looking at cons markers and DE between species

```r
setwd("/Users/vh3/Documents/PfMCA/ANALYSIS_2/pbpf")
library(scater, quietly = TRUE)
library(scmap)
library(Seurat)
library(scater)
library(scran)
library(cowplot)
library(RColorBrewer)
library(pheatmap)
library(scales)
```

```r
p.combined <- readRDS("pbpfcombo_SeuratInt_20200702.rds")
p.sce <- as.SingleCellExperiment(p.combined)
mf <- read.csv("male_female_specific.csv", header = TRUE)
mf <- mf[, 1:2]
```

```r
male.markers <- read.csv("/Users/vh3/Documents/PfMCA/ANALYSIS_2/pbpf/markers/male_markers.csv",
    header = TRUE, row.names = 1)
male.markers$stage <- rep("male", length(male.markers$Pbe_p_val))

female.markers <- read.csv("/Users/vh3/Documents/PfMCA/ANALYSIS_2/pbpf/markers/female_markers.csv",
    header = TRUE, row.names = 1)
female.markers$stage <- rep("female", length(female.markers$Pbe_p_val))

ook.markers <- read.csv("/Users/vh3/Documents/PfMCA/ANALYSIS_2/pbpf/markers/ook_markers.csv",
    header = TRUE, row.names = 1)
ook.markers$stage <- rep("ook", length(ook.markers$Pbe_p_val))

spz.markers <- read.csv("/Users/vh3/Documents/PfMCA/ANALYSIS_2/pbpf/markers/spz1_markers.csv",
    header = TRUE, row.names = 1)
spz.markers$stage <- rep("spz", length(spz.markers$Pbe_p_val))

m5 <- male.markers[1:5, ]
f5 <- female.markers[1:5, ]
o5 <- ook.markers[1:5, ]
s5 <- spz.markers[1:5, ]

all <- rbind(m5, f5)
all <- rbind(all, o5)
all <- rbind(all, s5)
```

```r
p.sce.m <- p.sce[rownames(p.sce) %in% rownames(all), ]
```

```r
lcpm_mat <- as.data.frame(assays(p.sce.m)[["logcounts"]])
tmat <- t(lcpm_mat)
tmat <- as.data.frame(tmat)
```

```r
tmat$seurat_cluster <- as.factor(p.sce$seurat_clusters)
tmat <- as.data.frame(tmat)
# lcpm_mat_sub <- lcpm_mat[!is.na(lcpm_mat$Cluster_k15), ]

clustmean <- aggregate(tmat[, 1:20], by = list(as.factor(as.character(tmat$seurat_cluster))),
    mean)
rownames(clustmean) <- clustmean$Group.1

clustmean2 <- clustmean[, 2:21]

tcm2 <- as.data.frame(t(clustmean2))

pheatmap(tcm2, show_colnames = TRUE, show_rownames = TRUE, cluster_cols = FALSE,
    color = brewer.pal(9, "PuRd"))
```
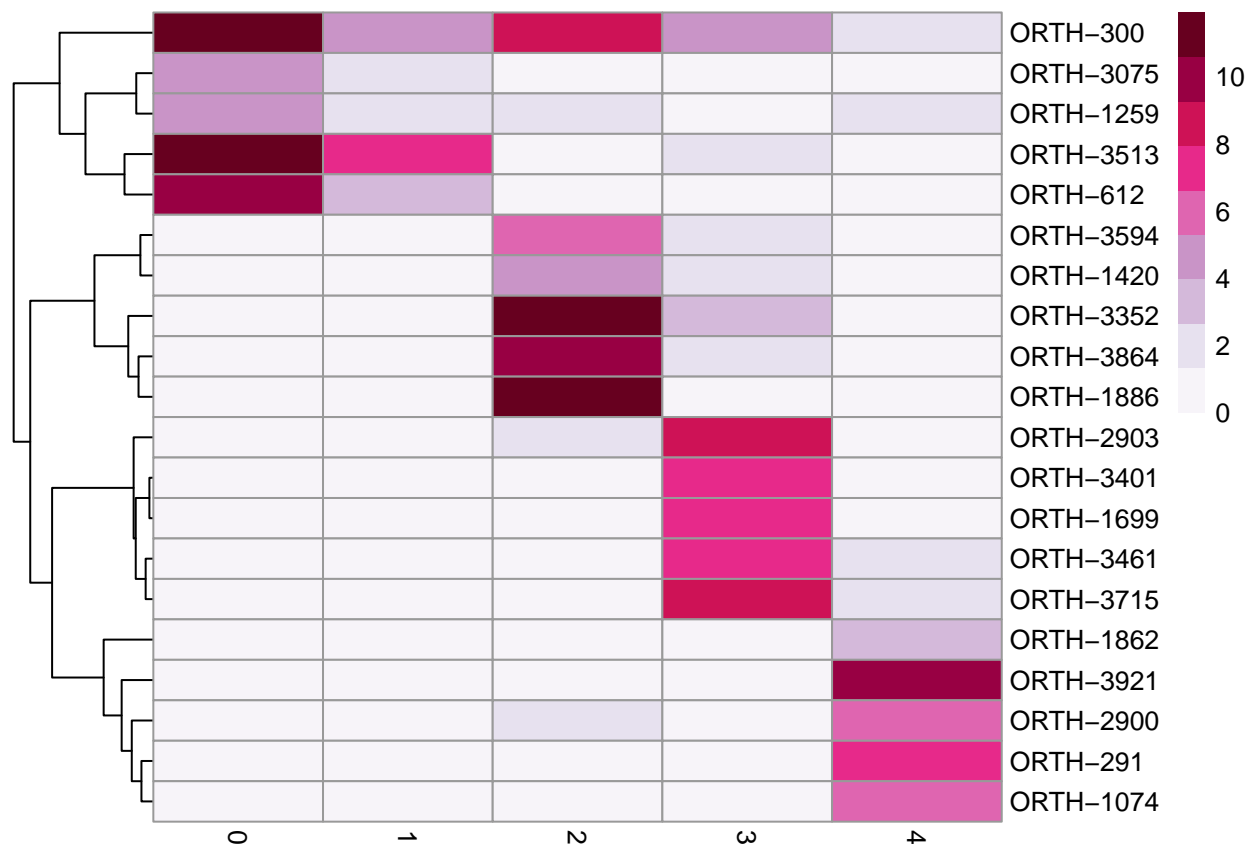


```r
