Bulk RNA-seq Cheatsheet

The tables below consist of valuable codes that will help you through this module. Each table represents a different library/tool and its corresponding commands.

Please note that these tables are not intended to tell you all the information you need to know about each command. The hyperlinks found in each piece of code will take you to the documentation for further information on the usage of each command.

Library/Package	Piece of Code	What it's called	What it does
Base R	`list.files()`	List files	Produces a character vector of files or directories in the specified directory
Base R	`names()`	Names	Gets or sets the names of an object
Base R	`colnames()`	Column names	Gets or sets the column names of a matrix or data frame
Base R	`all.equal()`	All equal	Checks if two R objects are nearly equal
Base R	`attr()`	Object Attributes	Gets or sets the attributes of an object
Base R	`rowSums()`	Row Sums	Returns the sum of the rows in a numeric matrix-like object (i.e.. a matrix, data.frame, etc.)
Base R	`relevel()`	Relevel	Reorders the levels of a factor as specified
Base R	`summary()`	Object Summary	Returns a result summary of an object
Base R	`as.data.frame()`	Data Frame	Checks if an object is a data.frame, and transforms the object into one, if possible

Library/Package	Piece of Code	What it's called	What it does
DESeq2	`rlog()`	Regularized Log Transformation	Returns log2 transformed count data
DESeq2	`vst()`	Variance Stabilizing Transformation	Estimates the dispersion trend of count data
DESeq2	`DESeqDataSetFromTximport()`	DESeqDataSet constructor	Creates a DESeqDataSet object
DESeq2	`DESeq()`	Differential Expression Analysis Based on the Negative Binomial Distribution	Estimates size factors, estimates dispersion, and performs negative binomial fitting and Wald statistics as steps in the default differential expression analysis
DESeq2	`plotPCA()`	PCA plot	Produces a principal component analysis plot for transformed data, while checking for batch effects

Library/Package	Piece of Code	What it's called	What it does
DESeq2	`counts()`	Counts	Saves the count data as a matrix
DESeq2	`results()`	Results	Returns the results table from a DESeq2 analysis
DESeq2	`assay()`	Assay	Returns count values in a matrix

Library/Package	Piece of Code	What it's called	What it does
fastp	`fastp`	FASTQ preprocessor	Preprocesses FASTQ files through adapter trimming, quality filtering, length filtering, and a number of additional option
FastQC	`fastqc`	FASTQC (Quality Control)	Performs quality control checks on raw sequence data and outputs a QC(quality control) report

Library/Package	Piece of Code	What it's called	What it does
ggplot2	`ggsave()`	GG Save	Saves the last plot in working directory
ggplot2	`last_plot()`	Last plot	Returns the last plot produced
ggplot2	`geom_point()`	Geom point	Creates a scatterplot (when added to the `ggplot()` function)
ggplot2	`xlab(); ylab()`	X Axis Labels; Y Axis Labels	Modifies the labels on the x axis and on the y axis, respectively
ggplot2	`coord_fixed()`	Cartesian Coordinates with Fixed Aspect Ratio	Coerces the coordinates on the plot to represent a fixed specified ratio

Library/Package	Piece of Code	What it's called	What it does
Salmon	`salmon index`	Salmon index	Builds a transcriptome index which is required for Salmon quantification
Salmon	`salmon quant`	Salmon quantification	Runs Salmon's quantification of transcript expression

Library/Package	Piece of Code	What it's called	What it does
stringr	`word()`	Word	Extracts words from a character vector
readr	`write_rds()`	Write RDS	Writes data to a .rds output file

Library/Package	Piece of Code	What it's called	What it does
tximport	`tximport()`	Tximport	Imports transcript-level estimates and summarizes abundances, counts, and transcript lengths to the gene-level by default
dplyr	`pull()`	Pull	Extracts a variable as a vector
pheatmap	`pheatmap()`	Pheatmap	Plots clustered heatmaps

Useful command line commands: