Comparative functional enrichment analysis between cell lines using only proteins measured in all three cell lines

Konstantin Kahnert

2025-05-13

Load libraries

```r
library(tidyverse)
library(clusterProfiler)
library(enrichplot)
library(org.Hs.eg.db)
library(Cairo)
library(svglite)
library(VennDiagram)
library(grid)
library(simplifyEnrichment)
```

Define function

```r
# map_cluster_number:
#   - x: enrichResult or compareClusterResult (from clusterProfiler)
#   - df: simplifyGO result data.frame with columns ID (term IDs) and Cluster_num
#   - comp: if TRUE, use x@compareClusterResult, else x@result
# Returns a data.frame of all enrich columns + term_size + Cluster_num
map_cluster_number <- function(x,
                               df,
                               comp = FALSE) {
  ## 1. Standardize the input DF's column names
  colnames(df) <- c("ID", "Cluster_num")

  ## 2. Select the correct slot from the clusterProfiler object
  if (comp) {
    # from compareClusterResult
    results <- x@compareClusterResult
  } else {
    # from a single enrichResult
    results <- x@result
  }

  ## 3. Get term_size (the numerator of the "BgRatio" string)
  ##    e.g. "12/21273" → 12
  results$term_size <- as.numeric(
    sapply(strsplit(results$BgRatio, "/"),
           function(parts) as.numeric(parts[1]))
  )

  ## 4. Merge the enrichment results with the cluster assignments
  merged_results <- merge(
    results,
    df,
    by    = "ID",
    all.x = TRUE
  )

  ## 5. Return the processed data frame
  return(merged_results)
}


# Function to map Ensembl IDs to gene symbols for a single row
map_geneID_to_symbol <- function(geneID_str) {
  geneIDs <- unlist(strsplit(geneID_str, "/"))
  mapped_ids <- bitr(geneIDs,
                     fromType = "ENSEMBL",
                     toType = "SYMBOL",
                     OrgDb = org.Hs.eg.db)

  # drop duplicates in ensembl column and keep first occurence
  mapped_ids <- mapped_ids[!duplicated(mapped_ids$ENSEMBL),]

  gene_symbols <- mapped_ids$SYMBOL
  return(paste(gene_symbols, collapse = "/"))
}


# Function to map Ensembl IDs to gene symbols for a single row
map_geneID_to_name <- function(geneIDs) {
  mapped_ids <- bitr(geneIDs,
                     fromType = "ENSEMBL",
                     toType = c("GENENAME", "SYMBOL"),
                     OrgDb = org.Hs.eg.db)

  # drop duplicates in ensembl column and keep first occurence
  mapped_ids <- mapped_ids[!duplicated(mapped_ids$ENSEMBL),]

  return(mapped_ids)
}
```

Set input and output paths

```
# set input and output paths
in_path <- "/mnt/Data/Projects/Cilia/revision/Restricted/data/"
out_path <- "/mnt/Data/Projects/Cilia/revision/Restricted/analysis/GO_BP/"
```

Load data

```
# Load the data
df_all <- read.delim(paste0(in_path, "Restricted_files_combined_as_cytoscape_input.csv"), sep = "\t", he
head(df_all)
```

```
##                   BB_ BB_ASC52telo BB_hTERT_RPE1_serum_starved BB_RPTEC_TERT1
## ENSG00000089289  True        False                        True           True
## ENSG00000266826  True        False                        True           True
## ENSG00000101004 False        False                       False          False
## ENSG00000102218 False        False                       False          False
## ENSG00000198553  True        False                        True          False
## ENSG00000112144 False        False                       False          False
##                 PrimaryCilia_ PrimaryCilia_ASC52telo
## ENSG00000089289          True                   True
## ENSG00000266826          True                   True
## ENSG00000101004          True                  False
## ENSG00000102218         False                  False
## ENSG00000198553         False                  False
## ENSG00000112144          True                   True
##                 PrimaryCilia_hTERT_RPE1_serum_starved PrimaryCilia_RPTEC_TERT1
## ENSG00000089289                                  True                     True
## ENSG00000266826                                  True                     True
## ENSG00000101004                                  True                     True
## ENSG00000102218                                 False                    False
## ENSG00000198553                                 False                    False
## ENSG00000112144                                  True                     True
##                 PrimaryCiliaTip_ PrimaryCiliaTip_ASC52telo
## ENSG00000089289            False                     False
## ENSG00000266826            False                     False
## ENSG00000101004            False                     False
## ENSG00000102218            False                     False
## ENSG00000198553            False                     False
## ENSG00000112144             True                      True
##                 PrimaryCiliaTip_hTERT_RPE1_serum_starved
## ENSG00000089289                                    False
## ENSG00000266826                                    False
## ENSG00000101004                                    False
## ENSG00000102218                                    False
## ENSG00000198553                                    False
## ENSG00000112144                                     True
##                 PrimaryCiliaTip_RPTEC_TERT1 PrimaryCiliaTZ_
## ENSG00000089289                       False           False
## ENSG00000266826                       False           False
## ENSG00000101004                       False            True
## ENSG00000102218                       False           False
## ENSG00000198553                       False           False
## ENSG00000112144                       False           False
##                 PrimaryCiliaTZ_ASC52telo
## ENSG00000089289                    False
## ENSG00000266826                    False
## ENSG00000101004                    False
## ENSG00000102218                    False
## ENSG00000198553                    False
## ENSG00000112144                    False
##                 PrimaryCiliaTZ_hTERT_RPE1_serum_starved
## ENSG00000089289                                   False
## ENSG00000266826                                   False
## ENSG00000101004                                    True
## ENSG00000102218                                   False
## ENSG00000198553                                   False
## ENSG00000112144                                   False
##                 PrimaryCiliaTZ_RPTEC_TERT1 Nucleus Mitotic Membrane Cytoplasm
## ENSG00000089289                      False   False    True    False      True
## ENSG00000266826                      False   False    True    False      True
## ENSG00000101004                      False    True   False     True      True
## ENSG00000102218                      False    True   False     True      True
## ENSG00000198553                      False   False   False     True     False
## ENSG00000112144                      False   False   False     True      True
##                 BB_num PrimaryCilia_num PrimaryCiliaTip_num PrimaryCiliaTZ_num
## ENSG00000089289      2                3                   0                  0
## ENSG00000266826      2                3                   0                  0
## ENSG00000101004      0                2                   0                  1
## ENSG00000102218      0                0                   0                  0
## ENSG00000198553      1                0                   0                  0
## ENSG00000112144      0                3                   2                  0
##                 ASC52telo hTERT_RPE1_serum_starved RPTEC_TERT1
## ENSG00000089289      True                     True        True
## ENSG00000266826      True                     True        True
## ENSG00000101004     False                     True        True
## ENSG00000102218     False                    False       False
## ENSG00000198553     False                     True       False
## ENSG00000112144      True                     True        True
```

```
# Identify columns that contain "num" in their names
num_columns <- grep("num", names(df_all), value = TRUE)

# Convert all other columns from string to logical
df_all <- df_all %>%
  mutate(across(-all_of(num_columns), ~ as.logical(.)))

# map gene IDs to gene names
mapped_ids <- map_geneID_to_name(rownames(df_all))

# index by Ensembl
rownames(mapped_ids) <- mapped_ids$ENSEMBL
```

```
# pull out exactly one entry per row of df_all
df_all$GeneSymbol <- mapped_ids[ rownames(df_all), "SYMBOL"   ]
df_all$GeneName   <- mapped_ids[ rownames(df_all), "GENENAME" ]

# add rownames as column and reset index
df_all$Ensembl_ID <- rownames(df_all)
rownames(df_all) <- NULL

# reorder columns to have Ensembl_ID first, Symbol and Gene name first and then all other columns
df_all <- df_all %>%  dplyr::select(Ensembl_ID, GeneSymbol, GeneName, everything())
```

Compare biological themes for different cell lines (all locations pooled)

Split data by cell line

```
# Perform filtering again
df_ASC52telo <- df_all %>% filter(ASC52telo == TRUE)
df_hTERT <- df_all %>% filter(hTERT_RPE1_serum_starved == TRUE)
df_RPTEC_TERT1 <- df_all %>% filter(RPTEC_TERT1 == TRUE)

# save input data as csv file
write.csv(df_ASC52telo, file = paste0(out_path, "Restricted_comparison_cell_lines_GO_BP_input_ASC52telo
write.csv(df_hTERT, file = paste0(out_path, "Restricted_comparison_cell_lines_GO_BP_input_hTERT_RPE1_se
write.csv(df_RPTEC_TERT1, file = paste0(out_path, "Restricted_comparison_cell_lines_GO_BP_input_RPTEC_T

# filter gene_id by cell line
gene_id_ASC52telo <- df_ASC52telo$Ensembl_ID
gene_id_hTERT <- df_hTERT$Ensembl_ID
gene_id_RPTEC_TERT1 <- df_RPTEC_TERT1$Ensembl_ID
```

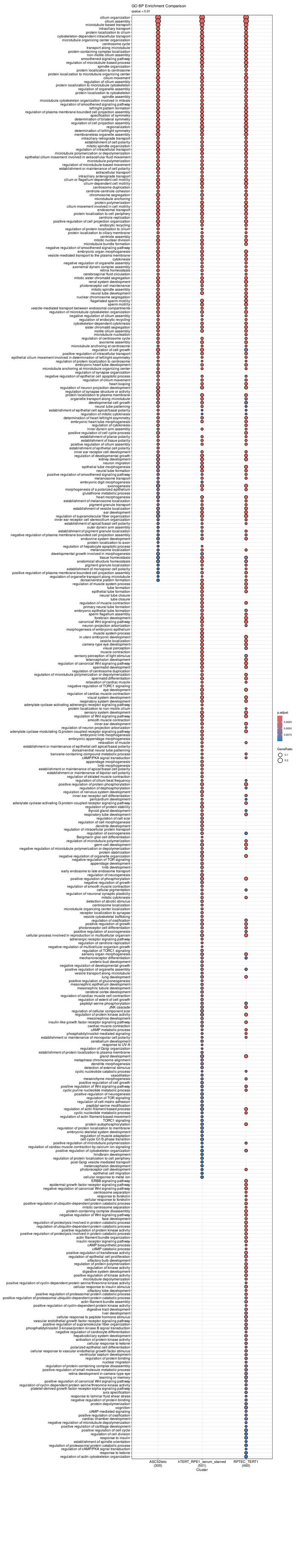
```
# prepare input data
input_genes <- list(
  ASC52telo = gene_id_ASC52telo,
  hTERT_RPE1_serum_starved = gene_id_hTERT,
  RPTEC_TERT1 = gene_id_RPTEC_TERT1
)
```

GO BP enrichment analysis

```
# Perform the compareCluster analysis
comp <- compareCluster(geneCluster = input_genes,
                    fun = "enrichGO",
                    OrgDb = org.Hs.eg.db,
                    keyType       = 'ENSEMBL',
                    ont           = "BP",
                    pAdjustMethod = "BH",
                    pvalueCutoff  = 0.01,
                    qvalueCutoff  = 0.01)
```