# stage <- as.data.frame(colData(mca.qc.sub)['stage']) cluster <-
# as.data.frame(colData(mca.qc.sub)['seurat_clusters'])


# pheatmap(clustmean2, annotation_col = stage, show_colnames = FALSE,
# show_rownames = TRUE, color = brewer.pal(9, 'PuRd'))

all$pbpfgene <- paste(all$pfgene, all$pbgene, sep = "    ")
all$orthname <- rownames(all)

p.sce.m <- p.sce[rownames(p.sce) %in% rownames(all), ]
```

```r
rowData(p.sce.m)$pbpfgene <- all[match(rownames(rowData(p.sce.m)), all[, 17]), 16]

rowData(p.sce.m)$pfgene <- all[match(rownames(rowData(p.sce.m)), all[, 17]), 14]

geneanno <- read.csv("/Users/vh3/Documents/allpf_geneanno.csv", header = TRUE)
rowData(p.sce.m)$gene_name <- geneanno[match(rowData(p.sce.m)$pfgene, geneanno[, 1]), 5]
rowData(p.sce.m)$product_desc <- geneanno[match(rowData(p.sce.m)$pfgene, geneanno[, 1]), 3]
rowData(p.sce.m)$genesanno <- paste(rowData(p.sce.m)$pbpfgene, rowData(p.sce.m)$gene_name, sep="   ")
rowData(p.sce.m)$name_desc <- paste(rowData(p.sce.m)$pfgene, rowData(p.sce.m)$gene_name, rowData(p.sce.

rd <- read.csv("pscemrowdata_mananno.csv", header=TRUE, row.names = 1)

rowData(p.sce.m) <- rd

lcpm_mat <- as.data.frame(assays(p.sce.m)[["logcounts"]])
