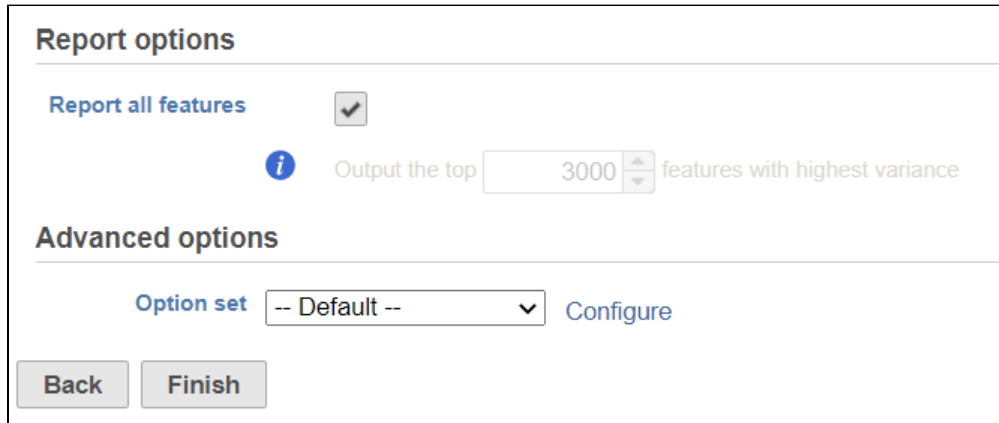


SCTransform

SC transform task performs the variance stabilizing normalization proposed in [1]. The task's interface follows that of SCTransform() function in R [2]. SCTransform v2 [3] provides the ability to perform downstream differential expression analyses besides the improvements on running speed and memory consumption. v2 is the default method in Flow.

We recommend performing the normalization on a single cell raw count data node. Select SCTransform task in Normalization and scaling section on the pop-up menu to invoke the dialog (Figure 1).

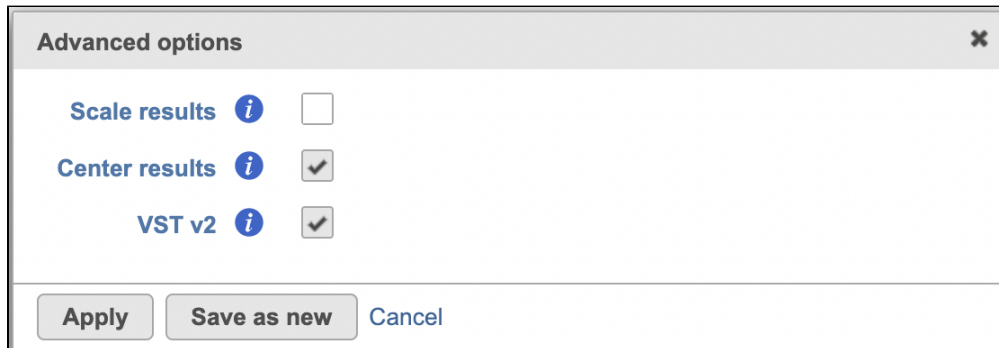


The dialog box titled "Report options" has a section "Report all features" with a checked checkbox. Below it, an information icon is followed by the text "Output the top" and a text input field containing "3000", followed by "features with highest variance". Below this is a section "Advanced options" with an "Option set" dropdown menu showing "-- Default --" and a "Configure" link. At the bottom are "Back" and "Finish" buttons.

Figure 1. When a data node containing a single cell count matrix is selected, SCTransform is available in the toolbox

By default, it will generate report on all the input features. Unchecking the *Report all features*, user can limit the results to a certain number of features with highest variance.

In Advanced options, users can click Configure to change the default settings (Figure 2).



The dialog box titled "Advanced options" has a close button (X) in the top right. It contains three options: "Scale results" with an information icon and an unchecked checkbox; "Center results" with an information icon and a checked checkbox; and "VST v2" with an information icon and a checked checkbox. At the bottom are "Apply", "Save as new", and "Cancel" buttons.

Figure 2. Advanced configure options

Scale results: Whether to scale residuals to have unit variance; default is FALSE

Center results: When set to Yes, center all the transformed features to have zero mean expression. Default is TRUE.

VST v2: Default is TRUE. When set to 'v2', it sets method = glmGamPoi_offset, n_cells=2000, and exclude_poisson = TRUE which causes the model to learn theta and intercept only besides excluding poisson genes from learning and regularization; If default is unchecked, it uses the original sctransform model (v1), it will only generate SC scaled data node.

There are two data nodes generated from this task (if VST v2 option is checked as default):

SC scaled data: it is a matrix of normalized values (residuals) that by default has the same size as the input data set. This data node is used to perform downstream exploratory analysis e.g. PCA, Seurat3 integration etc (Figure 3), this data node is not recommend to use for differential analysis.

SC corrected data: is equivalent to the 'corrected counts' in data slot generated after PrepSCTFindMarkers task in the SCT assay in Seurat object. It is used for downstream differential expression(DE) analyses (Figure 3).

Note: When perform DE analysis with Hurdle, the 'shrinkage of error term variance' option might need to turn off depending on the dataset. Similarly, the 'Lognormal with shrinkage/voom' option needs to turn off when run DE with GSA.

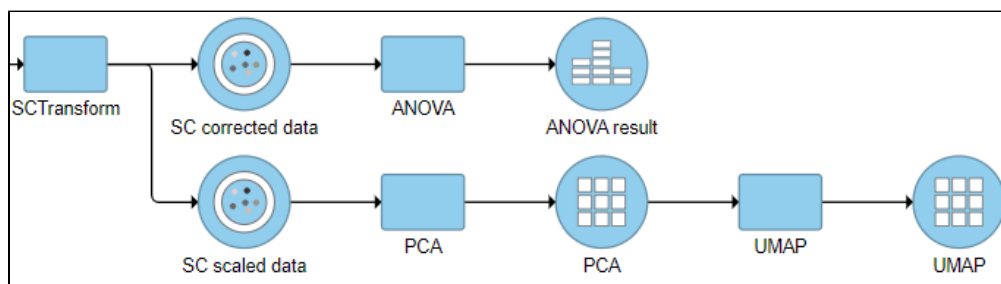


Figure 3. Downstream analysis after SCTransform v2 task in Flow.

References

1. Christoph Hafemeister, Rahul Satija. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. <https://doi.org/10.1101/576827>
2. SCTransform() documentation <https://www.rdocumentation.org/packages/Seurat/versions/3.1.4/topics/SCTransform>
3. Saket Choudhary, Rahul Satija. Comparison and evaluation of statistical error models for scRNA-seq. <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-021-02584-9>

Additional Assistance

If you need additional assistance, please visit [our support page](#) to submit a help ticket or find phone numbers for regional support.



Your Rating: ☆☆☆☆☆ Results: ★★★★★ 12 rates