Geometrical Unsupervised Learning of Breast Cancer SubTyping

2024-05-09

Abstract

Molecular subtyping of cancer is recognized as a critical and challenging step towards individua lized therapy. Breast cancer molecular subtyping has revolutionized the treatment and lowered th e mortality rate.

In the study, we further analyse breast cancer's subtyping using geometrical unsupervised learni ng. We aim to find a new way to examine the gene expression data by applying on it a different t opology. The topology is applied by using different distance function on the sample space.

In the study we used 5 metrical functions - Pairwise Correlation, Cosine Similarity, Kernel Dist ance, PCA and Euclidean distance for the control analyses.
Pairwise correlation measures the linear relationship between two variables.
Cosine similarity measures the cosine of the angle between two non-zero vectors.
Kernel distance measures nonlinear relationship using a kernel function.
Principal Component Analysis (PCA) is a dimensionality reduction technique, preserving as much v ariance as possible. After the reduction of the samples to one dimensional variable we measured the distance between the samples using the Euclidean distance.

The research contains RNA sequence samples taken from TCGA dataset. The data contains 313 sample s. The samples are divided by PAM50 classification to 35 “Normal” samples, 68 “Her2“ samples and 70 “Luminal A”, “Luminal B” and “Basal” samples.

The study pipeline is Desq2 analyses, heatmap and unsupervised hierarchical clustering, survival analyses and pathways enrichments analyses.
Each of the 5 metrics processed through the pipeline. The Euclidean distance, Pairwise Correlati on, Cosine Similarity had similar results. Those results are consistent with the PAM50 common su btyping. The kernel distance did not form clusters. The PCA had different clustering results.

The PCA resulted 3 clusters. Those clusters received the lowest p value in the survival analyse s, yet nonsignificant.
The pathway enrichment analyses didn’t contribute difference to any of the clustering of all met rics.

A deeper understanding of the PCA clustering can be done to understand if the results are signif icant.

```
setwd('C:/technion/Bio/project/')
library(TCGAbiolinks) # download the TCGA data

library(DESeq2)
```

```
## Warning: package 'DESeq2' was built under R version 4.3.3
```

```
## Loading required package: S4Vectors
```

```
## Loading required package: stats4
```

```
## Loading required package: BiocGenerics
```

```
##
## Attaching package: 'BiocGenerics'
```

```
## The following objects are masked from 'package:stats':
##
##     IQR, mad, sd, var, xtabs
```

```
## The following objects are masked from 'package:base':
##
##     anyDuplicated, aperm, append, as.data.frame, basename, cbind,
##     colnames, dirname, do.call, duplicated, eval, evalq, Filter, Find,
##     get, grep, grepl, intersect, is.unsorted, lapply, Map, mapply,
##     match, mget, order, paste, pmax, pmax.int, pmin, pmin.int,
##     Position, rank, rbind, Reduce, rownames, sapply, setdiff, sort,
##     table, tapply, union, unique, unsplit, which.max, which.min
```

```
##
## Attaching package: 'S4Vectors'
```

```
## The following object is masked from 'package:utils':
##
##     findMatches
```

```
## The following objects are masked from 'package:base':
##
##     expand.grid, I, unname
```

```
## Loading required package: IRanges
```

```
##
## Attaching package: 'IRanges'
```

```
## The following object is masked from 'package:grDevices':
##
##     windows
```

```
## Loading required package: GenomicRanges
```

```
## Loading required package: GenomeInfoDb
```

```
## Loading required package: SummarizedExperiment
```

```
## Loading required package: MatrixGenerics
```

```
## Loading required package: matrixStats
```

```
##
## Attaching package: 'MatrixGenerics'
```

```
## The following objects are masked from 'package:matrixStats':
##
##     colAlls, colAnyNAs, colAnys, colAvgsPerRowSet, colCollapse,
##     colCounts, colCummaxs, colCummins, colCumprods, colCumsums,
##     colDiffs, colIQRDiffs, colIQRs, colLogSumExps, colMadDiffs,
##     colMads, colMaxs, colMeans2, colMedians, colMins, colOrderStats,
##     colProds, colQuantiles, colRanges, colRanks, colSdDiffs, colSds,
##     colSums2, colTabulates, colVarDiffs, colVars, colWeightedMads,
##     colWeightedMeans, colWeightedMedians, colWeightedSds,
##     colWeightedVars, rowAlls, rowAnyNAs, rowAnys, rowAvgsPerColSet,
##     rowCollapse, rowCounts, rowCummaxs, rowCummins, rowCumprods,
##     rowCumsums, rowDiffs, rowIQRDiffs, rowIQRs, rowLogSumExps,
##     rowMadDiffs, rowMads, rowMaxs, rowMeans2, rowMedians, rowMins,
##     rowOrderStats, rowProds, rowQuantiles, rowRanges, rowRanks,
##     rowSdDiffs, rowSds, rowSums2, rowTabulates, rowVarDiffs, rowVars,
##     rowWeightedMads, rowWeightedMeans, rowWeightedMedians,
##     rowWeightedSds, rowWeightedVars
```

```
## Loading required package: Biobase
```

```
## Welcome to Bioconductor
##
##     Vignettes contain introductory material; view with
##     'browseVignettes()'. To cite Bioconductor, see
##     'citation("Biobase")', and for packages 'citation("pkgname")'.
```

```
##
## Attaching package: 'Biobase'
```

```
## The following object is masked from 'package:MatrixGenerics':
##
##     rowMedians
```

```
## The following objects are masked from 'package:matrixStats':
##
##     anyMissing, rowMedians
```

```
library(SummarizedExperiment)

library(ComplexHeatmap) # Heatmap (works with personalize distance function)
```

```
## Loading required package: grid
```

```
## ========================================
## ComplexHeatmap version 2.18.0
## Bioconductor page: http://bioconductor.org/packages/ComplexHeatmap/
## Github page: https://github.com/jokergoo/ComplexHeatmap
## Documentation: http://jokergoo.github.io/ComplexHeatmap-reference
##
## If you use it in published research, please cite either one:
## - Gu, Z. Complex Heatmap Visualization. iMeta 2022.
## - Gu, Z. Complex heatmaps reveal patterns and correlations in multidimensional
##     genomic data. Bioinformatics 2016.
##
##
## The new InteractiveComplexHeatmap package can directly export static
## complex heatmaps into an interactive Shiny app with zero effort. Have a try!
##
## This message can be suppressed by:
##   suppressPackageStartupMessages(library(ComplexHeatmap))
## ========================================
```

```
library(dendsort) # heatmap reordering columns
```

```
## Warning: package 'dendsort' was built under R version 4.3.3
```

```
library(factoextra) # elbow method
```

```
## Loading required package: ggplot2
```

```
## Warning: package 'ggplot2' was built under R version 4.3.3
```

```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

```
library(survival) # survival analyses
library(survminer) # survival analyses
```

```
## Loading required package: ggpubr
```

```
##
## Attaching package: 'survminer'
```

```
## The following object is masked from 'package:survival':
##
##     myeloma
```

```
library(msigdbr) # enrichment analysis
library(clusterProfiler) # enrichment analysis
```

```
## Warning: package 'clusterProfiler' was built under R version 4.3.3
```

```
##
```

```
## clusterProfiler v4.10.1  For help: https://yulab-smu.top/biomedical-knowledge-mining-book/
##
## If you use clusterProfiler in published research, please cite:
## T Wu, E Hu, S Xu, M Chen, P Guo, Z Dai, T Feng, L Zhou, W Tang, L Zhan, X Fu, S Liu, X Bo, an
d G Yu. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. The Innova
tion. 2021, 2(3):100141
```

```
##
## Attaching package: 'clusterProfiler'
```

```
## The following object is masked from 'package:IRanges':
##
##     slice
```

```
## The following object is masked from 'package:S4Vectors':
##
##     rename
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:Biobase':
##
##     combine
```

```
## The following object is masked from 'package:matrixStats':
##
##     count
```

```
## The following objects are masked from 'package:GenomicRanges':
##
##     intersect, setdiff, union
```

```
## The following object is masked from 'package:GenomeInfoDb':
##
##     intersect
```

```
## The following objects are masked from 'package:IRanges':
##
##     collapse, desc, intersect, setdiff, slice, union
```

```
## The following objects are masked from 'package:S4Vectors':
##
##     first, intersect, rename, setdiff, setequal, union
```

```
## The following objects are masked from 'package:BiocGenerics':
##
##     combine, intersect, setdiff, union
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(matrixStats)
library(circlize) # Heatmaps color
```

```
## Warning: package 'circlize' was built under R version 4.3.3
```

```
## ========================================
## circlize version 0.4.16
## CRAN page: https://cran.r-project.org/package=circlize
## Github page: https://github.com/jokergoo/circlize
## Documentation: https://jokergoo.github.io/circlize_book/book/
##
## If you use it in published research, please cite:
## Gu, Z. circlize implements and enhances circular visualization
##   in R. Bioinformatics 2014.
##
## This message can be suppressed by:
##   suppressPackageStartupMessages(library(circlize))
## ========================================
```

```
library(RColorBrewer)
library(scales)
```

```
## Warning: package 'scales' was built under R version 4.3.3
```

Download TCGA Data of BRCA RNA-seq

```
query_TCGA = GDCquery(
  project = "TCGA-BRCA",
  data.category = "Transcriptome Profiling",
  experimental.strategy = "RNA-Seq",
  workflow.type = "STAR - Counts", #raw counts
  data.type = 'Gene Expression Quantification',
  sample.type = c("Primary Tumor", "Solid Tissue Normal"))#removed metastatic tumor
```

```
## -------------------------------------
```

```
## o GDCquery: Searching in GDC database
```

```
## -------------------------------------
```

```
## Genome of reference: hg38
```

```
## -----------------------------------------
```

```
## oo Accessing GDC. This might take a while...
```

```
## -----------------------------------------
```

```
## ooo Project: TCGA-BRCA
```

```
## --------------------
```

```
## oo Filtering results
```

```
## --------------------
```

```
## ooo By experimental.strategy
```

```
## ooo By data.type
```

```
## ooo By workflow.type
```

```
## ooo By sample.type
```

```
## ----------------
```

```
## oo Checking data
```

```
## ----------------
```

```
## ooo Checking if there are duplicated cases
```

```
## ooo Checking if there are results for the query
```

```
## -------------------
```

```
## o Preparing output
```

```
## -------------------
```

```
GDCdownload(query = query_TCGA)
```

```
## Downloading data for project TCGA-BRCA
```

```
## Of the 1224 files for download 1224 already exist.
```

```
## All samples have been already downloaded
```

```
data = GDCprepare(query_TCGA)
```

```
## |                                                                      |  0%                                                                  |
|0.08169935% ~3 m remaining|                                                                  |0.1633987% ~2 m
remaining |                                                                       |0.245098% ~2 m remaining  |
|0.3267974% ~2 m remaining |                                                                  |0.4084967% ~2 m
remaining |                                                                       |0.4901961% ~2 m remaining |
|0.5718954% ~2 m remaining |                                                                  |0.6535948% ~2 m
remaining |                                                                       |0.7352941% ~2 m remaining |
|0.8169935% ~3 m remaining |                                                                  |0.8986928% ~2 m
remaining |                                                                       |0.9803922% ~2 m remaining |
|1.062092% ~2 m remaining  |                                                                  |1.143791% ~2 m
remaining  |                                                                      |1.22549% ~2 m remaining   |
|1.30719% ~2 m remaining   |                                                                  |1.388889% ~2 m
remaining  |                                                                      |1.470588% ~2 m remaining  |
|1.552288% ~2 m remaining  |                                                                  |1.633987% ~2 m
remaining  |                                                                      |1.715686% ~2 m remaining  |
|1.797386% ~2 m remaining  |                                                                  |1.879085% ~2 m
remaining  |=                                                                     |1.960784% ~2 m remaining  |=
|2.042484% ~2 m remaining  |=                                                                 |2.124183% ~2 m
remaining  |=                                                                     |2.205882% ~2 m remaining  |=
|2.287582% ~2 m remaining  |=                                                                 |2.369281% ~2 m
remaining  |=                                                                     |2.45098% ~2 m remaining   |=
|2.53268% ~2 m remaining   |=                                                                 |2.614379% ~2 m
remaining  |=                                                                     |2.696078% ~2 m remaining  |=
|2.777778% ~2 m remaining  |=                                                                 |2.859477% ~2 m
remaining  |=                                                                     |2.941176% ~2 m remaining  |=
|3.022876% ~2 m remaining  |=                                                                 |3.104575% ~2 m
remaining  |=                                                                     |3.186275% ~2 m remaining  |=
|3.267974% ~2 m remaining  |=                                                                 |3.349673% ~2 m
remaining  |=                                                                     |3.431373% ~2 m remaining  |=
|3.513072% ~2 m remaining  |=                                                                 |3.594771% ~2 m
remaining  |=                                                                     |3.676471% ~2 m remaining  |=
|3.75817% ~2 m remaining   |=                                                                 |3.839869% ~2 m
remaining  |==                                                                    |3.921569% ~2 m remaining  |==
|4.003268% ~2 m remaining  |==                                                                |4.084967% ~2 m
remaining  |==                                                                    |4.166667% ~2 m remaining  |==
|4.248366% ~2 m remaining  |==                                                                |4.330065% ~2 m
remaining  |==                                                                    |4.411765% ~2 m remaining  |==
|4.493464% ~2 m remaining  |==                                                                |4.575163% ~2 m
remaining  |==                                                                    |4.656863% ~2 m remaining  |==
|4.738562% ~2 m remaining  |==                                                                |4.820261% ~2 m
remaining  |==                                                                    |4.901961% ~2 m remaining  |==
|4.98366% ~2 m remaining   |==                                                                |5.065359% ~2 m
remaining  |==                                                                    |5.147059% ~2 m remaining  |==
|5.228758% ~2 m remaining  |==                                                                |5.310458% ~2 m
remaining  |==                                                                    |5.392157% ~2 m remaining  |==
|5.473856% ~2 m remaining  |==                                                                |5.555556% ~2 m
remaining  |==                                                                    |5.637255% ~2 m remaining  |==
|5.718954% ~2 m remaining  |===                                                               |5.800654% ~2 m
remaining  |===                                                                   |5.882353% ~2 m remaining  |===
|5.964052% ~2 m remaining  |===                                                               |6.045752% ~2 m
remaining  |===                                                                   |6.127451% ~2 m remaining  |===
|6.20915% ~2 m remaining   |===                                                               |6.29085% ~2 m r
```

```
emaining   |===                                                
```

```
|12.3366% ~2 m remaining   |======                                  |12.4183% ~2 m r
emaining   |======                                  |12.5% ~2 m remaining      |====
==                                  |12.5817% ~2 m remaining   |======
|12.6634% ~2 m remaining   |======                                  |12.7451% ~2 m r
emaining   |======                                  |12.8268% ~2 m remaining   |====
==                                  |12.9085% ~2 m remaining   |======
|12.9902% ~2 m remaining   |======                                  |13.0719% ~2 m r
emaining   |======                                  |13.15359% ~2 m remaining  |====
==                                  |13.23529% ~2 m remaining  |======
|13.31699% ~2 m remaining  |======                                  |13.39869% ~2 m
remaining  |=======                                 |13.48039% ~2 m remaining  |====
===                                 |13.56209% ~2 m remaining  |=======
|13.64379% ~2 m remaining  |=======                                 |13.72549% ~2 m
remaining  |=======                                 |13.80719% ~2 m remaining  |====
===                                 |13.88889% ~2 m remaining  |=======
|13.97059% ~2 m remaining  |=======                                 |14.05229% ~2 m
remaining  |=======                                 |14.13399% ~2 m remaining  |====
===                                 |14.21569% ~2 m remaining  |=======
|14.29739% ~2 m remaining  |=======                                 |14.37908% ~2 m
remaining  |=======                                 |14.46078% ~2 m remaining  |====
===                                 |14.54248% ~2 m remaining  |=======
|14.62418% ~2 m remaining  |=======                                 |14.70588% ~2 m
remaining  |=======                                 |14.78758% ~2 m remaining  |====
===                                 |14.86928% ~2 m remaining  |=======
|14.95098% ~2 m remaining  |=======                                 |15.03268% ~2 m
remaining  |=======                                 |15.11438% ~2 m remaining  |====
===                                 |15.19608% ~2 m remaining  |=======
|15.27778% ~2 m remaining  |=======                                 |15.35948% ~2 m
remaining  |========                                |15.44118% ~2 m remaining  |====
====                                |15.52288% ~2 m remaining  |========
|15.60458% ~2 m remaining  |========                                |15.68627% ~2 m
remaining  |========                                |15.76797% ~2 m remaining  |====
====                                |15.84967% ~2 m remaining  |========
|15.93137% ~2 m remaining  |========                                |16.01307% ~2 m
remaining  |========                                |16.09477% ~2 m remaining  |====
====                                |16.17647% ~2 m remaining  |========
|16.25817% ~2 m remaining  |========                                |16.33987% ~2 m
remaining  |========                                |16.42157% ~2 m remaining  |====
====                                |16.50327% ~2 m remaining  |========
|16.58497% ~2 m remaining  |========                                |16.66667% ~2 m
remaining  |========                                |16.74837% ~2 m remaining  |====
====                                |16.83007% ~2 m remaining  |========
|16.91176% ~2 m remaining  |========                                |16.99346% ~2 m
remaining  |========                                |17.07516% ~2 m remaining  |====
====                                |17.15686% ~2 m remaining  |========
|17.23856% ~2 m remaining  |=========                               |17.32026% ~2 m
remaining  |=========                               |17.40196% ~2 m remaining  |====
=====                               |17.48366% ~2 m remaining  |=========
|17.56536% ~2 m remaining  |=========                               |17.64706% ~2 m
remaining  |=========                               |17.72876% ~2 m remaining  |====
=====                               |17.81046% ~2 m remaining  |=========
|17.89216% ~2 m remaining  |=========                               |17.97386% ~2 m
```

```
remaining  |=========                                        |18.05556% ~2 m remaining  |====
=====                                          |18.13725% ~2 m remaining  |=========
|18.21895% ~2 m remaining  |=========                                        |18.30065% ~2 m
remaining  |=========                                        |18.38235% ~2 m remaining  |====
=====                                          |18.46405% ~2 m remaining  |=========
|18.54575% ~2 m remaining  |=========                                        |18.62745% ~2 m
remaining  |=========                                        |18.70915% ~2 m remaining  |====
=====                                          |18.79085% ~2 m remaining  |=========
|18.87255% ~2 m remaining  |=========                                        |18.95425% ~2 m
remaining  |=========                                        |19.03595% ~2 m remaining  |====
=====                                          |19.11765% ~2 m remaining  |=========
|19.19935% ~2 m remaining  |==========                                       |19.28105% ~2 m
remaining  |==========                                       |19.36275% ~2 m remaining  |====
======                                         |19.44444% ~2 m remaining  |==========
|19.52614% ~2 m remaining  |==========                                       |19.60784% ~2 m
remaining  |==========                                       |19.68954% ~2 m remaining  |====
======                                         |19.77124% ~2 m remaining  |==========
|19.85294% ~2 m remaining  |==========                                       |19.93464% ~2 m
remaining  |==========                                       |20.01634% ~2 m remaining  |====
======                                         |20.09804% ~2 m remaining  |==========
|20.17974% ~2 m remaining  |==========                                       |20.26144% ~2 m
remaining  |==========                                       |20.34314% ~2 m remaining  |====
======                                         |20.42484% ~2 m remaining  |==========
|20.50654% ~2 m remaining  |==========                                       |20.58824% ~2 m
remaining  |==========                                       |20.66993% ~2 m remaining  |====
======                                         |20.75163% ~2 m remaining  |==========
|20.83333% ~2 m remaining  |==========                                       |20.91503% ~2 m
remaining  |==========                                       |20.99673% ~2 m remaining  |====
======                                         |21.07843% ~2 m remaining  |===========
|21.16013% ~2 m remaining  |===========                                      |21.24183% ~2 m
remaining  |===========                                      |21.32353% ~2 m remaining  |====
=======                                        |21.40523% ~2 m remaining  |===========
|21.48693% ~2 m remaining  |===========                                      |21.56863% ~2 m
remaining  |===========                                      |21.65033% ~2 m remaining  |====
=======                                        |21.73203% ~2 m remaining  |===========
|21.81373% ~2 m remaining  |===========                                      |21.89542% ~2 m
remaining  |===========                                      |21.97712% ~2 m remaining  |====
=======                                        |22.05882% ~2 m remaining  |===========
|22.14052% ~1 m remaining  |===========                                      |22.22222% ~1 m
remaining  |===========                                      |22.30392% ~1 m remaining  |====
=======                                        |22.38562% ~1 m remaining  |===========
|22.46732% ~1 m remaining  |===========                                      |22.54902% ~1 m
remaining  |===========                                      |22.63072% ~1 m remaining  |====
=======                                        |22.71242% ~1 m remaining  |===========
|22.79412% ~1 m remaining  |===========                                      |22.87582% ~1 m
remaining  |===========                                      |22.95752% ~1 m remaining  |====
=======                                        |23.03922% ~1 m remaining  |============
|23.12092% ~1 m remaining  |============                                     |23.20261% ~1 m
remaining  |============                                     |23.28431% ~1 m remaining  |====
========                                       |23.36601% ~1 m remaining  |============
|23.44771% ~1 m remaining  |============                                     |23.52941% ~1 m
remaining  |============                                     |23.61111% ~1 m remaining  |====
```