rownames(lcpm_mat) <- rowData(p.sce.m)$man_anno
tmat <- t(lcpm_mat)
tmat <- as.data.frame(tmat)

tmat$seurat_cluster <- as.factor(p.sce.m$seurat_clusters)
tmat <- as.data.frame(tmat)
#lcpm_mat_sub <- lcpm_mat[!is.na(lcpm_mat$Cluster_k15), ]

#colnames(tmat) <- rowData(p.sce.m)$pbpfgene

clustmean <- aggregate(tmat[, 1:20], by = list(as.factor(as.character(tmat$seurat_cluster))),
    mean)
rownames(clustmean) <- clustmean$Group.1

clustmean2 <- clustmean[, 2:21]

tcm2 <- as.data.frame(t(clustmean2))


pheatmap(tcm2, show_colnames = TRUE, show_rownames = TRUE, cluster_cols = FALSE,
    color = brewer.pal(9, "PuRd"))
```
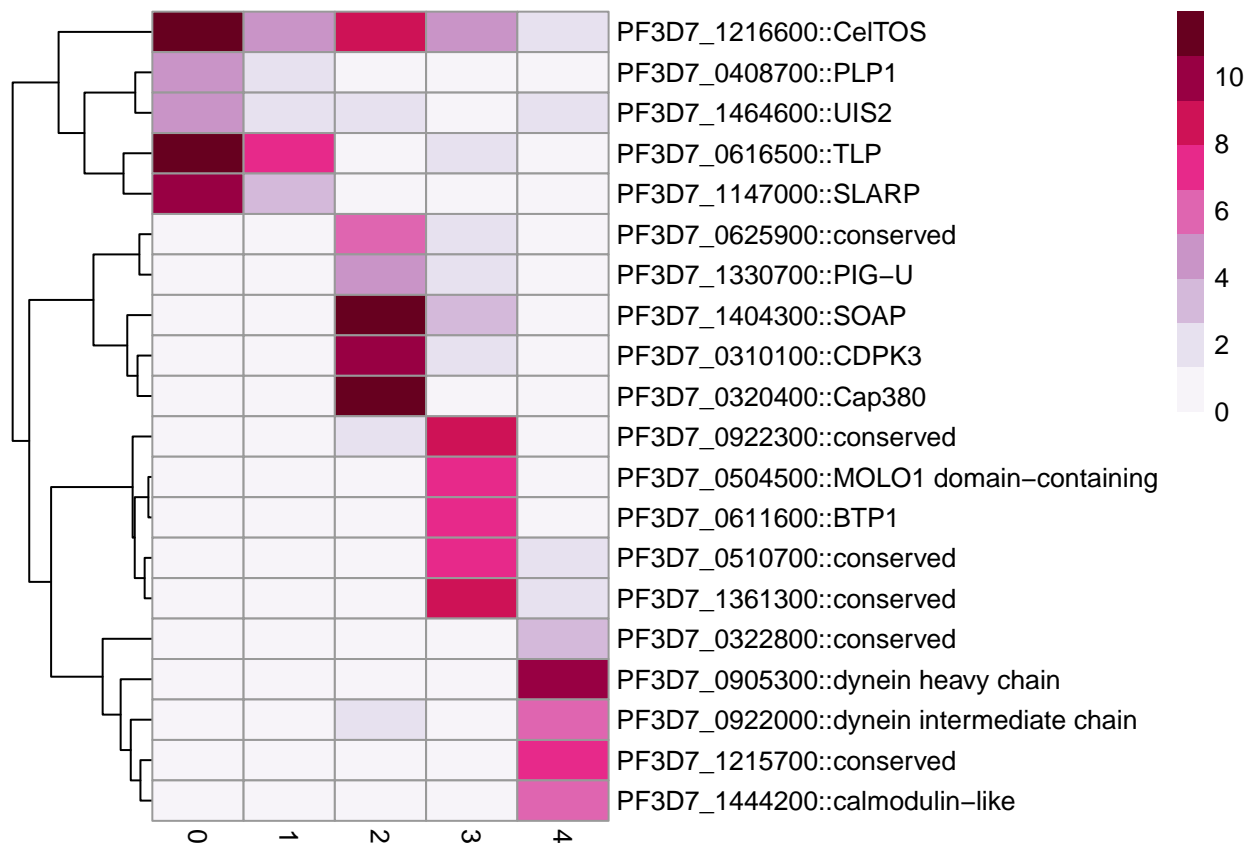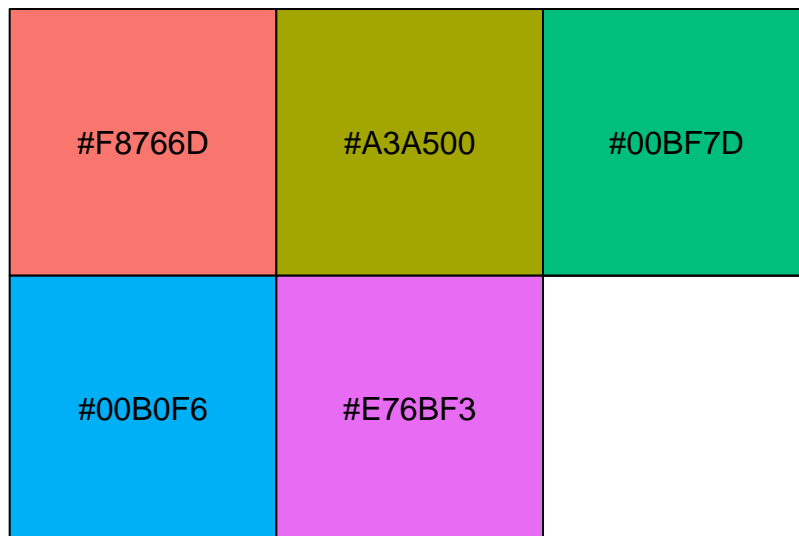
```
show_col(hue_pal()(5))
```



```r
cols <- c("spz"="#F8766D", "ook"="#00BF7D", "female"="#00B0F6", "male"="#E76BF3")

rowData(p.sce.m)$stage <- all[match(rownames(rowData(p.sce.m)), all[, 17]), 15]
#rowData(p.sce.m)$orthname <- rownames(rowData(p.sce.m))
rownames(rowData(p.sce.m)) <- rowData(p.sce.m)$man_anno

rd <- as.data.frame(rowData(p.sce.m))
rownames(rd) <- rd$man_anno
```
