

Gene Set Building and CAMML Analysis of GSE72056 Melanoma Data

Courtney Schiebout, H. Robert Frost

Load Libraries

Libraries “CAMML” (Schiebout and Frost 2022) and “Seurat” (Satija et al. 2015) need to be loaded to carry out this vignette, in addition to several other libraries for data processing and gene set development (Robinson, McCarthy, and Smyth 2010; Carlson 2023; Liberzon et al. 2011). Packages will also load additional libraries they depend on.

```
library(CAMML)
library(Seurat)
library(edgeR)
library(org.Hs.eg.db)
library(msigdb)
```

Load and Process the GSE72056 Dataset

The methods used to create and classify the data used in this vignette can be found at the following DOI: 10.1126/science.aad0501 (Tirosh et al. 2016). The data can be accessed using the Gene Expression Omnibus (GEO) at accession number GSE72056 (Tirosh et al. 2016; Edgar, Domrachev, and Lash 2002). This data is not originally formatted in the matrix style required for Seurat and must be slightly altered prior to analysis. The altered data structure can then be processed and normalized using the Seurat pipeline (Satija et al. 2015).

```
## Warning: Feature names cannot have underscores ('_'), replacing with dashes
## ('-')

## Warning: Data is of class data.frame. Coercing to dgCMatrx.

## Normalizing layer: counts

## Finding variable features for layer counts

## Centering and scaling data matrix

## PC_ 1
## Positive:  CST3, SERPING1, GSN, CTSL1, FN1, PMP22, TMEM176B, CD9, TGFBI, TSPAN4
##           TIMP1, THBS1, ENG, ALDH2, GPNMB, NNMT, S100A16, IER3, IGFBP7, MMP2
##           FSTL1, FCGRT, TUBB6, TMEM176A, SEPP1, CD14, PLAUR, GLUL, PPIC, CNN3
## Negative:  CCL5, CXCR4, PTPRCAP, CD52, CD69, ARHGDIB, TSC22D3, CORO1A, TXNIP, RPL13AP5
##           CCR7, CD3D, CD37, IL7R, MS4A1, LTB, SELL, LAPTM5, CD2, BANK1
##           CD79A, RHOH, RPL13A, CD19, LY9, PIK3IP1, DGKA, RAC2, CD22, TXK

## PC_ 2
## Positive:  CD2, IL32, CD3D, NKG7, CST7, CD8A, PRF1, KLRK1, TIGIT, GZMA
##           PDCD1, GZMK, ITM2A, CD8B, CTSW, SRGN, LCK, SIRPG, CCL4, ZAP70
##           PYHIN1, CCL4L2, CCL4L1, SH2D1A, IFITM1, CTLA4, CD27, HCST, IFNG, CD247
```