```
========                                  |23.69281% ~1 m remaining  |============
|23.77451% ~1 m remaining  |============                                  |23.85621% ~1 m
remaining  |============                                |23.93791% ~1 m remaining  |====
========                                  |24.01961% ~1 m remaining  |============
|24.10131% ~1 m remaining  |============                                  |24.18301% ~1 m
remaining  |============                                |24.26471% ~1 m remaining  |====
========                                  |24.34641% ~1 m remaining  |============
|24.4281% ~1 m remaining   |============                                  |24.5098% ~1 m r
emaining   |============                                |24.5915% ~1 m remaining   |====
========                                  |24.6732% ~1 m remaining   |============
|24.7549% ~1 m remaining   |============                                  |24.8366% ~1 m r
emaining   |============                                |24.9183% ~1 m remaining   |====
=========                                 | 25% ~1 m remaining       |=============
|25.0817% ~1 m remaining   |=============                                 |25.1634% ~1 m r
emaining   |=============                               |25.2451% ~1 m remaining   |====
=========                                 |25.3268% ~1 m remaining   |=============
|25.4085% ~1 m remaining   |=============                                 |25.4902% ~1 m r
emaining   |=============                               |25.5719% ~1 m remaining   |====
=========                                 |25.65359% ~1 m remaining  |=============
|25.73529% ~1 m remaining  |=============                                 |25.81699% ~1 m
remaining  |=============                               |25.89869% ~1 m remaining  |====
=========                                 |25.98039% ~1 m remaining  |=============
|26.06209% ~1 m remaining  |=============                                 |26.14379% ~1 m
remaining  |=============                               |26.22549% ~1 m remaining  |====
=========                                 |26.30719% ~1 m remaining  |=============
|26.38889% ~1 m remaining  |=============                                 |26.47059% ~1 m
remaining  |=============                               |26.55229% ~1 m remaining  |====
=========                                 |26.63399% ~1 m remaining  |=============
|26.71569% ~1 m remaining  |=============                                 |26.79739% ~1 m
remaining  |=============                               |26.87908% ~1 m remaining  |====
==========                                |26.96078% ~1 m remaining  |==============
|27.04248% ~1 m remaining  |==============                                |27.12418% ~1 m
remaining  |==============                              |27.20588% ~1 m remaining  |====
==========                                |27.28758% ~1 m remaining  |==============
|27.36928% ~1 m remaining  |==============                                |27.45098% ~1 m
remaining  |==============                              |27.53268% ~1 m remaining  |====
==========                                |27.61438% ~1 m remaining  |==============
|27.69608% ~1 m remaining  |==============                                |27.77778% ~1 m
remaining  |==============                              |27.85948% ~1 m remaining  |====
==========                                |27.94118% ~1 m remaining  |==============
|28.02288% ~1 m remaining  |==============                                |28.10458% ~1 m
remaining  |==============                              |28.18627% ~1 m remaining  |====
==========                                |28.26797% ~1 m remaining  |==============
|28.34967% ~1 m remaining  |==============                                |28.43137% ~1 m
remaining  |==============                              |28.51307% ~1 m remaining  |====
==========                                |28.59477% ~1 m remaining  |==============
|28.67647% ~1 m remaining  |==============                                |28.75817% ~1 m
remaining  |==============                              |28.83987% ~1 m remaining  |====
===========                               |28.92157% ~1 m remaining  |===============
|29.00327% ~1 m remaining  |===============                               |29.08497% ~1 m
remaining  |===============                             |29.16667% ~1 m remaining  |====
===========                               |29.24837% ~1 m remaining  |===============
```

```
|29.33007% ~1 m remaining  |==============                                  |29.41176% ~1 m
remaining  |==============                                  |29.49346% ~1 m remaining  |====
==========                                  |29.57516% ~1 m remaining  |==============
|29.65686% ~1 m remaining  |==============                                  |29.73856% ~1 m
remaining  |==============                                  |29.82026% ~1 m remaining  |====
==========                                  |29.90196% ~1 m remaining  |==============
|29.98366% ~1 m remaining  |==============                                  |30.06536% ~1 m
remaining  |==============                                  |30.14706% ~1 m remaining  |====
==========                                  |30.22876% ~1 m remaining  |==============
|30.31046% ~1 m remaining  |==============                                  |30.39216% ~1 m
remaining  |==============                                  |30.47386% ~1 m remaining  |====
==========                                  |30.55556% ~1 m remaining  |==============
|30.63725% ~1 m remaining  |==============                                  |30.71895% ~1 m
remaining  |===============                                 |30.80065% ~1 m remaining  |====
===========                                 |30.88235% ~1 m remaining  |===============
|30.96405% ~1 m remaining  |===============                                 |31.04575% ~1 m
remaining  |===============                                 |31.12745% ~1 m remaining  |====
===========                                 |31.20915% ~1 m remaining  |===============
|31.29085% ~1 m remaining  |===============                                 |31.37255% ~1 m
remaining  |===============                                 |31.45425% ~1 m remaining  |====
===========                                 |31.53595% ~1 m remaining  |===============
|31.61765% ~1 m remaining  |===============                                 |31.69935% ~1 m
remaining  |===============                                 |31.78105% ~1 m remaining  |====
===========                                 |31.86275% ~1 m remaining  |===============
|31.94444% ~1 m remaining  |===============                                 |32.02614% ~1 m
remaining  |===============                                 |32.10784% ~1 m remaining  |====
===========                                 |32.18954% ~1 m remaining  |===============
|32.27124% ~1 m remaining  |===============                                 |32.35294% ~1 m
remaining  |===============                                 |32.43464% ~1 m remaining  |====
===========                                 |32.51634% ~1 m remaining  |===============
|32.59804% ~1 m remaining  |===============                                 |32.67974% ~1 m
remaining  |================                                |32.76144% ~1 m remaining  |====
============                                |32.84314% ~1 m remaining  |================
|32.92484% ~1 m remaining  |================                                |33.00654% ~1 m
remaining  |================                                |33.08824% ~1 m remaining  |====
============                                |33.16993% ~1 m remaining  |================
|33.25163% ~1 m remaining  |================                                |33.33333% ~1 m
remaining  |================                                |33.41503% ~1 m remaining  |====
============                                |33.49673% ~1 m remaining  |================
|33.57843% ~1 m remaining  |================                                |33.66013% ~1 m
remaining  |================                                |33.74183% ~1 m remaining  |====
============                                |33.82353% ~1 m remaining  |================
|33.90523% ~1 m remaining  |================                                |33.98693% ~1 m
remaining  |================                                |34.06863% ~1 m remaining  |====
============                                |34.15033% ~1 m remaining  |================
|34.23203% ~1 m remaining  |================                                |34.31373% ~1 m
remaining  |================                                |34.39542% ~1 m remaining  |====
============                                |34.47712% ~1 m remaining  |================
|34.55882% ~1 m remaining  |=================                               |34.64052% ~1 m
remaining  |=================                               |34.72222% ~1 m remaining  |====
=============                               |34.80392% ~1 m remaining  |=================
|34.88562% ~1 m remaining  |=================                               |34.96732% ~1 m
```

```
remaining  |==================                                    |35.04902% ~1 m remaining  |====
==============                                    |35.13072% ~1 m remaining  |==================
|35.21242% ~1 m remaining  |==================                                    |35.29412% ~1 m
remaining  |==================                                    |35.37582% ~1 m remaining  |====
==============                                    |35.45752% ~1 m remaining  |==================
|35.53922% ~1 m remaining  |==================                                    |35.62092% ~1 m
remaining  |==================                                    |35.70261% ~1 m remaining  |====
==============                                    |35.78431% ~1 m remaining  |==================
|35.86601% ~1 m remaining  |==================                                    |35.94771% ~1 m
remaining  |==================                                    |36.02941% ~1 m remaining  |====
==============                                    |36.11111% ~1 m remaining  |==================
|36.19281% ~1 m remaining  |==================                                    |36.27451% ~1 m
remaining  |==================                                    |36.35621% ~1 m remaining  |====
==============                                    |36.43791% ~1 m remaining  |==================
|36.51961% ~1 m remaining  |==================                                    |36.60131% ~1 m
remaining  |==================                                    |36.68301% ~1 m remaining  |====
==============                                    |36.76471% ~1 m remaining  |==================
|36.84641% ~1 m remaining  |==================                                    |36.9281% ~1 m r
emaining   |==================                                    |37.0098% ~1 m remaining   |====
==============                                    |37.0915% ~1 m remaining   |==================
|37.1732% ~1 m remaining   |==================                                    |37.2549% ~1 m r
emaining   |==================                                    |37.3366% ~1 m remaining   |====
==============                                    |37.4183% ~1 m remaining   |==================
|37.5% ~1 m remaining      |==================                                    |37.5817% ~1 m r
emaining   |==================                                    |37.6634% ~1 m remaining   |====
==============                                    |37.7451% ~1 m remaining   |==================
|37.8268% ~1 m remaining   |==================                                    |37.9085% ~1 m r
emaining   |==================                                    |37.9902% ~1 m remaining   |====
==============                                    |38.0719% ~1 m remaining   |==================
|38.15359% ~1 m remaining  |==================                                    |38.23529% ~1 m
remaining  |==================                                    |38.31699% ~1 m remaining  |====
==============                                    |38.39869% ~1 m remaining  |==================
|38.48039% ~1 m remaining  |==================                                    |38.56209% ~1 m
remaining  |==================                                    |38.64379% ~1 m remaining  |====
================                                  |38.72549% ~1 m remaining  |==================
|38.80719% ~1 m remaining  |==================                                    |38.88889% ~1 m
remaining  |==================                                    |38.97059% ~1 m remaining  |====
================                                  |39.05229% ~1 m remaining  |==================
|39.13399% ~1 m remaining  |==================                                    |39.21569% ~1 m
remaining  |==================                                    |39.29739% ~1 m remaining  |====
================                                  |39.37908% ~1 m remaining  |==================
|39.46078% ~1 m remaining  |==================                                    |39.54248% ~1 m
remaining  |==================                                    |39.62418% ~1 m remaining  |====
================                                  |39.70588% ~1 m remaining  |==================
|39.78758% ~1 m remaining  |==================                                    |39.86928% ~1 m
remaining  |==================                                    |39.95098% ~1 m remaining  |====
================                                  |40.03268% ~1 m remaining  |==================
|40.11438% ~1 m remaining  |==================                                    |40.19608% ~1 m
remaining  |==================                                    |40.27778% ~1 m remaining  |====
================                                  |40.35948% ~1 m remaining  |==================
=                               |40.44118% ~1 m remaining  |==================
|40.52288% ~1 m remaining  |==================                                    |40.60458% ~1 m
```

```
remaining  |====================                                |40.68627% ~1 m remaining  |====
==================                                |40.76797% ~1 m remaining  |====================
=                                |40.84967% ~1 m remaining  |====================
|40.93137% ~1 m remaining  |====================                                |41.01307% ~1 m
remaining  |====================                                |41.09477% ~1 m remaining  |====
==================                                |41.17647% ~1 m remaining  |====================
=                                |41.25817% ~1 m remaining  |====================
|41.33987% ~1 m remaining  |====================                                |41.42157% ~1 m
remaining  |====================                                |41.50327% ~1 m remaining  |====
==================                                |41.58497% ~1 m remaining  |====================
=                                |41.66667% ~1 m remaining  |====================
|41.74837% ~1 m remaining  |====================                                |41.83007% ~1 m
remaining  |====================                                |41.91176% ~1 m remaining  |====
==================                                |41.99346% ~1 m remaining  |====================
=                                |42.07516% ~1 m remaining  |====================
|42.15686% ~1 m remaining  |====================                                |42.23856% ~1 m
remaining  |====================                                |42.32026% ~1 m remaining  |====
==================                                |42.40196% ~1 m remaining  |====================
==                                |42.48366% ~1 m remaining  |====================
|42.56536% ~1 m remaining  |====================                                |42.64706% ~1 m
remaining  |====================                                |42.72876% ~1 m remaining  |====
==================                                |42.81046% ~1 m remaining  |====================
==                                |42.89216% ~1 m remaining  |====================
|42.97386% ~1 m remaining  |====================                                |43.05556% ~1 m
remaining  |====================                                |43.13725% ~1 m remaining  |====
==================                                |43.21895% ~1 m remaining  |====================
==                                |43.30065% ~1 m remaining  |====================
|43.38235% ~1 m remaining  |====================                                |43.46405% ~1 m
remaining  |====================                                |43.54575% ~1 m remaining  |====
==================                                |43.62745% ~1 m remaining  |====================
==                                |43.70915% ~1 m remaining  |====================
|43.79085% ~1 m remaining  |====================                                |43.87255% ~60 s
remaining |====================                                |43.95425% ~60 s remaining |=====
==================                                |44.03595% ~60 s remaining |====================
=                                |44.11765% ~60 s remaining |====================
|44.19935% ~60 s remaining |====================                                |44.28105% ~60 s
remaining |====================                                |44.36275% ~59 s remaining |=====
==================                                |44.44444% ~59 s remaining |====================
==                                |44.52614% ~59 s remaining |====================
|44.60784% ~59 s remaining |====================                                |44.68954% ~59 s
remaining |====================                                |44.77124% ~59 s remaining |=====
==================                                |44.85294% ~59 s remaining |====================
==                                |44.93464% ~59 s remaining |====================
|45.01634% ~60 s remaining |====================                                |45.09804% ~60 s
remaining |====================                                |45.17974% ~59 s remaining |=====
==================                                |45.26144% ~59 s remaining |====================
==                                |45.34314% ~59 s remaining |====================
|45.42484% ~59 s remaining |====================                                |45.50654% ~59 s
remaining |====================                                |45.58824% ~59 s remaining |=====
==================                                |45.66993% ~59 s remaining |====================
==                                |45.75163% ~59 s remaining |====================
|45.83333% ~59 s remaining |====================                                |45.91503% ~58 s
```

```
remaining |======================                              |45.99673% ~58 s remaining |=====
==================                                 |46.07843% ~58 s remaining |==================
===                              |46.16013% ~58 s remaining |======================
|46.24183% ~58 s remaining |======================                              |46.32353% ~58 s
remaining |======================                              |46.40523% ~58 s remaining |=====
==================                                 |46.48693% ~58 s remaining |==================
===                              |46.56863% ~58 s remaining |======================
|46.65033% ~58 s remaining |======================                              |46.73203% ~57 s
remaining |======================                              |46.81373% ~57 s remaining |=====
==================                                 |46.89542% ~57 s remaining |==================
===                              |46.97712% ~57 s remaining |======================
|47.05882% ~57 s remaining |======================                              |47.14052% ~57 s
remaining |======================                              |47.22222% ~57 s remaining |=====
==================                                 |47.30392% ~57 s remaining |==================
===                              |47.38562% ~56 s remaining |======================
|47.46732% ~56 s remaining |======================                              |47.54902% ~56 s
remaining |======================                              |47.63072% ~56 s remaining |=====
==================                                 |47.71242% ~56 s remaining |==================
===                              |47.79412% ~56 s remaining |======================
|47.87582% ~56 s remaining |======================                              |47.95752% ~56 s
remaining |======================                              |48.03922% ~56 s remaining |=====
==================                                 |48.12092% ~56 s remaining |==================
====                             |48.20261% ~55 s remaining |======================
|48.28431% ~55 s remaining |======================                              |48.36601% ~55 s
remaining |======================                              |48.44771% ~55 s remaining |=====
==================                                 |48.52941% ~55 s remaining |==================
====                             |48.61111% ~55 s remaining |======================
|48.69281% ~55 s remaining |======================                              |48.77451% ~55 s
remaining |======================                              |48.85621% ~55 s remaining |=====
==================                                 |48.93791% ~55 s remaining |==================
====                             |49.01961% ~54 s remaining |======================
|49.10131% ~54 s remaining |======================                              |49.18301% ~54 s
remaining |======================                              |49.26471% ~54 s remaining |=====
==================                                 |49.34641% ~54 s remaining |==================
====                             |49.4281% ~54 s remaining  |======================
|49.5098% ~54 s remaining  |======================                              |49.5915% ~54 s
remaining  |======================                               |49.6732% ~54 s remaining  |====
==================                                 |49.7549% ~54 s remaining  |==================
=====                             |49.8366% ~53 s remaining  |======================
|49.9183% ~53 s remaining  |======================                              | 50% ~53 s rema
ining      |======================                               |50.0817% ~53 s remaining  |====
==================                                 |50.1634% ~53 s remaining  |==================
======                             |50.2451% ~53 s remaining  |======================
|50.3268% ~53 s remaining  |======================                              |50.4085% ~53 s
remaining  |======================                               |50.4902% ~53 s remaining  |====
==================                                 |50.5719% ~53 s remaining  |==================
======                             |50.65359% ~52 s remaining |======================
|50.73529% ~52 s remaining |======================                              |50.81699% ~52 s
remaining |======================                              |50.89869% ~52 s remaining |=====
==================                                 |50.98039% ~52 s remaining |==================
=====                             |51.06209% ~52 s remaining |======================
|51.14379% ~52 s remaining |======================                              |51.22549% ~52 s
```

```
remaining |==========================                        |51.30719% ~52 s remaining |=====
=====================                        |51.38889% ~52 s remaining |=====================
=====                        |51.47059% ~51 s remaining |==========================
|51.55229% ~51 s remaining |==========================                        |51.63399% ~51 s
remaining |==========================                        |51.71569% ~51 s remaining |=====
=====================                        |51.79739% ~51 s remaining |=====================
=====                        |51.87908% ~51 s remaining |===========================
|51.96078% ~51 s remaining |===========================                        |52.04248% ~51 s
remaining |===========================                        |52.12418% ~51 s remaining |=====
=====================                        |52.20588% ~51 s remaining |=====================
======                        |52.28758% ~50 s remaining |===========================
|52.36928% ~50 s remaining |===========================                        |52.45098% ~50 s
remaining |===========================                        |52.53268% ~50 s remaining |=====
=====================                        |52.61438% ~50 s remaining |=====================
======                        |52.69608% ~50 s remaining |===========================
|52.77778% ~50 s remaining |===========================                        |52.85948% ~50 s
remaining |===========================                        |52.94118% ~50 s remaining |=====
=====================                        |53.02288% ~50 s remaining |=====================
======                        |53.10458% ~50 s remaining |===========================
|53.18627% ~49 s remaining |===========================                        |53.26797% ~49 s
remaining |===========================                        |53.34967% ~49 s remaining |=====
=====================                        |53.43137% ~49 s remaining |=====================
======                        |53.51307% ~49 s remaining |===========================
|53.59477% ~49 s remaining |===========================                        |53.67647% ~49 s
remaining |===========================                        |53.75817% ~49 s remaining |=====
=====================                        |53.83987% ~49 s remaining |=====================
=======                        |53.92157% ~49 s remaining |============================
|54.00327% ~49 s remaining |============================                        |54.08497% ~48 s
remaining |============================                        |54.16667% ~48 s remaining |=====
=====================                        |54.24837% ~48 s remaining |=====================
=======                        |54.33007% ~48 s remaining |============================
|54.41176% ~48 s remaining |============================                        |54.49346% ~48 s
remaining |============================                        |54.57516% ~48 s remaining |=====
=====================                        |54.65686% ~48 s remaining |=====================
=======                        |54.73856% ~48 s remaining |============================
|54.82026% ~48 s remaining |============================                        |54.90196% ~48 s
remaining |============================                        |54.98366% ~48 s remaining |=====
=====================                        |55.06536% ~48 s remaining |=====================
=======                        |55.14706% ~48 s remaining |============================
|55.22876% ~48 s remaining |============================                        |55.31046% ~48 s
remaining |============================                        |55.39216% ~48 s remaining |=====
=====================                        |55.47386% ~48 s remaining |=====================
=======                        |55.55556% ~47 s remaining |============================
|55.63725% ~47 s remaining |============================                        |55.71895% ~47 s
remaining |============================                        |55.80065% ~47 s remaining |=====
=======================                        |55.88235% ~47 s remaining |=====================
========                        |55.96405% ~47 s remaining |=============================
|56.04575% ~47 s remaining |=============================                        |56.12745% ~47 s
remaining |=============================                        |56.20915% ~47 s remaining |=====
=======================                        |56.29085% ~47 s remaining |=====================
========                        |56.37255% ~46 s remaining |============================
|56.45425% ~46 s remaining |============================                        |56.53595% ~46 s
```