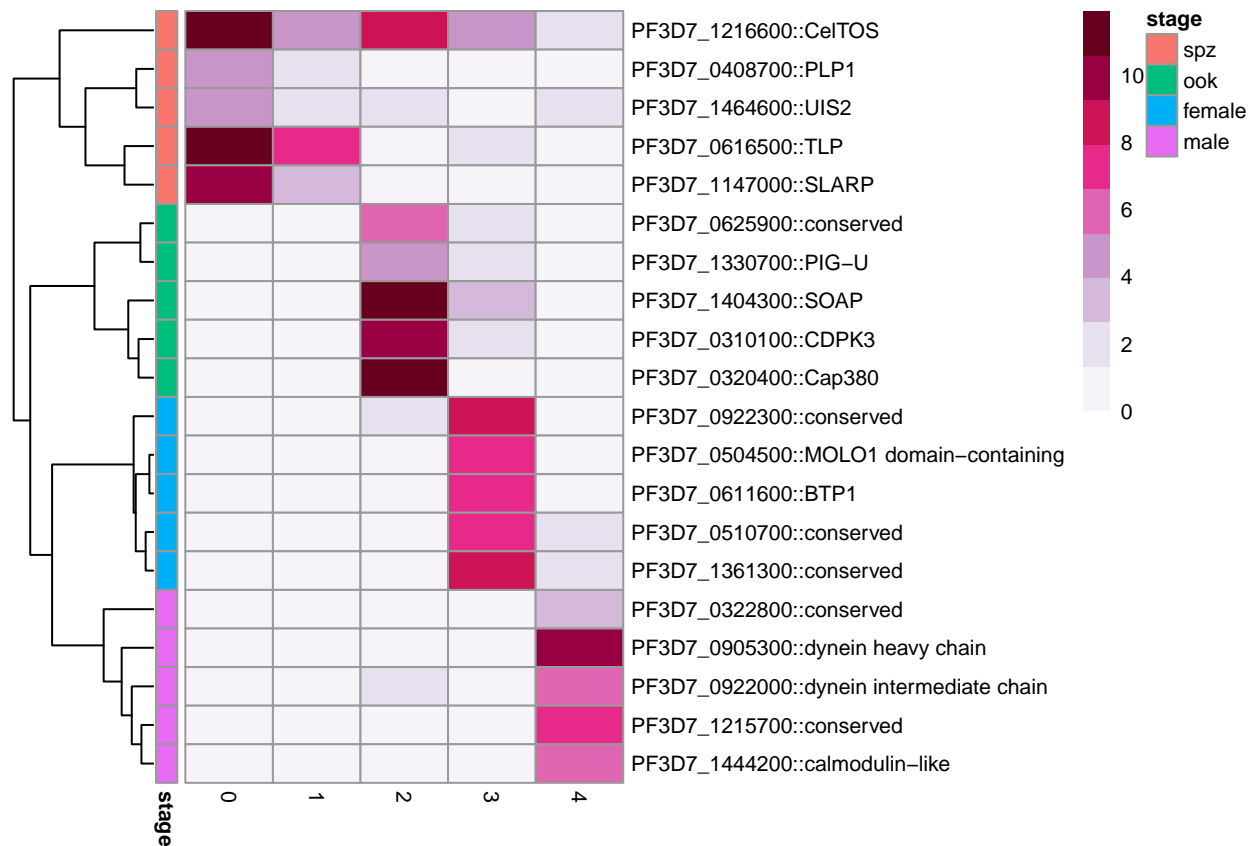
```r
stage <- rd["stage"]

ann_c <- list(stage = cols)

pheatmap(tcm2, show_colnames = TRUE, show_rownames = TRUE, cluster_cols = FALSE,
    color = brewer.pal(9, "PuRd"), annotation_row=stage, fontsize = 8, annotation_colors = ann_c)
```



```r
male.markers <- read.csv("/Users/vh3/Documents/PfMCA/ANALYSIS_2/pbpf/markers/male_pf_specific.csv", head
male.markers$stage <- rep("male", length(male.markers$p_val))
female.markers <- read.csv("/Users/vh3/Documents/PfMCA/ANALYSIS_2/pbpf/markers/female_pf_specific.csv",
female.markers$stage <- rep("female", length(female.markers$p_val))

ook.markers <- read.csv("/Users/vh3/Documents/PfMCA/ANALYSIS_2/pbpf/markers/ook_pf_specific.csv", header
ook.markers$stage <- rep("ook", length(ook.markers$p_val))

spz.markers <- read.csv("/Users/vh3/Documents/PfMCA/ANALYSIS_2/pbpf/markers/spz1_pf_specific.csv", heade
spz.markers$stage <- rep("spz", length(spz.markers$p_val))

m5 <- male.markers[1:5, ]
f5 <- female.markers[1:5, ]
o5 <- ook.markers[1:5, ]
s5 <- spz.markers[1:1, ]

all <- rbind(s5, o5)
all <- rbind(all, f5)
all <- rbind (all, m5)
