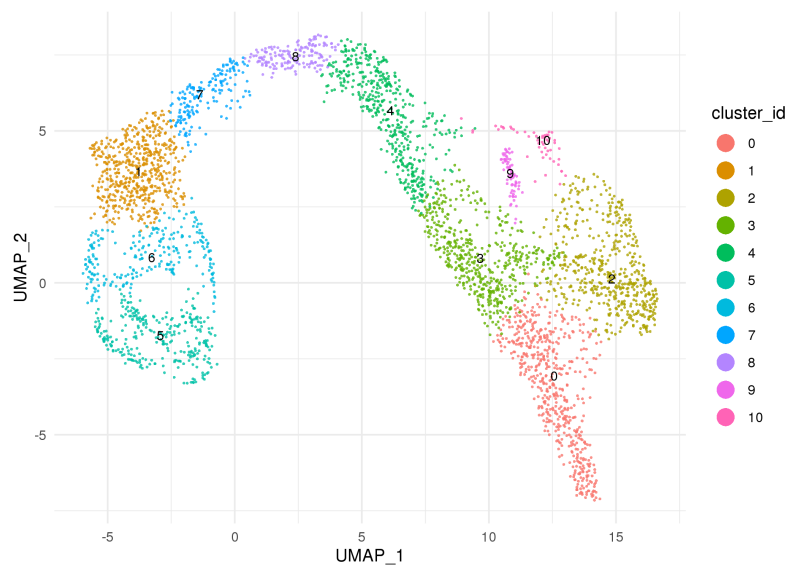


pipeline_seurat.py: summary report

Pancreatic endocrinogenesis (e15.5; scvelo example dataset)

Sansom group

November 5, 2020



Sample: e15

Run specs: no. components: 30, cluster resolution: 0.7, cluster algorithm: leiden, de test: wilcox

Code: <https://github.com/sansomlab/tenx>

Contents

1	Introduction	3
1.1	Optional tasks	3
2	Data quality control	4
2.1	Quality assessment and removal of low-quality cells	4
3	Removal of unwanted variation	6
3.1	Removal of unwanted variation (data normalisation)	6
3.2	Summary statistics	6
4	Dimension reduction	7
4.1	Scree (elbow) plot	7
4.2	Component heatmaps	8
5	Exploration of cell cycle effects	9
5.1	Visualisation of cell cycle effects	9
6	Highly variable gene identification	9
6.1	Identification of variable genes	9
6.2	JackStraw analysis	11
7	Visualisation of clusters and factors of interest	12
7.1	umap.mindist_0 plot colored by cluster_id	12
7.2	umap.mindist_0.1 plot colored by cluster_id	13
7.3	umap.mindist_0.3 plot colored by cluster_id	14
7.4	umap.mindist_0.5 plot colored by cluster_id	15
7.5	umap.mindist_0.7 plot colored by cluster_id	16
7.6	umap plot colored by nCount_RNA	17
7.7	umap plot colored by percent.mito	18
8	singleR	19
8.1	ImmGenData	19
8.2	MouseRNAseqData	21
9	Plots of summary statistics	23
9.1	Cells by cluster	23
9.2	Percent cells by cluster	23
9.3	Number of genes per cell per cluster	24
9.4	Number of umi per cell per cluster	25
10	Cluster dissimilarity	26
10.1	Dissimilarity by gene expression	26
11	Cluster resolution analysis	28
12	Diffusion map	29
12.0.1	Diffusion map plots (first 3 dimensions, different rotations) colored by cluster	29
13	Phate maps	30
14	Partion-based graph abstraction (PAGA)	32
14.1	UMAP	33
14.2	Force directed graphs	34

15 RNA velocity	36
15.1 reduced dimension maps	36
15.2 PAGA initialised FA2 graphs	39
15.3 PHATE graphs	42
16 Identification of cluster marker genes	45
17 Top cluster marker genes	46
18 Marker genes by cluster	47
18.0.1 Summary of numbers of DE genes per-cluster	47
18.1 Cluster 0: summary plots	48
18.2 Cluster 0 violin plots: positive marker genes	49
18.3 Cluster 0 violin plots: negative marker genes	50
18.4 Cluster 1: summary plots	51
18.5 Cluster 1 violin plots: positive marker genes	52
18.6 Cluster 1 violin plots: negative marker genes	53
18.7 Cluster 2: summary plots	54
18.8 Cluster 2 violin plots: positive marker genes	55
18.9 Cluster 2 violin plots: negative marker genes	56
18.10 Cluster 3: summary plots	57
18.11 Cluster 3 violin plots: positive marker genes	58
18.12 Cluster 3 violin plots: negative marker genes	59
18.13 Cluster 4: summary plots	60
18.14 Cluster 4 violin plots: positive marker genes	61
18.15 Cluster 4 violin plots: negative marker genes	62
18.16 Cluster 5: summary plots	63
18.17 Cluster 5 violin plots: positive marker genes	64
18.18 Cluster 5 violin plots: negative marker genes	65
18.19 Cluster 6: summary plots	66
18.20 Cluster 6 violin plots: positive marker genes	67
18.21 Cluster 6 violin plots: negative marker genes	68
18.22 Cluster 7: summary plots	69
18.23 Cluster 7 violin plots: positive marker genes	70
18.24 Cluster 7 violin plots: negative marker genes	71
18.25 Cluster 8: summary plots	72
18.26 Cluster 8 violin plots: positive marker genes	73
18.27 Cluster 8 violin plots: negative marker genes	74
18.28 Cluster 9: summary plots	75
18.29 Cluster 9 violin plots: positive marker genes	76
18.30 Cluster 9 violin plots: negative marker genes	77
18.31 Cluster 10: summary plots	78
18.32 Cluster 10 violin plots: positive marker genes	79
18.33 Cluster 10 violin plots: negative marker genes	80
19 Geneset analysis	81
19.1 GO.BP	81
19.2 GO.MF	93
19.3 GO.CC	94
19.4 KEGG	95
19.5 msigdb_biocarta	96
19.6 msigdb_reactome	97
19.7 msigdb_canonical_pathways	98
19.8 msigdb_tf_motifs	99
19.9 msigdb_immunological_signatures	100
19.10 Top genesets by cluster	101

1 Introduction

The core of the data analysis was performed using [Seurat](#) and [scanpy](#):

- The construction of the nearest neighbor graph, clustering and UMAP computation were performed using scanpy (or scvelo for use of hnsplib).
- The differential expression analysis was performed using Seurat.
- The geneset analysis was performed using [gsfisher](#)
- Please see <https://github.com/sansomlab/tenx> for more details.

The key parameter choices used for this analysis were:

- The number of pca components: 30
- The number of nearest neighbors: 30
- The distance metric used for the nearest neighbor graph: euclidean
- The method used for construction of the nearest neighbor graph: hnsw
- The resolution of the clustering: 0.7
- The clustering algorithm: leiden
- The differential expression test: wilcox

1.1 Optional tasks

This table summarises the status of the optional tasks. Tasks set to “True” were run.

task	run
explore_hvg_and_cell_cycle	True
singleR	True
jackstraw	True
compare_clusters	True
characterise_markers	True
top_marker_heatmap	True
extra_cluster_marker_plots	True
diffusionmap	True
phate	True
paga	True
velocity	True
knownmarkers	False
marker_report	True
exprsreport	False
genesets	True
cellbrowser	False

2 Data quality control

2.1 Quality assessment and removal of low-quality cells

Figure 1: Basic QC plots

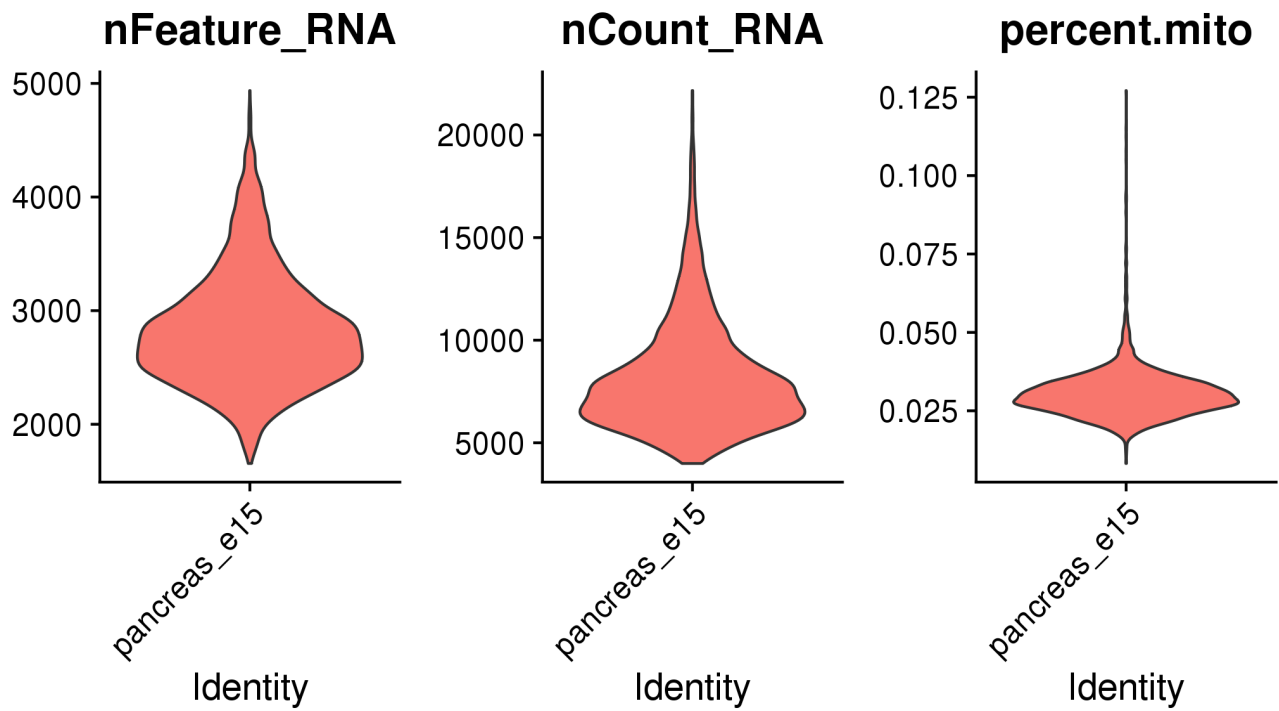


Figure 2: QC: violin plots

The dataset was filtered to remove (1) cells expressing fewer than 500 genes per cell and (2) cells with a fraction of mitochondrial reads greater than 0.05. Genes expressed in less than 3 cells were removed from the analysis.

3 Removal of unwanted variation

3.1 Removal of unwanted variation (data normalisation)

- The type of normalization applied was: log-normalization.
- A linear model was used to regress out the latent variables [percent.mito] before further analysis.
- The type of cell cycle regression applied was: none.

3.2 Summary statistics

	x
no_cells	3696.00
qc_min_gene_threshold	500.00
qc_min_percent_mito_threshold	0.00
qc_max_percent_mito_threshold	0.05
no_cells_after_qc	3628.00

Table 1: Run statistics

	input	after_qc_filters
pancreas_e15	3696	3628

Table 2: Numbers of cells

4 Dimension reduction

4.1 Scree (elbow) plot

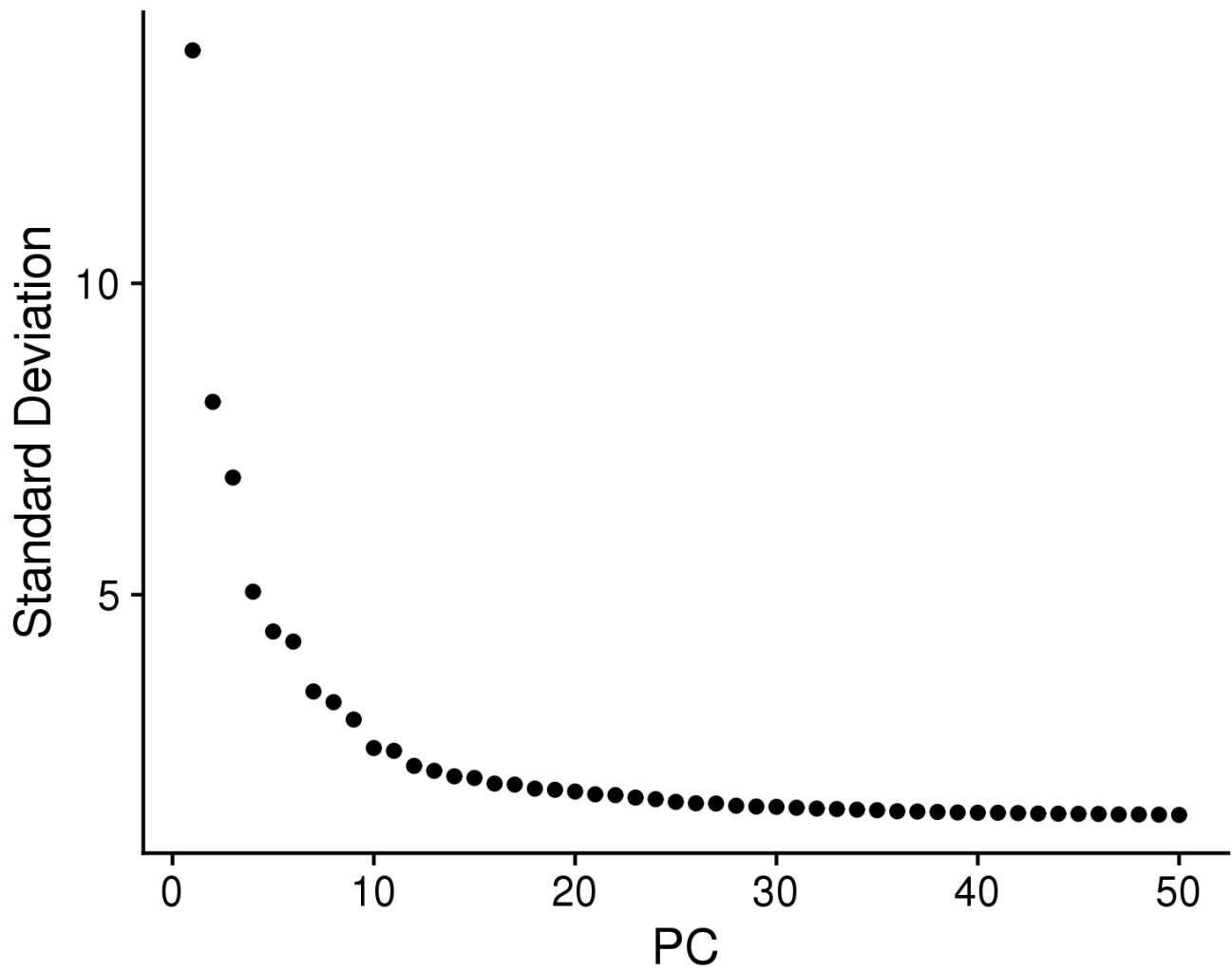


Figure 3: Scree plot showing proportion of variance explained by each PCA component