```
remaining |============================                      |56.61765% ~46 s remaining |=====
=======================                      |56.69935% ~46 s remaining |====================
========                      |56.78105% ~46 s remaining |============================
|56.86275% ~46 s remaining |============================                      |56.94444% ~46 s
remaining |============================                      |57.02614% ~46 s remaining |=====
=======================                      |57.10784% ~46 s remaining |====================
========                      |57.18954% ~45 s remaining |============================
|57.27124% ~45 s remaining |============================                      |57.35294% ~45 s
remaining |============================                      |57.43464% ~45 s remaining |=====
=======================                      |57.51634% ~45 s remaining |====================
========                      |57.59804% ~45 s remaining |============================
|57.67974% ~45 s remaining |============================                      |57.76144% ~45 s
remaining |============================                      |57.84314% ~45 s remaining |=====
=======================                      |57.92484% ~45 s remaining |====================
=========                      |58.00654% ~44 s remaining |============================
|58.08824% ~44 s remaining |============================                      |58.16993% ~44 s
remaining |============================                      |58.25163% ~44 s remaining |=====
=======================                      |58.33333% ~44 s remaining |====================
=========                      |58.41503% ~44 s remaining |============================
|58.49673% ~44 s remaining |============================                      |58.57843% ~44 s
remaining |============================                      |58.66013% ~44 s remaining |=====
=======================                      |58.74183% ~44 s remaining |====================
=========                      |58.82353% ~44 s remaining |============================
|58.90523% ~44 s remaining |============================                      |58.98693% ~43 s
remaining |============================                      |59.06863% ~43 s remaining |=====
=======================                      |59.15033% ~43 s remaining |====================
=========                      |59.23203% ~43 s remaining |============================
|59.31373% ~43 s remaining |============================                      |59.39542% ~43 s
remaining |============================                      |59.47712% ~43 s remaining |=====
=======================                      |59.55882% ~43 s remaining |====================
==========                      |59.64052% ~43 s remaining |============================
|59.72222% ~43 s remaining |============================                      |59.80392% ~43 s
remaining |============================                      |59.88562% ~43 s remaining |=====
=======================                      |59.96732% ~42 s remaining |====================
==========                      |60.04902% ~42 s remaining |============================
|60.13072% ~42 s remaining |============================                      |60.21242% ~42 s
remaining |============================                      |60.29412% ~42 s remaining |=====
=======================                      |60.37582% ~42 s remaining |====================
==========                      |60.45752% ~42 s remaining |============================
|60.53922% ~42 s remaining |============================                      |60.62092% ~42 s
remaining |============================                      |60.70261% ~42 s remaining |=====
=======================                      |60.78431% ~42 s remaining |====================
==========                      |60.86601% ~42 s remaining |============================
|60.94771% ~41 s remaining |============================                      |61.02941% ~41 s
remaining |============================                      |61.11111% ~41 s remaining |=====
=======================                      |61.19281% ~41 s remaining |====================
==========                      |61.27451% ~41 s remaining |============================
|61.35621% ~41 s remaining |============================                      |61.43791% ~41 s
remaining |============================                      |61.51961% ~41 s remaining |=====
=======================                      |61.60131% ~41 s remaining |====================
===========                      |61.68301% ~41 s remaining |============================
|61.76471% ~41 s remaining |============================                      |61.84641% ~41 s
```

```
remaining |================================                  |61.9281% ~41 s remaining  |=====
==========================                  |62.0098% ~41 s remaining  |====================
==========                  |62.0915% ~40 s remaining  |================================
|62.1732% ~40 s remaining  |================================                  |62.2549% ~40 s
remaining  |================================                  |62.3366% ~40 s remaining  |====
==========================                  |62.4183% ~40 s remaining  |====================
==========                  |62.5% ~40 s remaining     |================================
|62.5817% ~40 s remaining  |================================                  |62.6634% ~40 s
remaining  |================================                  |62.7451% ~40 s remaining  |====
==========================                  |62.8268% ~40 s remaining  |====================
==========                  |62.9085% ~40 s remaining  |================================
|62.9902% ~40 s remaining  |================================                  |63.0719% ~39 s
remaining  |================================                  |63.15359% ~39 s remaining |====
==========================                  |63.23529% ~39 s remaining |====================
==========                  |63.31699% ~39 s remaining |================================
|63.39869% ~39 s remaining |================================                  |63.48039% ~39 s
remaining |================================                  |63.56209% ~39 s remaining |=====
==========================                  |63.64379% ~39 s remaining |====================
==========                  |63.72549% ~39 s remaining |================================
|63.80719% ~39 s remaining |================================                  |63.88889% ~39 s
remaining |================================                  |63.97059% ~39 s remaining |=====
==========================                  |64.05229% ~38 s remaining |====================
==========                  |64.13399% ~38 s remaining |================================
|64.21569% ~38 s remaining |================================                  |64.29739% ~38 s
remaining |================================                  |64.37908% ~38 s remaining |=====
==========================                  |64.46078% ~38 s remaining |====================
==========                  |64.54248% ~38 s remaining |================================
|64.62418% ~38 s remaining |================================                  |64.70588% ~38 s
remaining |================================                  |64.78758% ~38 s remaining |=====
==========================                  |64.86928% ~37 s remaining |====================
==========                  |64.95098% ~37 s remaining |================================
|65.03268% ~37 s remaining |================================                  |65.11438% ~37 s
remaining |================================                  |65.19608% ~37 s remaining |=====
==========================                  |65.27778% ~37 s remaining |====================
==========                  |65.35948% ~37 s remaining |================================
|65.44118% ~37 s remaining |================================                  |65.52288% ~37 s
remaining |================================                  |65.60458% ~37 s remaining |=====
==========================                  |65.68627% ~37 s remaining |====================
==========                  |65.76797% ~37 s remaining |================================
|65.84967% ~36 s remaining |================================                  |65.93137% ~36 s
remaining |================================                  |66.01307% ~36 s remaining |=====
==========================                  |66.09477% ~36 s remaining |====================
==========                  |66.17647% ~36 s remaining |================================
|66.25817% ~36 s remaining |================================                  |66.33987% ~36 s
remaining |================================                  |66.42157% ~36 s remaining |=====
==========================                  |66.50327% ~36 s remaining |====================
==========                  |66.58497% ~36 s remaining |================================
|66.66667% ~36 s remaining |================================                  |66.74837% ~36 s
remaining |================================                  |66.83007% ~36 s remaining |=======
====================                  |66.91176% ~36 s remaining |==========================
======                  |66.99346% ~36 s remaining |================================
|67.07516% ~36 s remaining |================================                  |67.15686% ~35 s
```

```
remaining |==================================                |67.23856% ~35 s remaining |=====
==============================                |67.32026% ~35 s remaining |====================
==============                |67.40196% ~35 s remaining |==================================
|67.48366% ~35 s remaining |==================================                |67.56536% ~35 s
remaining |==================================                |67.64706% ~35 s remaining |=====
==============================                |67.72876% ~35 s remaining |====================
==============                |67.81046% ~35 s remaining |==================================
|67.89216% ~35 s remaining |==================================                |67.97386% ~35 s
remaining |==================================                |68.05556% ~34 s remaining |=====
==============================                |68.13725% ~34 s remaining |====================
==============                |68.21895% ~34 s remaining |==================================
|68.30065% ~34 s remaining |==================================                |68.38235% ~34 s
remaining |==================================                |68.46405% ~34 s remaining |=====
==============================                |68.54575% ~34 s remaining |====================
==============                |68.62745% ~34 s remaining |==================================
|68.70915% ~34 s remaining |==================================                |68.79085% ~34 s
remaining |==================================                |68.87255% ~33 s remaining |=====
==============================                |68.95425% ~33 s remaining |====================
==============                |69.03595% ~33 s remaining |==================================
|69.11765% ~33 s remaining |==================================                |69.19935% ~33 s
remaining |==================================                |69.28105% ~33 s remaining |=====
==============================                |69.36275% ~33 s remaining |====================
==============                |69.44444% ~33 s remaining |==================================
|69.52614% ~33 s remaining |==================================                |69.60784% ~33 s
remaining |==================================                |69.68954% ~33 s remaining |=====
==============================                |69.77124% ~32 s remaining |====================
==============                |69.85294% ~32 s remaining |==================================
|69.93464% ~32 s remaining |==================================                |70.01634% ~32 s
remaining |==================================                |70.09804% ~32 s remaining |=====
==============================                |70.17974% ~32 s remaining |====================
==============                |70.26144% ~32 s remaining |==================================
|70.34314% ~32 s remaining |==================================                |70.42484% ~32 s
remaining |==================================                |70.50654% ~32 s remaining |=====
==============================                |70.58824% ~32 s remaining |====================
==============                |70.66993% ~31 s remaining |==================================
|70.75163% ~31 s remaining |==================================                |70.83333% ~31 s
remaining |==================================                |70.91503% ~31 s remaining |=====
==============================                |70.99673% ~31 s remaining |====================
==============                |71.07843% ~31 s remaining |==================================
|71.16013% ~31 s remaining |==================================                |71.24183% ~31 s
remaining |==================================                |71.32353% ~31 s remaining |=====
==============================                |71.40523% ~31 s remaining |====================
==============                |71.48693% ~31 s remaining |==================================
|71.56863% ~30 s remaining |==================================                |71.65033% ~30 s
remaining |==================================                |71.73203% ~30 s remaining |=====
==============================                |71.81373% ~30 s remaining |====================
==============                |71.89542% ~30 s remaining |==================================
|71.97712% ~30 s remaining |==================================                |72.05882% ~30 s
remaining |==================================                |72.14052% ~30 s remaining |=====
==============================                |72.22222% ~30 s remaining |====================
==============                |72.30392% ~30 s remaining |==================================
|72.38562% ~30 s remaining |==================================                |72.46732% ~29 s
```

```
remaining |=====================================              |72.54902% ~29 s remaining |=====
================================              |72.63072% ~29 s remaining |====================
================              |72.71242% ~29 s remaining |=====================================
|72.79412% ~29 s remaining |=====================================              |72.87582% ~29 s
remaining |=====================================              |72.95752% ~29 s remaining |=====
================================              |73.03922% ~29 s remaining |====================
=================              |73.12092% ~29 s remaining |=====================================
=              |73.20261% ~29 s remaining |=====================================              |
73.28431% ~29 s remaining |=====================================              |73.36601% ~28 s
remaining |=====================================              |73.44771% ~28 s remaining |=====
================================              |73.52941% ~28 s remaining |====================
=================              |73.61111% ~28 s remaining |=====================================
=              |73.69281% ~28 s remaining |=====================================              |
73.77451% ~28 s remaining |=====================================              |73.85621% ~28 s
remaining |=====================================              |73.93791% ~28 s remaining |=====
================================              |74.01961% ~28 s remaining |====================
=================              |74.10131% ~28 s remaining |=====================================
=              |74.18301% ~28 s remaining |=====================================              |
74.26471% ~27 s remaining |=====================================              |74.34641% ~27 s
remaining |=====================================              |74.4281% ~27 s remaining  |=====
================================              |74.5098% ~27 s remaining  |====================
=================              |74.5915% ~27 s remaining  |=====================================
=              |74.6732% ~27 s remaining  |=====================================              |
74.7549% ~27 s remaining  |=====================================              |74.8366% ~27 s r
emaining  |=====================================              |74.9183% ~27 s remaining  |=====
================================              | 75% ~27 s remaining      |====================
=================              |75.0817% ~27 s remaining  |=====================================
==              |75.1634% ~26 s remaining  |=====================================              |
75.2451% ~26 s remaining  |=====================================              |75.3268% ~26 s r
emaining  |=====================================              |75.4085% ~26 s remaining  |=====
================================              |75.4902% ~26 s remaining  |====================
=================              |75.5719% ~26 s remaining  |=====================================
==              |75.65359% ~26 s remaining |=====================================              |
75.73529% ~26 s remaining |=====================================              |75.81699% ~26 s
remaining |=====================================              |75.89869% ~26 s remaining |=====
================================              |75.98039% ~26 s remaining |====================
=================              |76.06209% ~25 s remaining |=====================================
==              |76.14379% ~25 s remaining |=====================================              |
76.22549% ~25 s remaining |=====================================              |76.30719% ~25 s
remaining |=====================================              |76.38889% ~25 s remaining |=====
================================              |76.47059% ~25 s remaining |====================
=================              |76.55229% ~25 s remaining |=====================================
==              |76.63399% ~25 s remaining |=====================================              |
76.71569% ~25 s remaining |=====================================              |76.79739% ~25 s
remaining |=====================================              |76.87908% ~25 s remaining |=====
================================              |76.96078% ~24 s remaining |====================
=================              |77.04248% ~24 s remaining |=====================================
===             |77.12418% ~24 s remaining |=====================================              |
77.20588% ~24 s remaining |=====================================              |77.28758% ~24 s
remaining |=====================================              |77.36928% ~24 s remaining |=====
================================              |77.45098% ~24 s remaining |====================
=================              |77.53268% ~24 s remaining |=====================================
```

```
===            |77.61438% ~24 s remaining |=======================================           |
77.69608% ~24 s remaining |=======================================           |77.77778% ~24 s
remaining |=======================================           |77.85948% ~23 s remaining |=====
==================================           |77.94118% ~23 s remaining |====================
==================           |78.02288% ~23 s remaining |=====================================
===            |78.10458% ~23 s remaining |=======================================           |
78.18627% ~23 s remaining |=======================================           |78.26797% ~23 s
remaining |=======================================           |78.34967% ~23 s remaining |=====
==================================           |78.43137% ~23 s remaining |====================
==================           |78.51307% ~23 s remaining |=====================================
===            |78.59477% ~23 s remaining |=======================================           |
78.67647% ~23 s remaining |=======================================           |78.75817% ~23 s
remaining |=======================================           |78.83987% ~22 s remaining |=====
==================================           |78.92157% ~22 s remaining |====================
====================           |79.00327% ~22 s remaining |=====================================
====           |79.08497% ~22 s remaining |=======================================           |
79.16667% ~22 s remaining |=======================================           |79.24837% ~22 s
remaining |=======================================           |79.33007% ~22 s remaining |=====
==================================           |79.41176% ~22 s remaining |====================
====================           |79.49346% ~22 s remaining |=====================================
====           |79.57516% ~22 s remaining |=======================================           |
79.65686% ~22 s remaining |=======================================           |79.73856% ~21 s
remaining |=======================================           |79.82026% ~21 s remaining |=====
==================================           |79.90196% ~21 s remaining |====================
====================           |79.98366% ~21 s remaining |=====================================
====           |80.06536% ~21 s remaining |=======================================           |
80.14706% ~21 s remaining |=======================================           |80.22876% ~21 s
remaining |=======================================           |80.31046% ~21 s remaining |=====
==================================           |80.39216% ~21 s remaining |====================
====================           |80.47386% ~21 s remaining |=====================================
====           |80.55556% ~21 s remaining |=======================================           |
80.63725% ~21 s remaining |=======================================           |80.71895% ~21 s
remaining |=======================================           |80.80065% ~21 s remaining |=====
==================================           |80.88235% ~20 s remaining |====================
====================           |80.96405% ~20 s remaining |=====================================
=====           |81.04575% ~20 s remaining |=======================================           |
81.12745% ~20 s remaining |=======================================           |81.20915% ~20 s
remaining |=======================================           |81.29085% ~20 s remaining |=====
==================================           |81.37255% ~20 s remaining |====================
====================           |81.45425% ~20 s remaining |=====================================
=====           |81.53595% ~20 s remaining |=======================================           |
81.61765% ~20 s remaining |=======================================           |81.69935% ~20 s
remaining |=======================================           |81.78105% ~19 s remaining |=====
==================================           |81.86275% ~19 s remaining |====================
====================           |81.94444% ~19 s remaining |=====================================
=====           |82.02614% ~19 s remaining |=======================================           |
82.10784% ~19 s remaining |=======================================           |82.18954% ~19 s
remaining |=======================================           |82.27124% ~19 s remaining |=====
==================================           |82.35294% ~19 s remaining |====================
====================           |82.43464% ~19 s remaining |=====================================
=====           |82.51634% ~19 s remaining |=======================================           |
82.59804% ~19 s remaining |=======================================           |82.67974% ~18 s
```

```
remaining |=========================================         |82.76144% ~18 s remaining |=====
=====================================         |82.84314% ~18 s remaining |====================
====================         |82.92484% ~18 s remaining |=====================================
======         |83.00654% ~18 s remaining |=========================================         |
83.08824% ~18 s remaining |=========================================         |83.16993% ~18 s
remaining |=========================================         |83.25163% ~18 s remaining |=====
=====================================         |83.33333% ~18 s remaining |====================
====================         |83.41503% ~18 s remaining |=====================================
======         |83.49673% ~18 s remaining |=========================================         |
83.57843% ~18 s remaining |=========================================         |83.66013% ~17 s
remaining |=========================================         |83.74183% ~17 s remaining |=====
=====================================         |83.82353% ~17 s remaining |====================
====================         |83.90523% ~17 s remaining |=====================================
======         |83.98693% ~17 s remaining |=========================================         |
84.06863% ~17 s remaining |=========================================         |84.15033% ~17 s
remaining |=========================================         |84.23203% ~17 s remaining |=====
=====================================         |84.31373% ~17 s remaining |====================
====================         |84.39542% ~17 s remaining |=====================================
======         |84.47712% ~17 s remaining |=========================================         |
84.55882% ~16 s remaining |=========================================         |84.64052% ~16 s
remaining |=========================================         |84.72222% ~16 s remaining |=====
=====================================         |84.80392% ~16 s remaining |====================
====================         |84.88562% ~16 s remaining |=====================================
=======         |84.96732% ~16 s remaining |=========================================         |
85.04902% ~16 s remaining |=========================================         |85.13072% ~16 s
remaining |=========================================         |85.21242% ~16 s remaining |=====
=====================================         |85.29412% ~16 s remaining |====================
====================         |85.37582% ~16 s remaining |=====================================
=======         |85.45752% ~15 s remaining |=========================================         |
85.53922% ~15 s remaining |=========================================         |85.62092% ~15 s
remaining |=========================================         |85.70261% ~15 s remaining |=====
=====================================         |85.78431% ~15 s remaining |====================
====================         |85.86601% ~15 s remaining |=====================================
=======         |85.94771% ~15 s remaining |=========================================         |
86.02941% ~15 s remaining |=========================================         |86.11111% ~15 s
remaining |=========================================         |86.19281% ~15 s remaining |=====
=====================================         |86.27451% ~15 s remaining |====================
====================         |86.35621% ~14 s remaining |=====================================
=======         |86.43791% ~14 s remaining |=========================================         |
86.51961% ~14 s remaining |=========================================        |86.60131% ~14 s
remaining |=========================================        |86.68301% ~14 s remaining |=====
=====================================        |86.76471% ~14 s remaining |====================
====================        |86.84641% ~14 s remaining |=====================================
========        |86.9281% ~14 s remaining  |=========================================        |
87.0098% ~14 s remaining  |=========================================        |87.0915% ~14 s r
emaining  |=========================================        |87.1732% ~14 s remaining  |=====
=====================================        |87.2549% ~13 s remaining  |====================
====================        |87.3366% ~13 s remaining  |=====================================
========        |87.4183% ~13 s remaining  |=========================================        |
87.5% ~13 s remaining     |=========================================        |87.5817% ~13 s r
emaining  |=========================================        |87.6634% ~13 s remaining  |=====
=====================================        |87.7451% ~13 s remaining  |====================
```

```
=======================       |87.8268% ~13 s remaining  |=================================
========       |87.9085% ~13 s remaining  |=========================================       |
87.9902% ~13 s remaining  |=========================================       |88.0719% ~13 s r
emaining  |=========================================       |88.15359% ~13 s remaining |=====
=====================================       |88.23529% ~12 s remaining |====================
=======================       |88.31699% ~12 s remaining |=================================
========       |88.39869% ~12 s remaining |=========================================      |
88.48039% ~12 s remaining |=========================================      |88.56209% ~12 s
remaining |=========================================      |88.64379% ~12 s remaining |=====
=====================================      |88.72549% ~12 s remaining |====================
========================      |88.80719% ~12 s remaining |=================================
=========      |88.88889% ~12 s remaining |=========================================      |
88.97059% ~12 s remaining |=========================================      |89.05229% ~12 s
remaining |=========================================      |89.13399% ~11 s remaining |=====
=====================================      |89.21569% ~11 s remaining |====================
========================      |89.29739% ~11 s remaining |=================================
=========      |89.37908% ~11 s remaining |=========================================      |
89.46078% ~11 s remaining |=========================================      |89.54248% ~11 s
remaining |=========================================      |89.62418% ~11 s remaining |=====
=====================================      |89.70588% ~11 s remaining |====================
========================      |89.78758% ~11 s remaining |=================================
=========      |89.86928% ~11 s remaining |=========================================      |
89.95098% ~11 s remaining |=========================================      |90.03268% ~10 s
remaining |=========================================      |90.11438% ~10 s remaining |=====
=====================================      |90.19608% ~10 s remaining |====================
========================      |90.27778% ~10 s remaining |=================================
=========      |90.35948% ~10 s remaining |=========================================     |
90.44118% ~10 s remaining |=========================================     |90.52288% ~10 s
remaining |=========================================     |90.60458% ~10 s remaining |=====
=====================================     |90.68627% ~10 s remaining |====================
========================     |90.76797% ~10 s remaining |=================================
==========     |90.84967% ~10 s remaining |=========================================     |
90.93137% ~10 s remaining |=========================================     |91.01307% ~9 s r
emaining  |=========================================     |91.09477% ~9 s remaining  |=====
=====================================     |91.17647% ~9 s remaining  |====================
========================     |91.25817% ~9 s remaining  |=================================
==========     |91.33987% ~9 s remaining  |=========================================     |
91.42157% ~9 s remaining  |=========================================     |91.50327% ~9 s r
emaining  |=========================================     |91.58497% ~9 s remaining  |=====
=====================================     |91.66667% ~9 s remaining  |====================
========================     |91.74837% ~9 s remaining  |=================================
==========     |91.83007% ~9 s remaining  |=========================================     |
91.91176% ~9 s remaining  |=========================================     |91.99346% ~8 s r
emaining  |=========================================     |92.07516% ~8 s remaining  |=====
=====================================     |92.15686% ~8 s remaining  |====================
========================     |92.23856% ~8 s remaining  |=================================
==========    |92.32026% ~8 s remaining  |=========================================    |
92.40196% ~8 s remaining  |=========================================    |92.48366% ~8 s r
emaining  |=========================================    |92.56536% ~8 s remaining  |=====
=====================================    |92.64706% ~8 s remaining  |====================
========================    |92.72876% ~8 s remaining  |=================================
==========    |92.81046% ~8 s remaining  |=========================================    |
```

