Step-by-step Guide to Preprocess Expression Data using FlexStat Pipeline - Sample data

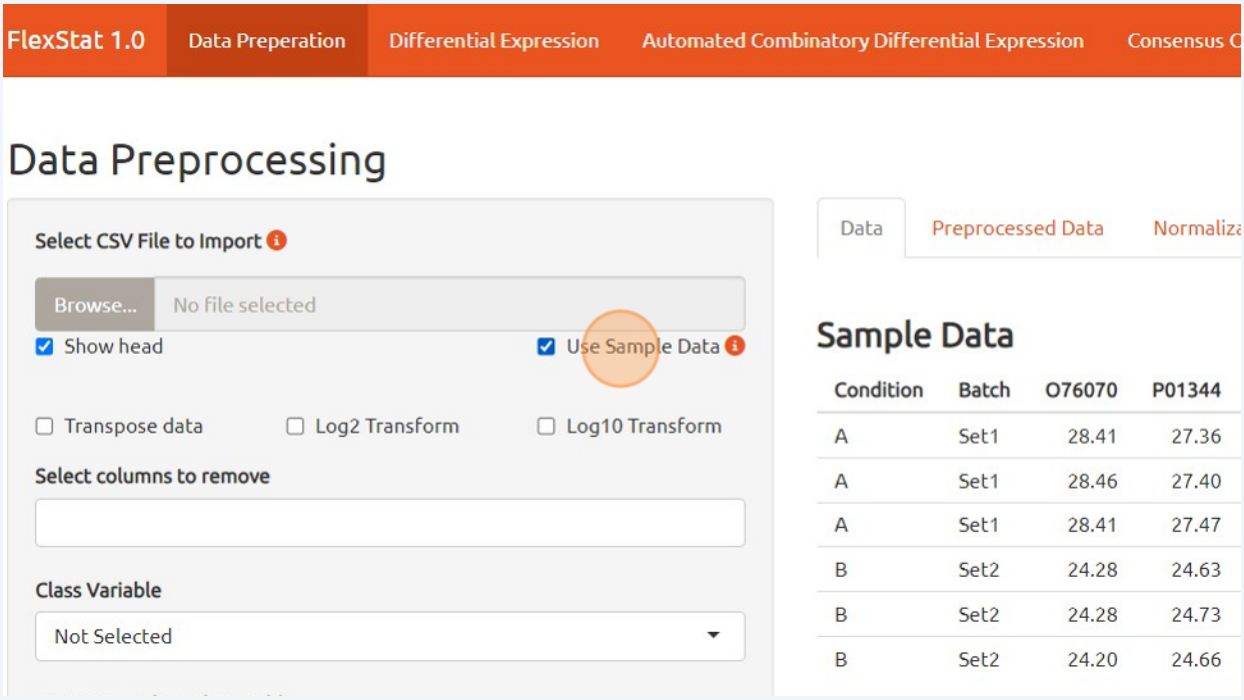
This feature facilitates preprocessing expression data with experimentally generated data. It involves missing value imputation and data normalization where users can specify the algorithm, and method to be used.

This tutorial is based on sample data into the application.