4.2 Component heatmaps

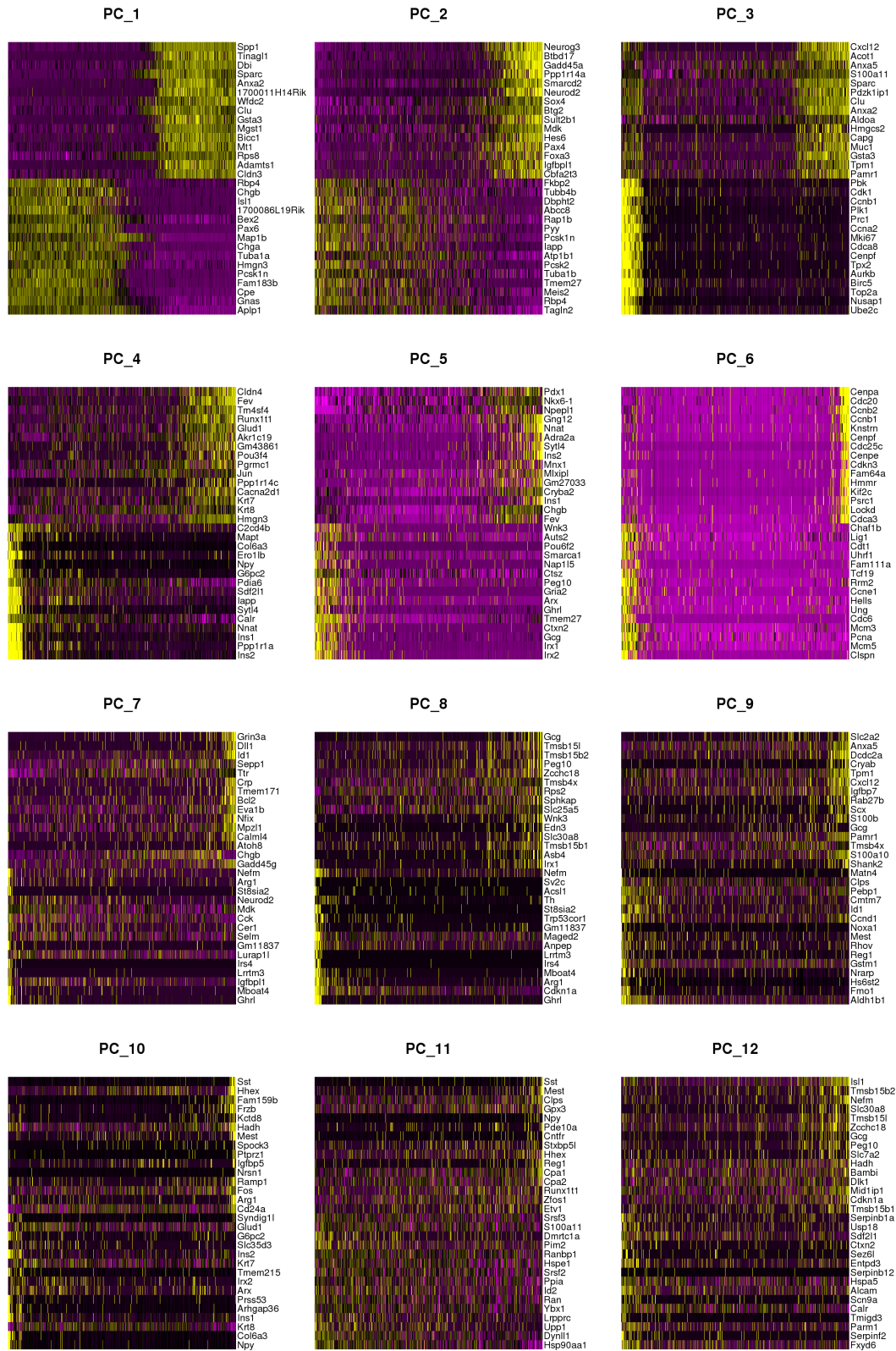


Figure 4: Heatmaps of the top genes for each PCA component

5 Exploration of cell cycle effects

5.1 Visualisation of cell cycle effects

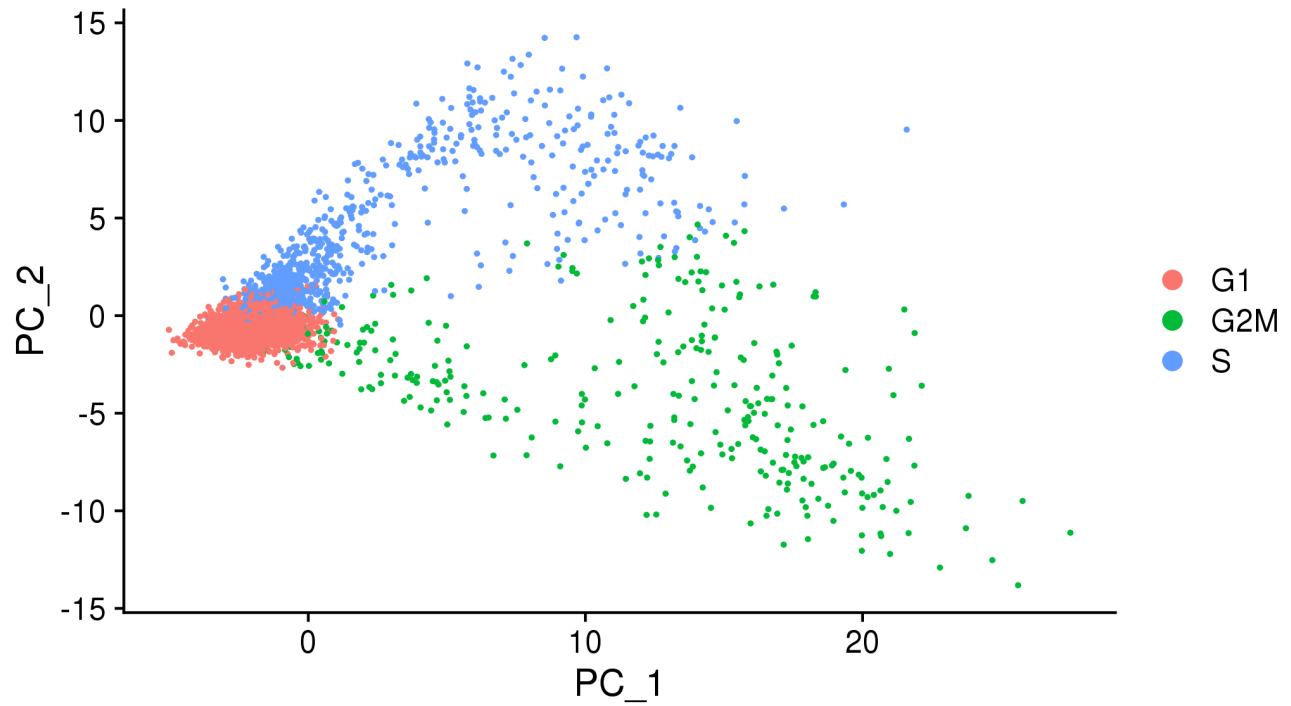


Figure 5: PCA analysis of cells based on expression of cell cycle genes (without regression of cell-cycle effects)