```
92.89216% ~7 s remaining  |=============================================    |92.97386% ~7 s r
emaining  |=============================================    |93.05556% ~7 s remaining  |=====
=========================================    |93.13725% ~7 s remaining  |====================
========================    |93.21895% ~7 s remaining  |=====================================
==========    |93.30065% ~7 s remaining  |=============================================    |
93.38235% ~7 s remaining  |=============================================    |93.46405% ~7 s r
emaining  |=============================================    |93.54575% ~7 s remaining  |=====
=========================================    |93.62745% ~7 s remaining  |====================
========================    |93.70915% ~7 s remaining  |=====================================
==========    |93.79085% ~7 s remaining  |=============================================    |
93.87255% ~6 s remaining  |=============================================    |93.95425% ~6 s r
emaining  |=============================================    |94.03595% ~6 s remaining  |=====
=========================================    |94.11765% ~6 s remaining  |====================
========================    |94.19935% ~6 s remaining  |=====================================
===========   |94.28105% ~6 s remaining  |=============================================   |
94.36275% ~6 s remaining  |=============================================   |94.44444% ~6 s r
emaining  |=============================================   |94.52614% ~6 s remaining  |=====
=========================================   |94.60784% ~6 s remaining  |====================
========================   |94.68954% ~6 s remaining  |=====================================
===========   |94.77124% ~5 s remaining  |=============================================   |
94.85294% ~5 s remaining  |=============================================   |94.93464% ~5 s r
emaining  |=============================================   |95.01634% ~5 s remaining  |=====
=========================================   |95.09804% ~5 s remaining  |====================
========================   |95.17974% ~5 s remaining  |=====================================
===========   |95.26144% ~5 s remaining  |=============================================   |
95.34314% ~5 s remaining  |=============================================   |95.42484% ~5 s r
emaining  |=============================================   |95.50654% ~5 s remaining  |=====
=========================================   |95.58824% ~5 s remaining  |====================
========================   |95.66993% ~5 s remaining  |=====================================
===========   |95.75163% ~4 s remaining  |=============================================   |
95.83333% ~4 s remaining  |=============================================   |95.91503% ~4 s r
emaining  |=============================================   |95.99673% ~4 s remaining  |=====
=========================================   |96.07843% ~4 s remaining  |====================
========================  |96.16013% ~4 s remaining  |=====================================
============  |96.24183% ~4 s remaining  |============================================  |
96.32353% ~4 s remaining  |============================================  |96.40523% ~4 s r
emaining  |============================================  |96.48693% ~4 s remaining  |=====
=========================================  |96.56863% ~4 s remaining  |====================
========================  |96.65033% ~4 s remaining  |=====================================
============  |96.73203% ~3 s remaining  |============================================  |
96.81373% ~3 s remaining  |============================================  |96.89542% ~3 s r
emaining  |============================================  |96.97712% ~3 s remaining  |=====
=========================================  |97.05882% ~3 s remaining  |====================
========================  |97.14052% ~3 s remaining  |=====================================
============  |97.22222% ~3 s remaining  |============================================  |
97.30392% ~3 s remaining  |============================================  |97.38562% ~3 s r
emaining  |============================================  |97.46732% ~3 s remaining  |=====
=========================================  |97.54902% ~3 s remaining  |====================
========================  |97.63072% ~3 s remaining  |=====================================
============  |97.71242% ~2 s remaining  |============================================  |
97.79412% ~2 s remaining  |============================================  |97.87582% ~2 s r
emaining  |============================================  |97.95752% ~2 s remaining  |=====
```

```
==========================================  |98.03922% ~2 s remaining  |=====================
============================== |98.12092% ~2 s remaining  |=====================================
============== |98.20261% ~2 s remaining  |================================================= |
98.28431% ~2 s remaining  |================================================= |98.36601% ~2 s r
emaining  |================================================= |98.44771% ~2 s remaining  |=====
============================================= |98.52941% ~2 s remaining  |=====================
============================== |98.61111% ~1 s remaining  |=====================================
============== |98.69281% ~1 s remaining  |================================================= |
98.77451% ~1 s remaining  |================================================= |98.85621% ~1 s r
emaining  |================================================= |98.93791% ~1 s remaining  |=====
============================================= |99.01961% ~1 s remaining  |=====================
============================== |99.10131% ~1 s remaining  |=====================================
============== |99.18301% ~1 s remaining  |================================================= |
99.26471% ~1 s remaining  |================================================= |99.34641% ~1 s r
emaining  |================================================= |99.4281% ~1 s remaining   |=====
============================================= |99.5098% ~1 s remaining   |=====================
============================== |99.5915% ~0 s remaining   |=====================================
============== |99.6732% ~0 s remaining   |================================================= |
99.7549% ~0 s remaining   |================================================= |99.8366% ~0 s re
maining   |================================================= |99.9183% ~0 s remaining   |=====
============================================|100% ~0 s remaining       |=====================
```

```
## Starting to add information to samples
```

```
##  => Add clinical information to samples
```

```
##  => Adding TCGA molecular information from marker papers
```

```
##  => Information will have prefix 'paper_'
```

```
## brca subtype information from:doi.org/10.1016/j.ccell.2018.03.014
```

```
## Available assays in SummarizedExperiment :
##   => unstranded
##   => stranded_first
##   => stranded_second
##   => tpm_unstrand
##   => fpkm_unstrand
##   => fpkm_uq_unstrand
```

Extract metadata and Counts matrix

```
# ---------- metadata
metadata <- as.data.frame(colData(data)[, c("gender", "vital_status", "age_at_index", "days_to_l
ast_follow_up", "paper_pathologic_stage", "paper_BRCA_Subtype_PAM50", "tissue_type")])

colnames(metadata) <- c("Gender", "SurvivalStatus", "Age", "DaysToLastFollowUp", "PathologicStag
e", "PAM50", "TissueType")
metadata <- na.omit(metadata)
metadata <- metadata %>% mutate(SurvivalStatus = ifelse(SurvivalStatus == "Alive", FALSE, Surviv
alStatus)) %>%
  mutate(SurvivalStatus = ifelse(SurvivalStatus == "Dead", TRUE, SurvivalStatus))
```

-The metadata is biased to some Cancer sub types: Basal Her2 LumA LumB Normal 176 68 529 186 35 for having a more reliable dataset we will choose an unbiased subset dataset. in total the research contains 313 samples: 35 Normal, 68 Her2 and 70 LumA, LumB and Basal.

```
normal_data <- metadata[metadata$PAM50 == "Normal", ]
her2_data <- metadata[metadata$PAM50 == "Her2", ]
basal_data <- metadata[metadata$PAM50 == "Basal", ]
luma_data <- metadata[metadata$PAM50 == "LumA", ]
lumb_data <- metadata[metadata$PAM50 == "LumB", ]

metadata <- rbind(normal_data, her2_data, basal_data[1:70, ], luma_data[1:70, ], lumb_data[1:70,
])
```

```
# ---------- gene names
gene_ids_to_names <- rowData(data)[, c("gene_id", "gene_name")]

# ---------- counts matrix
counts <- as.data.frame(assay(data, "unstranded"))
counts <- counts[, colnames(counts) %in% rownames(metadata)] #filter
counts <- counts[, rownames(metadata)] # reorder

# Make sure the counts matrix corresponds to the metadata
all(rownames(metadata) == colnames(counts))
```

```
## [1] TRUE
```

Prepare and Run DESeq2

- Constructs a DESeq2 object

```
# Set Factors for dds analyses
metadata$PathologicStage <- as.factor(metadata$PathologicStage)
metadata$SurvivalStatus <- as.factor(metadata$SurvivalStatus)

dds <- DESeqDataSetFromMatrix(countData=counts,
                              colData=metadata,
                              design= ~ PathologicStage + SurvivalStatus)
dds <- DESeq(dds)
```

```
## estimating size factors
```

```
## estimating dispersions
```

```
## gene-wise dispersion estimates
```

```
## mean-dispersion relationship
```

```
## final dispersion estimates
```

```
## fitting model and testing
```

```
## -- replacing outliers and refitting for 11573 genes
## -- DESeq argument 'minReplicatesForReplace' = 7
## -- original counts are preserved in counts(dds)
```

```
## estimating dispersions
```

```
## fitting model and testing
```

Normalize Results

```
res <- results(dds)
significant_genes <- length(which(res$padj < 0.05)) # number if significant genes
print(paste(significant_genes, " significant genes where found."))
```

```
## [1] "432  significant genes where found."
```

```
counts.vst <- vst(dds) # normal the data
counts.vst <- assay(counts.vst)

var_per_gene <- apply(counts.vst, 1, var)  # Calculate the variance per gene
selectedGenes <- names(var_per_gene[order(var_per_gene, decreasing = T)][1:50]) # Take the first
50 genes

counts.vst.significant <- counts.vst[selectedGenes,]  # Construct a new matrix only for the sign
ificant genes

# Make names informative names
genes_names <- gene_ids_to_names$gene_name[gene_ids_to_names$gene_id %in% selectedGenes]
rownames(counts.vst.significant) <-genes_names
colnames(counts.vst.significant) <- paste0(metadata$Gender, "_", metadata$Age, "_", metadata$Pat
hologicStage, "_seq", seq_along(metadata$Gender))
```

Heatmaps by Distance functions

In this section we will plot for each distance function the Heatmap graph. The data is the normalized counts results. For each plot we added annotation for the PAM50 BRCA subtypes.

```
# Heatmap Colors
col_fun = colorRamp2(c(0, 5, 10, 15, 20), brewer.pal(n = 5, name = "RdBu"))
col_PAM50 = c("Basal" = "#FB9A99", "Her2"="#B2DF8A", "LumA"="#A6CEE3", "LumB"="#1F78B4", "Norma
l"="#FDBF6F")

# ---------------------------- Heat Maps ----------------------------

# ---------------------------- Euclidean Distance

euc_dist <- dist(counts.vst.significant, method = 'euclidean') # Create euclidean distance matri
x for heatmap
euc_row_dend = dendsort(hclust(dist(counts.vst.significant))) # for row ordering
euc_col_dend = dendsort(hclust(dist(t(counts.vst.significant)))) # for column ordering

# the heatmap
ht_euc = ComplexHeatmap::Heatmap(counts.vst.significant,
                      column_title = "Euclidean Distance",
                      name = " ",
                      show_row_names = TRUE, show_column_names = FALSE,
                      col = col_fun,
                      clustering_distance_rows = function(x) euc_dist, # calculates the heatmap
by distance matrix
                      clustering_distance_columns = function(x) euc_dist,
                      clustering_method_rows = "complete",
                      clustering_method_columns = "complete",
                      column_dend_height = unit(2, "cm"),
                      cluster_rows = euc_row_dend, cluster_columns = euc_col_dend,
                      row_names_rot = -45,
                      row_names_gp = grid::gpar(fontsize = 6),
                      top_annotation = ComplexHeatmap::HeatmapAnnotation(PAM50 = c(metadata$PAM5
0), col=list(PAM50 = col_PAM50)),
                      heatmap_width = unit(15, "cm")
                      )

# ---------------------------- Pair wise Correlation

# A function that calculates pairwise distance for matrix
cor_distance <- function(mat) {
  as.dist(1 - cor(t(mat)))
}

cor_row_dend = dendsort(hclust(cor_distance(counts.vst.significant))) # for row ordering
cor_col_dend = dendsort(hclust(cor_distance(t(counts.vst.significant))))  # for column ordering

# the heatmap
ht_cor = ComplexHeatmap::Heatmap(counts.vst.significant,
                      column_title = "Correlation Pairwise Distance",
                      name = " ",
                      show_row_names = TRUE, show_column_names = FALSE,
                      col = col_fun,
                      clustering_distance_rows = function(x, y) 1 - cor(x, y), # calculates the
heatmap by function for each sample
```

```
                      clustering_distance_columns = function(x, y ) 1 - cor(x, y),
                      clustering_method_rows = "complete",
                      clustering_method_columns = "complete",
                      column_dend_height = unit(2, "cm"),
                      cluster_rows = cor_row_dend, cluster_columns = cor_col_dend,
                      row_names_rot = -45,
                      row_names_gp = grid::gpar(fontsize = 6),
                      top_annotation = ComplexHeatmap::HeatmapAnnotation(PAM50 = c(metadata$PAM5
0), col=list(PAM50 = col_PAM50)),
                      heatmap_width = unit(15, "cm")
                      )

# ---------------------------- Cosine Similarity
library(lsa) # for cosine similarity function
```

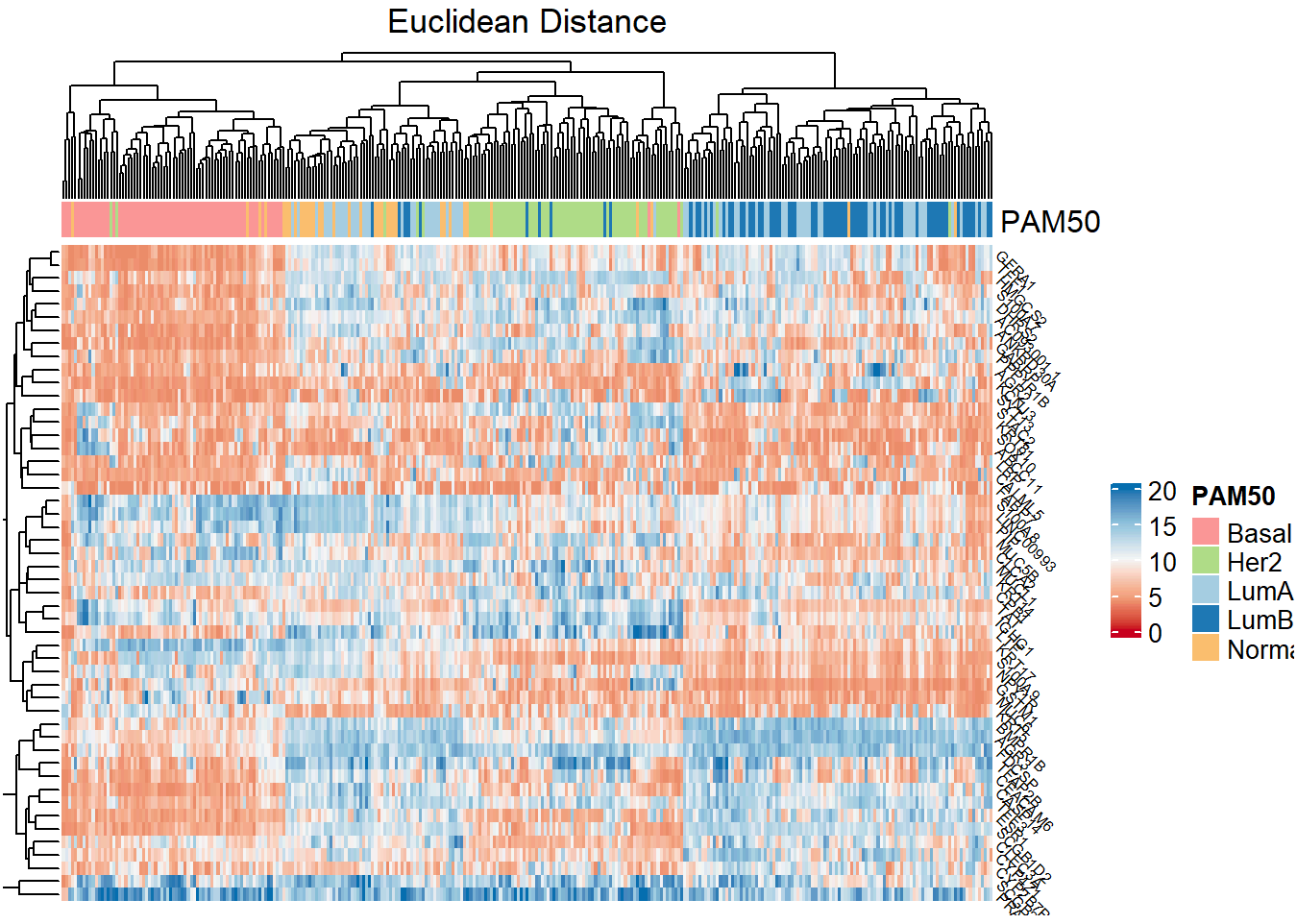
```
## Warning: package 'lsa' was built under R version 4.3.3
```

```
## Loading required package: SnowballC
```

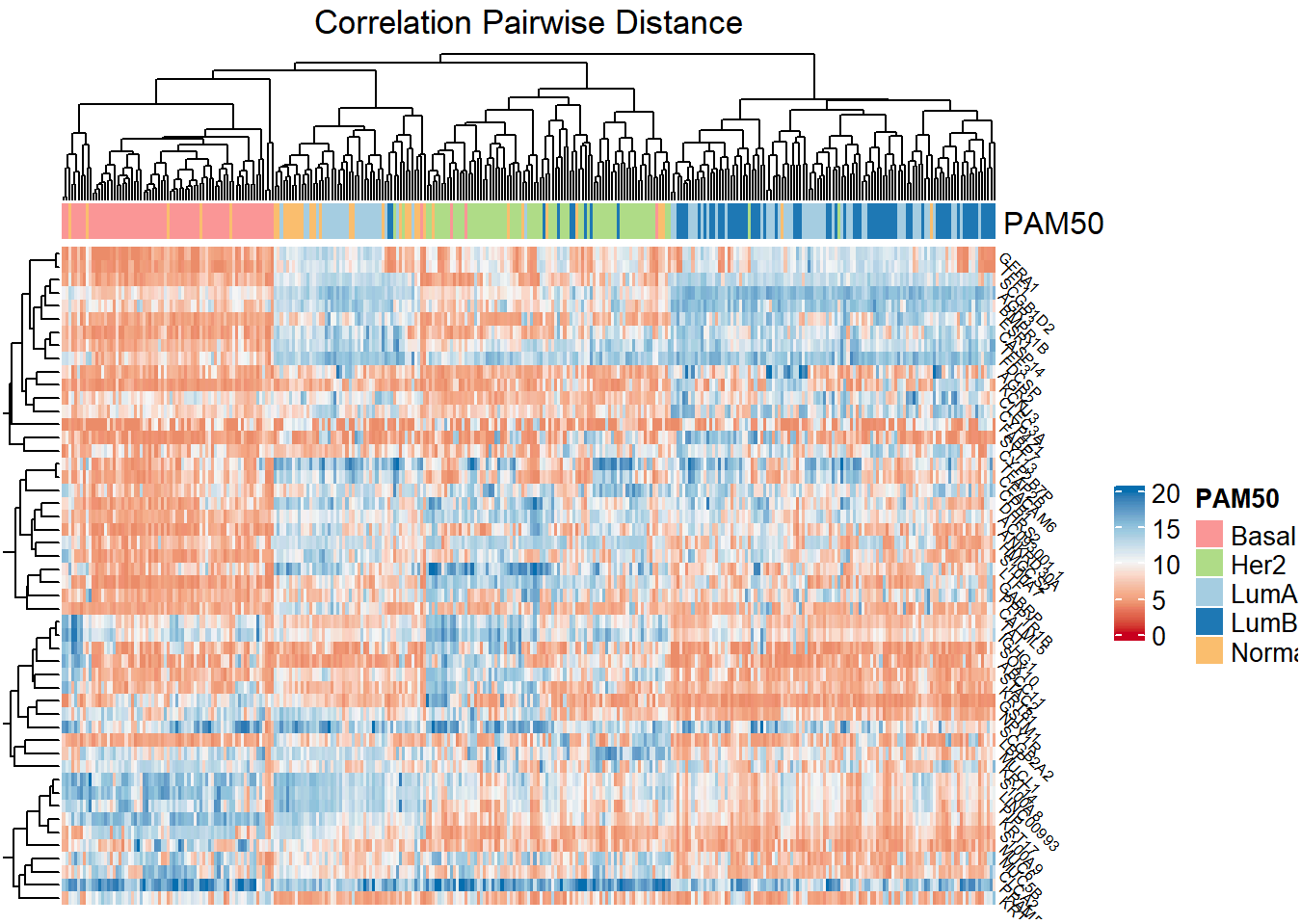
```
# Convert similarity to distance so it can be used to create the clustering
 cos_dist <- as.dist(1 - cosine(counts.vst.significant))