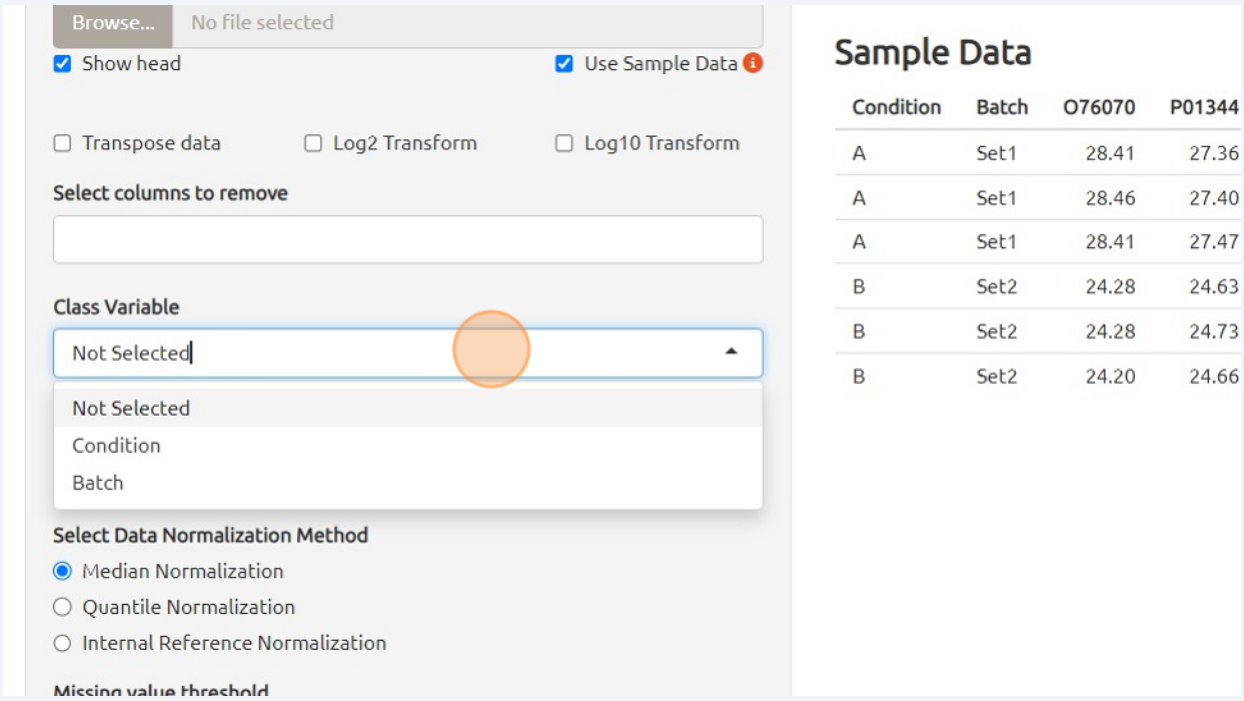
1 Navigate to https://jglab.shinyapps.io/flexstatv1-pipeline-only/

2 Go to "Data Preparation" tab.

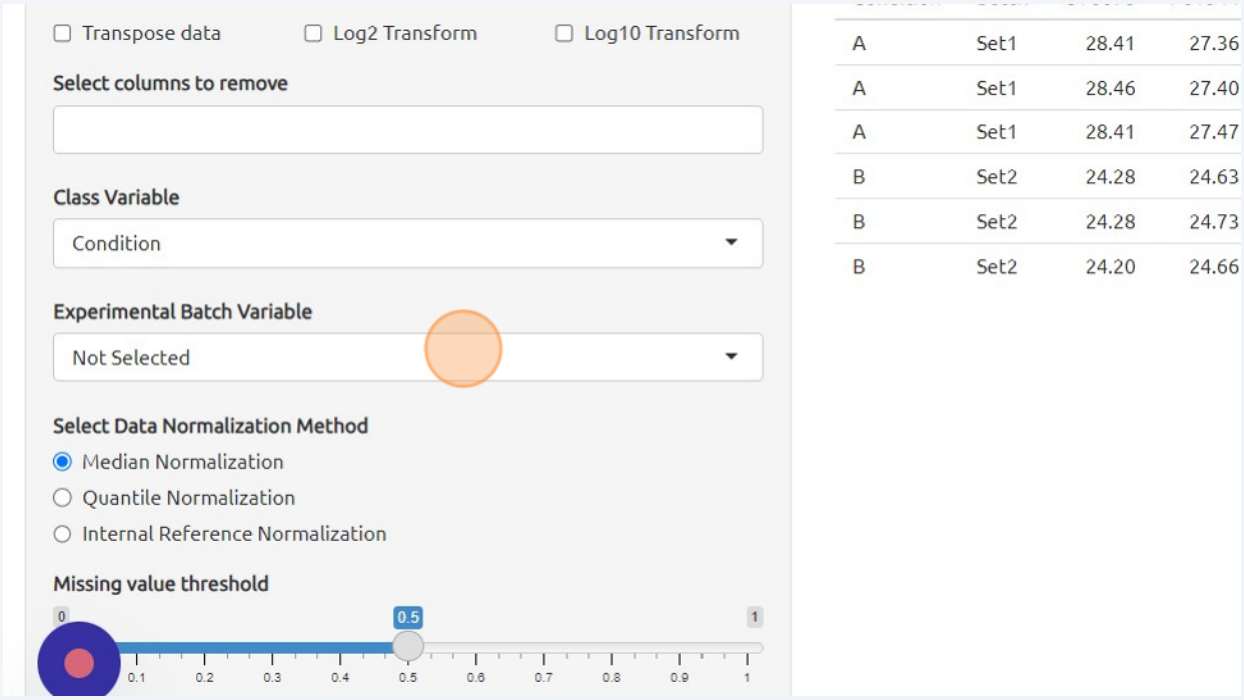
3 Click "Use Sample Data"



