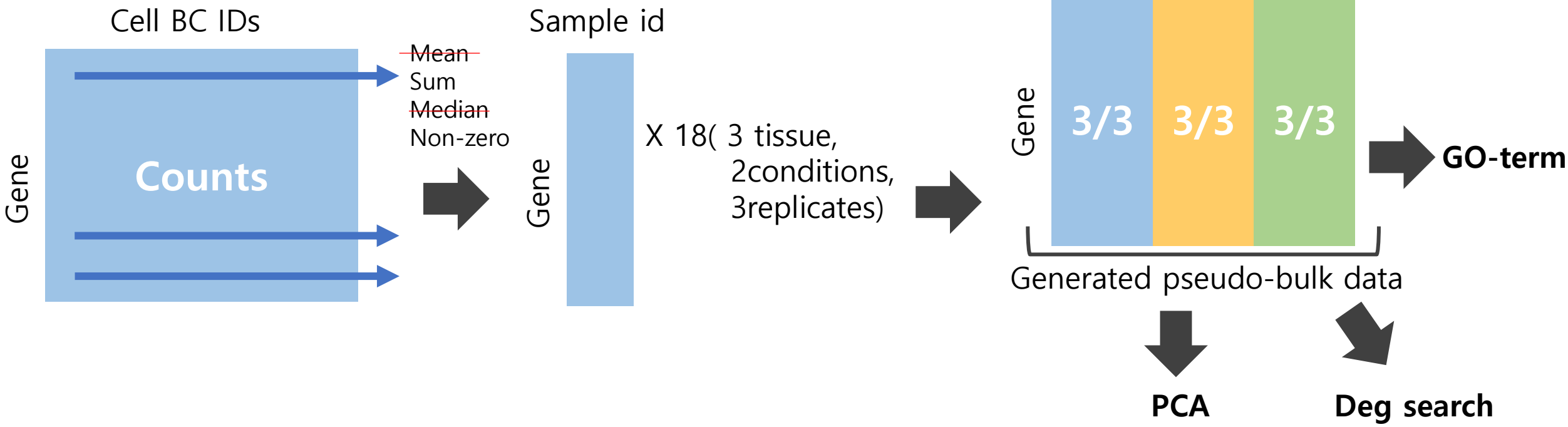
In silico bulk(Pseudo-bulk) using muscat

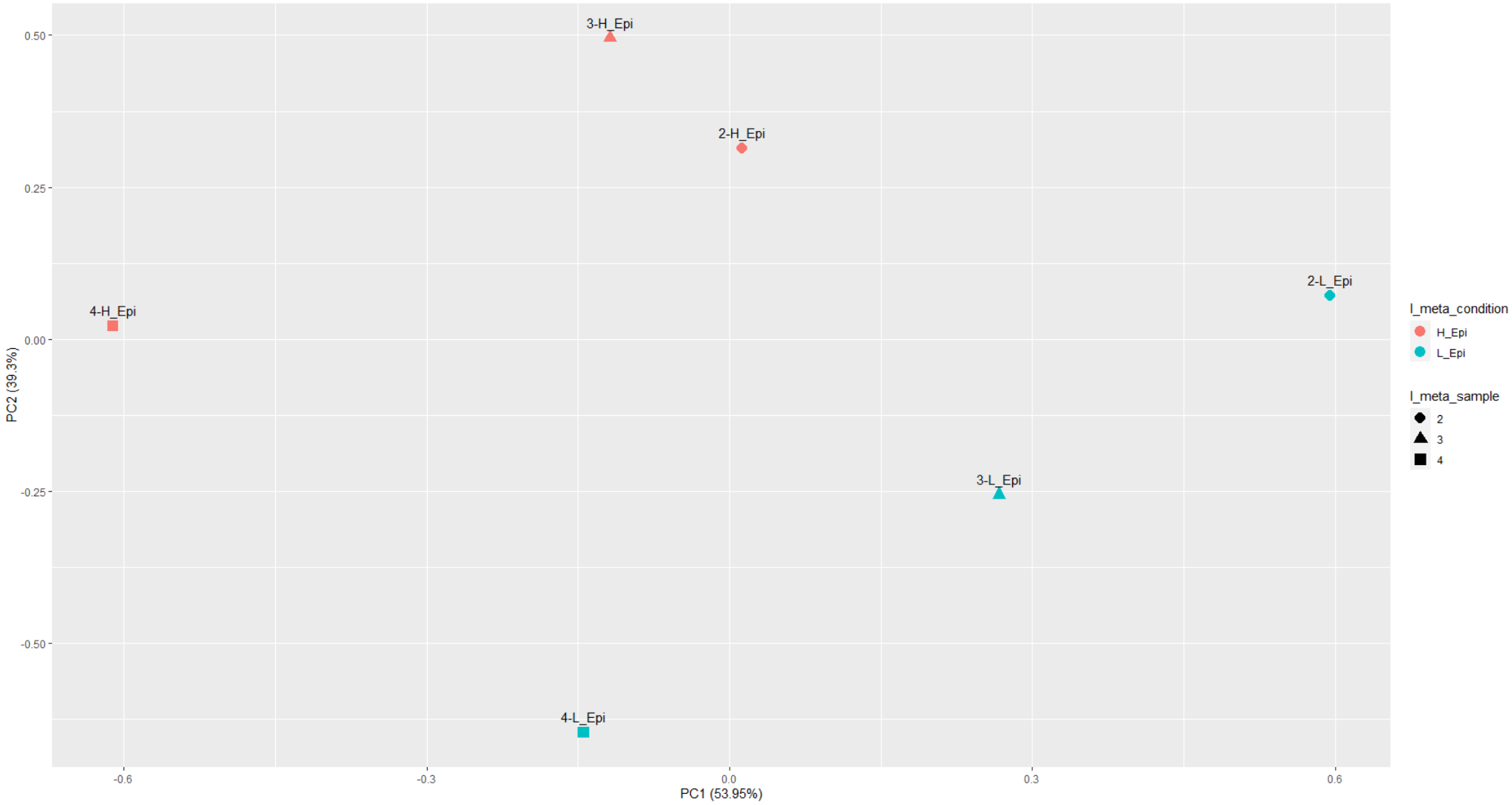
`muscat` (**Mu**lti-sample **mu**lti-group **sc**RNA-seq **a**nalysis **t**ools) provides various methods for *Differential State* (DS) analyses in multi-sample, multi-group, multi-(cell-)subpopulation scRNA-seq data, as elaborated in our preprint:

> Crowell HL, Soneson C*, Germain P-L*,
> Calini D, Collin L, Raposo C, Malhotra D & Robinson MD:
> On the discovery of population-specific state transitions from
> multi-sample multi-condition single-cell RNA sequencing data.
> *bioRxiv* **713412** (July, 2019). doi: 10.1101/713412

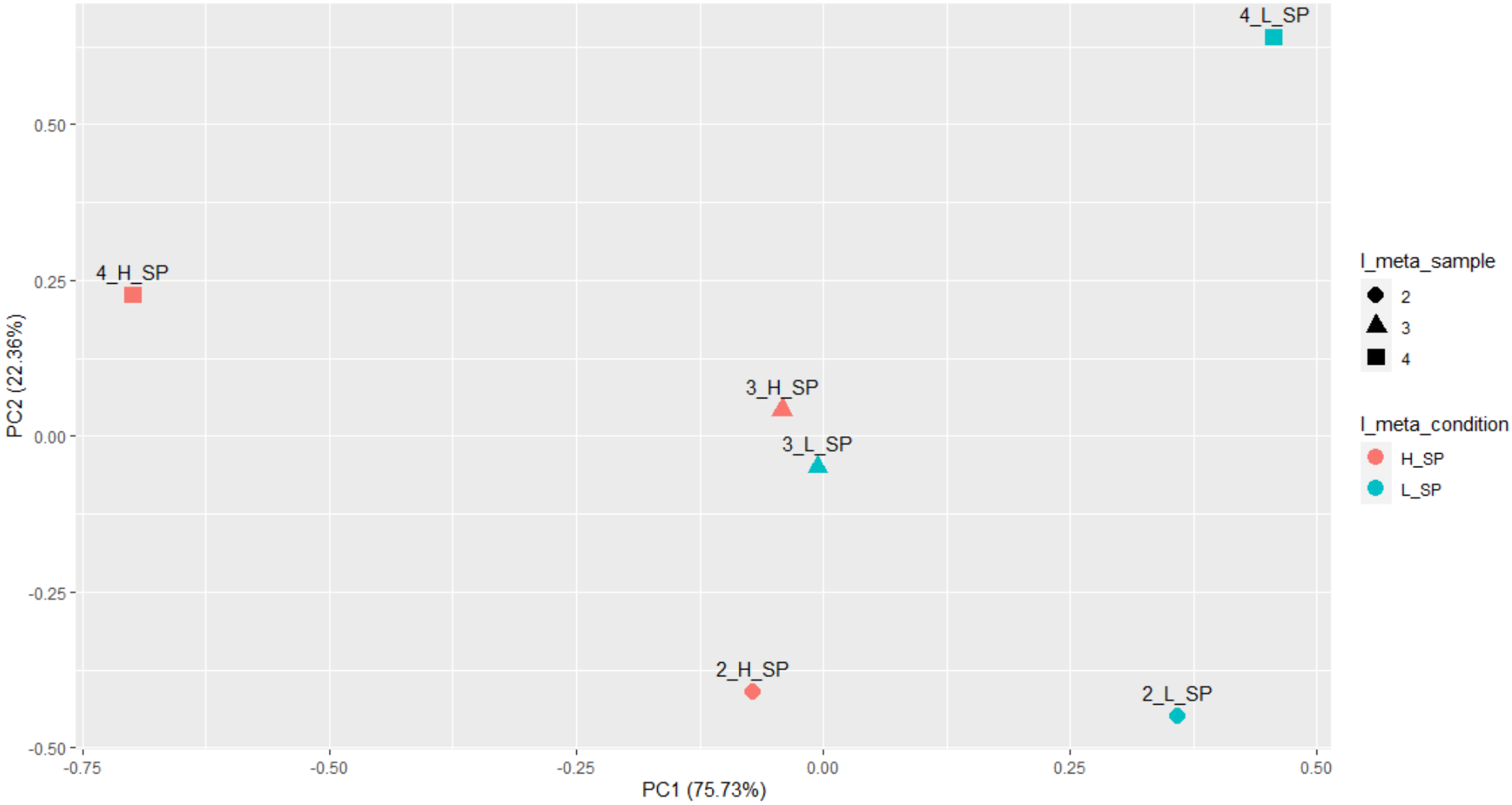
*These authors contributed equally.



