Hcrt analysis - males only

```
library(Seurat)
library(ggplot2)
library(sctransform)
library(cowplot)
```

```
##
## ************************************************************
## Note: As of version 1.0.0, cowplot does not change the
##   default ggplot2 theme anymore. To recover the previous
##   behavior, execute:
##   theme_set(theme_cowplot())
## ************************************************************
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##     filter, lag
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(scrattch.hicat)
```

load only Hcrt core cells, with doublets removed

```
load('~/postdoc2/Shibin_hypocretin/hypocretin/hcrtcorecells.RData')
```

relabel clusters

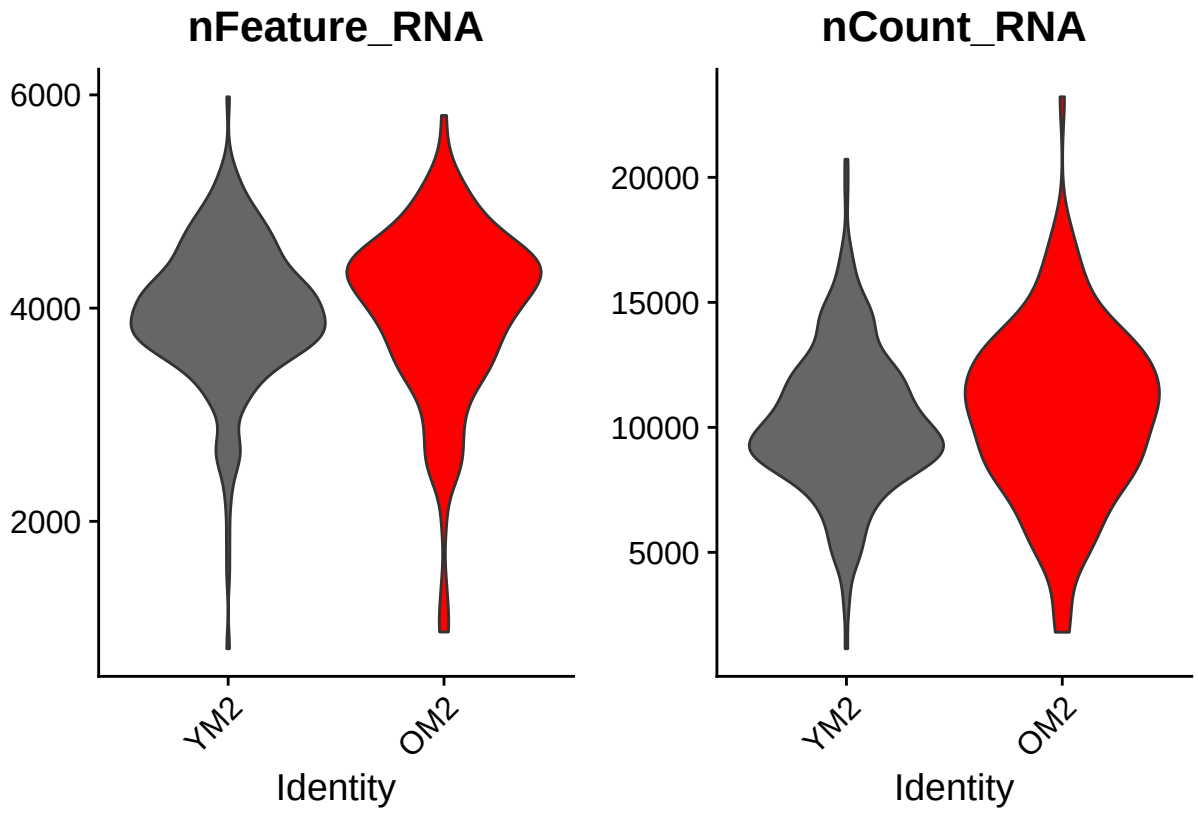
```
hcrt<-RenameIdents(hcrt,"4"="1")
hcrt<-RenameIdents(hcrt,"10"="2")
hcrt<-RenameIdents(hcrt,"9"="3")
hcrt$merged.res.2.renamed<-Idents(hcrt)
hcrt$merged.res.2.renamed<-factor(hcrt$merged.res.2.renamed,levels=c("0","1","2","3"))
Idents(hcrt)<-'merged.res.2.renamed'
```

split by sex - males here:

```
sex<-as.factor(hcrt$librarynames %in% c("YF2","OF2"))
levels(sex)<-c('male','female')
hcrt$sex<-sex
Idents(hcrt)<-'sex'
hcrt<-subset(hcrt,idents='male')
```

check genes and umis detected per library – data quality

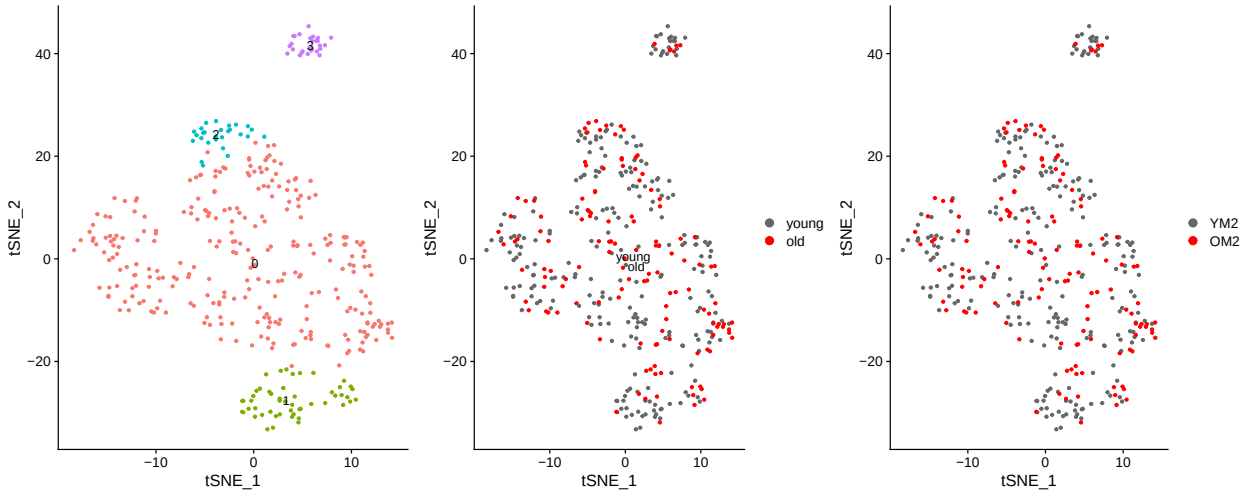
```
VlnPlot(hcrt,features=c("nFeature_RNA","nCount_RNA"),group.by='librarynames',pt.size = 0,cols=c("gray40
```




```
p1<-DimPlot(hcrt, label = TRUE,group.by='merged.res.2.renamed') + NoLegend()
```