# the heatmap
ht_cos = ComplexHeatmap::Heatmap(counts.vst.significant,
                      column_title = "Cosine Simalarity",
                      name = " ",
                      show_row_names = TRUE, show_column_names = FALSE,
                      col = col_fun,
                      clustering_distance_rows = function(x,y) 1- cosine(x, y), # calculates the
heatmap by function for each sample
                      clustering_distance_columns = function(x,y) 1- cosine(x, y),
                      clustering_method_rows = "complete",
                      clustering_method_columns = "complete",
                      column_dend_height = unit(2, "cm"),
                      row_names_rot = -45,
                      row_names_gp = grid::gpar(fontsize = 6),
                      top_annotation = ComplexHeatmap::HeatmapAnnotation(PAM50 = c(metadata$PAM5
0), col=list(PAM50 = col_PAM50)),
                      heatmap_width = unit(15, "cm")
                      )


# ---------------------------- Kernel Similarity
# A function that calculates kernel similarity for matrix
ker_dis_matrix <- function(mat) {
    product_matrix <- t(mat) %*% mat
    result_matrix <- (product_matrix + 1)^2
    return(result_matrix)
}

ker_row_dend = dendsort(hclust(cos_dist)) # for row ordering
ker_col_dend = dendsort(hclust(t(cos_dist)))  # for column ordering

# the heatmap
ht_ker = ComplexHeatmap::Heatmap(counts.vst.significant,
                      column_title = "Kernel based distance (for nonlinear distance)",
                      name = " ",
                      show_row_names = TRUE, show_column_names = FALSE,
                      col = col_fun,
                      clustering_distance_rows = function(x, y) (t(x)%*%y + 1)^2, # calculates t
he heatmap by function for each sample
                      clustering_distance_columns = function(x, y) (t(x)%*%y + 1)^2,
                      clustering_method_rows = "complete",
                      clustering_method_columns = "complete",
                      column_dend_height = unit(2, "cm"),
                      row_names_rot = -45,
                      row_names_gp = grid::gpar(fontsize = 6),
                      top_annotation = ComplexHeatmap::HeatmapAnnotation(PAM50 = c(metadata$PAM5
0), col=list(PAM50 = col_PAM50)),
                      heatmap_width = unit(15, "cm")
                      )
```

```
# ---------------------------- Mahalanobis Distance

rescaled_data <- rescale(counts.vst.significant) # normalize for smaller values for the mahalano
bis distance function, so it wont exceed the max int value of the computer

# data for the mahalanobis_distance function
means <- colMeans(rescaled_data)
cov_matrix <- cov(rescaled_data) + diag(1e-4, ncol(rescaled_data)) # add a very small number so
the matrix will be diagonalizable
n_features <- ncol(rescaled_data)

# A function that calculates mahalanobis distance for samples
mahalanobis_distance <- function(x, y) {
  x <- matrix(x, nrow = 1, ncol = n_features)
  y <- matrix(y, nrow = 1, ncol = n_features)

  # Check if x or y need padding
  if (length(x) < n_features) {
    x <- x + rep(0, n_features - length(x))
  }
  if (length(y) < n_features) {
    y <- y + rep(0, n_features - length(y))
  }
  # Calculate distance
  return(as.numeric(sqrt((x - y)%*% solve(cov_matrix) %*% t(x - y))))
}

# the heatmap
ht_mah = ComplexHeatmap::Heatmap(counts.vst.significant,
                      column_title = "Mahalanobis distance",
                      name = " ",
                      show_row_names = TRUE, show_column_names = FALSE,
                      col = col_fun,
                      clustering_distance_rows = function(x,y) mahalanobis_distance(x,y), # calc
ulates the heatmap by function for each sample
                      clustering_distance_columns = function(x,y) mahalanobis_distance(x,y),
                      clustering_method_rows = "complete",
                      clustering_method_columns = "complete",
                      column_dend_height = unit(2, "cm"),
                      row_names_rot = -45,
                      row_names_gp = grid::gpar(fontsize = 6),
                      top_annotation = ComplexHeatmap::HeatmapAnnotation(PAM50 = c(metadata$PAM5
0), col=list(PAM50 = col_PAM50)),
                      heatmap_width = unit(15, "cm")
                      )

# ---------------------------- PCA

pcaResults = prcomp(t(counts.vst.significant))
pca_hc = ComplexHeatmap::Heatmap(counts.vst.significant,
```

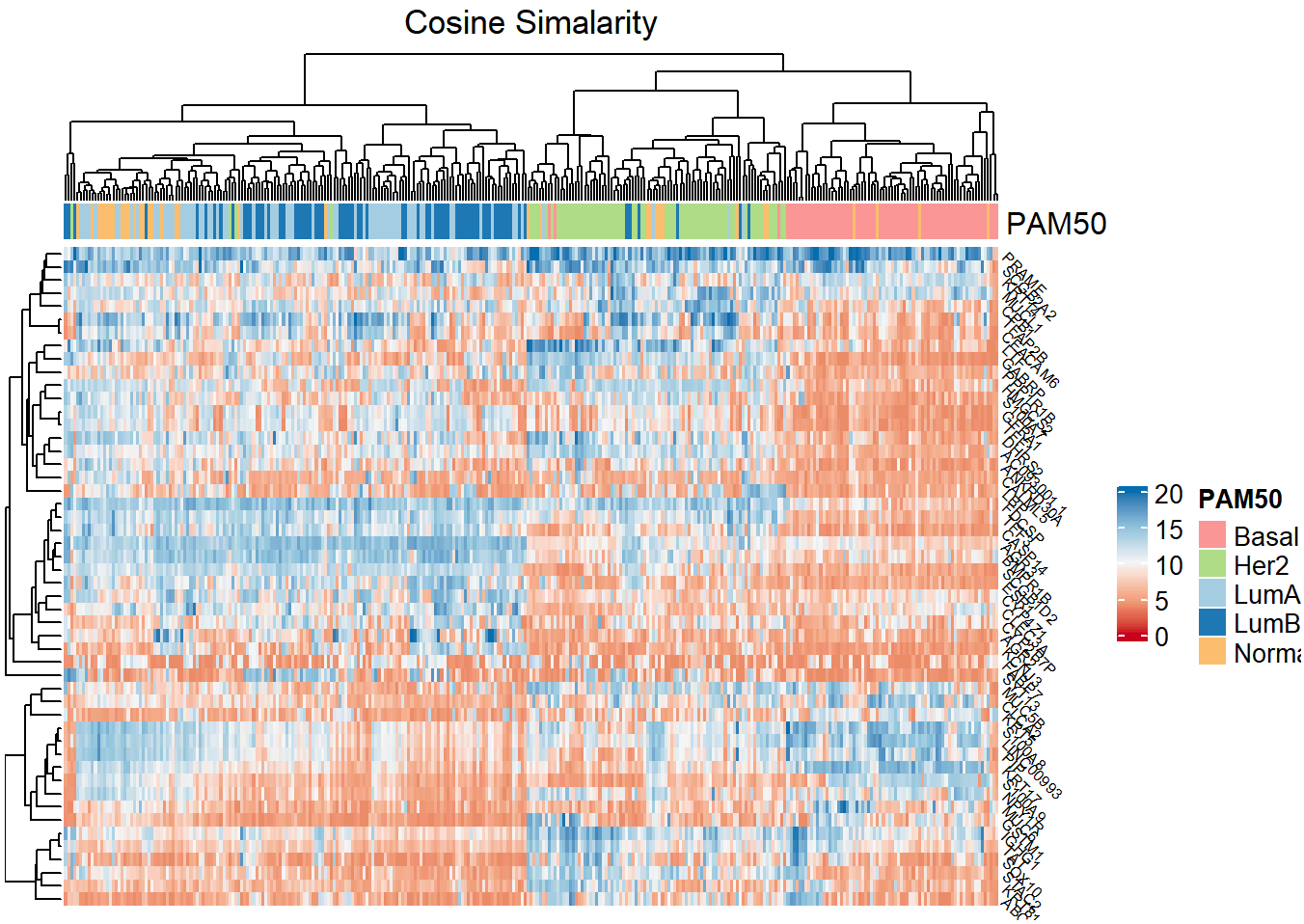
```
                        column_title = "distance by PCA",
                        name = " ",
                        show_row_names = TRUE, show_column_names = FALSE,
                        col = col_fun,
                        clustering_distance_rows = function(x, y) sqrt(sum((pcaResults$x[x, ] -
pcaResults$x[y, ])^2)), # calculates the heatmap by function for each sample
                        clustering_distance_columns = function(x, y) sqrt(sum((pcaResults$x[x, ] -
pcaResults$x[y, ])^2)),
                        clustering_method_rows = "complete",
                        clustering_method_columns = "complete",
                        column_dend_height = unit(2, "cm"),
                        row_names_rot = -45,
                        row_names_gp = grid::gpar(fontsize = 6),
                        top_annotation = ComplexHeatmap::HeatmapAnnotation(PAM50 = c(metadata$PAM5
0), col=list(PAM50 = col_PAM50)),
                        heatmap_width = unit(15, "cm")
                        )


# prints the plots
ht_euc
```



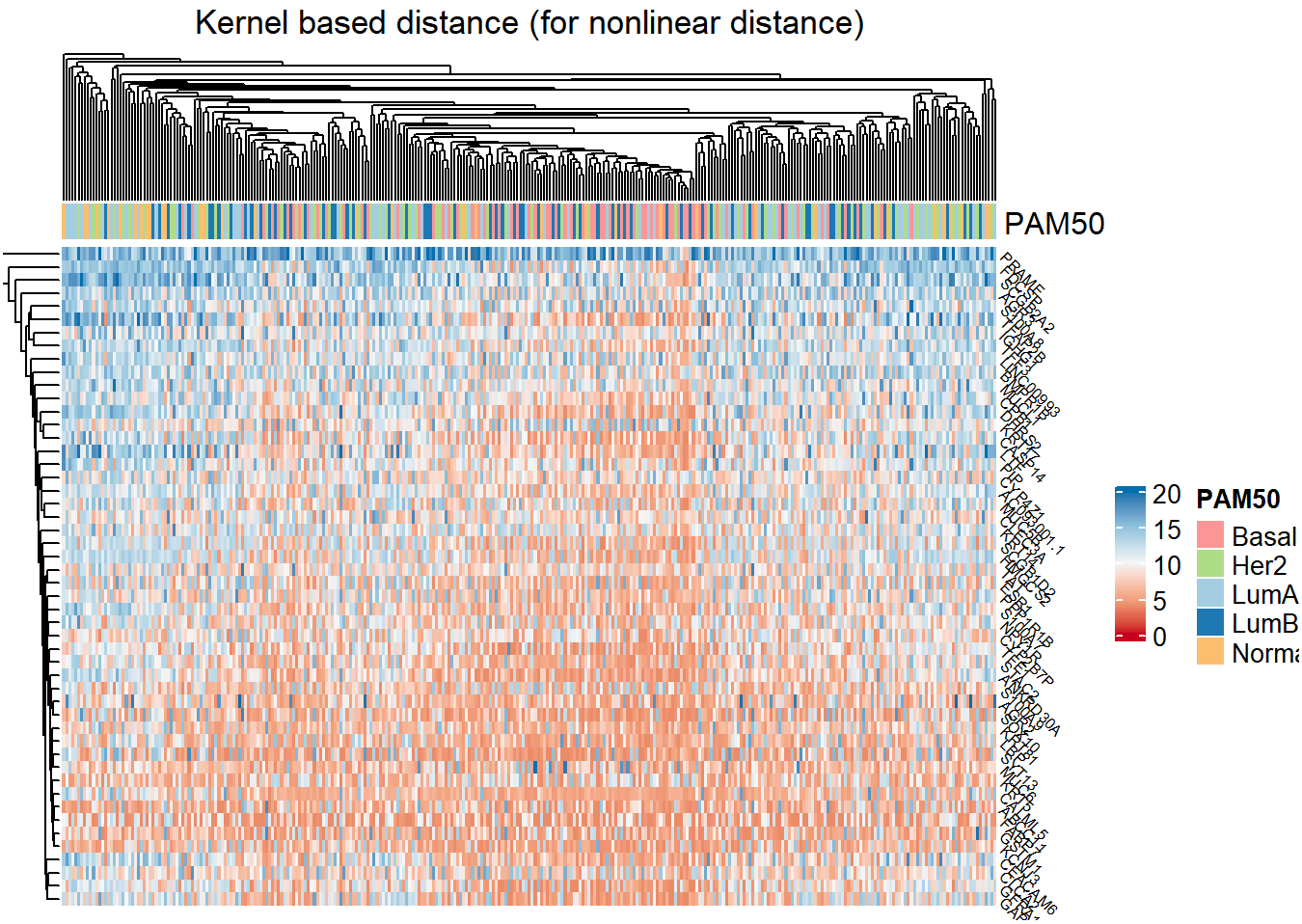

```
ht_cor
```


Correlation Pairwise Distance
PAM50
20
15
10
5
0
PAM50
Basal
Her2
LumA
LumB
Norma

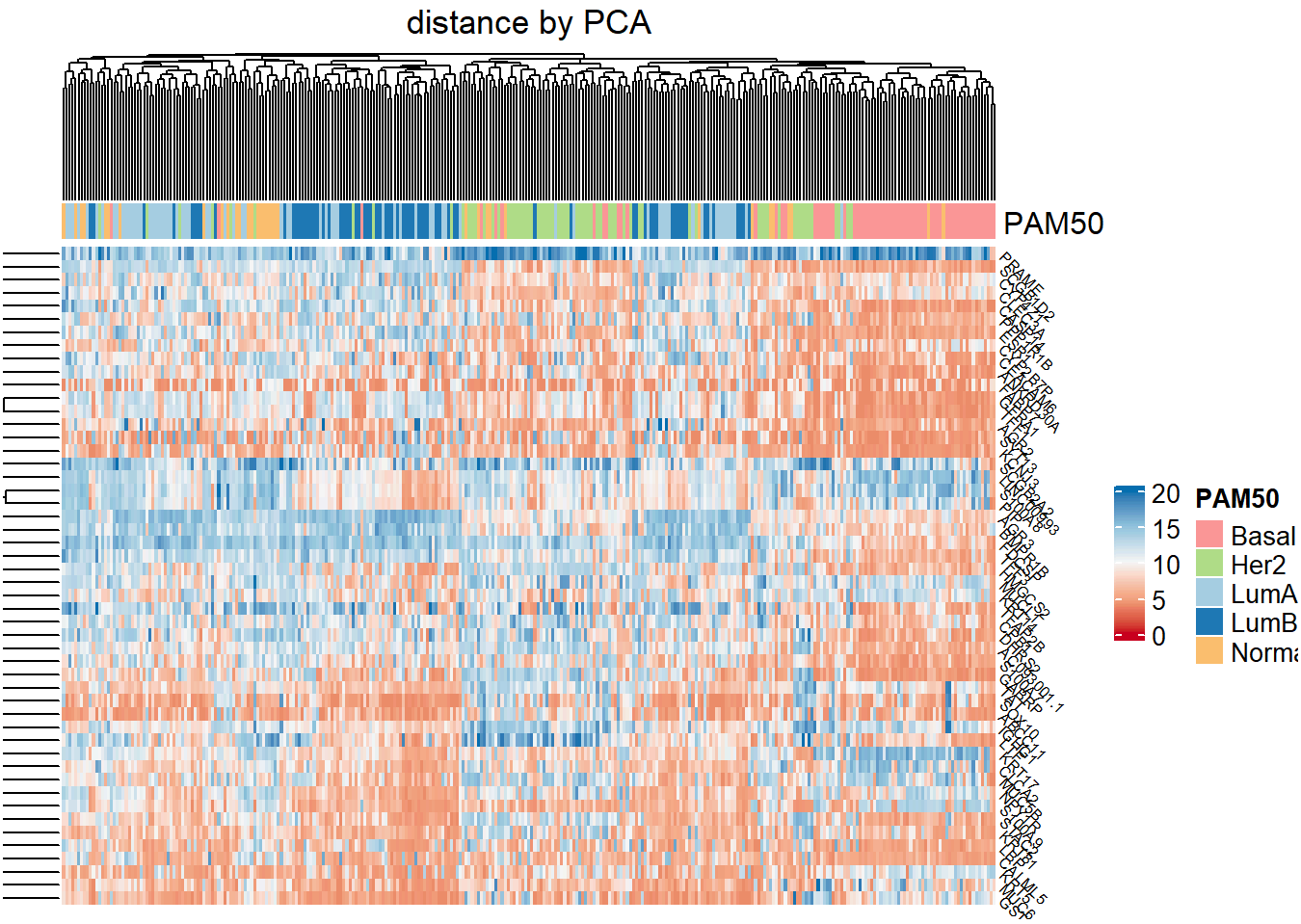
```
ht_cos
```


Cosine Simalarity
PAM50
20
15
10
5
0
PAM50
Basal
Her2
LumA
LumB
Norma

