Table S1. A list of ANPELA methods for compensation, transformation, normalization, and signal clean in the SCP data processing. ANPELA includes a total of 6 compensation methods, 15 transformation methods, 6 normalization methods, and 4 signal cleaning methods. Each method is represented by a unique three-letter abbreviation (Abb.) for use in the SCP data processing steps. If no method is applied at a given step, the code '**NON**' is used to indicate the absence of any method. Both the introduction, limitation(s) and research application(s) of each method are included. FC, flow cytometry; MC, mass cytometry.

Abb.	Method Name	Brief Introduction, Application(s) and Limitation(s) of Each Method
		Step1. Data Compensation
ATS	AutoSpill	**Method's Introduction**: This method automated gates cells to eliminate cellular debris and other non-cellular contamination, then calculates an initial spillover matrix based on robust linear regression and iteratively refines to reduce error. **Method's Limitation**: Applicable only to FC; Requires single-color controls. **Research Application**: It was used to compensate the flow cytometry data in pathological research for elucidating the immunopathology underlying COVID-19 severity.
CAT	CATALYST	**Method's Introduction**: This method calculates spillover matrix based on the spillover observed for bead-based single-stained populations, then applies the compensation matrix from the solved linear system ("NNLS" or "classical") to the bead and cell samples. **Method's Limitation**: Applicable only to MC; Requires single-stained beads controls. **Research Application**: It was applied to compensate the mass cytometry data in a type 2 diabetes mellitus study to quantify proteins of pancreatic exocrine, immune cells and stromal components.
CTS	CytoSpill	**Method's Introduction**: This method leverages the understanding of spillover sources to effectively constrain the estimation of the spillover matrix, and employs a non-negative least squares algorithm, to obtain the real (compensated) data that better reflects the true biological signals. **Method's Limitation**: Applicable only to MC.

		Research Application: It was utilized to compensate the mass cytometry data in the single-cell profiling of systemic immune and inflammatory responses study to correct for signal overlap between markers.
FLC	FlowCore	**Method's Introduction**: This method provides a systematic approach for estimating the spillover matrix by analyzing single-color control samples, then inverts this spillover matrix to derive a compensation matrix and compensates the experimental dataset based on the compensation matrix. **Method's Limitation**: Applicable only to FC; Requires single-color controls or spillover matrix. **Research Application**: It was used to compensate the flow cytometry data in an immunological study for investigating memory B cells and humoral responses in 145 subjects after vaccine administration.
MTC	MetaCyto	**Method's Introduction**: This method uses the pre-calculated spillover matrix of each FCS file to compensate the corresponding data, effectively correcting for fluorescence spillover and accommodating variations across samples and experimental runs. **Method's Limitation**: Applicable only to FC; Requires pre-calculated spillover matrix of FCS file(s). **Research Application**: It was employed to process the data of an immunomes project to ensure the high quality of high-dimensional flow cytometry.
SPR	spillR	**Method's Introduction**: This method develops a nonparametric finite mixture model based on an expectation-maximization algorithm, utilizing the complete bead distributions to model the spillover probability instead of estimating a spillover matrix. **Method's Limitation**: Applicable only to MC; Requires single-stained beads controls. **Research Application**: It is an original method which was implemented on datasets from CATALYST, proving that it can reduce the risk of introducing biases that could affect downstream analyses.
	Step2. Data Transformation	
ACS	Arcsinh Transformation	**Method's Introduction**: This method first scales the data by dividing it by a specified factor, and then applies an inverse hyperbolic sine (arcsinh) transformation to allow the data to retain linearity near zerowhile gradually adopting a logarithmic-like behavior as values move away from zero.

		Method's Limitation: No significant limitations. **Research Application**: It was applied in a fibroblast study for revealing stromal heterogeneity and identifying coordinated relationships between mesenchymal and immune cell subsets.
ANN	Asinh with Non-negative Value	**Method's Introduction**: This method represents a modified version of the arcsinh transformation, which begins by subtracting a specified constant from the raw data to shift the entire dataset, then sets negative values to zero and uses arcsinh transformation to process the data. **Method's Limitation**: No significant limitations. **Research Application**: It is the suggested transformation method of X-shift which processes data using k-nearest-neighbor estimation of cell density and arranges populations by marker-based classification.
ARN	Asinh with Randomized Negative Value	**Method's Introduction**: This method is a modified version of arcsinh transformation that first subtracts a specified constant from the raw data to adjust the baseline and randomizes negative values to a normal distribution, then uses arcsinh transformation to process the data. **Method's Limitation**: No significant limitations. **Research Application**: It is the recommended transformation method for PhenoGraph, which models the data space using a nearest-neighbor graph and detects density of nodes to isolate diverse subpopulations.
BEP	Biexponential Transformation	**Method's Introduction**: This method is an over-parameterized inverse of the hyperbolic sine, and can provide a linear representation of data around zero and a logarithmic representation of the data at higher intensity values, with a smooth transition between the two extremes. **Method's Limitation**: No significant limitations. **Research Application**: It was integrated into a high-dimensional flow cytometry analytical protocol to diminish the unwanted variation of flow cytometry data.
BOX	Box-Cox Transformation	**Method's Introduction**: This method is a type of power transformation, which is designed to stabilize variance and invert convert data into a more symmetric distribution, thereby making the data more normally distributed and enhancing the validity of downstream results. **Method's Limitation**: No significant limitations.