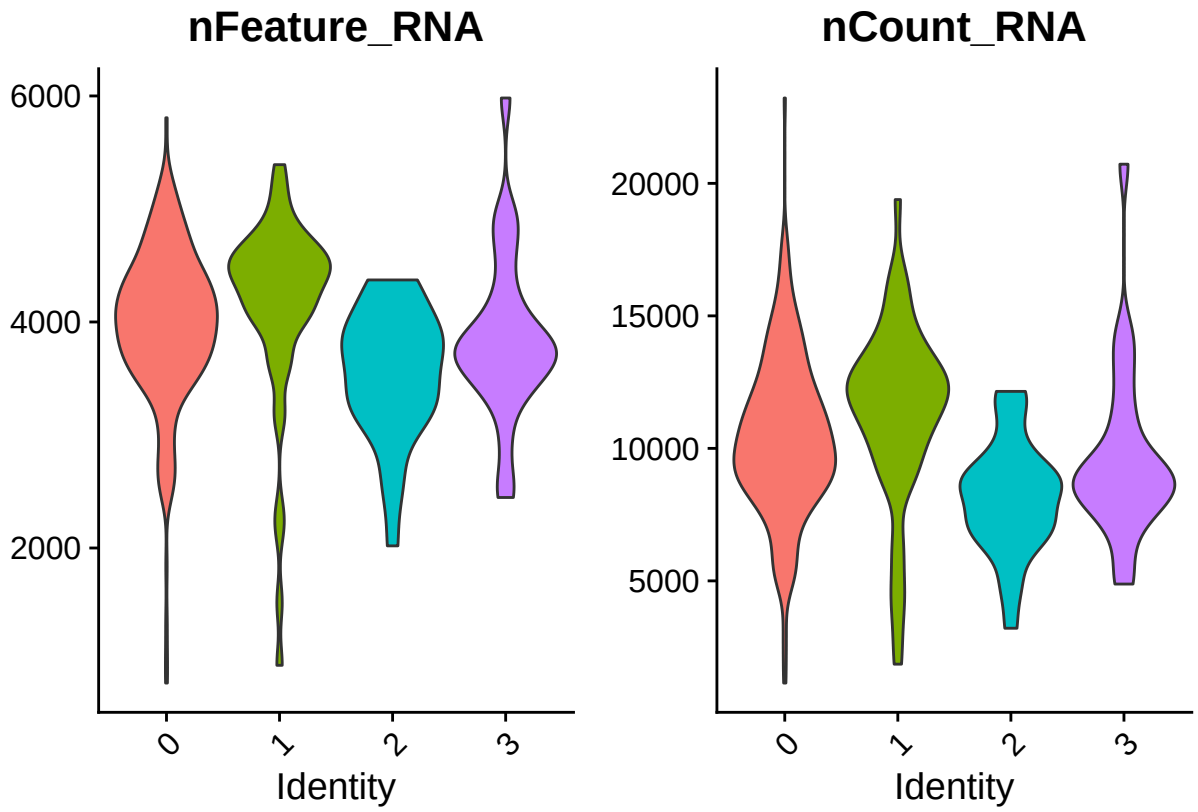
```
## Warning: Using `as.character()` on a quosure is deprecated as of rlang 0.3.0.
## Please use `as_label()` or `as_name()` instead.
## This warning is displayed once per session.
```

```
p2<-DimPlot(hcrt, label = TRUE,group.by='age',cols=c("gray40","red"))
p3<-DimPlot(hcrt, label = FALSE,group.by='librarynames',cols=c("gray40","red"))
plot_grid(p1,p2,p3,ncol=3)
```



check genes and umis detected per cluster – data quality + some notion of how large a cell we are looking at.

```
VlnPlot(hcrt,features=c("nFeature_RNA","nCount_RNA"),group.by='merged.res.2.renamed',pt.size = 0)
```



Look for cluster specific genes:

```
Idents(hcrt)<-'merged.res.2.renamed'
markers<-FindAllMarkers(hcrt,logfc.threshold = log(2))
```

```
## Calculating cluster 0
## Calculating cluster 1
```

```
## Calculating cluster 2
```

```
## Calculating cluster 3
```

```
markers %>% group_by(cluster) %>% top_n(5,avg_logFC)
```

```
## # A tibble: 20 x 7
## # Groups:   cluster [4]
##       p_val avg_logFC pct.1 pct.2 p_val_adj cluster gene
##       <dbl>     <dbl> <dbl> <dbl>     <dbl> <fct>   <chr>
##  1 1.63e-32     0.767 1     0.952  2.61e-28 0       Rgs7
##  2 1.02e-28     1.06  0.971 0.533  1.62e-24 0       Hmcn1
##  3 4.75e-28     0.891 0.978 0.59   7.57e-24 0       Cntnap4
##  4 2.94e-18     0.835 0.92  0.581  4.68e-14 0       Pde3a
##  5 1.82e-17     0.943 0.993 0.905  2.91e-13 0       Hcrt
##  6 1.74e-24     1.36  0.98  0.427  2.77e-20 1       Cntn4
##  7 4.92e-21     0.942 0.882 0.315  7.84e-17 1       Fam189a1
##  8 2.56e-19     0.932 1     0.615  4.08e-15 1       B3galt1
##  9 4.50e-18     0.923 0.706 0.188  7.17e-14 1       Efna5
## 10 4.48e- 9     0.954 0.882 0.679  7.14e- 5 1       Trpm3
## 11 8.16e-16     1.30  1     0.949  1.30e-11 2       Adgrb3
## 12 1.26e-15     1.46  1     0.554  2.02e-11 2       Oxr1
## 13 2.27e-15     1.04  0.889 0.288  3.61e-11 2       Tmem108
## 14 4.18e-14     1.24  0.926 0.446  6.67e-10 2       Ldb2
## 15 1.20e- 3     0.931 0.519 0.251  1.00e+ 0 2       Erbb4
## 16 8.79e-23     1.70  0.741 0.093  1.40e-18 3       Hs3st4
## 17 6.89e-22     2.13  1     0.331  1.10e-17 3       Sgcd
## 18 7.25e-17     1.73  0.926 0.367  1.16e-12 3       Thsd7b
## 19 5.05e-16     1.80  1     0.856  8.06e-12 3       Ntm
## 20 2.75e-14     1.72  0.741 0.198  4.39e-10 3       Frmpd4
```

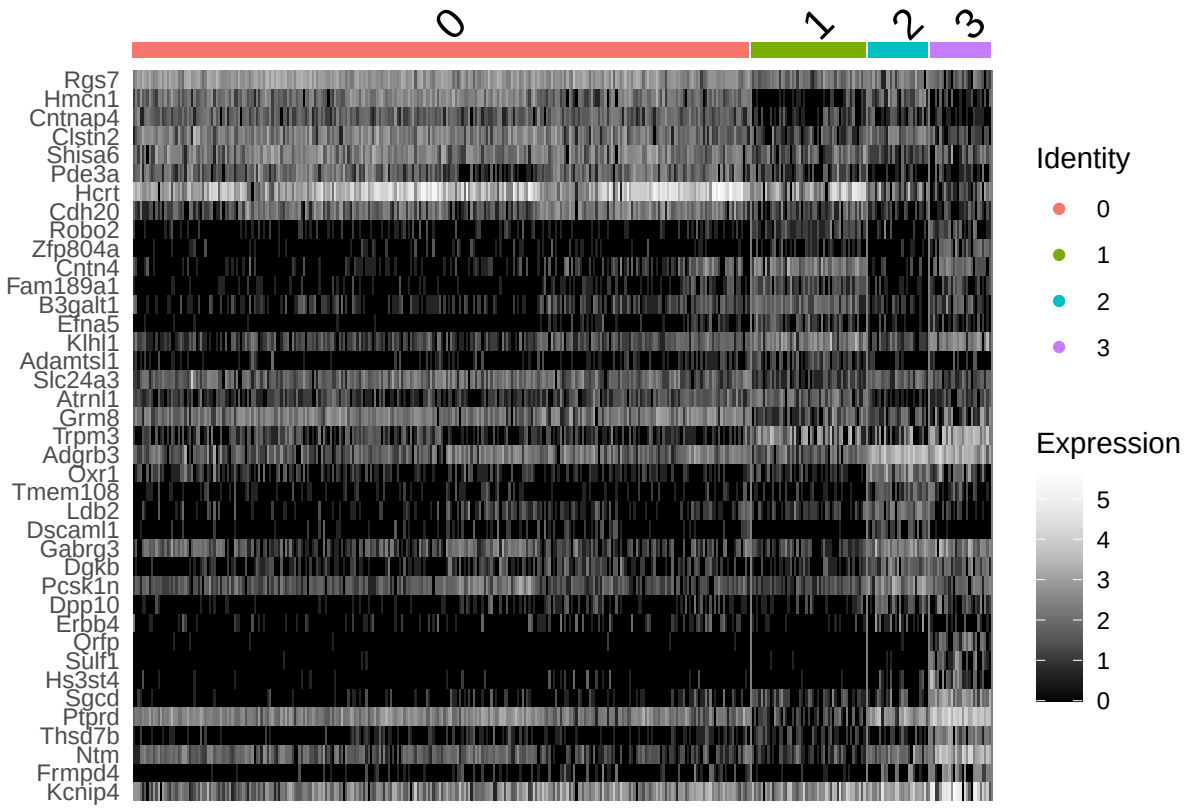
```
#markers.filtered<-markers[markers$p_val_adj<0.05,]
#write.csv(markers.filtered,file='~/postdoc2/Shibin_hypocretin/hypocretin/males_cluster_DEGs.csv')
```

plot a heatmap

```
markers %>% group_by(cluster) %>% top_n(10,avg_logFC) ->top10
DoHeatmap(hcrt, features = top10$gene,slot='data')+
  scale_fill_gradientn(colors = rev(RColorBrewer::brewer.pal(n = 10, name = "Greys")))
```

```
## Warning in RColorBrewer::brewer.pal(n = 10, name = "Greys"): n too large, allowed maximum for palett
## Returning the palette you asked for with that many colors
```