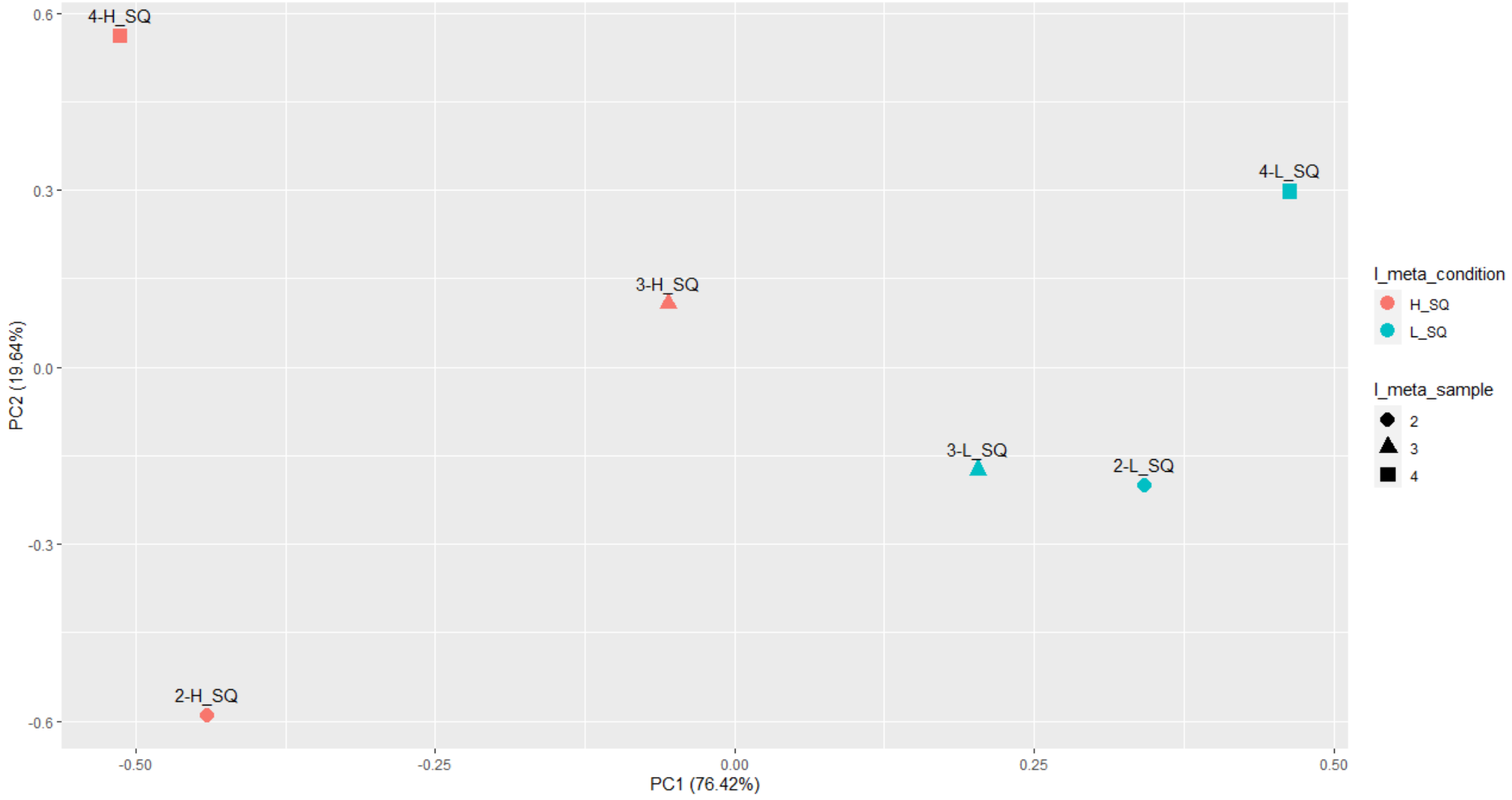
Pseudo-bulk for 6 **Epi** using **sum** aggregated



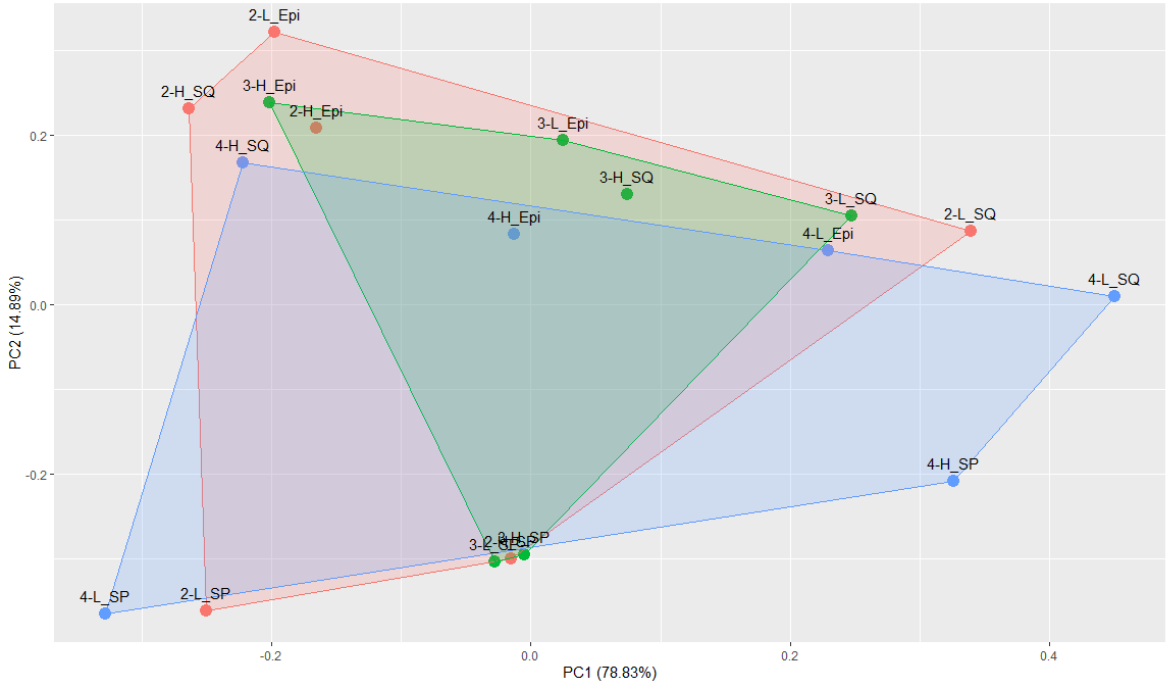
Pseudo-bulk for 6 **SP** using **sum** aggregated



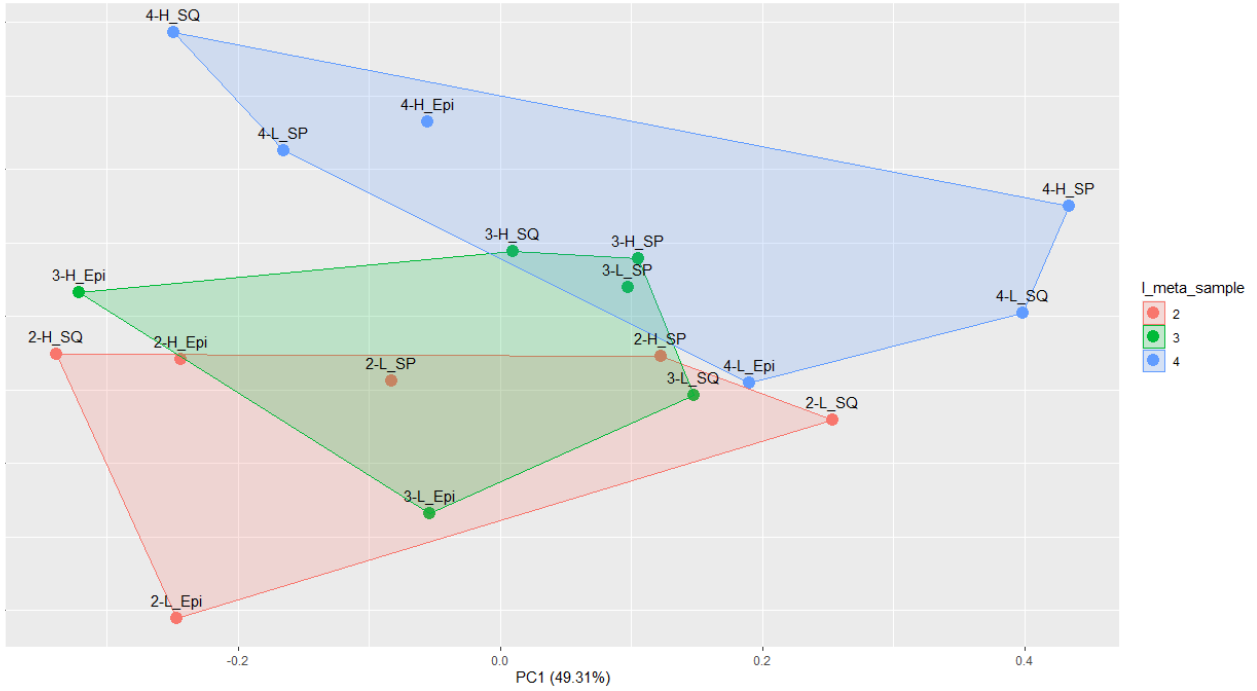
Pseudo-bulk for 6 **SQ** using **sum** aggregated