```

```r
p.sce.m <- p.sce[rownames(p.sce) %in% rownames(all), ]

lcpm_mat <- as.data.frame(assays(p.sce.m)[["logcounts"]])
tmat <- t(lcpm_mat)
tmat <- as.data.frame(tmat)

tmat$seurat_cluster <- as.factor(p.sce$ident)
tmat <- as.data.frame(tmat)
#lcpm_mat_sub <- lcpm_mat[!is.na(lcpm_mat$Cluster_k15), ]

clustmean <- aggregate(tmat[, 1:16], by = list(as.factor(as.character(tmat$seurat_cluster))),
    mean)
rownames(clustmean) <- clustmean$Group.1

clustmean2 <- clustmean[, 2:17]

tcm2 <- as.data.frame(t(clustmean2))


pheatmap(tcm2, show_colnames = TRUE, show_rownames = TRUE, cluster_cols = FALSE,
    color = brewer.pal(9, "PuRd"))
```
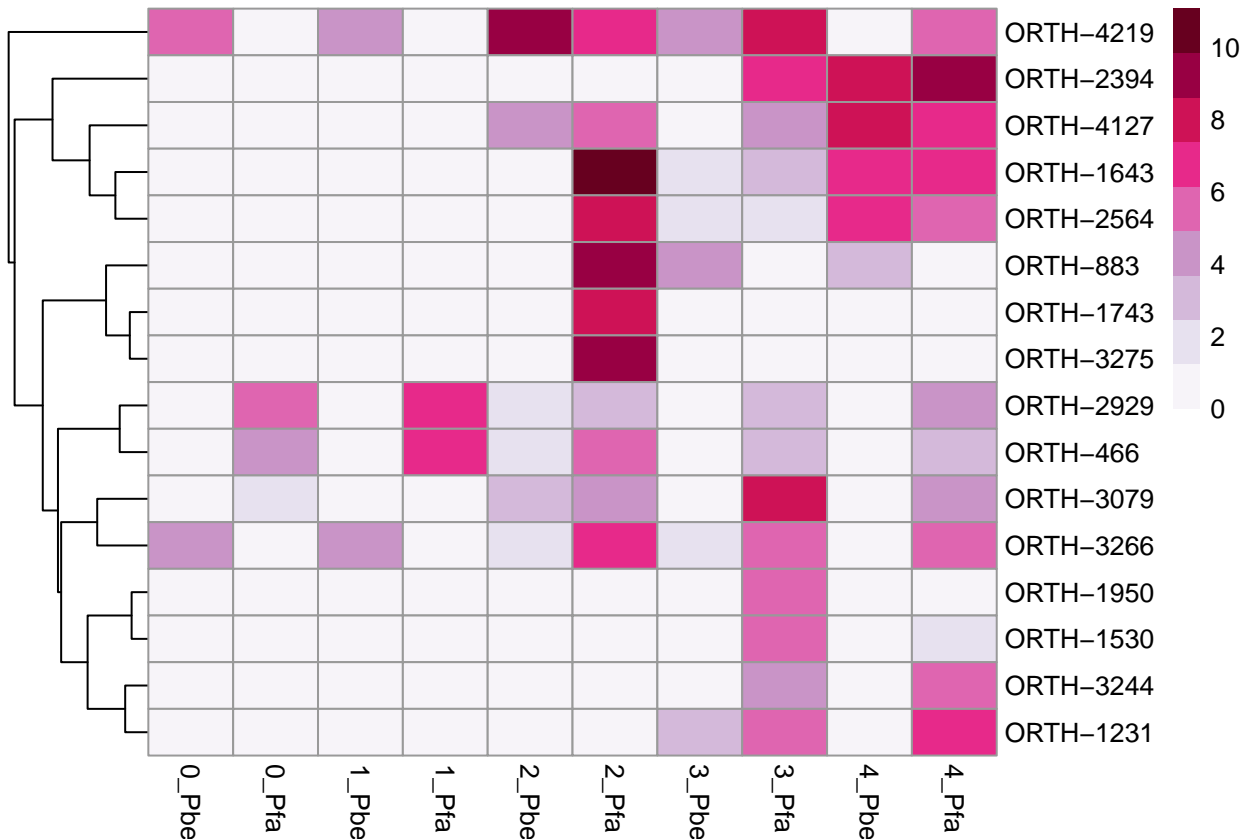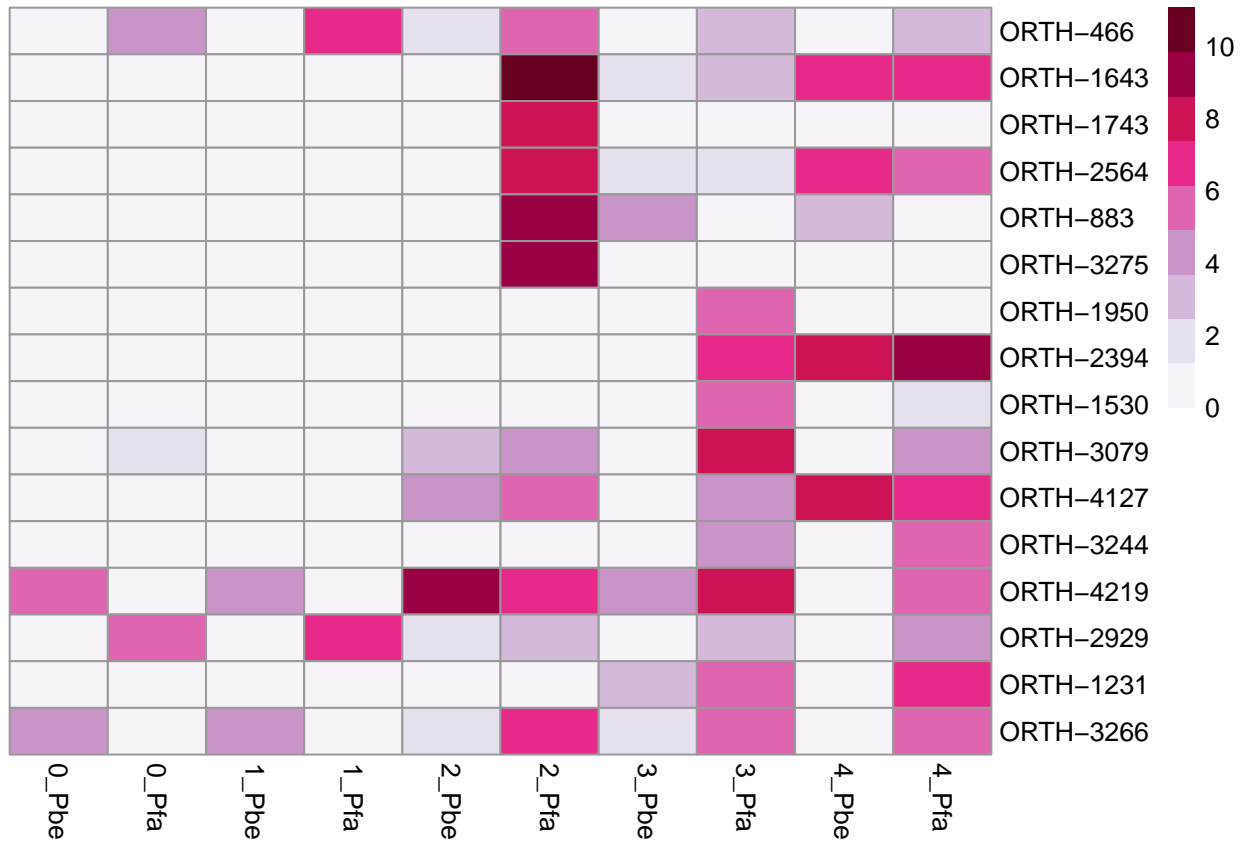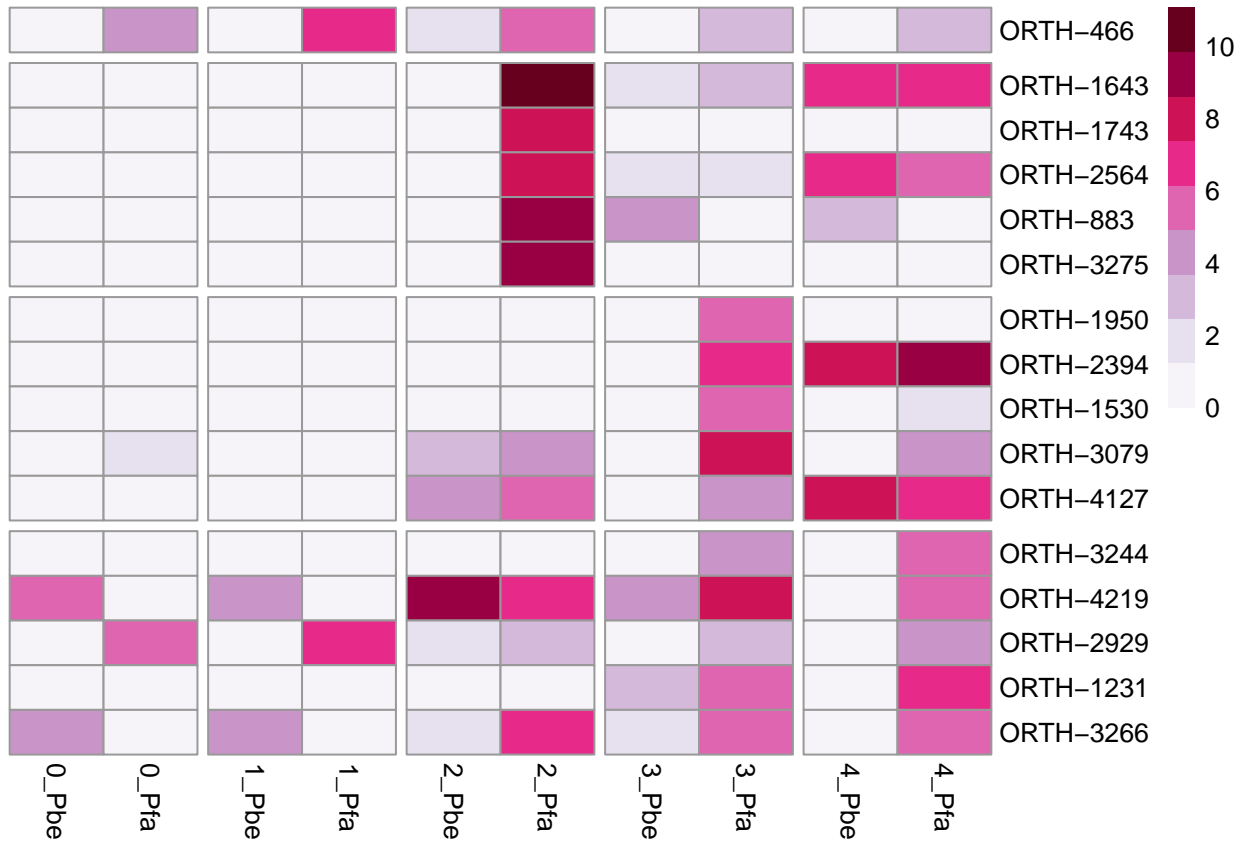


```r
tcm3 <- tcm2[rownames(all), ]
pheatmap(tcm3, show_colnames = TRUE, show_rownames = TRUE, cluster_cols = FALSE, cluster_rows = FALSE,
    color = brewer.pal(9, "PuRd"))
```

```
group <- all["stage"]
test <- which(group$stage != dplyr::lag(group$stage))
test2 <- test-1
pheatmap(tcm3, show_colnames = TRUE, show_rownames = TRUE, cluster_cols = FALSE, cluster_rows = FALSE,