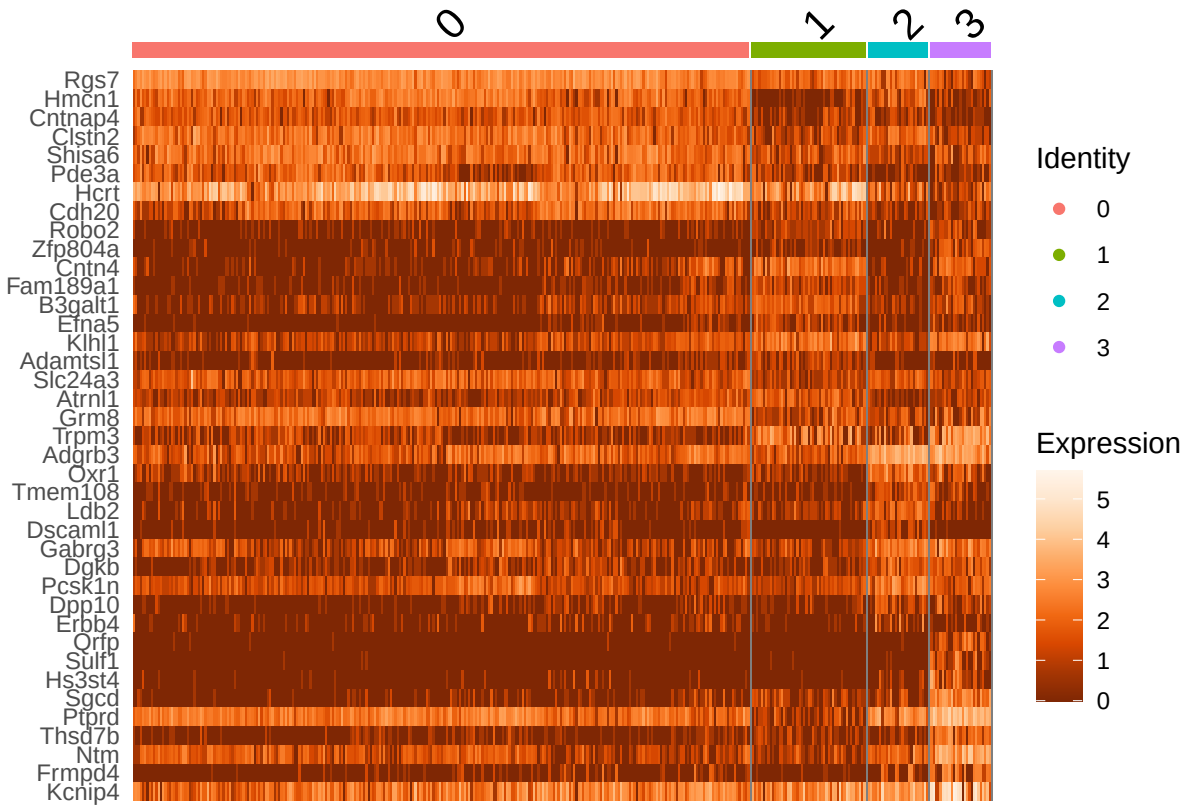
```
## Scale for 'fill' is already present. Adding another scale for 'fill', which
## will replace the existing scale.
```




```
DoHeatmap(hcrt, features = top10$gene,slot='data')+
  scale_fill_gradientn(colors = rev(RColorBrewer::brewer.pal(n = 10, name = "Oranges")))
```

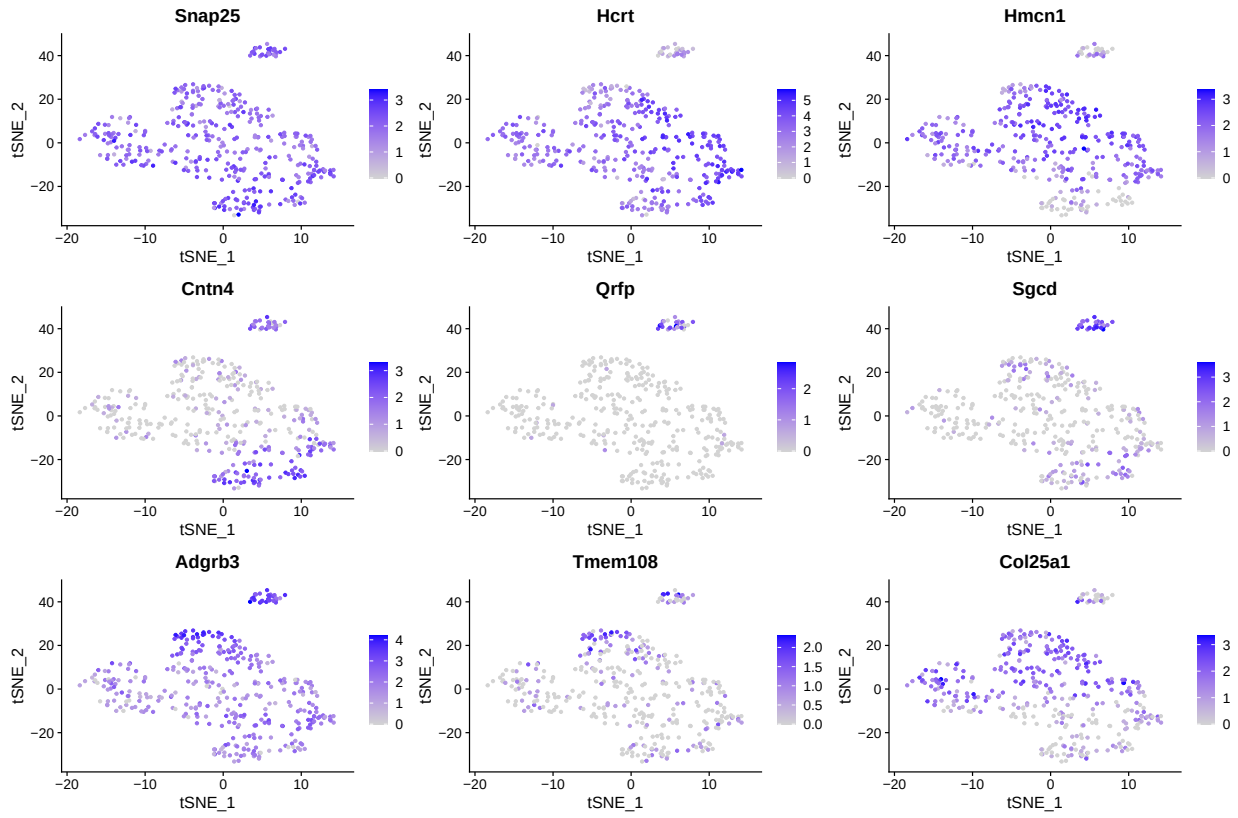
```
## Warning in RColorBrewer::brewer.pal(n = 10, name = "Oranges"): n too large, allowed maximum for pale
## Returning the palette you asked for with that many colors
```

```
## Scale for 'fill' is already present. Adding another scale for 'fill', which
## will replace the existing scale.
```

