

Functional Annotation Data Mining

657.001 Transcriptomics (WS 2018/19)



Current status of annotation

- ▶ We now have
 - Lists of DE transcripts (with IDs only)
 - Functional annotation for all transcripts
 - ▶ Still want to know
 - Which functions are more expressed in one sample than in the other
- ➔ Must combine the two information sources

Annotation data mining

- ▶ Extract annotation information only for DE transcripts
 - Annotation for condition 1
 - Annotation for condition 2
- ▶ Focus on top annotation terms per condition
 - Order by transcript (gene) count
- ▶ Compare top annotation terms between conditions
 - How are they different?
 - Do the differences make sense?

Web-based annotation: PANTHER

- ▶ One of the most widely used gene list functional annotation tools
- ▶ Annotations include
 - GO terms (mostly)
 - Also
 - PANTHER protein classes
 - PANTHER pathways



Protein ANnotation THrough Evolutionary Relationship

Web-based annotation: DAVID

- ▶ Uses several annotation sources
 - GO, pathways, and many others
- ▶ Annotation presented in several ways
 - Functional annotation chart
 - Functional annotation table
 - Annotation clusters
- ▶ Lots of information provided by hyperlinks
 - Mostly useful, sometimes redundant
 - Finding what you want is not easy!



Database for Annotation, Visualization and Integrated Discovery

Lab overview