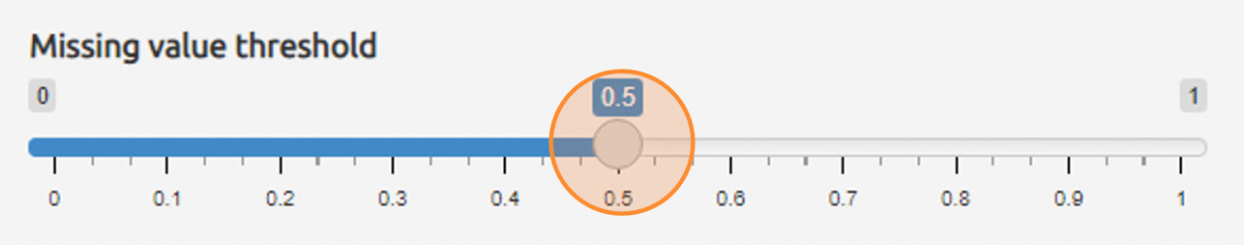
4 Select 'Class' variable to generate quality control plots.



5 Select "Experimental batch variable" to be used in normalization.

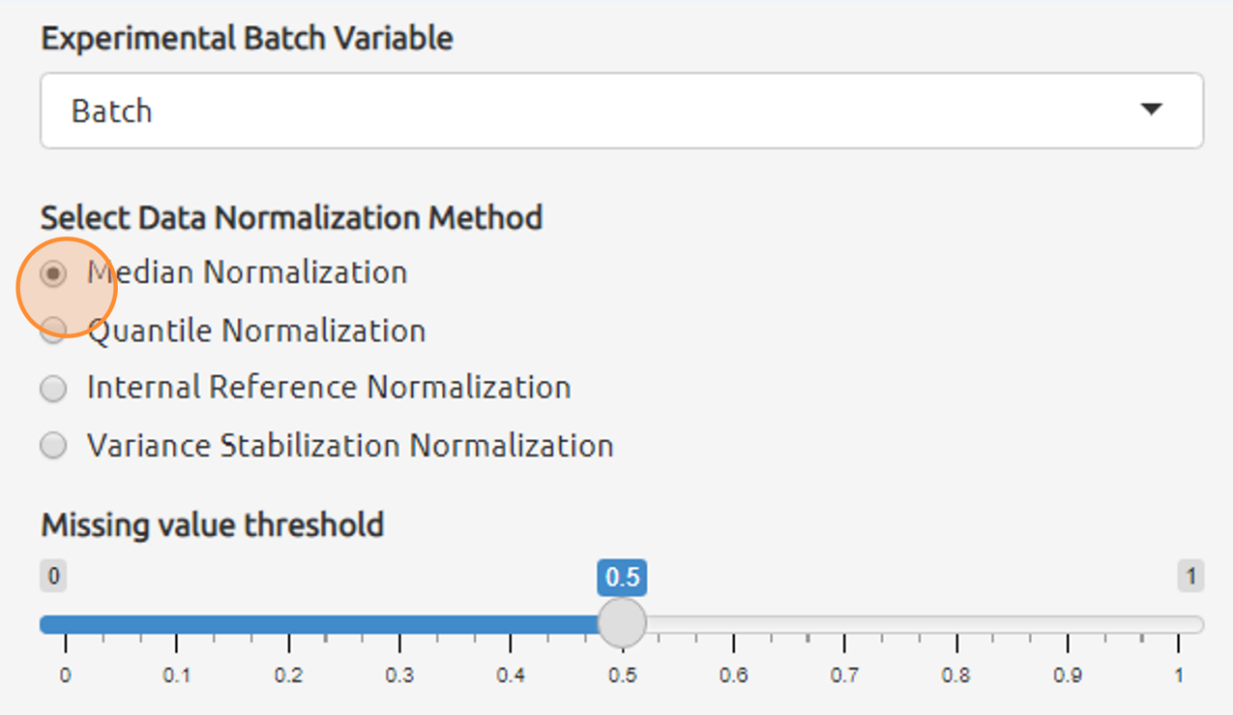


6 Select missing value threshold. Default is 0.5 which means proteins with 50% or less missing values will be considered for imputation.



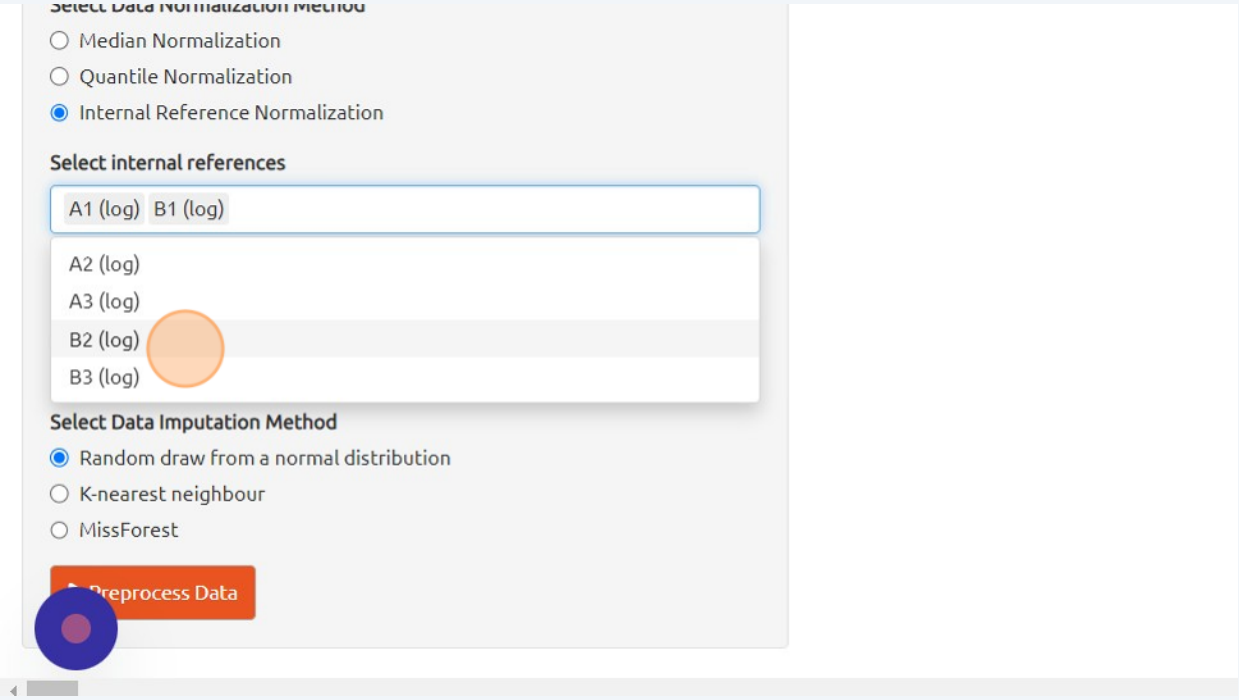
7 Select Data Normalization method

1. Median normalization
2. Quantile normalization
3. Internal Reference Normalization: Recommended to use for label-based data
4. if selected, select the corresponding internal reference to be used
5. Variance Stabilization Normalization: Recommended to use for label-free data

