De-regulated genes shared between Mouse and Humans Adult and Pediatric

Once the analysis of de-regulated genes during CHIKV infection in mouse and adult and pediatric humans is ready it is possibly to compare the list to see which genes are show the same patter of expression between these samples and those who has different patterns.

For the comparison with mouse genes is necessary first to identify the correspondent orthologs genes.

Venn Diagram comparing downregulated genes between Humans Pediatric and Adult

```r
rm(list = ls())
library(RColorBrewer)
library(VennDiagram)
```

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

```
## Loading required package: futile.logger
```

```r
library(GeneOverlap)
library(eulerr)
```

```
## Registered S3 method overwritten by 'eulerr':
##   method    from
##   plot.venn gplots
```

```r
orth = read.delim('diff_expression_mouse/hsapiens_genes_orthologs_mouse.txt')

pediatric = read.delim('diff_expression_pediatric_cohort/diff_express_analysis_pediatric_cohort.tsv')
adult = read.delim('diff_expression_adult_cohort/diff_express_analysis_adult_cohort.txt')
mousefeet  = read.delim('diff_expression_mouse/diff_express_analysis_ctrl_f2.vs.inf_f2_with_orth')
mouseln  = read.delim('diff_expression_mouse/diff_express_analysis_ctrl_ln2.vs.inf_ln2_with_orth')

pediatricdown = pediatric[pediatric$logFC <= -1 & pediatric$FDR < 0.05,]
adultdown = adult[adult$logFC <= -1 & adult$FDR < 0.05,]
mousefeetdown = mousefeet[mousefeet$logFC <= -1 & mousefeet$FDR < 0.05,]
mouselndown = mouseln[mouseln$logFC <= -1 & mouseln$FDR < 0.05,]

pediatricup = pediatric[pediatric$logFC >= 1 & pediatric$FDR < 0.05,]
adultup = adult[adult$logFC >= 1 & adult$FDR < 0.05,]
mousefeetup = mousefeet[mousefeet$logFC >= 1 & mousefeet$FDR < 0.05,]
mouselnup = mouseln[mouseln$logFC >= 1 & mouseln$FDR < 0.05,]

#n = length(which((rownames(adultdown) %in% rownames(pediatricdown)) == T))

#Downregulated genes
set1down = rownames(pediatricdown)
set2down = rownames(adultdown)
set3down = as.vector(mousefeetdown$Gene.stable.ID)
set3down[is.na(set3down)] = '-'
set4down = as.vector(mouselndown$Gene.stable.ID)
set4down[is.na(set4down)] = '-'

#venn.diagram(x = list(set1down,set2down,set3down),
#             category.names = c('Pediatric','Adult','Mouse - Feet'),
```

```
#               main = 'Downregulated genes',
#               imagetype="png",
#               filename = 'downregulated_venn_diagram_Adult_Child_Mousefeet.png',
#               output=T,
#               lty = 2,
#               fill = brewer.pal(3, "Pastel2"))

#venn.diagram(x = list(set1down,set2down,set4down),
#               category.names = c('Pediatric','Adult','Mouse - LN'),
#               main = 'Downregulated genes',
#               imagetype="png",
#               filename = 'downregulated_venn_diagram_Adult_Child_MouseLN.png',
#               output=T,
#               lty = 2,
#               fill = brewer.pal(3, "Pastel2"))

#Upregulated genes
set1up = rownames(pediatricup)
set2up = rownames(adultup)
set3up = as.vector(mousefeetup$Gene.stable.ID)
set3up[is.na(set3up)] = '-'
set4up = as.vector(mouselnup$Gene.stable.ID)
set4up[is.na(set4up)] = '-'

#venn.diagram(x = list(set1up,set2up,set3up),
#               category.names = c('Pediatric','Adult','Mouse - Feet'),
#               main = 'Upregulated genes',
#               imagetype="png",
#               filename = 'upregulated_venn_diagram_Adult_Child_Mousefeet.png',
#               output=T,
#               lty = 2,
#               fill = brewer.pal(3, "Pastel2"))

#venn.diagram(x = list(set1up,set2up,set4up),
#               category.names = c('Pediatric','Adult','Mouse - LN'),
#               main = 'Upregulated genes',
#               imagetype="png",
#               filename = 'upregulated_venn_diagram_Adult_Child_MouseLN.png',
#               output=T,
#               lty = 2,
#               fill = brewer.pal(3, "Pastel2"))
```