Pseudo-bulk PCA for 18samples for detecting sample-batch

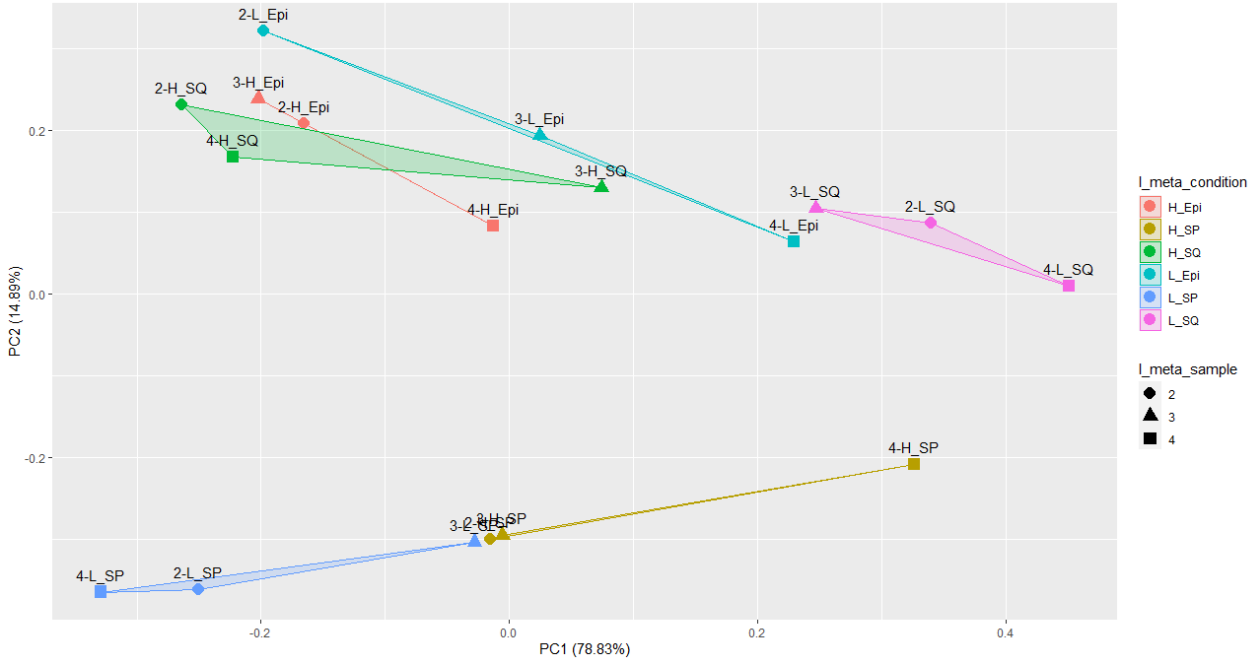


Num-detected

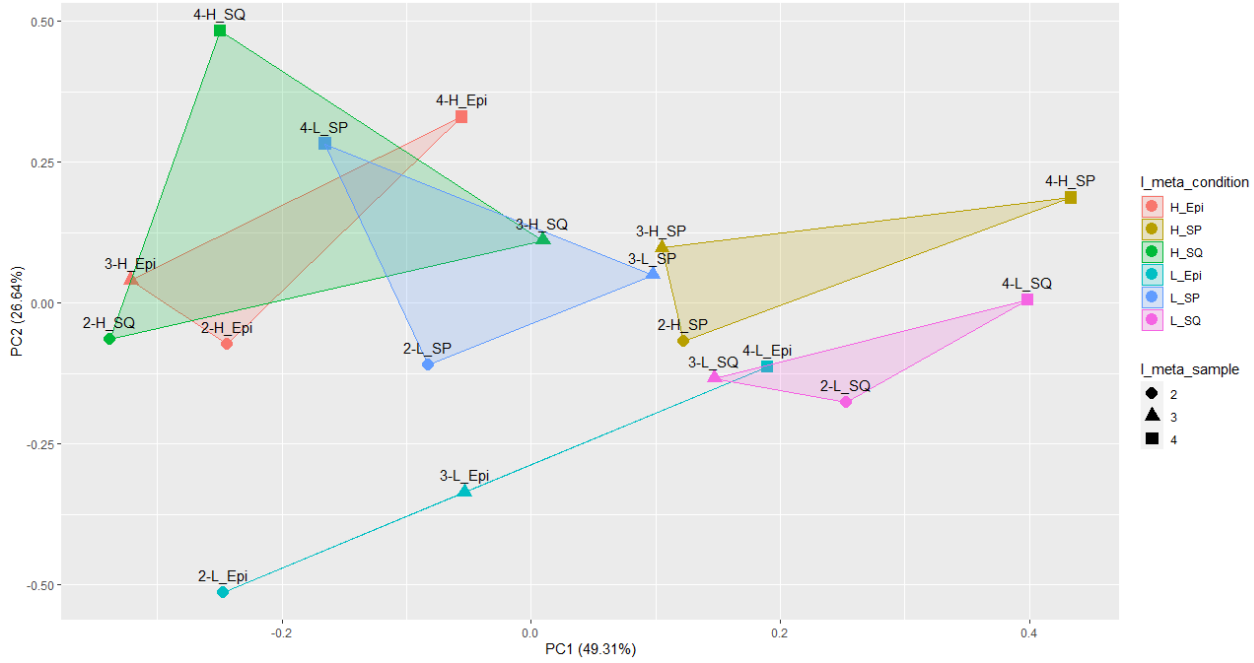


Sum

Pseudo-bulk PCA for 18samples, colored by condition



Num-detected



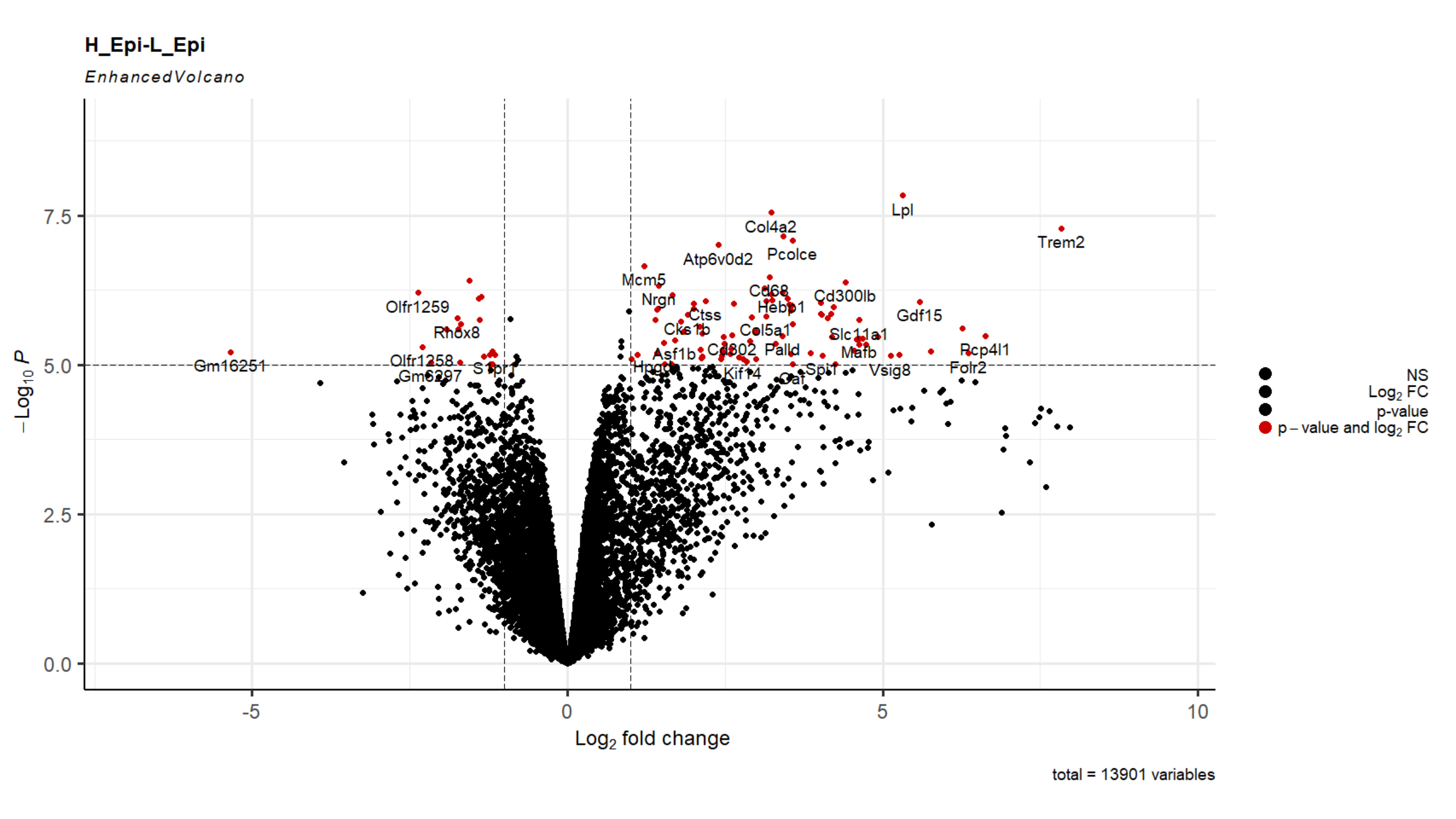
Sum

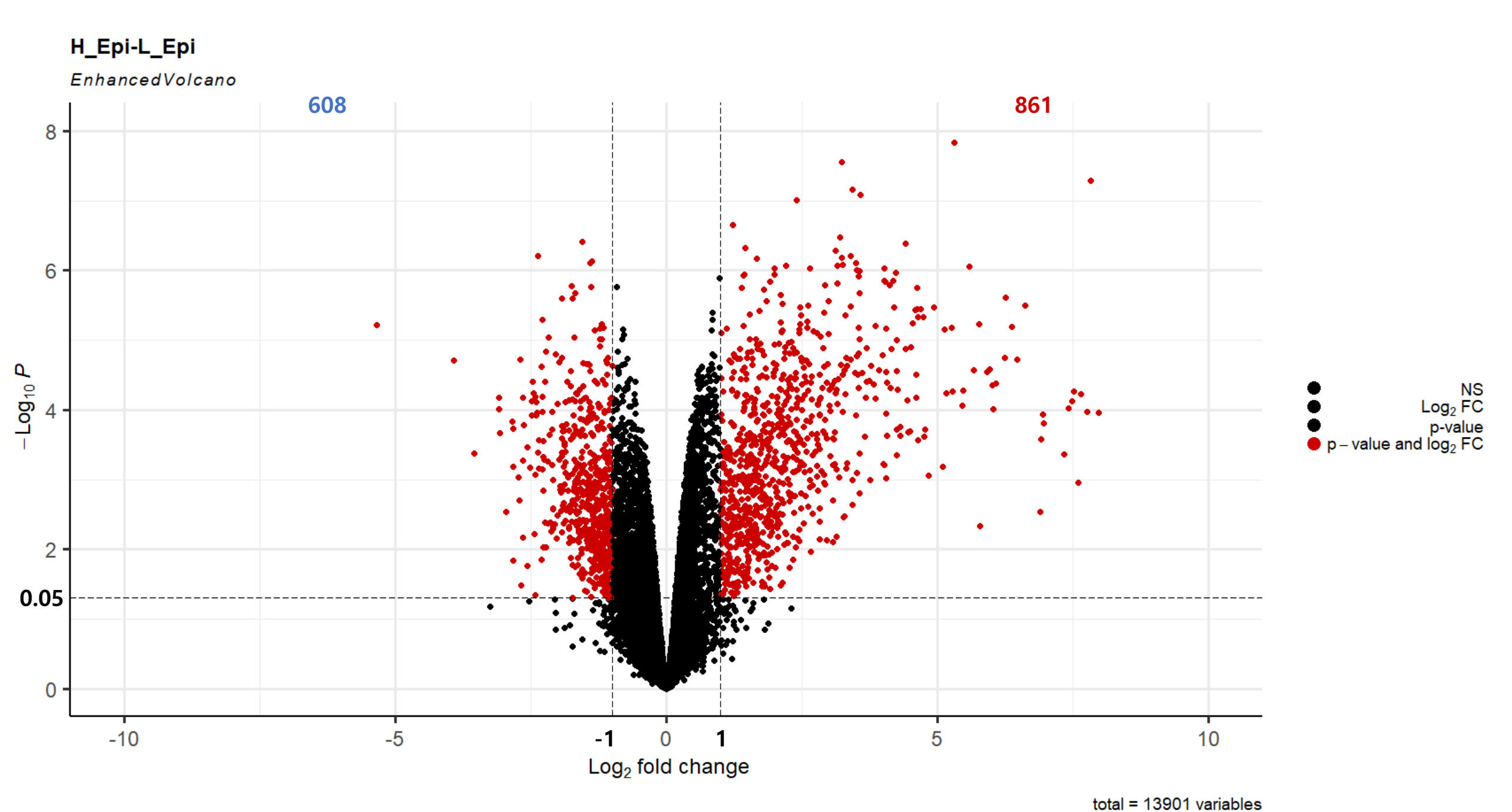
4월1일 이후 추가내용

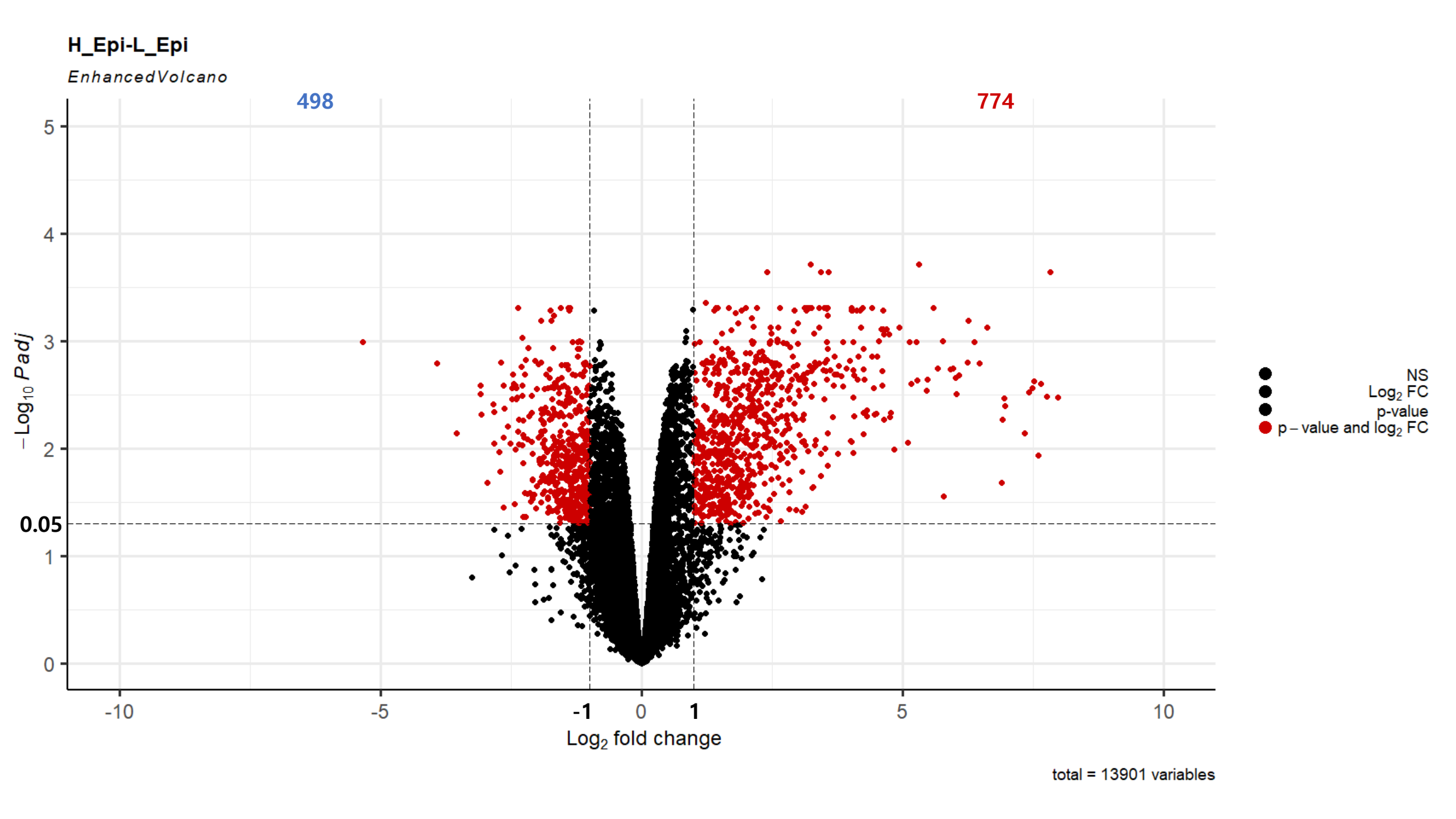
Muscat 수정해서 edgeR 제대로 돌림. <- interactive plot만듦(검색가능하게 바꾸기)
~~Deg파일 정리해서 ppt에 올려놓겠습니다~~
~~Volcano 보기 좋게~~

Trem2는 아마 nk cell 아닌듯? <- Lpl 0이 아닌 세포랑, 0인 세포 개수 맞춰서 랜덤으로 비교?

