bar graphs and other methods of data representation. The latter requires the addition of ggplot tools, which can be generated by R.

FIGURE AND TABLE LEGENDS:

Figure 1: Workflow of *Cytofast* package. The data were generated by mass cytometry from a tumor 3 days after treatment with immunotherapy or left untreated. Two different clustering techniques were compared: *Cytosplore* and *FlowSOM*. *Cytofast* was used to visualize differences between the two techniques.

Figure 2: Cluster overview and cluster abundance per group as analyzed by *Cytofast* following *Cytosplore*. Heatmap of all NK cell clusters (CD161^{+} cells defined automatically by *Cytosplore*), which were identified 3 days after immunotherapy (PD-L1). Data shown is based on *Cytosplore* clustering and pooled from the untreated and PD-L1 treated groups. Levels of ArcSinh5-transformed expression marker are displayed on a rainbow scale. On the lower panel, the relative abundance of each sample is represented by the green-to-purple scale. The dendrogram on the right represents the similarity between samples based on subset

frequencies. The frequency scale represents the dispersion of the mean. A low or a high frequency is represented by a green or purple color, respectively.

Figure 3: Cluster overview and cluster abundance per group as analyzed by *Cytofast* following *FlowSOM*. Heatmap of all NK cell clusters (pre-gated on $CD161^+$ events), which were identified 3 days after immunotherapy (PD-L1). Data shown is based on *FlowSOM* clustering and pooled from the untreated and PD-L1 treated groups. Levels of ArcSinh5-transformed expression marker are displayed on a rainbow scale. On the lower panel, the relative abundance of each sample is represented by the green-to-purple scale. The dendrogram on the right represents the similarity between samples based on subset frequencies. The frequency scale represents the dispersion of the mean. A low or a high frequency is represented by a green or purple color, respectively.

Figure 4: *Cytofast* representation with boxplots of the clusters defined by *Cytosplore*. The frequency of each cluster is represented in a boxplot, separated into the two groups (PBS and PD-L1). One individual dot corresponds to one mouse.

Figure 5: *Cytofast* representation with boxplots of the clusters defined by *FlowSOM*. The frequency of each cluster is represented in a boxplot, separated into the two groups (PBS and PD-L1). One individual dot corresponds to one mouse.

Figure 6: Distribution of signal intensity plots from NK cells automatically gated by *Cytosplore*. Distribution of signal intensities are shown in a histogram for three specific markers: CD45, CD11c, and CD54.

Figure 7: Distribution of signal intensity plots from NK cells automatically gated by *FlowSOM*. Distribution of signal intensities are shown in a histogram for three specific markers: CD45, CD11c, and CD54, segregated by groups PBS and PD-L1.

DISCUSSION:

Cytofast is a rapid computational tool that provides a quick and global exploration of cytometric data by highlighting and quantifying treatment-specific cellular subsets. The protocol described aims to further process clustering analyses with *Cytosplore* or *FlowSOM*. Other clustering analysis tools are suitable for *Cytofast*, but this requires the use of *Cytofast* to assign each cell to a subset. *Cytofast*, however, is not a clustering method, and therefore requires clustering procedures before use.

The analysis performed here showed that certain $CD161^+$ NK cell subsets in the tumor microenvironment were sensitive to a PD-L1 blockade. This was evidenced by changes in their phenotype and abundance, which were observed using both *Cytosplore* and *FlowSOM* as clustering methods. Both methods distinguished the main NK cell cluster ($CD11b^+$ $NKG2A^+$) with slightly different frequencies (15%–20% for *Cytosplore*, 30%–40% for *FlowSOM*). The differences in abundance and this approximation did not affect the global pattern, because both dendrograms displayed in the right panels of **Figure 2** and **Figure 3** showed similar results. By using *Cytofast*, it is thus possible (independent from the clustering method chosen) to segregate PD-L1-treated and untreated mice based on analyses of NK cell cluster phenotype and abundance.

Depending on the recorded parameters, modifications to the protocol are needed. Specifically, certain parameters such as time and background must be removed while performing the clustering analysis. In addition, it is important that each cell is assigned to a subset. The cfData function will simply add the raw cell counts per cluster per sample into the cfList. From this step, the cytoheatmap can be built as explained in section 3.

Cytofast has been successfully used as a visualization and quantification tool to compare different clustering methods[13]. This R package is also compatible with advanced features, such as the *globaltest*[14], which can test associations between groups of clusters using clinical variables. In the future, the *globaltest* tool and other algorithms can be integrated with *Cytofast* for more in-depth visualization and quantification.

ACKNOWLEDGMENTS:
We acknowledge funding from the European Commission of a H2020 MSCA award under proposal number 675743 (ISPIC). We thank Tetje van der Sluis and Iris Pardieck for testing the protocol.

DISCLOSURES:
The authors have nothing to disclose.

REFERENCES:
1. Anchang, B. et al. Visualization and cellular hierarchy inference of single-cell data using SPADE. *Nature Protocols.* **11** (7), 1264-1279 (2016).
2. Zunder, E. R., Lujan, E., Goltsev, Y., Wernig, M., Nolan, G. P. A continuous molecular roadmap to iPSC reprogramming through progression analysis of single-cell mass cytometry. *Cell Stem Cell.* **16** (3), 323-337 (2015).
3. Van Gassen, S. et al. *FlowSOM*: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry A.* **87** (7), 636-645 (2015).
4. Levine, J. H. et al. Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell.* **162** (1), 184-197 (2015).
5. Samusik, N., Good, Z., Spitzer, M. H., Davis, K. L., Nolan, G. P. Automated mapping of phenotype space with single-cell data. *Nature Methods.* **13** (6), 493-496 (2016).
6. Spitzer, M. H. et al. IMMUNOLOGY. An interactive reference framework for modeling a dynamic immune system. *Science.* **349** (6244), 1259425, (2015).
7. Hotelling, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology.* **24** (6), 417-441 (1933).
8. van der Maaten, L., Hinton, G. Visualizing Data using t-SNE. *Journal of Machine Learning Research.* (2008).
9. Pezzotti, N., Hollt, T., Lelieveldt, B., Eisemann, E., Vilanova, A. Hierarchical Stochastic Neighbor Embedding. *Computer Graphics Forum.* **35** (3), 21-30 (2016).
10. Becht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nature Biotechnology.* **37**, 38 (2018).
11. Haghverdi, L., Buettner, F., Theis, F. J. Diffusion maps for high-dimensional single-cell analysis of differentiation data. *Bioinformatics.* **31** (18), 2989-2998 (2015).

12. Beyrend, G., Stam, K., Höllt, T., Ossendorp, F., Arens, R. *Cytofast*: A workflow for visual and quantitative analysis of flow and mass cytometry data to discover immune signatures and correlations. *Computational and Structural Biotechnology Journal.* **16** 435-442 (2018).
13. Beyrend, G. et al. PD-L1 blockade engages tumor-infiltrating lymphocytes to co-express targetable activating and inhibitory receptors. *Journal for ImmunoTherapy of Cancer.* **7** (1), 217 (2019).
14. Goeman, J. J., van de Geer, S. A., de Kort, F., van Houwelingen, H. C. A global test for groups of genes: testing association with a clinical outcome. *Bioinformatics.* **20** (1), 93-99 (2004).

Figure 1

Figure 1, Beyrend et al.

