Hcrt analysis - females 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

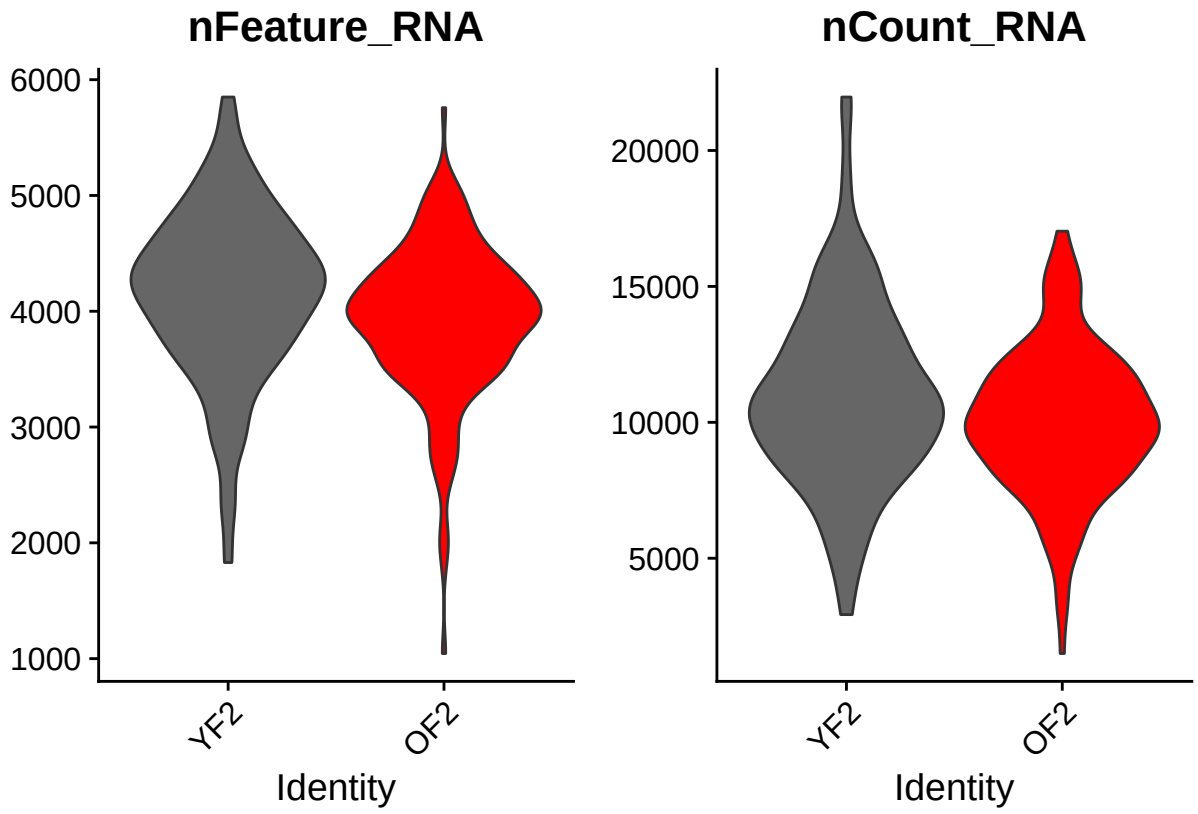
```
hcrt<-RenameIdents(hcrt,"4"="1")
hcrt<-RenameIdents(hcrt,"10"="2")
hcrt<-RenameIdents(hcrt,"9"="3")
hcrt$merged.res.2.renamed<-Idents(hcrt)
hcrt$merged.res.2.renamed<-factor(hcrt$merged.res.2.renamed,levels=c("0","1","2","3"))
Idents(hcrt)<-'merged.res.2.renamed'
```

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='female')
```

check genes and umis detected per library – data quality

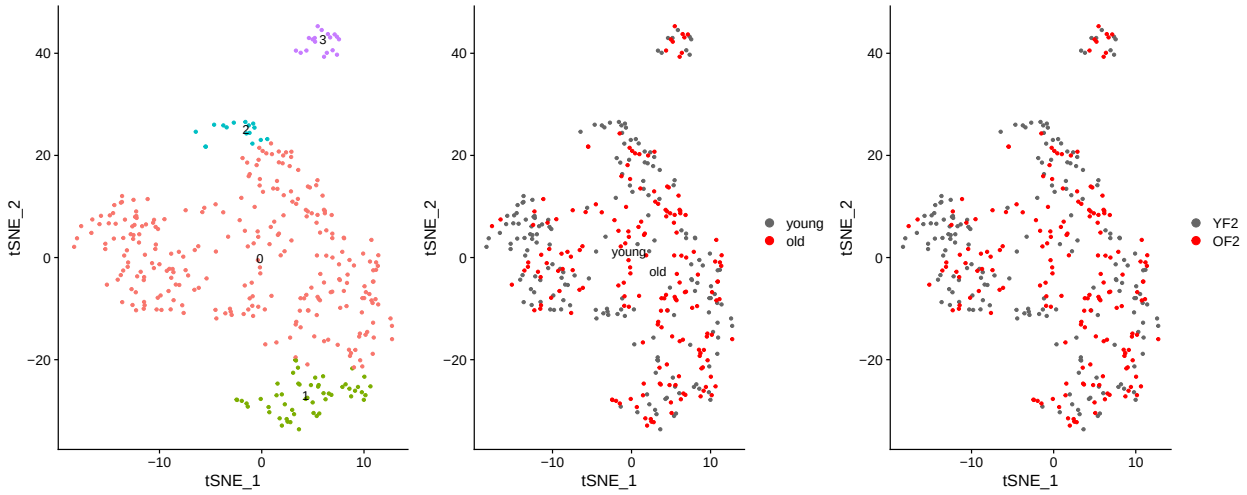
```
VlnPlot(hcrt,features=c("nFeature_RNA","nCount_RNA"),group.by='librarynames',pt.size = 0, cols=c("gray4
```




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

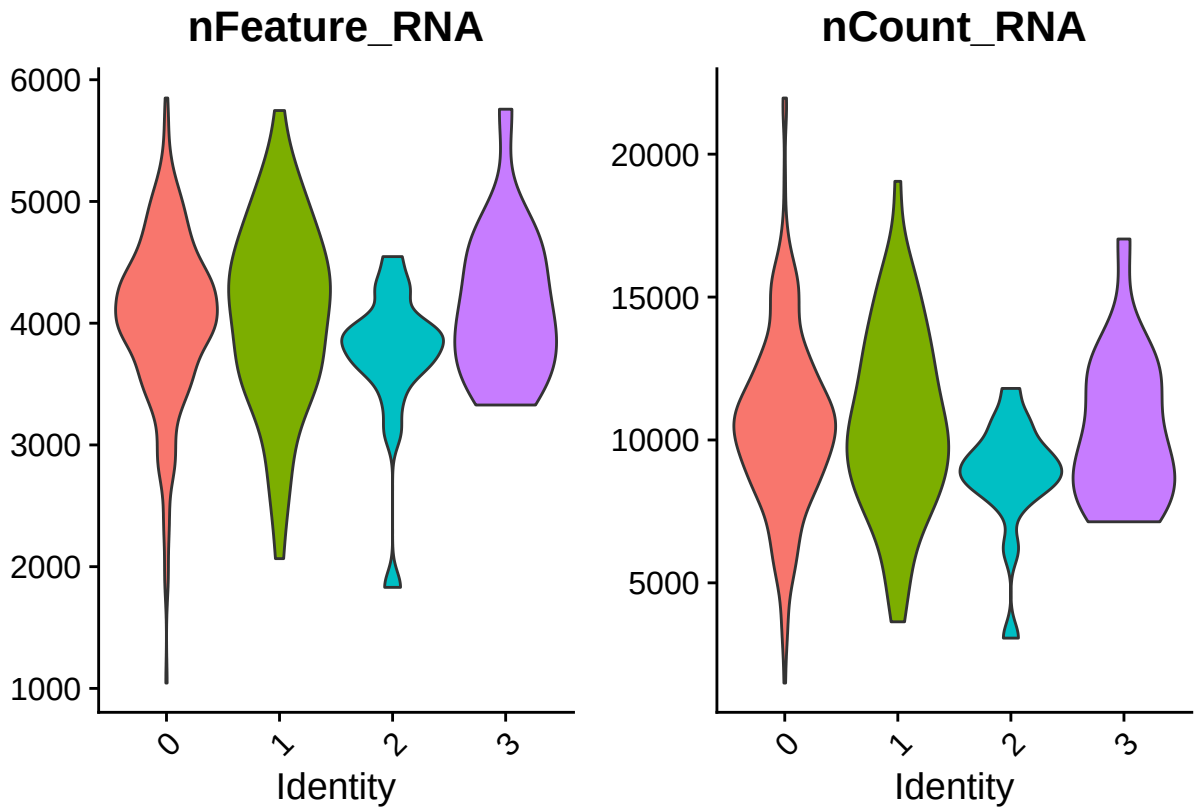
```
## Warning: Using `as.character()` on a quosure is deprecated as of rlang 0.3.0.
## Please use `as_label()` or `as_name()` instead.
## This warning is displayed once per session.
```

```
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 3.34e-25     0.742 1     0.977  5.32e-21 0       Rgs7
##  2 1.50e-20     0.910 0.974 0.529  2.39e-16 0       Hmcn1
##  3 3.29e-20     0.778 0.989 0.851  5.25e-16 0       Hs3st5
##  4 6.36e-20     0.760 0.962 0.586  1.01e-15 0       Cntnap4
##  5 1.46e- 9     0.822 0.748 0.425  2.33e- 5 0       Col25a1
##  6 3.99e-29     1.48  1     0.36   6.36e-25 1       Cntn4
##  7 2.93e-24     0.951 0.887 0.257  4.67e-20 1       Fam189a1
##  8 1.81e-23     1.07  1     0.567  2.88e-19 1       B3galt1
##  9 1.32e-19     0.873 0.774 0.237  2.11e-15 1       Adamtsl1
## 10 7.55e- 7     0.925 0.774 0.66   1.20e- 2 1       Trpm3
## 11 1.81e-10     1.52  1     0.97   2.88e- 6 2       Adgrb3
## 12 6.47e- 9     1.26  1     0.570  1.03e- 4 2       Ntng1
## 13 4.61e- 8     1.15  1     0.967  7.35e- 4 2       Kcnip4
## 14 6.52e- 8     1.19  0.938 0.709  1.04e- 3 2       Unc5d
## 15 3.67e- 7     1.23  0.938 0.516  5.85e- 3 2       Oxr1
## 16 2.90e-33     2.09  0.778 0.036  4.63e-29 3       Qrfp
## 17 3.51e-17     2.13  1     0.293  5.60e-13 3       Sgcd
## 18 2.48e-15     1.94  0.833 0.173  3.96e-11 3       Frmpd4
## 19 7.80e-15     1.92  0.944 0.319  1.24e-10 3       Thsd7b
## 20 7.44e-10     1.81  1     0.66   1.19e- 5 3       Trpm3
```

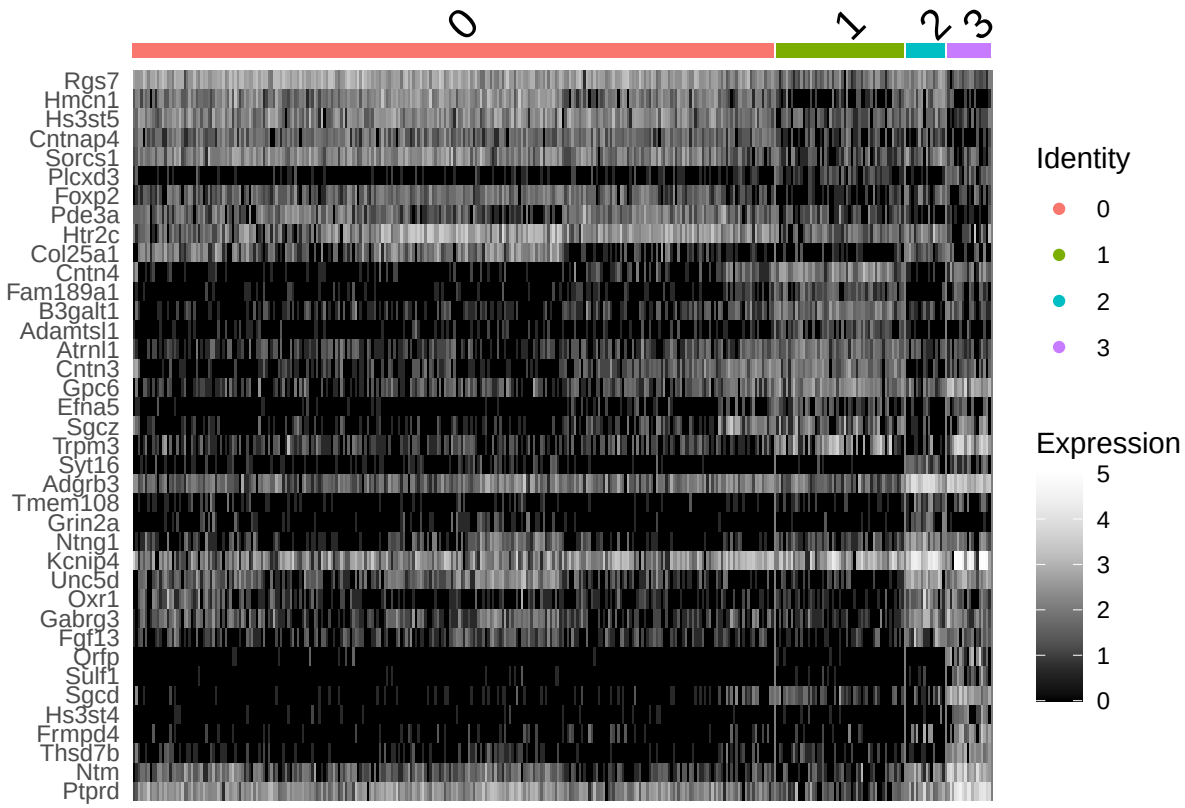
```
markers.filtered<-markers[markers$p_val_adj<0.05,]
#write.csv(markers.filtered,file='~/postdoc2/Shibin_hypocretin/hypocretin/females_cluster_DEGs.csv')
```

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
```

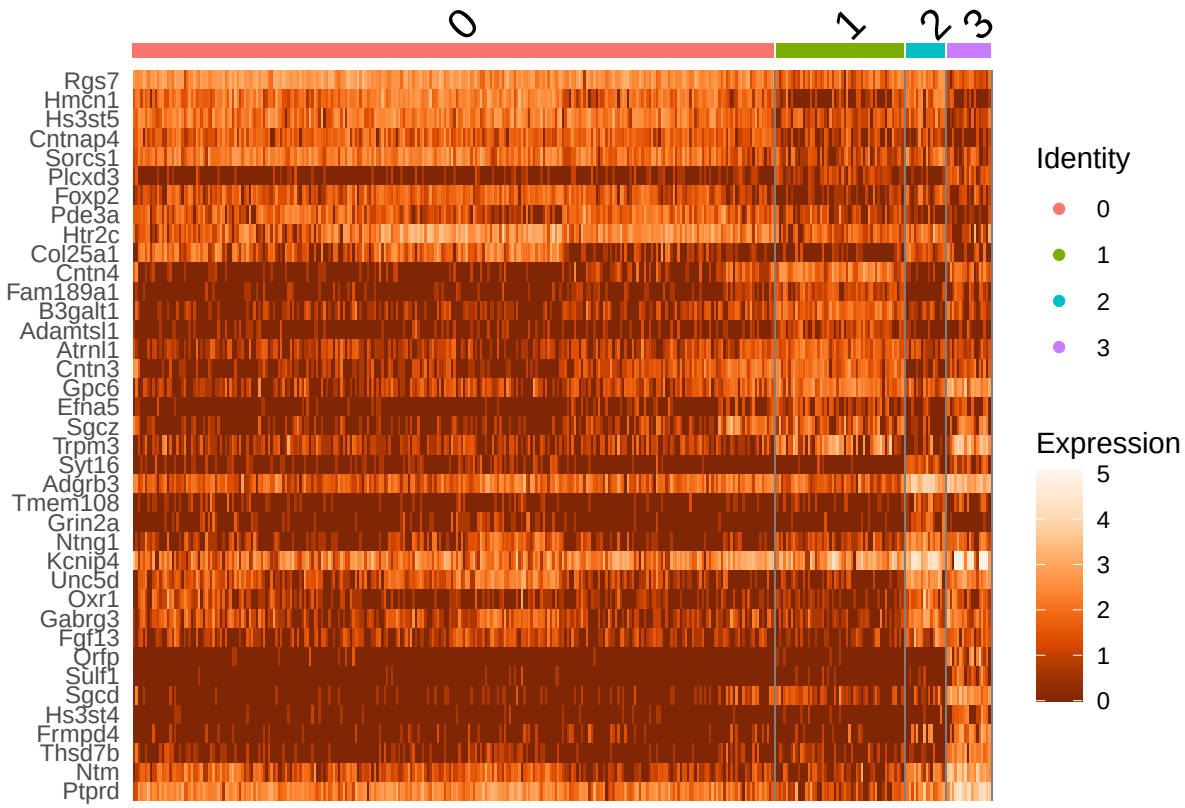
```
## Scale for 'fill' is already present. Adding another scale for 'fill', which
## will replace the existing scale.
```




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

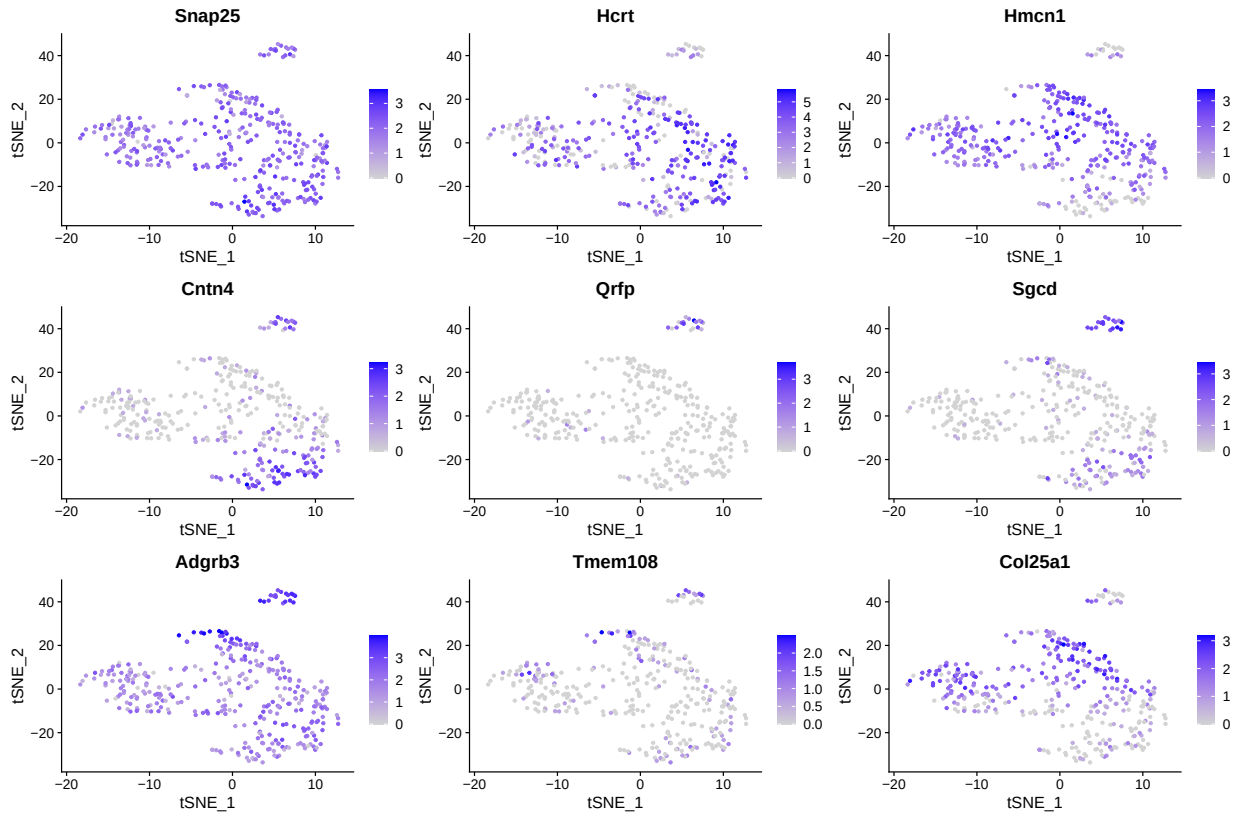
```
## Warning in RColorBrewer::brewer.pal(n = 10, name = "Oranges"): n too large, allowed maximum for pale
## Returning the palette you asked for with that many colors
```

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

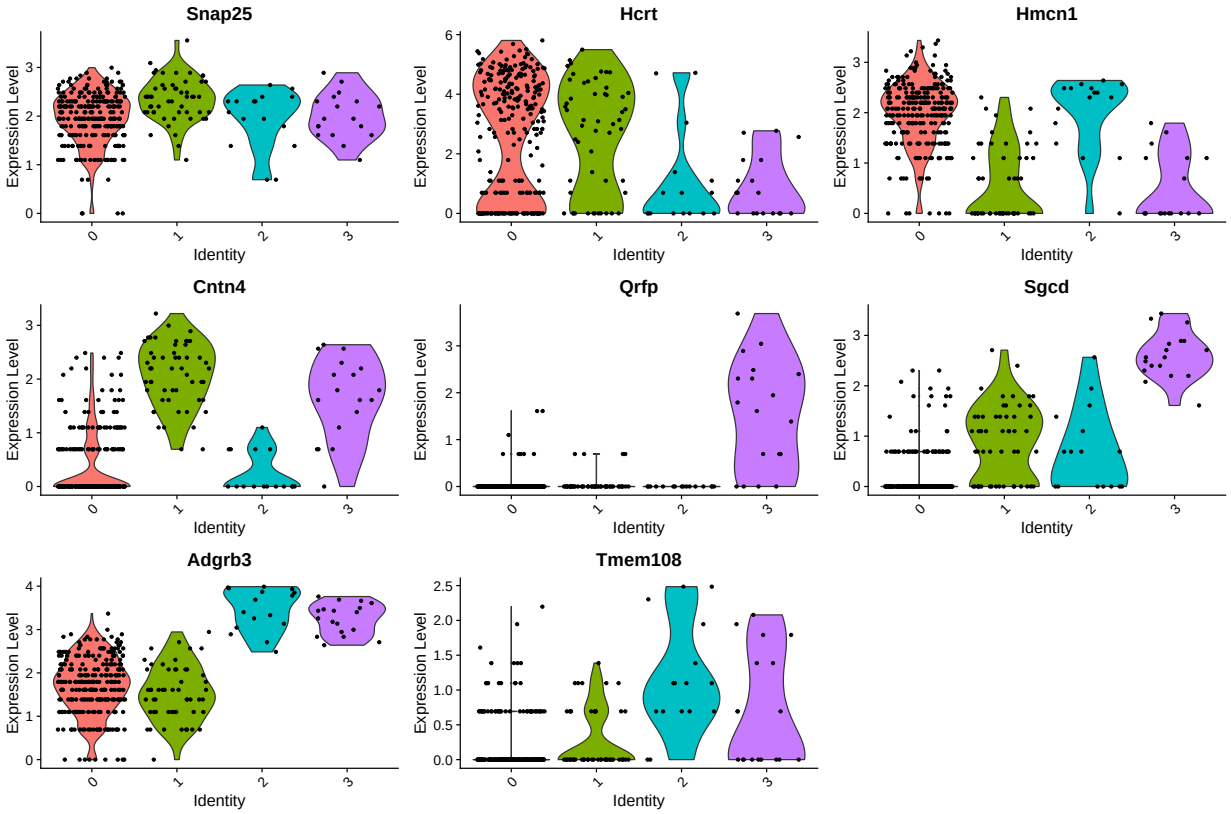


Featureplots of marker genes

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

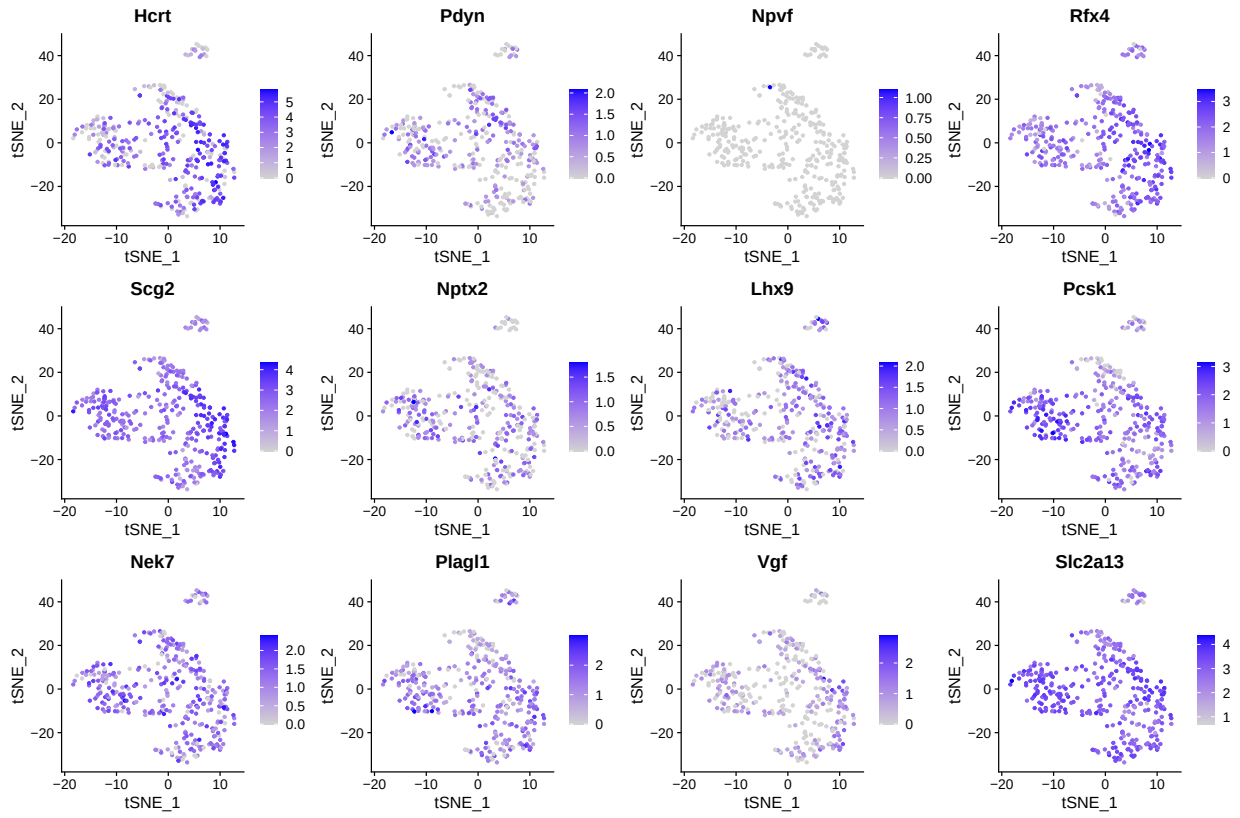



```
VlnPlot(hcrt,features=c("Snap25",'Hcrt',"Hmcn1","Cntn4","Qrfp","Sgcd","Adgrb3","Tmem108"),group.by = "m
```



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   0 135   0 131
##    1   0  22   0  31
##    2   0  13   0   3
##    3   0   9   0   9
```

percentage.

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

```
##
##      YM2        YF2 OM2        OF2
##    0     0.75418994     0.75287356
##    1     0.12290503     0.17816092
##    2     0.07262570     0.01724138
##    3     0.05027933     0.05172414
```

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)
```

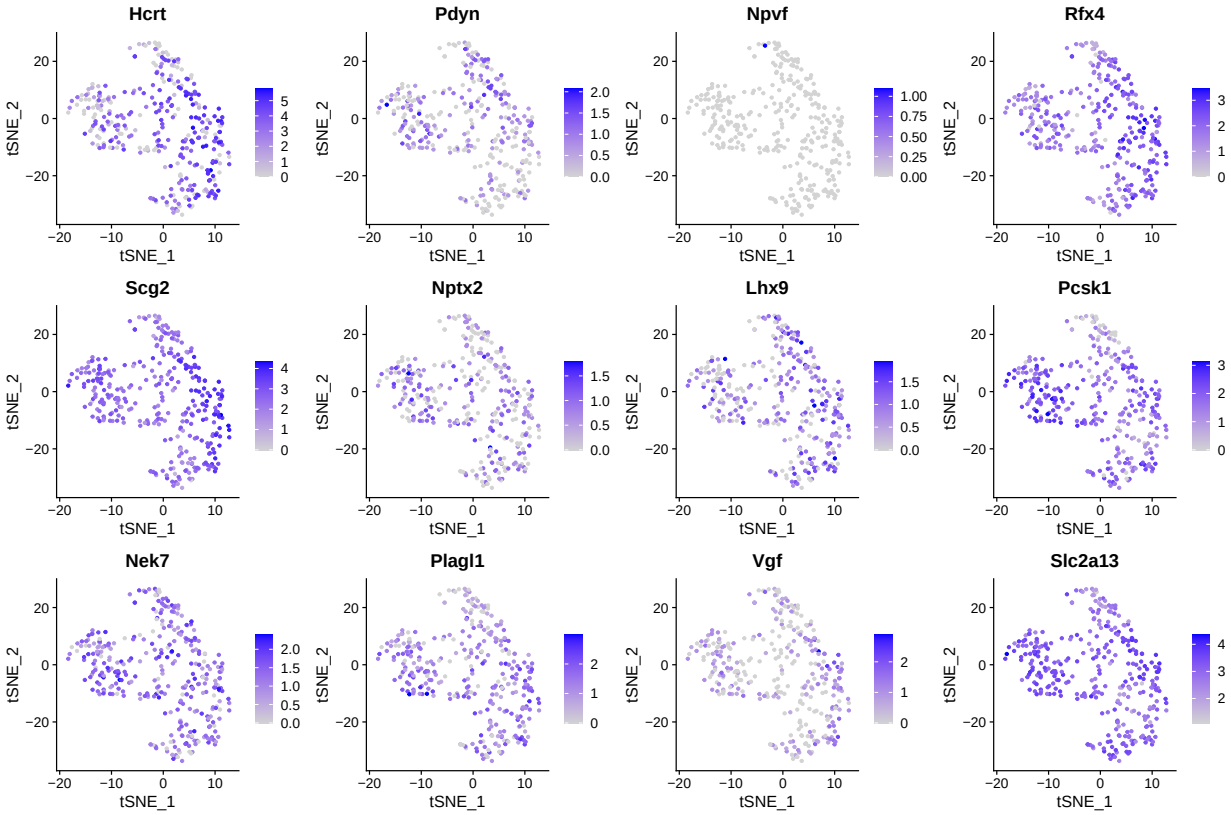
```
##
## young   old
##   170   165
```

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

```
##
## YM2 YF2 OM2 OF2
##   0 170   0 165
```

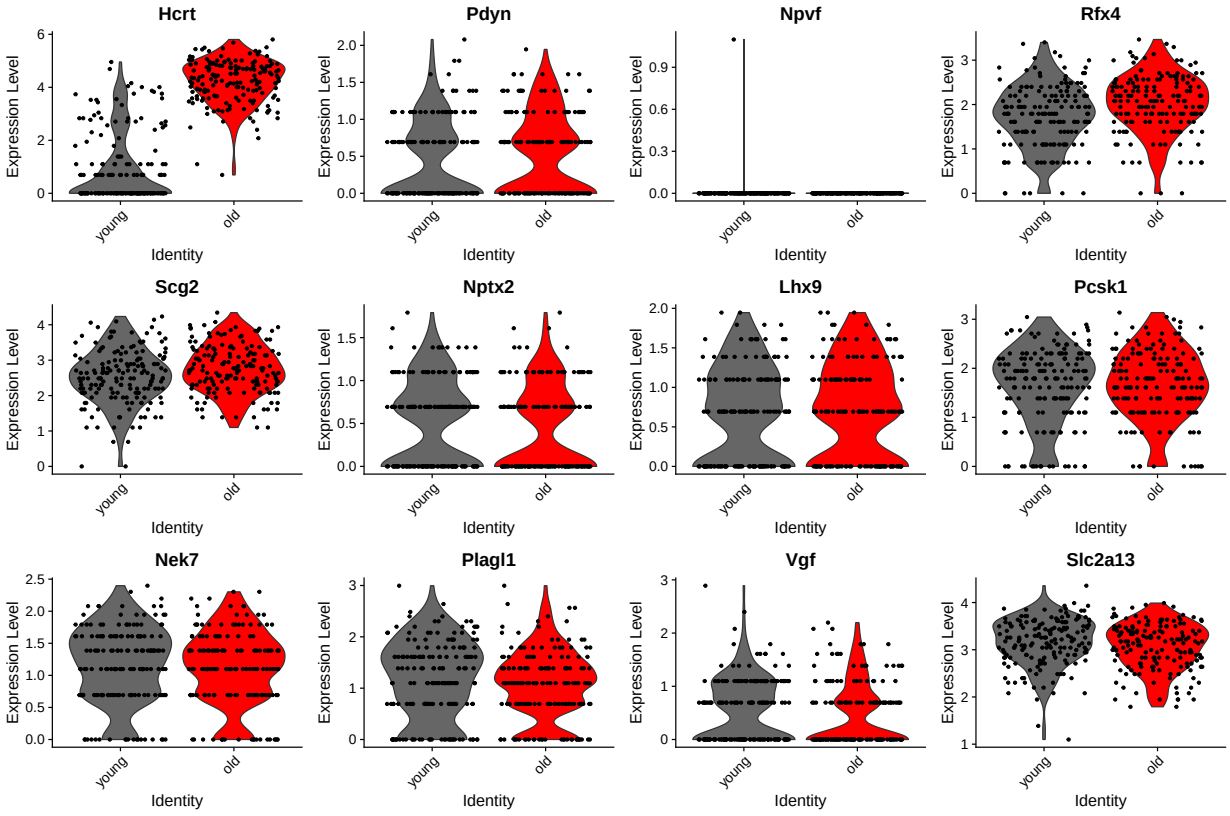
first though, Mickelsen's genes:

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



and split by age

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

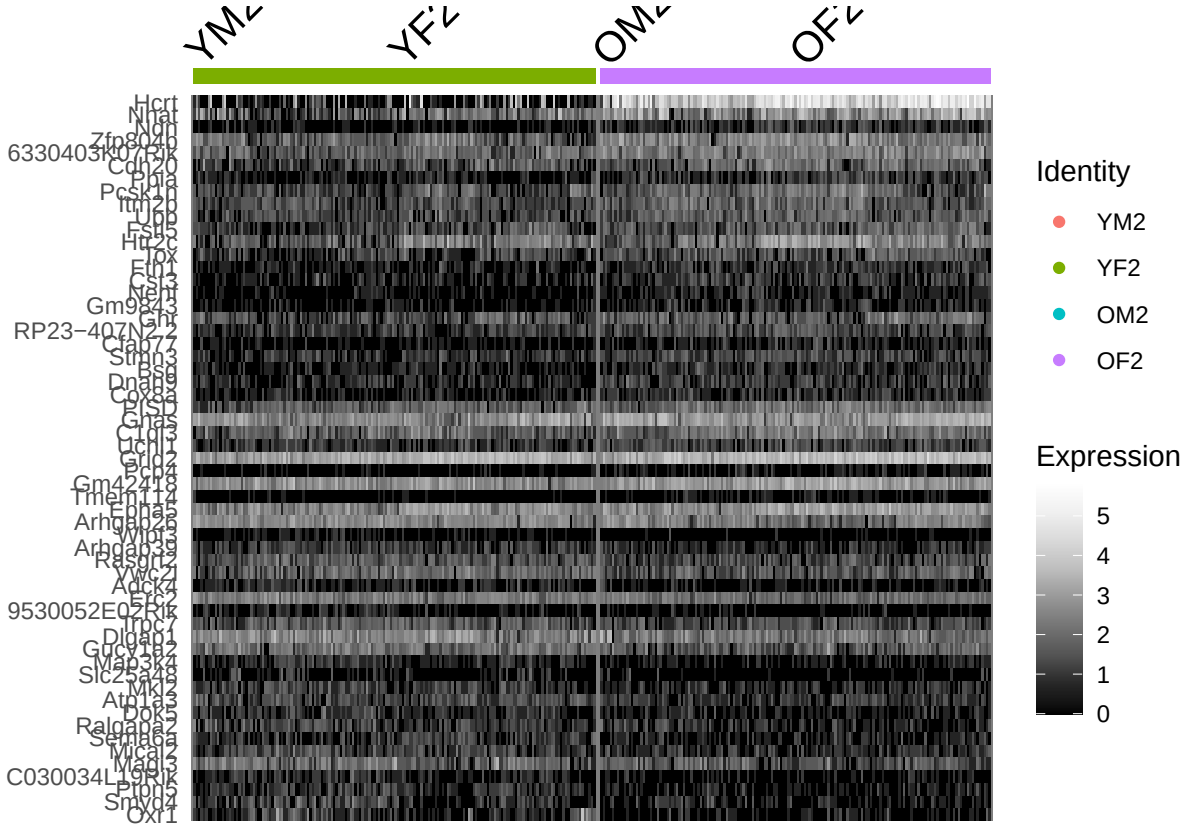



```
agemarkers<-FindMarkers(hcrtcore,ident.2="young",ident.1="old")
#agemarkers
```

plot heatmap of sorted genes.

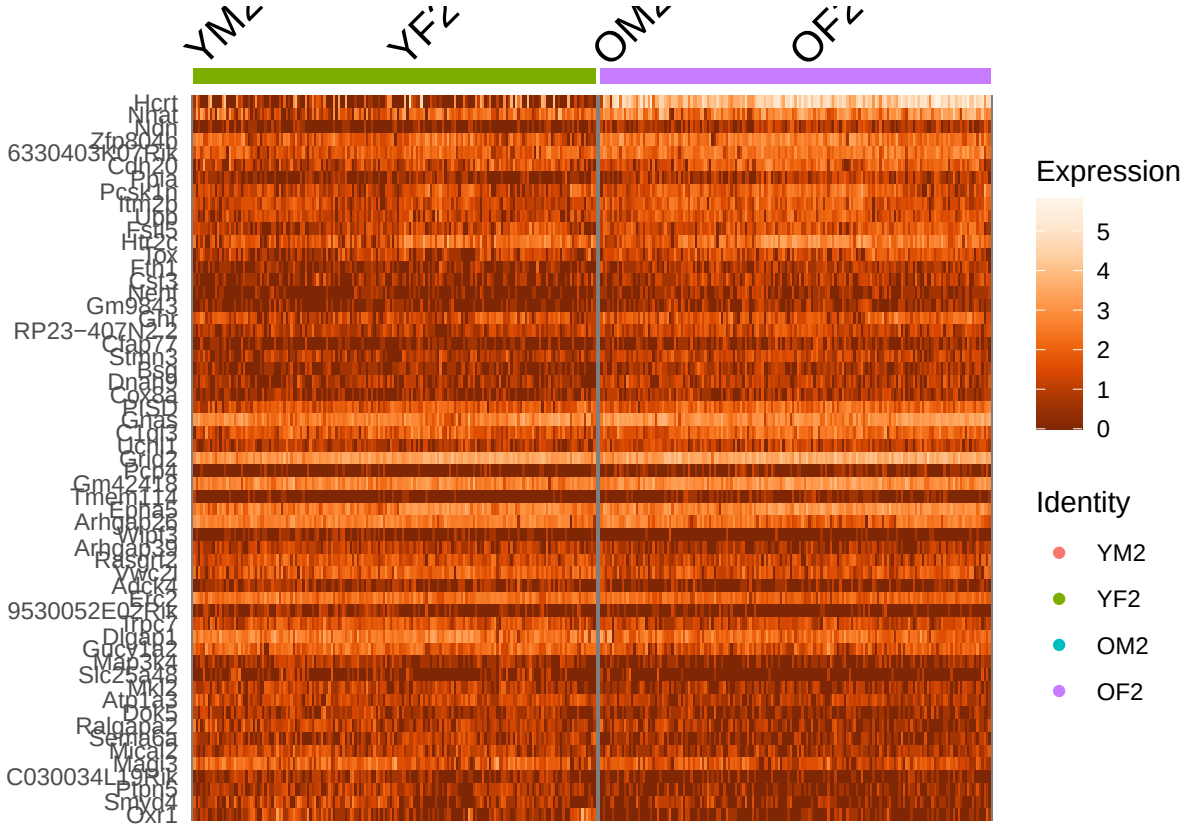
```
a<-agemarkers[agemarkers$p_val_adj<0.05,]
b<-agemarkers[order(a$avg_logFC,decreasing=T),]
DoHeatmap(hcrtcore,features=rownames(b),group.by='librarynames',slot='data')+
  scale_fill_gradientn(colors = rev(RColorBrewer::brewer.pal(n = 9, name = "Greys")))
```

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



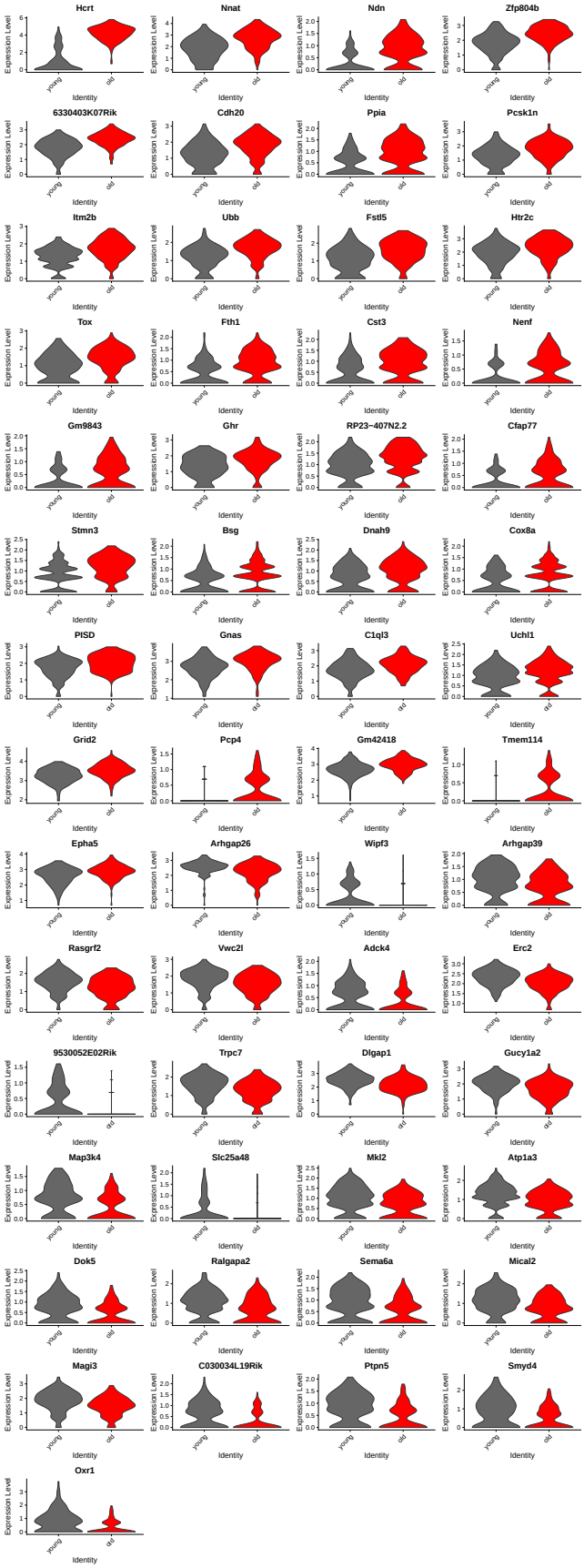

```
DoHeatmap(hcrtcore,features=rownames(b),group.by='librarynames',slot='data')+
  scale_fill_gradientn(colors = rev(RColorBrewer::brewer.pal(n = 9, name = "Oranges")))
```

```
## 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"))
```


Hcrt
Nnat
Ndn
Zfp804b
6330403K07Rik
Cdh20
Ppia
Pcsk1n
Itm2b
Ubb
Fstl5
Htr2c
Tox
Fth1
Cst3
Nenf
Gm9843
Ghr
RP23−407N2.2
Cfap77
Stmn3
Bsg
Dnah9
Cox8a
PISD
Gnas
C1ql3
Uchl1
Grid2
Pcp4
Gm42418
Tmem114
Epha5
Arhgap26
Wipf3
Arhgap39
Rasgrf2
Vwc2l
Adck4
Erc2
9530052E02Rik
Trpc7
Dlgap1
Gucy1a2
Map3k4
Slc25a48
Mkl2
Atp1a3
Dok5
Ralgapa2
Sema6a
Mical2
Magi3
C030034L19Rik
Ptpn5
Smyd4
Oxr1
Expression Level
Identity
young
old