6 Highly variable gene identification

6.1 Identification of variable genes

The selection method used to identify highly variable genes was: vst

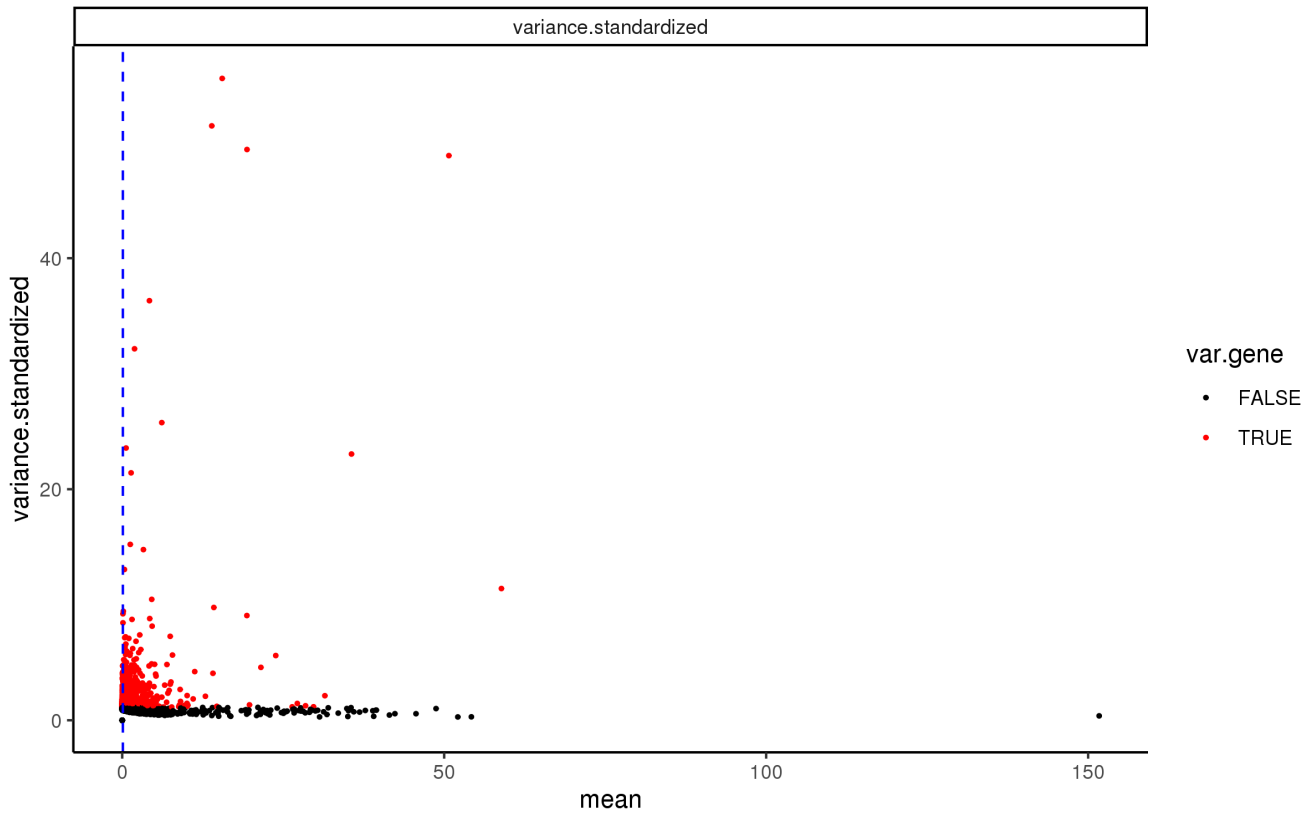


Figure 6: Mean dispersion relationship

6.2 JackStraw analysis

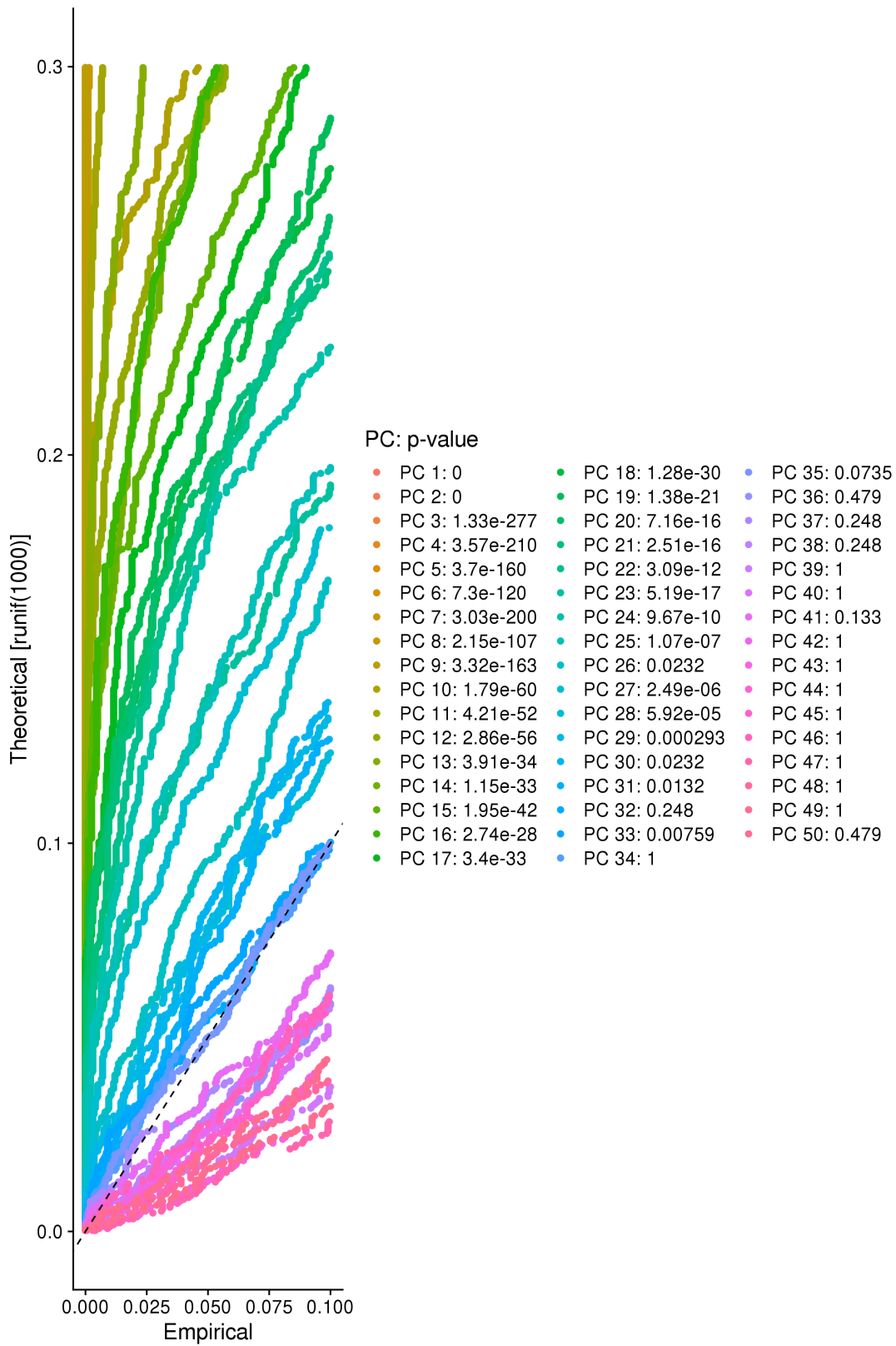


Figure 7: Seurat Jack Straw analysis of PCA components

7 Visualisation of clusters and factors of interest

7.1 umap.mindist_0 plot colored by cluster_id

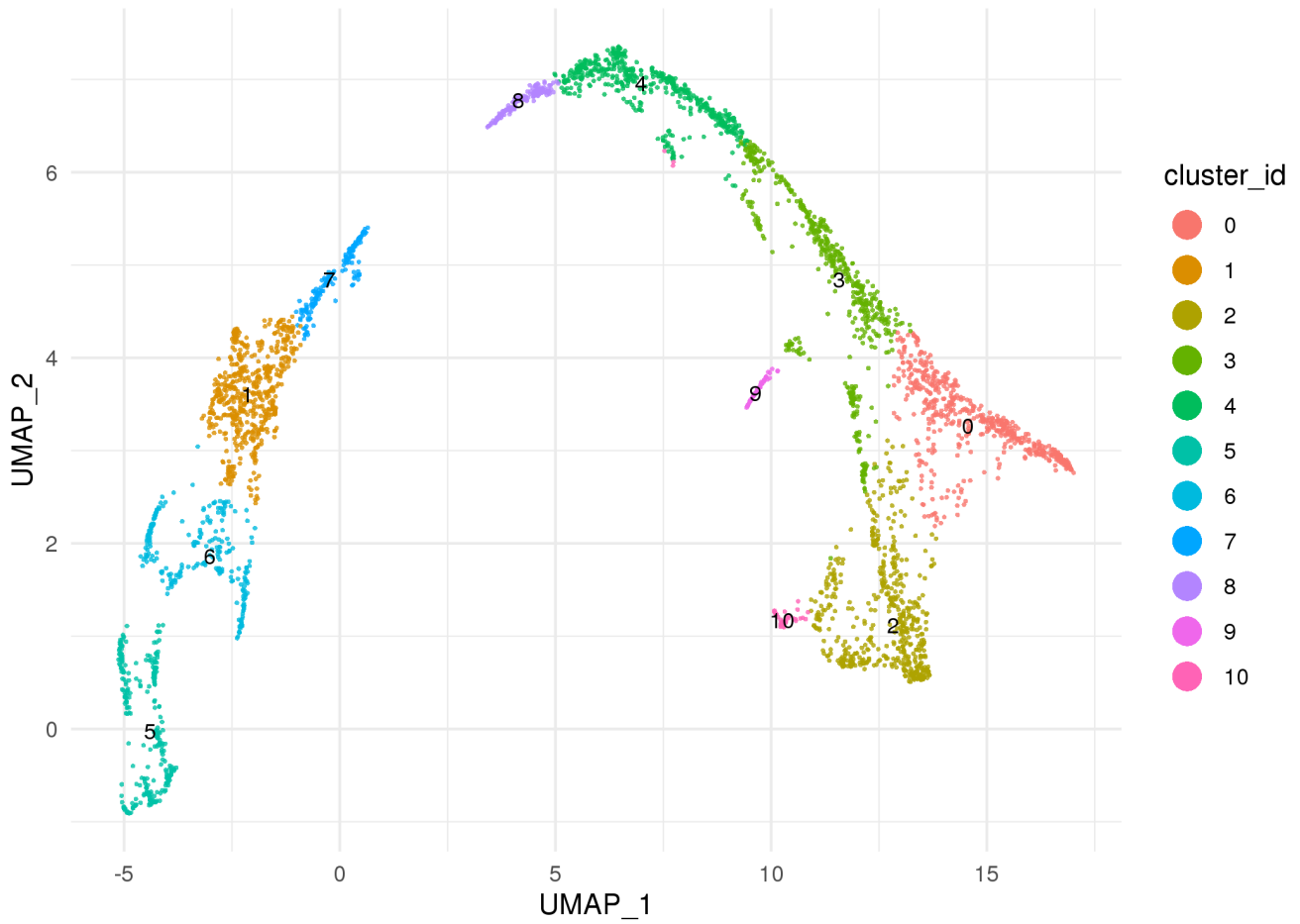


Figure 8: umap.mindist_0 plot colored by cluster_id

7.2 umap.mindist_0.1 plot colored by cluster_id

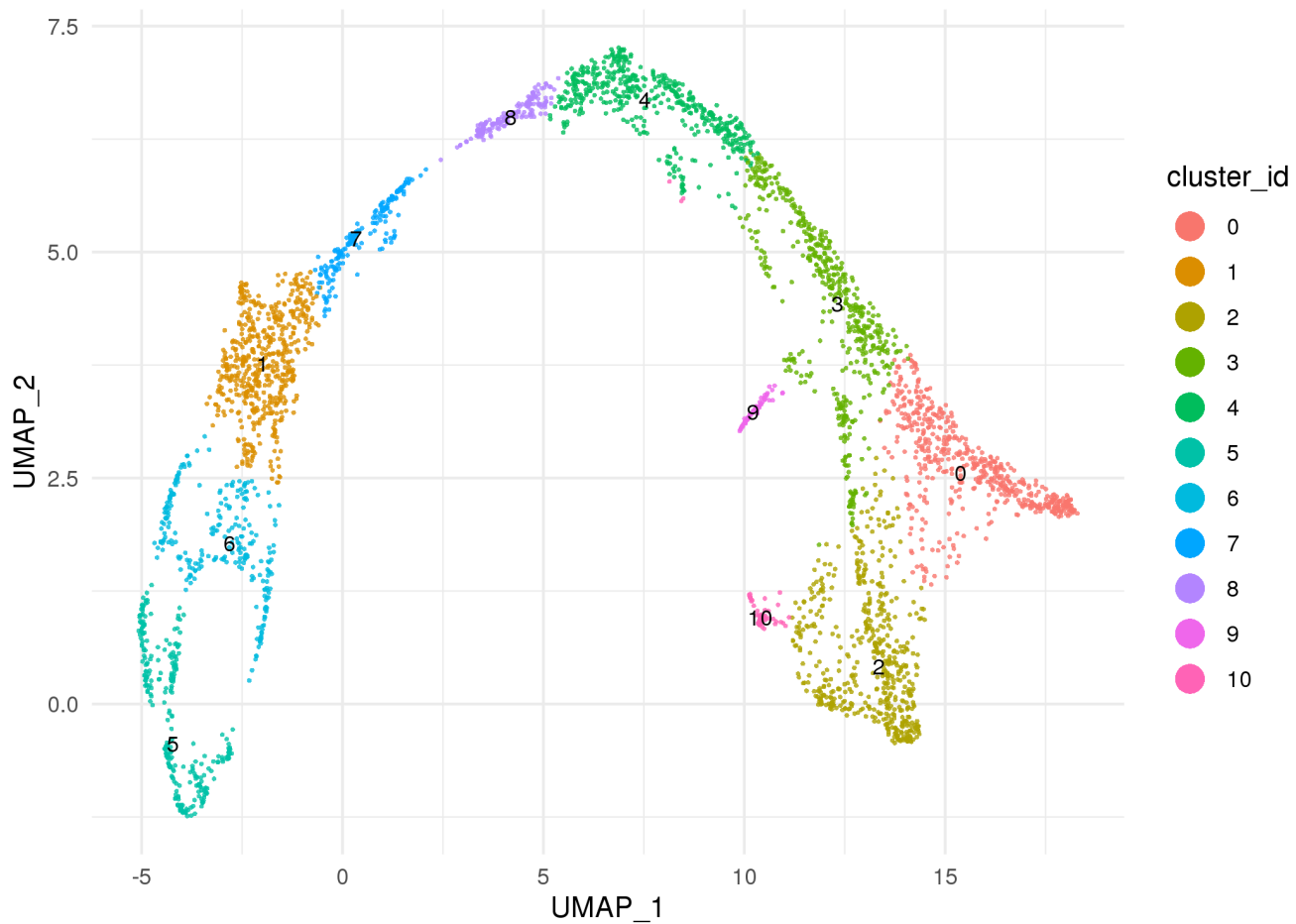


Figure 9: umap.mindist_0.1 plot colored by cluster_id

7.3 umap.mindist_0.3 plot colored by cluster_id

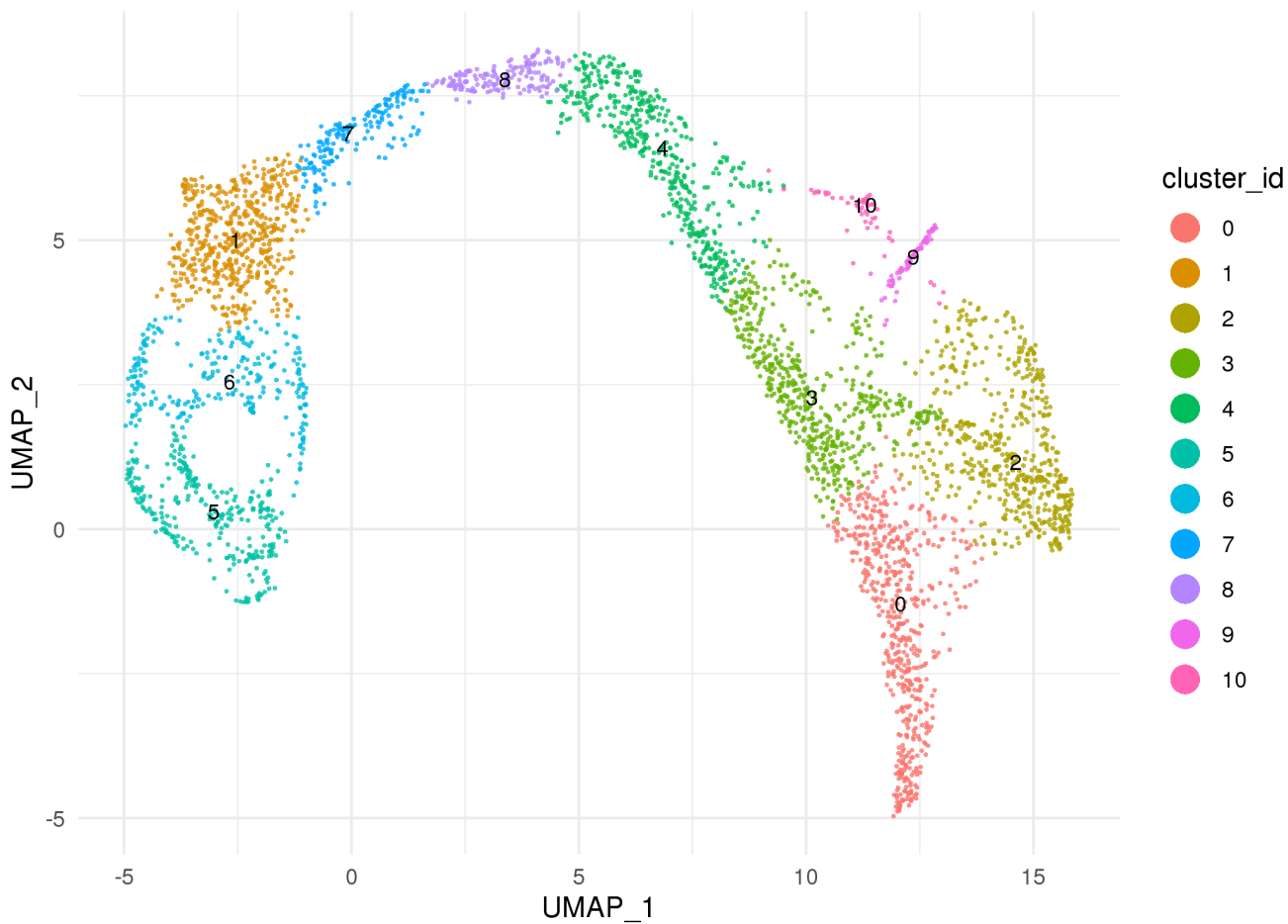


Figure 10: umap.mindist_0.3 plot colored by cluster_id

7.4 umap.mindist_0.5 plot colored by cluster_id

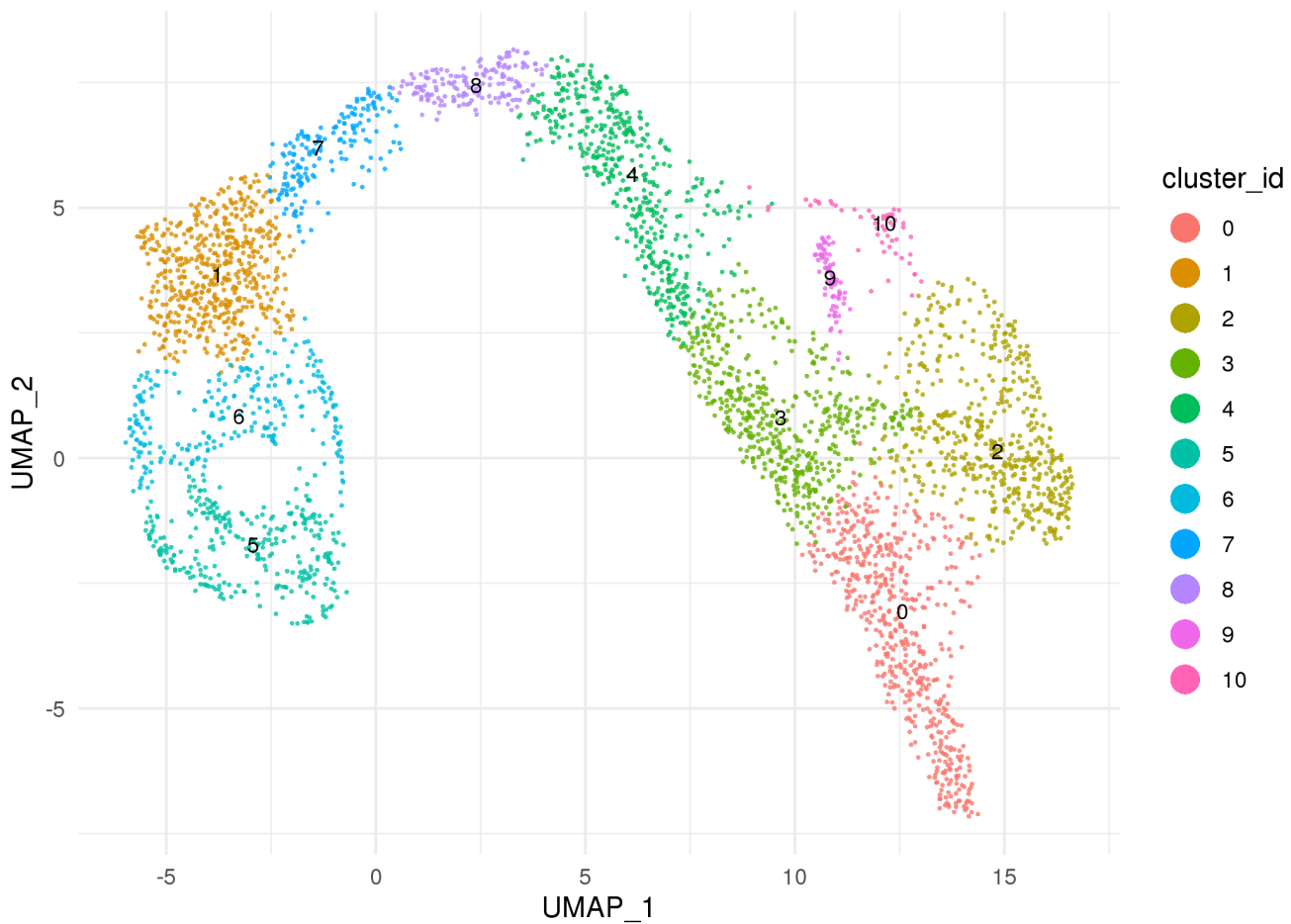


Figure 11: umap.mindist_0.5 plot colored by cluster_id

7.5 umap.mindist_0.7 plot colored by cluster_id

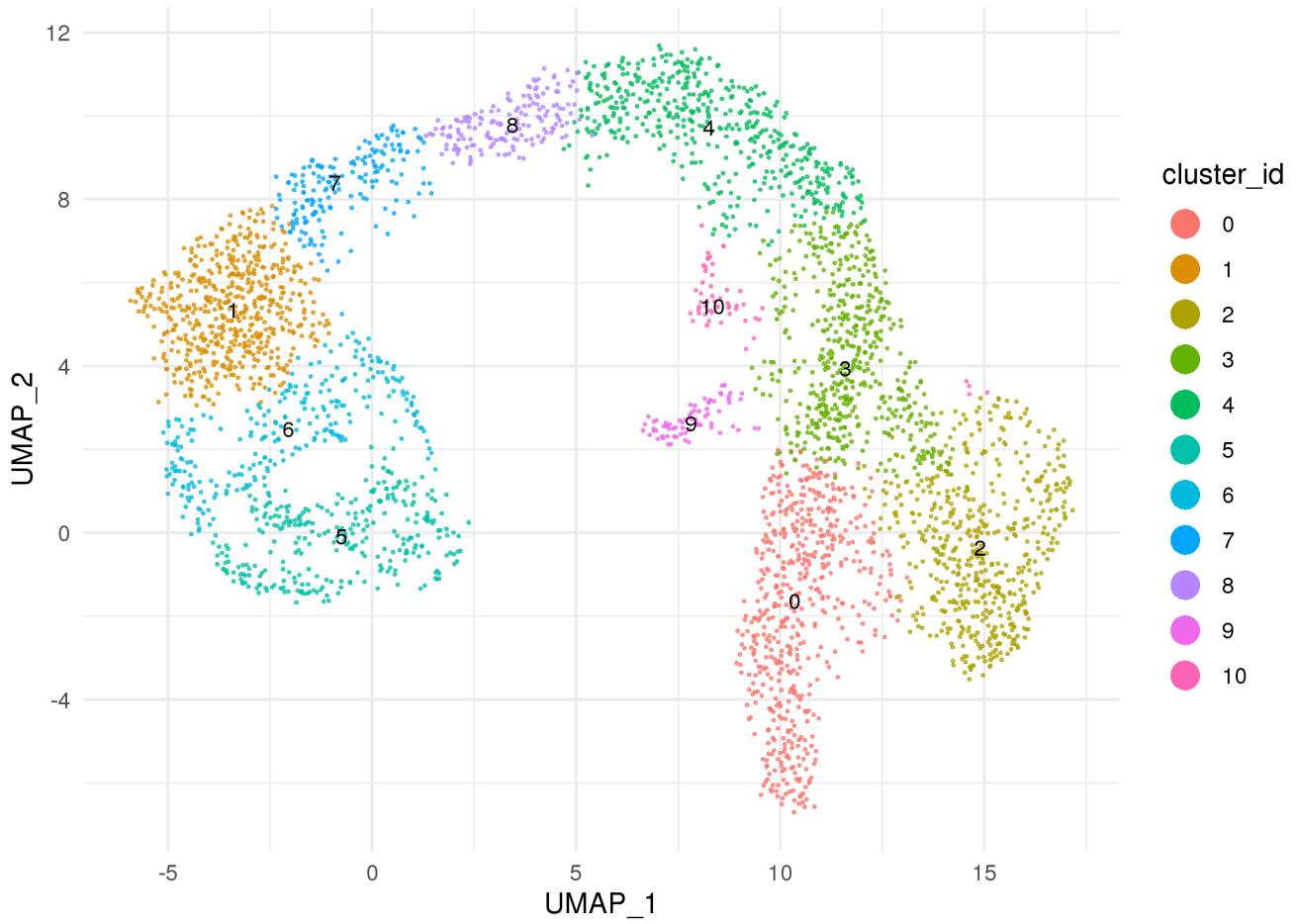


Figure 12: umap.mindist_0.7 plot colored by cluster_id

7.6 umap plot colored by nCount_RNA

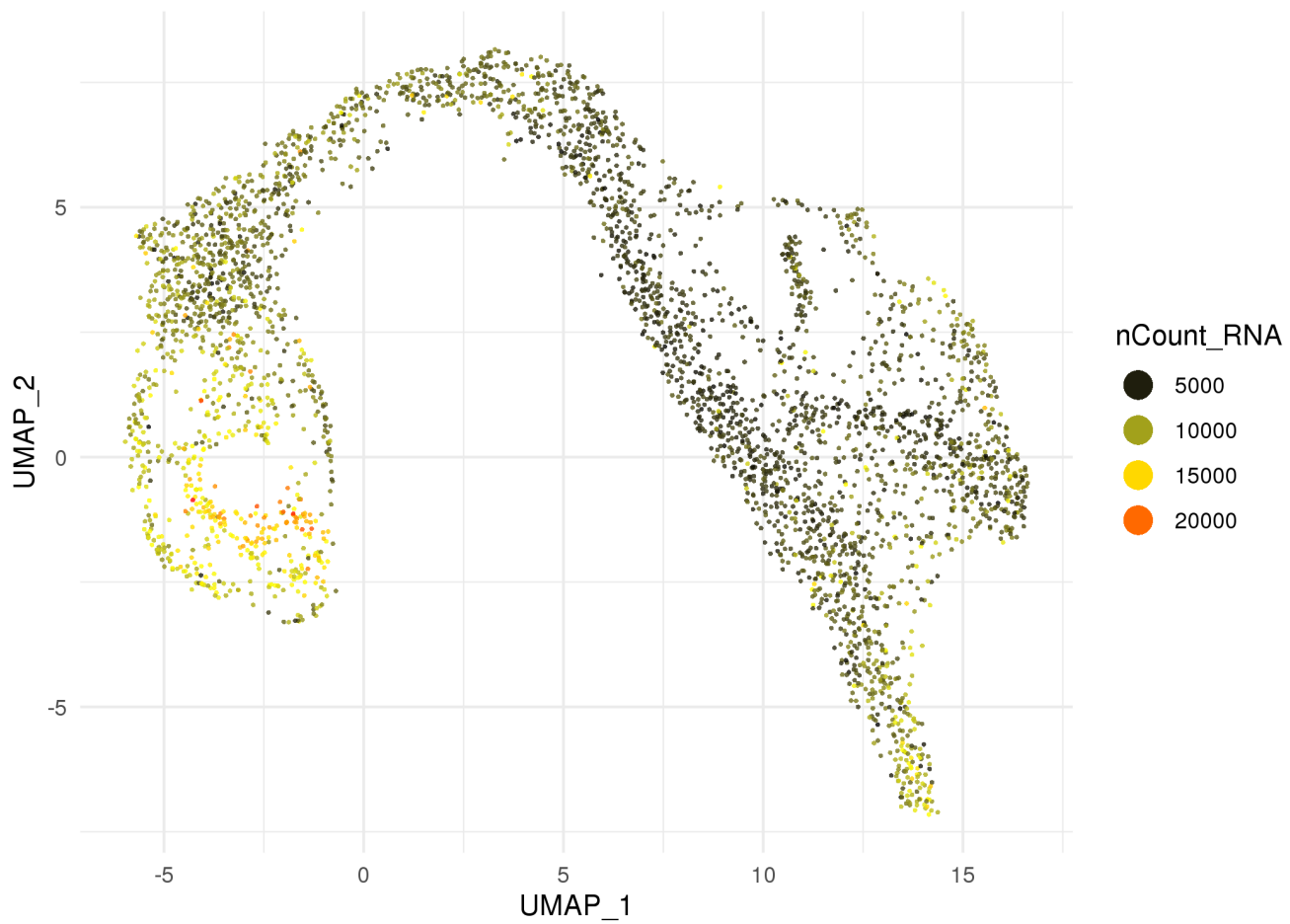


Figure 13: umap plot colored by nCount_RNA

7.7 umap plot colored by percent.mito

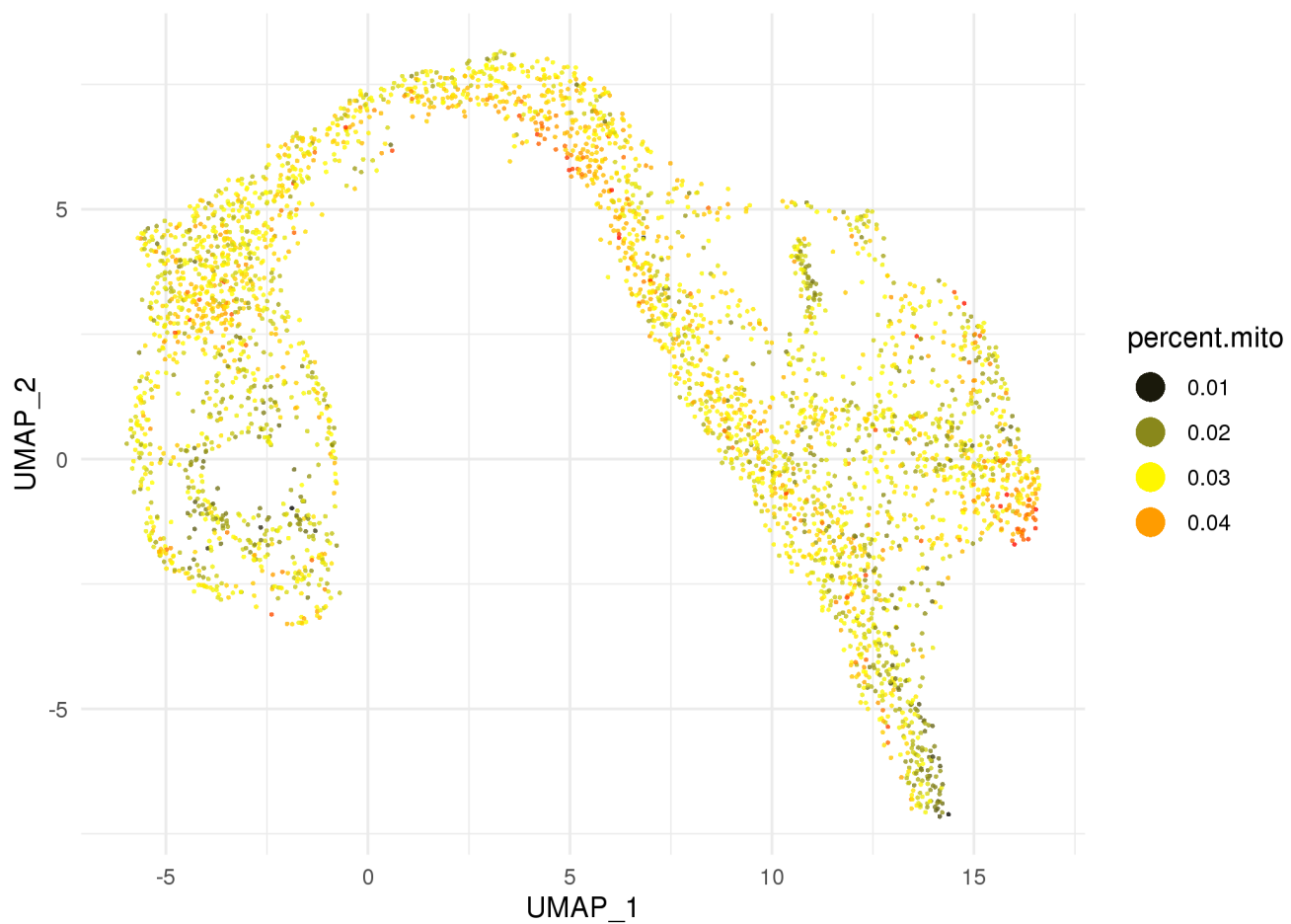


Figure 14: umap plot colored by percent.mito

8 singleR

See the [singleR paper](#) and [bioconductor package](#) for more information.

8.1 ImmGenData

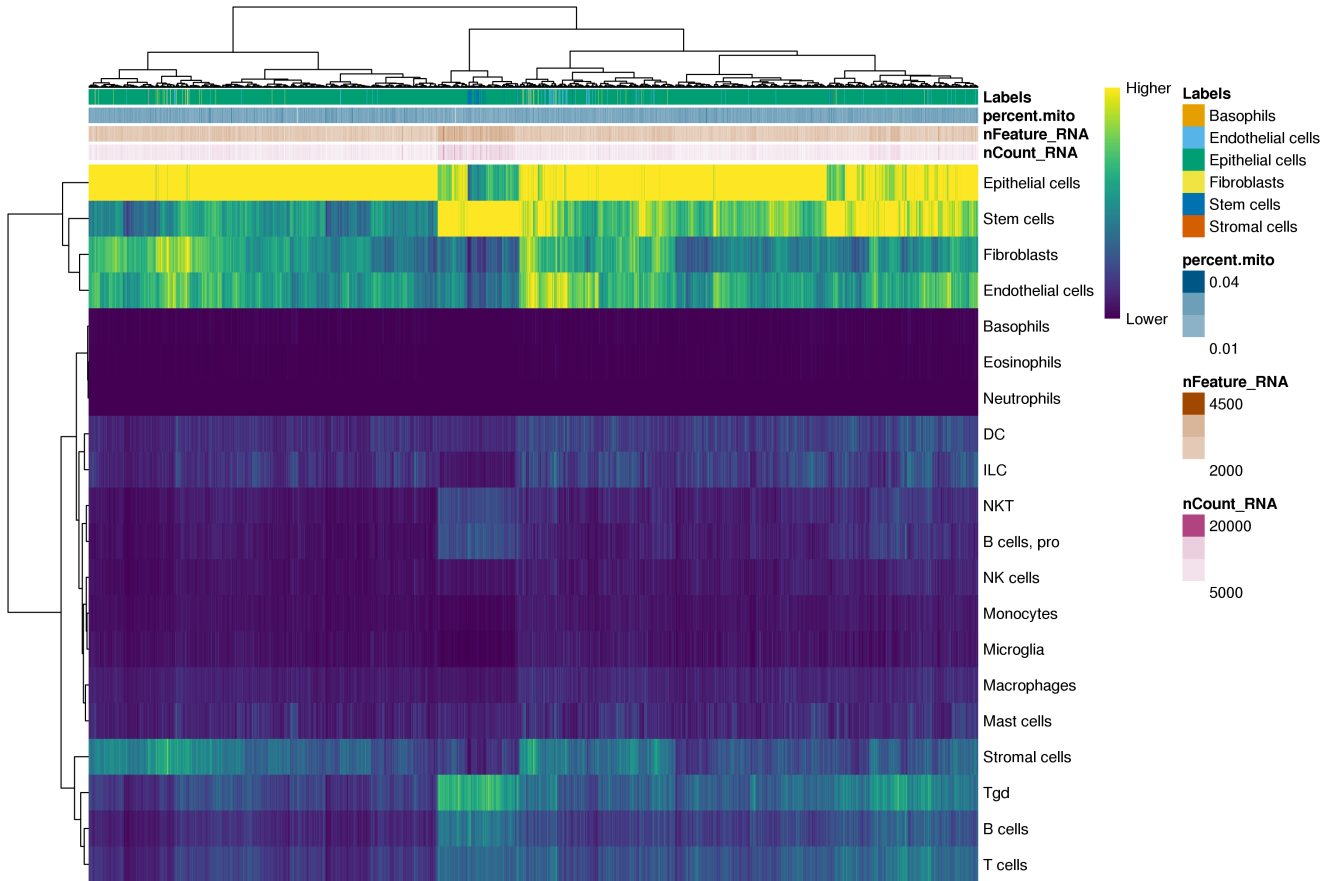


Figure 15: singleR predictions (ImmGenData)