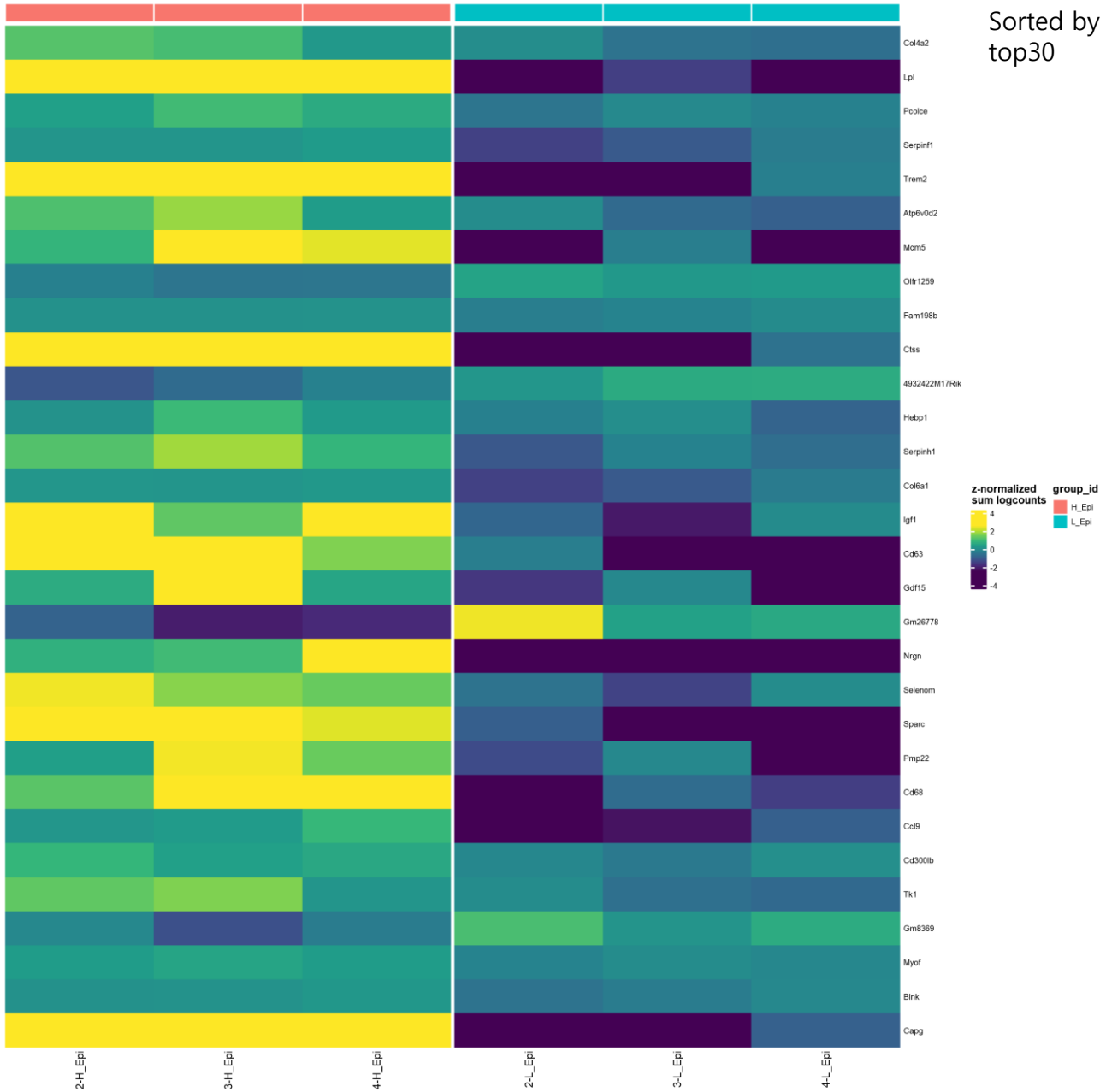
```
> df_pb_sum[grep('Trem2', rownames(df_pb_sum)),]
      2-H_Epi 2-L_Epi 3-H_Epi 3-L_Epi 4-H_Epi 4-L_Epi
Trem2    1327       8    1541       8    2038       0
> df_pb_nz[grep('Trem2', rownames(df_pb_nz)),]
      2-H_Epi 2-L_Epi 3-H_Epi 3-L_Epi 4-H_Epi 4-L_Epi
Trem2     100       4      63       1     128       0
```


H_Epi-L_Epi
EnhancedVolcano
Lpl
Trem2
Col4a2
Pcolce
Atp6v0d2
Mcm5
Nrgn
Cd68
Cd300lb
Hebp1
Gdf15
Olfr1259
Ctss
Cks1b
Col5a1
Slc11a1
Rho6x8
Asf1b
Cd802
Palld
Mafb
Rcp4l1
Olfr1258
Gm16251
Gm6297
S1pr1
Hpgds
Kif14
Cenf
Spi1
Vsig8
Folr2
7.5
5.0
2.5
0.0
-5
0
5
10
$-\text{Log}_{10}\,P$
Log_2 fold change
NS
Log_2 FC
p-value
p – value and log_2 FC
total = 13901 variables


H_Epi-L_Epi
EnhancedVolcano
608
861
8
6
4
2
0.05
0
−Log10 P
-10
-5
-1
0
1
5
10
Log2 fold change
NS
Log2 FC
p-value
p – value and log2 FC
total = 13901 variables


H_Epi-L_Epi
EnhancedVolcano
498
774
5
4
3
2
0.05
1
0
−Log10 Padj
-10
-5
-1
0
1
5
10
Log2 fold change
total = 13901 variables
NS
Log2 FC
p-value
p – value and log2 FC

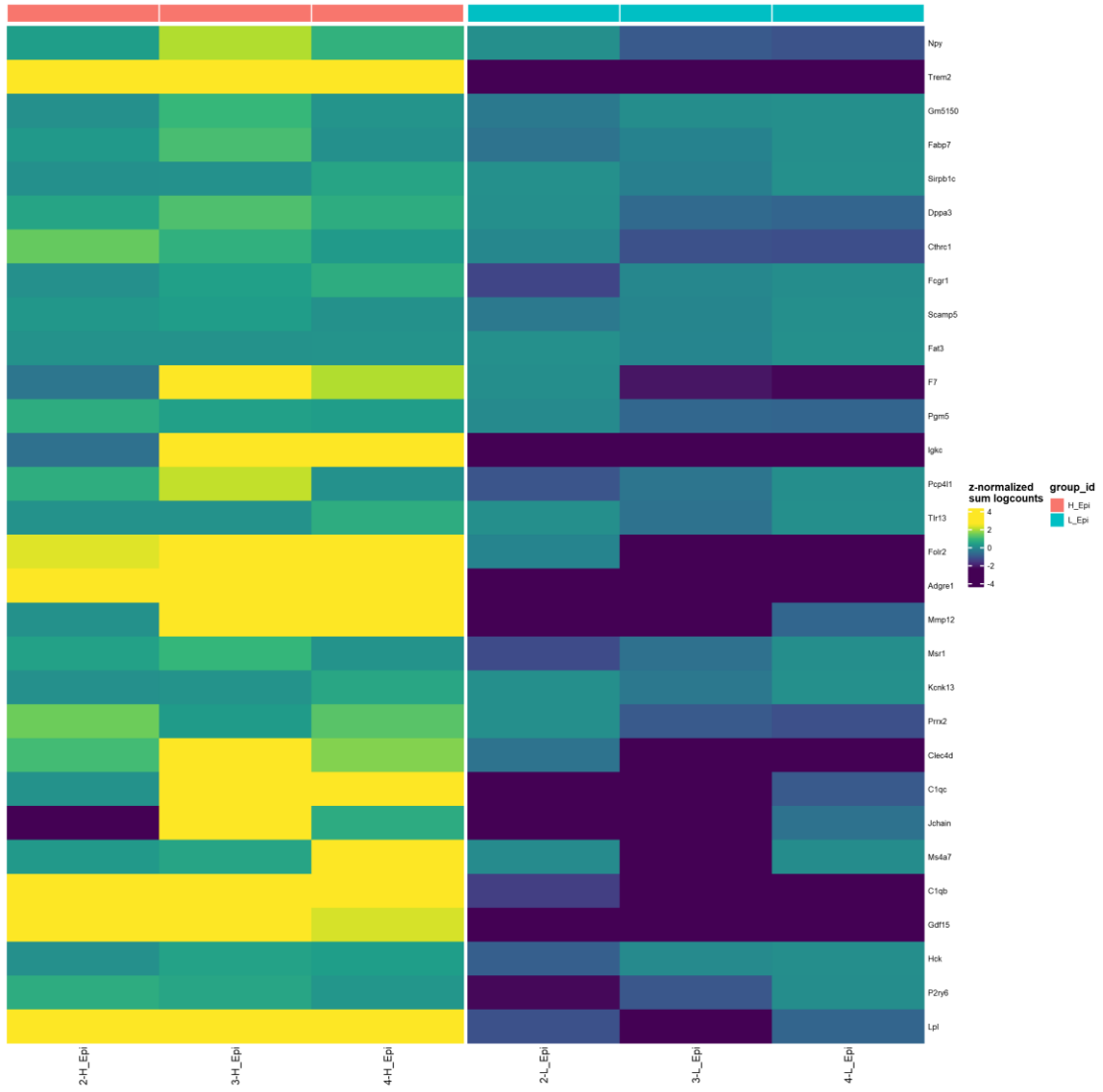
Pseudo-bulk(sum) for Epi, colored by condition

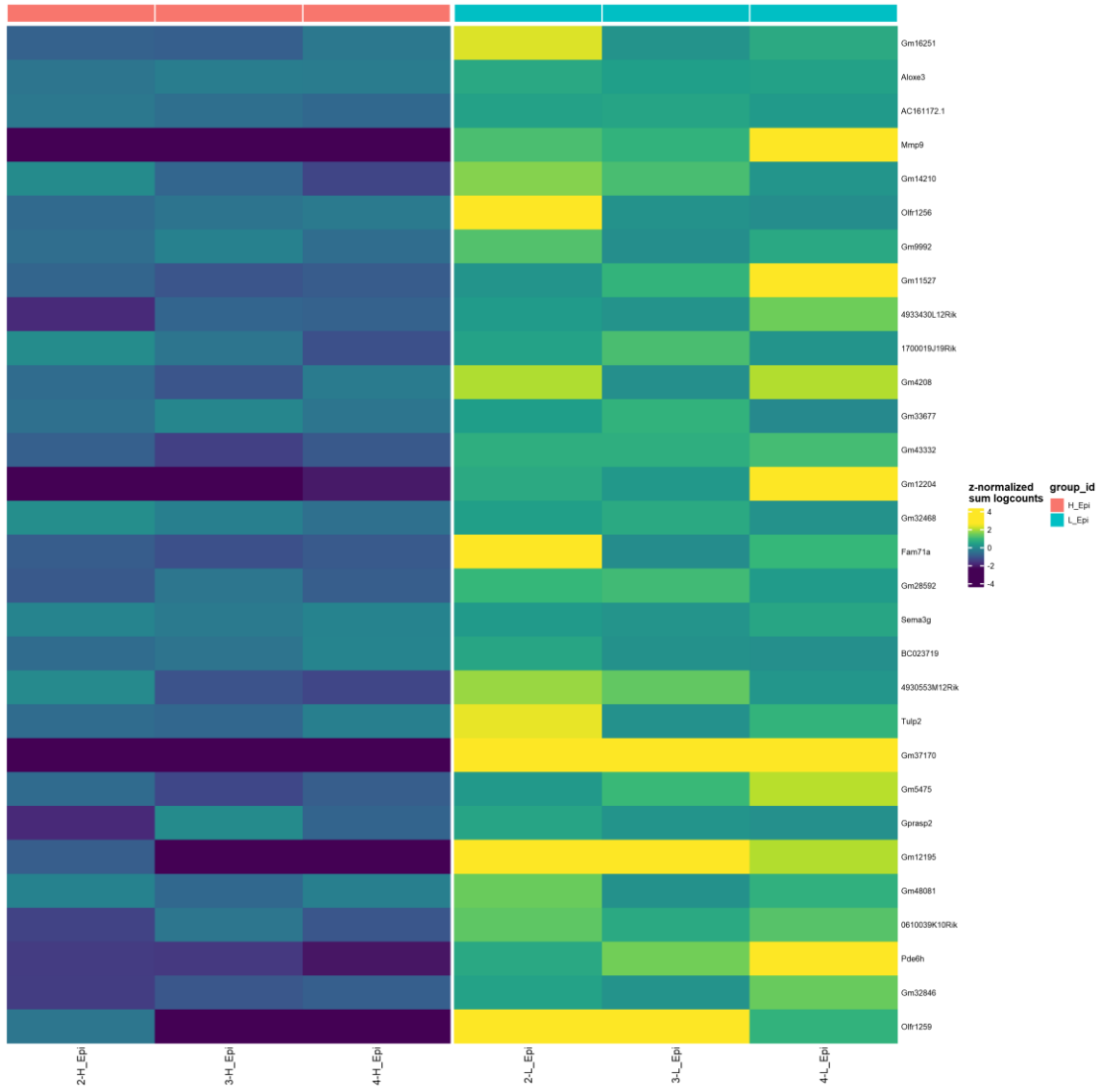


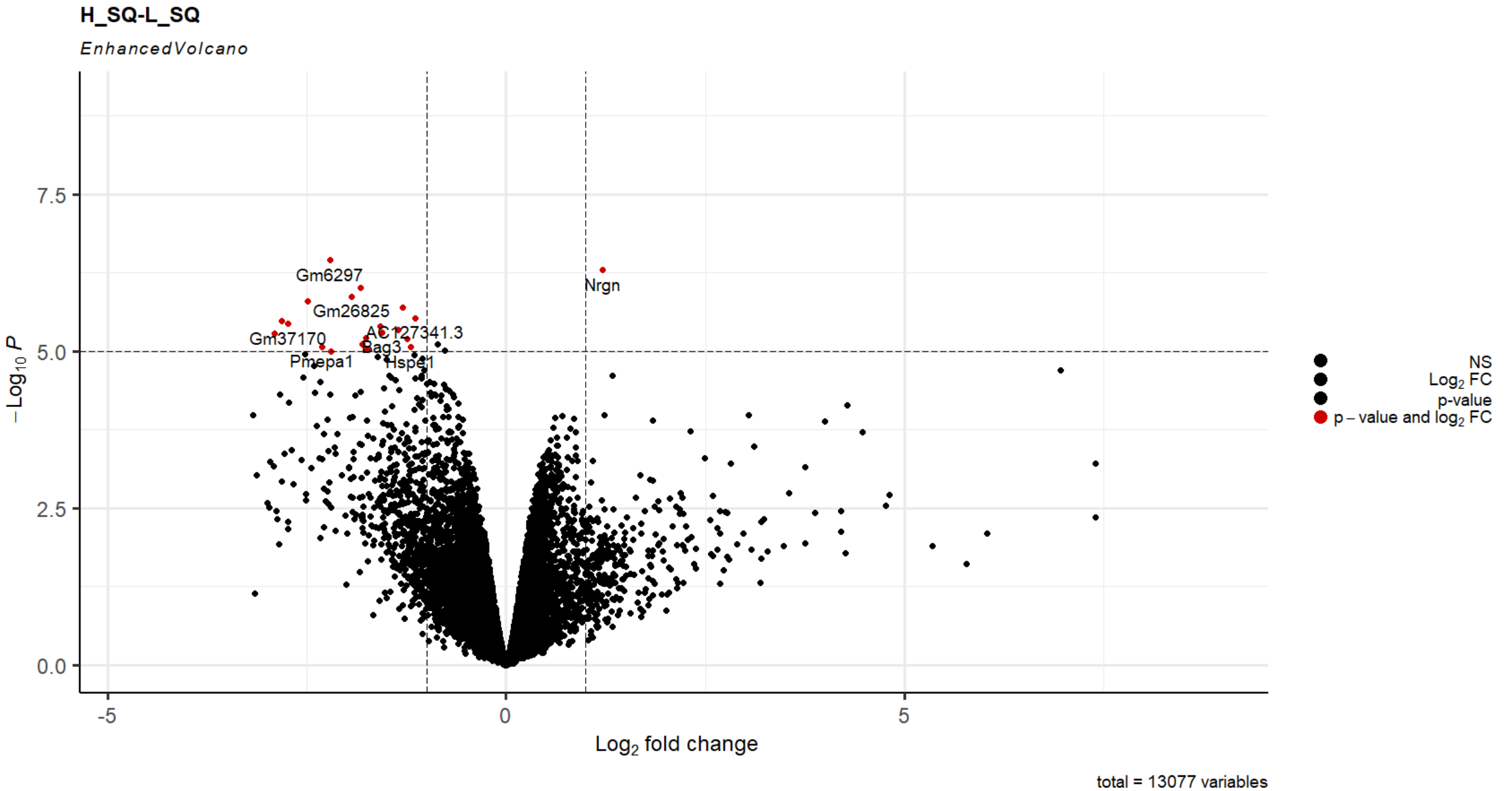
Sorted by p_adj.loc (FDR < 0.05)
top30

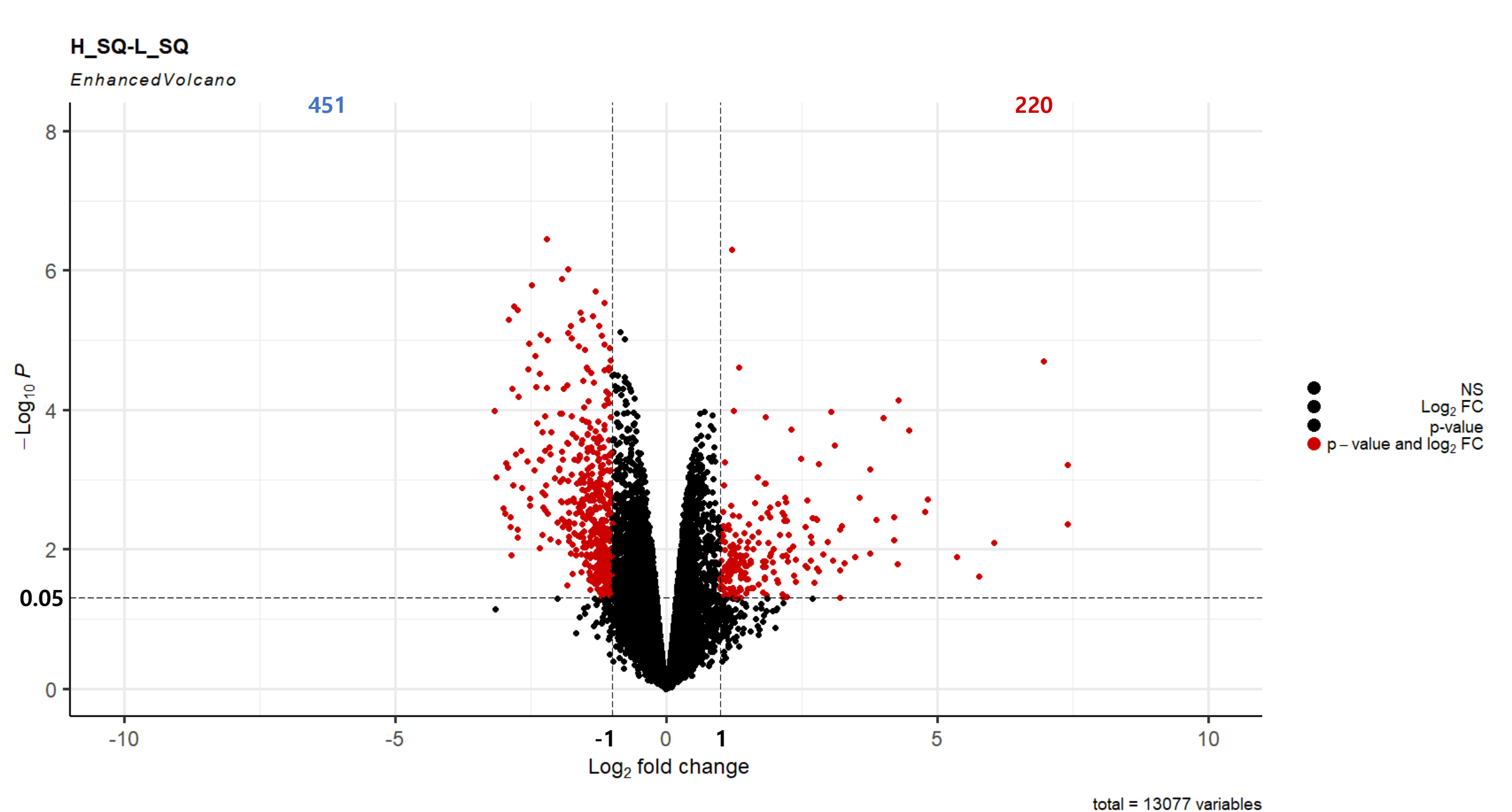
Pseudo-bulk(sum) for Epi, colored by condition

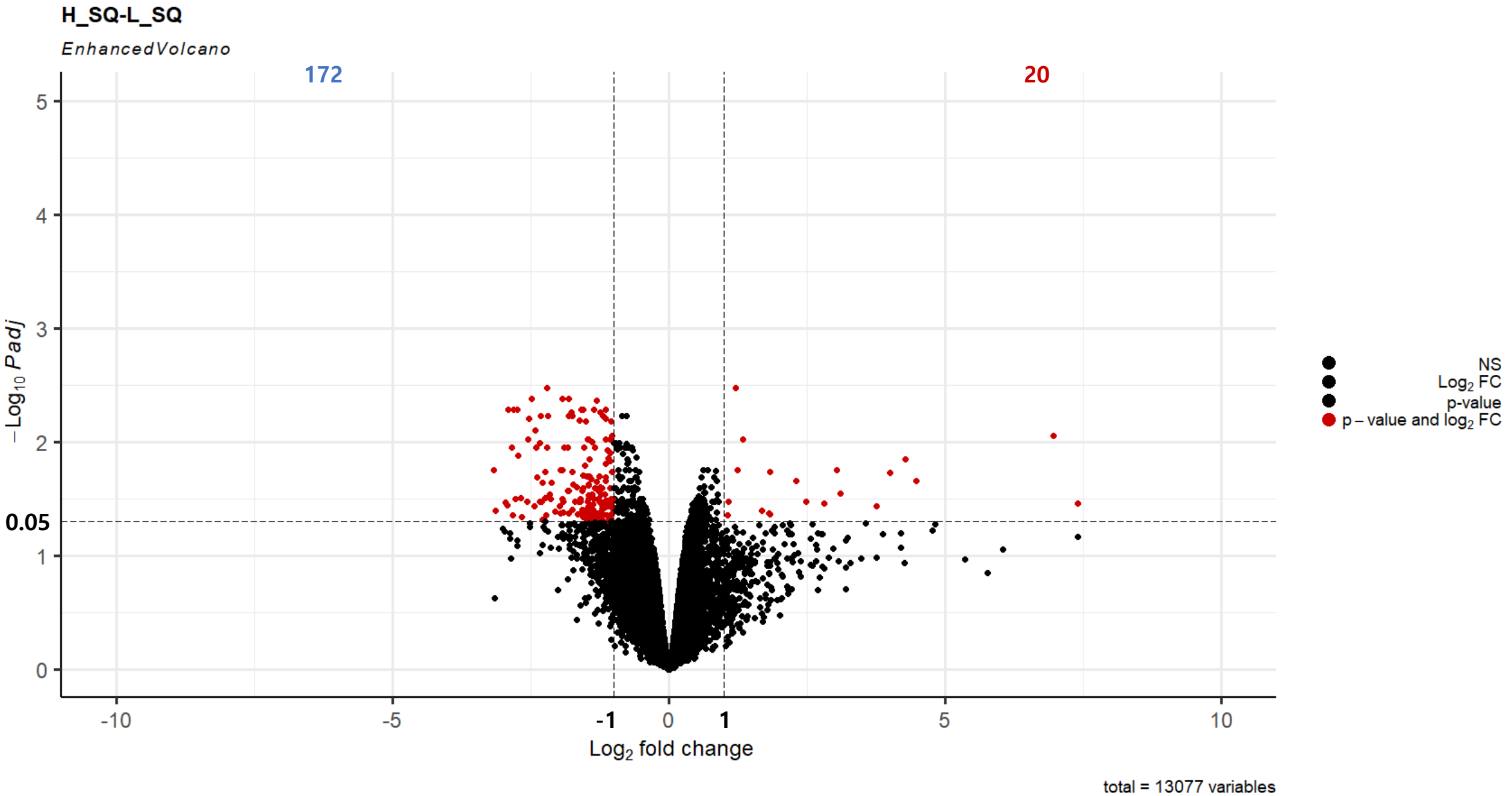
Sorted by logFC (FDR < 0.05)
top30



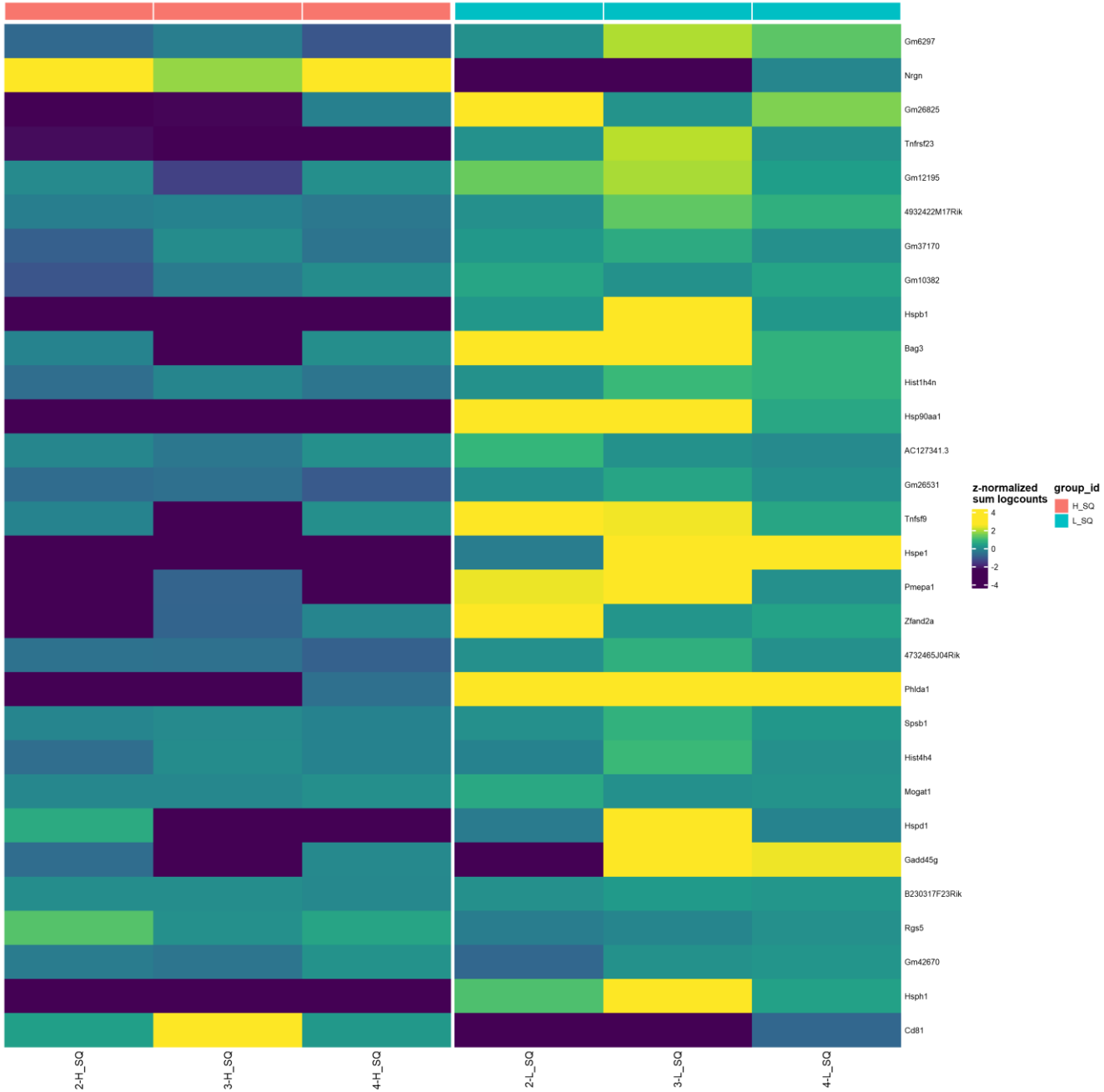



H_SQ-L_SQ
EnhancedVolcano
7.5
5.0
2.5
0.0
−Log10 P
-5
0
5
Log2 fold change
Gm6297
Gm26825
Gm37170
AC127341.3
Bag3
Pmepa1
Hspe1
Nrgn
NS
Log2 FC
p-value
p – value and log2 FC
total = 13077 variables


H_SQ-L_SQ
EnhancedVolcano
451
220
8
6
4
2
0.05
0
−Log10 P
-10
-5
-1
0
1
5
10
Log2 fold change
NS
Log2 FC
p-value
p – value and log2 FC
total = 13077 variables


H_SQ-L_SQ
EnhancedVolcano
172
20
5
4
3
2
0.05
1
0
−Log10 Padj
-10
-5
-1
0
1
5
10
Log2 fold change
NS
Log2 FC
p-value
p – value and log2 FC
total = 13077 variables

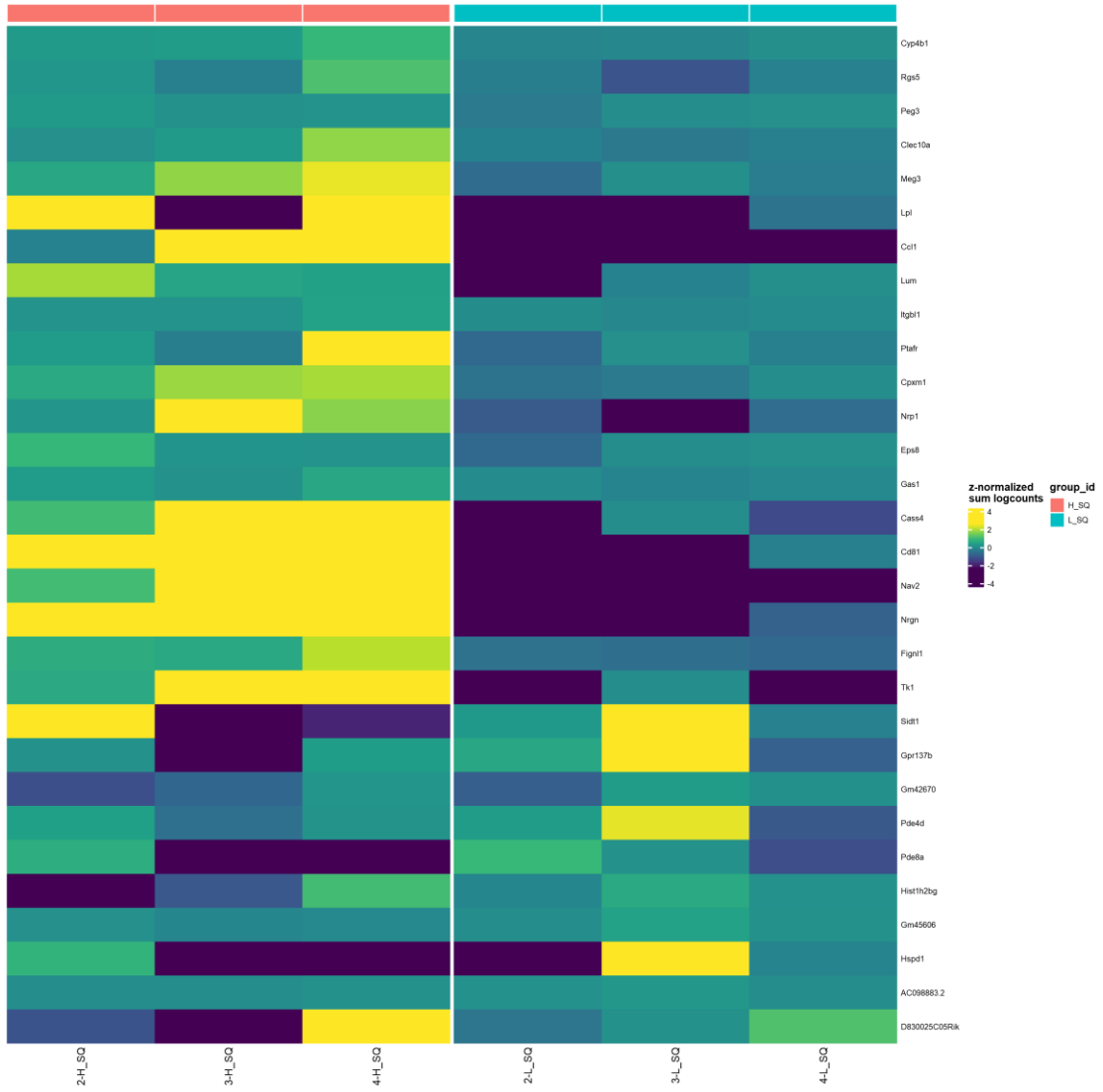
Pseudo-bulk(sum) for SQ, colored by condition

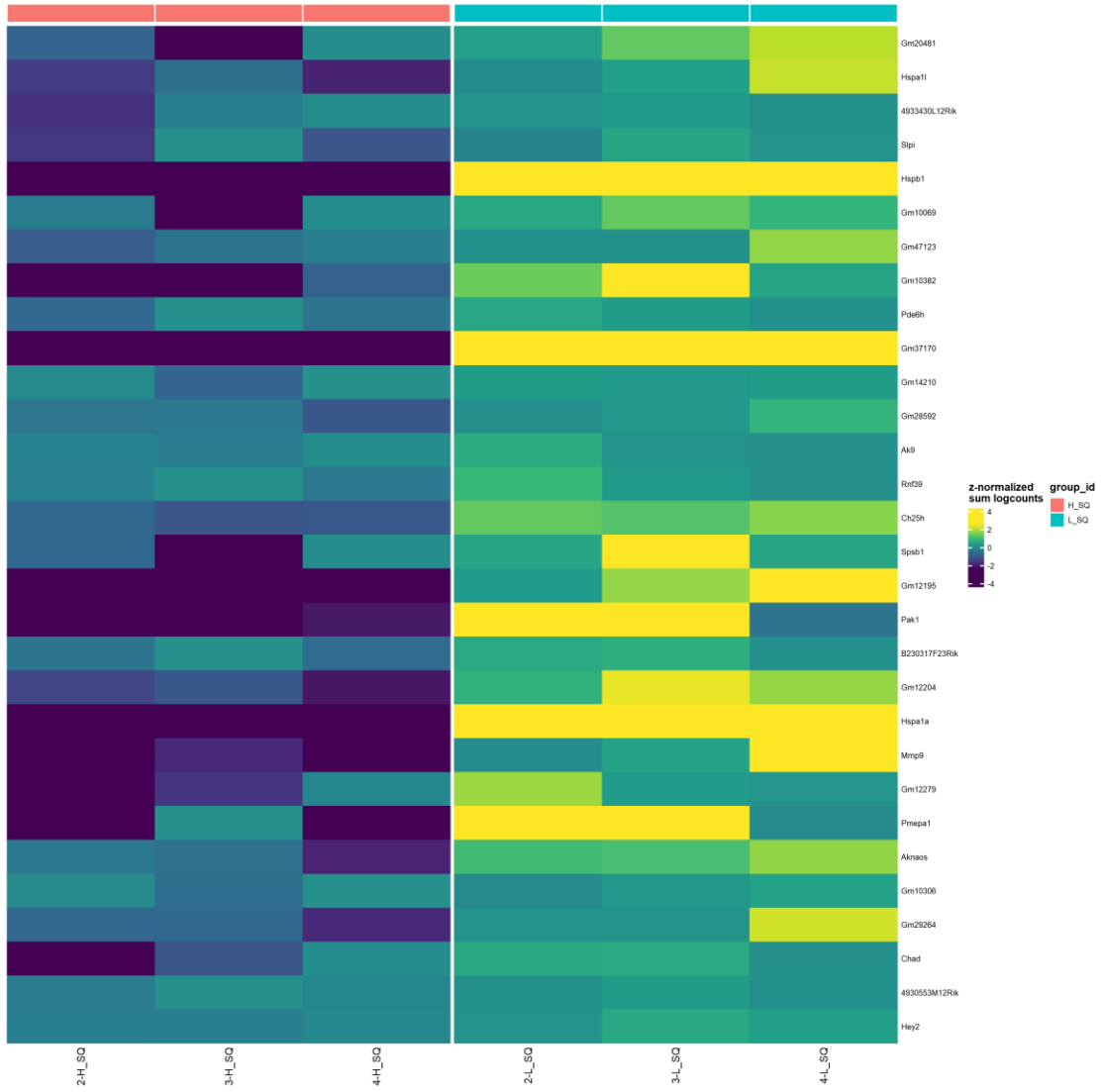


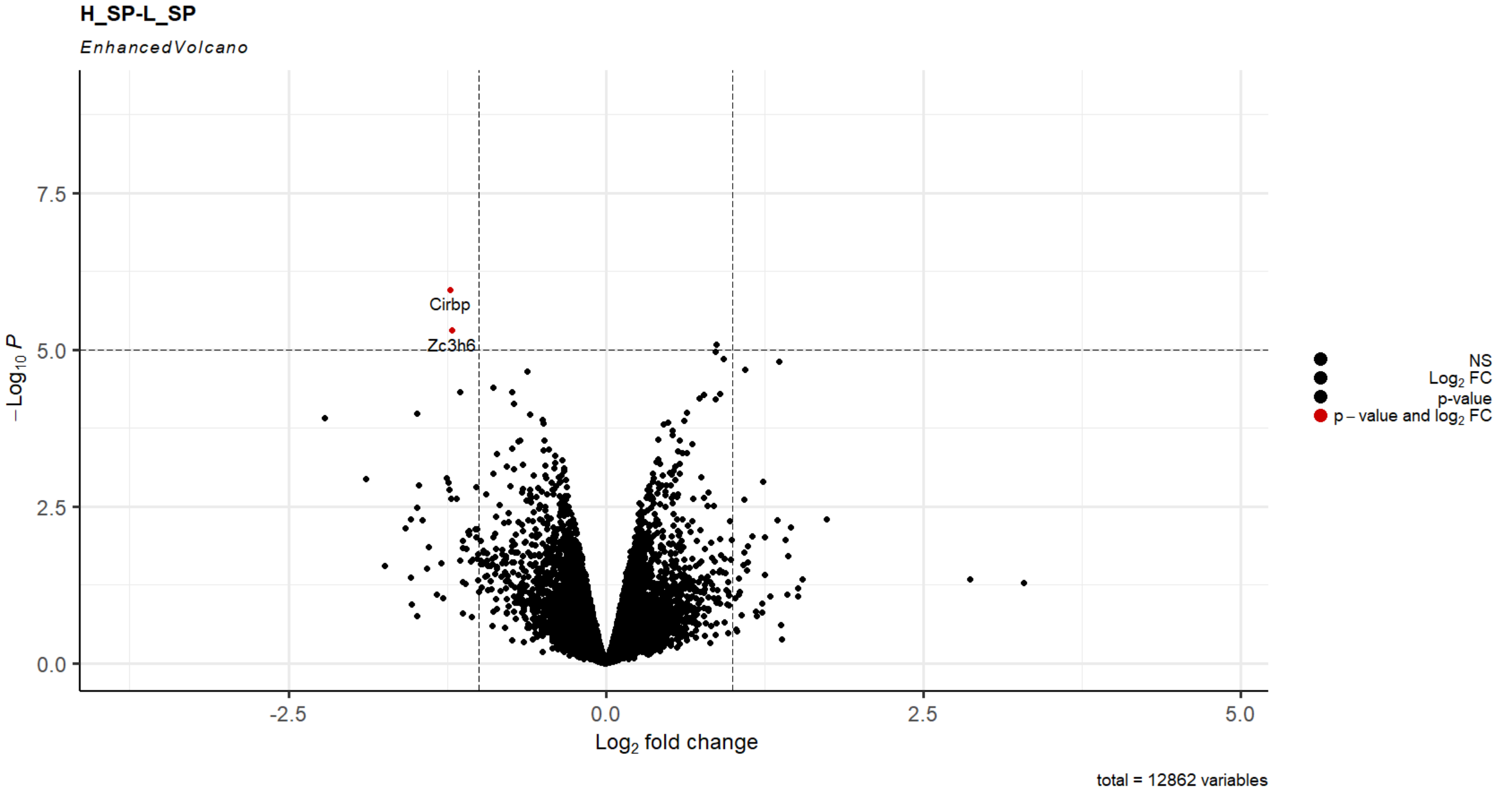
p_adj.loc (FDR < 0.05)

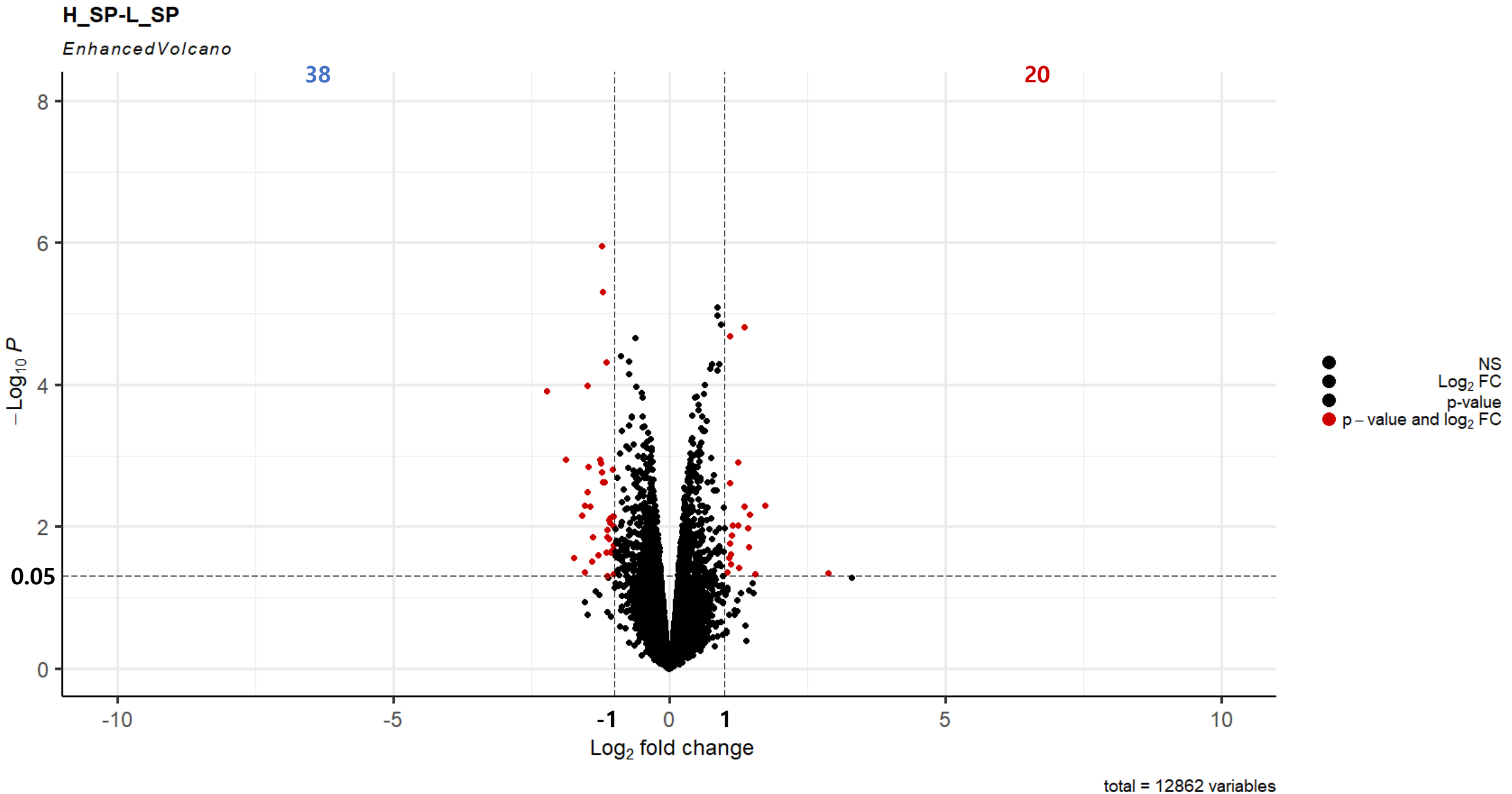
Pseudo-bulk(sum) for SQ, colored by condition

