

Data analysis tools for RNA-Seq

Goal	Software Program	Why you need it
Acquire annotated genome data	NCBI	Annotation information is used to map reads to specific genes
Differential gene expression (DE)	Rockhopper	Imports FASTQ data, maps reads to open reading frames, normalizes data, determines statistically significant DE using q-values, manually calculate fold change (FC)
Create database containing gene names, identifiers, protein function, FC values and other important information	Filemaker Pro	Can import Excel (.csv) files with their data into one database from a variety of sources
Learn more about the function of a gene	Ecocyc	This source is well curated for <i>E. coli</i> K-12 and can link you to other databases
Pathway analysis/enrichment	Ecocyc Cellular Overview	This can allow you to paint your data over pathways to visualize which genes change in expression
Gene ontology enrichment	Ecocyc and other databases	This allows you to determine if genes with particular functions are more commonly seen in your dataset
Transcription factor enrichment	Ecocyc and other databases	This allows you to determine which transcription factors are more commonly seen in your dataset.
Transcriptional regulons	Ecocyc Regulatory Overview Regulon DB	This gives you lists of which genes are controlled by specific transcription factors