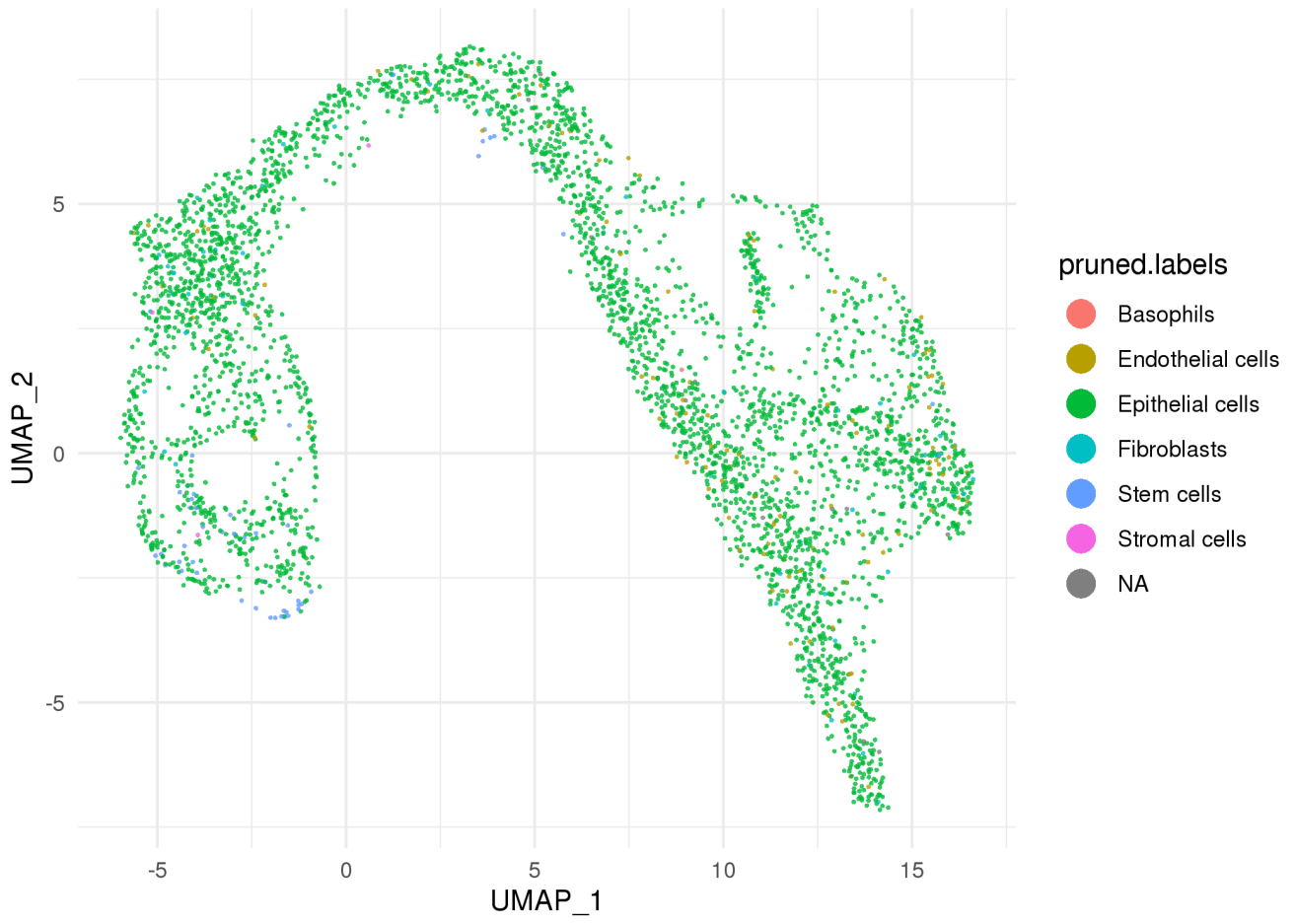


Figure 16: pruned singleR predictions (ImmGenData)

8.2 MouseRNAseqData

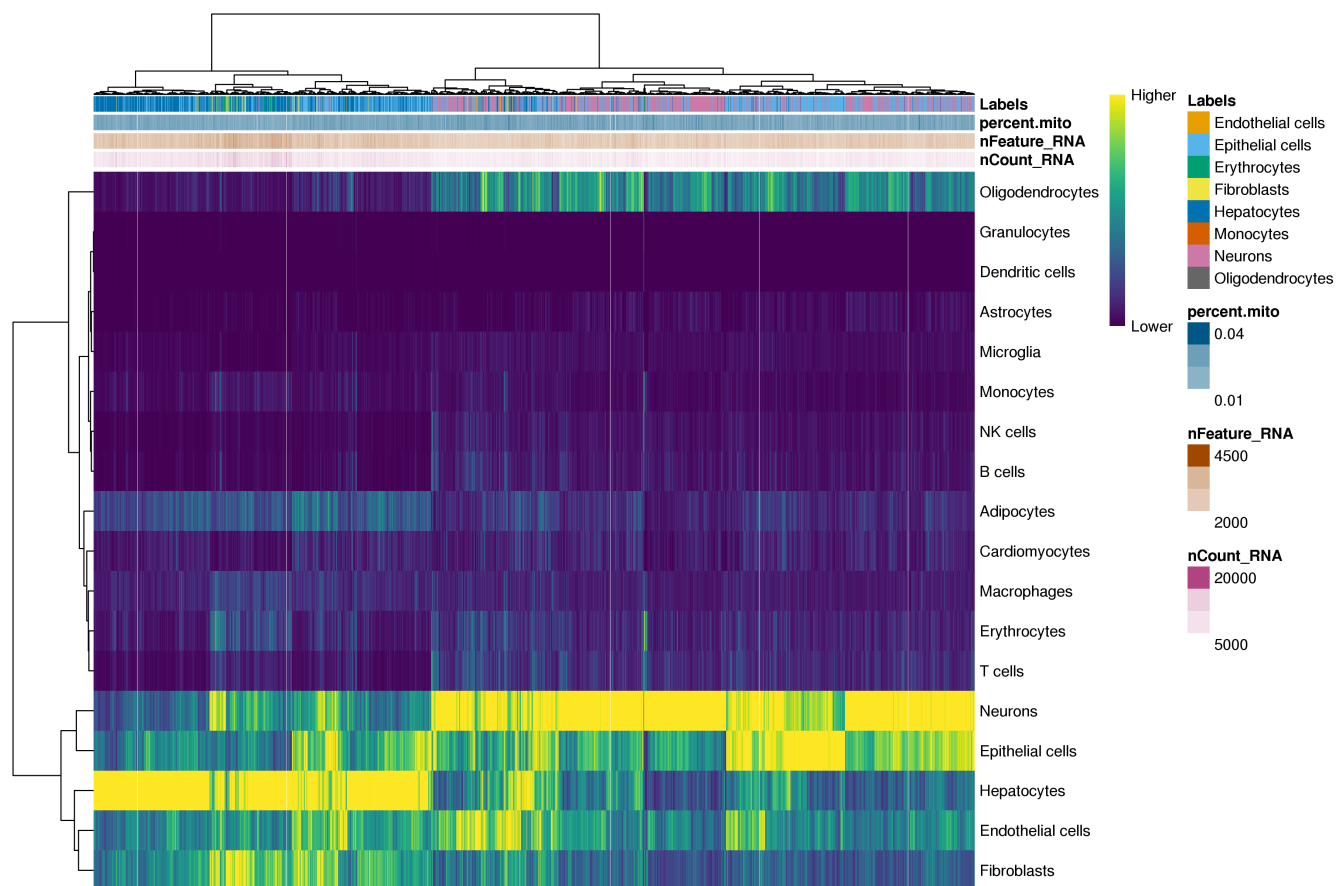


Figure 17: singleR predictions (MouseRNAseqData)

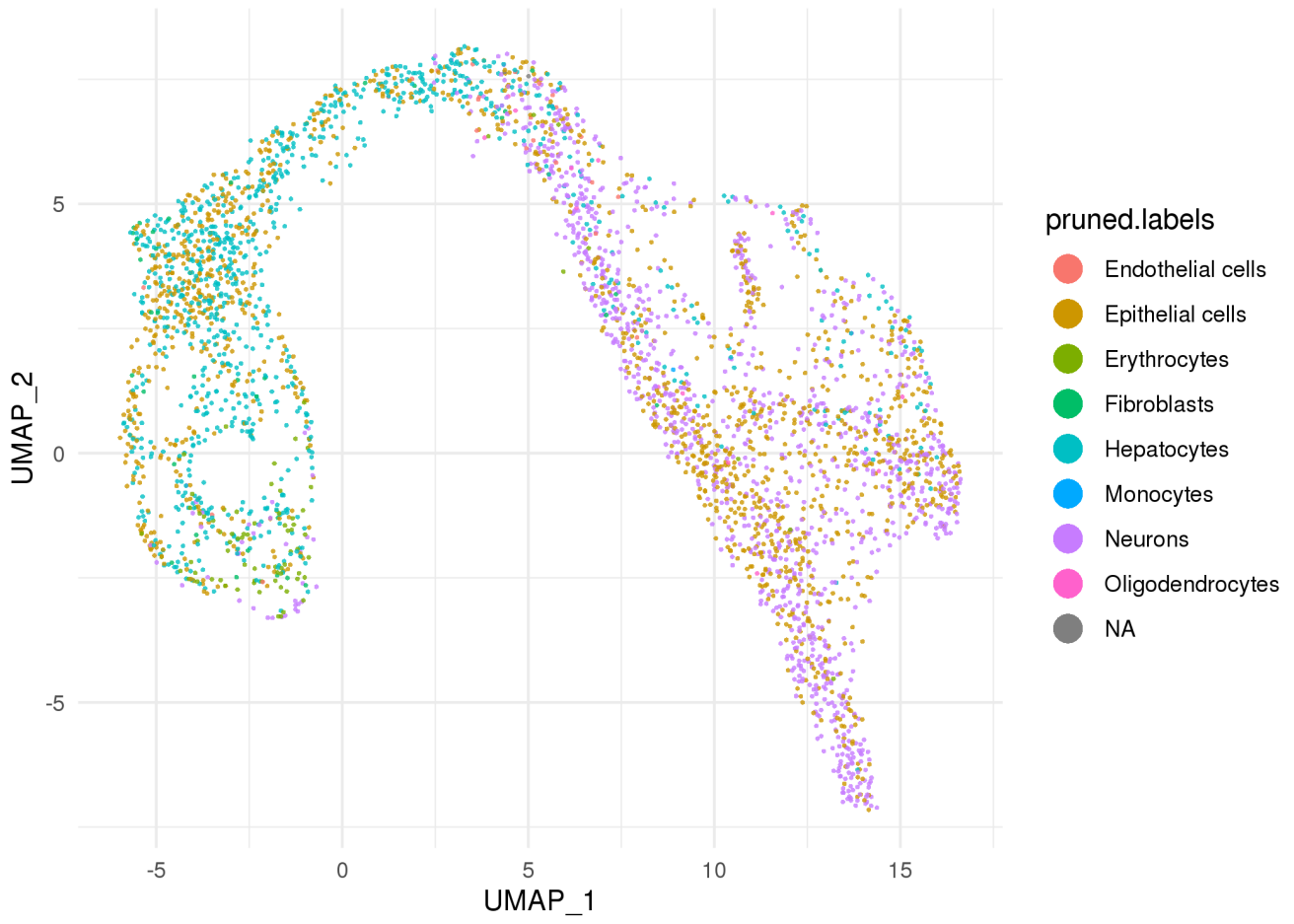


Figure 18: pruned singleR predictions (MouseRNAseqData)

9 Plots of summary statistics

Plots of summary statistics (e.g. cell number) by factor of interest (e.g. cluster)

9.1 Cells by cluster

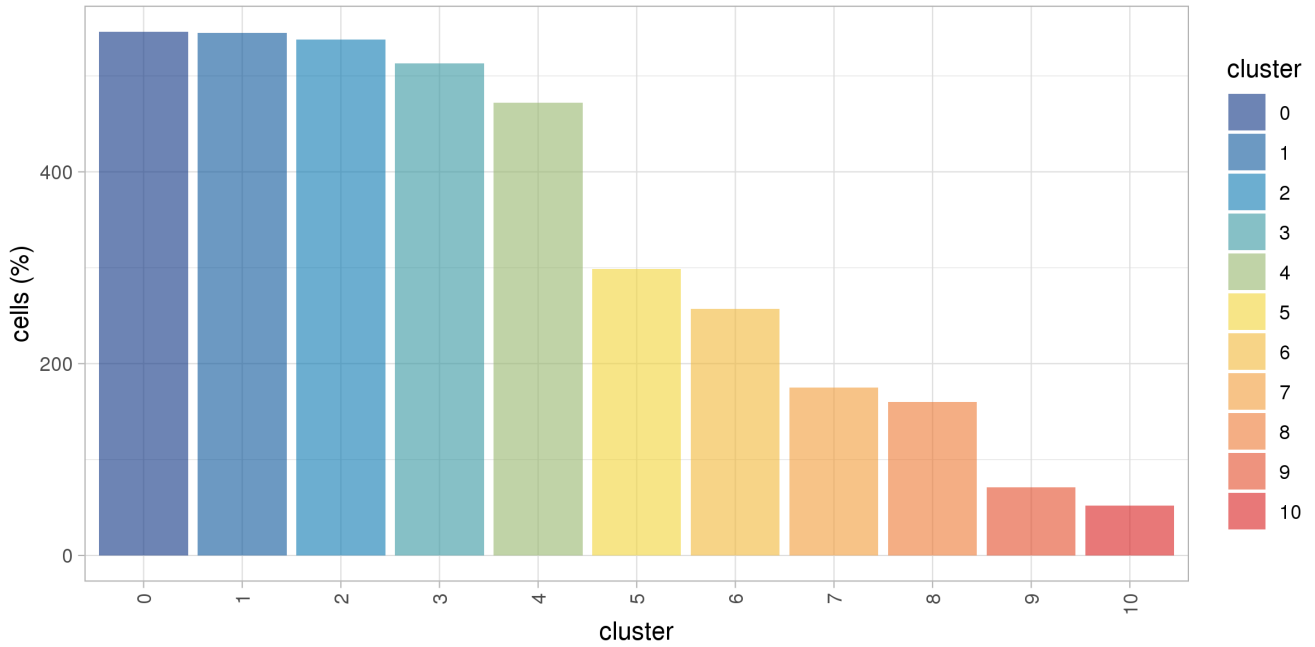


Figure 19: Cells by cluster

9.2 Percent cells by cluster

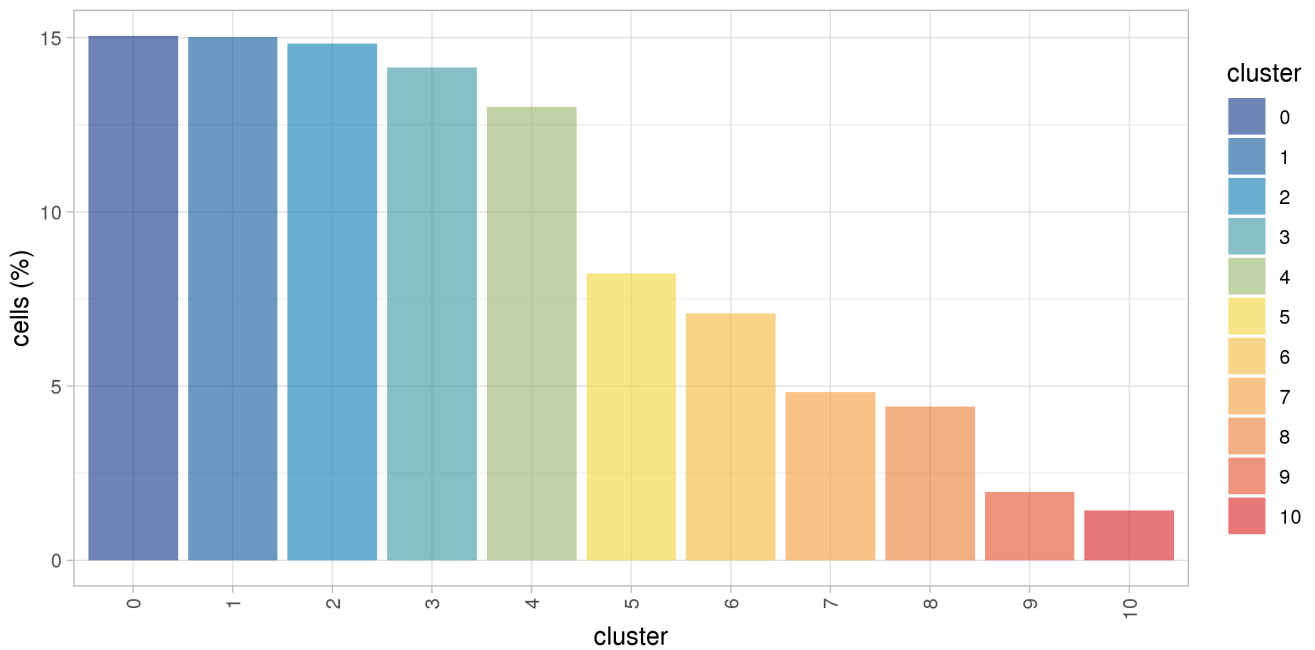


Figure 20: Percent cells by cluster

9.3 Number of genes per cell per cluster

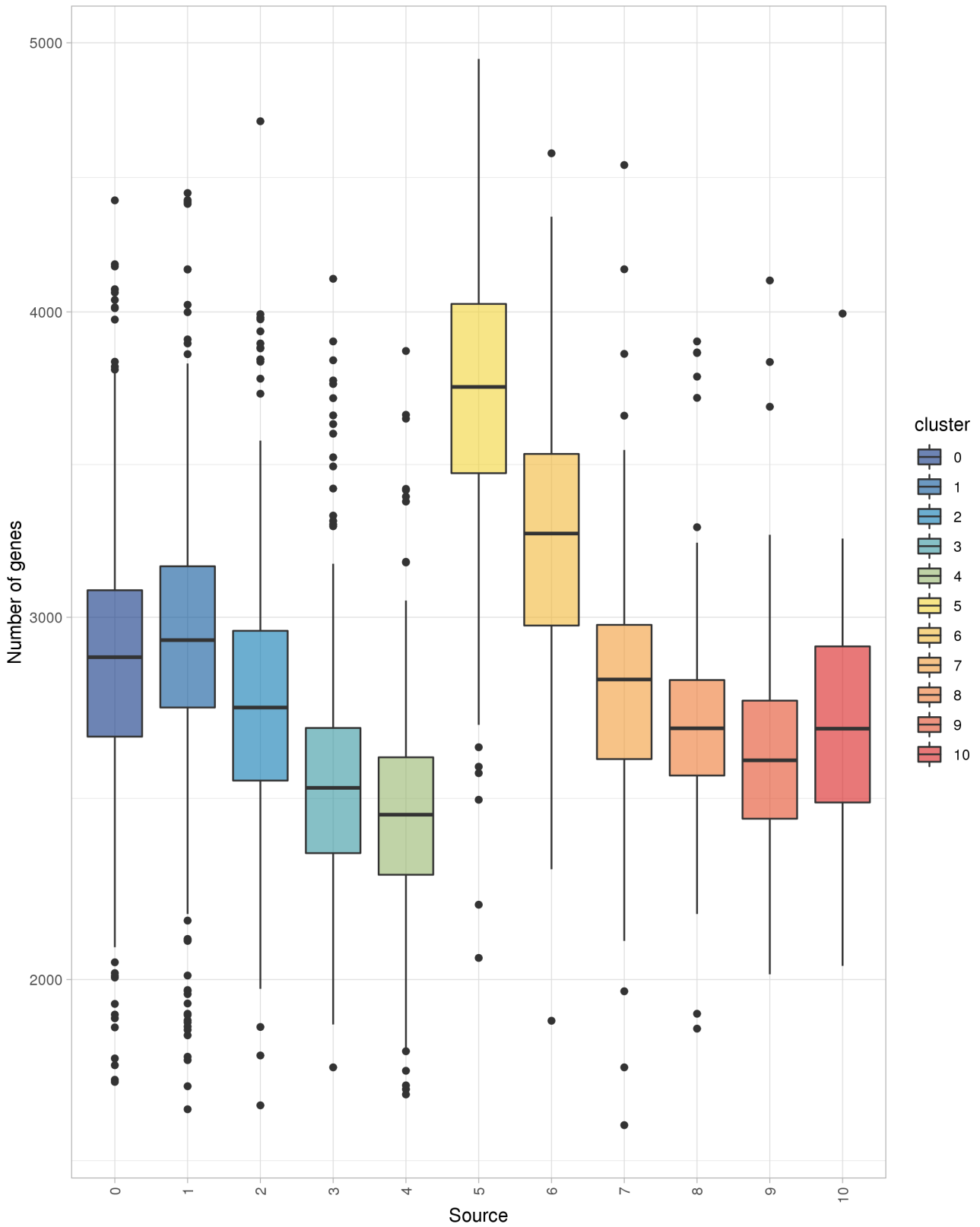


Figure 21: Number of genes per cell per cluster

9.4 Number of umi per cell per cluster

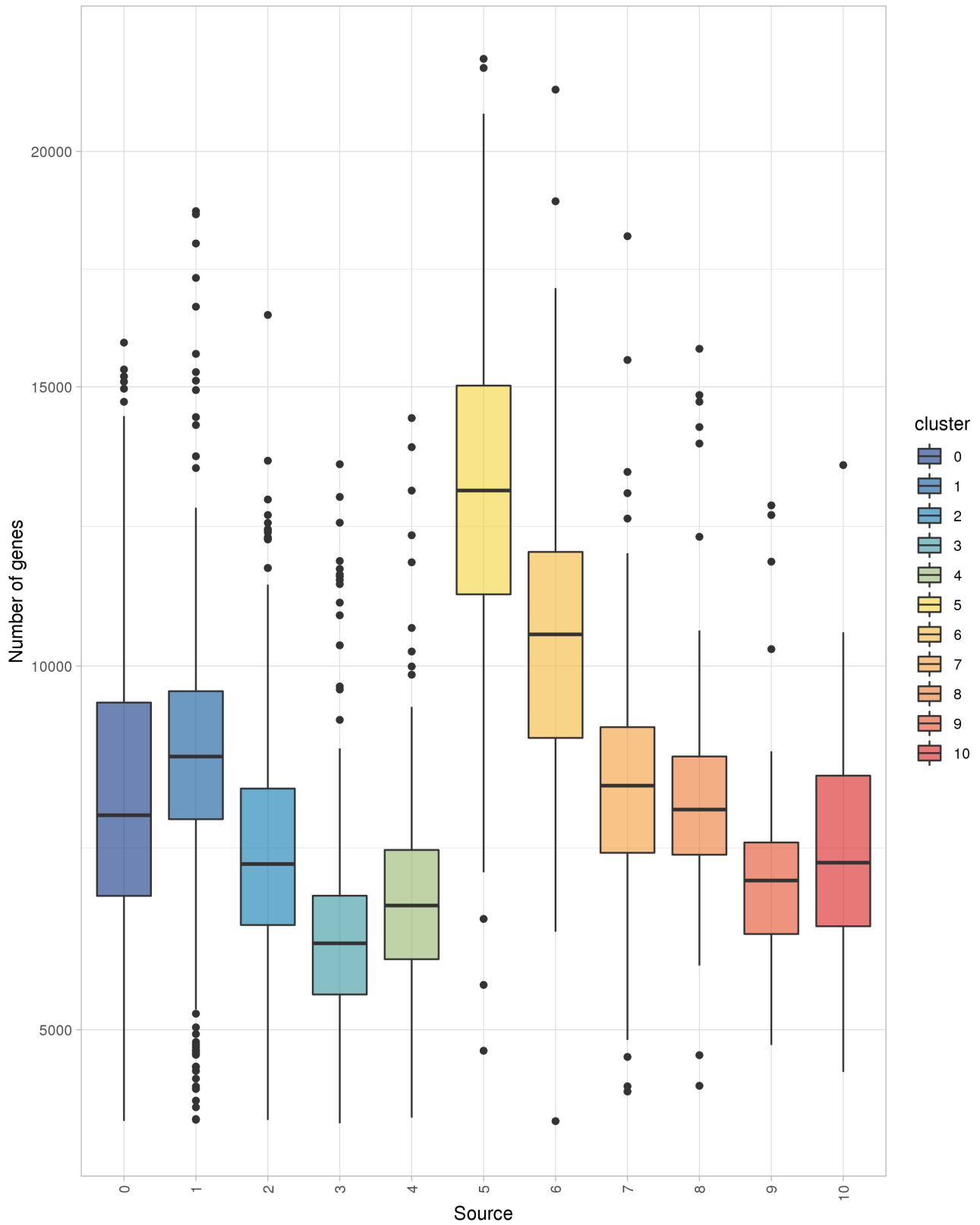


Figure 22: Number of umi per cell per cluster

10 Cluster dissimilarity

10.1 Dissimilarity by gene expression

The distances between the clusters was assessed using the “BuildClusterTree” function in the Seurat package, which “constructs a phylogenetic tree relating the “average” cell from each identity class”.

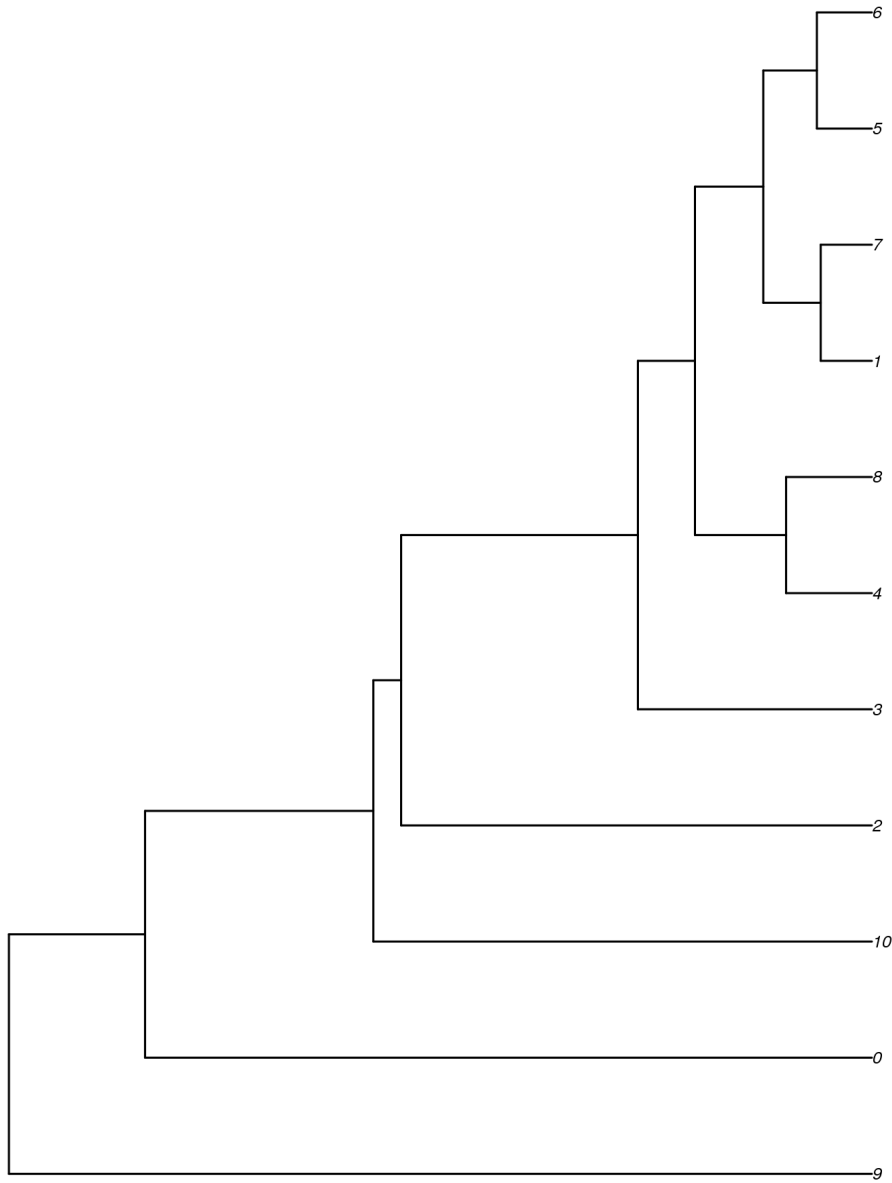


Figure 23: Visualisation of inter-cluster distances (cluster average, gene-based)

11 Cluster resolution analysis

The [clustree algorithm](#) is used to compare the different clustering resolutions.