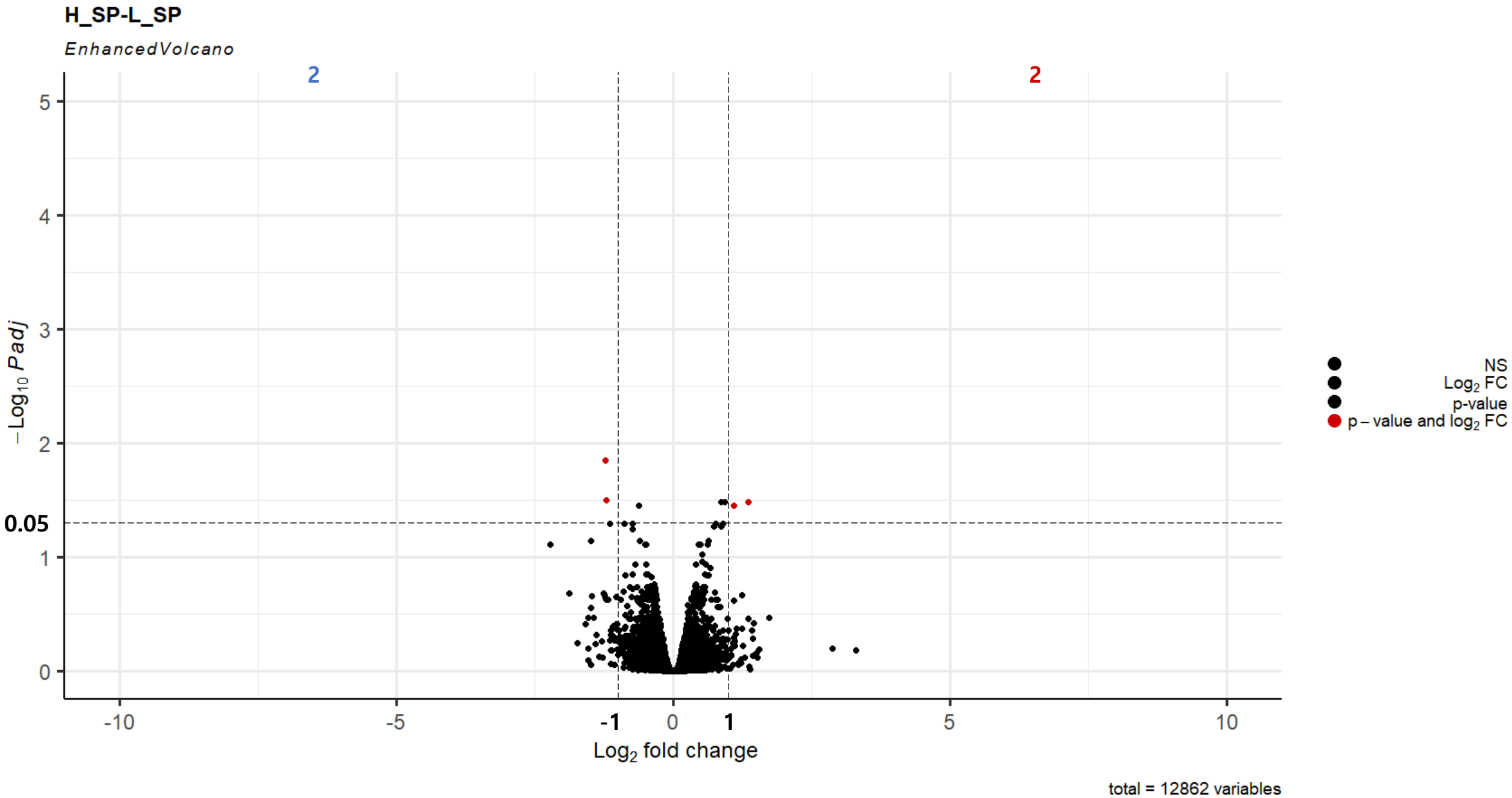
Sorted by logFC (FDR < 0.05)
top30



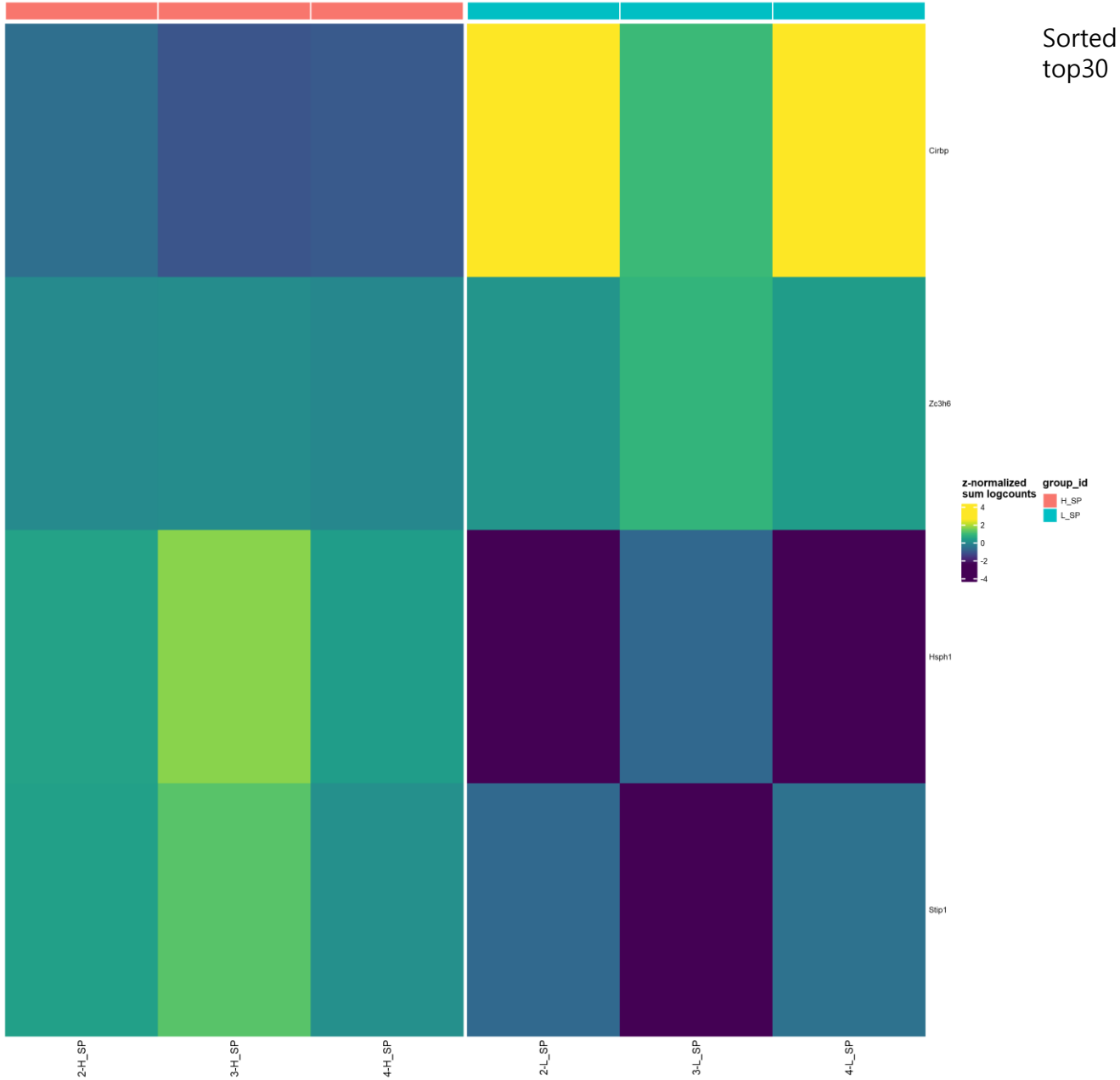



H_SP-L_SP
EnhancedVolcano
Cirbp
Zc3h6
7.5
5.0
2.5
0.0
−Log10 P
-2.5
0.0
2.5
5.0
Log2 fold change
NS
Log2 FC
p-value
p – value and log2 FC
total = 12862 variables

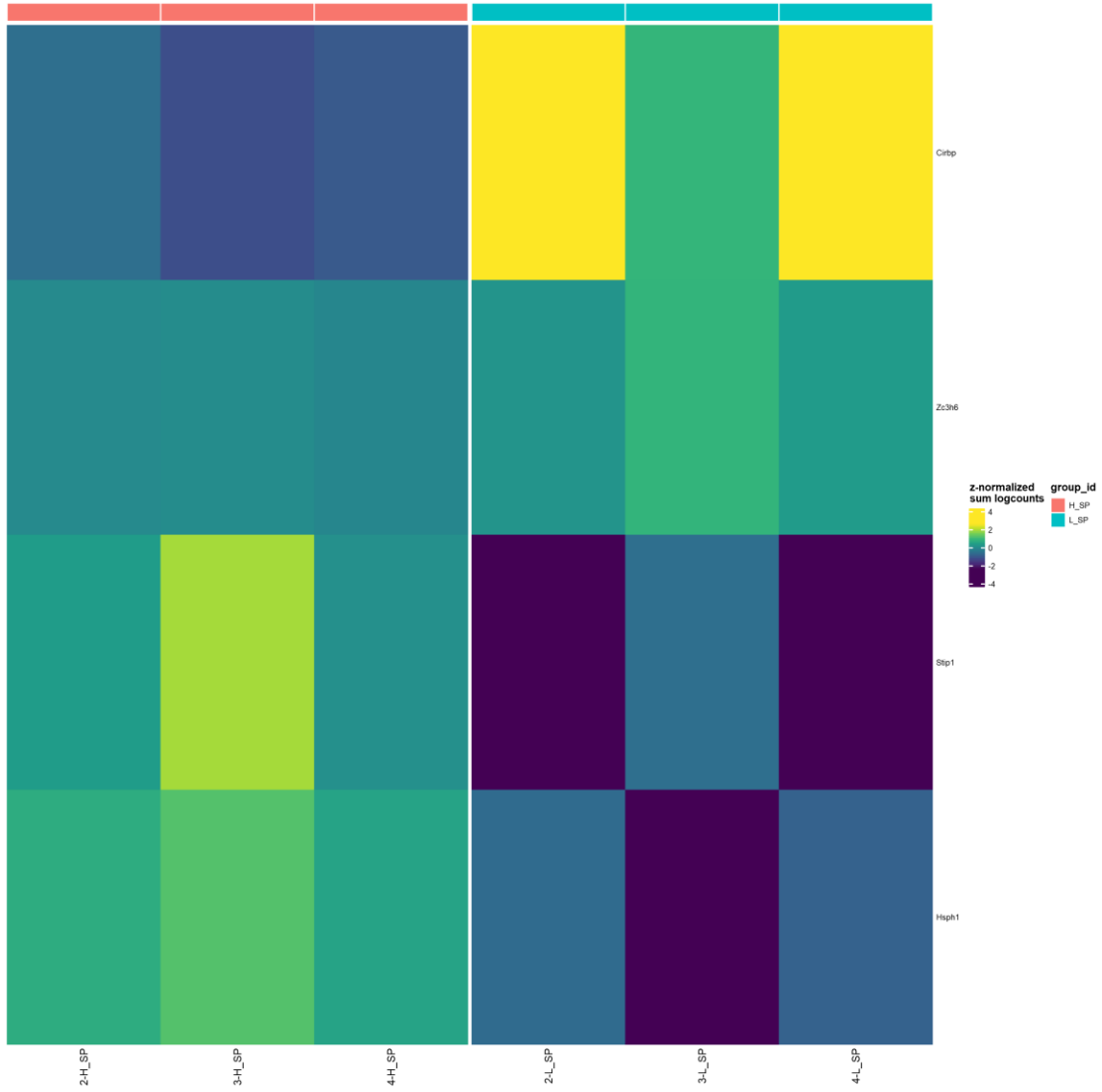
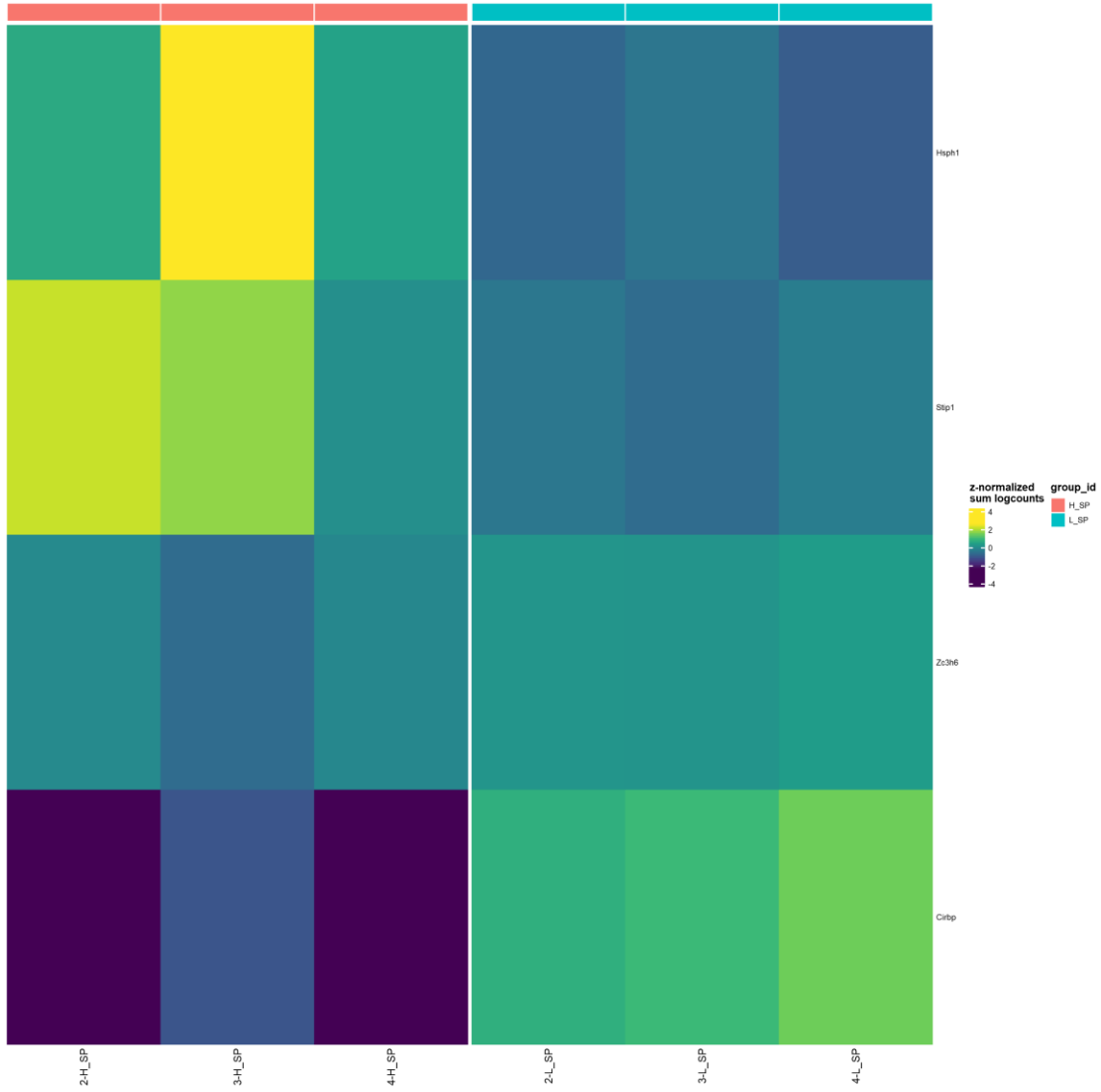

H_SP-L_SP
EnhancedVolcano
38
20
8
6
4
2
0.05
0
−Log10 P
-10
-5
-1
0
1
5
10
Log2 fold change
NS
Log2 FC
p-value
p – value and log2 FC
total = 12862 variables


H_SP-L_SP
EnhancedVolcano
2
2
5
4
3
2
1
0
0.05
−Log10 Padj
-10
-5
-1
0
1
5
10
Log2 fold change
NS
Log2 FC
p-value
p – value and log2 FC
total = 12862 variables

Pseudo-bulk(sum) for SP, colored by condition



Sorted by p_adj.loc (FDR < 0.05)
top30

Pseudo-bulk(sum) for SP, colored by condition

Sorted by logFC (FDR < 0.05)
top30