Consideration after plot Venn Diagrams Venn Diagrams are ready but they don't show well the proportion of diff expressed genes at each dataset and the proportion of overlap between datasets (besides that the number seems to be wrong). I prepared new Euler diagrams with the same information in order to better represent the proportions.

```
#Upregulated genes Mouse Feet vs Adult vs Child
area1 = length(set1up)
area2 = length(set2up)
area3 = length(set3up)
n12 = length(intersect(set1up, set2up))
n23 = length(intersect(set2up, set3up))
n13 = length(intersect(set1up, set3up))
```

```r
n123 = length(intersect(intersect(set1up, set2up),set3up))

fit = euler(c('A' = area1-((n12+n13)-n123),
              'B' = area2-((n12+n23)-n123),
              'C' = area3-((n13+n23)-n123),
              'A&B' = n12 - n123,
              'B&C' = n23 - n123,
              'A&C'= n13 - n123,
              'A&B&C'= n123),
            shape='ellipse',
            quantities=T)
png('upregulated_euler_diagram_Adult_Child_Mousefeet.png',width=865,height = 575)
plot(fit,
     lty = 2,
     fills = brewer.pal(3, "Pastel2"),
     quantities = list(fontsize = 13),
     labels = c('Child','Adult','Mouse - Feet'),
     cex =2,
     main = 'Upregulated Genes')
dev.off()
```

```
## pdf
##   2
```

```r
#Upregulated genes Mouse Linfonode vs Adult vs Child
area1 = length(set1up)
area2 = length(set2up)
area3 = length(set4up)
n12 = length(intersect(set1up, set2up))
n23 = length(intersect(set2up, set4up))
n13 = length(intersect(set1up, set4up))
n123 = length(intersect(intersect(set1up, set2up),set4up))

fit = euler(c('A' = area1-((n12+n13)-n123),
              'B' = area2-((n12+n23)-n123),
              'C' = area3-((n13+n23)-n123),
              'A&B' = n12 - n123,
              'B&C' = n23 - n123,
              'A&C'= n13 - n123,
              'A&B&C'= n123),
            shape='ellipse',
            quantities=T)
png('upregulated_euler_diagram_Adult_Child_MouseLN.png',width=865,height = 575)
plot(fit,
     lty = 2,
     fills = brewer.pal(3, "Pastel2"),
     quantities = list(fontsize = 13),
     labels = c('Child','Adult','Mouse - LN'),
     cex =2,
     main = 'Upregulated Genes')
dev.off()
```

```
## pdf
##   2
```

```r
#Downregulated genes Mouse Feet vs Adult vs Child
area1 = length(set1down)
area2 = length(set2down)
area3 = length(set3down)
n12 = length(intersect(set1down, set2down))
n23 = length(intersect(set2down, set3down))
n13 = length(intersect(set1down, set3down))
n123 = length(intersect(intersect(set1down, set2down),set3down))

fit = euler(c('A' = area1-((n12+n13)-n123),
              'B' = area2-((n12+n23)-n123),
              'C' = area3-((n13+n23)-n123),
              'A&B' = n12 - n123,
              'B&C' = n23 - n123,
              'A&C'= n13 - n123,
              'A&B&C'= n123),
            shape='ellipse',
            quantities=T)
png('downregulated_euler_diagram_Adult_Child_Mousefeet.png',width=865,height = 575)
plot(fit,
     lty = 2,
     fills = brewer.pal(3, "Pastel2"),
     quantities = list(fontsize = 13),
     labels = c('Child','Adult','Mouse - Feet'),
     cex =2,
     main = 'Downregulated Genes')
dev.off()
```

```
## pdf
##   2
```

```r
#Downregulated genes Mouse Linfonode vs Adult vs Child
area1 = length(set1down)
area2 = length(set2down)
area3 = length(set4down)
n12 = length(intersect(set1down, set2down))
n23 = length(intersect(set2down, set4down))
n13 = length(intersect(set1down, set4down))
n123 = length(intersect(intersect(set1down, set2down),set4down))

fit = euler(c('A' = area1-((n12+n13)-n123),
              'B' = area2-((n12+n23)-n123),
              'C' = area3-((n13+n23)-n123),
              'A&B' = n12 - n123,
              'B&C' = n23 - n123,
              'A&C'= n13 - n123,
              'A&B&C'= n123),
            shape='ellipse',
            quantities=T)
png('downregulated_euler_diagram_Adult_Child_MouseLN.png',width=865,height = 575)
plot(fit,
     lty = 2,
     fills = brewer.pal(3, "Pastel2"),
     quantities = list(fontsize = 13),
```

```
    labels = c('Child','Adult','Mouse - LN'),
    cex =2,
    main = 'Downregulated Genes')
dev.off()
```

```
## pdf
##    2
```

