BECKMAN COULTER
Life Sciences

Leveraging the Combined Power of Kaluza and the Cytobank Platform

This Tech Note shows how to accelerate your data analysis workflows by combining the interactive sliders for compensation and data transformation in Kaluza Analysis* and the power of cloud computing for machine learning algorithms on the Cytobank platform*. Learn how you can upload data from Kaluza directly into your Cytobank account.

We will also show how to export fcs files from the Cytobank platform that contain derived parameters from machine learning algorithms and visualize the data in Kaluza Analysis.

In recent years computational approaches such as dimensionality reduction and clustering techniques have gained popularity in the flow cytometry community.

This Tech Note explains how the functionalities of Kaluza Analysis Software that allow the fast development of a gating strategy, intuitive scale adjustments and easy compensation corrections can be combined with the advanced algorithms offered by the Cytobank platform.

Kaluza Analysis offers easy to use real time data exploration tools and supports QC tracking with a dedicated module. Different data visualization tools support the exploration of low to medium complexity data. Data reporting functionality enables the rapid generation of summary results.

Cytobank is a cloud-based analysis platform with integrated machine learning based analysis algorithms, as well as a structured and secure content management system for flow cytometry and other single cell data.

Cytobank's clustering, dimensionality reduction, and visualization tools (**SPADE**, **viSNE**, **CITRUS**, **FlowSOM**) leverage the scalable compute and collaborative power of the cloud, allowing large analyses to be done quickly, and the cloud-based storage provides the capability to automatically archive and easily share these data securely and safely[1].

Combining desktop Kaluza Analysis Software and cloud-based Cytobank software allows the user to choose which solution to use at each step (Figure 1).

Figure 1: Steps in the flow cytometry analysis pipeline covered by Kaluza and the Cytobank platform. The Cytobank platform offers automatic compensation calculation but manual adjustment of the compensation matrix is more interactive in Kaluza. Gating steps on biaxial plots work well in both systems. For exploratory data analysis, clustering and data visualization Cytobank is the software of choice.

Analysis of multi-dimensional flow cytometry data is done in several stages[2]. Early stages of data QC and cleanup have lower requirements for compute performance and may be performed on your own laptop or PC. The first stage of data pre-processing is then followed by exploration and analysis. The exploration stage provides an overview of the cells in the samples. For low to mid-complexity marker panels, biparametric plots supplemented with tools such as Radar Plots and Tree Plots available in Kaluza can be sufficient.

For high dimensional data, dimensionality reduction techniques such as viSNE allow the information from all parameters to be summarized in a single two dimensional display[4]. In the next stage of data analysis, clustering algorithms such as FlowSOM and SPADE can be used to automatically identify cells of similar phenotype in an unbiased manner[5,6]. All of these algorithms are integrated on the Cytobank platform.

These later stages of data analysis benefit from cloud computing to free up local equipment and allow for parallel runs. While it is possible to perform viSNE, SPADE and FlowSOM calculations using desktop software, the computer cannot be used for other tasks during this time. Also, only one run at a time can be performed, slowing down optimization of algorithm settings. Cloud-based data analysis does not suffer from these limitations.

In order to combine the comfort of desktop based real time data interaction with cloud-based machine learning tools, a solution for smooth data transfer is required.

The Kaluza Cytobank Plugin bridges this gap by enabling you to compensate and transform your data using Kaluza's innovative user interface and real time feedback and directly upload the data to your Cytobank account where the cloud based system will allow high dimensional data to be analyzed and visualized with high impact graphics.

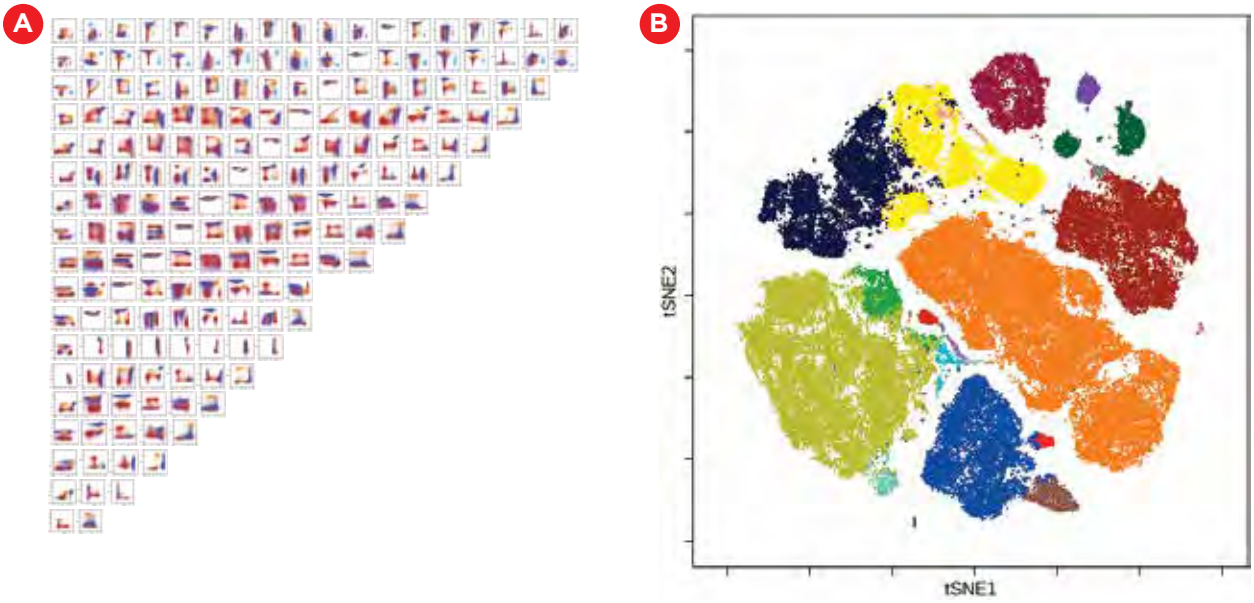

Figure 2: Comparison of biaxial plots with machine learning analysis of a 20-color panel cytometry data. A) N by N plot view of a 20-color panel cytometry data resulting in 190 plots. B) viSNE (or t-SNE) map view of the same data that visualize the 24-parameter information on one single map.

Note: viSNE in Cytobank uses the Barnes-Hut implementation of the t-SNE algorithm.

For details refer to the Use Machine Learning Algorithms to Explore the Potential of Your High Dimensional Flow Cytometry Data Example of a 20-color Panel on CytoFLEX LX Tech Note.