    color = brewer.pal(9, "PuRd"), gaps_row = test2, gaps_col = c(2, 4, 6, 8))
```

ORTH–466
ORTH–1643
ORTH–1743
ORTH–2564
ORTH–883
ORTH–3275
ORTH–1950
ORTH–2394
ORTH–1530
ORTH–3079
ORTH–4127
ORTH–3244
ORTH–4219
ORTH–2929
ORTH–1231
ORTH–3266

10
8
6
4
2
0

0_Pbe 0_Pfa 1_Pbe 1_Pfa 2_Pbe 2_Pfa 3_Pbe 3_Pfa 4_Pbe 4_Pfa

```r
all$pbpfgene <- paste(all$pfgene, all$pbgene, sep="   ")
all$orthname <- rownames(all)

rowData(p.sce.m)$pbpfgene <- all[match(rownames(rowData(p.sce.m)), all[, 10]), 9]

rowData(p.sce.m)$pfgene <- all[match(rownames(rowData(p.sce.m)), all[, 10]), 7]

geneanno <- read.csv("/Users/vh3/Documents/allpf_geneanno.csv", header = TRUE)
rowData(p.sce.m)$gene_name <- geneanno[match(rowData(p.sce.m)$pfgene, geneanno[, 1]), 5]
rowData(p.sce.m)$product_desc <- geneanno[match(rowData(p.sce.m)$pfgene, geneanno[, 1]), 3]
rowData(p.sce.m)$genesanno <- paste(rowData(p.sce.m)$pbpfgene, rowData(p.sce.m)$gene_name, sep="   ")
rowData(p.sce.m)$name_desc <- paste(rowData(p.sce.m)$pfgene, rowData(p.sce.m)$gene_name, rowData(p.sce.m