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
```

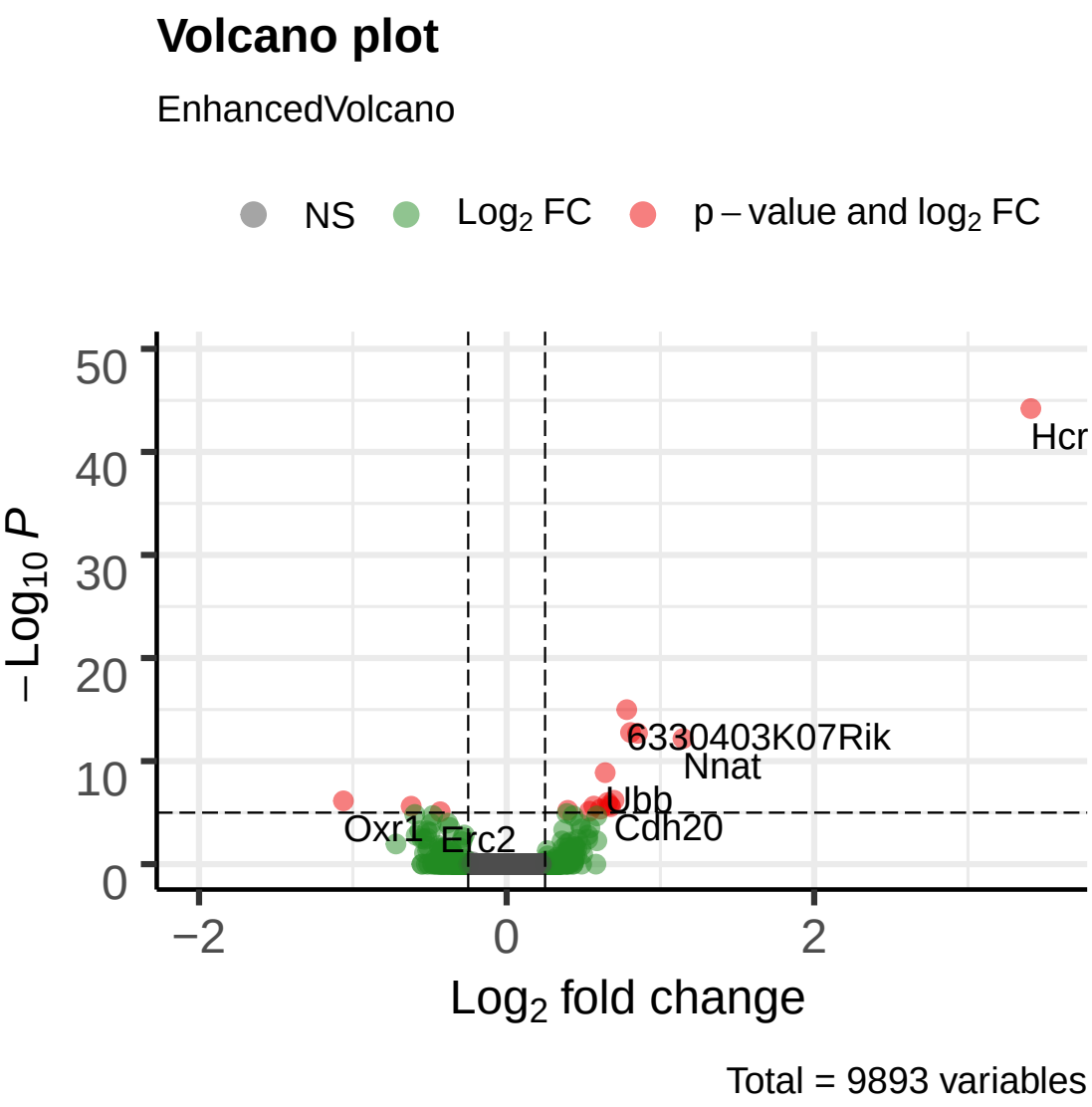
```
# convert from ln to log2 fold change.
agemarkers2$avg_log2FC<-agemarkers2$avg_logFC*log2(exp(1))
library(EnhancedVolcano)
```

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

```
EnhancedVolcano(agemarkers2,
                lab=rownames(agemarkers2),
                x='avg_log2FC',
                y='p_val_adj',
                xlim=c(-2,3.5),
                FCcutoff=0.25,
                transcriptPointSize=3,
                transcriptLabSize = 5)
```

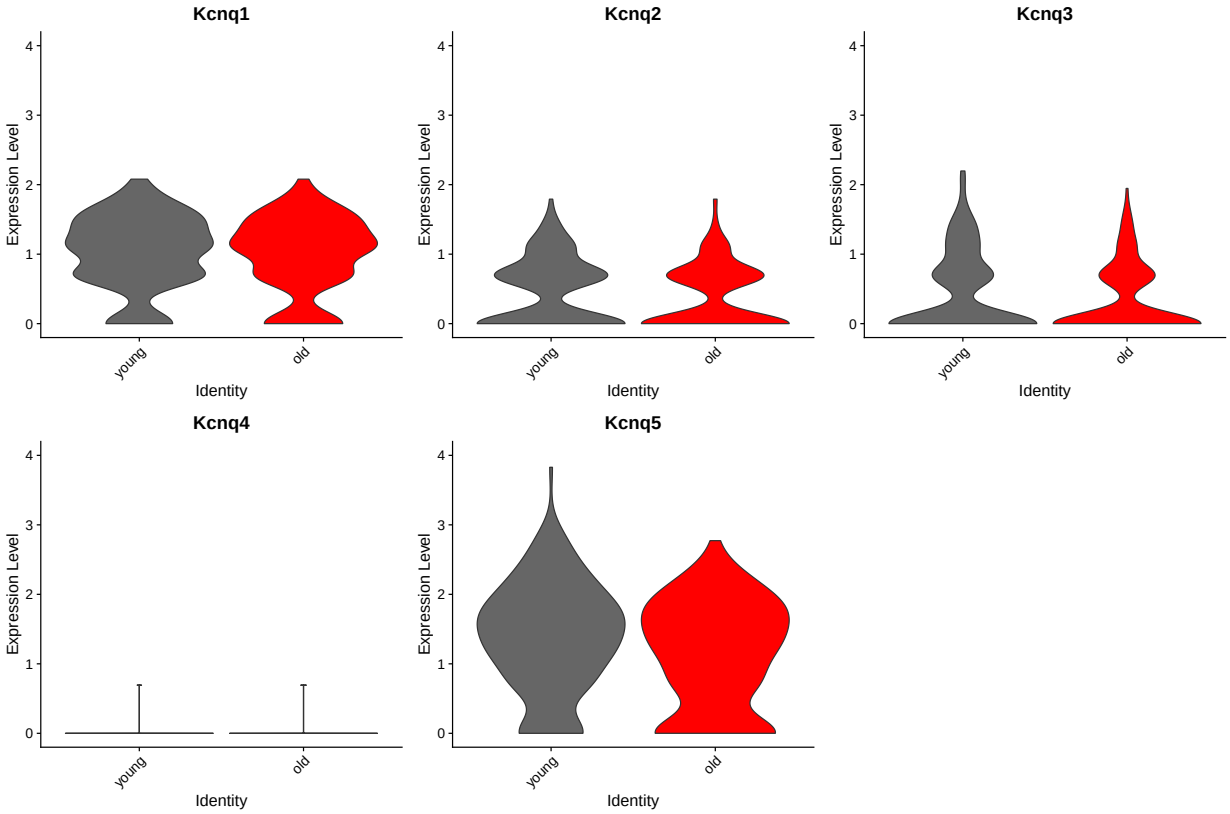
```
## Warning in EnhancedVolcano(agemarkers2, lab = rownames(agemarkers2), x =
## "avg_log2FC", : transcriptPointSize argument deprecated in v1.4 - please use
## pointSize
```

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



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 86.470588
## Kcnq2          Kcnq2 young 54.117647
## Kcnq3          Kcnq3 young 44.705882
## Kcnq4          Kcnq4 young  2.941176
## Kcnq5          Kcnq5 young 86.470588
## Kcnq11         Kcnq1   old 82.424242
## Kcnq21         Kcnq2   old 47.272727
## Kcnq31         Kcnq3   old 36.969697
## Kcnq41         Kcnq4   old  3.636364
## Kcnq51         Kcnq5   old 75.757576
```

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