

GAGE analysis of mouse dds output

Load required libraries

```
library(plyr)
library(dplyr)

##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:plyr':
##
##      arrange, count, desc, failwith, id, mutate, rename, summarise,
##      summarize
##
## The following objects are masked from 'package:stats':
##
##      filter, lag
##
## The following objects are masked from 'package:base':
##
##      intersect, setdiff, setequal, union
library(tibble)
library(stringr)
library(biomaRt)
library(genefilter)
library(DESeq2)

## Loading required package: S4Vectors
## Loading required package: stats4
## Loading required package: BiocGenerics
## Loading required package: parallel
##
## Attaching package: 'BiocGenerics'
##
## The following objects are masked from 'package:parallel':
##
##      clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,
##      clusterExport, clusterMap, parApply, parCapply, parLapply,
##      parLapplyLB, parRapply, parSapply, parSapplyLB
##
## The following objects are masked from 'package:dplyr':
##
##      combine, intersect, setdiff, union
##
## The following objects are masked from 'package:stats':
##
##      IQR, mad, sd, var, xtabs
##
## The following objects are masked from 'package:base':
##
##      anyDuplicated, append, as.data.frame, basename, cbind,
##      colMeans, colnames, colSums, dirname, do.call, duplicated,
##      eval, evalq, Filter, Find, get, grep, grepl, intersect,
```

```

##      is.unsorted, lapply, lengths, Map, mapply, match, mget, order,
##      paste, pmax, pmax.int, pmin, pmin.int, Position, rank, rbind,
##      Reduce, rowMeans, rownames, rowSums, sapply, setdiff, sort,
##      table, tapply, union, unique, unsplit, which, which.max,
##      which.min

##
## Attaching package: 'S4Vectors'

## The following objects are masked from 'package:dplyr':
##
##      first, rename

## The following object is masked from 'package:plyr':
##
##      rename

## The following object is masked from 'package:base':
##
##      expand.grid

## Loading required package: IRanges

##
## Attaching package: 'IRanges'

## The following objects are masked from 'package:dplyr':
##
##      collapse, desc, slice

## The following object is masked from 'package:plyr':
##
##      desc

## Loading required package: GenomicRanges
## Loading required package: GenomeInfoDb
## Loading required package: SummarizedExperiment
## 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)".

## Loading required package: DelayedArray
## Loading required package: matrixStats

##
## Attaching package: 'matrixStats'

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

## The following objects are masked from 'package:genefilter':
##
##      rowSds, rowVars

```

```

## The following object is masked from 'package:dplyr':
##
##     count
## The following object is masked from 'package:plyr':
##
##     count
## Loading required package: BiocParallel
##
## Attaching package: 'DelayedArray'
## The following objects are masked from 'package:matrixStats':
##
##     colMaxs, colMins, colRanges, rowMaxs, rowMins, rowRanges
## The following objects are masked from 'package:base':
##
##     aperm, apply
library(gplots)

##
## Attaching package: 'gplots'
## The following object is masked from 'package:IRanges':
##
##     space
## The following object is masked from 'package:S4Vectors':
##
##     space
## The following object is masked from 'package:stats':
##
##     lowess
library(ggplot2)
library(RColorBrewer)
library(stringr)
library(devtools)
library(ggrepel)
library(reshape2)
library(data.table)

##
## Attaching package: 'data.table'
## The following objects are masked from 'package:reshape2':
##
##     dcast, melt
## The following object is masked from 'package:SummarizedExperiment':
##
##     shift
## The following object is masked from 'package:GenomicRanges':
##
##     shift