rd <- read.csv("rdfordespeciesmananno.csv", header=TRUE, row.names = 1)

rowData(p.sce.m) <- rd

rowData(p.sce.m)$stage <- all[match(rownames(rowData(p.sce.m)), all[, 10]), 8]
rowData(p.sce.m)$orthname <- rownames(rowData(p.sce.m))
rownames(rowData(p.sce.m)) <- rowData(p.sce.m)$man_anno

rd <- as.data.frame(rowData(p.sce.m))
#rownames(rd) <- rd$man_anno
```
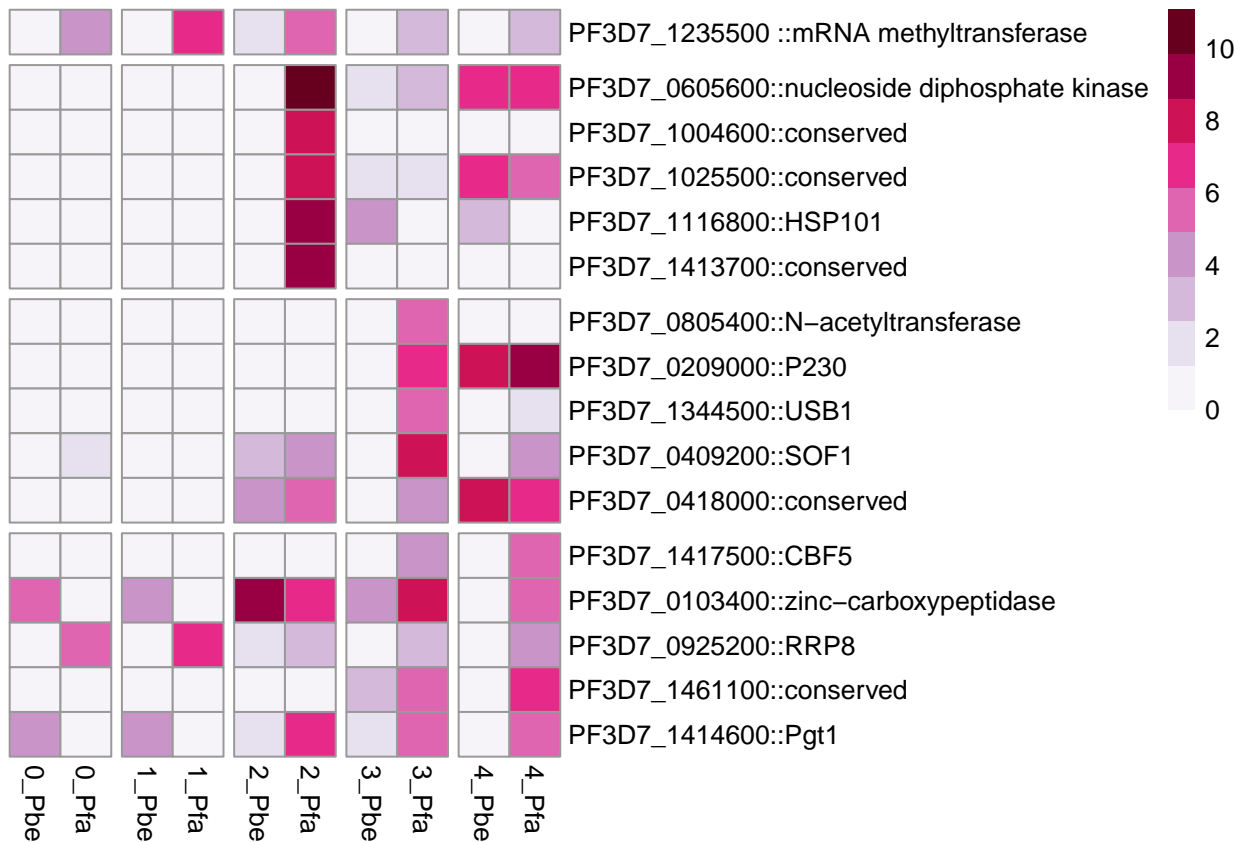
```r
ann_c <- list(stage = cols)

tcm4 <- tcm3

all$man_anno <- rd[match(rownames(all), rownames(rd)), 7]

rownames(tcm4) <- all[match(rownames(tcm4), all[, 10]), 11]

group <- all["stage"]
test <- which(group$stage != dplyr::lag(group$stage))
test2 <- test-1
pheatmap(tcm4, show_colnames = TRUE, show_rownames = TRUE, cluster_cols = FALSE, cluster_rows = FALSE,
    color = brewer.pal(9, "PuRd"), gaps_row = test2, gaps_col = c(2, 4, 6, 8))
```



```r
ann_c <- list(stage = cols)


rownames(rd) <- rd$man_anno
stage <- rd["stage"]


pheatmap(tcm4, show_colnames = TRUE, show_rownames = TRUE, cluster_cols = FALSE, cluster_rows = FALSE,
    color = brewer.pal(9, "PuRd"), gaps_row = test2, gaps_col = c(2, 4, 6, 8), annotation_row=stage, fo
```