Tips for Success

- To be able to analyze data files together in Cytobank, make sure that the parameter names and axis labels are consistent. Both can easily be adjusted in Kaluza.
- To process a large number of samples in Kaluza, automate steps like applying parameter values and applying a previously determined compensation matrix, by utilizing the Batch Process Activity.

Protocol

The following steps will be described:

Kaluza Analysis	Kaluza Cytobank Plugin	Cytobank platform
• Importing files • Adjusting Parameters Adjusting compensation • Adjusting data transformation • Pre-gating of the population of interest	• Exporting compensated and transformed data as fcs files • Editing or removing fcs keywords	• Data scaling • Advanced analyses • Data export • Generation of high impact graphics

1. Importing files into Kaluza Analysis:

There are multiple ways to import files into Kaluza Analysis. The preferred approach is to drag & drop files into the analysis list. Please refer to Kaluza IFU C10986 for additional options.

1.1. Locate the files you wish to include in the Analysis List.

1.2. Select the files, drag them into the list and release the mouse button (Figure 3). The Analysis List automatically fills once the mouse button is released. By default, the original file names appear in the Data Set column.

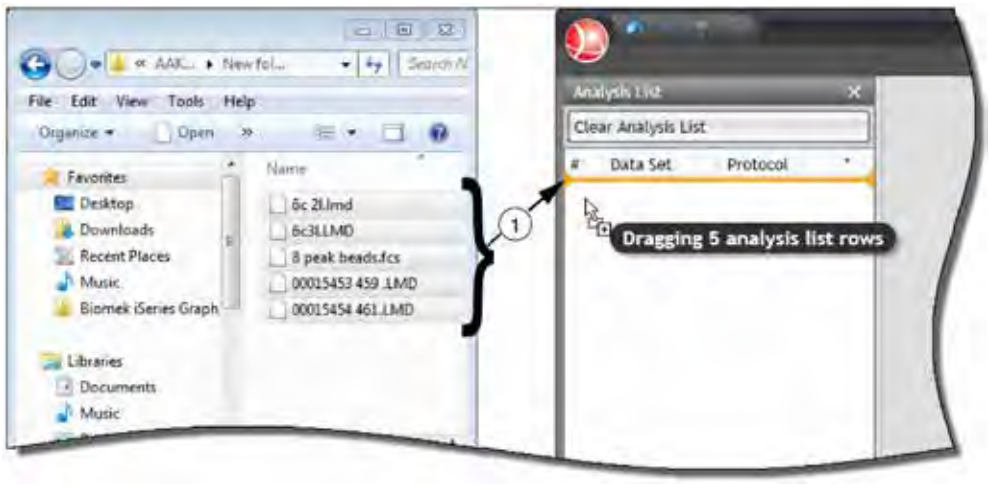

Figure 3: Open files in Kaluza using Drag&Drop

For detailed instructions on how to use Kaluza Analysis software, refer to the Kaluza IFU C10986.

2. Check and adjust Parameter Values, as needed:

The Parameters pane (Figure 4) contains a list of the parameters collected in the original Data Set. This pane enables you to edit parameter names, descriptions, types, detectors, and measurement type.

2.1. Check and if desired change Name (corresponds to fcs keyword $PnN) and Description (corresponds to fcs keyword $PnS). Note that the raw data file keywords remain unchanged, however, the keywords in fcs files exported using the Kaluza Cytobank Plugin will be updated.

2.2. To update the Name and Description fields highlight the text in the field that you wish to change.

2.3. Press the (Delete) key on your keyboard.

2.4. Type the new content into the field and press (Enter). There is a limit of: 5,000 characters for parameter name, 5,000 characters for parameter description and 5,000 characters for parameter detector.

Figure 4: Example of Parameters pane in Kaluza showing buttons to edit the parameter selection, values for parameter name and the associated description.

3. Add All Plots and adjust compensation:

Fluorescence spillover of fluorochromes into channels other than their designated detection channel is corrected by fluorescence compensation. Compensation artifacts can negatively impact downstream analysis if not corrected[3]. Post-acquisition compensation can free up the instrument by shortening time required for sample acquisition. Even if samples are acquired with the compensation matrix already applied, it is advisable to check and correct the compensation for all samples.

The Add All Plots function allows you to create biparametric plots for all fluorescence parameter combinations with a single click. This is a great starting point for making adjustments to the compensation or data transformation.

3.1. From the Plots & Tables ribbon tab, click the drop-down arrow on the "Add All Plots" icon to customize which plots will be included on the plot sheet. To follow the workflow outlined in this application note, make selections as shown in Figure 5.

Figure 5: Add All Plots Options selections used for the example shown here. Include gating plot with plot type: Dot and Gate Type: Ellipse, chose FS Linear and SS Linear as X axis and Y axis detectors, respectively.

3.2. Click on Add All Plots to create plots.

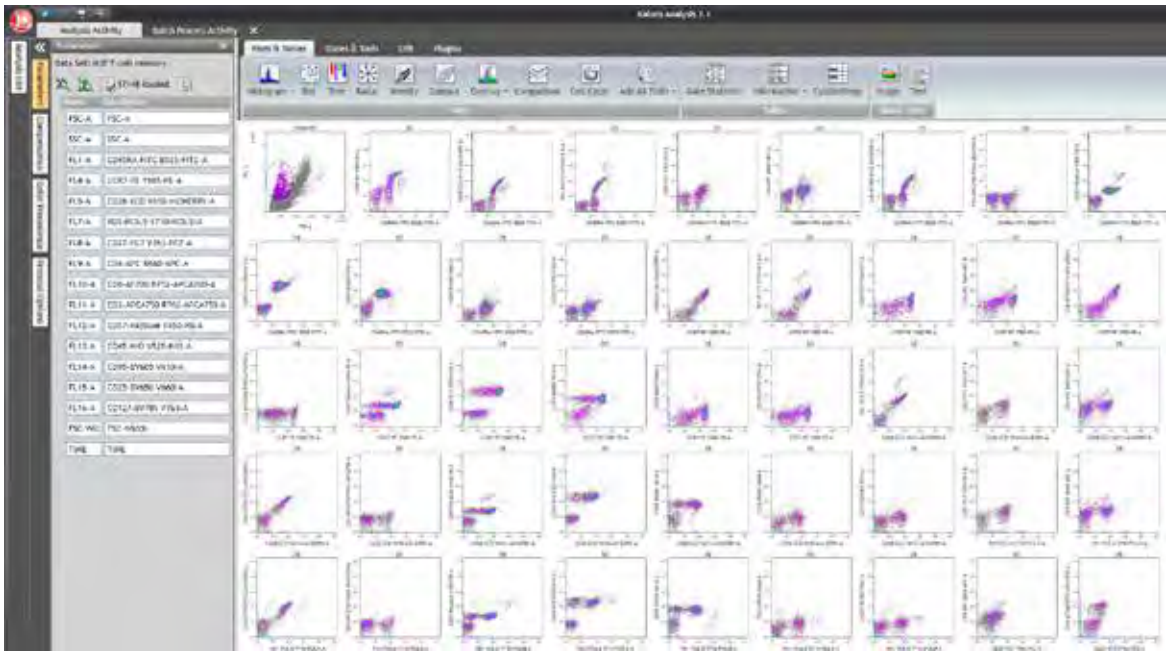

Figure 6: Plots created after selecting "Add All Plots".

3.3. The plot in the upper left corner is the gating plot. Double click to enlarge and adjust region A so that it includes cell population(s) of interest that will allow you to evaluate the compensation.

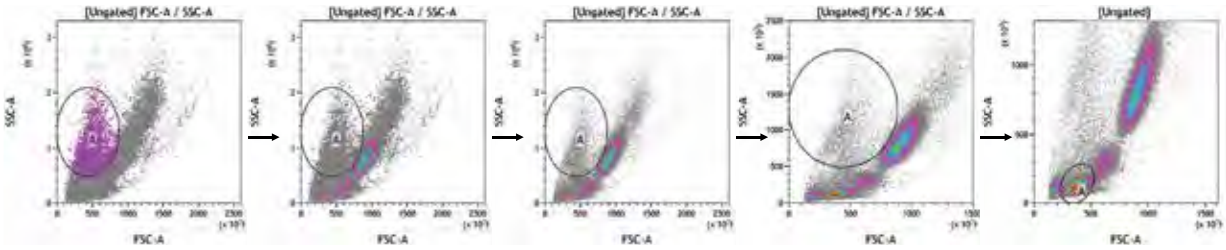

Figure 7: Gating plot before and after adjustments were made to switch from dot to density plot, increase plot resolution, zoom in on the population and gate on lymphocytes. Data was generated using a normal whole blood sample stained with DuraClone IM T Cell Subsets Tube* (PN: B53328), acquired on a CytoFLEX LX cytometer* (PN: C40324) and analyzed using Kaluza Analysis Software*. Plots are for illustration purposes only.

3.4. If you selected the logicle option when setting up the Add All Plots function, logicle sliders will appear when you click on a plot. Select and drag the slider to adjust the scale to display negative values. As demonstrated on the Y-axis in Figure 8, the numerical value appears and changes as you move the slider. This adjustment may be helpful when judging the quality of the compensation settings as it helps visualize events below zero.

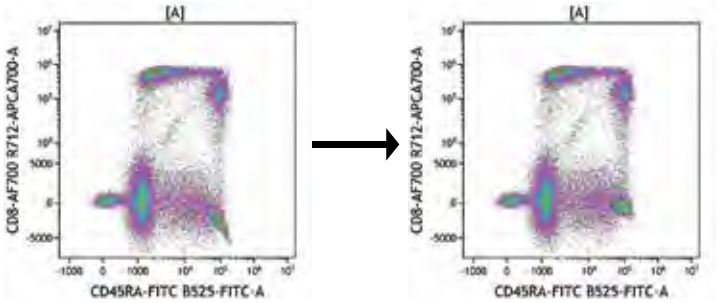

Figure 8: Logicle transformation of data in the CCR-PE Y585-PE-A channel (Y axis). Data was generated using a normal whole blood sample stained with DuraClone IM T Cell Subsets Tube* (PN: B53328), acquired on a CytoFLEX LX cytometer* (PN: C40324) and analyzed using Kaluza Analysis Software*. Plots are for illustration purposes only.

3.5. To make adjustments to the compensation matrix, create Spillover Sliders on all applicable plots that allow you to compensate for fluorescence Spillover by using real-time visual cues on plots. To use the Spillover Sliders select the Compensation icon for the Gates & Tools ribbon tab. Applicable plots are now equipped with plot sliders (Figure 9).

3.6. To update Spillover using the Spillover Sliders, click one of the sliders and slide in the direction of the desired change. The Spillover value is displayed next to the slider. The Spillover Sliders allow values between 0 and 100%. Adjustment values outside of this range must be entered manually into the Spillover matrix.

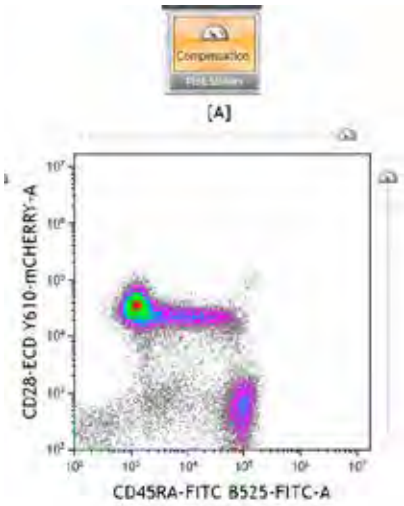

Figure 9: Compensation Icon from Gates & Tolls ribbon tab and exemplary density plot with Spillover Sliders.

Machine learning algorithms are unable to identify this artefact such as Improperly scaled data, over- or undercompensation and as a result may infer differences between phenotypically similar calls. This can prevent the viSNE map from converging and may affect coloring by channel. It is therefore important to properly adjust compensation and transformation (Figure 10, Figure 11).

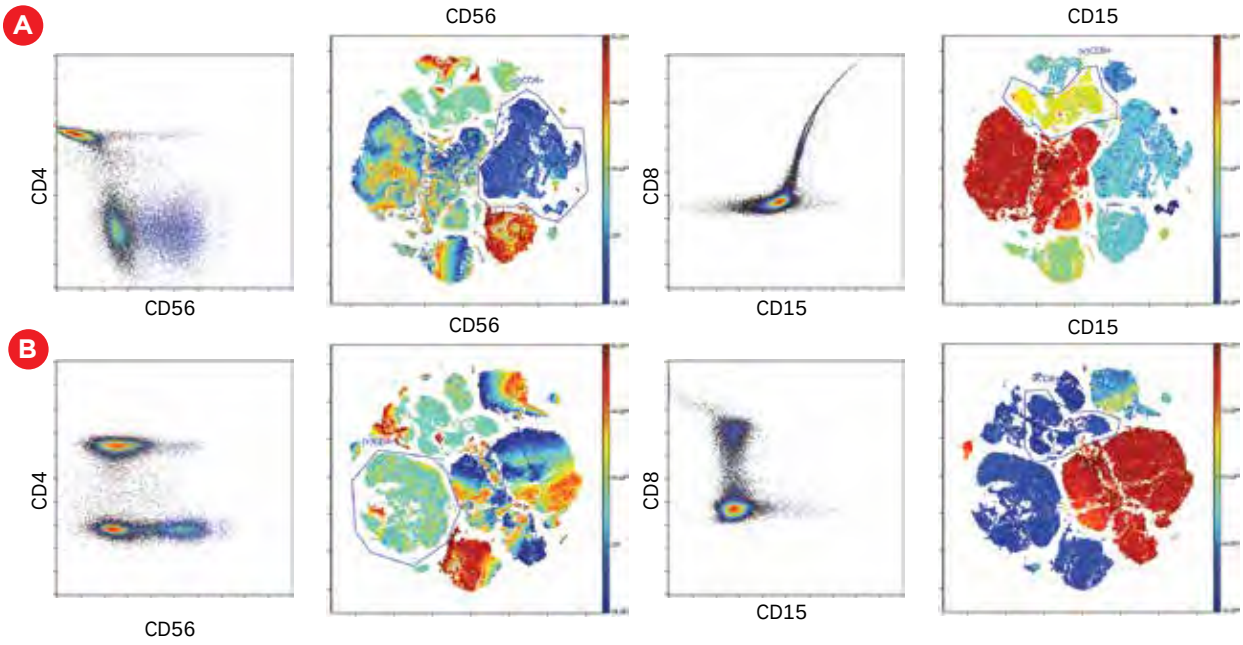

Figure 10: Effect of fluorescence spillover compensation on viSNE visualization of 20-color flow cytometry data. viSNE analysis was performed on the Cytobank platform.

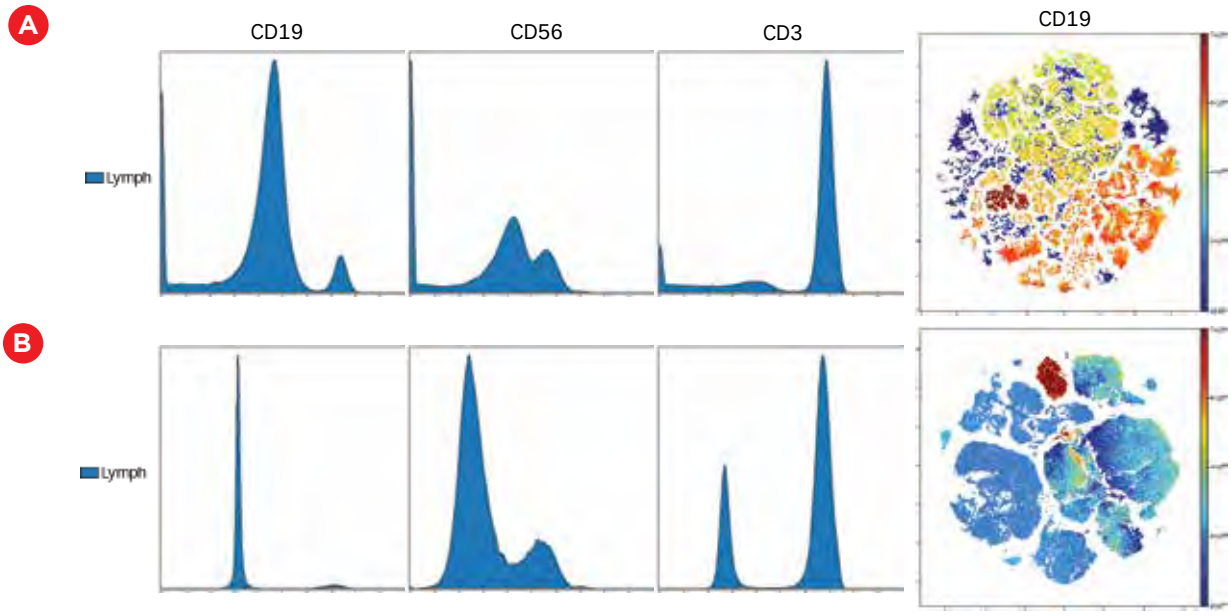

Figure 11: Effect of data transformation on viSNE visualization of 20-color flow cytometry data. Representative histograms for lymphocyte markers show the data transformation. A shows suboptimal data transformation resulting in a poorly converged viSNE map and falsly high coloring for CD19. B shows the same data with optimized transformation settings. viSNE analysis was performed on the Cytobank platform.

4. Pre-gate population of interest

Debris, doublets and often also dead cells or unwanted events detected in a dump channel should be removed before applying machine learning algorithms. These events add unwanted noise to the downstream analysis, could negatively impact data display and confound statistical results if not identified and excluded appropriately.

The stability of the sample acquisition process can be evaluated by displaying scatter and fluorescence versus time data. Gaps in data acquisition may indicate disruptions of the sample flow caused by blockages or bubbles. This display allows the user to restrict the analysis to parts of the acquisition when sample flow and signal detection were stable. On pressurized analyzers the sample running dry often results in a burst of artefact events across all parameters. Gating on a TIME plot allows the exclusion of such artefacts.

Depending on the desired data visualization and research questions, it may be of value to pre-gate the population of interest for further analysis.

These gating steps may also be performed in the Gating function on the Cytobank Platform.

For detailed instructions on how to create plots and gates in Kaluza, refer to the Kaluza IFU C10986.

4.1. Optional step: To provide a better overview of the data and the gating strategy it is suggested to create a new plot sheet. Select the New Sheet...hyperlink (Figure 12). A drop-down appears, enabling you to choose between adding a new plot sheet or report sheet. New sheets are added as the active sheet. You can add up to 1024 sheets per Protocol.

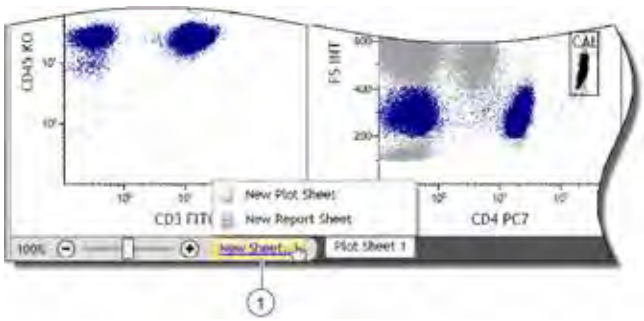

Figure 12: Adding a New Sheet

4.2. One option for adding new plots is to choose items from the Plots & Tables ribbon tab. To add a new item:

- To the bottom of the sheet, select the icon corresponding to the type of item to include in the Protocol.