Test Gene Overlap between datasets using GeneOverlap R package

```
adult = list(AdultUp = set2up, AdultDown = set2down)
child = list(ChildUp = set1up, ChildDown = set1down)

adultchild = list(ChildUp = set1up, AdultUp = set2up, ChildDown = set1down, AdultDown = set2down)
mouse  = list(MouseFeetUp = set3up, MouseLNUP = set4up, MouseFeetDown = set3down, MouseLNDown = set4down

### Mouse vs Adult+Child
#number of human genes with a ortholog in mouse
n = length(unique((orth[orth$Mouse.gene.stable.ID != '',1])))
gom.obj <- newGOM(adultchild, mouse, n )
drawHeatmap(gom.obj, adj.p = T)
```
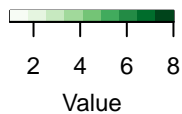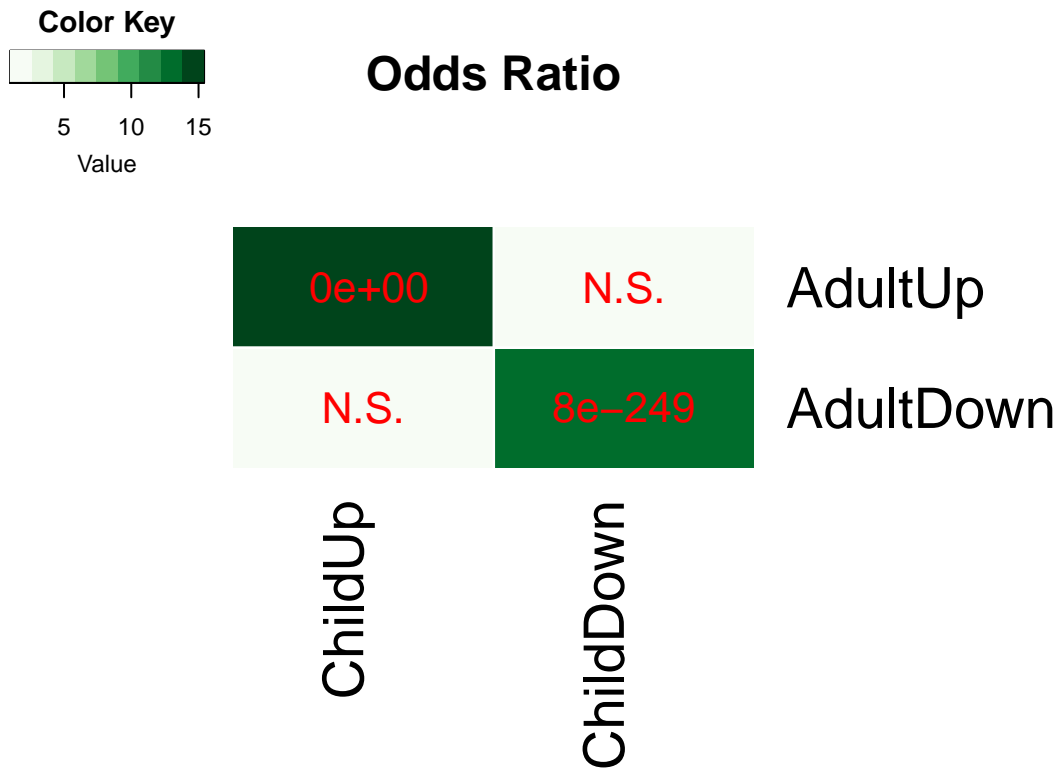


```
### Ault vs Child
# number of genes in Human annotation file Homo_sapiens.GRCh38.89.gtf
n = 58233
gom.obj <- newGOM(adult, child, n )
drawHeatmap(gom.obj, adj.p = T)
```

Color Key

Odds Ratio

0e+00	N.S.	AdultUp
N.S.	8e−249	AdultDown
ChildUp	ChildDown	

N.S.: Not Significant; −−: Ignored

Preliminar Conclusion

Two analysis were conducted qith diff expressed genes between the diferent datasets (Child, Adult, Mouse/feet and Mouse Linfonode). First I prepared Venn diagrams to show the overlap between de-regulated genes between datasets, based on that I also performed a calculation of association between the diff express genes with GeneOverlap Package. There is no association between downregulated genes in mouse samples and adult or child samples. It possible to detect a valid association in upregulated genes between mouse and human samples, and this association is stronger betwen mouse and child samples, and also between adult and child data when considering up and down regulated genes.

Important consideration Until now I'm just performing analysis with diff expressed genes, identified with analysis based on different control samples. Variations on controls can lead a variations on relative expression results. The next step should be compare the normalized expression values of genes between the datasets.