Charlee Cobb - Transcriptomics, Exercise 12

Charlee Cobb

2023-04-10

M12_v1_description) Bioinformatician's goals M12_v1_1) What are three main services that Bioinformaticians provide?

The three main services bioinformatics can provide is Data Integration, Data Analysis, and Data Visualization.

M12_v2_description) Common plots part 1 M12_v2_1) What plot can be helpful in determining if your normalization process was successful? Explain what it shows

The best plot for determining normalization process are Box Plots. They're great at comparing data values and the change in distribution after normalization.

M12_v3_description) Common plots part 2 M12_v3_1) What is the purpose of a Venn Diagram and what is its limitation?

A Venn Diagram is great for comparing elements in a list, like overlapping differentially expressed genes. However, it's limited when you want to compare more than three lists as the graphs become more complex.

M12_v4_description) Genomic Data Visualization part 1 M12_v4_1) How can we use Genome browsers to visualize RNA-seq data?

The Genome browser can visualize RNA-seq data by showing the nucleotides in the given RNA sequence and revealing the gaps, exons, and splice sites. It also provides a dendrogram for phlogenomics, and can show you where the reads align against the reference genome.

M12_v4_2) How do you create a dot plot?

Is made with the genomic viewer data.

M12_v5_description) Genomic Data Visualization part 2 M12_v5_1) How is the height in a Sequence Logo graph calculated?

You calculate the entropy measure against the variability of each position.

M12_v5_2) How can you avoid making mistakes in interpreting dendrograms?

You need to know how your graph is rooted, and make assupmtions about the distance using the x axis.

M12_v6_description) Genomic Data Visualization part 3 M12_v6_1) What can you conclude by clustering samples instead of genes?

Clustering by samples can tell you how closely related the samples are to one another. CLustering by genes may show you expression patterns, but it won't differentiate the samples.

M12_v6_2) Why do we calculate the -log10() of the p-value when creating a volcano plot?

We calculate the -log10() of the pvalue so on the graph we see things that are more significant at the top of the graph and can clearly see genes with a significant fold change and pvalue.

M12_v6_3) How can we use heatmaps to interpret our data?

Heatmaps help us understand paterns in our data by using clustering to seperate the data. Heatmaps show more dimensions in patterns than a dendrogram would.

M12_v7_description) Principal Component Analysis M12_v7_1) What is the relationship between principal components and the variation in the data?

Principal components reduce dimiensionality so we can better see the amount of variation and the range of variation there is in the data.

M12_v8_description) Gene Networks M12_v8_1) Give three examples of edges in a gene network and the nodes that connect them.

One edge type is a metabolic reaction with nodes being the metabolite and the metabolic gene. Another edge type is Protein Protein interaction with the nodes being two proteins, and a third edge is microRNA/RNA relations where the nodes are miRNA and Target RNA.

M12_v8_2) What kind of edges can you draw using RNA-seq experimental data?

With RNA-seq experimental data, we can draw edges of regulatory interactions or biological processes.