Click here to access/download;Figure;Figure 1 Overview.eps

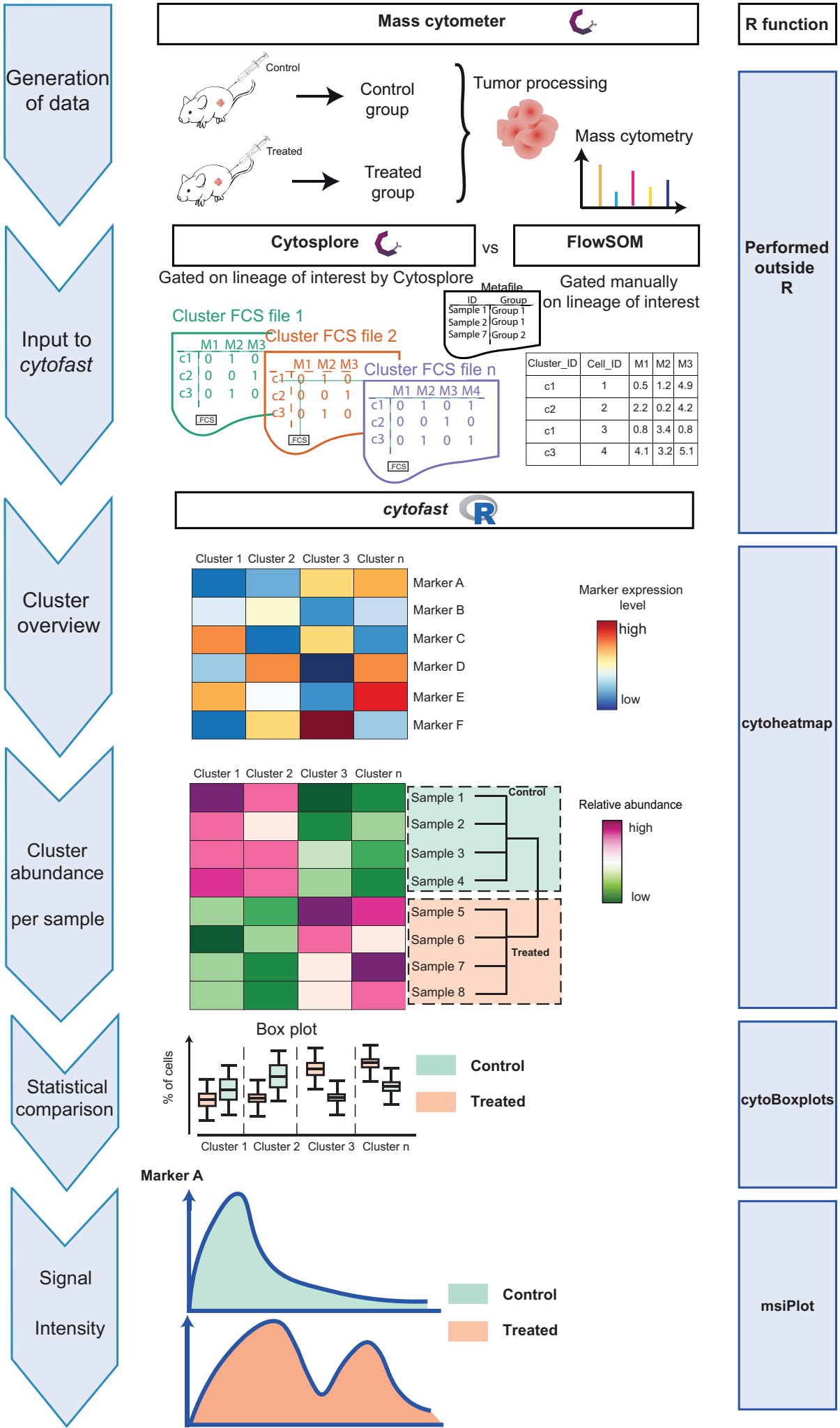


Figure 2

Click here to access/download;Figure;Figure 2.eps

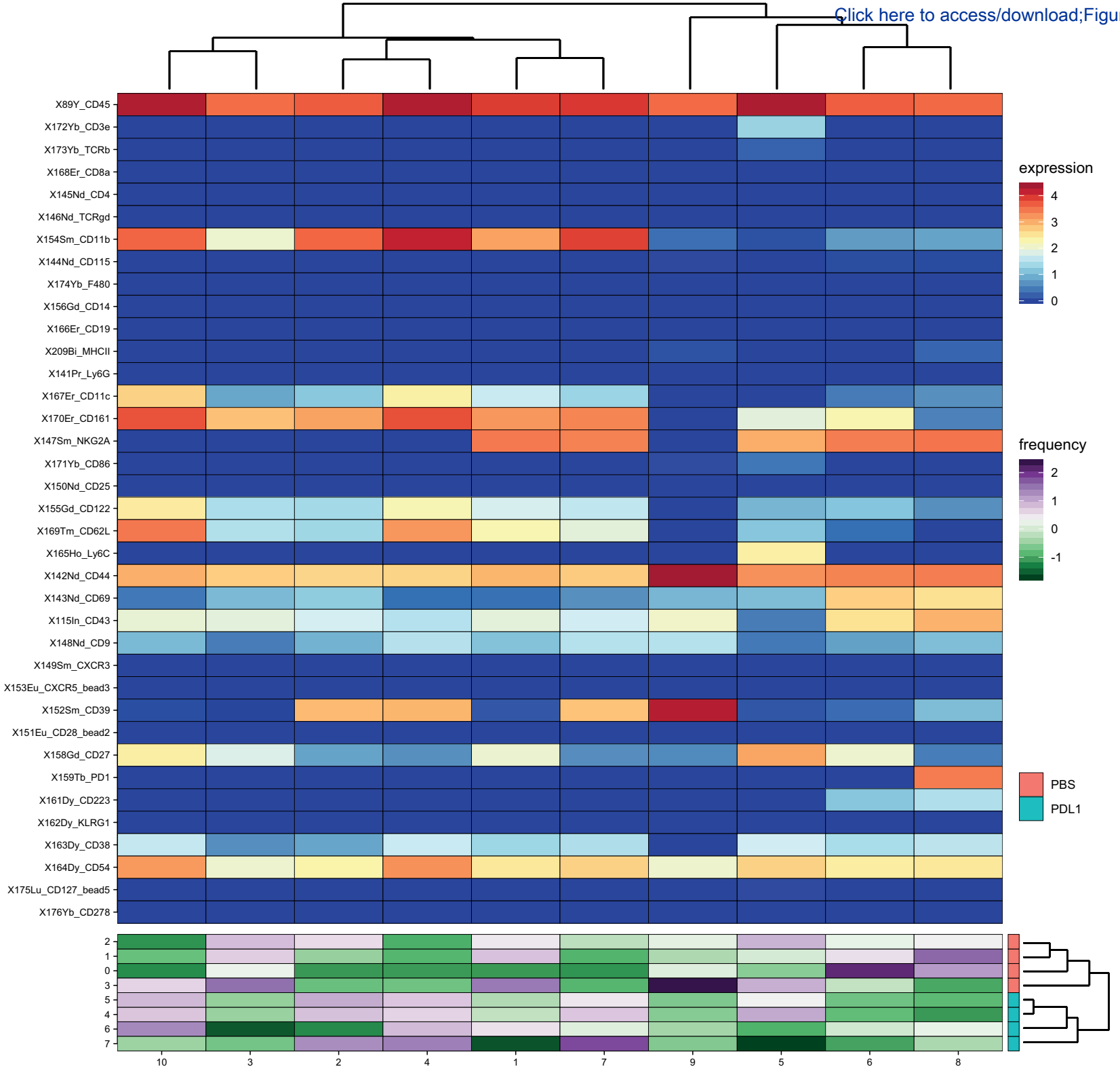


Figure 3

Click here to access/download;Figure;Figure 3.eps

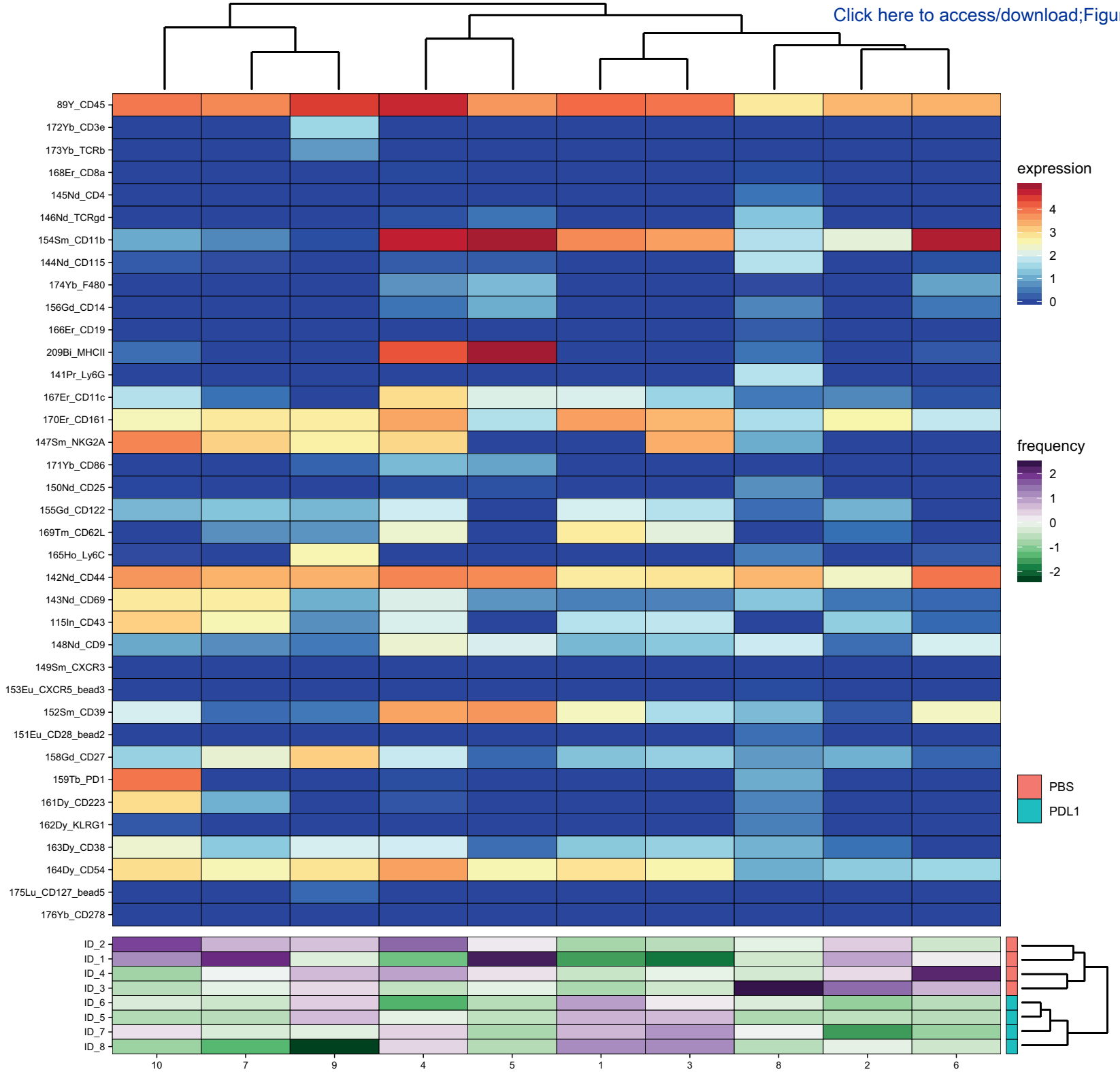


Figure 4

Click here to access/download;Figure;Figure 4.eps

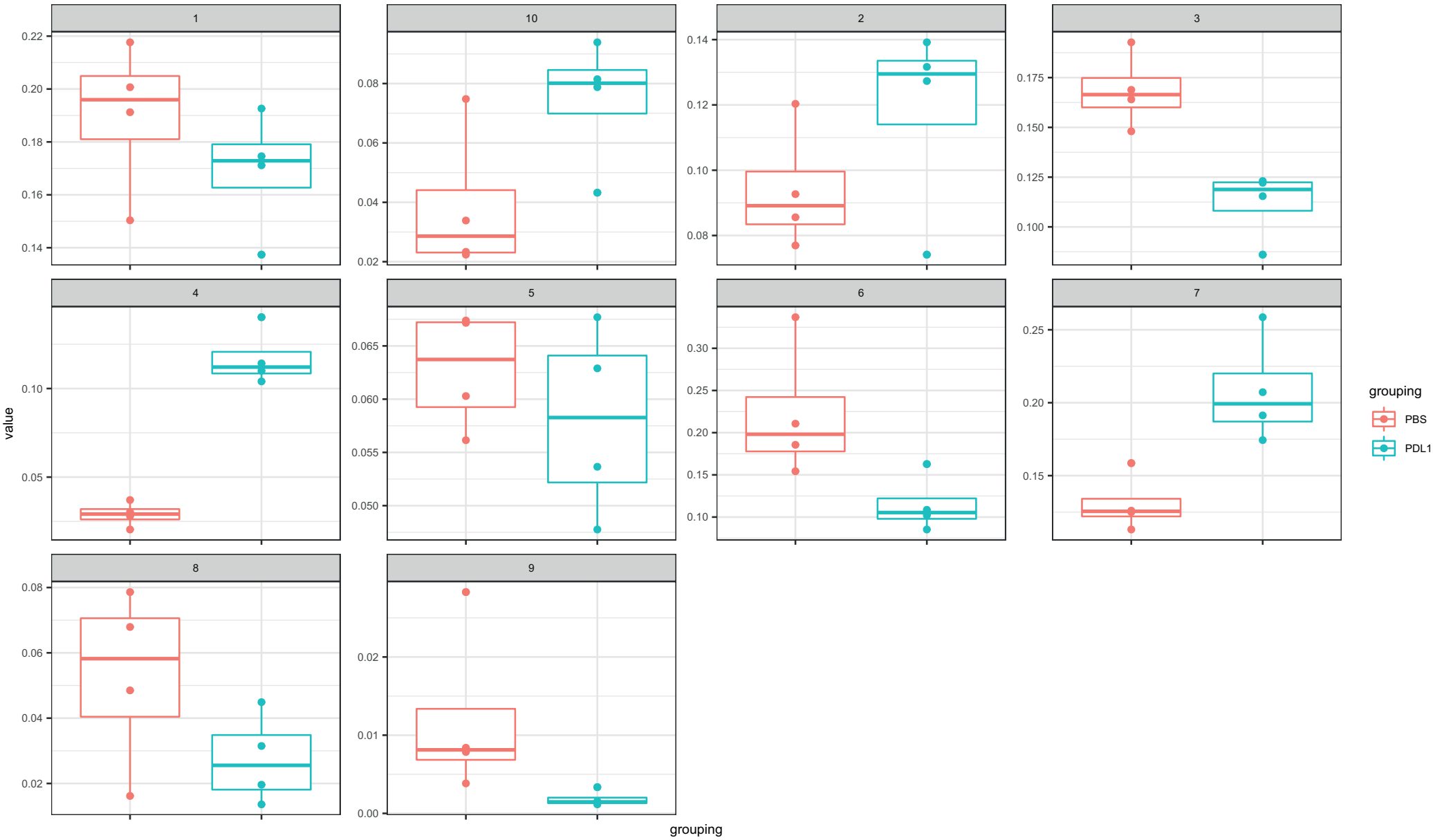


Figure 5

Click here to access/download;Figure;Figure 5.eps

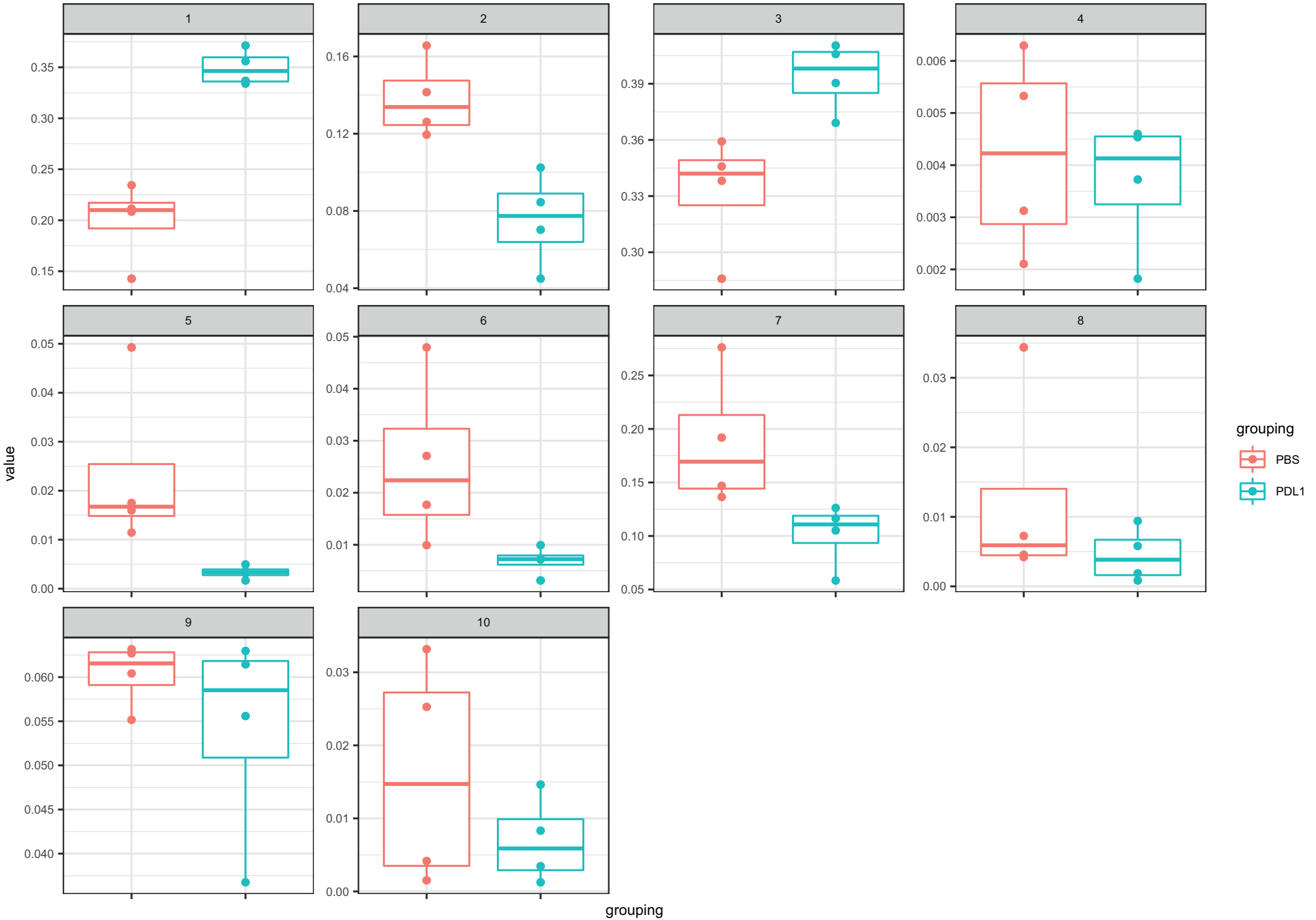


Figure 6

Click here to access/download;Figure;Figure 6.eps

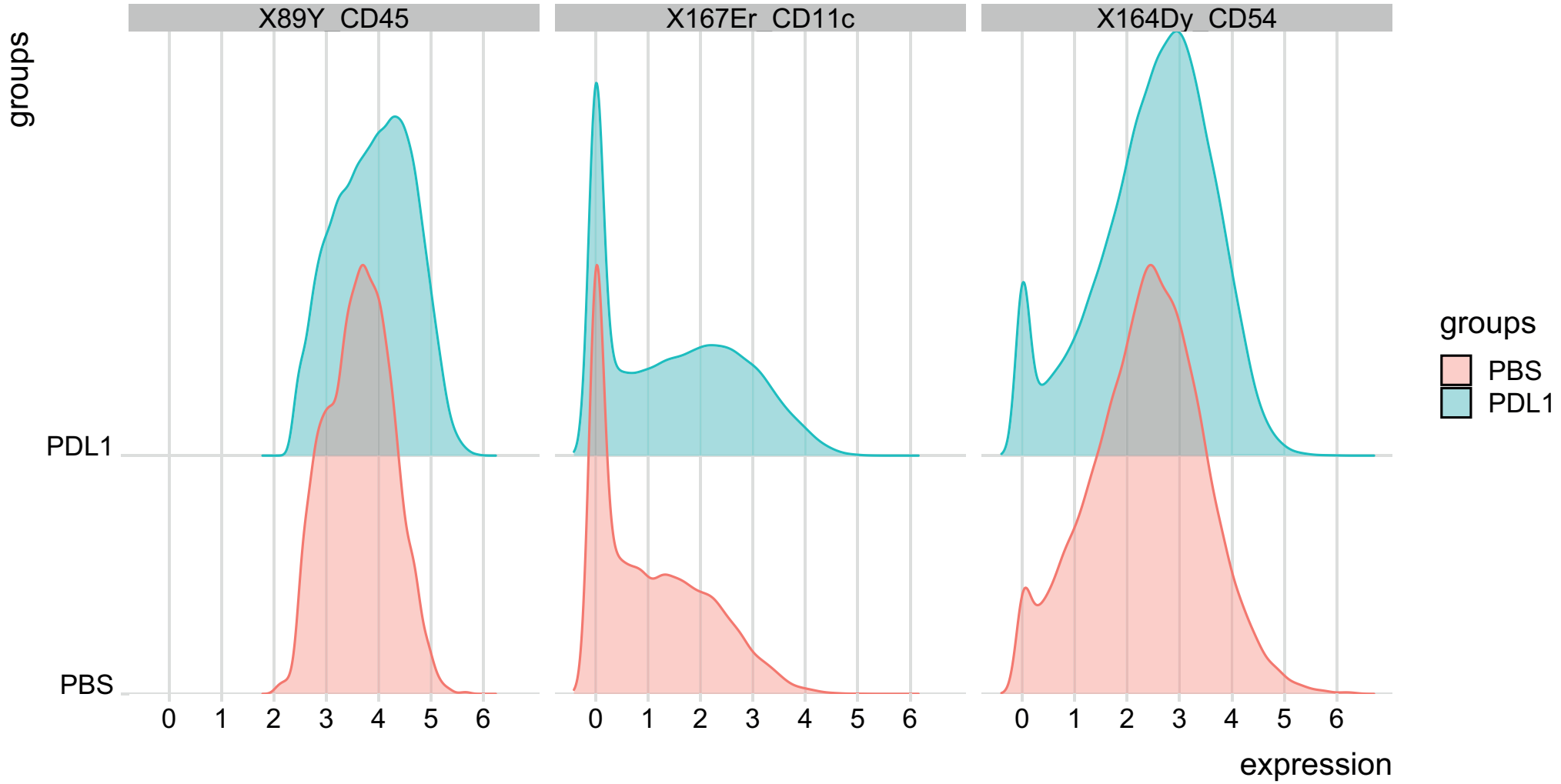


Figure 7

Click here to access/download;Figure;Figure 7.eps

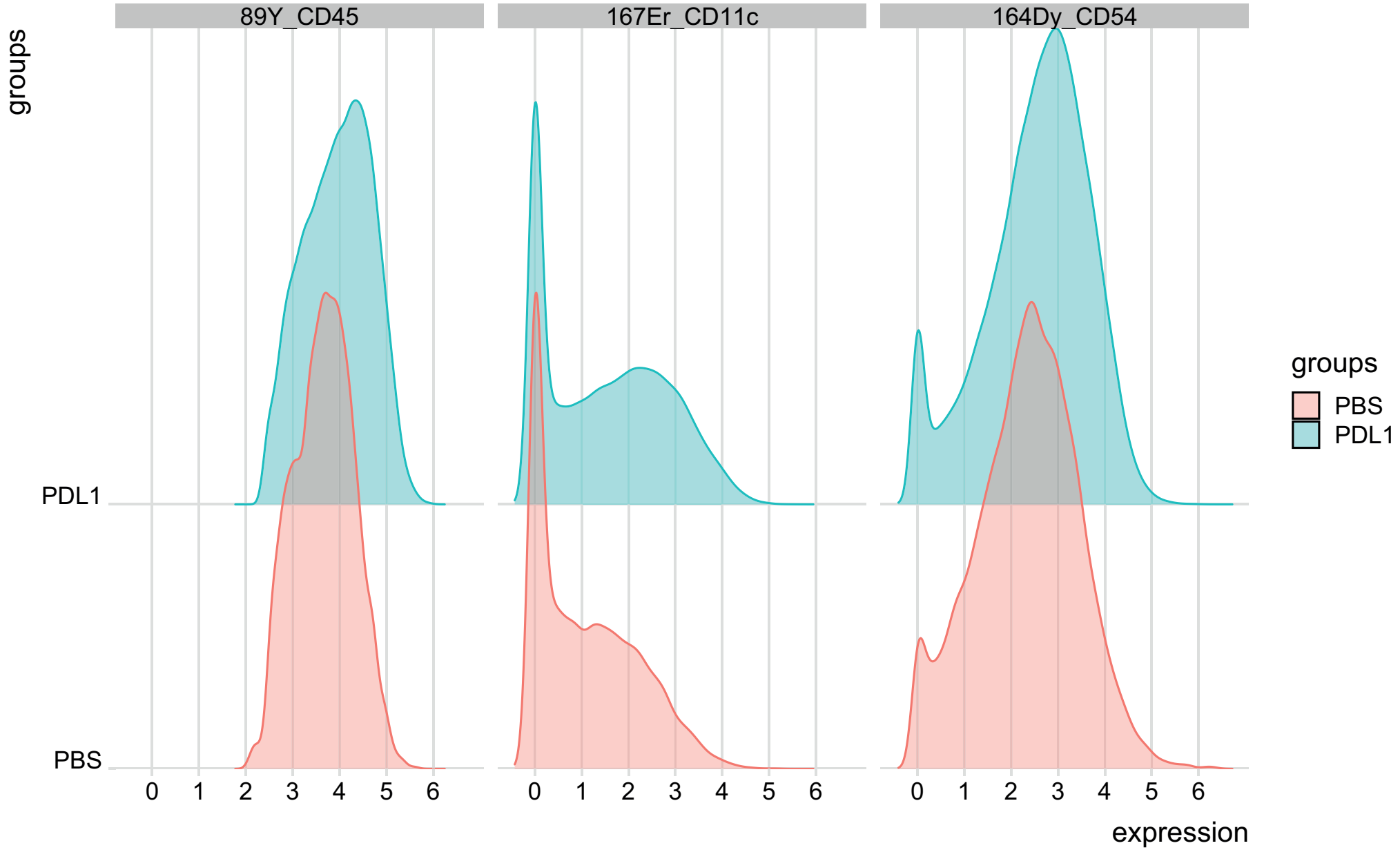


Table of Materials

Name of Material/ Equipment	Company	Catalog Number	Comments/Description
Computer	Dell	NA	NA

LEIDS UNIVERSITAIR MEDISCH CENTRUM

sender Dr. R. Arens
department Immunohematology and Blood Transfusion
address Albinusdreef 2, 2333 ZA Leiden
phone +31 71 526 3434
e-mail r.arens@lumc.nl
date September 3st, 2019