Save results as csv file

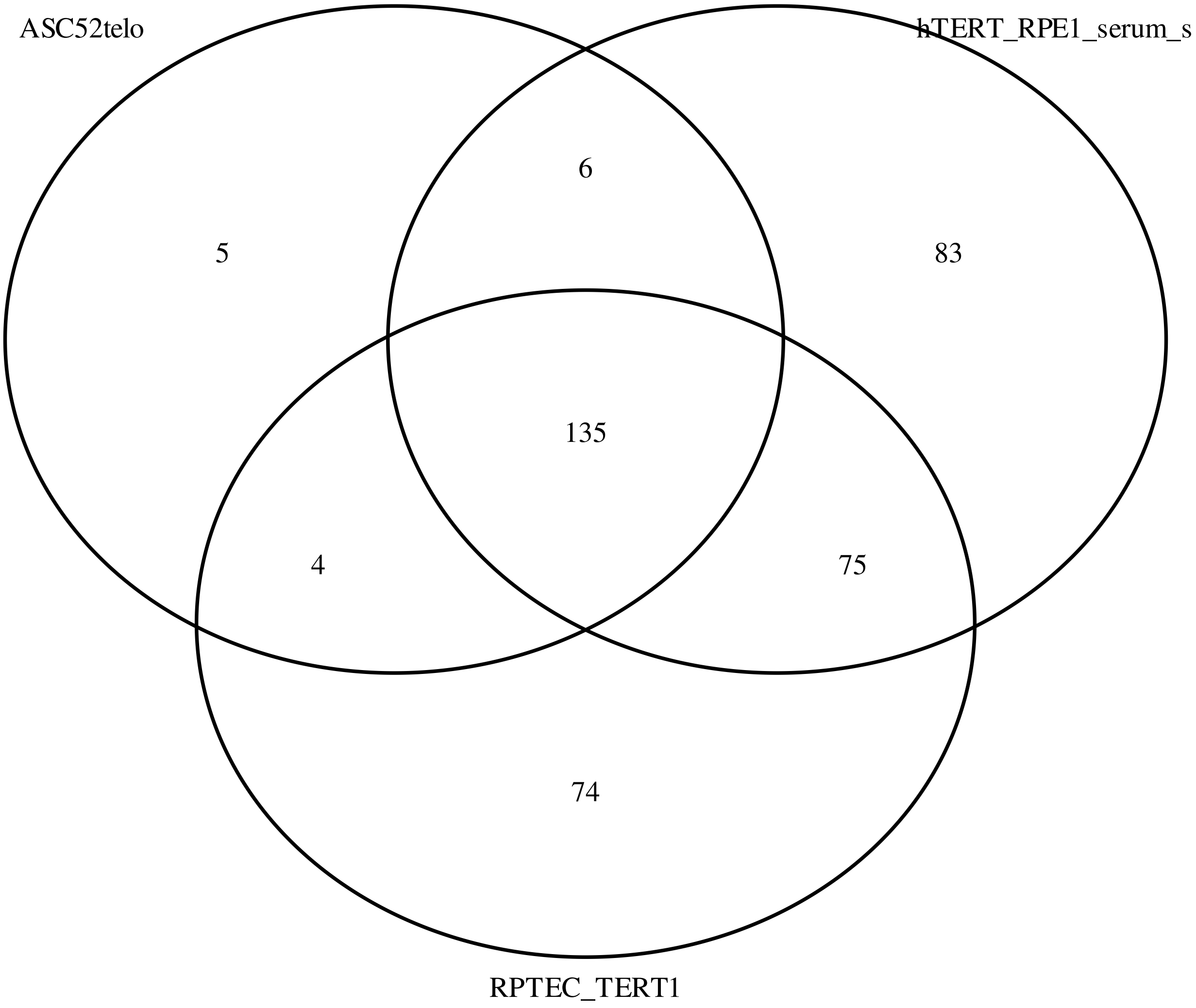
```
dotplot(comp, showCategory = NULL) + ggtitle("GO BP Enrichment Comparison", subtitle = "qvalue < 0.01")
  scale_y_discrete(labels = function(x) str_wrap(x, width = 100))
```


GO BP Enrichment Comparison
qvalue < 0.01
cilium organization
cilium assembly
microtubule-based transport
intraciliary transport
protein localization to cilium
cytoskeleton-dependent intracellular transport
microtubule organizing center organization
centrosome cycle
transport along microtubule
protein-containing complex localization
non-motile cilium assembly
smoothened signaling pathway
regulation of microtubule-based process
spindle organization
protein localization to centrosome
protein localization to microtubule organizing center
cilium movement
regulation of cilium assembly
protein localization to microtubule cytoskeleton
regulation of organelle assembly
protein localization to cytoskeleton
spindle assembly
microtubule cytoskeleton organization involved in mitosis
regulation of smoothened signaling pathway
left/right pattern formation
regulation of plasma membrane bounded cell projection assembly
specification of symmetry
determination of bilateral symmetry
regulation of cell projection assembly
regionalization
determination of left/right symmetry
membraneless organelle assembly
intraciliary retrograde transport
establishment of cell polarity
mitotic spindle organization
regulation of intracellular transport
microtubule polymerization or depolymerization
epithelial cilium movement involved in extracellular fluid movement
microtubule polymerization
regulation of microtubule-based movement
establishment or maintenance of cell polarity
extracellular transport
intraciliary anterograde transport
cilium or flagellum-dependent cell motility
cilium-dependent cell motility
centrosome duplication
centriole-centriole cohesion
chromosome segregation
microtubule anchoring
protein polymerization
cilium movement involved in cell motility
endosomal transport
protein localization to cell periphery
centriole replication
positive regulation of cell projection organization
endocytic recycling
regulation of protein localization to cilium
protein localization to ciliary membrane
centriole assembly
mitotic nuclear division
microtubule bundle formation
negative regulation of smoothened signaling pathway
embryonic organ morphogenesis
vesicle-mediated transport to the plasma membrane
cytokinesis
negative regulation of organelle assembly
axonemal dynein complex assembly
retina homeostasis
cerebrospinal fluid circulation
mitotic sister chromatid segregation
renal system development
photoreceptor cell maintenance
mitotic spindle assembly
neural tube development
nuclear chromosome segregation
flagellated sperm motility
sperm motility
vesicle-mediated transport between endosomal compartments
regulation of microtubule cytoskeleton organization
negative regulation of cilium assembly
regulation of endocytic recycling
cytoskeleton-dependent cytokinesis
sister chromatid segregation
motile cilium assembly
microtubule nucleation
regulation of centrosome cycle
axoneme assembly
microtubule anchoring at centrosome
regulation of cell growth
positive regulation of intracellular transport
epithelial cilium movement involved in determination of left/right asymmetry
regulation of protein localization to centrosome
embryonic heart tube development
microtubule anchoring at microtubule organizing center
regulation of synapse organization
negative regulation of epithelial cell apoptotic process
regulation of cilium movement
heart looping
regulation of neuron projection development
regulation of synapse structure or activity
protein localization to plasma membrane
organelle transport along microtubule
developmental cell growth
neural tube patterning
establishment of epithelial cell apical/basal polarity
regulation of mitotic cytokinesis
determination of heart left/right asymmetry
embryonic heart tube morphogenesis
regulation of cytokinesis
inner dynein arm assembly
positive regulation of cell cycle process
establishment of planar polarity
establishment of tissue polarity
positive regulation of cilium assembly
establishment of epithelial cell polarity
inner ear receptor cell development
regulation of developmental growth
kidney development
neuron migration
epithelial tube morphogenesis
neural tube formation
positive regulation of smoothened signaling pathway
melanosome transport
embryonic digit morphogenesis
axonogenesis
morphogenesis of a polarized epithelium
glutathione metabolic process
heart morphogenesis
establishment of melanosome localization
pigment granule transport
establishment of vesicle localization
ear development
regulation of supramolecular fiber organization
inner ear receptor cell stereocilium organization
establishment of apical/basal cell polarity
outer dynein arm assembly
establishment of pigment granule localization
negative regulation of plasma membrane bounded cell projection assembly
endocrine system development
protein localization to axon
regulation of hepatocyte apoptotic process
melanosome localization
developmental growth involved in morphogenesis
tissue homeostasis
anatomical structure homeostasis
pigment granule localization
establishment of monopolar cell polarity
positive regulation of plasma membrane bounded cell projection assembly
regulation of organelle transport along microtubule
dorsal/ventral pattern formation
regulation of muscle system process
tube formation
epithelial tube formation
neural tube closure
tube closure
regulation of muscle contraction
primary neural tube formation
embryonic epithelial tube formation
sperm flagellum assembly
forebrain development
canonical Wnt signaling pathway
neuron projection arborization
morphogenesis of embryonic epithelium
muscle system process
in utero embryonic development
vesicle localization
camera-type eye development
visual perception
muscle contraction
sensory perception of light stimulus
telencephalon development
regulation of canonical Wnt signaling pathway
spermatid development
regulation of centrosome duplication
regulation of microtubule polymerization or depolymerization
spermatid differentiation
relaxation of cardiac muscle
negative regulation of TORC1 signaling
eye development
regulation of cardiac muscle contraction
visual system development
respiratory system development
adenylate cyclase-activating adrenergic receptor signaling pathway
protein localization to non-motile cilium
sensory system development
regulation of Wnt signaling pathway
smooth muscle contraction
inner ear development
regulation of neuron projection arborization
adenylate cyclase-modulating G protein-coupled receptor signaling pathway
embryonic limb morphogenesis
embryonic appendage morphogenesis
relaxation of muscle
establishment or maintenance of epithelial cell apical/basal polarity
dorsal/ventral neural tube patterning
benzene-containing compound metabolic process
cAMP/PKA signal transduction
appendage morphogenesis
limb morphogenesis
establishment or maintenance of apical/basal cell polarity
establishment or maintenance of bipolar cell polarity
regulation of striated muscle contraction
regulation of cilium beat frequency
positive regulation of protein phosphorylation
regulation of dephosphorylation
regulation of nervous system development
inner ear receptor cell differentiation
pericardium development
adenylate cyclase-activating G protein-coupled receptor signaling pathway
regulation of protein stability
thyroid gland development
respiratory tube development
regulation of cell size
regulation of cell morphogenesis
dendrite development
regulation of intracellular protein transport
regulation of axonogenesis
Bergmann glial cell differentiation
regulation of microtubule polymerization
germ cell development
negative regulation of microtubule polymerization or depolymerization
protein stabilization
negative regulation of organelle organization
negative regulation of TOR signaling
appendage development
limb development
early endosome to late endosome transport
regulation of neurogenesis
positive regulation of phosphorylation
negative regulation of growth
regulation of smooth muscle contraction
cellular pigmentation
regulation of neuronal synaptic plasticity
mitotic cytokinesis
detection of abiotic stimulus
centrosome localization
microtubule organizing center localization
receptor localization to synapse
vesicle cytoskeletal trafficking
regulation of ossification
positive regulation of growth
photoreceptor cell differentiation
positive regulation of axonogenesis
cellular process involved in reproduction in multicellular organism
adrenergic receptor signaling pathway
regulation of centriole replication
negative regulation of multicellular organism growth
regulation of TORC1 signaling
sensory organ morphogenesis
mechanoreceptor differentiation
ureteric bud development
negative regulation of developmental growth
positive regulation of organelle assembly
vesicle transport along microtubule
lung development
positive regulation of gluconeogenesis
mesonephric epithelium development
mesonephric tubule development
cerebral cortex development
regulation of cardiac muscle cell contraction
regulation of extent of cell growth
peptidyl-serine phosphorylation
JNK cascade
regulation of cellular component size
regulation of protein kinase activity
mesonephros development
insulin-like growth factor receptor signaling pathway
cardiac muscle contraction
cAMP metabolic process
phosphatidylinositol-mediated signaling
establishment or maintenance of monopolar cell polarity
cerebellum development
response to UV-A
regulation of Golgi organization
establishment of protein localization to plasma membrane
gland development
metaphase chromosome alignment
dendrite morphogenesis
detection of external stimulus
cyclic nucleotide catabolic process
vasodilation
mesenchyme morphogenesis
positive regulation of cell growth
positive regulation of Wnt signaling pathway
cyclic purine nucleotide metabolic process
positive regulation of neurogenesis
regulation of TOR signaling
regulation of cell-matrix adhesion
peptidyl-serine modification
regulation of actin filament-based process
cyclic nucleotide metabolic process
regulation of actin filament-based movement
TORC1 signaling
protein autophosphorylation
regulation of protein localization to membrane
embryonic skeletal system development
regulation of muscle adaptation
cell cycle G1/S phase transition
positive regulation of microtubule polymerization
regulation of cardiac muscle contraction by calcium ion signaling
positive regulation of cytoskeleton organization
hindbrain development
regulation of protein localization to cell periphery
post-Golgi vesicle-mediated transport
metencephalon development
photoreceptor cell development
epithelial cell migration
cellular response to metal ion
ERBB signaling pathway
epidermal growth factor receptor signaling pathway
negative regulation of canonical Wnt signaling pathway
centrosome separation
response to forskolin
cellular response to forskolin
positive regulation of ubiquitin-dependent protein catabolic process
mitotic centrosome separation
protein-containing complex disassembly
negative regulation of Wnt signaling pathway
face development
regulation of proteolysis involved in protein catabolic process
regulation of ubiquitin-dependent protein catabolic process
positive regulation of protein kinase activity
positive regulation of proteolysis involved in protein catabolic process
actin filament bundle organization
insulin receptor signaling pathway
cAMP biosynthetic process
cAMP catabolic process
positive regulation of transferase activity
regulation of epithelial cell proliferation
olfactory bulb development
regulation of protein polymerization
regulation of kinase activity
digestive system development
positive regulation of kinase activity
microtubule depolymerization
positive regulation of cyclin-dependent protein serine/threonine kinase activity
cellular response to insulin stimulus
olfactory lobe development
positive regulation of proteasomal protein catabolic process
positive regulation of proteasomal ubiquitin-dependent protein catabolic process
actin filament bundle assembly
positive regulation of cyclin-dependent protein kinase activity
digestive tract development
liver development
cellular response to peptide hormone stimulus
vascular endothelial growth factor receptor signaling pathway
positive regulation of supramolecular fiber organization
phosphatidylinositol 3-kinase/protein kinase B signal transduction
negative regulation of cardiocyte differentiation
hepaticobiliary system development
activation of protein kinase activity
cellular response to ketone
polarized epithelial cell differentiation
cellular response to vascular endothelial growth factor stimulus
ventricular septum development
regulation of protein binding
nuclear migration
regulation of protein-containing complex disassembly
positive regulation of small molecule metabolic process
retina development in camera-type eye
learning or memory
positive regulation of canonical Wnt signaling pathway
regulation of cyclin-dependent protein serine/threonine kinase activity
platelet-derived growth factor receptor-alpha signaling pathway
axis specification
response to laminar fluid shear stress
negative regulation of protein binding
protein depolymerization
cognition
cAMP-mediated signaling
positive regulation of ossification
cardiac chamber development
negative regulation of microtubule depolymerization
positive regulation of cartilage development
positive regulation of cell cycle
regulation of cell division
response to insulin
establishment of spindle orientation
regulation of proteasomal protein catabolic process
regulation of cAMP/PKA signal transduction
response to ketone
regulation of actin cytoskeleton organization
ASC52telo
(300)
hTERT_RPE1_serum_starved
(501)
RPTEC_TERT1
(460)
Cluster
p.adjust
0.0025
0.0050
0.0075
GeneRatio
0.1
0.2

```
# save dotplot as svg file
ggsave(paste0(out_path, "Restricted_comparison_cell_lines_GO_BP_dotplot.svg"), plot = last_plot(), devi
```

Visualize overlap of terms

```
# extract results
results <- comp@compareClusterResult

# prepare input data
input_genes <- list(
  ASC52telo = gene_id_ASC52telo,
  hTERT_RPE1_serum_starved = gene_id_hTERT,
  RPTEC_TERT1 = gene_id_RPTEC_TERT1
)

# split by location
ASC52telo <- results[results$Cluster == "ASC52telo",]
hTERT_RPE1_serum_starved <- results[results$Cluster == "hTERT_RPE1_serum_starved",]
RPTEC_TERT1 <- results[results$Cluster == "RPTEC_TERT1",]

# Create a list of the four sets
go_lists <- list(
  ASC52telo = ASC52telo$ID,
  hTERT_RPE1_serum_starved = hTERT_RPE1_serum_starved$ID,
  RPTEC_TERT1 = RPTEC_TERT1$ID
)

# Plot the Venn diagram
venn.plot <- venn.diagram(
  x = go_lists,
  category.names = c("ASC52telo", "hTERT_RPE1_serum_starved", "RPTEC_TERT1"),
  filename = NULL,
  output = TRUE
)

grid.newpage()
grid.draw(venn.plot)
```



```
# Save the captured plot as an SVG file
svglite(paste0(out_path, "Restricted_comparison_cell_lines_GO_BP_venn.svg"), width = 7, height = 6)
grid.draw(venn.plot)
dev.off()
```

```
## cairo_pdf
##         2
```

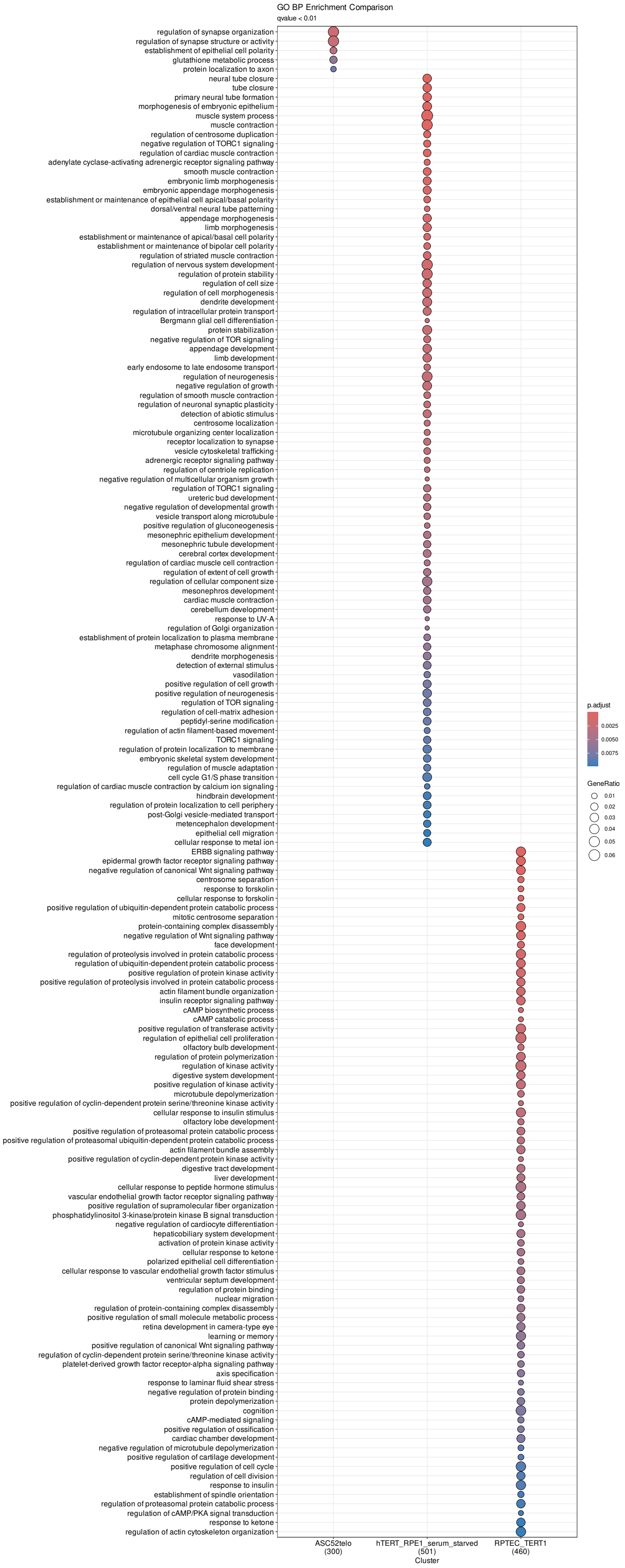
Filter for terms only enriched for one of the locations

```
 # get results as data frame
comp_results <- comp@compareClusterResult

# Step 2: Count occurrences
term_counts <- table(comp_results$ID)

# Step 3: Filter proteins that appear at least twice
```

```
terms_at_least_twice <- names(term_counts[term_counts >= 2])

# remove terms that are enriched in more than one cell line
unspecific_terms <- comp_results[comp_results$ID %in% terms_at_least_twice,]
specific_terms <- comp_results[!comp_results$ID %in% terms_at_least_twice,]

# create a copy of comp
comp_filtered <- comp

# update results in comp
comp_filtered@compareClusterResult <- specific_terms
```

Dot plot of uniquely enriched terms

```
# plot dotplot
dotplot(comp_filtered, showCategory = NULL) + ggtitle("GO BP Enrichment Comparison", subtitle = "qvalue
  scale_y_discrete(labels = function(x) str_wrap(x, width = 100))
```


GO BP Enrichment Comparison
qvalue < 0.01
regulation of synapse organization
regulation of synapse structure or activity
establishment of epithelial cell polarity
glutathione metabolic process
protein localization to axon
neural tube closure
tube closure
primary neural tube formation
morphogenesis of embryonic epithelium
muscle system process
muscle contraction
regulation of centrosome duplication
negative regulation of TORC1 signaling
regulation of cardiac muscle contraction
adenylate cyclase-activating adrenergic receptor signaling pathway
smooth muscle contraction
embryonic limb morphogenesis
embryonic appendage morphogenesis
establishment or maintenance of epithelial cell apical/basal polarity
dorsal/ventral neural tube patterning
appendage morphogenesis
limb morphogenesis
establishment or maintenance of apical/basal cell polarity
establishment or maintenance of bipolar cell polarity
regulation of striated muscle contraction
regulation of nervous system development
regulation of protein stability
regulation of cell size
regulation of cell morphogenesis
dendrite development
regulation of intracellular protein transport
Bergmann glial cell differentiation
protein stabilization
negative regulation of TOR signaling
appendage development
limb development
early endosome to late endosome transport
regulation of neurogenesis
negative regulation of growth
regulation of smooth muscle contraction
regulation of neuronal synaptic plasticity
detection of abiotic stimulus
centrosome localization
microtubule organizing center localization
receptor localization to synapse
vesicle cytoskeletal trafficking
adrenergic receptor signaling pathway
regulation of centriole replication
negative regulation of multicellular organism growth
regulation of TORC1 signaling
ureteric bud development
negative regulation of developmental growth
vesicle transport along microtubule
positive regulation of gluconeogenesis
mesonephric epithelium development
mesonephric tubule development
cerebral cortex development
regulation of cardiac muscle cell contraction
regulation of extent of cell growth
regulation of cellular component size
mesonephros development
cardiac muscle contraction
cerebellum development
response to UV-A
regulation of Golgi organization
establishment of protein localization to plasma membrane
metaphase chromosome alignment
dendrite morphogenesis
detection of external stimulus
vasodilation
positive regulation of cell growth
positive regulation of neurogenesis
regulation of TOR signaling
regulation of cell-matrix adhesion
peptidyl-serine modification
regulation of actin filament-based movement
TORC1 signaling
regulation of protein localization to membrane
embryonic skeletal system development
regulation of muscle adaptation
cell cycle G1/S phase transition
regulation of cardiac muscle contraction by calcium ion signaling
hindbrain development
regulation of protein localization to cell periphery
post-Golgi vesicle-mediated transport
metencephalon development
epithelial cell migration
cellular response to metal ion
ERBB signaling pathway
epidermal growth factor receptor signaling pathway
negative regulation of canonical Wnt signaling pathway
centrosome separation
response to forskolin
cellular response to forskolin
positive regulation of ubiquitin-dependent protein catabolic process
mitotic centrosome separation
protein-containing complex disassembly
negative regulation of Wnt signaling pathway
face development
regulation of proteolysis involved in protein catabolic process
regulation of ubiquitin-dependent protein catabolic process
positive regulation of protein kinase activity
positive regulation of proteolysis involved in protein catabolic process
actin filament bundle organization
insulin receptor signaling pathway
cAMP biosynthetic process
cAMP catabolic process
positive regulation of transferase activity
regulation of epithelial cell proliferation
olfactory bulb development
regulation of protein polymerization
regulation of kinase activity
digestive system development
positive regulation of kinase activity
microtubule depolymerization
positive regulation of cyclin-dependent protein serine/threonine kinase activity
cellular response to insulin stimulus
olfactory lobe development
positive regulation of proteasomal protein catabolic process
positive regulation of proteasomal ubiquitin-dependent protein catabolic process
actin filament bundle assembly
positive regulation of cyclin-dependent protein kinase activity
digestive tract development
liver development
cellular response to peptide hormone stimulus
vascular endothelial growth factor receptor signaling pathway
positive regulation of supramolecular fiber organization
phosphatidylinositol 3-kinase/protein kinase B signal transduction
negative regulation of cardiocyte differentiation
hepaticobiliary system development
activation of protein kinase activity
cellular response to ketone
polarized epithelial cell differentiation
cellular response to vascular endothelial growth factor stimulus
ventricular septum development
regulation of protein binding
nuclear migration
regulation of protein-containing complex disassembly
positive regulation of small molecule metabolic process
retina development in camera-type eye
learning or memory
positive regulation of canonical Wnt signaling pathway
regulation of cyclin-dependent protein serine/threonine kinase activity
platelet-derived growth factor receptor-alpha signaling pathway
axis specification
response to laminar fluid shear stress
negative regulation of protein binding
protein depolymerization
cognition
cAMP-mediated signaling
positive regulation of ossification
cardiac chamber development
negative regulation of microtubule depolymerization
positive regulation of cartilage development
positive regulation of cell cycle
regulation of cell division
response to insulin
establishment of spindle orientation
regulation of proteasomal protein catabolic process
regulation of cAMP/PKA signal transduction
response to ketone
regulation of actin cytoskeleton organization
ASC52telo
(300)
hTERT_RPE1_serum_starved
(501)
RPTEC_TERT1
(460)
Cluster
p.adjust
0.0025
0.0050
0.0075
GeneRatio
0.01
0.02
0.03
0.04
0.05
0.06

```r
# save dotplot as svg file
ggsave(paste0(out_path, "Restricted_comparison_cell_lines_GO_BP_dotplot_specific_terms_only.svg"), plot
```

Clustering of enriched GO BP terms

```r
# Subset for pc_tip
ASC52telo <- comp_filtered@compareClusterResult[
  comp_filtered@compareClusterResult$Cluster == "ASC52telo",
]

# Subset for bb_tz
hTERT_RPE1_serum_starved <- comp_filtered@compareClusterResult[
  comp_filtered@compareClusterResult$Cluster == "hTERT_RPE1_serum_starved",
]

# Subset for bb_tz
RPTEC_TERT1 <- comp_filtered@compareClusterResult[
  comp_filtered@compareClusterResult$Cluster == "RPTEC_TERT1",
]

# Create new compareClusterResult objects for each subset
comp_filtered_ASC52telo <- comp_filtered
comp_filtered_hTERT_RPE1_serum_starved <- comp_filtered
comp_filtered_RPTEC_TERT1 <- comp_filtered

comp_filtered_ASC52telo@compareClusterResult <- ASC52telo
comp_filtered_hTERT_RPE1_serum_starved@compareClusterResult <- hTERT_RPE1_serum_starved
comp_filtered_RPTEC_TERT1@compareClusterResult <- RPTEC_TERT1
```

Cluster results - ASC52telo

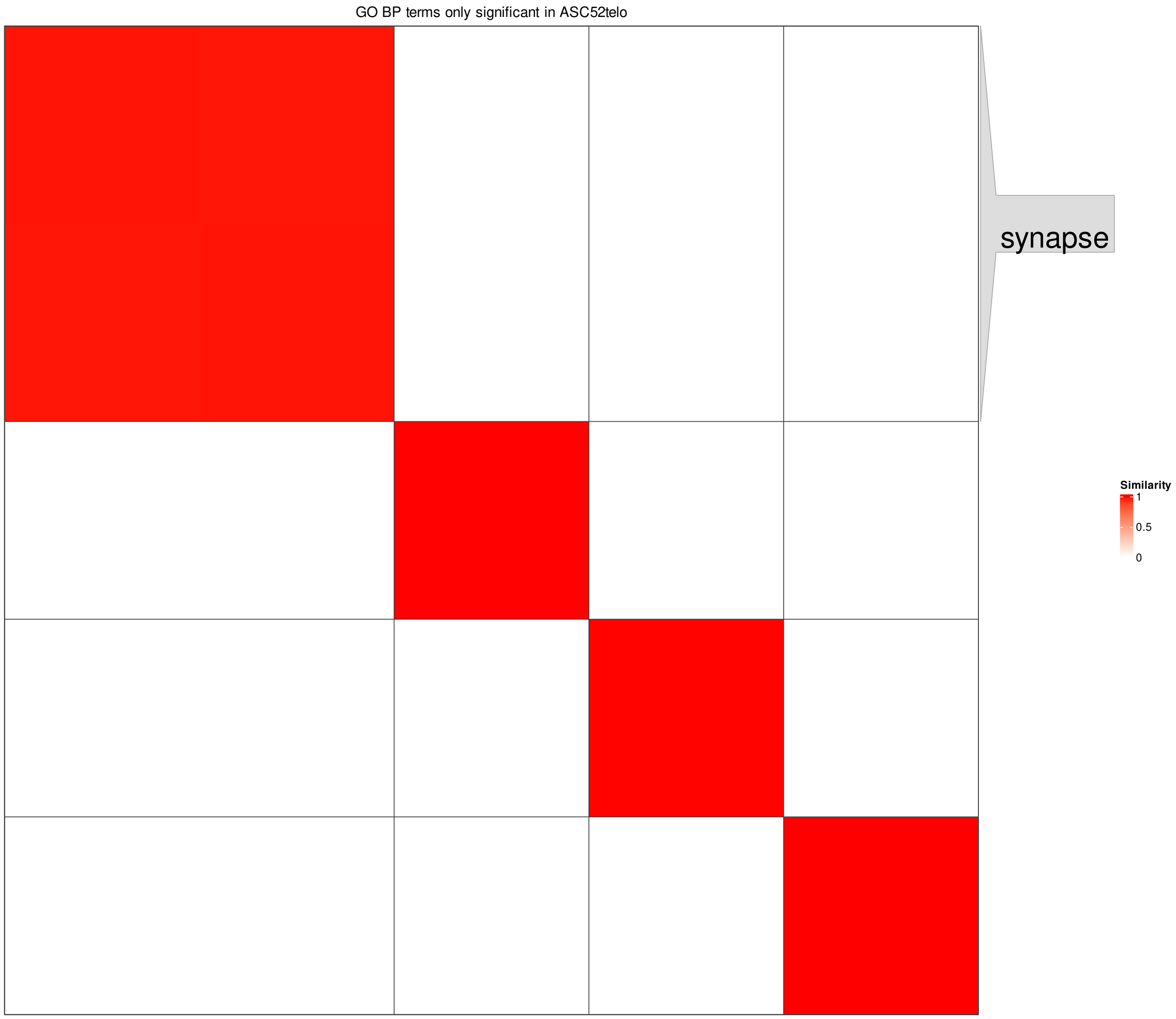
```r
go_id = comp_filtered_ASC52telo@compareClusterResult$ID
mat = GO_similarity(go_id,
                    ont = 'BP',
                    db = 'org.Hs.eg.db',
                    measure = "Sim_Relevance_2006")
```

```r
# Capture the plot
heatmap_plot <- grid.grabExpr({
 df <- simplifyGO(mat,
             method = 'binary_cut',
             plot = TRUE,
             column_title = "GO BP terms only significant in ASC52telo",
             use_raster = FALSE,
             order_by_size = TRUE,
             fontsize_range = c(18, 36),
              max_words = 6,
              word_cloud_grob_param = list(col = 'black',
                                            max_width = unit(200, "mm")))
})

# Save the captured plot as an SVG file
svglite(paste0(out_path, "Restricted_comparison_cell_lines_cilia_only_combined_GO_BP_dotplot_specific_t
grid.draw(heatmap_plot)
dev.off()
```

Plot cluster heatmap

```
## cairo_pdf
##         2
```

```
grid.newpage()
grid.draw(heatmap_plot)
```



```
# add cluster number from GO term clustering
results <- map_cluster_number(comp_filtered_ASC52telo,
                              df = df,
                              comp = TRUE
)

# Apply the function to each row of the DataFrame
results$GeneSymbol <- sapply(results$geneID, map_geneID_to_symbol)

# save results as csv file
write.csv(results, file = paste0(out_path, "Restricted_comparison_cell_lines_cilia_only_combined_GO_BP_c
```

Process and save results

Cluster results - hTERT_RPE1_serum_starved

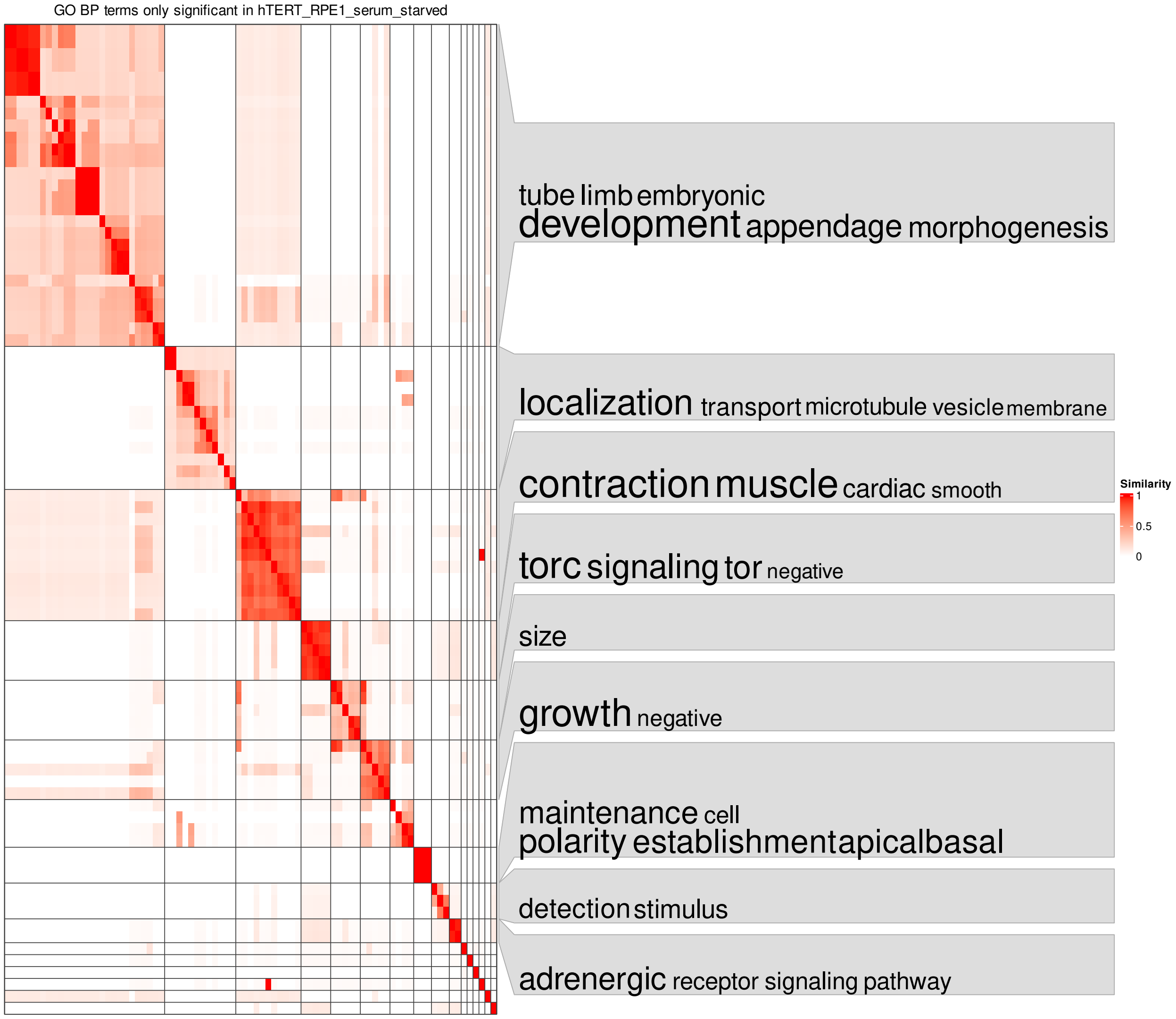
```
go_id = comp_filtered_hTERT_RPE1_serum_starved@compareClusterResult$ID
mat = GO_similarity(go_id,
                    ont = 'BP',
                    db = 'org.Hs.eg.db',
                    measure = "Sim_Relevance_2006")
```

```
# Capture the plot
heatmap_plot <- grid.grabExpr({
 df <- simplifyGO(mat,
             method = 'binary_cut',
             plot = TRUE,
             column_title = "GO BP terms only significant in hTERT_RPE1_serum_starved",
             use_raster = FALSE,
             order_by_size = TRUE,
             fontsize_range = c(18, 36),
             max_words = 6,
             word_cloud_grob_param = list(col = 'black',
                                          max_width = unit(200, "mm")))
})

# Save the captured plot as an SVG file
svglite(paste0(out_path, "Restricted_comparison_cell_lines_cilia_only_combined_GO_BP_dotplot_specific_te
grid.draw(heatmap_plot)
dev.off()
```

Plot cluster heatmap

```
## cairo_pdf
##         2
```

```r
grid.newpage()
grid.draw(heatmap_plot)
```



```r
# add cluster number from GO term clustering
results <- map_cluster_number(comp_filtered_hTERT_RPE1_serum_starved,
                              df = df,
                              comp = TRUE
)

# Apply the function to each row of the DataFrame
results$GeneSymbol <- sapply(results$geneID, map_geneID_to_symbol)

# save results as csv file
write.csv(results, file = paste0(out_path, "Restricted_comparison_cell_lines_cilia_only_combined_GO_BP_
```

Process and save results

Cluster results - RPTEC_TERT1

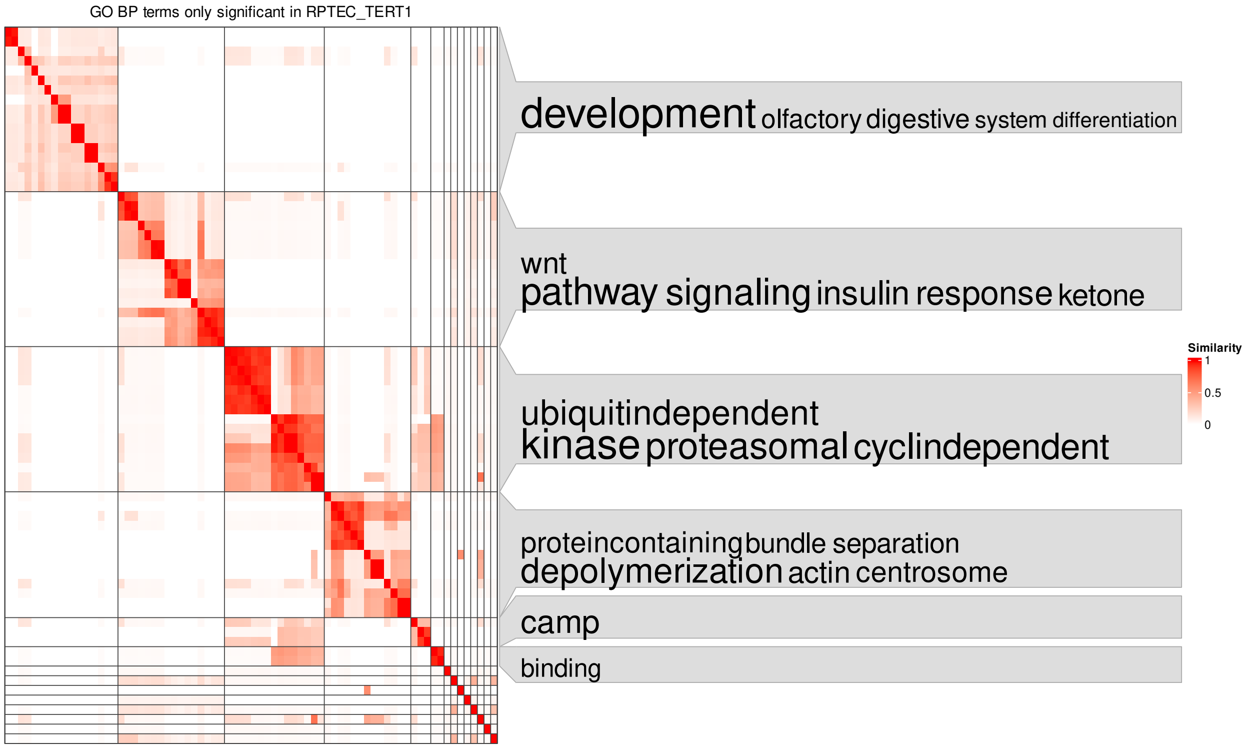
```r
go_id = comp_filtered_RPTEC_TERT1@compareClusterResult$ID
mat = GO_similarity(go_id,
                    ont = 'BP',
                    db = 'org.Hs.eg.db',
                    measure = "Sim_Relevance_2006")
```

```r
# Capture the plot
heatmap_plot <- grid.grabExpr({
 df <- simplifyGO(mat,
             method = 'binary_cut',
             plot = TRUE,
             column_title = "GO BP terms only significant in RPTEC_TERT1",
             use_raster = FALSE,
             order_by_size = TRUE,
             fontsize_range = c(18, 36),
             max_words = 6,
             word_cloud_grob_param = list(col = 'black',
                                          max_width = unit(200, "mm")))
})

# Save the captured plot as an SVG file
svglite(paste0(out_path, "Restricted_comparison_cell_lines_cilia_only_combined_GO_BP_dotplot_specific_t
grid.draw(heatmap_plot)
dev.off()
```

Plot cluster heatmap

```
## cairo_pdf
##         2
```

```r
grid.newpage()
grid.draw(heatmap_plot)
```



```
# add cluster number from GO term clustering
results <- map_cluster_number(comp_filtered_RPTEC_TERT1,
                              df = df,
                              comp = TRUE
)

# Apply the function to each row of the DataFrame
results$GeneSymbol <- sapply(results$geneID, map_geneID_to_symbol)

# save results as csv file
write.csv(results, file = paste0(out_path, "Restricted_comparison_cell_lines_cilia_only_combined_GO_BP_
```

Process and save results

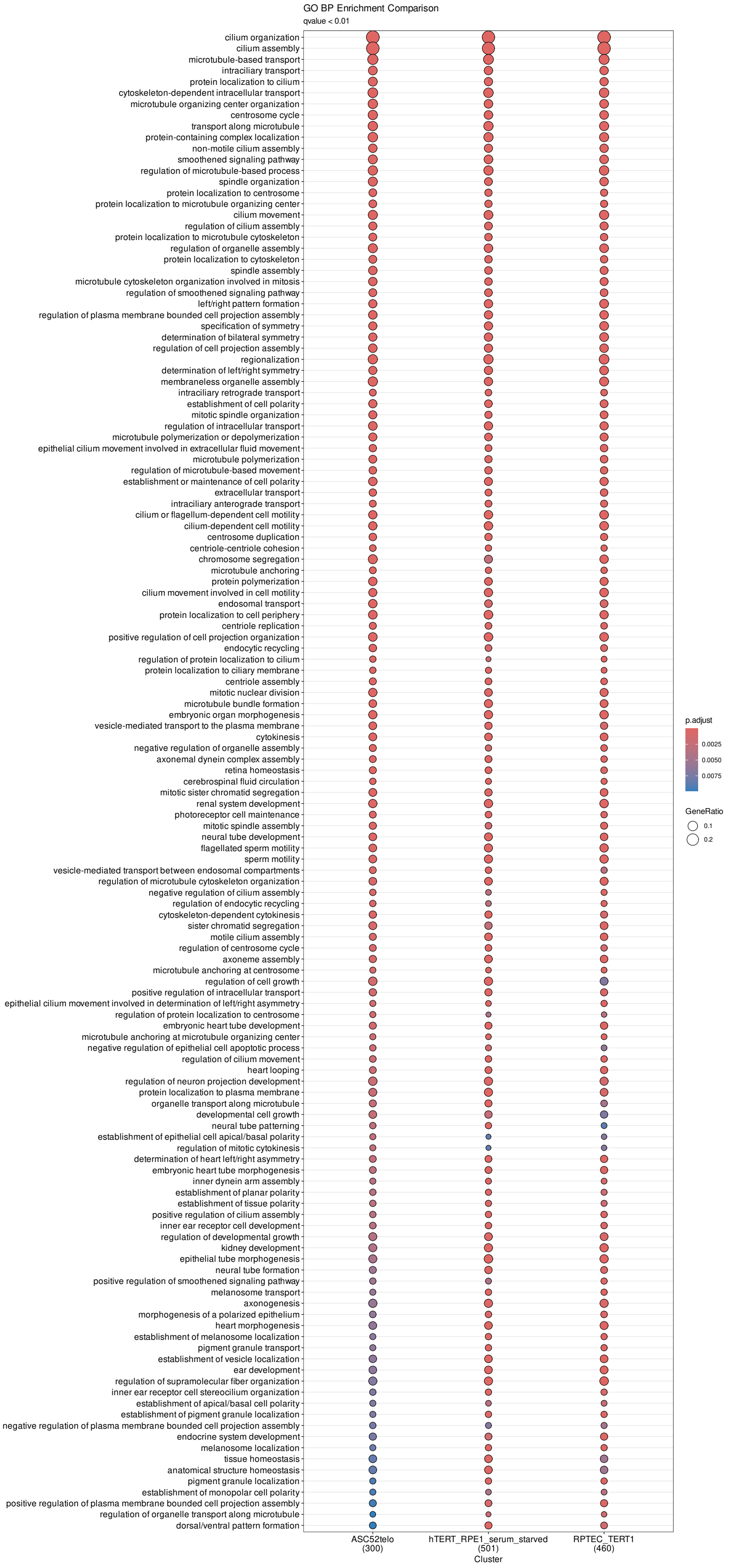
Filter for terms enriched in all three cell lines

```
 # get results as data frame
comp_results <- comp@compareClusterResult

# Count occurrences
term_counts <- table(comp_results$ID)

# get terms in all three cell lines
terms_in_all <- names(term_counts[term_counts == 3])
```

```
# remove all unspecific terms
unspecific_terms <- comp_results %>% filter(ID %in% terms_in_all)

# create a copy of comp
comp_filtered <- comp

# update results in comp
comp_filtered@compareClusterResult <- unspecific_terms
```

Dot plot of uniquely enriched terms

```
# plot dotplot
dotplot(comp_filtered, showCategory = NULL) + ggtitle("GO BP Enrichment Comparison", subtitle = "qvalue
  scale_y_discrete(labels = function(x) str_wrap(x, width = 100))
```


GO BP Enrichment Comparison
qvalue < 0.01
cilium organization
cilium assembly
microtubule-based transport
intraciliary transport
protein localization to cilium
cytoskeleton-dependent intracellular transport
microtubule organizing center organization
centrosome cycle
transport along microtubule
protein-containing complex localization
non-motile cilium assembly
smoothened signaling pathway
regulation of microtubule-based process
spindle organization
protein localization to centrosome
protein localization to microtubule organizing center
cilium movement
regulation of cilium assembly
protein localization to microtubule cytoskeleton
regulation of organelle assembly
protein localization to cytoskeleton
spindle assembly
microtubule cytoskeleton organization involved in mitosis
regulation of smoothened signaling pathway
left/right pattern formation
regulation of plasma membrane bounded cell projection assembly
specification of symmetry
determination of bilateral symmetry
regulation of cell projection assembly
regionalization
determination of left/right symmetry
membraneless organelle assembly
intraciliary retrograde transport
establishment of cell polarity
mitotic spindle organization
regulation of intracellular transport
microtubule polymerization or depolymerization
epithelial cilium movement involved in extracellular fluid movement
microtubule polymerization
regulation of microtubule-based movement
establishment or maintenance of cell polarity
extracellular transport
intraciliary anterograde transport
cilium or flagellum-dependent cell motility
cilium-dependent cell motility
centrosome duplication
centriole-centriole cohesion
chromosome segregation
microtubule anchoring
protein polymerization
cilium movement involved in cell motility
endosomal transport
protein localization to cell periphery
centriole replication
positive regulation of cell projection organization
endocytic recycling
regulation of protein localization to cilium
protein localization to ciliary membrane
centriole assembly
mitotic nuclear division
microtubule bundle formation
embryonic organ morphogenesis
vesicle-mediated transport to the plasma membrane
cytokinesis
negative regulation of organelle assembly
axonemal dynein complex assembly
retina homeostasis
cerebrospinal fluid circulation
mitotic sister chromatid segregation
renal system development
photoreceptor cell maintenance
mitotic spindle assembly
neural tube development
flagellated sperm motility
sperm motility
vesicle-mediated transport between endosomal compartments
regulation of microtubule cytoskeleton organization
negative regulation of cilium assembly
regulation of endocytic recycling
cytoskeleton-dependent cytokinesis
sister chromatid segregation
motile cilium assembly
regulation of centrosome cycle
axoneme assembly
microtubule anchoring at centrosome
regulation of cell growth
positive regulation of intracellular transport
epithelial cilium movement involved in determination of left/right asymmetry
regulation of protein localization to centrosome
embryonic heart tube development
microtubule anchoring at microtubule organizing center
negative regulation of epithelial cell apoptotic process
regulation of cilium movement
heart looping
regulation of neuron projection development
protein localization to plasma membrane
organelle transport along microtubule
developmental cell growth
neural tube patterning
establishment of epithelial cell apical/basal polarity
regulation of mitotic cytokinesis
determination of heart left/right asymmetry
embryonic heart tube morphogenesis
inner dynein arm assembly
establishment of planar polarity
establishment of tissue polarity
positive regulation of cilium assembly
inner ear receptor cell development
regulation of developmental growth
kidney development
epithelial tube morphogenesis
neural tube formation
positive regulation of smoothened signaling pathway
melanosome transport
axonogenesis
morphogenesis of a polarized epithelium
heart morphogenesis
establishment of melanosome localization
pigment granule transport
establishment of vesicle localization
ear development
regulation of supramolecular fiber organization
inner ear receptor cell stereocilium organization
establishment of apical/basal cell polarity
establishment of pigment granule localization
negative regulation of plasma membrane bounded cell projection assembly
endocrine system development
melanosome localization
tissue homeostasis
anatomical structure homeostasis
pigment granule localization
establishment of monopolar cell polarity
positive regulation of plasma membrane bounded cell projection assembly
regulation of organelle transport along microtubule
dorsal/ventral pattern formation
ASC52telo
(300)
hTERT_RPE1_serum_starved
(501)
RPTEC_TERT1
(460)
Cluster
p.adjust
0.0025
0.0050
0.0075
GeneRatio
0.1
0.2

```
# save dotplot as svg file
ggsave(paste0(out_path, "Restricted_comparison_celllines_GO_BP_dotplot_shared_terms_only.svg"), plot = 
```

