Analysis Plots

Once you have your bioinformatics files, your work is not done; you must analyze your data by visualizing it through analysis plots. There are many types of analysis plots, but it is important to choose an appropriate one for the question you are trying to answer.

Below is a list of common analysis plots and what information they provide.

Genome Browser Tracks

Genome browser tracks show the experimental signal across the genome. With an RNA-Seq experiment, this signal reflects gene expression, while with an RBP-eCLIP experiment, it shows binding sites for the protein. Tracks are created from bioinformatics files of a sequenced sample and visualized using a genome browser, such as the Integrative Genomics Viewer (IGV). The reference genome is along the x-axis and typically has gene models represented to show where exons and UTRs are located. Each row is a different sample and peaks indicate enrichment, such as where a protein binds.

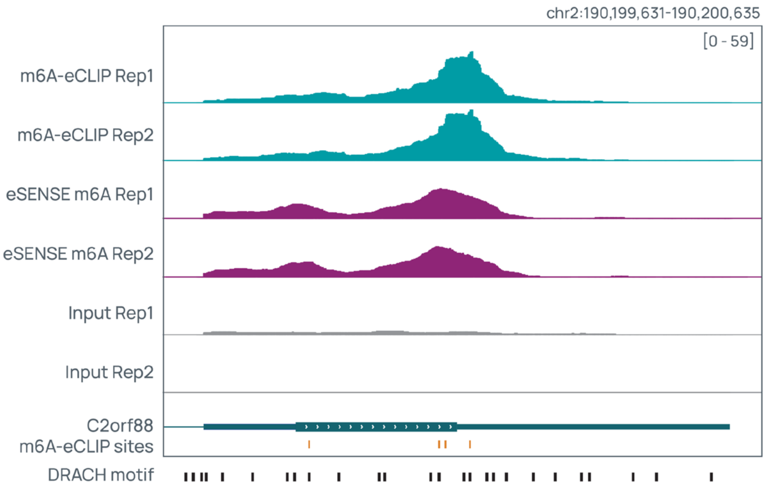

Figure 1: Example browser tracks showing regions of m6A deposits along C2of88. The specific methylated adenosines are marked in orange.

Box Plots

Box plots show the distribution of data including features such as gene expression or peak widths. Typically, samples are placed on the x-axis, and the data under examination on the y-axis. The box spans from the 1st to the 3rd quartile of the distribution. A horizontal line inside the box marks the median of the distribution, and the lines outside the box (whiskers) extend to 1.5 times the size of the box. If there are any points outside the whiskers, they are considered outliers and shown as individual dots.

By showing a direct visual comparison between sample distributions, box plots allow for an easy way to compare replicates or conditions to identify global changes.

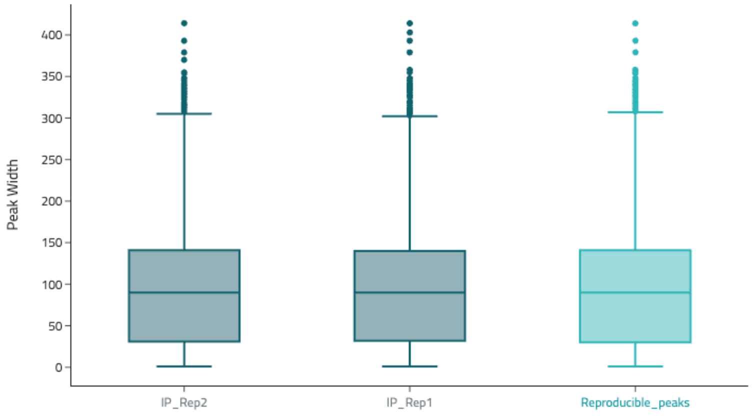

Figure 2: Example box plot showing the width of enriched peaks in an eCLIP experiment. Data for individual replicates and reproducible peaks are shown.

Volcano Plots

Volcano plots show how specific genes or binding patterns on genes change between two conditions. The ratio of how much gene expression or binding has changed (typically log-transformed fold change) is along with the x-axis, with decreasing expression on the left and increasing expression on the right. The y-axis represents the statistical significance of the expression changes. Each gene or binding site is represented by a dot. If the dot is far to the left and high up on the plot, then the gene is downregulated. If the dot is to the far right and high up, the gene is upregulated. If the dot is higher up the y-axis, then the change is more significant.