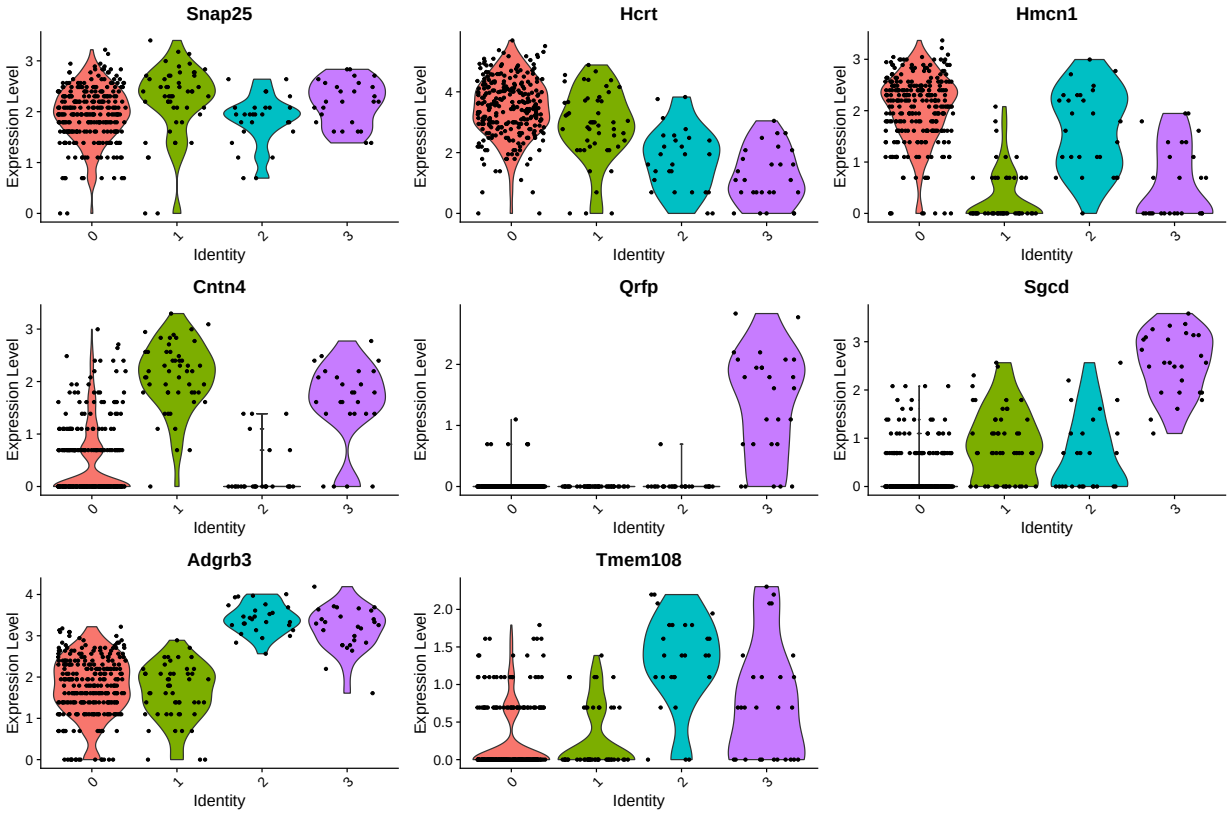


Featureplots of marker genes

```
FeaturePlot(hcrt,features=c("Snap25",'Hcrt',"Hmcn1","Cntn4","Qrfp","Sgcd","Adgrb3","Tmem108","Col25a1"))
```

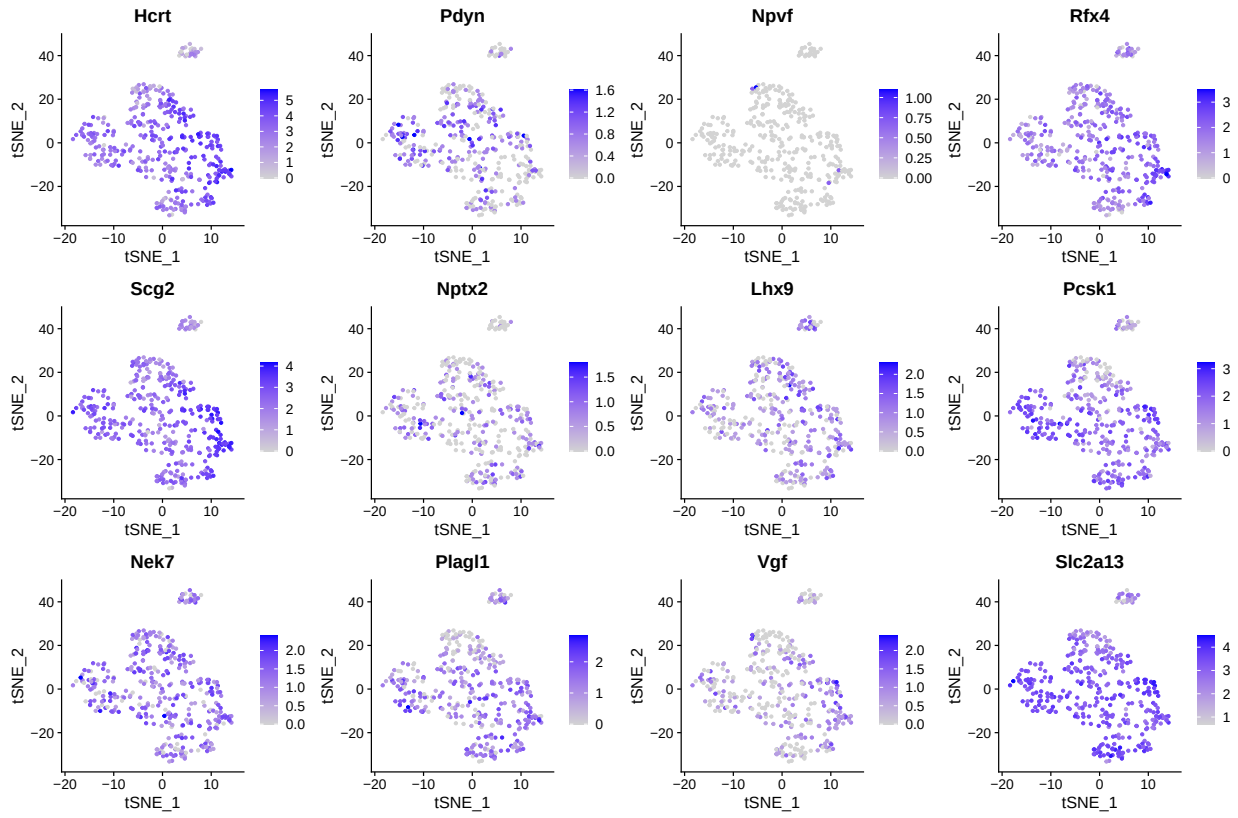



```
VlnPlot(hcrt,features=c("Snap25",'Hcrt',"Hmcn1","Cntn4","Qrfp","Sgcd","Adgrb3","Tmem108"),group.by = "m
```



Mickelsen's genes:

```
FeaturePlot(hcrt,c("Hcrt", "Pdyn", "Npvf", "Rfx4", "Scg2", "Nptx2", "Lhx9", "Pcsk1", "Nek7", "Plagl1",
```



do young and old cells contribute differently to the 4 clusters? total cell numbers.

```
table(hcrt$merged.res.2.renamed,hcrt$librarynames)
```

```
##
##      YM2 YF2 OM2 OF2
##    0 173   0 103   0
##    1  36   0  15   0
##    2  16   0  11   0
##    3  21   0   6   0
```

percentage.

```
t<-table(hcrt$merged.res.2.renamed,hcrt$librarynames)
prop.table(t,2) # cell percentages
```

```
##
##             YM2 YF2        OM2 OF2
##    0 0.70325203     0.76296296
##    1 0.14634146     0.11111111
##    2 0.06504065     0.08148148
##    3 0.08536585     0.04444444
```

the distribution of cells across the 4 clusters from each library is remarkably similar.