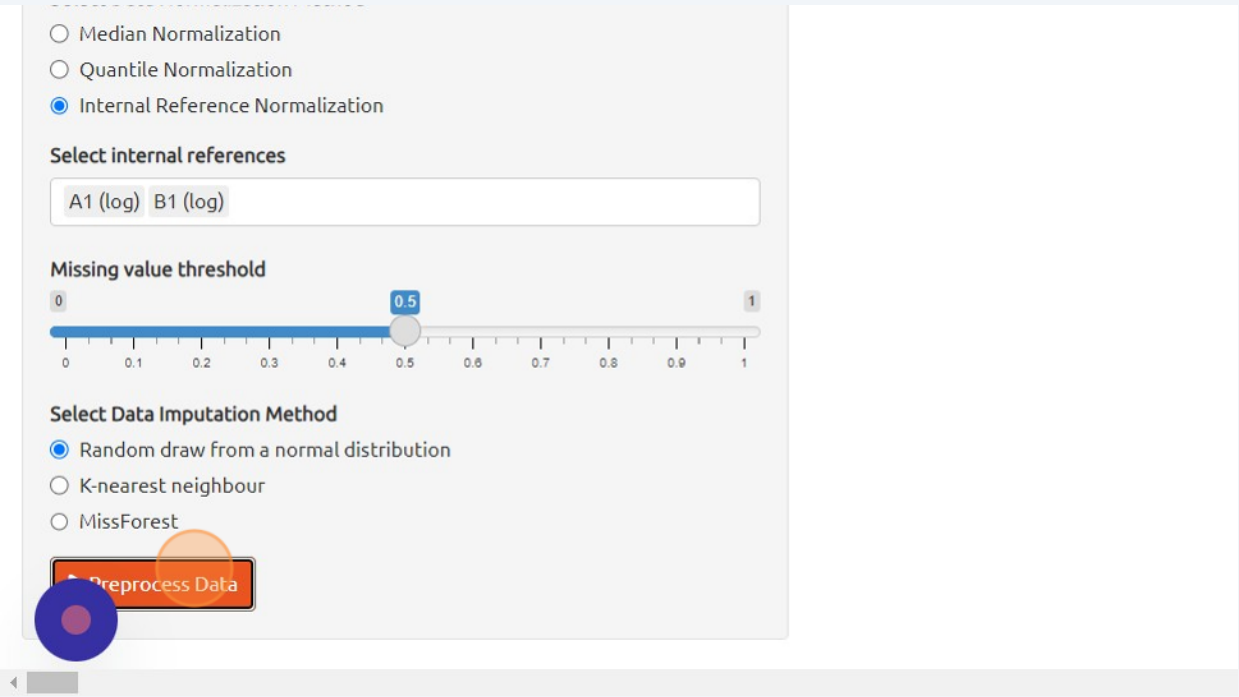


8 **Internal Reference Normalization:** if selected, then select the corresponding internal reference to be used.

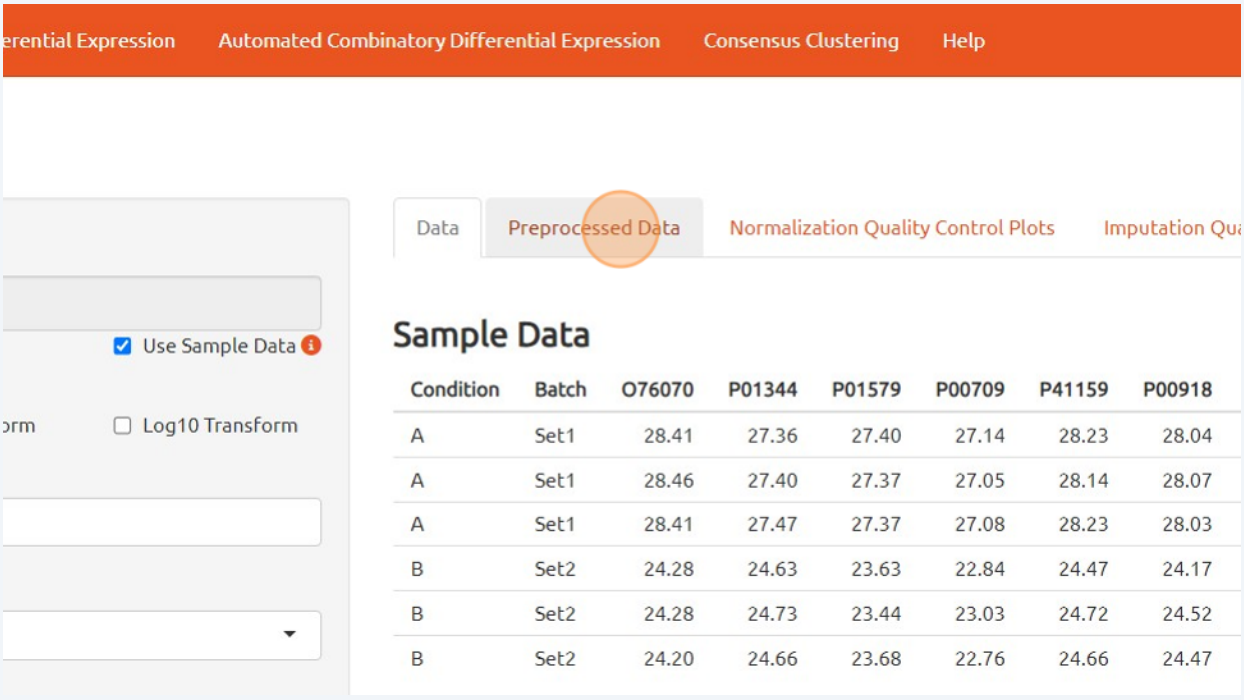
Select the corresponding internal reference to be used. Here we select "A1 (log)" and "B1 (log)".



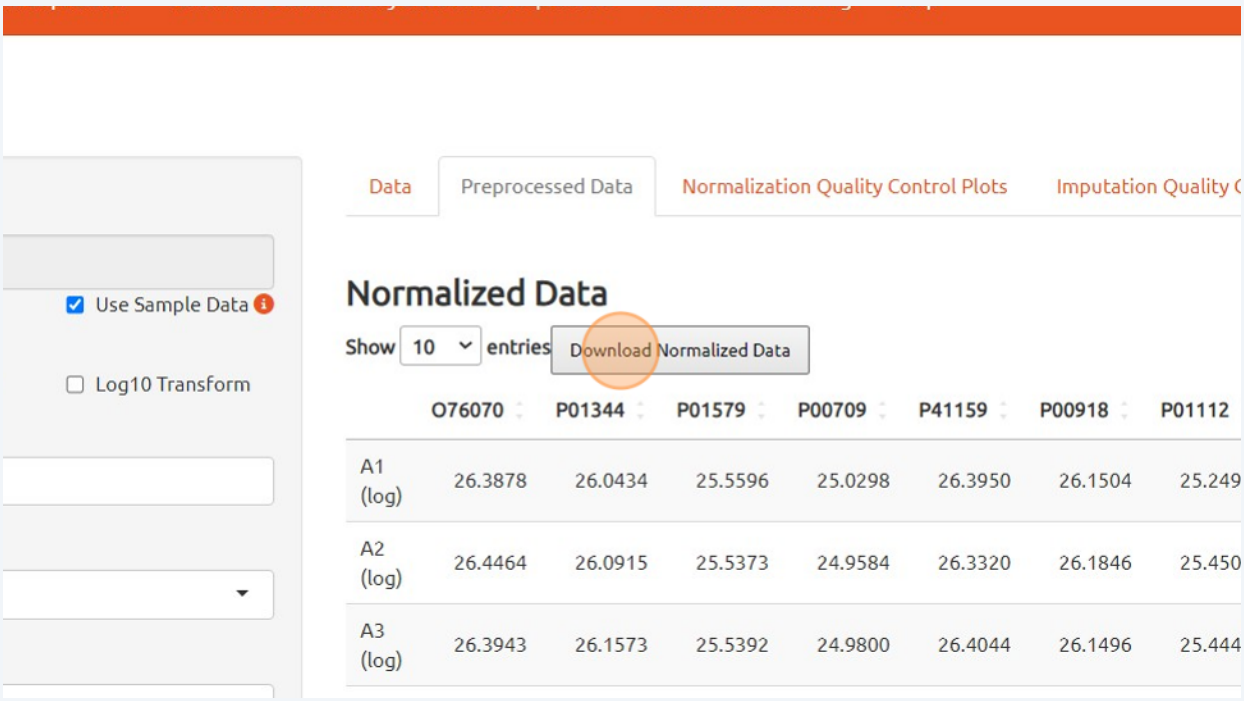
9 Click "Preprocess Data"



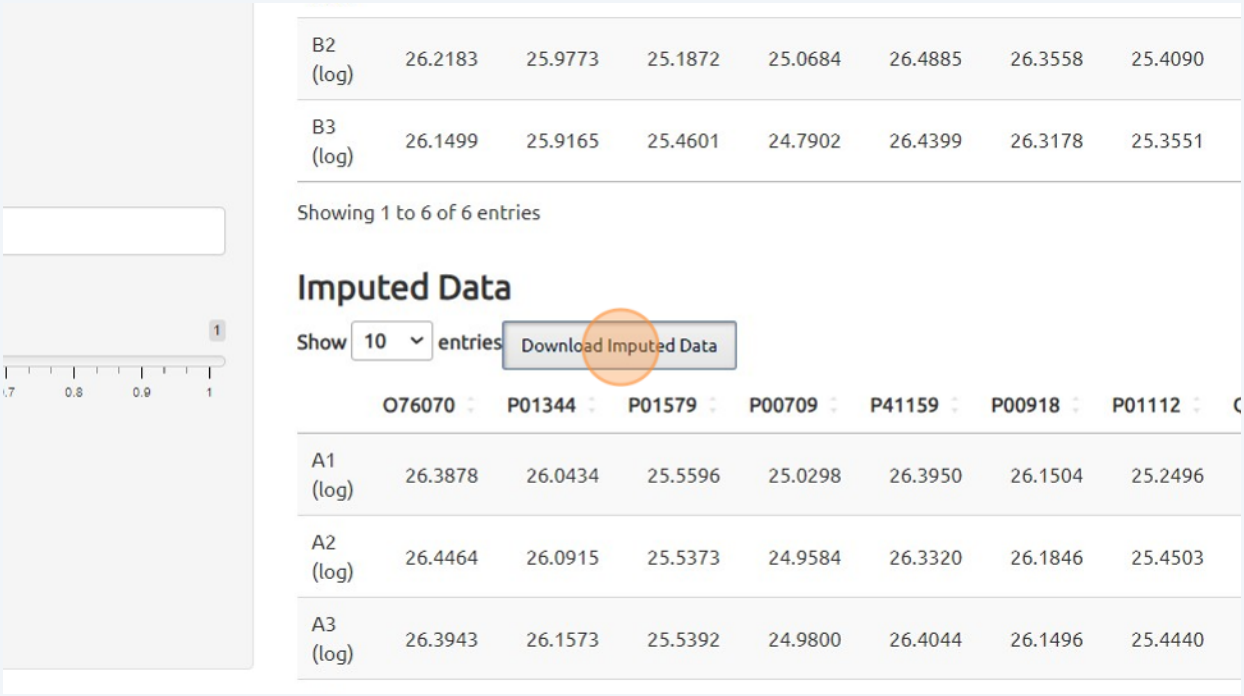
10 Click "Preprocessed Data" to check the preprocessing results.