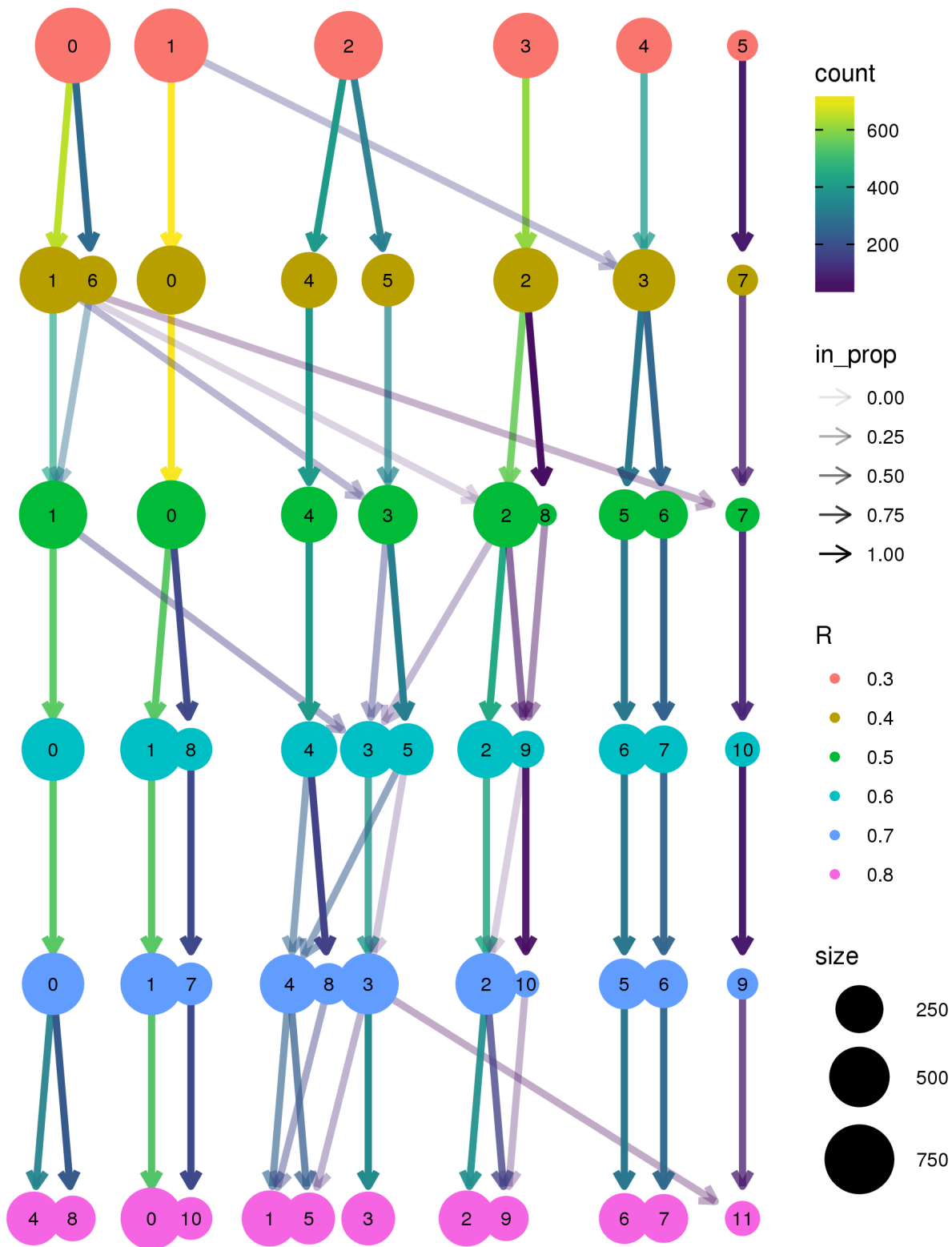


Figure 24: The relationships between the clusters identified at different resolutions

12 Diffusion map

12.0.1 Diffusion map plots (first 3 dimensions, different rotations) colored by cluster

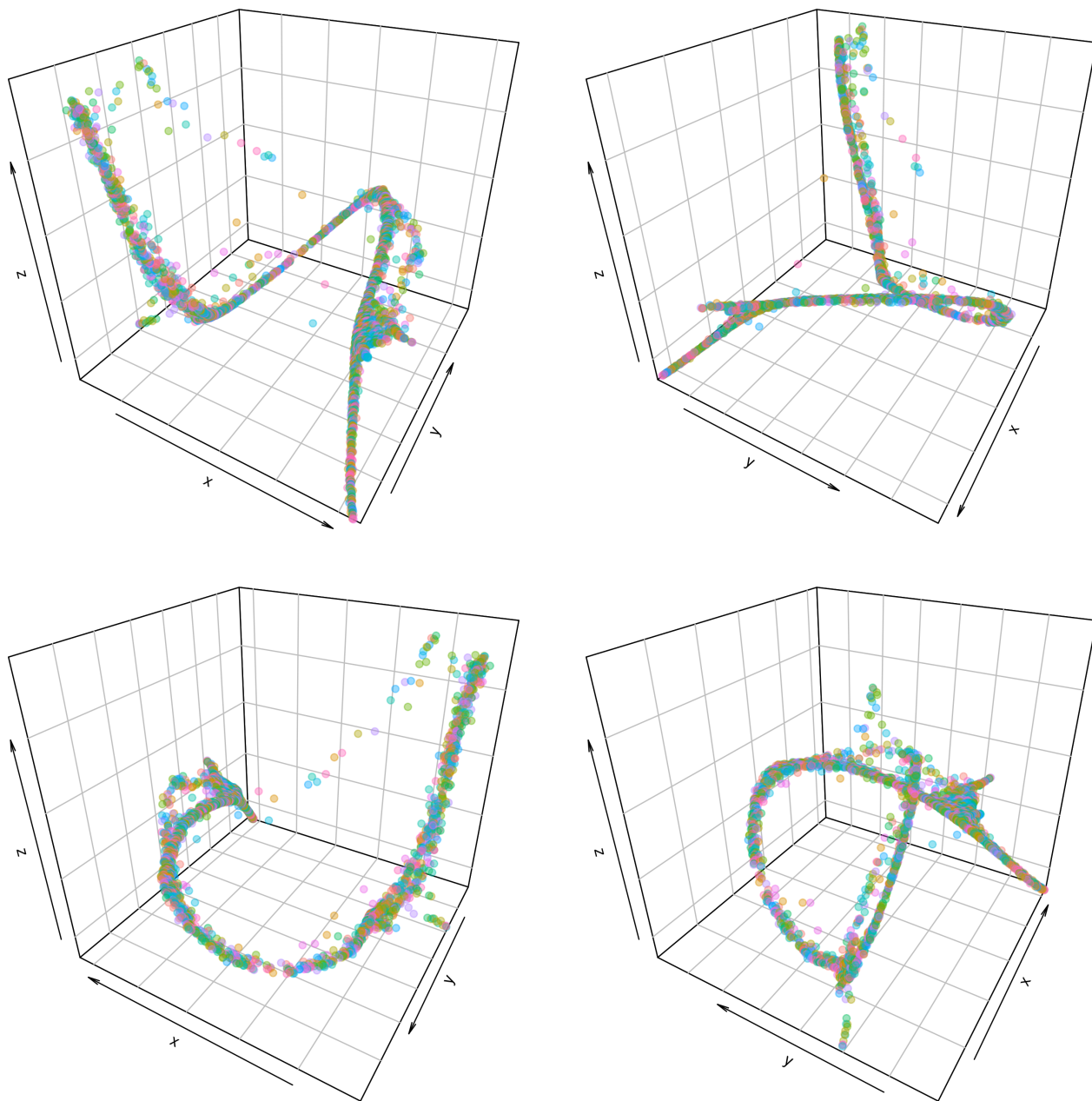


Figure 25: Diffusion map plots (first 3 dimensions, different rotations) colored by cluster

13 Phate maps

See: <https://www.nature.com/articles/s41587-019-0336-3>.

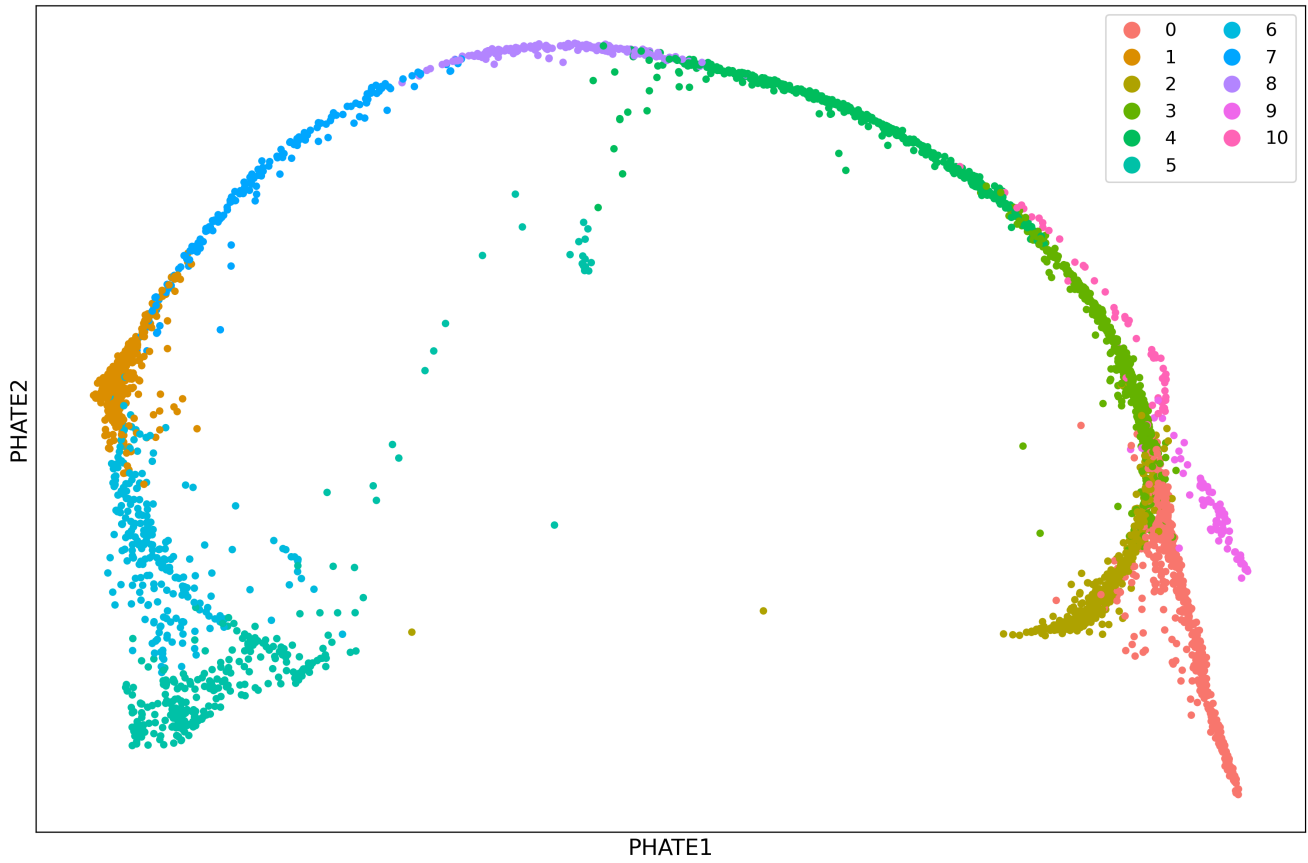


Figure 26: 2D Phate map

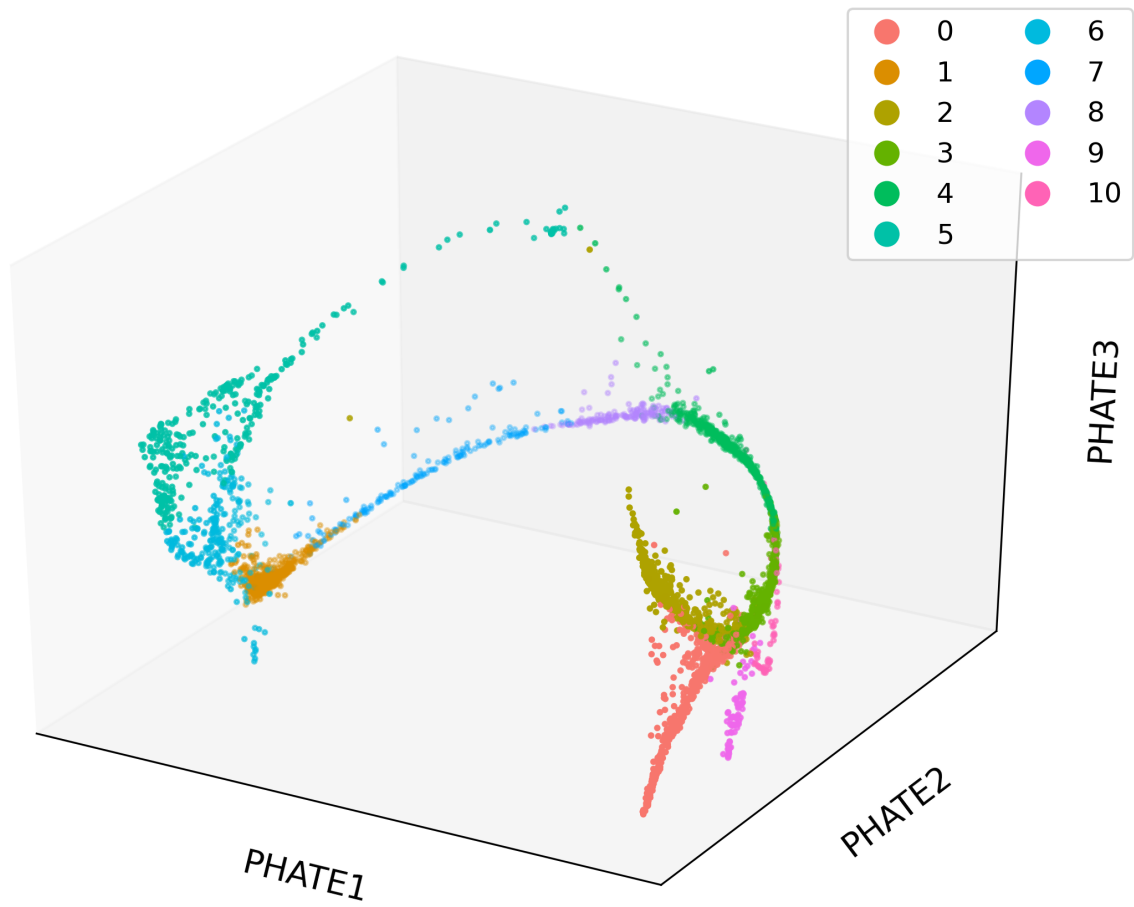


Figure 27: 3D Phate map

14 Partion-based graph abstraction (PAGA)

The original method is described here: [PAGA](#)