subject Cover letter, Beyrend et al. JoVE60525

Editorial Office of Journal of Visualized Experiments

Dear Editor,

Herewith we re-submit our manuscript entitled “Visualization and quantification of high-dimensional cytometry data by *Cytofast*” by Beyrend and colleagues for consideration of publication in the Journal of Visualized Experiments (JoVE60525).

We addressed the editorial and reviewer comments.

Reviewers' comments:

Reviewer #1: *In this short article, Beyrend et al introduce Cytofast, an R package to visualize results of clustering analyses. Cytofast uses cluster labels per cell and cell x samples relationships as input and can be used to plot heatmaps of clusters' abundances across samples, intensity distributions of markers within a cluster and across conditions, and boxplots of cluster frequencies across conditions. The package doesn't appear particularly new or groundbreaking but may be useful for those with little programming experience. Using Cytosplore in this context (which relies on a graphical user interface) may thus be easier for end-users than using FlowSOM which relies on a command line interface. The paper is easy to follow and one can envision how it can be accompanied by a video.*

Major point:Is the data analyzed in this manuscript published? I think it is crucial to exemplify such an analysis on a public dataset so that readers can download the data and follow and reproduce the analysis on their own workstations. If the data can not be made public for now, I think it would be necessary to choose another dataset.

The dataset will be published in Flow Repository, reference before submitting.

Minor points:

About cleaning channels: It would be good to show the end-user that may not be very familiar with R how to print the markers' names and to select markers by name. This would also be more informative in the manuscript

A new step has been added to the manuscript to specifically show how to select markers names (See section *2. Equipment and software for the data analysis*).

For FlowSOM clustering, doesn't nClus=20 translate to 20 clusters rather than 10 (used in the code versus the text)? The authors should clarify this point.

This was a mistake which has been corrected in the manuscript (See section *3.2. Clustering performed by FlowSOM*) .

The labels of figures 5 and 6 are misleading. It is the distribution of intensities rather than the median intensity that is displayed. Similarly the corresponding function may need to be renamed

I don't think the package described here is very original or powerful so I don't think that "deep exploration" (used in the discussion) is warranted, since the functions presented here are not doing anything particularly impressive.

These points have been revised in the manuscript

Reviewer #2

Manuscript Summary

In the manuscript by Beyrend et al, the authors describe use of their recently published R package, Cytofast, to facilitate analysis of mass cytometry data. Overall the text is thorough and will provide an important addition to the field of high-dimensional single cell data analysis.

Major Concerns:

1) It is critical to understand what events were put into the Cytosplore and FlowSOM analyses, as these appear discrepant. In its current form, it appears that Cytosplore uses all events from the original .fcs files, generates clusters at an undefined HSNE level, from which NK cells are selected? In contrast, FlowSOM appears to use pre-gated NK cells. While this is alluded to on lines 224-227, it is currently difficult to understand.

a. If an NK cell cluster was selected from Cytosplore for downstream analysis, clarify by what criteria this was selected. This could readily be incorporated around line 78/79, which refers to tSNE maps being generated at "the desired level". What level of HSNE was selected and how this decision was made?

Cytosplore uses all events from the original .fcs files. Clustering analysis was performed by Cytosplore, using 3 different HSNE levels. NK cells were then defined on the HSNE level 1, where the events were selected to zoom in the next level 2 and 3. Clustering ended to 10 different clusters, which have been exported at the second level.

FlowSOM used pre-gated NK cells based on $CD161^{+}$ cells.

We clarify these points in the manuscript, see section *3.1. Clustering performed by Cytosplore* and *3.2. Clustering performed by FlowSOM*.

b. Please clarify whether the analysis presented in Figure 2 and 3 both correspond to NK cell clusters, including a specific statement of input cell events in the legends for Figure 1 and 2.

We have clarify this point in the manuscript, see legends of Figure 2 and 3.

2) Correction? In Figure 2 legend, line 254, texts states "data shown is based on "FlowSOM clustering…", but the title states this is from Cytosplore?

This point has been corrected in the legend of Figure 3.

3) Clarify methods: i) What version of R was used for the analylsis? ii) What version of R is required for cytofast (and are there known incompatibilities)?

R should be upgraded to *Bioconductor* version 3.9; *R* version 3.6.0. is required for cytofast

4) Clarify methods: Clarify how to install cytofast in R; this could be taken directly from the Bioconductor page (which should also be referred to in the MS).

This has been added to the manuscript, see section *2. Equipment and software for the data analysis*

Minor Concerns:
1) MS would benefit from some minor copy-editing

The manuscript has been proofread.

2) Title: The title states that Cytofast is useful for flow and mass cytometry data, but the manuscript includes no mention of flow cytometry data beyond the title. Based on this, removing mention of flow cytometry data from the title would be more accurate.

We have changed the title, and use cytometry instead of flow and mass cytometry. We note however that Cytofast can be used for both flow and mass cytometry.

3) Order of methods and figures: Currently the manuscript describes Cytosplore then FlowSOM, however the order of these algorithms is reversed in the figures. It would make it easier to follow if there is a parallel structure in text and figures.

This has been corrected in the manuscript.

4) Additional reference: On line 53-54, it is worth including reference to UMAP as an additional dimensionality-reduction technique (e.g. reference Nat Biotechnol. 2018 Dec 3. doi: 10.1038/nbt.4314. Dimensionality reduction for visualizing single-cell data using UMAP. Becht E … Newell EW)

The paper is now mentioned in the manuscript line 52.

5) On line 82-83, the following sentence is difficult to understand "The analysis of the tumor micro environment is ending to ten phenotypically different subsets." Is this saying that there were 10 phenotypically different subsets identified in the tumor microenvironment dataset?

We have re-phrased this. The analysis of the tumor-microenvironment dataset has been performed and 10 phenotypically different subsets were identified by Cytosplore.

6) On line 109, it would be helpful to clarify how data were gated for NK cells (software and gating strategy).

The manuscript has been edited to include that detail line. $CD161^{+}$ events loaded in FlowSOM algorithm were pre-gated with FlowJo (see section *3.2. Clustering performed by FlowSOM*)

7) On lines 209-213, it would be helpful to clarify which Figure this is referring to. The description later in the paragraph makes it difficult to follow.

The figures 2 and 3 are now written as reference to this text.

8) In Figure legend 2 and 3, please clarify what the scale refers to for "frequency"—unclear what is meant by +2 to -2 in this context.

This refers to the dispersion around the mean. It can just be seen as low and high frequency. A description has been added to the Figures 2 and 3.

9) For Fig. 4 and 5, please clarify if individual dots correspond to individual samples/fcs files.

One individual dot is corresponding to one mouse, legends of the figures have been adapted accordingly.

10) To increase utility to readers, it may be helpful to identify any lines of code that can be used directly in R and any lines of code that may require modification for an individual's specific experiment. If this detail is not included in the text, it would be worth emphasizing which fields need to be modified in the video.

The position of channel like “Background” or “Time” is depending on the CyTOF panel being used. The cleaning steps described in the protocol need to be adjusted for each user. This has been indicated in the manuscript and will be emphasized during the video.

11) For those not familiar with R, it may be useful to provide a link to a resource that instructs (e.g. https://www.rstudio.com/online-learning/)

This has been added to the manuscript.

Reviewer #3:

Manuscript Summary:

In the paper, Beyrend and colleagues present Cytofast, an R package for the downstream analysis of cell clusters identified by automatic gating methods.

Major Concerns:

In the paper, Beyrend and colleagues present Cytofast, an R package for the downstream analysis of cell clusters identified by automatic gating methods. I think this tools is useful and valuable for the scientific community. However, I do not recommend this work for publication.

To me, there is a too strong overlap with the work presented in Beyrend, G., et al., Cytofast: A workflow for visual and quantitative analysis of flow and mass cytometry data to discover immune signatures and correlations. Computational and Structural Biotechnology Journal, 2018. 16: p. 435-44,event if the dataset used is not the same.

The work presented in Beyend et al, 2018, CSBJ, is the global presentation of the algorithm. The article submitted to JoVE is a protocol explaining step by step how to use Cytofast and compare the output of two different clustering analysis. The short abstract has been edited to clarify the aim of the work presented in JoVE.

In any cases, I think that there are multiple issues with the manuscript.
The introduction paragraph must be extended to present the outputs of the different algorithms. What are the limits of each algorithm that the author present? Why are the authors presenting some visualization algorithms "t-SNE, HSNE" if Cytofast cannot process them.

We clearly state in the discussion the uniqueness of Cytofast compared to other algorithms.

Cytofast can process HSNE and tSNE outputs as long as each cell is assigned to a cluster.

Is there some R packages similar to Cytofast publicly available? Why the algorithm outputs are difficult to understand ?

To our knowledge, there are no R packages similar to *Cytofast* publicly available. FlowSOM contains a visualization part, but limited to a few markers only (not more than 10) and therefore more suitable to flow cytometry.

I think the tutorial is not enough detailed in its current form. "cfData@expr <- cfData@expr[,-c(3,4,6,8:10,46:49,51:54)]" to what these columns corresponds to ?

This point has been made clearer line (see *3.2. Clustering performed by FlowSOM*).

I think that there is no way that a biologist with a limited knowledge in bioinformatics can understand and reproduce this pipeline. The whole tutorial/pipeline should be embedded in a formal R package in complement to the one already published. The tutorial paper should be presented in the form of an R markdown presentation.

It will be presented indeed as such in the video. Knowledge of R software is recommended and we have added the tutorial for this in the paragraph. It is beyond the scope of this manuscript to explain R.

The quality of the figures must be enhanced and the figure legends must be better presented (Figure 3,4,5,6...).

The quality of the figures has been improved.

What are the limits of the package developed by the authors? What remains to be developed ideally? Is there some statistical methods in their package to identify relevant clusters?

Cytofast is not a clustering analysis, but a method to visualize the output of several clustering algorithms assigning each cell to a cluster. The package is still in evolution and will contain new features in future. As an example the algorithm will be able to check the homogeneity of one cluster, assessing then the efficiency of the clustering method.