Lets look for differentially expressed genes between young and old in the supercluster of 0,1,2

```
hcrtcore<-subset(hcrt,idents=c("0","1","2"))
```

```
table(hcrtcore$age)
```

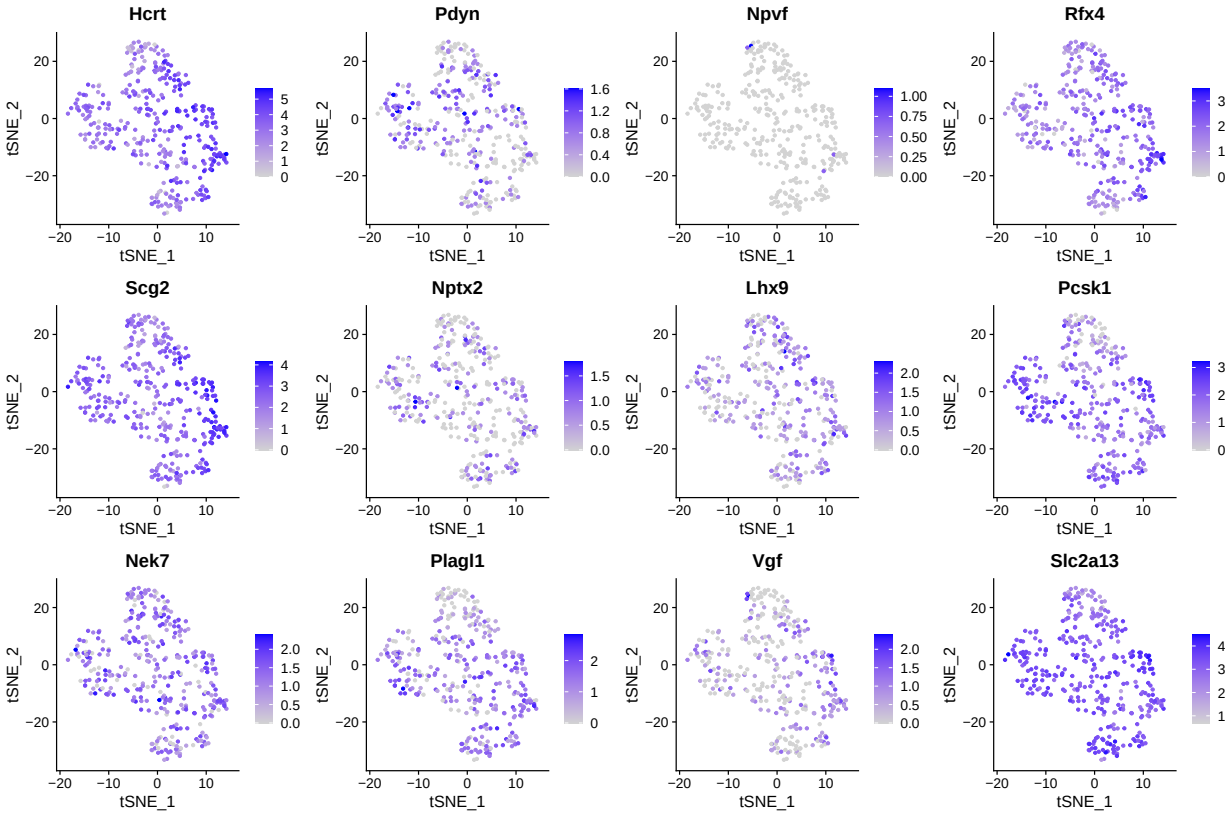
```
##
## young   old
##   225   129
```

```
table(hcrtcore$librarynames)
```

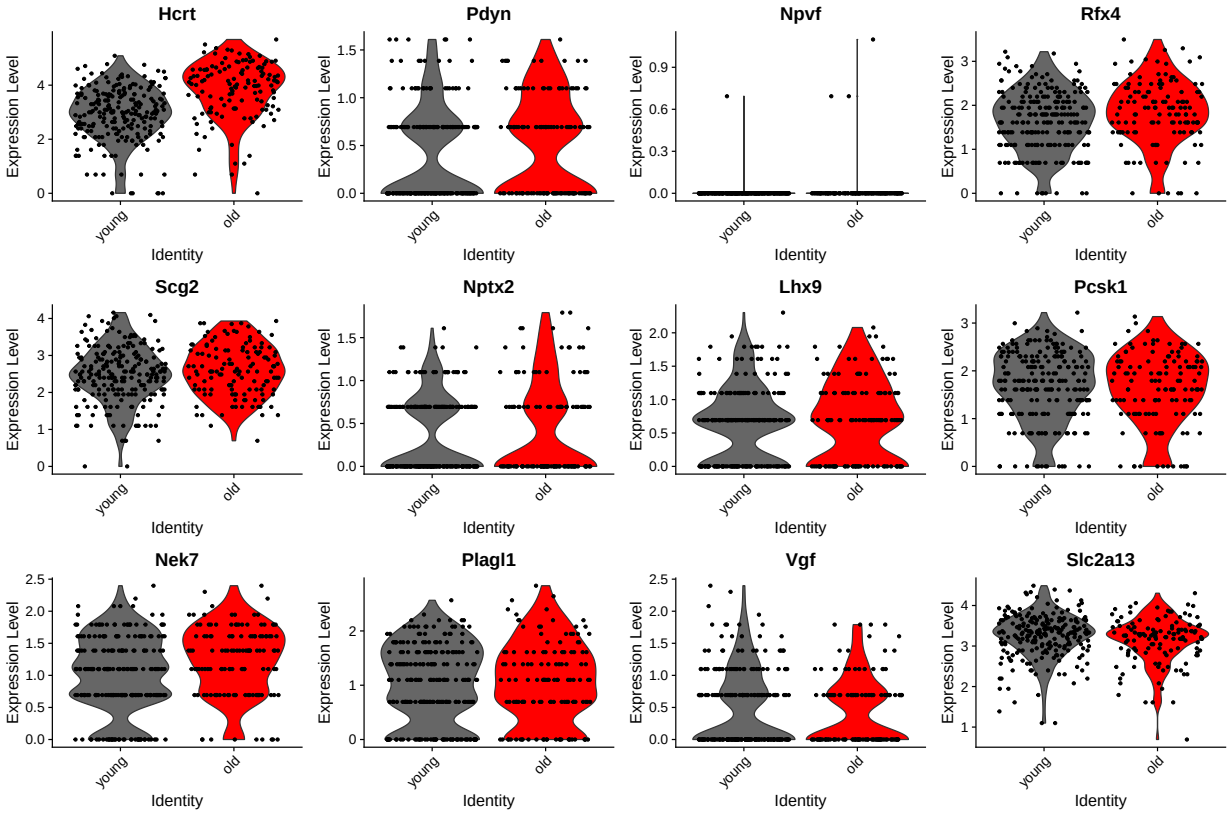
```
##
## YM2 YF2 OM2 OF2
## 225   0 129   0
```

first though, Mickelsen's genes:

```
FeaturePlot(hcrtcore,c("Hcrt", "Pdyn", "Npvf", "Rfx4", "Scg2", "Nptx2", "Lhx9", "Pcsk1", "Nek7", "Plagl
```