Cluster results - shared terms of all three cell lines

```
# Split by location
ASC52telo <- comp_filtered@compareClusterResult[comp_filtered@compareClusterResult$Cluster == "ASC52tel
hTERT_RPE1_serum_starved <- comp_filtered@compareClusterResult[comp_filtered@compareClusterResult$Cluste
RPTEC_TERT1 <- comp_filtered@compareClusterResult[comp_filtered@compareClusterResult$Cluster == "RPTEC_

# Create new compareClusterResult objects for each subset
comp_filtered_ASC52telo<- comp_filtered
comp_filtered_hTERT_RPE1_serum_starved <- comp_filtered
comp_filtered_RPTEC_TERT1 <- comp_filtered

comp_filtered_ASC52telo@compareClusterResult <- ASC52telo
comp_filtered_hTERT_RPE1_serum_starved@compareClusterResult <- hTERT_RPE1_serum_starved
comp_filtered_RPTEC_TERT1@compareClusterResult <- RPTEC_TERT1
```

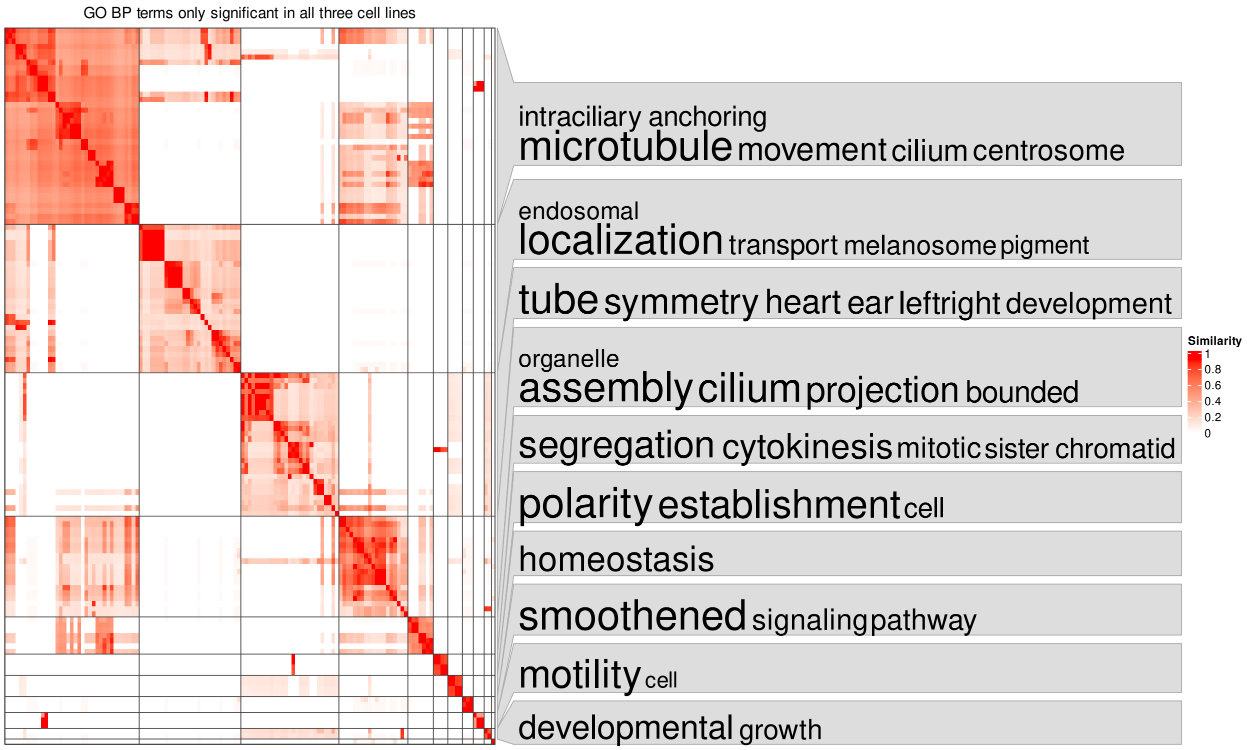
```
go_id = comp_filtered_ASC52telo@compareClusterResult$ID
mat = GO_similarity(go_id,
                    ont = 'BP',
                    db = 'org.Hs.eg.db',
                    measure = "Sim_Relevance_2006")
```

```


# Capture the plot
heatmap_plot <- grid.grabExpr({
 df <- simplifyGO(mat,
             method = 'binary_cut',
             plot = TRUE,
             column_title = "GO BP terms only significant in all three cell lines",
             use_raster = FALSE,
             order_by_size = TRUE,
             fontsize_range = c(18, 36),
             max_words = 6,
             word_cloud_grob_param = list(col = 'black',
                                          max_width = unit(200, "mm")))
})

# Save the captured plot as an SVG file
svglite(paste0(out_path, "Restricted_comparison_cell_lines_cilia_only_combined_GO_BP_dotplot_terms_in_al
grid.draw(heatmap_plot)
dev.off()
```

Plot cluster heatmap

```
## cairo_pdf
##         2
```

```
grid.newpage()
grid.draw(heatmap_plot)
```



```
# add cluster number from GO term clustering
results <- map_cluster_number(comp_filtered_ASC52telo,
                              df = df,
                              comp = TRUE
)

# Apply the function to each row of the DataFrame
results$GeneSymbol <- sapply(results$geneID, map_geneID_to_symbol)

# save results as csv file
write.csv(results, file = paste0(out_path, "Restricted_comparison_cell_lines_cilia_only_combined_GO_BP_c
```

```
# add cluster number from GO term clustering
results <- map_cluster_number(comp_filtered_hTERT_RPE1_serum_starved,
                              df = df,
                              comp = TRUE
)

# Apply the function to each row of the DataFrame
results$GeneSymbol <- sapply(results$geneID, map_geneID_to_symbol)
```

```r
# save results as csv file
write.csv(results, file = paste0(out_path, "Restricted_comparison_cell_lines_cilia_only_combined_GO_BP_c
```

```r


# add cluster number from GO term clustering
results <- map_cluster_number(comp_filtered_RPTEC_TERT1,
                              df = df,
                              comp = TRUE
)

# Apply the function to each row of the DataFrame
results$GeneSymbol <- sapply(results$geneID, map_geneID_to_symbol)

# save results as csv file
write.csv(results, file = paste0(out_path, "Restricted_comparison_cell_lines_cilia_only_combined_GO_BP_c
```

Process and save results

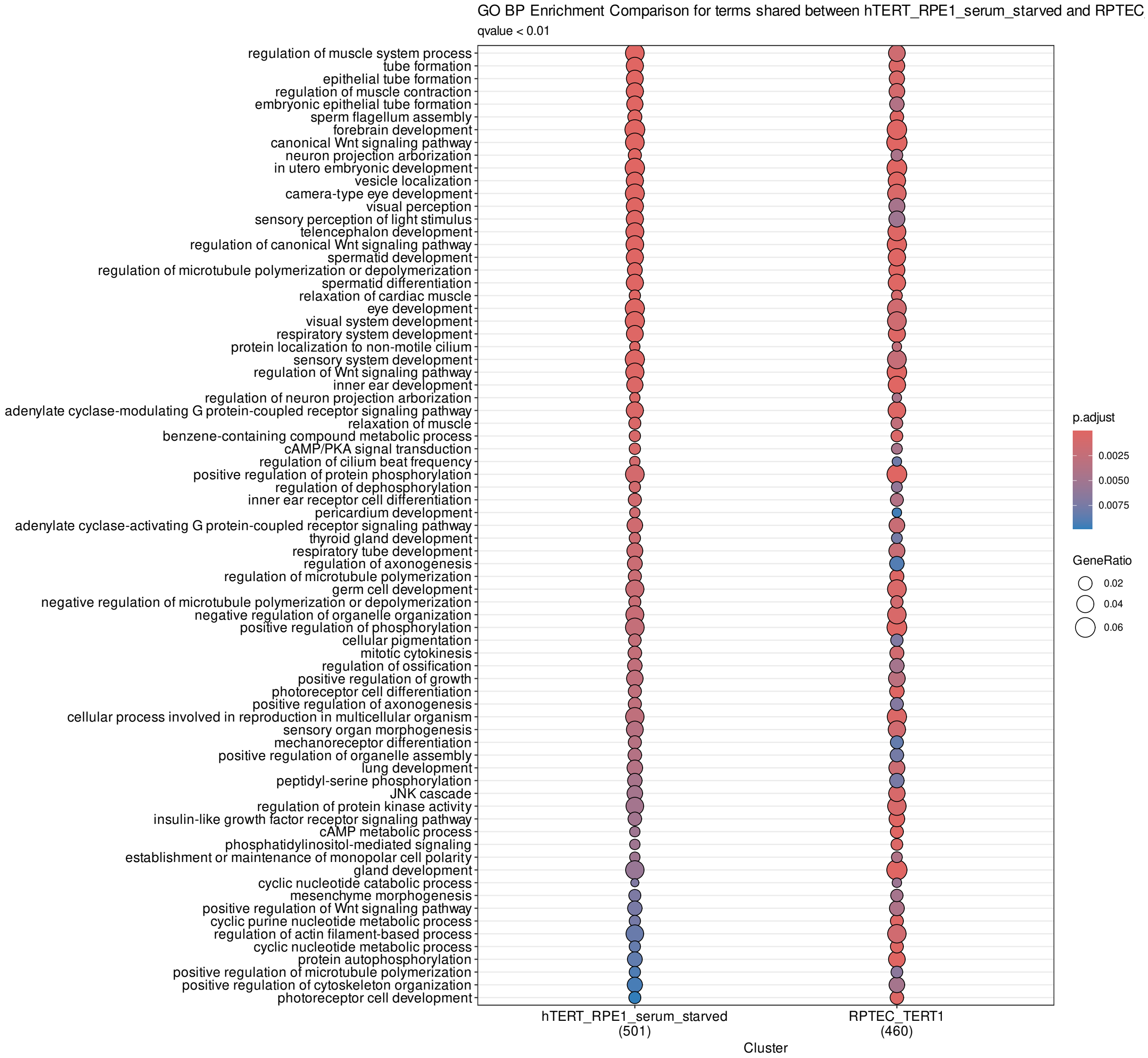
Filter for terms shared between hTERT__RPE1__serum__starved and RPTEC__TERT1

```r
 # get results as data frame
comp_results <- comp@compareClusterResult

# Count occurrences
term_counts <- table(comp_results$ID)

# get terms in all three cell lines
terms_in_all <- names(term_counts[term_counts == 3])

# Remove terms of ASC52telo
comp_results <- comp_results %>% filter(Cluster != "ASC52telo")

# Count occurrences
term_counts <- table(comp_results$ID)

# Filter proteins that appear at least twice
terms_at_least_twice <- names(term_counts[term_counts >= 2])

# remove terms that are enriched in all cell lines from terms_at_least_twice
terms_at_least_twice <- terms_at_least_twice[!terms_at_least_twice %in% terms_in_all]

# remove terms that are enriched in more than one cell line
unspecific_terms <- comp_results[comp_results$ID %in% terms_at_least_twice,]
specific_terms <- comp_results[!comp_results$ID %in% terms_at_least_twice,]

# create a copy of comp
comp_filtered <- comp

# update results in comp
comp_filtered@compareClusterResult <- unspecific_terms
```

```
comp_filtered_RPTEC_TERT1 <- comp_filtered
comp_filtered_hTERT_RPE1 <- comp_filtered

comp_filtered_hTERT_RPE1@compareClusterResult <- comp_filtered@compareClusterResult[comp_filtered@compa
comp_filtered_RPTEC_TERT1@compareClusterResult <- comp_filtered@compareClusterResult[comp_filtered@compa
```