```

```

## The following object is masked from 'package:IRanges':
##
##      shift
## The following objects are masked from 'package:S4Vectors':
##
##      first, second
## The following objects are masked from 'package:dplyr':
##
##      between, first, last
library(purrr)

##
## Attaching package: 'purrr'
## The following object is masked from 'package:data.table':
##
##      transpose
## The following object is masked from 'package:DelayedArray':
##
##      simplify
## The following object is masked from 'package:GenomicRanges':
##
##      reduce
## The following object is masked from 'package:IRanges':
##
##      reduce
## The following object is masked from 'package:plyr':
##
##      compact
library(gtools)
library(gage)
library(gageData)
library(org.Mm.eg.db)

## Loading required package: AnnotationDbi
##
## Attaching package: 'AnnotationDbi'
## The following object is masked from 'package:dplyr':
##
##      select
##
data(kegg.sets.mm)
data(sigmet.idx.mm)
kegg.sets.mm.s = kegg.sets.mm[sigmet.idx.mm]

```

Purpose

To compare the pathway enrichment within the DGE profiles for murine dermal fibroblasts following poly(I:C) transfection.

```
Mouse <- "Expanded_Design_Factor_Mice"
Mouse_DGEs <- basename(Sys.glob(file.path(Mouse, "*DGE_results.txt")))

exptcounts <- function(files) {
  d <- read.delim(files)
  d
}

Mouse_DGEs_read <- exptcounts(file.path(Mouse, Mouse_DGEs))
```

Performing GAGE analysis and generating output

```
output_dir <- "GAGE_output"

##GAGE analysis
gage_mixed <- function(input) {
  dd <- dplyr::select(input, log2FoldChange, padj, ENTREZID) %>%
    na.omit() %>%
    dplyr::select(log2FoldChange, ENTREZID) %>%
    dplyr::distinct(ENTREZID, .keep_all = TRUE)
  dd$ENTREZID <- gsub(pattern = ",.*", replacement = "", dd$ENTREZID)
  e = dd$log2FoldChange
  names(e) = dd$ENTREZID
  ##As per GAGE manual suggestion, looking at same.dir = FALSE since genes in pathways
  ##do not tend to just increase or just decrease.
  ef <- gage(na.omit(e), gsets = kegg.sets.mm.s, same.dir = FALSE)
  g <- ef$greater
  gg <- g[mixedorder(rownames(g), decreasing = TRUE),] %>%
    as.data.frame() %>%
    rownames_to_column(., var = "pathway")
  gg
}

GAGE_heatmap <- function(DGEs, name) {
  DGEs_GAGE <- gage_mixed(DGEs) %>%
    dplyr::select(pathway, q.val)

  write.table(DGEs_GAGE, file = file.path(output_dir, paste0(Sys.Date(),
    name, ".txt")), row.names=FALSE, sep="\t", quote=FALSE)

  DGEs_GAGE_sub <- dplyr::filter_at(DGEs_GAGE, vars(-pathway), any_vars(. <= 0.09))
  write.table(DGEs_GAGE_sub, file = file.path(output_dir, paste0(Sys.Date(),
    name, "Significant.txt")), row.names=FALSE, sep="\t", quote=FALSE)

  ##Now make heatmap
  gage.m <- melt(DGEs_GAGE_sub, id = 'pathway')
  gage.m$value <- -log10(gage.m$value)
  gage.m[order(gage.m$value, decreasing = TRUE)]
  ##the more significant, the higher the value.
```

```

##Color switch is is at q.val ~ 0.09
pdf(file = file.path(output_dir, paste0(Sys.Date(), name, "Heatmap.pdf")), height = 2, width = 3.5)
plot <- ggplot(gage.m, aes(gage.m$variable, gage.m$pathway))
p <- plot + geom_tile(color = 'black', aes(fill = gage.m$value)) +
  theme_bw(base_size = 6) +
  theme(axis.text.x = element_text(angle = 90, hjust = 1),
        axis.text.y = element_text(size = rel(1.5))) +
  scale_fill_gradient2(low = "#e66101", high = "#542788", mid = "white", midpoint = 1.046, limits = c(5
  ylab("") +
  xlab("\n")
print(p)
dev.off()
}

GAGE_heatmap(Mouse_DGEs_read, "MouseGAGE")

