

This folder contains the necessary files and R scripts needed to carry out cellular deconvolution of whole blood gene expression profiles to generate inferred neutrophil, lymphocyte, monocyte, and eosinophil counts via the PCA-based method used in the test analyses detailed in the Technical Validation section of O'Connell et al.

pca_deconvolution_script.r contains annotated R script used for deconvolution. *WBC4_126_marker_genes.csv* contains the WBC4.126 marker gene list. *whole_blood_raw_counts.csv* is an abbreviated example input file illustrating the format whole blood RNA sequencing data should be provided in. Data should be in the form of gene-level summarized read counts. The exact RNA sequencing data used for the analyses presented in the manuscript can be retrieved from the National Center for Biotechnology Information via BioProject accession number PRJNA949611.