Unlike box plots, volcano plots enable the identification of specific changes between two conditions.

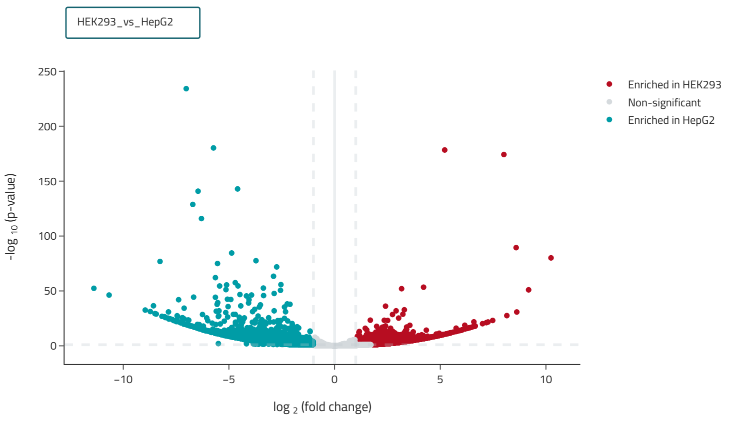

Figure 3: Example volcano plot showing miRNA binding differentials between two conditions.

Heatmap

Heatmaps use color variations to visualize changes in gene measurements, such as expression with RNA-Seq or base accessibility with eSHAPE. When visualizing changes in base accessibility, the position along the RNA is on the x-axis and samples are on the y-axis, although this can change based off the type of data visualized. Each value at a position on the plot is represented by a color along a spectrum between the minimum and maximum gene measurements. Heatmaps are useful for seeing patterns of changes, such as where there are common regions of base accessibility across conditions.

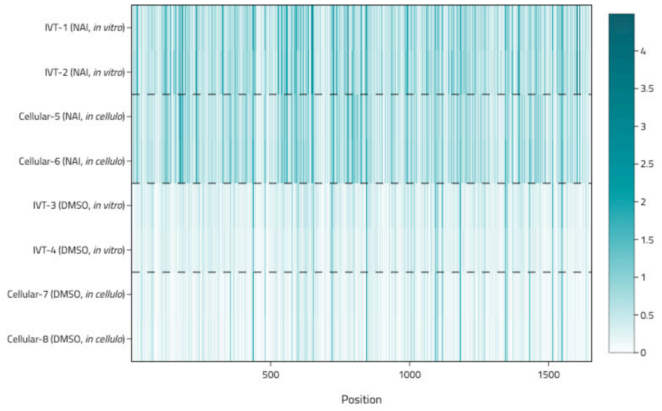

Figure 4: Example heatmap showing differences in base accessibility between different conditions.

Stacked bar chart

Stacked bar charts can compare the distribution of characteristics between different samples. Each sample is represented by a bar along the y-axis divided into different colors by the characteristic being analyzed. For example, the chart can show the distribution of gene features, such as the coding sequence or UTRs, that are bound by proteins or miRNAs.

Stacked bar charts are best used for showing how the distribution of multiple features compares between samples.

Interested in what you can discover from your bioinformatics data? Contact us to learn more.

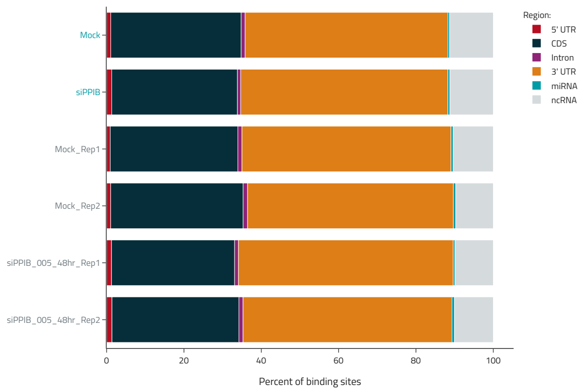

Figure 5: Example stacked bar chart showing what transcriptomic features were enriched in an eCLIP experiment.