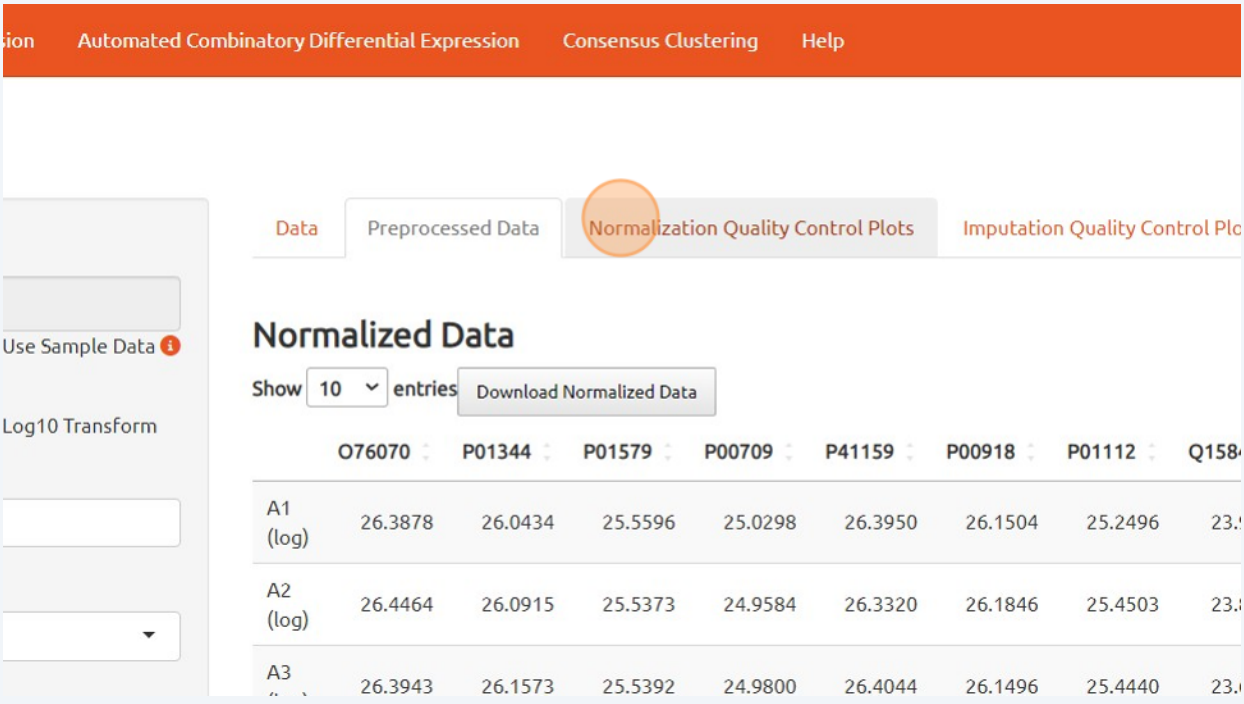
11 Click "Download Normalized Data" to download the normalized data.