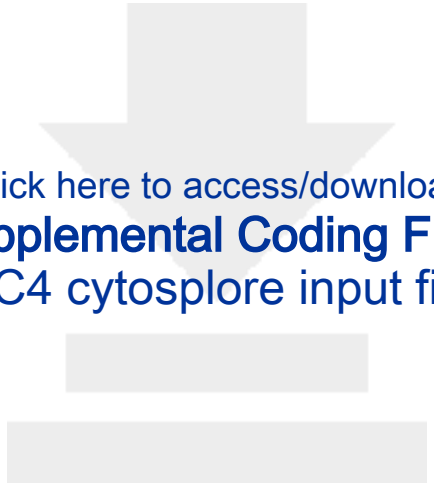
Based on the comments of the reviewer, we have improved our manuscript and look forward to your decision.

Sincerely,

Ramon Arens, PhD
Leiden University Medical Center,
The Netherlands

1. tocluster - Cytosplore input files

1. tocluster - Cytosplore input files



Click here to access/download

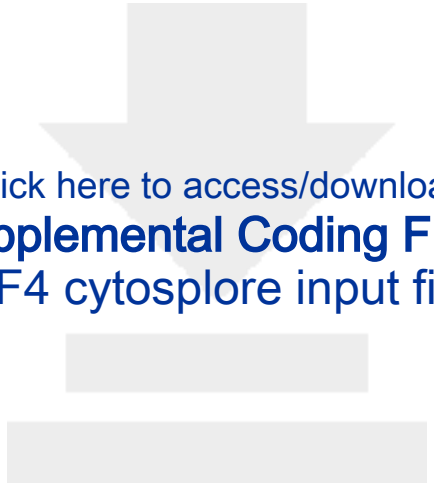
Supplemental Coding Files

1.2 - C4 cytosplore input file.fcs

1. tocluster - Cytosplore input files

1. tocluster - Cytosplore input files

1. tocluster - Cytosplore input files



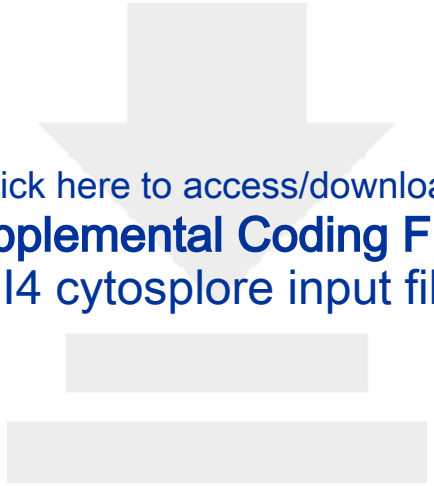
Click here to access/download

Supplemental Coding Files

1.5 - F4 cytosplore input file.fcs

1. tocluster - Cytosplore input files

1. tocluster - Cytosplore input files

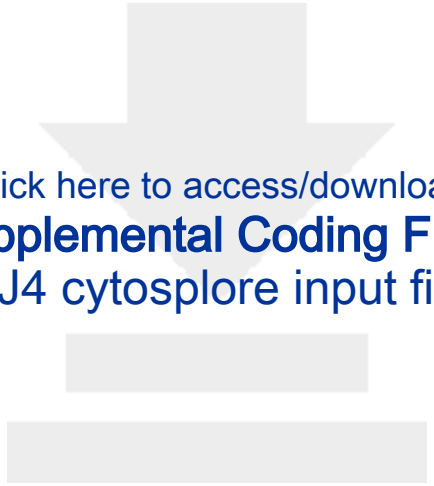


Click here to access/download

Supplemental Coding Files

1.7 - l4 cytosplore input file.fcs

1. tocluster - Cytosplore input files



Click here to access/download

Supplemental Coding Files

1.8 - J4 cytosplore input file.fcs

2. tostudy - Cytosplore output files

Click here to access/download

Supplemental Coding Files

2.1 nkcyto-1 cytosplore output.fcs

Click here to access/download

Supplemental Coding Files

2.3 nkcyto-3 cytosplore output.fcs

Click here to access/download

Supplemental Coding Files

2.4 nkcyto-4 cytosplore output.fcs

Click here to access/download

Supplemental Coding Files

2.5 nkcyto-5 cytosplore output.fcs

Click here to access/download

Supplemental Coding Files

2.6 nkcyto-6 cytosplore output.fcs

Click here to access/download

Supplemental Coding Files

2.7 nkcyto-7 cytosplore output.fcs

Click here to access/download

Supplemental Coding Files

2.8 nkcyto-8 cytosplore output.fcs

Click here to access/download

Supplemental Coding Files

2.9 nkcyto-9 cytosplore output.fcs

Click here to access/download

Supplemental Coding Files

2.10 nkcyto-10 cytosplore output.fcs

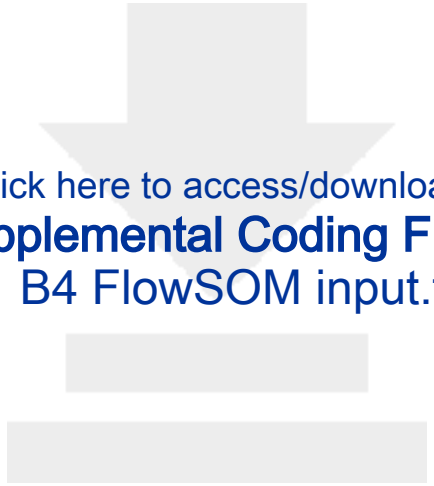
3. sample_id Cytosplore

Click here to access/download

Supplemental Coding Files

3. sample_id Cytosplore.xlsx

4. tocluster-raw data FlowSOM input

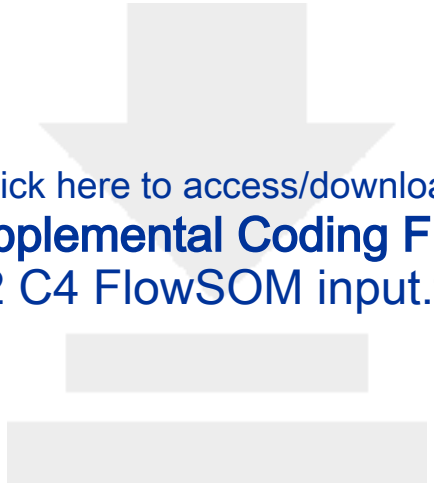


Click here to access/download

Supplemental Coding Files

4.1 B4 FlowSOM input.fcs

4. tocluster-raw data FlowSOM input

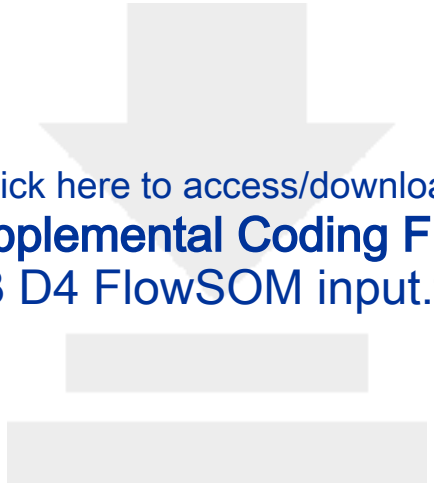


Click here to access/download

Supplemental Coding Files

4.2 C4 FlowSOM input.fcs

4. tocluster-raw data FlowSOM input

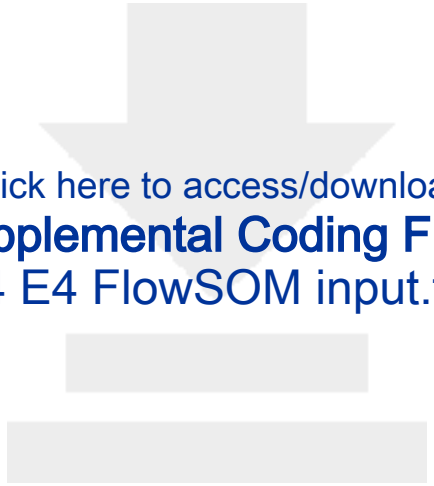


Click here to access/download

Supplemental Coding Files

4.3 D4 FlowSOM input.fcs

4. tocluster-raw data FlowSOM input

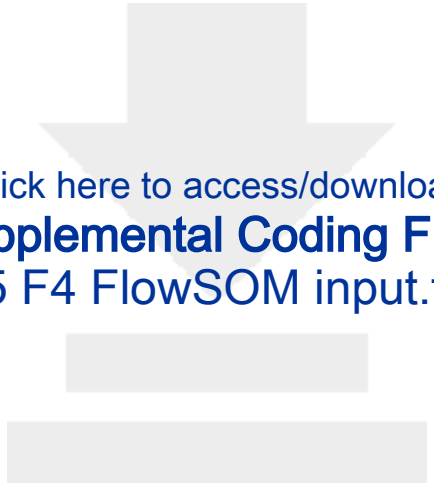


Click here to access/download

Supplemental Coding Files

4.4 E4 FlowSOM input.fcs

4. tocluster-raw data FlowSOM input

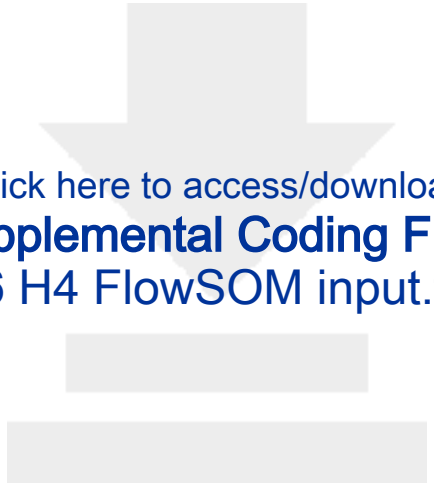


Click here to access/download

Supplemental Coding Files

4.5 F4 FlowSOM input.fcs

4. tocluster-raw data FlowSOM input

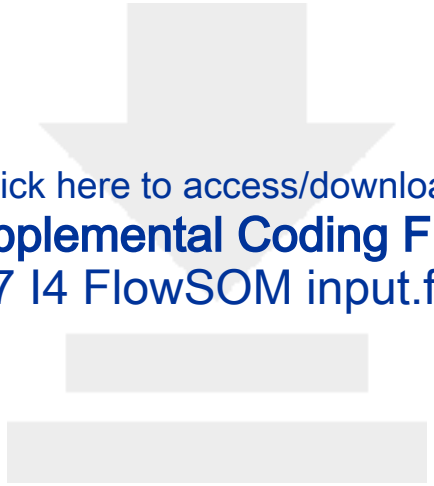


Click here to access/download

Supplemental Coding Files

4.6 H4 FlowSOM input.fcs

4. tocluster-raw data FlowSOM input



Click here to access/download

Supplemental Coding Files

4.7 I4 FlowSOM input.fcs

4. tocluster-raw data FlowSOM input

Click here to access/download

Supplemental Coding Files

4.8 J4 FlowSOM input.fcs

5. sample_id FlowSOM

Click here to access/download

Supplemental Coding Files

5. sample_id FlowSOM.xlsx

jove JOURNAL OF VISUALIZED EXPERIMENTS
1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

Title of Article:	Visualization and Quantification of High-Dimensional Cytometry Data using Cytofast and the upstream clustering methods FlowSOM and Cytosplore
Author(s):	Guillaume Beyrend, Koen Stam, Ferry Ossendorp, Ramon Arens

Item 1: The Author elects to have the Materials be made available (as described at http://www.jove.com/publish) via:

☐ Standard Access ☑ Open Access

Item 2: Please select one of the following items:

☑ The Author is **NOT** a United States government employee.

☐ The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.

☐ The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

ARTICLE AND VIDEO LICENSE AGREEMENT