  OR

- To a specific location on the sheet, click the specific icon, drag it to the location you prefer and release the mouse button.

4.3. To add a gate, select the Gates and Tolls ribbon tab. To add a polygonal gate, select the (Polygon) icon.

- Click your mouse where you wish to begin creating your gate.
- Determine the path you need for your gate and continue clicking your mouse at the location of each direction change. As you draw the gate, a new line will be added each time you click your mouse, and the default gate color will display, working as a guide to show your progress.
- Double click or select the initial point when the gate is completed.

4.4. To gate a plot select the hyperlink located at the top of a plot. A pop-up menu appears, which contains a list of gates, including recently created gates and gates by category (Common, Quadrant, and Boolean). Select the desired gate from the pop-up(s). The events on the plot are now filtered, based on the events in the selected gate.

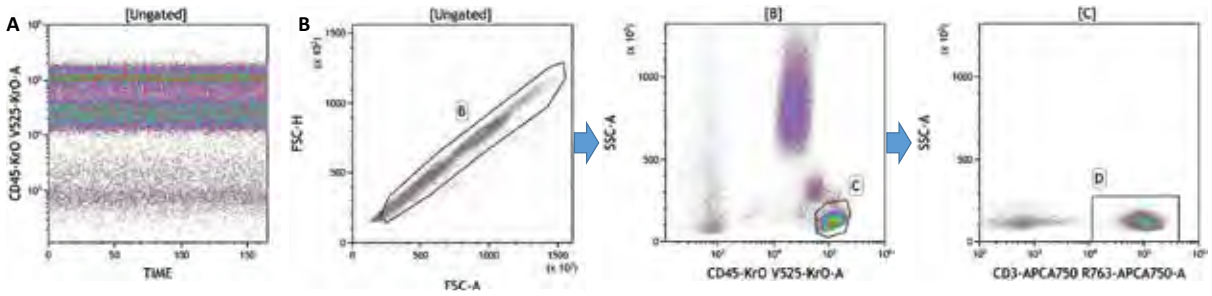

Figure 13: Exemplary gating strategy to check for instabilities (A), exclude doublets (B), identify CD45 bright SSC low cells (C) and to identify the CD3 positive subsets of population C (D). Data was generated using a normal whole blood sample stained with DuraClone IM T Cell Subsets Tube* (PN: B53328), acquired on a CytoFLEX LX cytometer* (PN: C40324) and analyzed using Kaluza Analysis Software*. Plots are for illustration purposes only.

4.5. Apply the changes you made to all the data sets in your experiment. For a detailed description of all options refer to the Kaluza IFU C10986.

- Applying a Protocol and logicle transformation settings to a Raw Data Set. One option is to drag the protocol from another data set. The logicle settings are part of the Protocol.
  i. From the Analysis List, select the Protocol you wish to use for another raw Data Set.
  ii. With the Protocol still selected, drag it to the cell in the Protocol column corresponding to the Data Set, and then and release the mouse button when the cell becomes highlighted in orange.
- Applying, Parameter names or Compensation to a Raw Data Set
  i. Select the Analysis List row containing the Protocol you wish to apply to another Data Set.
  ii. Right click on the row, and select Copy.
  iii. Right click on the Analysis List row(s) to which you wish to apply the copied information.
  iv. Select Paste Special > Protocol OR Data Set OR Compensation OR Parameters

5. Export data using the Kaluza Cytobank Plugin

Open the Kaluza Cytobank Plugin from the Plugins Tab and log in to your Cytobank account. For installation instructions see the Kaluza Cytobank IFU.

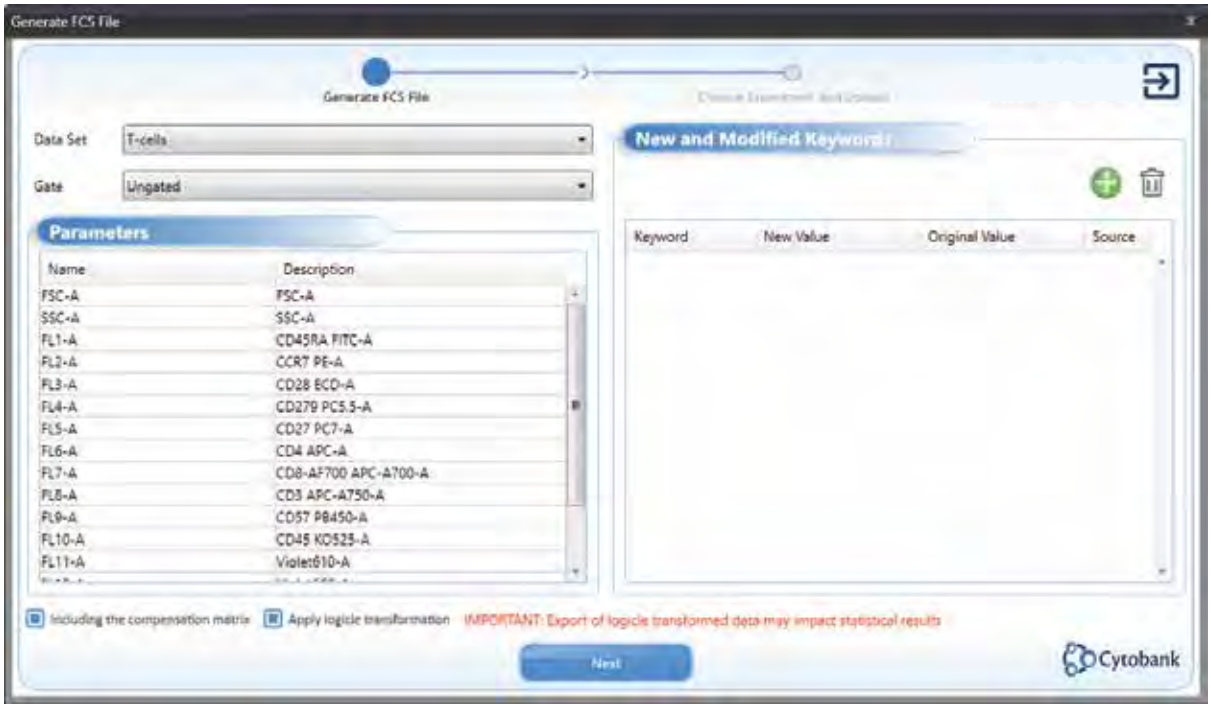

Figure 14: Kaluza Cytobank Plugin

5.1. If you are working in a Composite with several files, select the Data Set to be exported from the Data Set combo box.