```
ht_ker
```


Kernel based distance (for nonlinear distance)
PAM50
20
15
10
5
0
PAM50
Basal
Her2
LumA
LumB
Norma

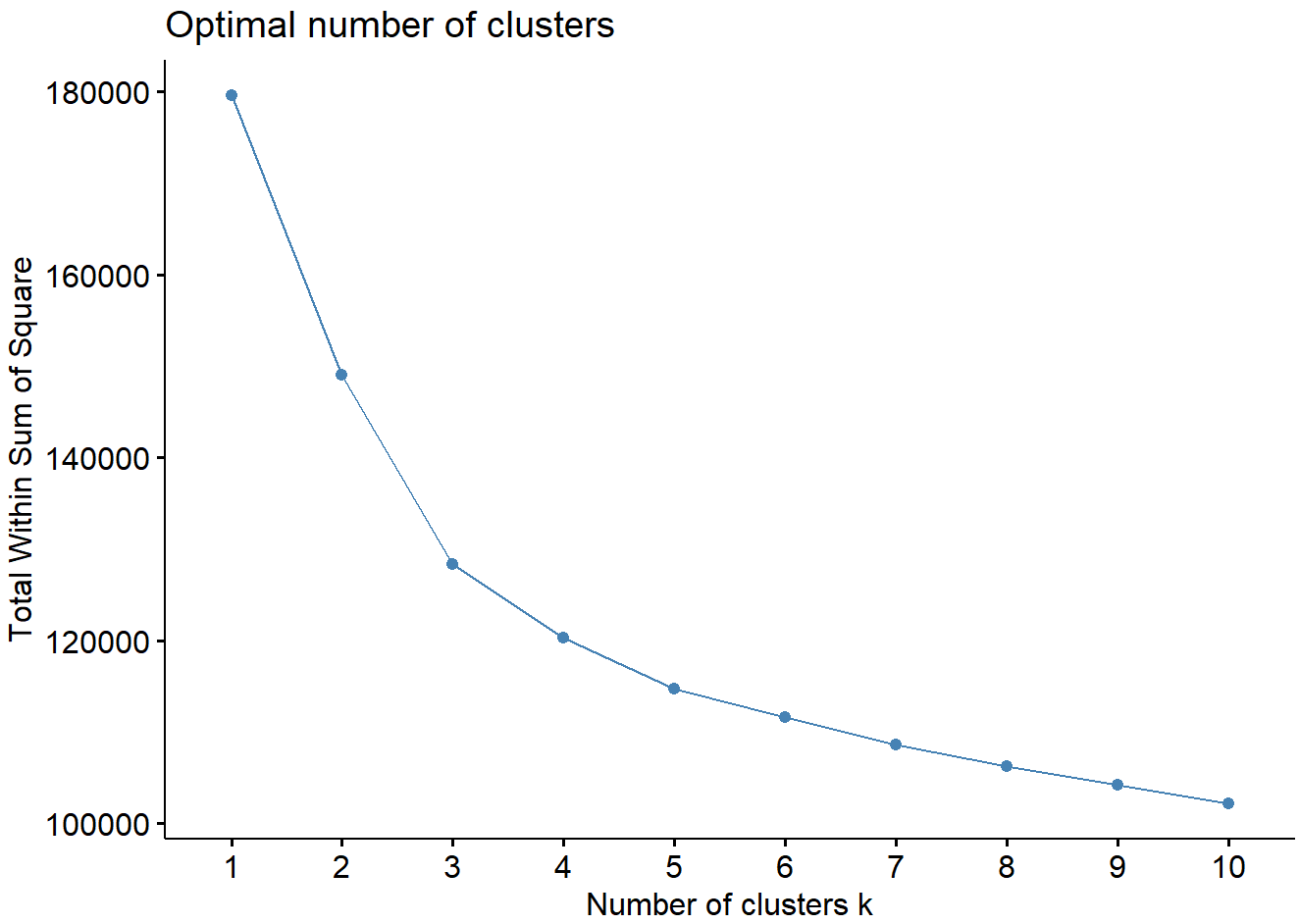
```
pca_hc
```



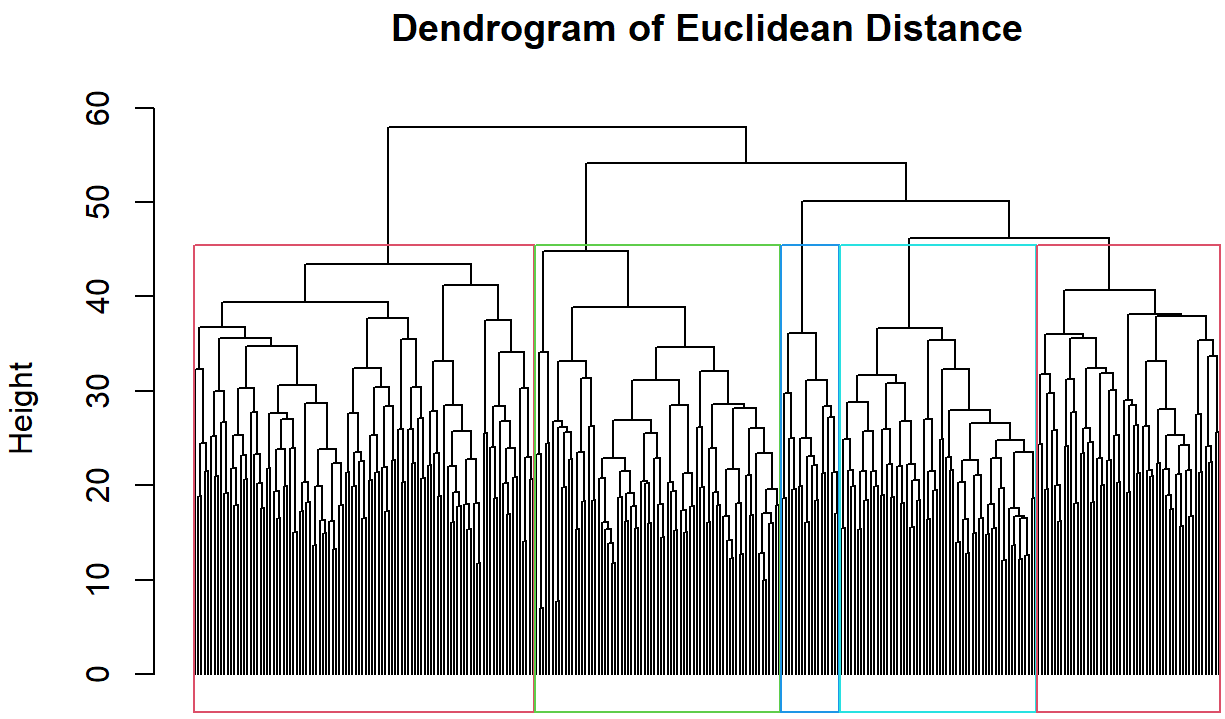
Clustring by Distance functions

In this section we will see the hierarchical clustering plot and add the clustering results for the metadata. Note that for the Kernel distance, function clustering where not created. So further analyses for this distance functions were not calculated

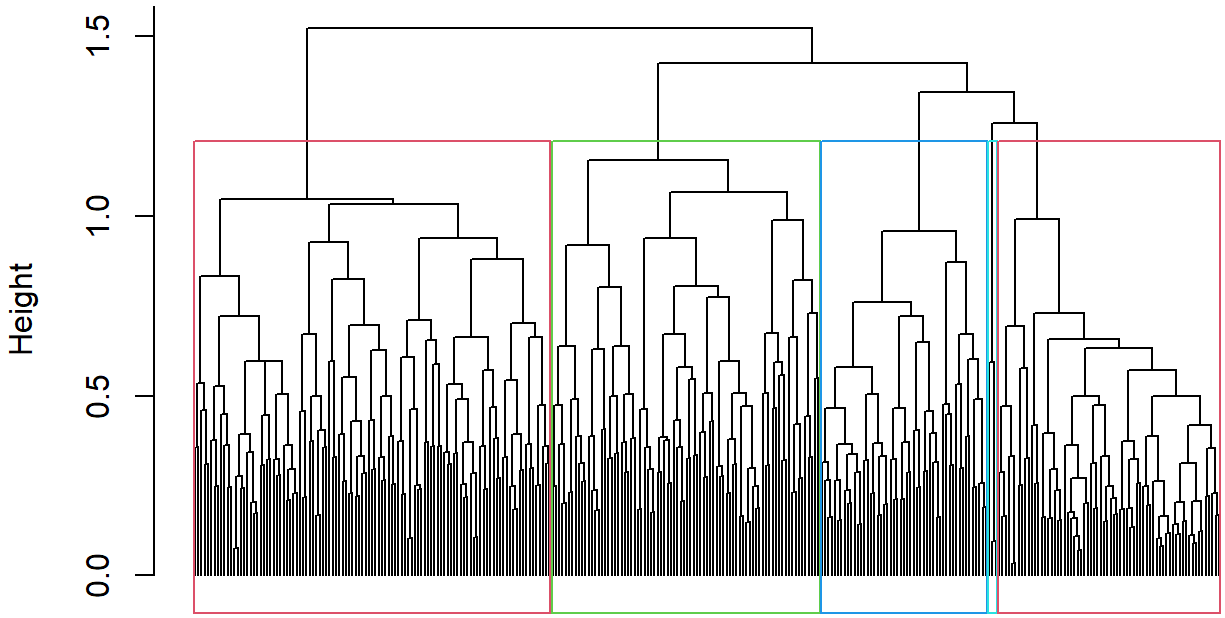
```
# Elbow Graph by Generic Clustering
fviz_nbclust(t(counts.vst.significant), FUN = hcut, method = "wss")
```



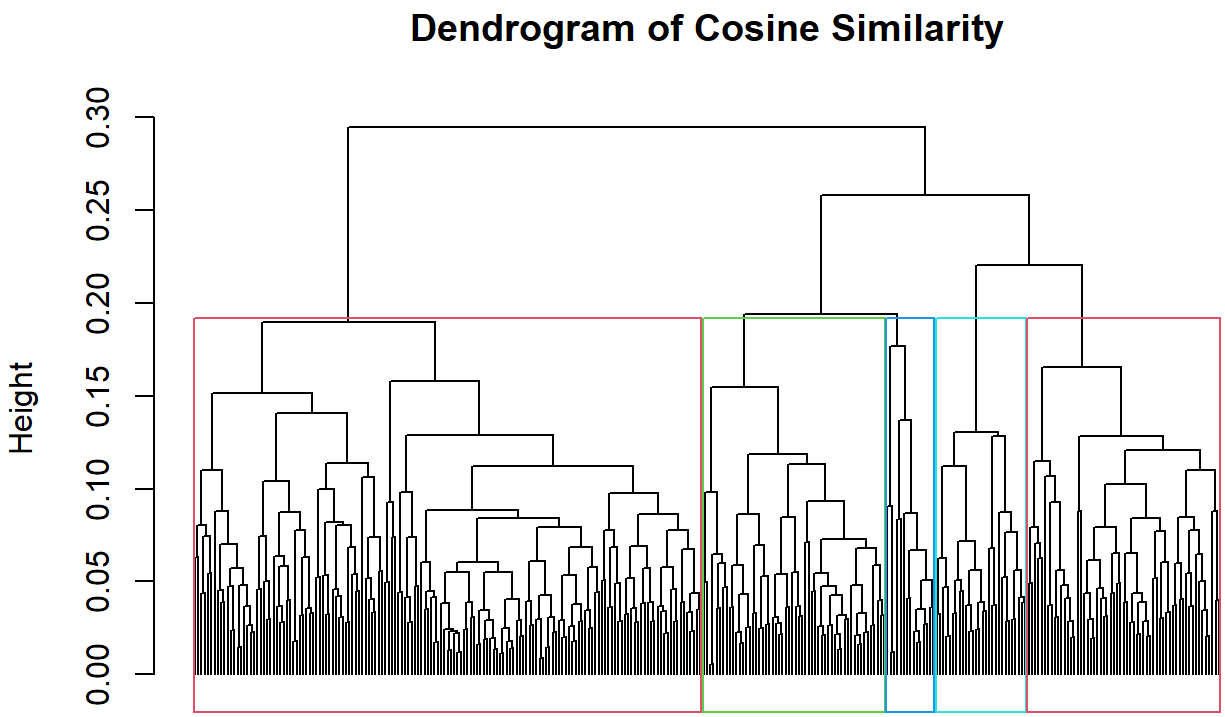

```
k <- 5

# Clustering By distance functions and add to metadata

# ---------------------------- Euclidean Distance
hclust_res_euc <- hclust(dist(t(counts.vst.significant))) # clustering
plot(hclust_res_euc, cex = 0.6, hang = -1, main = "Dendrogram of Euclidean Distance", labels=FAL
SE, xlab="", sub="") # plot
rect.hclust(hclust_res_euc, k = k, border = 2:5)
```



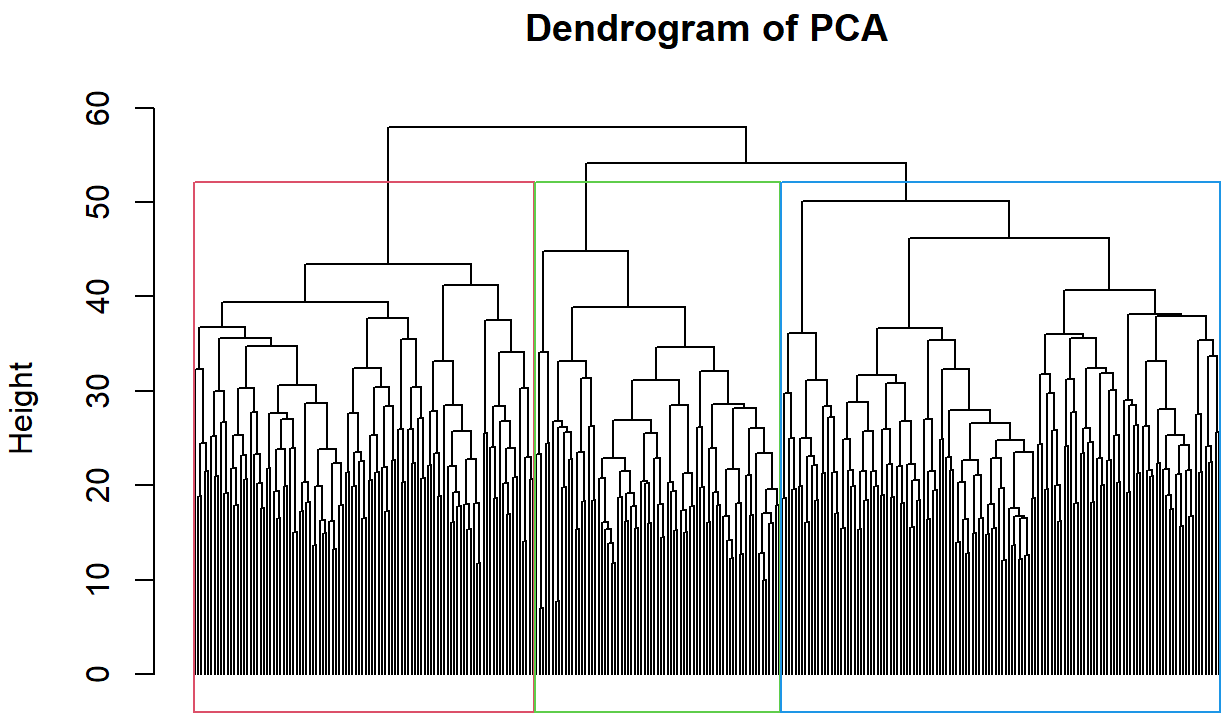

```
clusters_euc <- cutree(hclust_res_euc, k =k) # add to metadata
metadata$ClusterEuc <- clusters_euc

# ---------------------------- Pair wise Correlation
hclust_res_cor <- hclust(cor_distance(t(counts.vst.significant))) # clustering
plot(hclust_res_cor, cex = 0.6, hang = -1, main = "Dendrogram of Pair wise Correlation", labels=
FALSE, xlab="", sub="") # plot
rect.hclust(hclust_res_cor, k = k, border = 2:5)
```




```
clusters_cor <- cutree(hclust_res_cor, k =k) # add to metadata
metadata$ClusterCor <- clusters_cor

# ---------------------------- Cosine Similarity
hclust_res_cos <- hclust(t(cos_dist)) # clustering
plot(hclust_res_cos, cex = 0.6, hang = -1, main = "Dendrogram of Cosine Similarity", labels=FALS
E, xlab="", sub="") # plot
rect.hclust(hclust_res_cos, k = k, border = 2:5)
```




```
clusters_cos <- cutree(hclust_res_cos, k =k) # add to metadata
metadata$ClusterCos <- clusters_cos

# ---------------------------- Kernel Similarity
#not clustering where created.

# ---------------------------- PCA
pcaResults <- prcomp(t(counts.vst.significant))
hclust_res_pca <- hclust(dist(pcaResults$x)) # clustering
plot(hclust_res_pca, cex = 0.6, hang = -1, main = "Dendrogram of PCA", labels=FALSE, xlab="", su
b="") # plot
rect.hclust(hclust_res_pca, k = 3, border = 2:5)
```




```
clusters_pca <- cutree(hclust_res_pca, k =3) # add to metadata
metadata$ClusterPCA <- clusters_pca
```

Survival Analsys

In this section we will do survival analyses with “Kaplan-Meir” curve.

```
metadata$SurvivalStatus <- as.logical(metadata$SurvivalStatus)
# ---------------------------- PAM50
ggsurvplot(
    survfit(Surv(metadata$DaysToLastFollowUp, metadata$SurvivalStatus) ~ PAM50, data = metadat
a),
    data = metadata,
    xlab = "Months",
    xscale = 30.4,
    break.x.by = 182.4,
    surv.median.line = "hv",
    legend.title = "By PAM50",
    pval = TRUE,
    pval.coord = c(0.1, 0.825),
    xlim = c(0,4000),
    ylim=c(0.8, 1),
    ggtheme = theme_light()
)
```

```
## Warning: Removed 9 rows containing missing values or values outside the scale range
## (`geom_step()`).
```

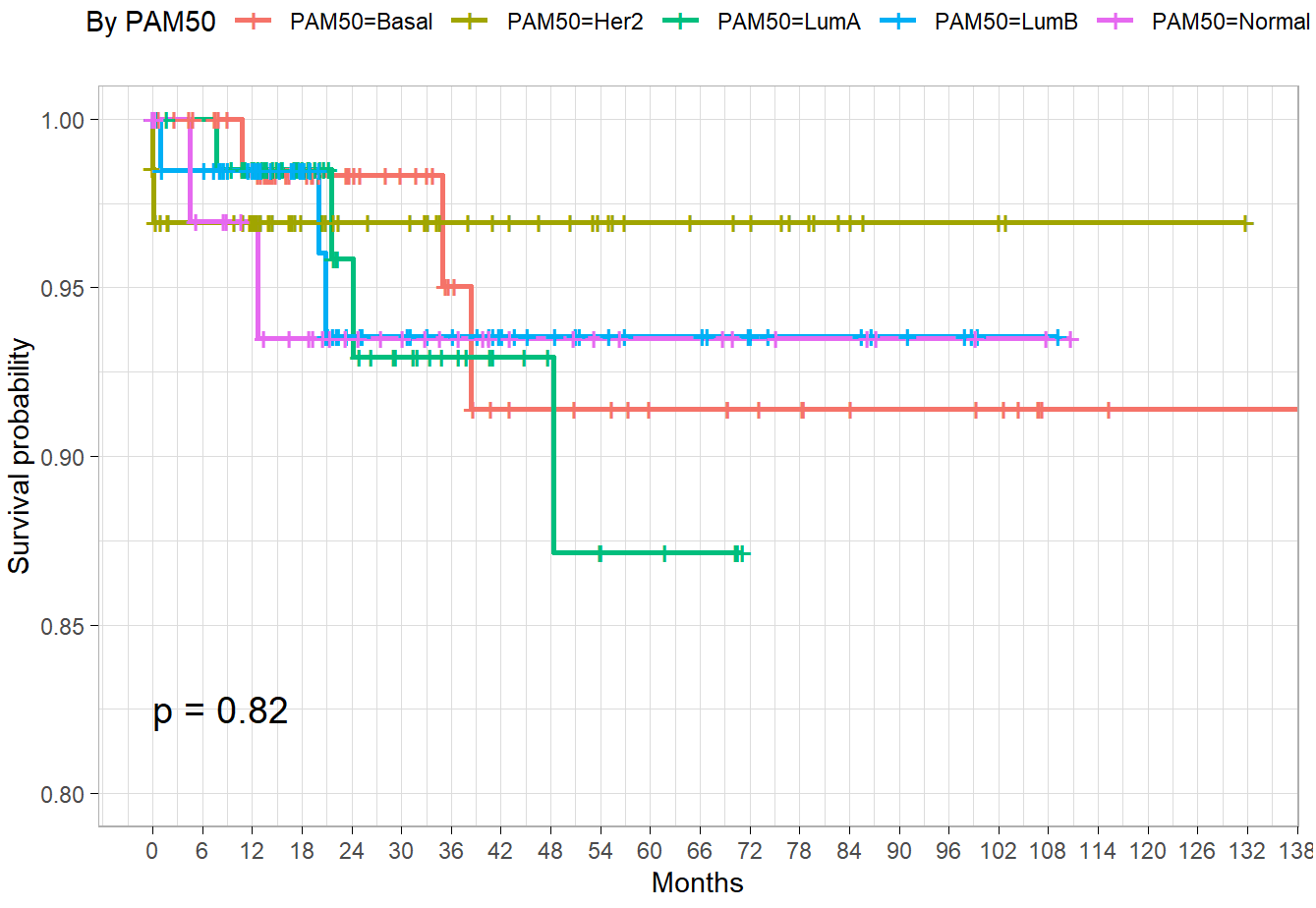
```
## Warning: Removed 7 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
## Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
```

```
## Warning: Removed 9 rows containing missing values or values outside the scale range
## (`geom_step()`).
```

```
## Warning: Removed 7 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
## Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
```

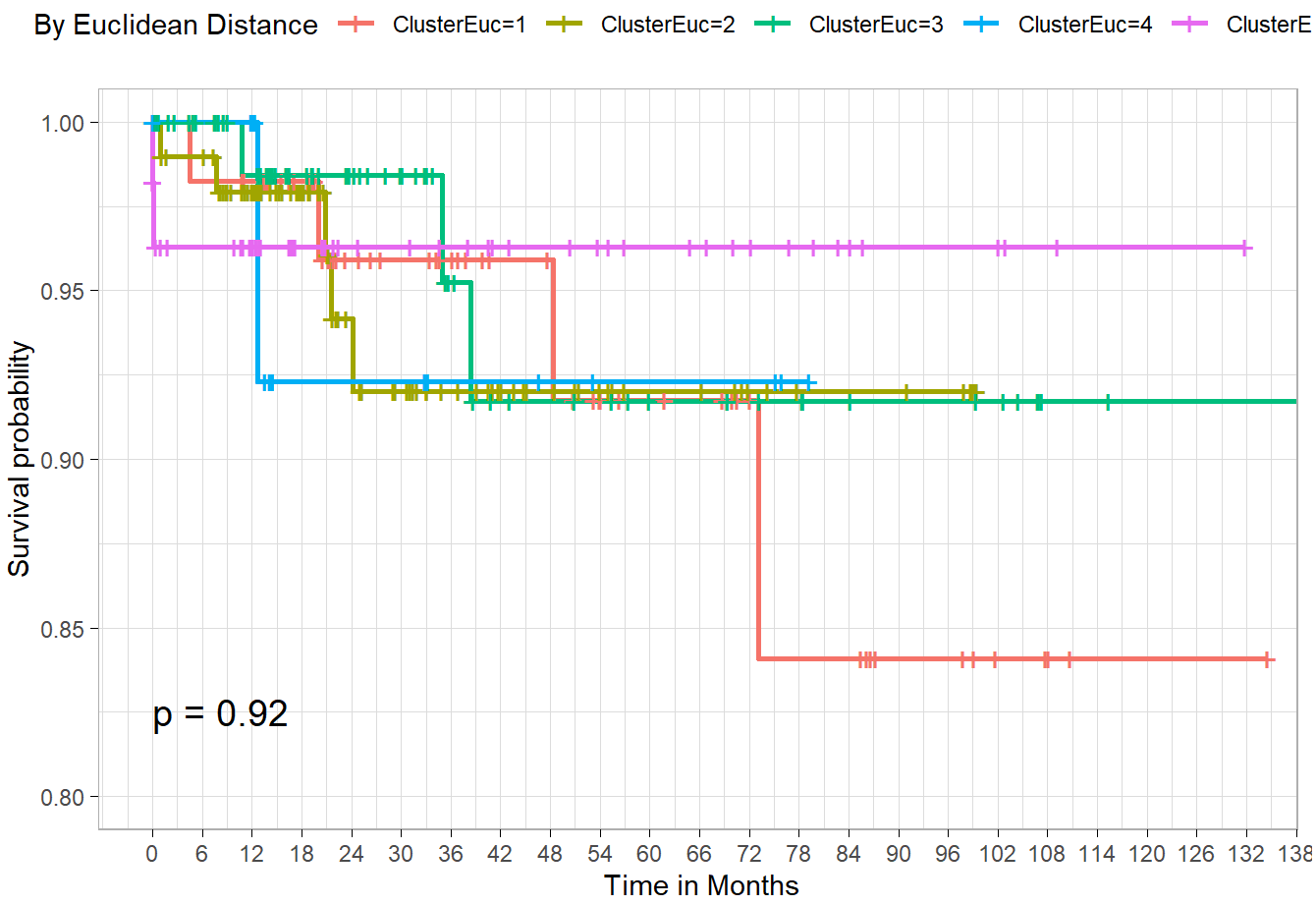