1. **Defined Terms.** As used in this Article and Video License Agreement, the following terms shall have the following meanings: "**Agreement**" means this Article and Video License Agreement; "**Article**" means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; "**Author**" means the author who is a signatory to this Agreement; "**Collective Work**" means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; "**CRC License**" means the Creative Commons Attribution-Non Commercial-No Derivs 3.0 Unported Agreement, the terms and conditions of which can be found at: http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode; "**Derivative Work**" means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; "**Institution**" means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; "**JoVE**" means MyJove Corporation, a Massachusetts corporation and the publisher of The Journal of Visualized Experiments; "**Materials**" means the Article and / or the Video; "**Parties**" means the Author and JoVE; "**Video**" means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion of the Article, and in which the Author may or may not appear.

2. **Background.** The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3. **Grant of Rights in Article.** In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to **Sections 4** and **7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and(c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License.

4. **Retention of Rights in Article.** Notwithstanding the exclusive license granted to JoVE in **Section 3** above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5. **Grant of Rights in Video – Standard Access.** This **Section 5** applies if the "Standard Access" box has been checked in **Item 1** above or if no box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, Subject to **Section 7** below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

6. **Grant of Rights in Video – Open Access.** This **Section 6** applies only if the "Open Access" box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to **Section 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. For any Video to which this **Section 6** is applicable, JoVE and the Author hereby grant to the public all such rights in the Video as provided in, but subject to all limitations and requirements set forth in, the CRC License.

7. **Government Employees.** If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in **Item 2** above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum rights permitted under such statute. In such case, all provisions contained herein that are not in conflict with such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

8. **Protection of the Work.** The Author(s) authorize JoVE to take steps in the Author(s) name and on their behalf if JoVE believes some third party could be infringing or might infringe the copyright of either the Author's Article and/or Video.

9. **Likeness, Privacy, Personality.** The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

10. **Author Warranties.** The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

11. **JoVE Discretion.** If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole

discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole discretion and without giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

12. **Indemnification.** The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in JoVE or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contamination due to the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

13. **Fees.** To cover the cost incurred for publication, JoVE must receive payment before production and publication of the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

14. **Transfer, Governing Law.** This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to me one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement is required per submission.

CORRESPONDING AUTHOR

Name:	Guillaume Beyrend
Department:	Department of Immunohematology and Blood Transfusion
Institution:	Leiden University Medical Center
Title:	MS.
Signature:	Ramon Arens
Date:	09/12/2019

Please submit a **signed** and **dated** copy of this license by one of the following three methods:

1. Upload an electronic version on the JoVE submission site
2. Fax the document to +1.866.381.2236
3. Mail the document to JoVE / Attn: JoVE Editorial / 1 Alewife Center #200 / Cambridge, MA 02140

612542.6 For questions, please contact us at submissions@jove.com or +1.617.945.9051.

Signature Certificate

Document Ref.: HC6DE-WMASL-RHNI3-XQSXK

Document signed by:

Ramon Arens

Verified E-mail:
r.arens@lumc.nl

IP: 145.88.209.33 | Date: 12 Sep 2019 07:25:44 UTC

Ramon Arens

Document completed by all parties on:
12 Sep 2019 07:25:44 UTC