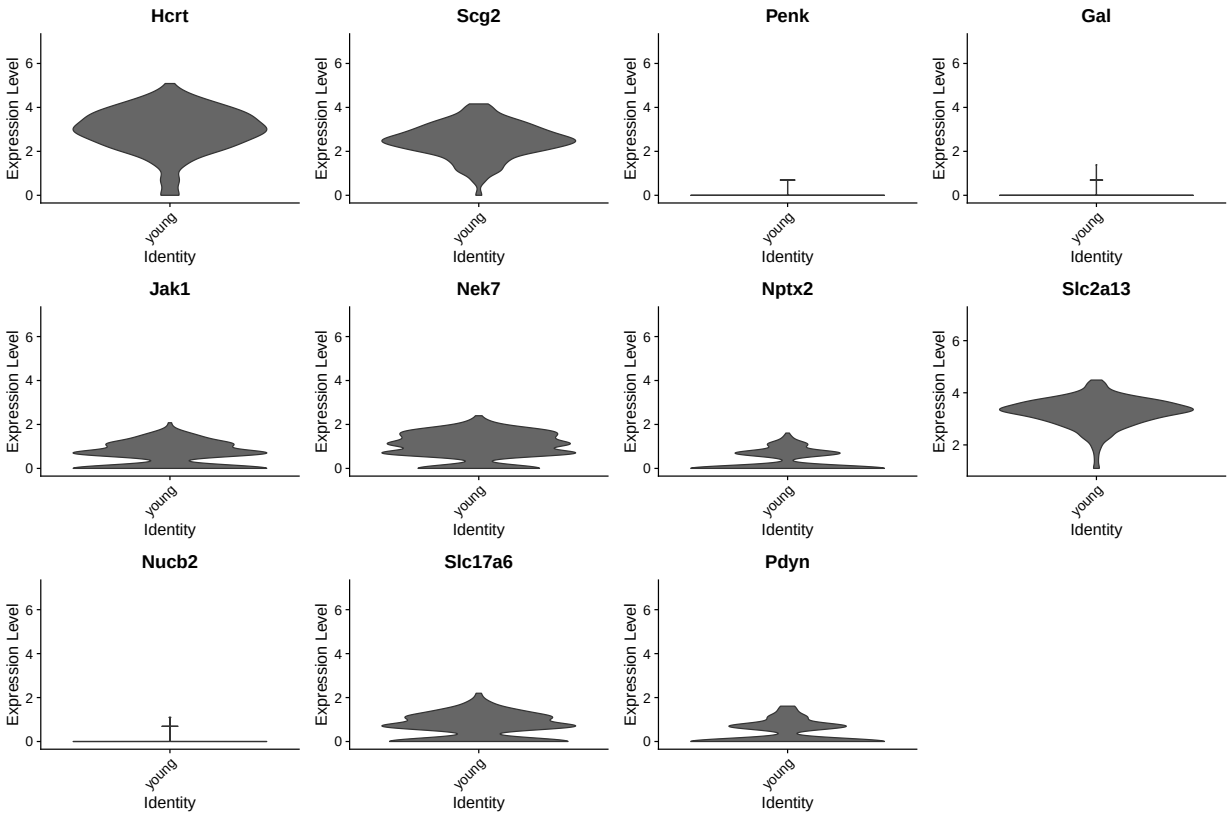



```
Idents(hcrtcore)<-'age'
VlnPlot(hcrtcore,c("Hcrt", "Pdyn", "Npvf", "Rfx4", "Scg2", "Nptx2", "Lhx9", "Pcsk1", "Nek7", "Plagl1",
```



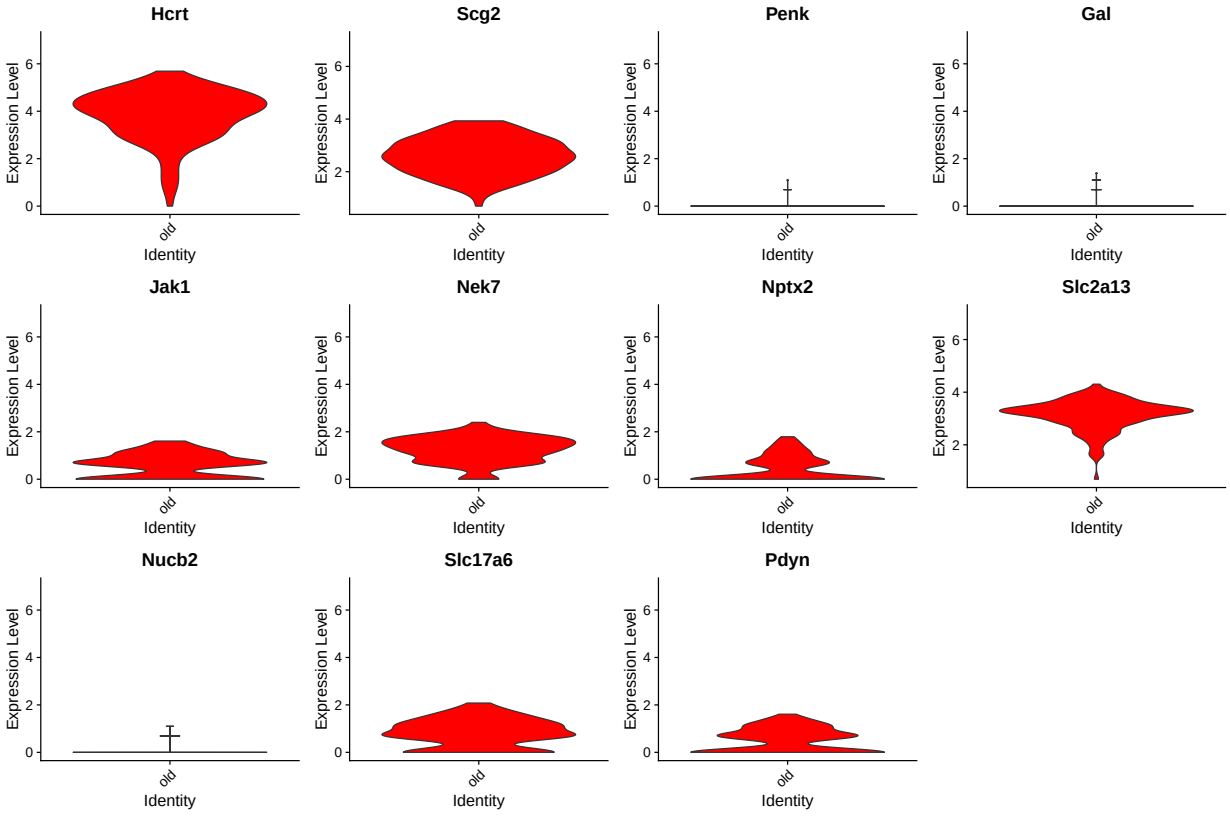
New plots for Shibin.

```
genesofinterest<-c("Hcrt","Scg2","Penk","Gal","Jak1","Nek7","Nptx2","Slc2a13","Nucb2","Slc17a6","Pdyn")
VlnPlot(hcrtcore,genesofinterest,idents = 'young',y.max=7,pt.size = 0,cols=c("gray40"))
```



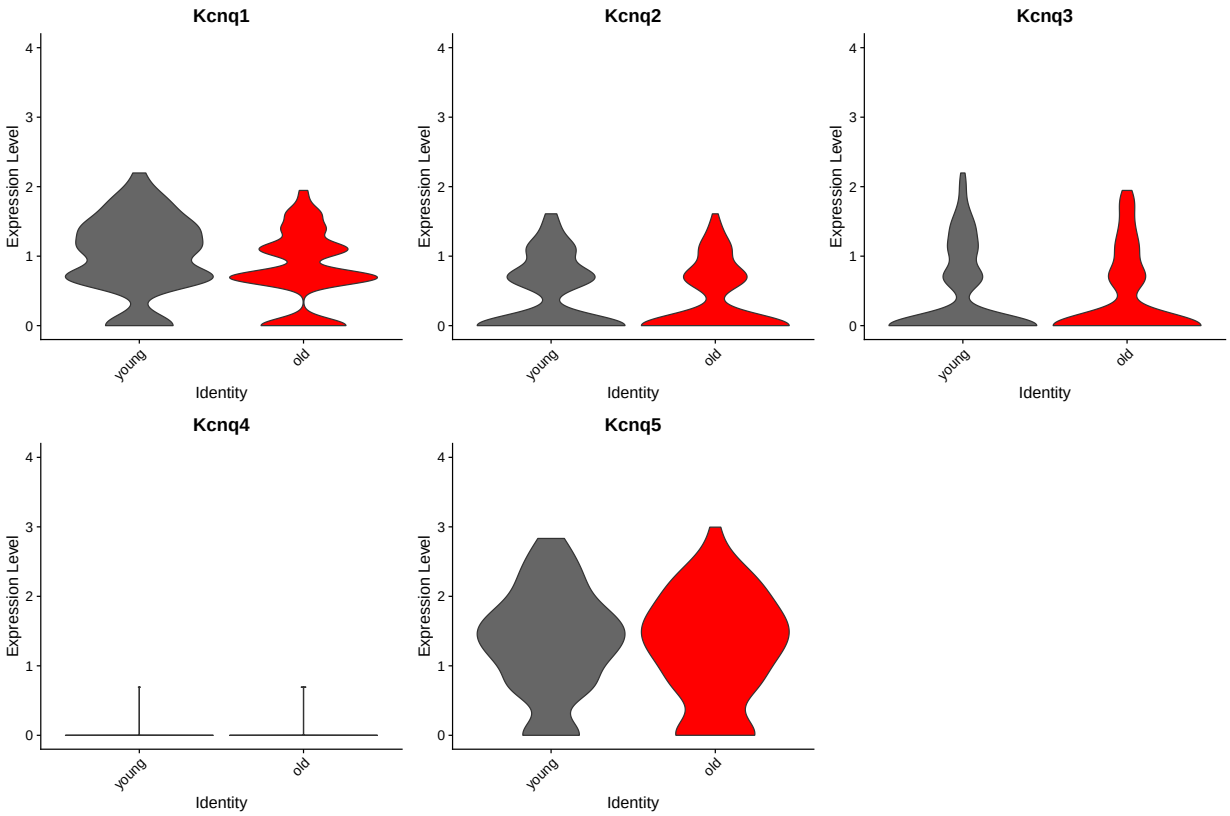
New plots for Shibin.

```
VlnPlot(hcrtcore,genesofinterest,idents = 'old',y.max=7,pt.size=0,cols=c("red"))
```