The Kaluza Cytobank Plugin will use the Parameter names from Kaluza and apply them to the FCS file that will be created. By default, the compensation matrix from Kaluza Analysis Software, including any adjustments you have made, will be included in the FCS file or applied to the data when exporting logicle transformed data. Uncheck the checkbox for „Including the compensation matrix“ if you wish to export uncompensated data.

In order to apply the logicle transformation setting as defined in Kaluza, check the „Apply logicle transformation“ box. If this box is checked both compensation and logicle transformation will be applied to the data. You will not be able to adjust the compensation or transformation in the Cytobank software.

5.2. To add or modify FCS keywords click the “+“ icon in the “New and Modified Keywords“ window to select keywords. Keywords that can be edited are displayed in green. Select a keyword to edit it and add the desired value or text to the “New Value“ column.

Select “Next“ to select the location where you want to save the FCS file.

Tips For Success

- The Kaluza Cytobank Plugin will export data as FCS files irrespective of the original data format.
- For FCS from some instruments, the Cytobank platform may not be able to run SPADE due to an issue with extracting compensation from these files. This can be circumvented by exporting data using the Kaluza Cytobank Plugin with logicle transformation applied.

5.3. Select location where the FCS files will be saved. This step allows you to store the newly generated FCS files locally and to manually several files to the Cytobank platform in one upload.

- Define the location and select “Save“.

Note: Adding files to an existing experiment that were generated another way than the files already in this experiment may result in the new files having a different marker panel and needing to be scaled differently. Therefore it may be beneficial to only add FCS files exported from Kaluza to Expriments containing FCS files exported from Kaluza, also keeping the application of the logicle transformation the same across files.

5.4. Select an existing experiment or create a new experiment on the Cytobank platform that you want to add the newly generated FCS file to.

- To select and experiment, select the experiment from the Experiment Selection list or search by ID, experiment name or purpose.

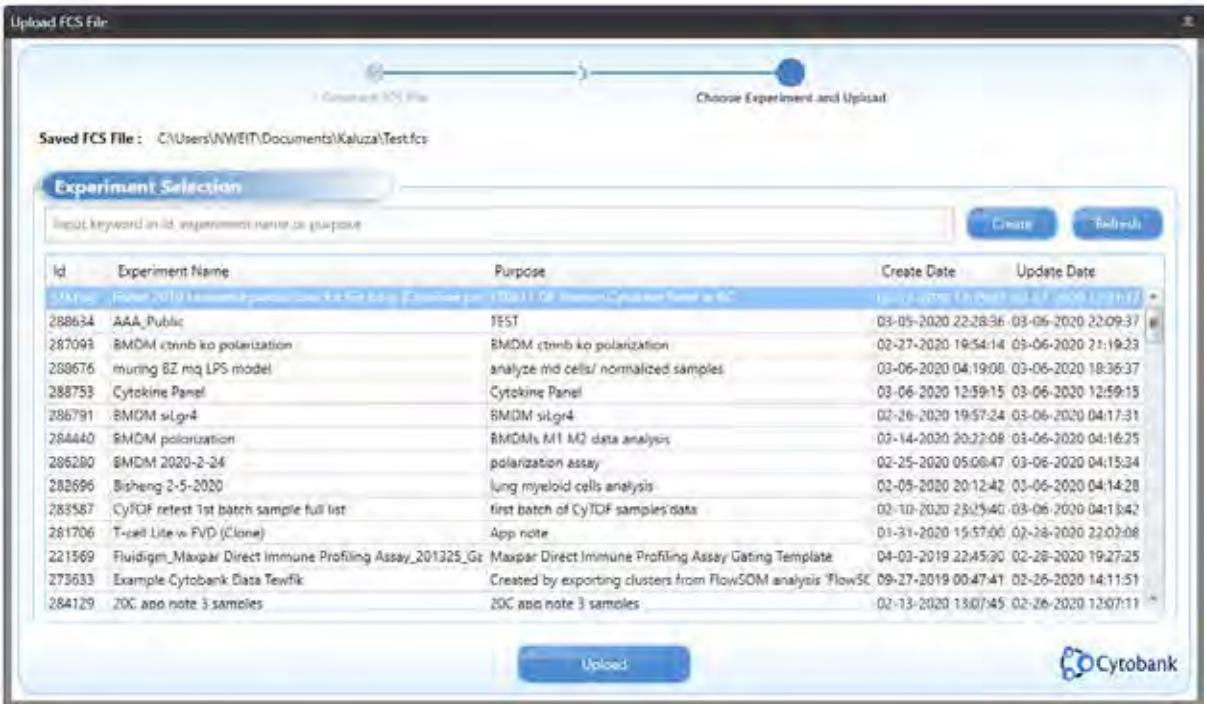

Figure 15: Kaluza Cytobank Plugin Experiment Selection

- To create a new experiment, select “Create“. You will then be prompted to enter an Experiment Name, a Purpose and Comments (optional). Select Save.
- Select Upload to upload your data to the Cytobank Platform.
- Confirm the upload request when prompted.

Your FCS file will now be uploaded to the Cytobank Platform.

Log in to your Cytobank account using a web browser to access the experiment with your newly created FCS file.

6. Analyze, collaborate on and manage your data on the Cytobank platform.

Log into your Cytobank account and select an Experiment from the Experiment Manager.

Figure 16: Cytobank Experiment Manager

6.1. Select the Experiment to open the Experiment Summary. Select Advanced Analysis or Gating to enter the experiment.

6.2. Select Scales and make the necessary adjustments to the data display. Data exported from Kaluza with the logicle transformation applied should be used with linear scaling on the Cytobank platform. No further adjustments of data transformation are required. For improved visualization for gating purposes, you may want to adjust minimum and maximum settings. This will not impact the results of the advanced algorithms.

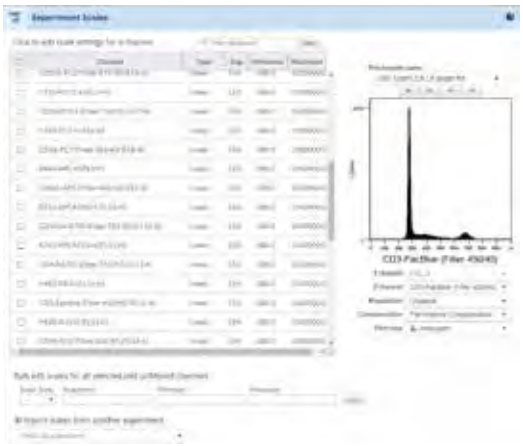

Figure 17: Experiment Scales allows adjustment of data transformation

For data exported from Kaluza without a transformation applied, please make the appropriate adjustments on the Cytobank platform.