		Research Application: It was utilized in a robust model-based clustering approach to handle the issues of transformation and outlier identification simultaneously.
CLR	Centered Log Ratio Transformation	**Method's Introduction**: This method calculates the natural logarithm of the ratio between each component and the geometric mean of all components within a sample, thereby converting relative proportions into a set of real numbers that can be analyzed using standard statistical techniques. **Method's Limitation**: Not applicable for dataset containing 0 or negative value. **Research Application**: It was used to convert cell population frequencies into an unconstrained real number scale, enhancing the sensitivity and accuracy in detecting fluorescence anomalies.
FVS	FlowVS Transformation	**Method's Introduction**: This method utilizes Bartlett's likelihood-ratio test to select the parameters of the transformation for each channel, and then processes the data using the inverse hyperbolic sine (arcsinh) transformation with corresponding parameter. **Method's Limitation**: No significant limitations. **Research Application**: It has been applied in an immunological study for profiling memory Th cell subsets with flow cytometry and using the machine learning algorithm FlowSOM to interpret the data.
HPL	Hyperlog Transformation	**Method's Introduction**: This method adds together a linear function and an exponential function to combine the characteristics of both linear and exponential functions to achieve a versatile scaling approach for data, addressing a wide range of data distributions. **Method's Limitation**: No significant limitations. **Research Application**: It was utilized to achieve visual purposes in a protein-protein interactions study, which enables a quantitative yeast two-hybrid system.
LGT	Logicle Transformation	**Method's Introduction**: This method provides a particular generalization of the hyperbolic sine function with an additional adjustable parameter than linear or logarithmic functions to make data asymptotically linear in the vicinity of zero and asymptotically logarithmic at higher (positive and negative) values. **Method's Limitation**: No significant limitations.

		Research Application: It was applied in a kidney allograft rejection study for investigating whether circulating EC could serve as an earlier and less invasive biomarker for allograft rejection.
LIN	Linear Transformation	**Method's Introduction**: This method enables the scaling and translation of raw data using a linear function by applying a linear equation of the form $y=ax+b$, where a represents the scaling factor and b denotes the translation constant. **Method's Limitation**: No significant limitations. **Research Application**: It was used to adjust the flow cytometry measured values to fit a desired range or format before stored in FCS files.
LNT	Ln Transformation	**Method's Introduction**: This method employs a logarithmic transformation based on the natural logarithm, denoted as ln(x), to convert data to a logarithmic scale to compress large values and enhance the visibility of smaller values. **Method's Limitation**: Not applicable for dataset containing 0 or negative value. **Research Application**: It was applied in a *Plasmodium falciparum* study for assessing the feasibility of *P. falciparum* quantification in natural infections of children by flow cytometry.
LOG	Log Transformation	**Method's Introduction**: This method performs a nonlinear conversion transforming the data to a logarithmic scale, which is useful to correct for heteroscedasticity and to change skewed distributions into more symmetric, Gaussian distributed peaks. **Method's Limitation**: Not applicable for dataset containing 0 or negative value. **Research Application**: It was employed in a microbial study to assess how iron associated with organic ligands influences the abundances and diversity of phytoplankton, microbes, and viruses.
QUA	Quadratic Transformation	**Method's Introduction**: This method uses a quadratic function to transform raw data based on a mathematical function of the form $y=ax2+bx+c$, where a, b, and c are coefficients that define the shape and position of the parabola. **Method's Limitation**: No significant limitations.