Dot plot of uniquely enriched terms

```
# plot dotplot
dotplot(comp_filtered, showCategory = NULL) + ggtitle("GO BP Enrichment Comparison for terms shared bet
  scale_y_discrete(labels = function(x) str_wrap(x, width = 100))
```



```
# save dotplot as svg file
ggsave(paste0(out_path, "Restricted_comparison_celllines_GO_BP_dotplot_shared_hTERT_RPTEC.svg"), plot =
```

Cluster results - shared between hTERT_RPE1_serum_starved and RPTEC_TERT1

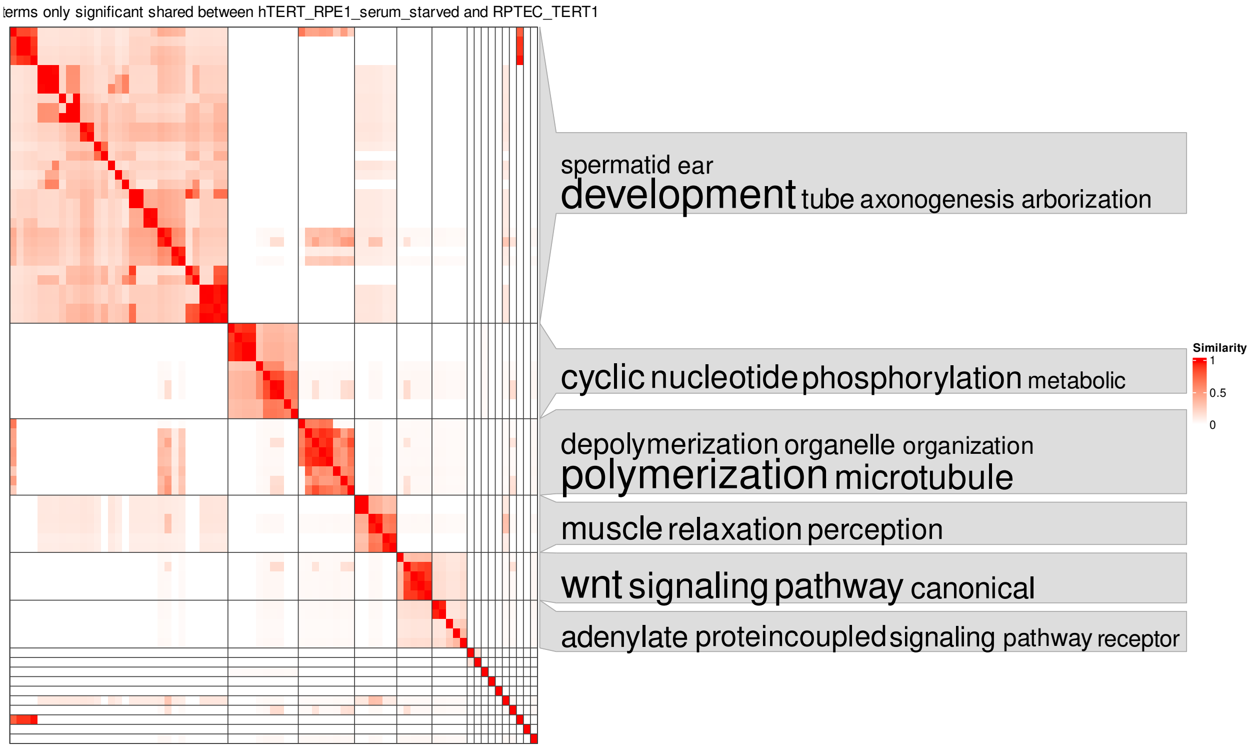
```
go_id = comp_filtered_hTERT_RPE1@compareClusterResult$ID
mat = GO_similarity(go_id,
                    ont = 'BP',
                    db = 'org.Hs.eg.db',
                    measure = "Sim_Relevance_2006")
```

```


# Capture the plot
heatmap_plot <- grid.grabExpr({
 df <- simplifyGO(mat,
             method = 'binary_cut',
             plot = TRUE,
             column_title = "GO BP terms only significant shared between hTERT_RPE1_serum_starved and RI
             use_raster = FALSE,
             order_by_size = TRUE,
             fontsize_range = c(18, 36),
             max_words = 6,
             word_cloud_grob_param = list(col = 'black',
                                          max_width = unit(200, "mm")))
})

# Save the captured plot as an SVG file
svglite(paste0(out_path, "Restricted_comparison_cell_lines_GO_BP_dotplot_shared_hTERT_RPTEC_heatmap.svg
grid.draw(heatmap_plot)
dev.off()
```

Plot cluster heatmap

```
## cairo_pdf
##         2
```

```
grid.newpage()
grid.draw(heatmap_plot)
```



```r
# add cluster number from GO term clustering
results <- map_cluster_number(comp_filtered_hTERT_RPE1,
                              df = df,
                              comp = TRUE
)

# Apply the function to each row of the DataFrame
results$GeneSymbol <- sapply(results$geneID, map_geneID_to_symbol)

# save results as csv file
write.csv(results, file = paste0(out_path, "Restricted_comparison_cell_lines_GO_BP_dotplot_shared_hTERT_
```

```r
# add cluster number from GO term clustering
results <- map_cluster_number(comp_filtered_RPTEC_TERT1,
                              df = df,
                              comp = TRUE
)

# Apply the function to each row of the DataFrame
results$GeneSymbol <- sapply(results$geneID, map_geneID_to_symbol)

# save results as csv file
write.csv(results, file = paste0(out_path, "Restricted_comparison_cell_lines_GO_BP_dotplot_shared_hTERT_
```

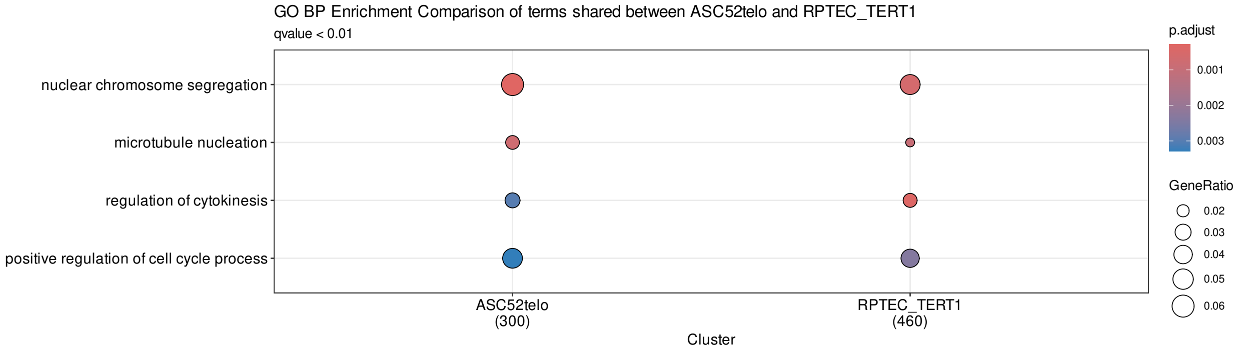
Process and save results

Filter for terms shared between ASC52telo and RPTEC_TERT1

```
 # get results as data frame
comp_results <- comp@compareClusterResult

# Count occurrences
term_counts <- table(comp_results$ID)

# get terms in all three cell lines
terms_in_all <- names(term_counts[term_counts == 3])

# Remove terms of hTERT_RPE1_serum_starved
comp_results <- comp_results %>% filter(Cluster != "hTERT_RPE1_serum_starved")

# Count occurrences
term_counts <- table(comp_results$ID)

# Filter proteins that appear at least twice
terms_at_least_twice <- names(term_counts[term_counts >= 2])

# remove terms that are enriched in all cell lines from terms_at_least_twice
terms_at_least_twice <- terms_at_least_twice[!terms_at_least_twice %in% terms_in_all]

# remove terms that are enriched in more than one cell line
unspecific_terms <- comp_results[comp_results$ID %in% terms_at_least_twice,]
specific_terms <- comp_results[!comp_results$ID %in% terms_at_least_twice,]

# create a copy of comp
comp_filtered <- comp

# update results in comp
comp_filtered@compareClusterResult <- unspecific_terms

comp_filtered_ASC52telo <- comp_filtered
comp_filtered_RPTEC_TERT1 <- comp_filtered

comp_filtered_ASC52telo@compareClusterResult <- comp_filtered@compareClusterResult[comp_filtered@compare
comp_filtered_RPTEC_TERT1@compareClusterResult <- comp_filtered@compareClusterResult[comp_filtered@compa
```

Dot plot of uniquely enriched terms

```
# plot dotplot
dotplot(comp_filtered, showCategory = NULL) + ggtitle("GO BP Enrichment Comparison of terms shared betwe
  scale_y_discrete(labels = function(x) str_wrap(x, width = 100))
```



```
# save dotplot as svg file
ggsave(paste0(out_path, "Restricted_comparison_celllines_GO_BP_dotplot_shared_ASC52telo_RPTEC.svg"), pl
```

Cluster results - shared between hTERT_RPE1_serum_starved and RPTEC_TERT1

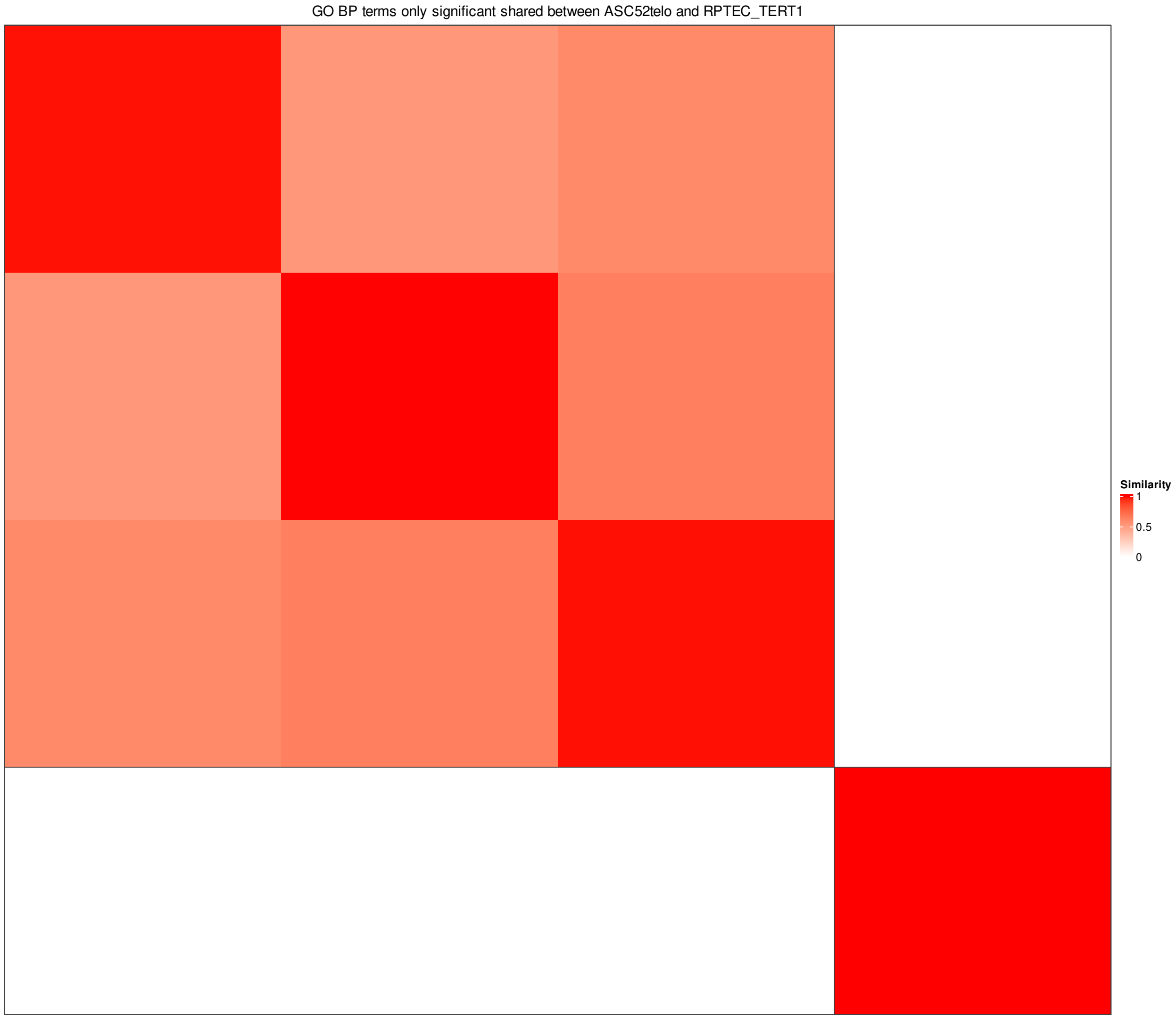
```
go_id = comp_filtered_ASC52telo@compareClusterResult$ID
mat = GO_similarity(go_id,
                    ont = 'BP',
                    db = 'org.Hs.eg.db',
                    measure = "Sim_Relevance_2006")
```

```
# Capture the plot
heatmap_plot <- grid.grabExpr({
 df <- simplifyGO(mat,
             method = 'binary_cut',
             plot = TRUE,
             column_title = "GO BP terms only significant shared between ASC52telo and RPTEC_TERT1",
             use_raster = FALSE,
             order_by_size = TRUE,
             fontsize_range = c(18, 36),
             max_words = 6,
             word_cloud_grob_param = list(col = 'black',
                                          max_width = unit(200, "mm")))
})

# Save the captured plot as an SVG file
svglite(paste0(out_path, "Restricted_comparison_cell_lines_GO_BP_dotplot_shared_ASC52telo_RPTEC_heatmap
grid.draw(heatmap_plot)
dev.off()
```

