

Intro To Bioinformatics Final Project - Workflow Submission

Name	ID
Michal Mano	204865687
Saar Ofek	204811403

Disease of choice: Asthma

Biological Question: We would like to explore the genetic differences between patients with mild and severe asthma, and if there is any correlation with the age of diagnosis. We are interested in this question because Saar was diagnosed with mild asthma at a late age (20), despite most cases being diagnosed in childhood.

How can gene expression data help? We can run an analysis on Bulk RNA-Seq data to compare gene expression in samples of healthy patients and those with varying degrees of asthma severity. It is common to categorize asthma as either Mild/Moderate and Severe. We can use the genes found to identify pathways that are associated with asthma, and if any are more strongly associated with more severe forms of it.

Other analysis: We will also use GWAS to determine if the differentially expressed genes found in the previous step are significant markers for asthma. GWAS can also help us look for genes that are differentially expressed in those with childhood onset asthma vs. adult onset.

Accession ID	Location	Data Type	Count Matrix	Groups
GSE136587	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE136587	Bulk RNA-Seq	raw counts	39 samples: 17 mild/moderate, 6 healthy, 16 severe
GSE145505	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE145505	Bulk RNA-Seq	raw counts	442 samples of children with 6 respiratory phenotypes
GSE207751	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE207751	Bulk RNA-Seq	raw counts	50 samples: 36 asthma, 14 healthy
GSE158752	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE158752	Bulk RNA-seq	Raw counts	67 Healthy: 17 Mild/Moderate:25 Severe:25
GSE119579	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE119579	Bulk RNA-Seq	RPKM normalized	7 non asthmatic, 6 asthmatic