- To change any value within a field within the scale editor, simply click on it. When done editing the field, press the "return" key on your keyboard or click outside of the field to finalize your changes.

For further information about data scaling on the Cytobank platform refer to

- **https://support.cytobank.org/hc/en-us/articles/206061527-View-and-adjust-scale-type-maximum-minimum-and-scale-argument-for-a-channel**

and

- **http://blog.cytobank.org/2012/03/30/making-beautiful-plots-data-display-basics/**

7. Advanced Data Analysis

Choose from the machine learning algorithms integrated in the Cytobank Platform for more advanced data analysis options.

In this example, viSNE and FlowSOM were run sequentially on the data set. This resulted in the generation of new fcs files with four additional derived parameters. All algorithms were run with default settings.

For details on the workflow, refer to the Use Machine Learning Algorithms to Explore the Potential of Your High Dimensional Flow Cytometry Data Example of a 20-color Panel on CytoFLEX LX Tech Note.

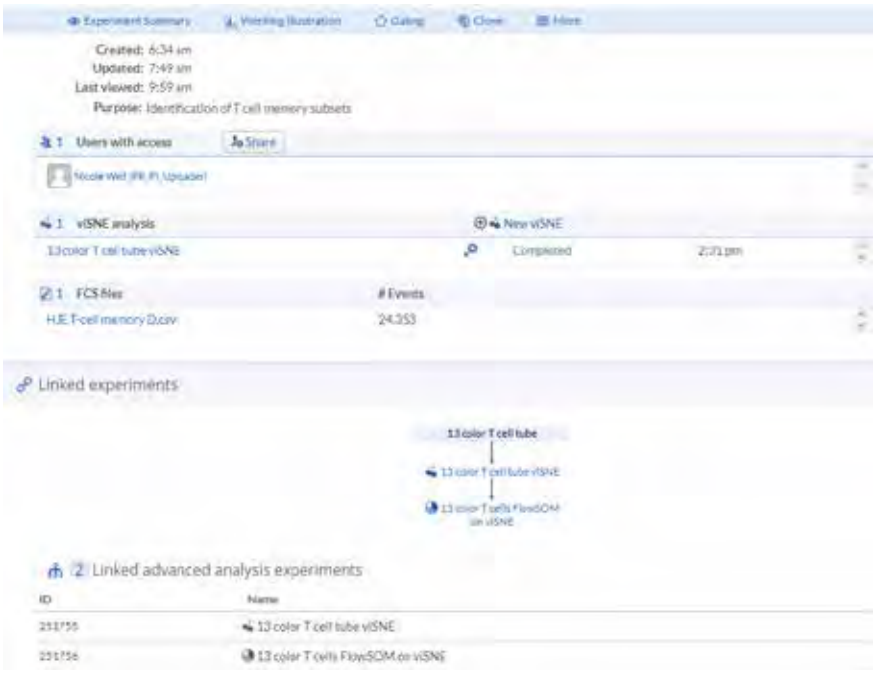

Figure 18: Experiment Manager view of the linked experiments conducted for this example.

In some instances, you might want to manually analyze the parameters created by one of the machine learning algorithms on the Cytobank platform in Kaluza.

8. Export fcs files from the Cytobank platform.

After running advanced algorithms like viSNE or FlowSOM fcs files containing additional information from the analysis are generated. These fcs files can be exported from the Cytobank platform to be stored offline or to be analyzed using other software packages. Open the viSNE or FlowSOM experiment to export these fcs files.

8.1. Navigate inside an experiment and then to Actions > Export > Download Files to enter the download interface. The file download interface allows the selection of files to download from a Cytobank experiment. The choices include FCS files, attachments, and Sample Tags from the experiment. Each FCS file can be downloaded individually, or any number of files can be zipped up to download all at once.

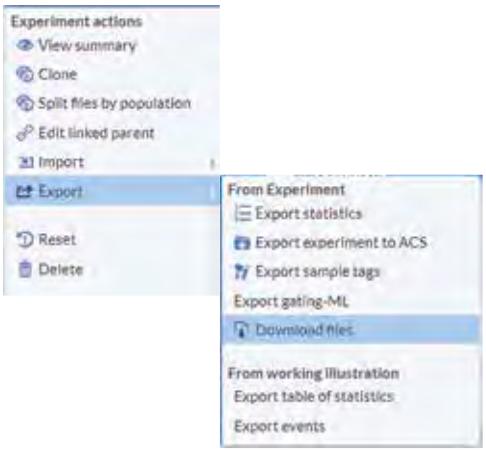

Figure 19: Experiment Action and From Experiment Window.

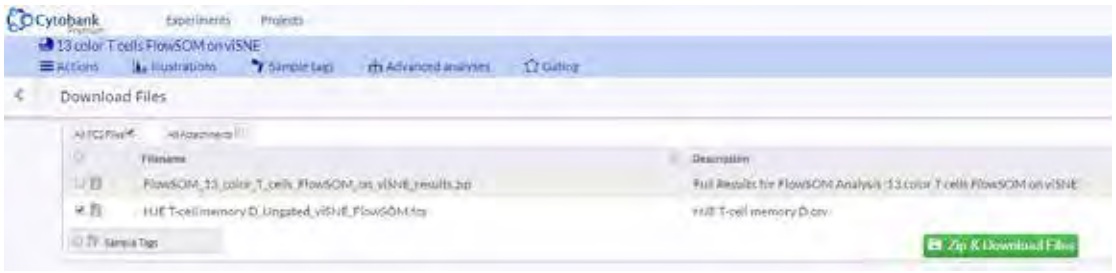

Figure 20: Download Files Window

9. Load files exported from the Cytobank platform in Kaluza.

9.1. Unzip the downloaded files and open fcs files in Kaluza (see Step 1 for detailed instructions) When the data is loaded into Kaluza, additional parameters are available (Figure 21).

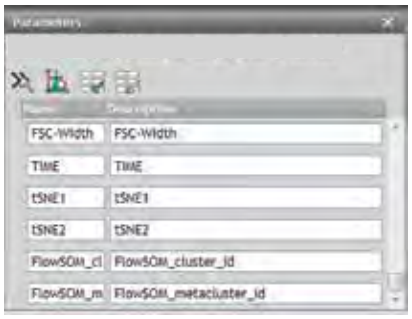

Figure 21: Parameter list as displayed in Kaluza including additional derived parameters from machine learning assisted algorithms added by the Cytobank platform.