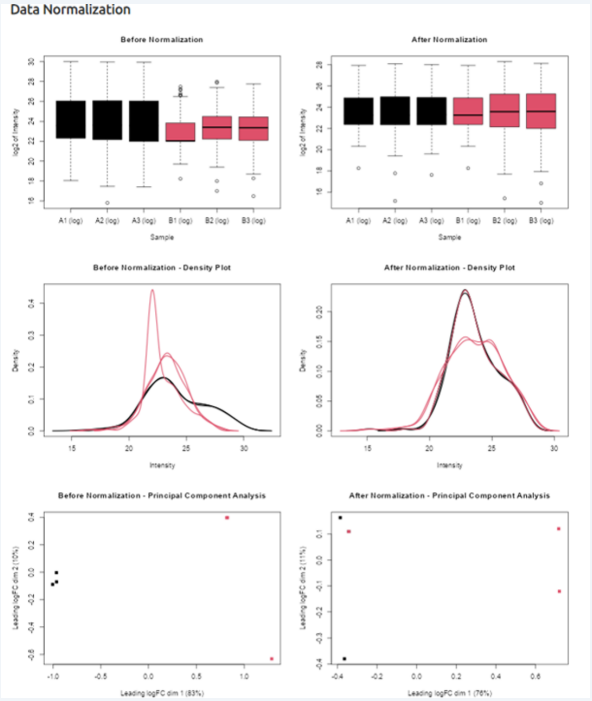
12 Click "Download Imputed Data" to download the imputed data.



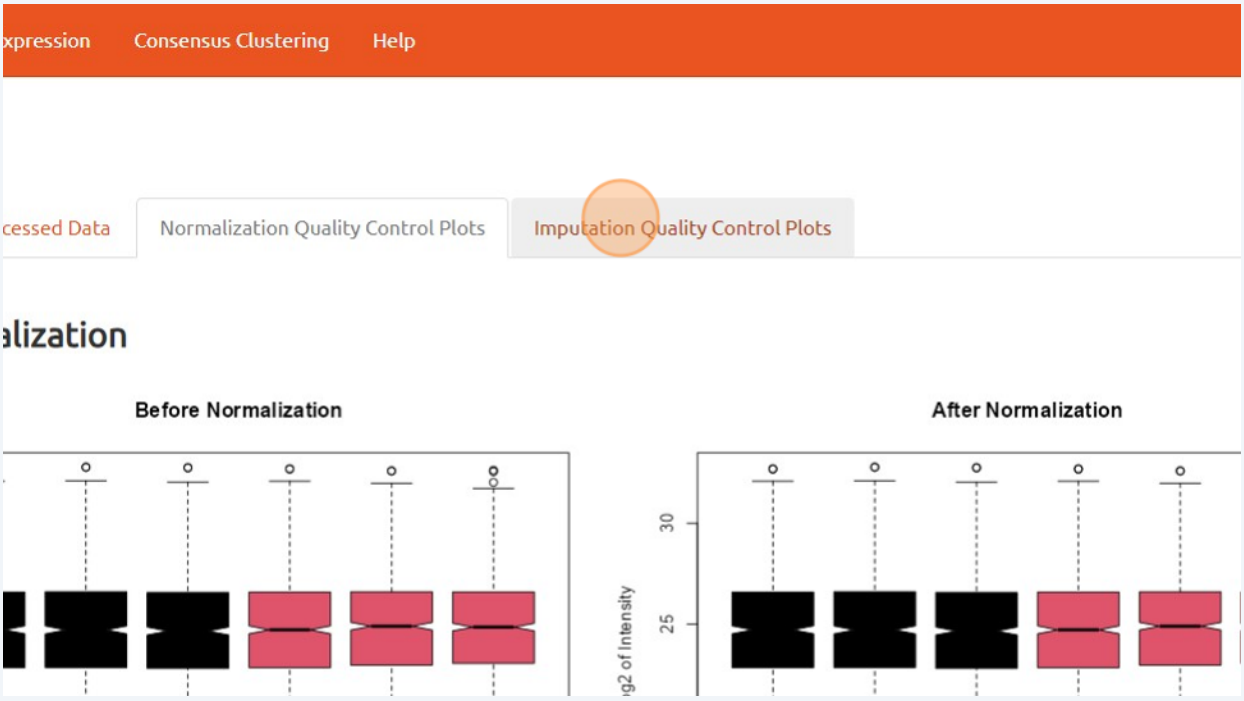
13 Click "Normalization Quality Control Plots" to view the outputs of normalization steps.



14 Click "Normalization Quality Control Plots" to view the outputs of normalization steps.



15 Click "Imputation Quality Control Plots"



16 Click "Imputation Quality Control Plots"