```

```

## pdf
## 2

```

Session Info

```
sessionInfo()
```

```

## R version 3.5.2 (2018-12-20)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: macOS Sierra 10.12.6
##
## Matrix products: default
## BLAS: /Library/Frameworks/R.framework/Versions/3.5/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/3.5/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] parallel stats4 stats graphics grDevices utils datasets
## [8] methods base
##
## other attached packages:
## [1] bindrcpp_0.2.2 org.Mm.eg.db_3.7.0
## [3] AnnotationDbi_1.44.0 gageData_2.20.0
## [5] gage_2.32.1 gtools_3.8.1
## [7] purrr_0.2.5 data.table_1.12.0
## [9] reshape2_1.4.3 ggrepel_0.8.0
## [11] usethis_1.4.0 devtools_2.0.1
## [13] RColorBrewer_1.1-2 ggplot2_3.1.0
## [15] gplots_3.0.1 DESeq2_1.22.2
## [17] SummarizedExperiment_1.12.0 DelayedArray_0.8.0
## [19] BiocParallel_1.16.5 matrixStats_0.54.0
## [21] Biobase_2.42.0 GenomicRanges_1.34.0
## [23] GenomeInfoDb_1.18.1 IRanges_2.16.0
## [25] S4Vectors_0.20.1 BiocGenerics_0.28.0
## [27] genefilter_1.64.0 biomaRt_2.38.0
## [29] stringr_1.3.1 tibble_2.0.1
## [31] dplyr_0.7.8 plyr_1.8.4

```

```
##
## loaded via a namespace (and not attached):
## [1] colorspace_1.4-0      rprojroot_1.3-2      htmlTable_1.13.1
## [4] XVector_0.22.0        base64enc_0.1-3      fs_1.2.6
## [7] rstudioapi_0.9.0      remotes_2.0.2        bit64_0.9-7
## [10] splines_3.5.2         geneplotter_1.60.0   knitr_1.21
## [13] pkgload_1.0.2         Formula_1.2-3        annotate_1.60.0
## [16] cluster_2.0.7-1       png_0.1-7            graph_1.60.0
## [19] compiler_3.5.2        httr_1.4.0           backports_1.1.3
## [22] assertthat_0.2.0      Matrix_1.2-15        lazyeval_0.2.1
## [25] cli_1.0.1             acepack_1.4.1        htmltools_0.3.6
## [28] prettyunits_1.0.2     tools_3.5.2          gtable_0.2.0
## [31] glue_1.3.0            GenomeInfoDbData_1.2.0 Rcpp_1.0.0
## [34] Biostrings_2.50.2     gdata_2.18.0         xfun_0.4
## [37] ps_1.3.0              XML_3.98-1.16        zlibbioc_1.28.0
## [40] scales_1.0.0          hms_0.4.2            yaml_2.2.0
## [43] memoise_1.1.0         gridExtra_2.3        rpart_4.1-13
## [46] latticeExtra_0.6-28  stringi_1.2.4        RSQLite_2.1.1
## [49] desc_1.2.0            checkmate_1.9.1      caTools_1.17.1.1
## [52] pkgbuild_1.0.2        rlang_0.3.1          pkgconfig_2.0.2
## [55] bitops_1.0-6          evaluate_0.12        lattice_0.20-38
## [58] bindr_0.1.1           labeling_0.3          htmlwidgets_1.3
## [61] bit_1.1-14            tidyselect_0.2.5     processx_3.2.1
## [64] magrittr_1.5          R6_2.3.0             Hmisc_4.1-1
## [67] DBI_1.0.0             pillar_1.3.1         foreign_0.8-71
## [70] withr_2.1.2           survival_2.43-3      KEGGREST_1.22.0
## [73] RCurl_1.95-4.11       nnet_7.3-12          crayon_1.3.4
## [76] KernSmooth_2.23-15    rmarkdown_1.11       progress_1.2.0
## [79] locfit_1.5-9.1        grid_3.5.2           blob_1.1.1
## [82] callr_3.1.1           digest_0.6.18        xtable_1.8-3
## [85] munsell_0.5.0         sessioninfo_1.1.1
```