```
# ---------------------------- Euclidean Distance
ggsurvplot(
    survfit(Surv(metadata$DaysToLastFollowUp, metadata$SurvivalStatus) ~ ClusterEuc, data = meta
data),
    data = metadata,
    xlab = " Time in Months",
    xscale = 30.4,
    xlim = c(0,4000),
    ylim=c(0.8, 1),
    break.x.by = 182.4,
    pval = TRUE,
    pval.coord = c(0.1, 0.825),
    surv.median.line = "hv",
    legend.title = "By Euclidean Distance",
    ggtheme = theme_light()
)
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_step()`).
## Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
## Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_step()`).
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
## Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
```

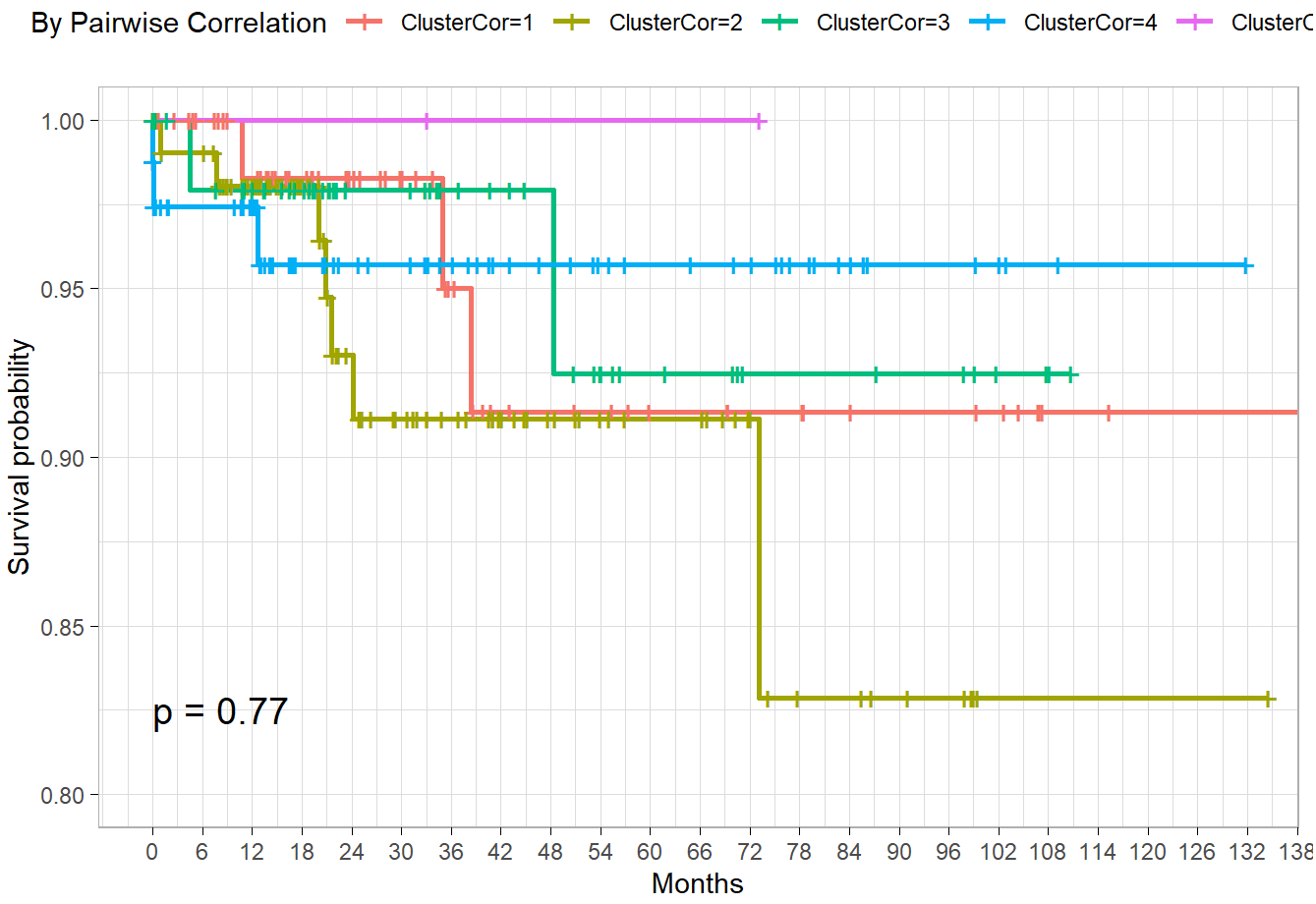



```
# ---------------------------- Pairwise Correlation
ggsurvplot(
    survfit(Surv(metadata$DaysToLastFollowUp, metadata$SurvivalStatus) ~ ClusterCor, data = meta
data),
    data = metadata,
    xlab = "Months",
    xscale = 30.4,
    break.x.by = 182.4,
    surv.median.line = "hv",
    legend.title = "By Pairwise Correlation",
    xlim = c(0,4000),
    ylim=c(0.8, 1),
    pval = TRUE,
    pval.coord = c(0.1, 0.825),
    ggtheme = theme_light()
)
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_step()`).
## Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
## Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_step()`).
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
## Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
```

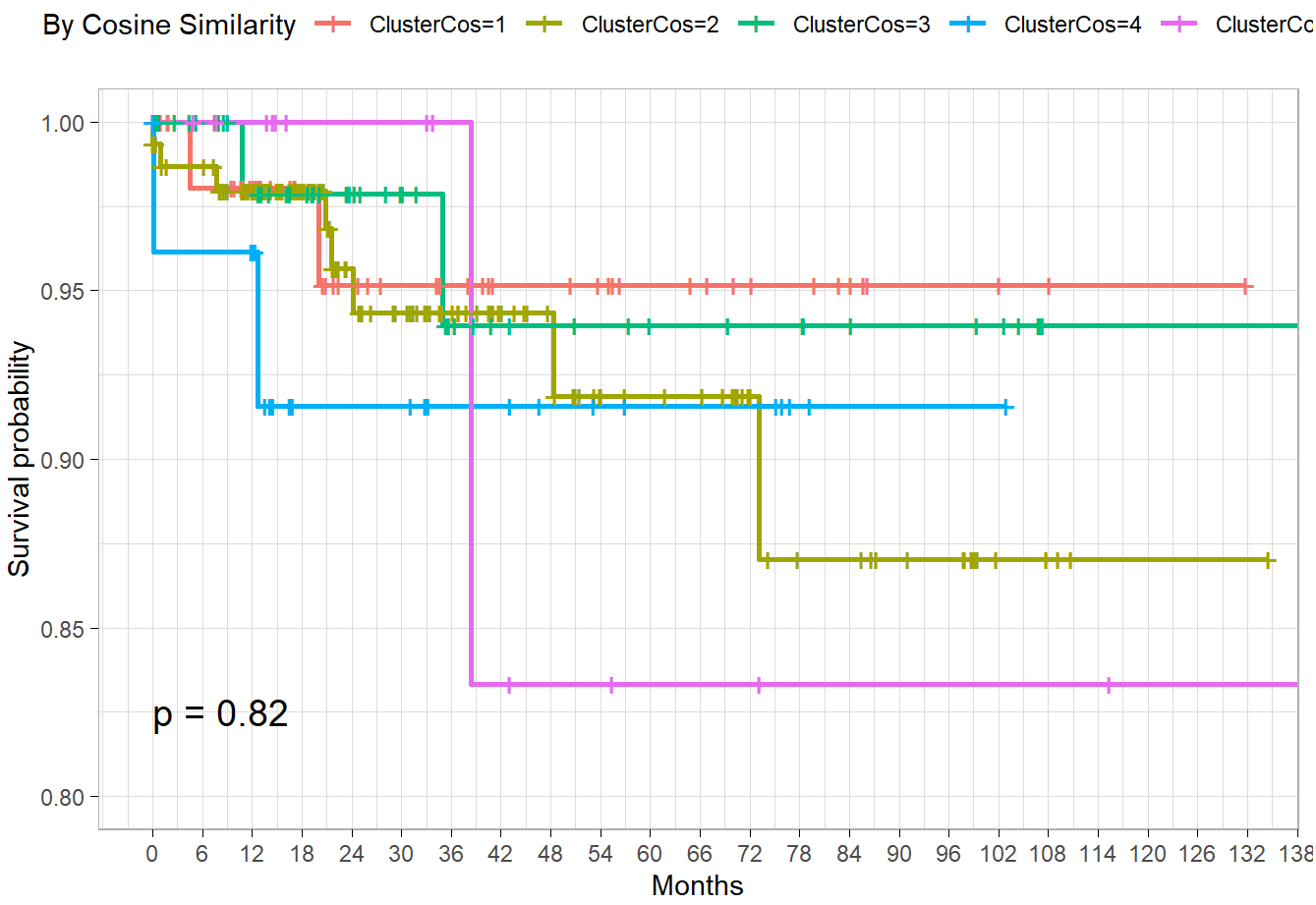


```
# ----------------------------  Cosine Similarity
ggsurvplot(
    survfit(Surv(metadata$DaysToLastFollowUp, metadata$SurvivalStatus) ~ ClusterCos, data = meta
data),
    data = metadata,
    xlab = "Months",
    xscale = 30.4,
    break.x.by = 182.4,
    surv.median.line = "hv",
    legend.title = "By Cosine Similarity",
    xlim = c(0,4000),
    ylim=c(0.8, 1),
    pval = TRUE,
    pval.coord = c(0.1, 0.825),
    ggtheme = theme_light()
)
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_step()`).
## Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
## Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_step()`).
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
## Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
```

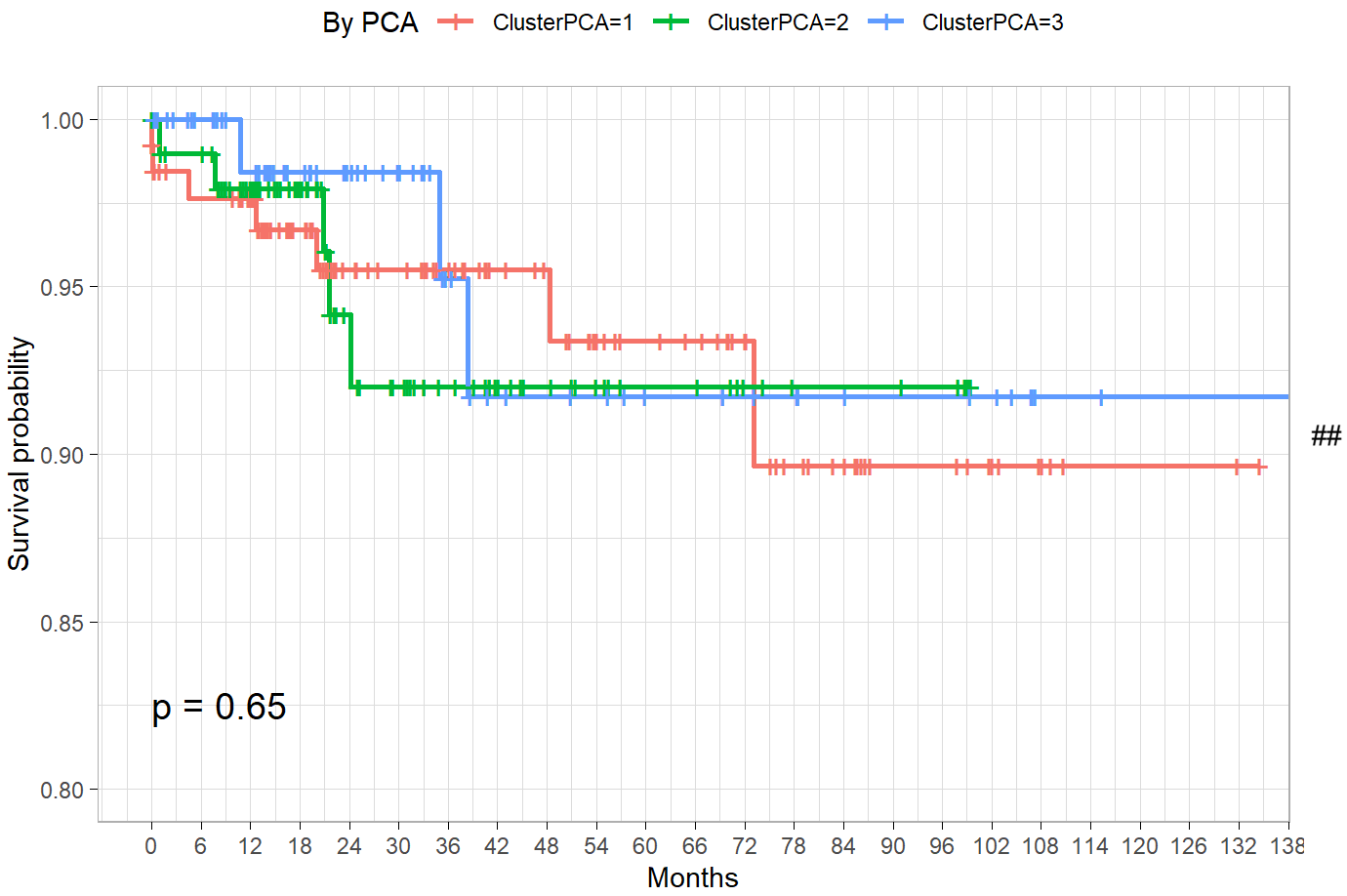



```
# ---------------------------- PCA
ggsurvplot(
    survfit(Surv(metadata$DaysToLastFollowUp, metadata$SurvivalStatus) ~ ClusterPCA, data = meta
data),
    data = metadata,
    xlab = "Months",
    xscale = 30.4,
    break.x.by = 182.4,
    surv.median.line = "hv",
    legend.title = "By PCA",
    xlim = c(0,4000),
    ylim=c(0.8, 1),
    pval = TRUE,
    pval.coord = c(0.1, 0.825),
    ggtheme = theme_light()
)
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_step()`).
## Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
## Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_step()`).
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
## Removed 1 row containing missing values or values outside the scale range
## (`geom_segment()`).
```



Enrichment Analysis This section divided to 2 parts. The first part is GSEA on the counts.vst results, where the results are analyses by cluster. In this section no significant pathways for any cluster of the distance functions. Because the similarity between the results, we only presents the euclidean distance and the PCA.

```
# prepare hallmarks
hallmarks <- msigdbr(species = "Homo sapiens", category = "H") %>%
  dplyr::select(gs_name, gene_symbol)
hallmarks <- msigdbr(species = "Homo sapiens", category = "H")
hallmarks <- hallmarks[,c('gs_name', 'gene_symbol')]

# ---------------------------- Euclidean Distance

results_list <- list()
for (i in unique(metadata$ClusterEuc)) {
  cluster_samples <- which(metadata$ClusterEuc == i)
  gene_list <- apply(counts.vst.significant[, cluster_samples], 1, var)
  names(gene_list) <- rownames(counts.vst.significant)
  gene_list <- sort(gene_list, decreasing = TRUE)

  gsea_results <- GSEA(gene_list, TERM2GENE = hallmarks, verbose = TRUE, scoreType = "pos")
  results_list[[paste("Cluster", i)]] <- gsea_results
}
```

```
## preparing geneSet collections...
```

```
## GSEA analysis...
```

```
## no term enriched under specific pvalueCutoff...
```

```
## preparing geneSet collections...
```

```
## GSEA analysis...
```

```
## no term enriched under specific pvalueCutoff...
```

```
## preparing geneSet collections...
```

```
## GSEA analysis...
```

```
## no term enriched under specific pvalueCutoff...
```

```
## preparing geneSet collections...
```

```
## GSEA analysis...
```

```
## no term enriched under specific pvalueCutoff...
```

```
## preparing geneSet collections...
```

```
## GSEA analysis...
```

```
## no term enriched under specific pvalueCutoff...
```

```
# Assuming the GSEA results are stored in 'results_list'
for (i in names(results_list)) {
  gsea_results <- results_list[[i]]

  # Check if the results are empty
  if (length(gsea_results) == 0 || nrow(gsea_results@result) == 0) {
    print(paste("No significant pathways for Euclidean", i))
  } else {
    print(paste("Results for", i))
    dotplot(gsea_results)
  }
}
```

```
## [1] "No significant pathways for Euclidean Cluster 1"
## [1] "No significant pathways for Euclidean Cluster 2"
## [1] "No significant pathways for Euclidean Cluster 3"
## [1] "No significant pathways for Euclidean Cluster 4"
## [1] "No significant pathways for Euclidean Cluster 5"
```

```
# ----------------------------  PCA

results_list <- list()
for (i in unique(metadata$ClusterPCA)) {
  cluster_samples <- which(metadata$ClusterPCA == i)
  gene_list <- apply(counts.vst.significant[, cluster_samples], 1, var)
  names(gene_list) <- rownames(counts.vst.significant)
  gene_list <- sort(gene_list, decreasing = TRUE)

  gsea_results <- GSEA(gene_list, TERM2GENE = hallmarks, verbose = TRUE, scoreType = "pos")
  results_list[[paste("Cluster", i)]] <- gsea_results
}
```

```
## preparing geneSet collections...
```

```
## GSEA analysis...
```

```
## no term enriched under specific pvalueCutoff...
```

```
## preparing geneSet collections...
```

```
## GSEA analysis...
```

```
## no term enriched under specific pvalueCutoff...
```

```
## preparing geneSet collections...
```

```
## GSEA analysis...
```

```
## no term enriched under specific pvalueCutoff...
```

```
# Assuming the GSEA results are stored in 'results_list'
for (i in names(results_list)) {
  gsea_results <- results_list[[i]]

  # Check if the results are empty
  if (length(gsea_results) == 0 || nrow(gsea_results@result) == 0) {
    print(paste("No significant pathways for PCA Cluster", i))
  } else {
    print(paste("Results for", i))
    dotplot(gsea_results)
  }
}
```

```
## [1] "No significant pathways for PCA Cluster Cluster 1"
## [1] "No significant pathways for PCA Cluster Cluster 2"
## [1] "No significant pathways for PCA Cluster Cluster 3"
```

The second part we repeated the DESeq analyses design by each distance function Clustering and did the GSEA by the new dds results.