New plots of KCNQ channels

```
genesofinterest<-c("Kcnq1","Kcnq2","Kcnq3","Kcnq4","Kcnq5")
VlnPlot(hcrtcore,genesofinterest,y.max=4,pt.size = 0,cols=c("gray40","red"))
```



Fraction expressed.

```
a <- DotPlot(object = hcrtcore, features = genesofinterest)
a$data[,c('features.plot','id','pct.exp')]
```

```
##        features.plot    id   pct.exp
## Kcnq1          Kcnq1 young 86.222222
## Kcnq2          Kcnq2 young 50.666667
## Kcnq3          Kcnq3 young 40.444444
## Kcnq4          Kcnq4 young  1.333333
## Kcnq5          Kcnq5 young 88.000000
## Kcnq11         Kcnq1   old 79.069767
## Kcnq21         Kcnq2   old 41.085271
## Kcnq31         Kcnq3   old 35.658915
## Kcnq41         Kcnq4   old  3.100775
## Kcnq51         Kcnq5   old 82.945736
```

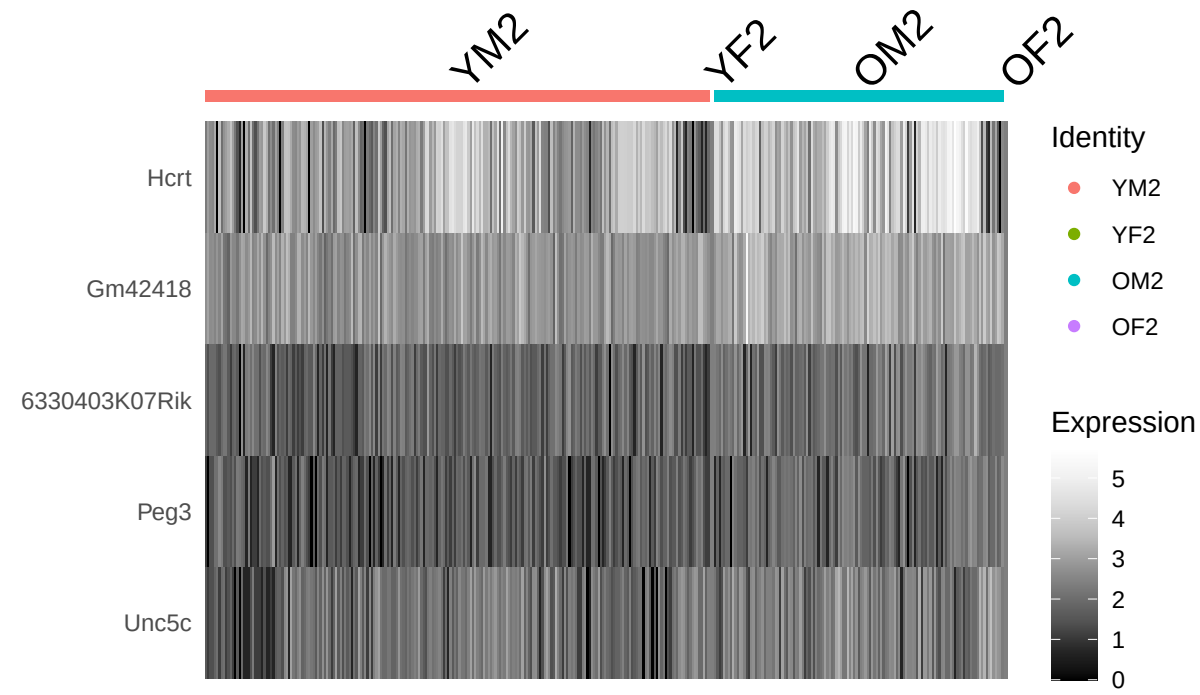
```
Idents(hcrtcore)<-'age'
agemarkers<-FindMarkers(hcrtcore,ident.2="young",ident.1="old")
#agemarkers
```

plot heatmap of sorted genes.

```
a<-agemarkers[agemarkers$p_val_adj<0.05,]
b<-agemarkers[order(a$avg_logFC,decreasing=T),]
DoHeatmap(hcrtcore,features=rownames(b),group.by='librarynames',slot='data')+
  scale_fill_gradientn(colors = rev(RColorBrewer::brewer.pal(n = 9, name = "Greys")))
```

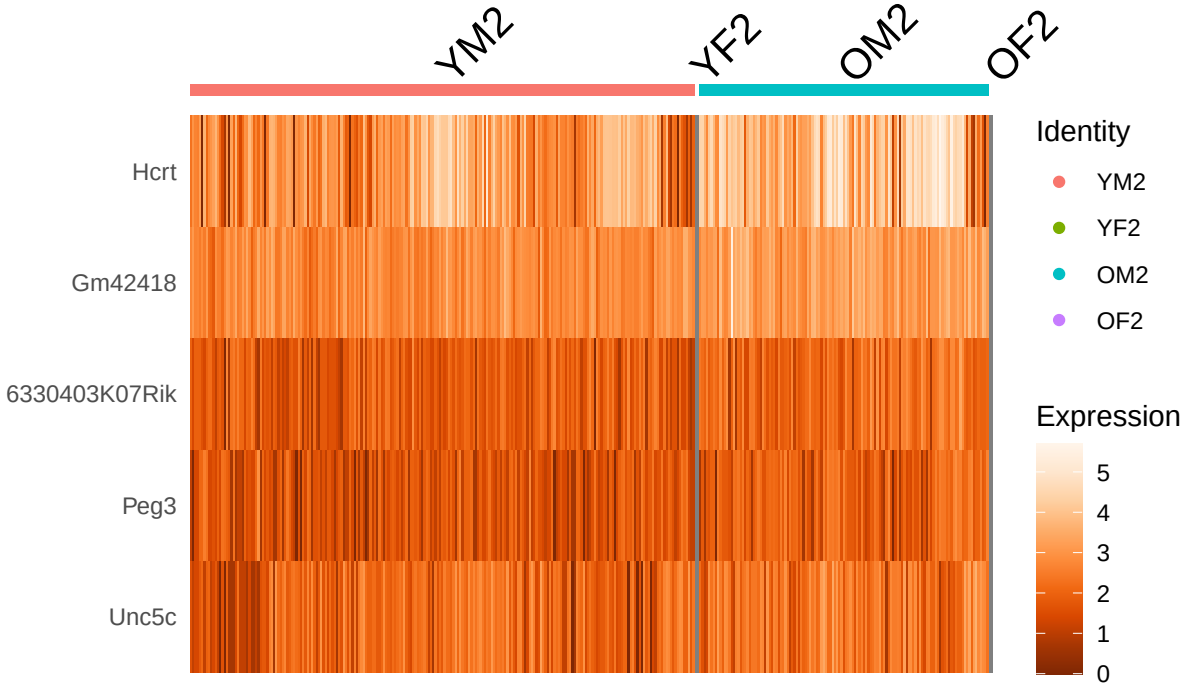
```
## Scale for 'fill' is already present. Adding another scale for 'fill', which
```

```
## will replace the existing scale.
```