9.2. Visualization of a viSNE map calculated on the Cytobank platform in Kaluza

- Create a dot or density plot and select tSNE 1 and tSNE 2 as parameters.
- Cytobank provides the tSNE parameters with linear scaling centered around 0. Kaluza cannot visualize negative values on a linear scale. To overcome this limitation, logicle scaling may be used with the logicle sliders pushed up and to the right to extend the linear part of the logicle scale across the whole plot.
- Adjust the minimum and maximum symmetrically around zero based on the data to aid in improving the visualization. Please note that distances may not be accurately reflected in the visualization and it is not recommended to perform further clustering analysis on the logicle scaled tSNE parameters on the logicle scaled tSNE parameters. This is intended for visualization purposes only.

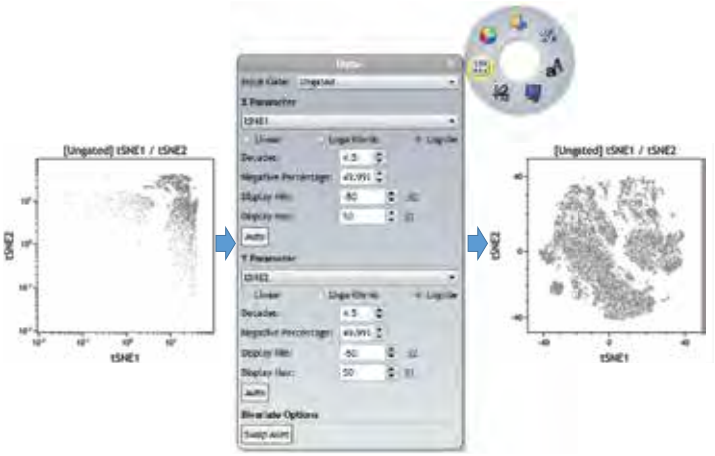

Figure 22: Dot plot showing tSNE parameters in Kaluza before and after. Final settings for the data transformation are indicated in the Data window.

9.3. FlowSOM cluster and metacluster IDs are stored as channel number corresponding to the number of the cluster or metacluster inside the fcs file. Plotting the cluster or metacluster channel in a histogram with linear scaling allows the user to gate on the population included in the cluster or metacluster.

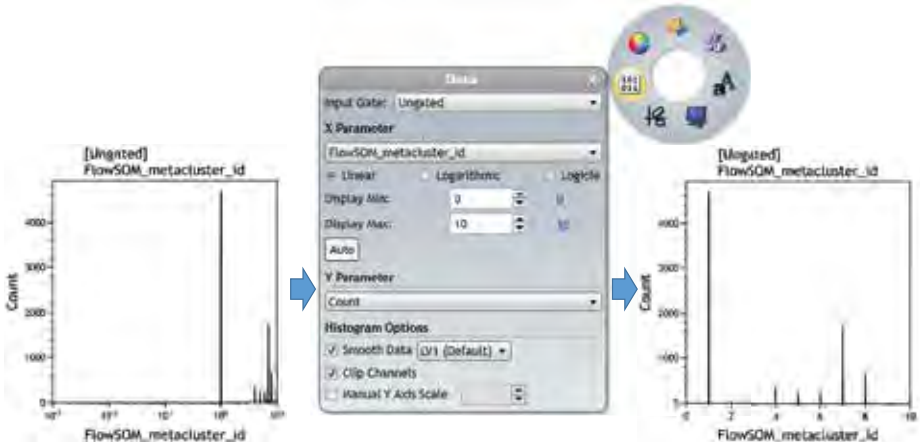

Figure 23: Histogram plot showing FlowSOM_metacluster_id parameter in Kaluza before and after. Final settings for the data transformation are indicated in the Data window.

Figure 20 shows the same Data Set visualized in Cytobank and in Kaluza.

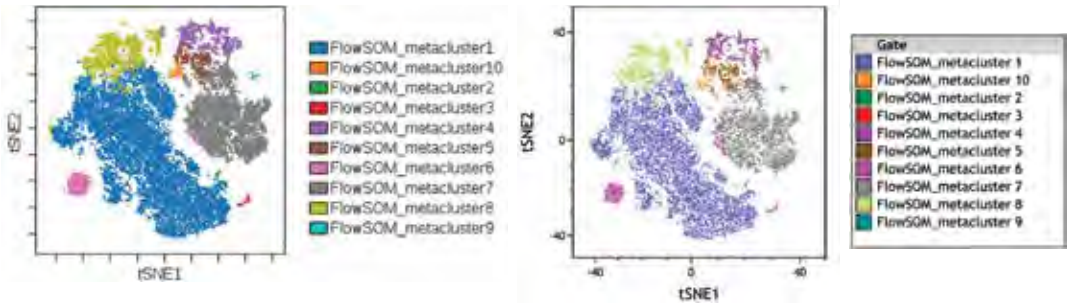

Figure 24: Comparison of data visualization on the Cytobank platform (A) and in Kaluza (B). Data was generated using a normal whole blood sample stained with DuraClone IM T Cell Subsets Tube* (PN: B53328), acquired on a CytoFLEX LX cytometer* (PN: C40324) and analyzed using Kaluza Analysis Software and Cytobank platform*. Plots are for illustration purposes only.

Conclusion

The Kaluza Cytobank Plugin allows the user to move compensated, transformed and anonymized data sets from the Kaluza desktop software to the Cytobank cloud-platform, taking advantage of the intuitive compensation adjustments supported by Kaluza and the machine learning algorithms integrated in the Cytobank platform. Combining both packages offers the user the flexibility to choose from a broad offering of features and functionalities and to customize their analysis pipeline to fit their research needs. Used in combination, Kaluza and the Cytobank platform offer a comprehensive data analysis solution.

*For detailed instructions on how to create plots and gates in Kaluza, refer to the Kaluza IFU C10986

References

1. Kotecha, N., Krutzik, P. O. & Irish, J. M. Web-Based Analysis and Publication of Flow Cytometry Experiments. Current Protocols in Cytometry 53, 10.17.1-10.17.24 (2010).
2. Rahim, A. et al. High throughput automated analysis of big flow cytometry data. Methods 134–135, 164–176 (2018).
3. Mair, F. et al. The end of gating? An introduction to automated analysis of high dimensional cytometry data: Highlights. European Journal of Immunology 46, 34–43 (2016).
4. Amir, E. D. et al. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. Nature Biotechnology 31, 545–552 (2013).
5. Qiu, P. et al. Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. Nature Biotechnology 29, 886–891 (2011).
6. Van Gassen, S. et al. FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data: FlowSOM. Cytometry 87, 636–645 (2015).

For Beckman Coulter's worldwide office locations and phone numbers, please visit Contact Us at beckman.com

FLOW-6614APP03.20