```
# ---------------------------- Euclidean Distance

metadata$ClusterEuc <- as.factor(metadata$ClusterEuc)
euclidean_dds <- DESeqDataSetFromMatrix(countData=counts,
                              colData=metadata,
                              design= ~ ClusterEuc)

euclidean_dds <- DESeq(euclidean_dds)
```

```
## estimating size factors
```

```
## estimating dispersions
```

```
## gene-wise dispersion estimates
```

```
## mean-dispersion relationship
```

```
## final dispersion estimates
```

```
## fitting model and testing
```

```
## -- replacing outliers and refitting for 9868 genes
## -- DESeq argument 'minReplicatesForReplace' = 7
## -- original counts are preserved in counts(dds)
```

```
## estimating dispersions
```

```
## fitting model and testing
```

```
euc.counts.vst <- vst(euclidean_dds)
euc.counts.vst <- assay(euc.counts.vst)

rownames(euc.counts.vst) <- gene_ids_to_names$gene_name[gene_ids_to_names$gene_id %in% rownames
(euc.counts.vst)]
var_per_gene <- apply(euc.counts.vst, 1, var)  # Calculate the variance per gene
var_per_gene <- var_per_gene[order(var_per_gene, decreasing = T)]
selectedGenes <- names(var_per_gene[order(var_per_gene, decreasing = T)])
selectedGenes <- gene_ids_to_names$gene_name[gene_ids_to_names$gene_id %in% selectedGenes]

hm <- GSEA(var_per_gene, TERM2GENE = hallmarks, pvalueCutoff = 0.05, eps = 0)
```

```
## preparing geneSet collections...
```

```
## GSEA analysis...
```

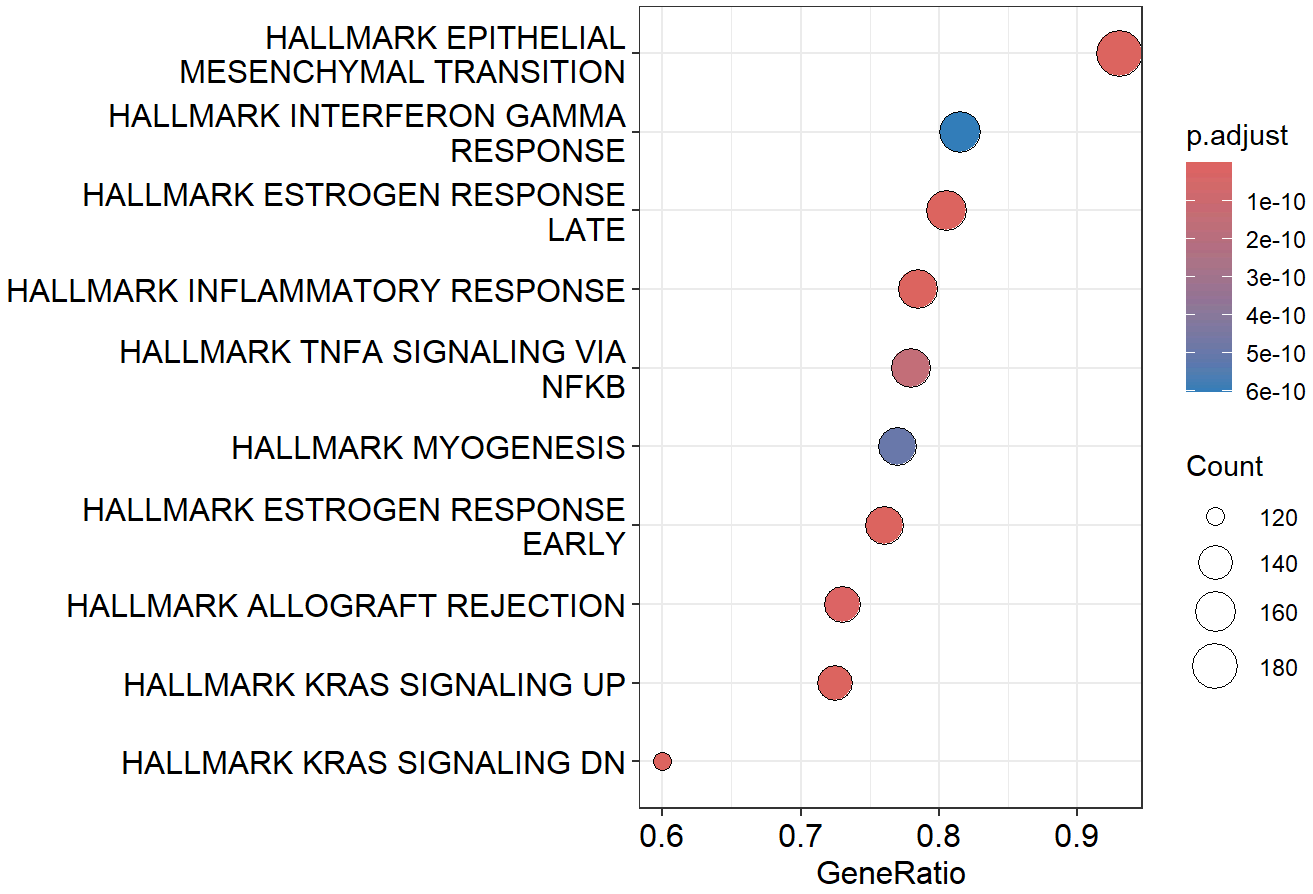
```
## Warning in preparePathwaysAndStats(pathways, stats, minSize, maxSize, gseaParam, : There are
ties in the preranked stats (1.02% of the list).
## The order of those tied genes will be arbitrary, which may produce unexpected results.
```

```
## Warning in preparePathwaysAndStats(pathways, stats, minSize, maxSize,
## gseaParam, : There are duplicate gene names, fgsea may produce unexpected
## results.
```

```
## leading edge analysis...
```

```
## done...
```

```
dotplot(hm)
```




```
# ---------------------------- Pairwise Correlation
metadata$ClusterCor <- as.factor(metadata$ClusterCor)
correlation_dds <- DESeqDataSetFromMatrix(countData=counts,
                              colData=metadata,
                              design= ~ ClusterCor)
correlation_dds <- DESeq(correlation_dds)
```

```
## estimating size factors
```

```
## estimating dispersions
```

```
## gene-wise dispersion estimates
```

```
## mean-dispersion relationship
```

```
## final dispersion estimates
```

```
## fitting model and testing
```

```
## -- replacing outliers and refitting for 4864 genes
## -- DESeq argument 'minReplicatesForReplace' = 7
## -- original counts are preserved in counts(dds)
```

```
## estimating dispersions
```

```
## fitting model and testing
```

```
cor.counts.vst <- vst(correlation_dds)
cor.counts.vst <- assay(cor.counts.vst)

rownames(cor.counts.vst) <- gene_ids_to_names$gene_name[gene_ids_to_names$gene_id %in% rownames
(cor.counts.vst)]
var_per_gene <- apply(cor.counts.vst, 1, var)  # Calculate the variance per gene
var_per_gene <- var_per_gene[order(var_per_gene, decreasing = T)]
selectedGenes <- names(var_per_gene[order(var_per_gene, decreasing = T)])
selectedGenes <- gene_ids_to_names$gene_name[gene_ids_to_names$gene_id %in% selectedGenes]

hallmarks <- msigdbr(species = "Homo sapiens", category = "H")
hallmarks <- hallmarks[,c('gs_name', 'gene_symbol')]
hm <- GSEA(var_per_gene, TERM2GENE = hallmarks, pvalueCutoff = 0.05, eps = 0)
```

```
## preparing geneSet collections...
```

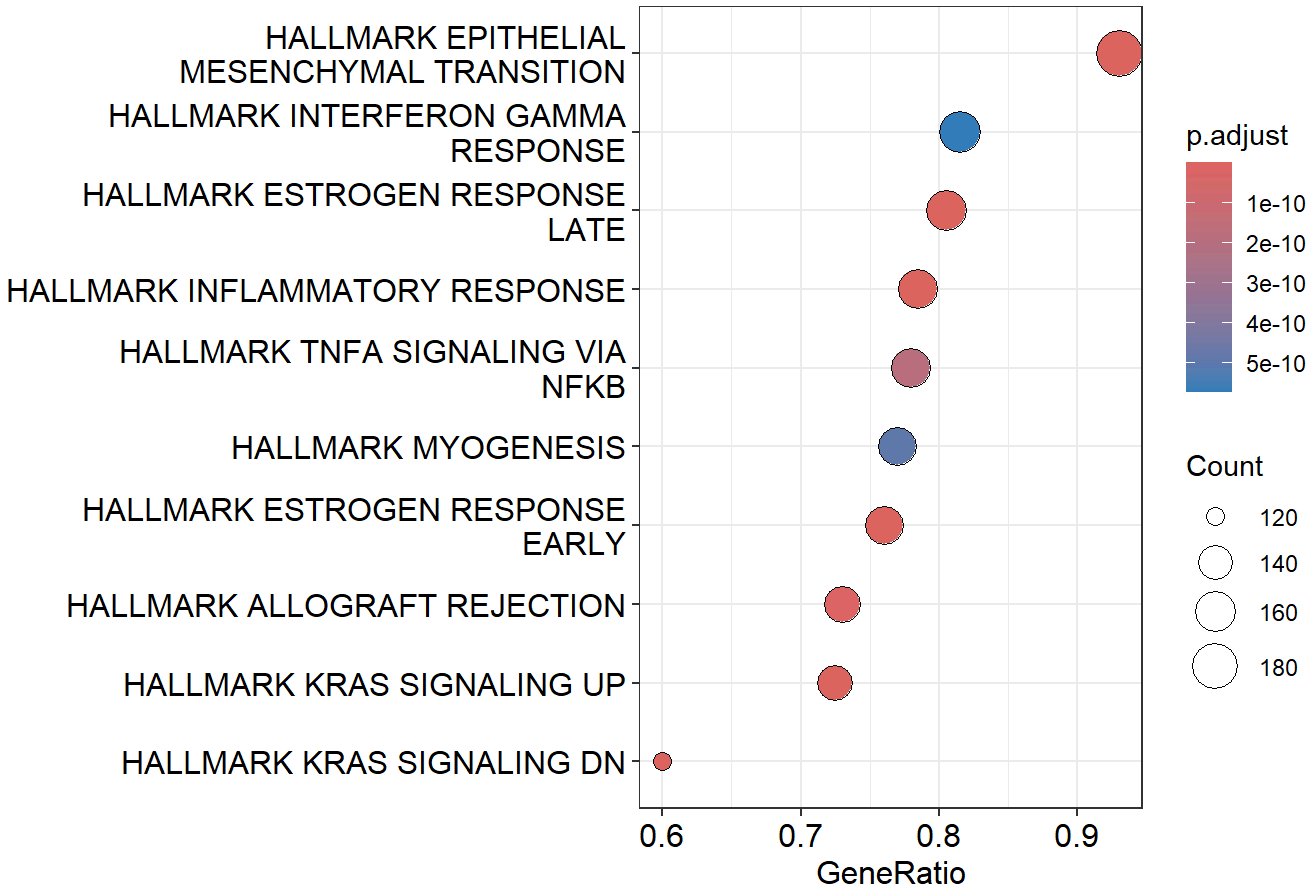
```
## GSEA analysis...
```

```
## Warning in preparePathwaysAndStats(pathways, stats, minSize, maxSize, gseaParam, : There are
ties in the preranked stats (1.02% of the list).
## The order of those tied genes will be arbitrary, which may produce unexpected results.
## Warning in preparePathwaysAndStats(pathways, stats, minSize, maxSize, gseaParam, : There are
duplicate gene names, fgsea may produce unexpected results.
```

```
## leading edge analysis...
```

```
## done...
```

```
dotplot(hm)
```




```
# ----------------------------  Cosine Similarity
metadata$ClusterCos <- as.factor(metadata$ClusterCos)
cosine_dds <- DESeqDataSetFromMatrix(countData=counts,
                              colData=metadata,
                              design= ~ ClusterCos)
cosine_dds <- DESeq(cosine_dds)
```

```
## estimating size factors
```

```
## estimating dispersions
```

```
## gene-wise dispersion estimates
```

```
## mean-dispersion relationship
```

```
## final dispersion estimates
```

```
## fitting model and testing
```

```
## -- replacing outliers and refitting for 9778 genes
## -- DESeq argument 'minReplicatesForReplace' = 7
## -- original counts are preserved in counts(dds)
```

```
## estimating dispersions
```

```
## fitting model and testing
```

```
cos.counts.vst <- vst(cosine_dds)
cos.counts.vst <- assay(cos.counts.vst)

rownames(cos.counts.vst) <- gene_ids_to_names$gene_name[gene_ids_to_names$gene_id %in% rownames
(cos.counts.vst)]
var_per_gene <- apply(cos.counts.vst, 1, var)  # Calculate the variance per gene
var_per_gene <- var_per_gene[order(var_per_gene, decreasing = T)]
selectedGenes <- names(var_per_gene[order(var_per_gene, decreasing = T)])
selectedGenes <- gene_ids_to_names$gene_name[gene_ids_to_names$gene_id %in% selectedGenes]

hallmarks <- msigdbr(species = "Homo sapiens", category = "H")
hallmarks <- hallmarks[,c('gs_name', 'gene_symbol')]
hm <- GSEA(var_per_gene, TERM2GENE = hallmarks, pvalueCutoff = 0.05, eps = 0)
```

```
## preparing geneSet collections...
```

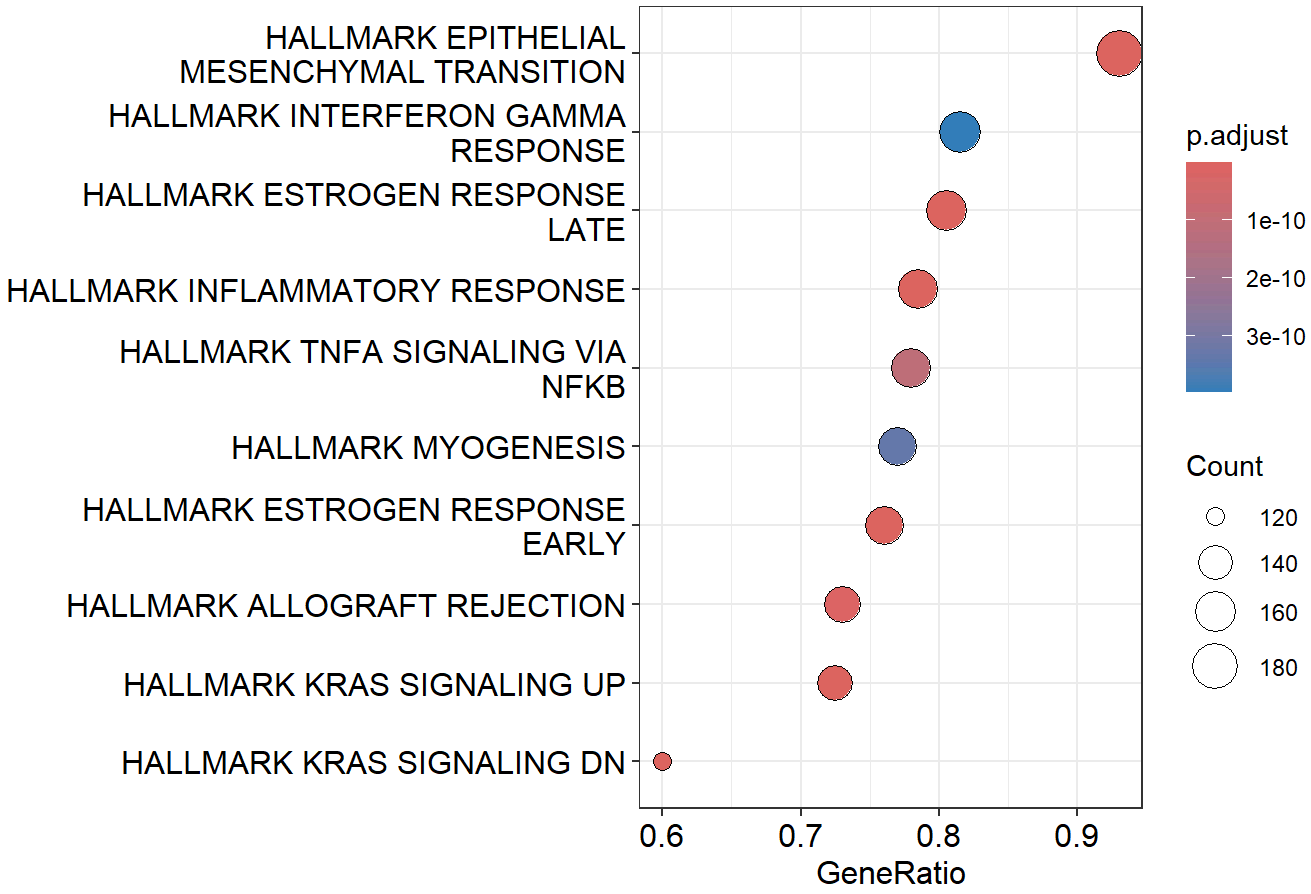
```
## GSEA analysis...
```

```
## Warning in preparePathwaysAndStats(pathways, stats, minSize, maxSize, gseaParam, : There are
ties in the preranked stats (1.02% of the list).
## The order of those tied genes will be arbitrary, which may produce unexpected results.
## Warning in preparePathwaysAndStats(pathways, stats, minSize, maxSize, gseaParam, : There are
duplicate gene names, fgsea may produce unexpected results.
```

```
## leading edge analysis...
```

```
## done...
```

```
dotplot(hm)
```




```
# ----------------------------  PCA
metadata$ClusterPCA <- as.factor(metadata$ClusterPCA)
pca_dds <- DESeqDataSetFromMatrix(countData=counts,
                              colData=metadata,
                              design= ~ ClusterPCA)
pca_dds <- DESeq(pca_dds)
```

```
## estimating size factors
```

```
## estimating dispersions
```

```
## gene-wise dispersion estimates
```

```
## mean-dispersion relationship
```

```
## final dispersion estimates
```

```
## fitting model and testing
```

```
## -- replacing outliers and refitting for 12224 genes
## -- DESeq argument 'minReplicatesForReplace' = 7
## -- original counts are preserved in counts(dds)
```

```
## estimating dispersions
```

```
## fitting model and testing
```

```
pca.counts.vst <- vst(pca_dds)
pca.counts.vst <- assay(pca.counts.vst)

rownames(pca.counts.vst) <- gene_ids_to_names$gene_name[gene_ids_to_names$gene_id %in% rownames
(pca.counts.vst)]
var_per_gene <- apply(pca.counts.vst, 1, var)  # Calculate the variance per gene
var_per_gene <- var_per_gene[order(var_per_gene, decreasing = T)]
selectedGenes <- names(var_per_gene[order(var_per_gene, decreasing = T)])
selectedGenes <- gene_ids_to_names$gene_name[gene_ids_to_names$gene_id %in% selectedGenes]

hallmarks <- msigdbr(species = "Homo sapiens", category = "H")
hallmarks <- hallmarks[,c('gs_name', 'gene_symbol')]
hm <- GSEA(var_per_gene, TERM2GENE = hallmarks, pvalueCutoff = 0.05, eps = 0)
```

```
## preparing geneSet collections...
```

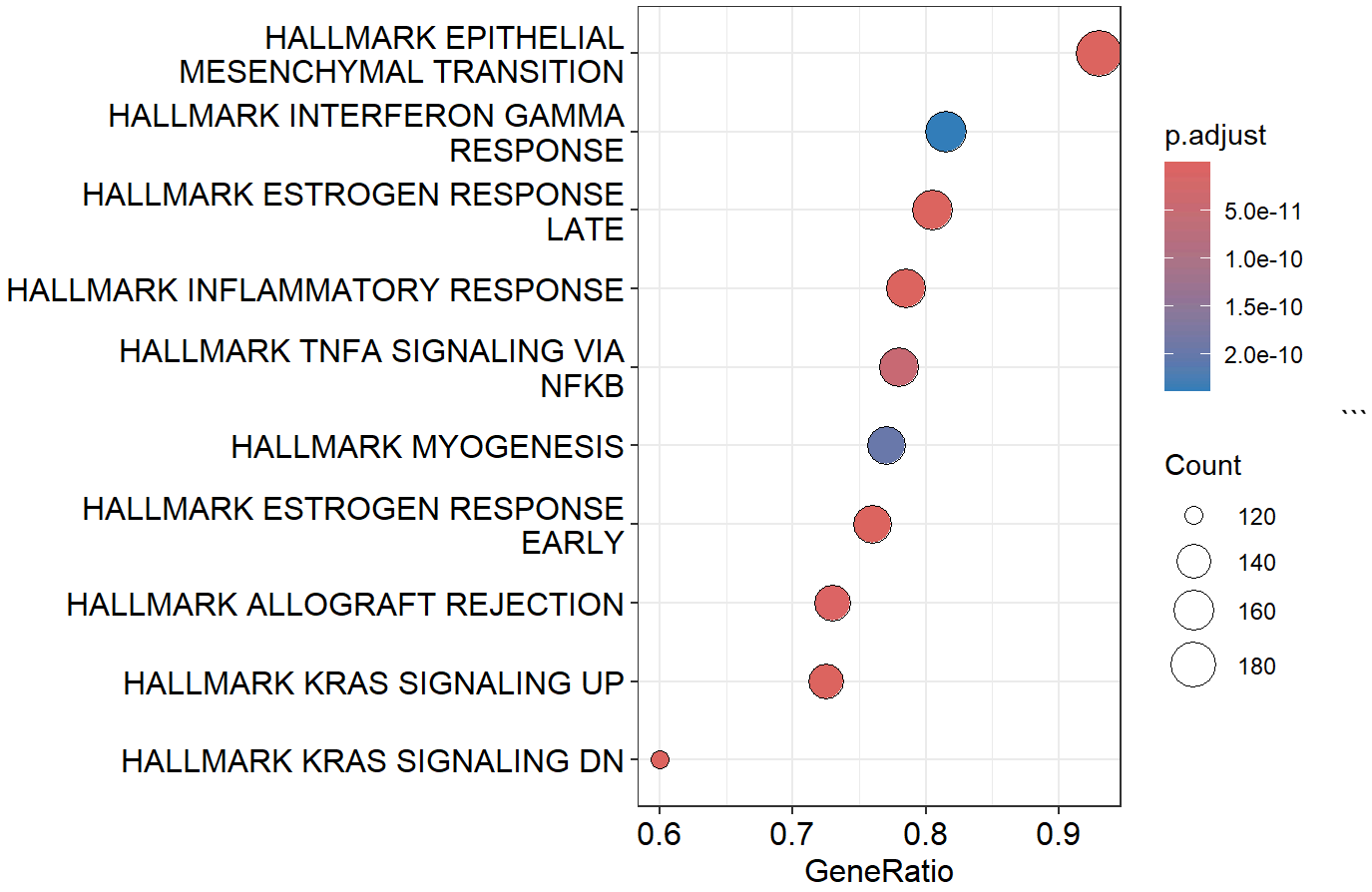
```
## GSEA analysis...
```

```
## Warning in preparePathwaysAndStats(pathways, stats, minSize, maxSize, gseaParam, : There are
ties in the preranked stats (1.02% of the list).
## The order of those tied genes will be arbitrary, which may produce unexpected results.
## Warning in preparePathwaysAndStats(pathways, stats, minSize, maxSize, gseaParam, : There are
duplicate gene names, fgsea may produce unexpected results.
```

```
## leading edge analysis...
```

```
## done...
```

```
dotplot(hm)
```


HALLMARK EPITHELIAL MESENCHYMAL TRANSITION
HALLMARK INTERFERON GAMMA RESPONSE
HALLMARK ESTROGEN RESPONSE LATE
HALLMARK INFLAMMATORY RESPONSE
HALLMARK TNFA SIGNALING VIA NFKB
HALLMARK MYOGENESIS
HALLMARK ESTROGEN RESPONSE EARLY
HALLMARK ALLOGRAFT REJECTION
HALLMARK KRAS SIGNALING UP
HALLMARK KRAS SIGNALING DN
0.6
0.7
0.8
0.9
GeneRatio
p.adjust
5.0e-11
1.0e-10
1.5e-10
2.0e-10
Count
120
140
160
180

```
```