Plot cluster heatmap

```
## cairo_pdf
##         2
```

```
grid.newpage()
grid.draw(heatmap_plot)
```



```
# add cluster number from GO term clustering
results <- map_cluster_number(comp_filtered_ASC52telo,
                              df = df,
                              comp = TRUE
)

# Apply the function to each row of the DataFrame
results$GeneSymbol <- sapply(results$geneID, map_geneID_to_symbol)

# save results as csv file
write.csv(results, file = paste0(out_path, "Restricted_comparison_cell_lines_GO_BP_dotplot_shared_ASC52t
# add cluster number from GO term clustering
results <- map_cluster_number(comp_filtered_RPTEC_TERT1,
                              df = df,
                              comp = TRUE
)

# Apply the function to each row of the DataFrame
results$GeneSymbol <- sapply(results$geneID, map_geneID_to_symbol)

# save results as csv file
write.csv(results, file = paste0(out_path, "Restricted_comparison_cell_lines_GO_BP_dotplot_shared_ASC52
```

Process and save results

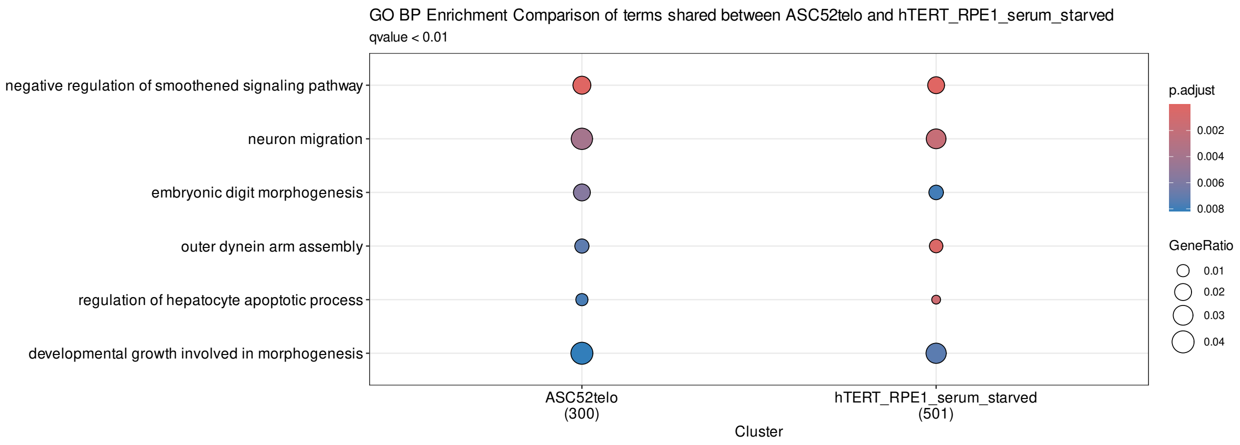
Filter for terms shared between ASC52telo and hTERT_RPE1_serum_starved

```
 # get results as data frame
comp_results <- comp@compareClusterResult

# Count occurrences
term_counts <- table(comp_results$ID)

# get terms in all three cell lines
terms_in_all <- names(term_counts[term_counts == 3])

# Remove terms of hTERT_RPE1_serum_starved
comp_results <- comp_results %>% filter(Cluster != "RPTEC_TERT1")

# Count occurrences
term_counts <- table(comp_results$ID)

# Filter proteins that appear at least twice
terms_at_least_twice <- names(term_counts[term_counts >= 2])

# remove terms that are enriched in all cell lines from terms_at_least_twice
terms_at_least_twice <- terms_at_least_twice[!terms_at_least_twice %in% terms_in_all]

# remove terms that are enriched in more than one cell line
unspecific_terms <- comp_results[comp_results$ID %in% terms_at_least_twice,]
specific_terms <- comp_results[!comp_results$ID %in% terms_at_least_twice,]

# create a copy of comp
comp_filtered <- comp

# update results in comp
comp_filtered@compareClusterResult <- unspecific_terms

comp_filtered_ASC52telo <- comp_filtered
comp_filtered_hTERT_RPE1 <- comp_filtered
```

```
comp_filtered_ASC52telo@compareClusterResult <- comp_filtered@compareClusterResult[comp_filtered@compare
comp_filtered_hTERT_RPE1@compareClusterResult <- comp_filtered@compareClusterResult[comp_filtered@compar
```

Dot plot of uniquely enriched terms

```
# plot dotplot
dotplot(comp_filtered, showCategory = NULL) + ggtitle("GO BP Enrichment Comparison of terms shared betwe
  scale_y_discrete(labels = function(x) str_wrap(x, width = 100))
```



```
# save dotplot as svg file
ggsave(paste0(out_path, "Restricted_comparison_celllines_GO_BP_dotplot_shared_ASC52telo_hTERT.svg"), pl
```

Cluster results - shared between hTERT_RPE1_serum_starved and RPTEC_TERT1

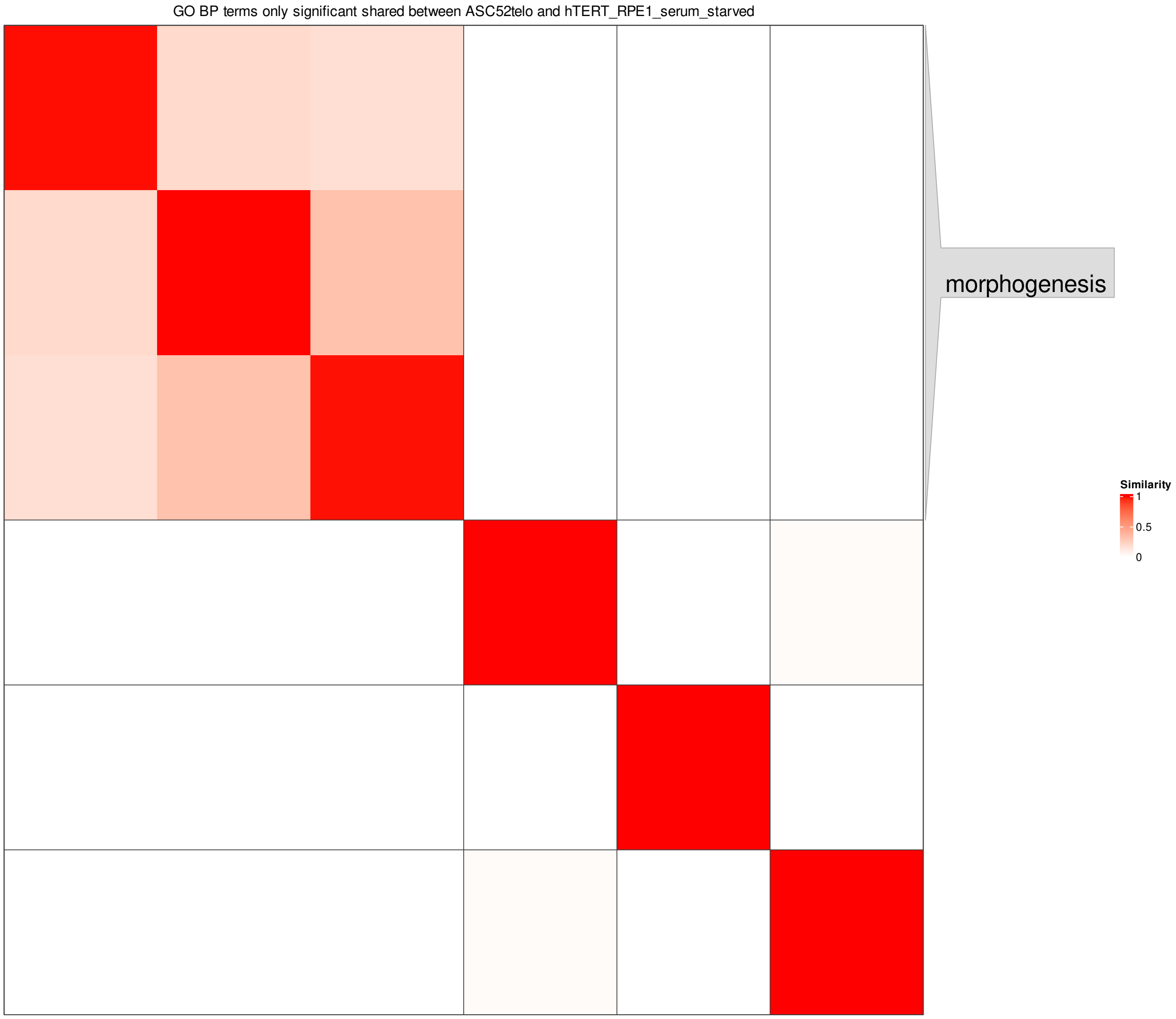
```
go_id = comp_filtered_ASC52telo@compareClusterResult$ID
mat = GO_similarity(go_id,
                    ont = 'BP',
                    db = 'org.Hs.eg.db',
                    measure = "Sim_Relevance_2006")
```

```
# Capture the plot
heatmap_plot <- grid.grabExpr({
 df <- simplifyGO(mat,
             method = 'binary_cut',
             plot = TRUE,
             column_title = "GO BP terms only significant shared between ASC52telo and hTERT_RPE1_serum
             use_raster = FALSE,
             order_by_size = TRUE,
             fontsize_range = c(18, 36),
             max_words = 6,
             word_cloud_grob_param = list(col = 'black',
                                                           max_width = unit(200, "mm")))
})

# Save the captured plot as an SVG file
svglite(paste0(out_path, "Restricted_comparison_cell_lines_GO_BP_dotplot_shared_ASC52telo_hTERT_heatmap
grid.draw(heatmap_plot)
dev.off()
```

Plot cluster heatmap

```
## cairo_pdf
##         2
```

```
grid.newpage()
grid.draw(heatmap_plot)
```



```
# add cluster number from GO term clustering
results <- map_cluster_number(comp_filtered_ASC52telo,
                              df = df,
                              comp = TRUE
)

# Apply the function to each row of the DataFrame
results$GeneSymbol <- sapply(results$geneID, map_geneID_to_symbol)

# save results as csv file
write.csv(results, file = paste0(out_path, "Restricted_comparison_cell_lines_GO_BP_dotplot_shared_ASC52
```

```
# add cluster number from GO term clustering
results <- map_cluster_number(comp_filtered_hTERT_RPE1,
                              df = df,
                              comp = TRUE
)

# Apply the function to each row of the DataFrame
results$GeneSymbol <- sapply(results$geneID, map_geneID_to_symbol)

# save results as csv file
write.csv(results, file = paste0(out_path, "Restricted_comparison_cell_lines_GO_BP_dotplot_shared_ASC52
```

Process and save results