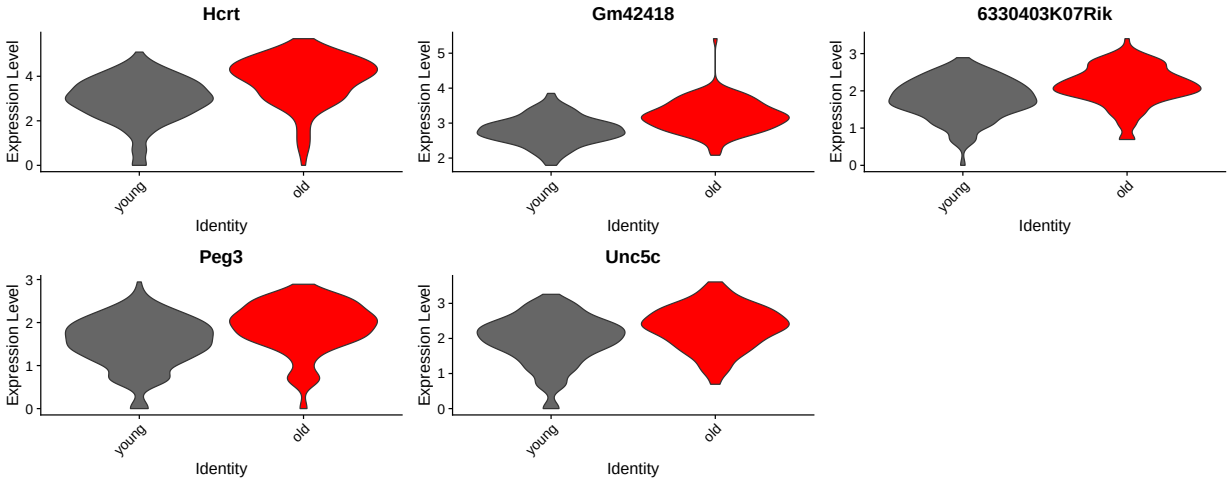

```
DoHeatmap(hcrtcore,features=rownames(b),group.by='librarynames',slot='data')+
  scale_fill_gradientn(colors = rev(RColorBrewer::brewer.pal(n = 9, name = "Oranges")))
```

```
## Scale for 'fill' is already present. Adding another scale for 'fill', which
## will replace the existing scale.
```



plot all significant makers.

```
Idents(hcrtcore)<-"age"
VlnPlot(hcrtcore,features=rownames(b),pt.size = 0,cols=c("gray40","red"))
```



what does this look like in a volcano plot?

```
#try DEseq DEG test
agemarkers2<-FindMarkers(hcrtcore,ident.2="young",ident.1="old",logfc.threshold = 0)
#agemarkers2
```

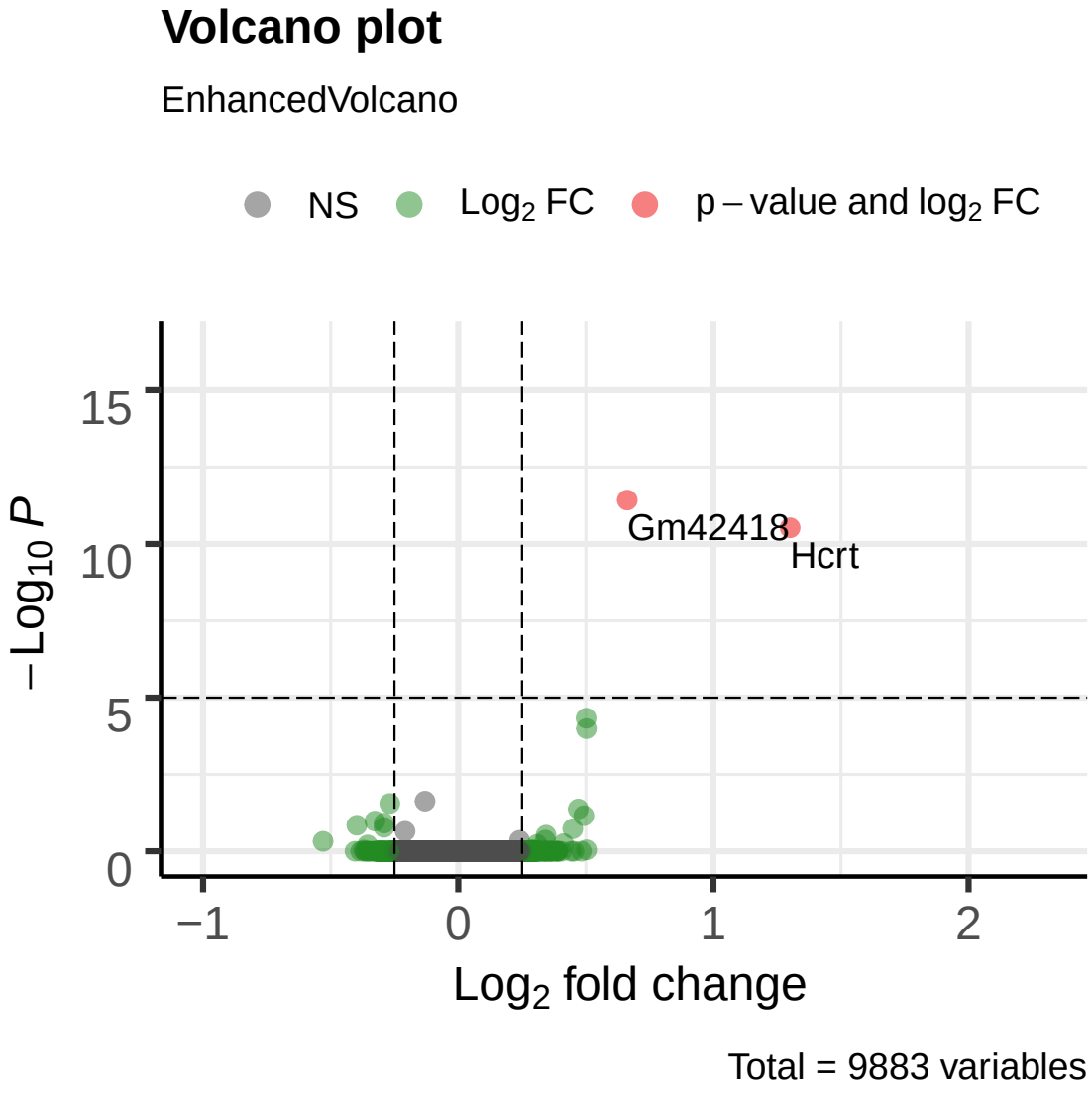
```r
# convert from ln to log2 fold change.
agemarkers2$avg_log2FC<-agemarkers2$avg_logFC*log2(exp(1))
library(EnhancedVolcano)
```

```
## Loading required package: ggrepel
```

```r
EnhancedVolcano(agemarkers2,
                lab=rownames(agemarkers2),
                x='avg_log2FC',
                y='p_val_adj',
                xlim=c(-1,2.3),
                FCcutoff=0.25,
                transcriptPointSize=3,
                transcriptLabSize = 5)
```

```
## Warning in EnhancedVolcano(agemarkers2, lab = rownames(agemarkers2), x =
## "avg_log2FC", : transcriptPointSize argument deprecated in v1.4 - please use
## pointSize
```

```
## Warning in EnhancedVolcano(agemarkers2, lab = rownames(agemarkers2), x =
## "avg_log2FC", : transcriptLabSize argument deprecated in v1.4 - please use
## labSize
```




```
write.csv(agemarkers2[agemarkers2$p_val_adj<0.05,],file = '~/postdoc2/Shibin_hypocretin/hypocretin/agema
```