		Research Application: It was applied in a proteomics study to investigate the relationship between time intervals and mass-to-charge ratios in order to map the abundance of particles detected by the plate.
SST	Split Scale Transformation	**Method's Introduction**: This method consists of a logarithmic scale at high values and a linear scale at low values with a fixed transition point chosen so that the slope (first derivative) of the transform is continuous at that point. **Method's Limitation**: Not applicable for dataset containing 0 or negative value. **Research Application**: It is one of the built-in data transformation methods in Gating-ML, and is considered highly valuable for display or analysis of cytometry data.
TRU	Truncate Transformation	**Method's Introduction**: This method is designed to manage the lower limit of the dynamic range of data by setting a predefined cutoff value by truncating any data points that fall below this cutoff and replacing those lower values with the cutoff itself. **Method's Limitation**: No significant limitations. **Research Application**: It was utilized to clean the erroneous measurements of flow cytometry data in an open-source software system.
		Step3. Data Normalization
BEN	Bead-based Normalization	**Method's Introduction**: This method identifies isotope-containing bead events, converts raw data to local medians, computes averages across all files, uses global means to calculate slopes for each time point, and multiplies these slopes by the corresponding data. **Method's Limitation**: Applicable only to MC. **Research Application**: It was applied to remove unwanted variance in a cancer study aiming to identify cancer cell-autonomous determinants of obesity-induced PM breast cancer risk.
GSN	GaussNorm	**Method's Introduction**: This method identifies landmarks based on the pre-determined maximum number M and kernel density estimation, then classifies the detected landmarks into M classes and transforms the landmarks of each sample independently to align the landmarks.

		Method's Limitation: No significant limitations. **Research Application**: It was used in a B cell study aiming to identify an MS-specific immune cell population by deep immune phenotyping and relate it to soluble signaling molecules in CSF.
MEA	Mean Normalization	**Method's Introduction**: This method adjusts the data of each column by subtracting the corresponding mean of each column and then dividing the result by the range (the difference between the maximum and minimum values), effectively centering the data around zero and scaling it to a fixed range. **Method's Limitation**: No significant limitations. **Research Application**: It was employed in a study investigating redox stress in live cells for a clearer comparison of oxidative stress responses across different conditions and cell lines.
MMN	Min-max Normalization	**Method's Introduction**: This method applies a linear transformation by subtracting the minimum value from each data point and then dividing the result by the range (the difference between the maximum and minimum values), which scales the data to a specified range, typically [0, 1]. **Method's Limitation**: No significant limitations. **Research Application**: It was utilized in a study on tetanus vaccination to scale the data to [0, 1], ensuring balanced feature contributions for clustering of plasmablast responses.
WPS	WarpSet	**Method's Introduction**: This method identifies landmarks based on the kernel density estimation, then estimates the most likely total number K of landmarks and performs K-means clustering to classify landmarks, finally transforms the data using the warping function estimated for each sample. **Method's Limitation**: No significant limitations. **Research Application**: It was integrated into an open source, extensible graphical user interface iFlow for normalization before comparisons can be made across samples.
ZSC	ZScore	**Method's Introduction**: This method standardizes the data by subtracting the mean from each value and then dividing by the standard deviation, eliminating differences in scale (or units) across the data and ensuring that all variables contribute equally to further analysis.

		Method's Limitation: No significant limitations. **Research Application**: It was applied in the peripheral B cell subset analysis, ensuring that each feature contributed equally to the clustering and classification of cell populations.
		Step4. Signal Clean
FAI	FlowAI	**Method's Introduction**: This method first removes surges deviations from the median flow rate, then eliminates the acquisition regions whose statistics are shifted from the most stable acquisition region and deletes outliers in the lower limit and margin events from the upper limit of the dynamic range. **Method's Limitation**: No significant limitations. **Research Application**: It was integrated into a practical workflow to properly prepare spectral flow cytometry data for high dimensional analysis.
FCL	FlowClean	**Method's Introduction**: This method divides the time of data into equally sized bins and employs changepoint analysis (CPA), a tool for studying trends in a data series, to flag cells that occur within time frames exhibiting unusual ratios of cell populations. **Method's Limitation**: Only useful for more than 30,000 cells; Applicable only to FC. **Research Application**: It was recommended to use for removing suspect events in a guideline for the use of flow cytometry and cell sorting in immunological studies.
FCU	FlowCut	**Method's Introduction**: This method separates events along the time axis into equally sized segments, utilizes three tests to assess the need for file cleaning, then calculates the mean, median, several percentiles, skewness, and variation of the flow signal and removes any erroneous measurements. **Method's Limitation**: No significant limitations. **Research Application**: It was utilized to check for aberrant signal patterns or events in optimised multicolor immunofluorescence panel (OMIP) study.
PQC	PeacoQC	**Method's Introduction**: This method performs automated quality control on cytometry data by identifying and removing low-quality events based on their proximity to density peaks, effectively

	cleaning data to ensure high-quality, reliable datasets for downstream analysis. **Method's Limitation**: No significant limitations. **Research Application**: It is the recommended signal clean method for FlowSOM, an unsupervised clustering algorithm that identifies cell populations in high-dimensional cytometry data.