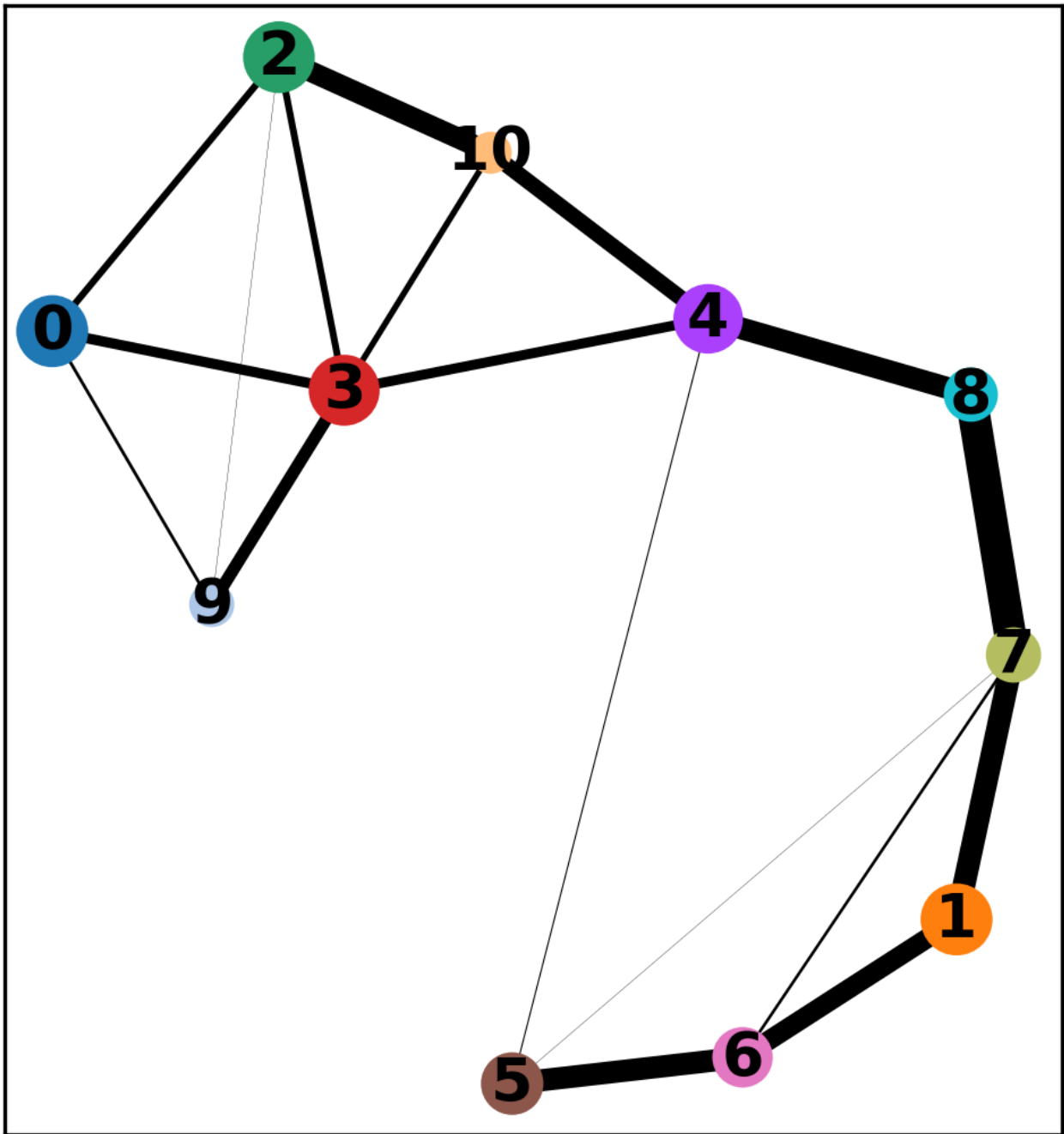


Figure 28: PAGA graph

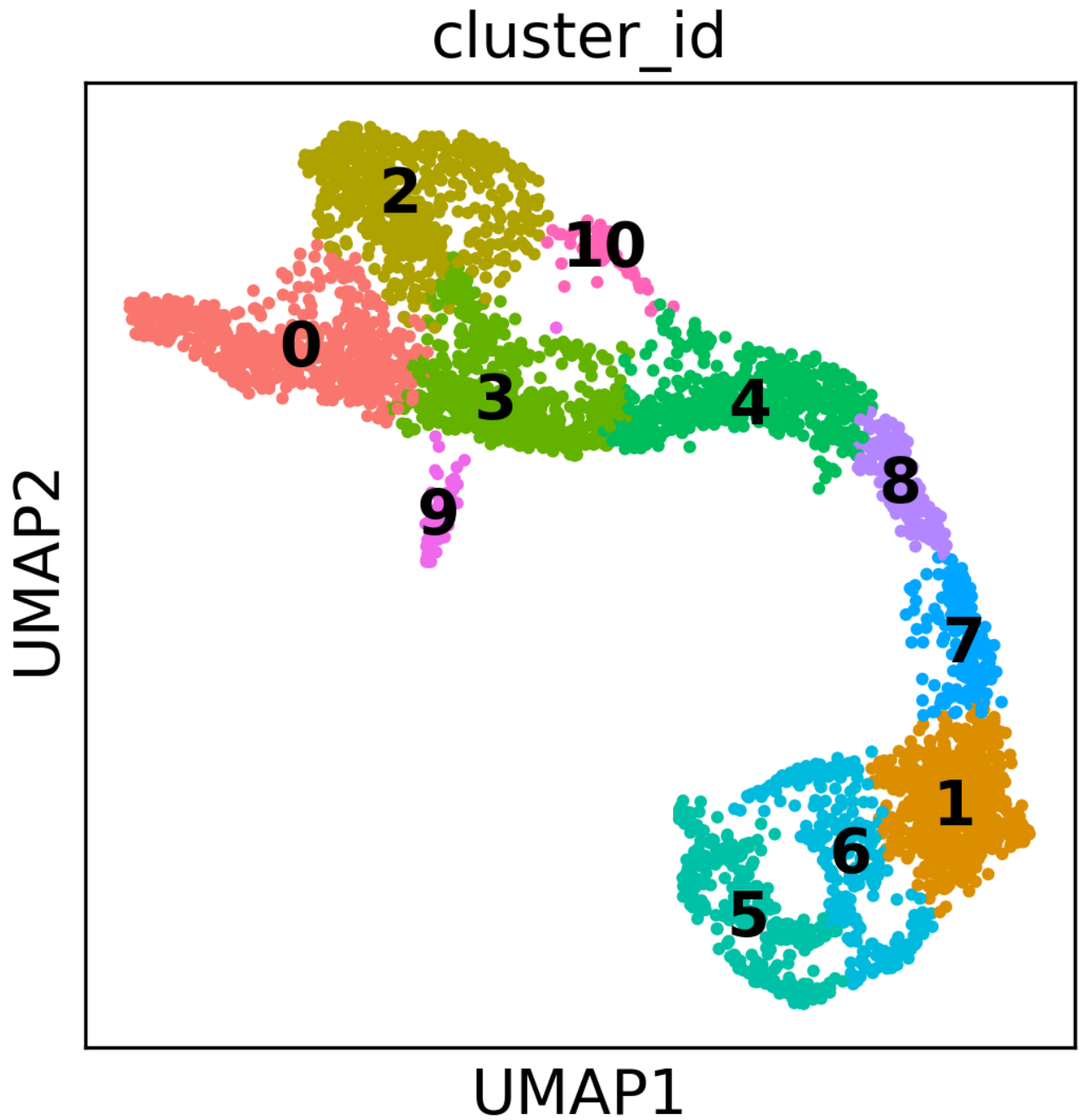


Figure 29: PAGA initialised UMAP

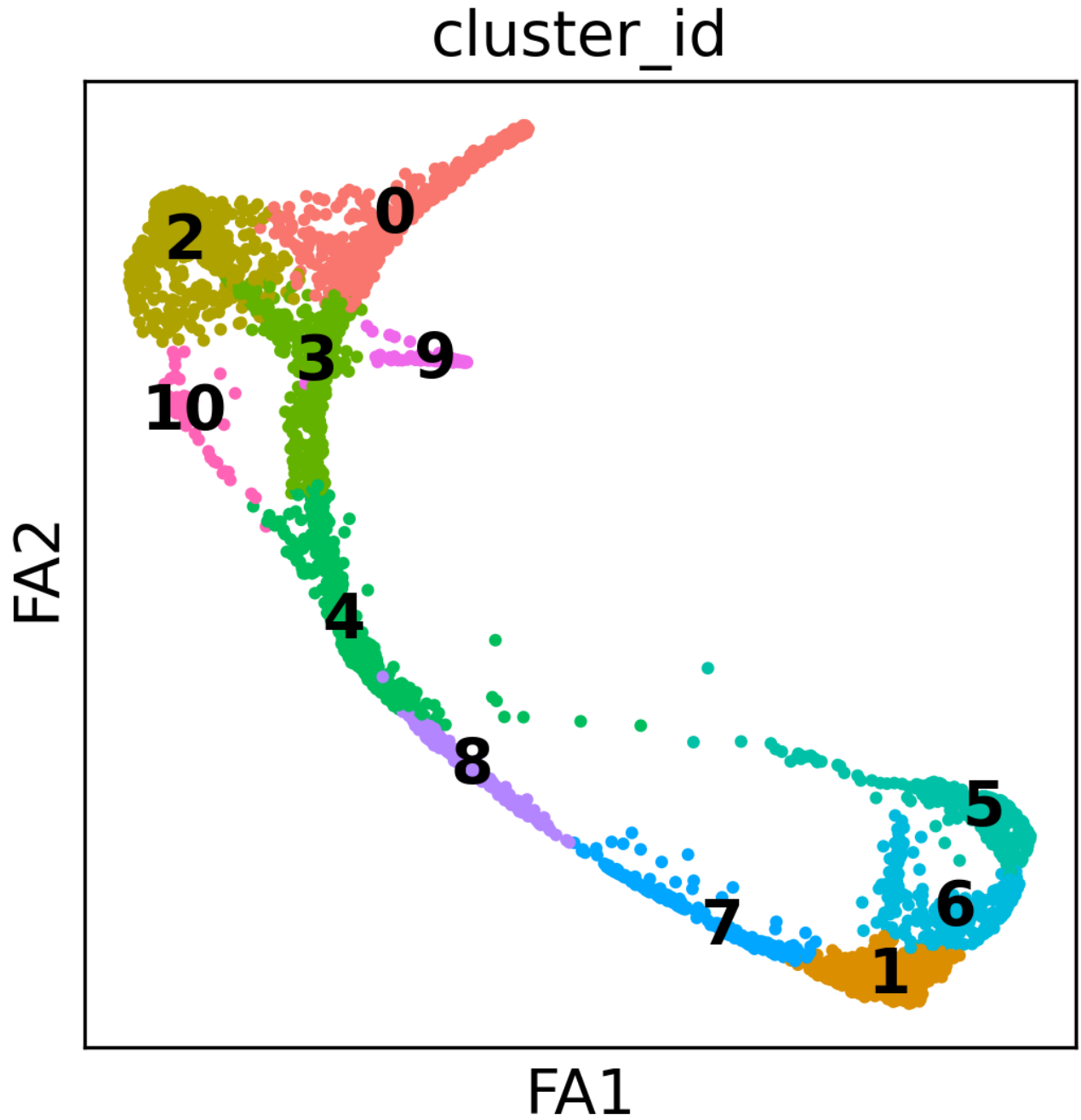


Figure 30: Force atlas 2 (FA2) force directed graph

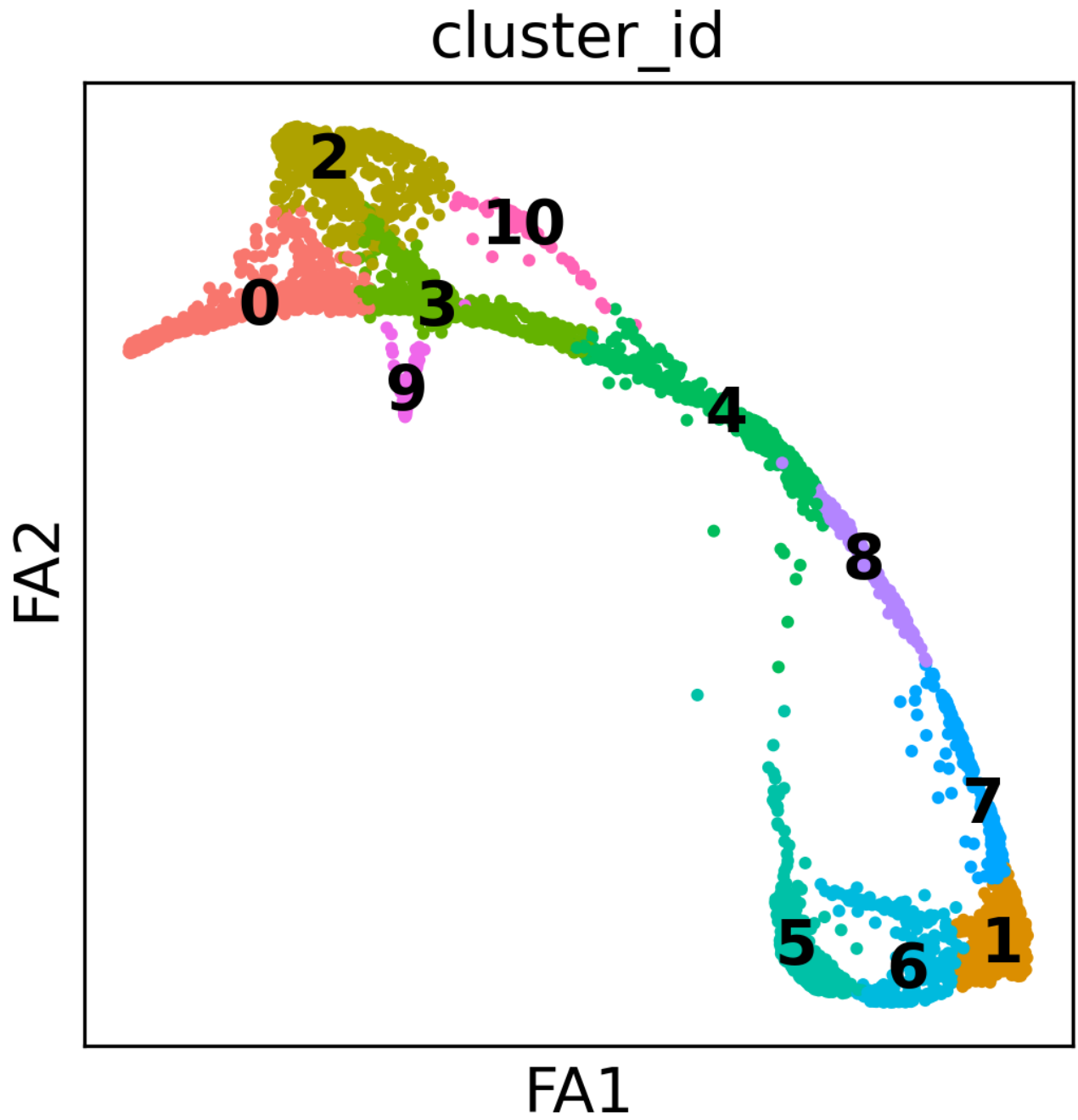


Figure 31: PAGA initialised FA2 force directed graph

15 RNA velocity

The original method is described here: [RNA velocity](#)

We now use scvelo [Generalizing RNA velocity to transient cell states through dynamical modeling](#)

15.1 reduced dimension maps

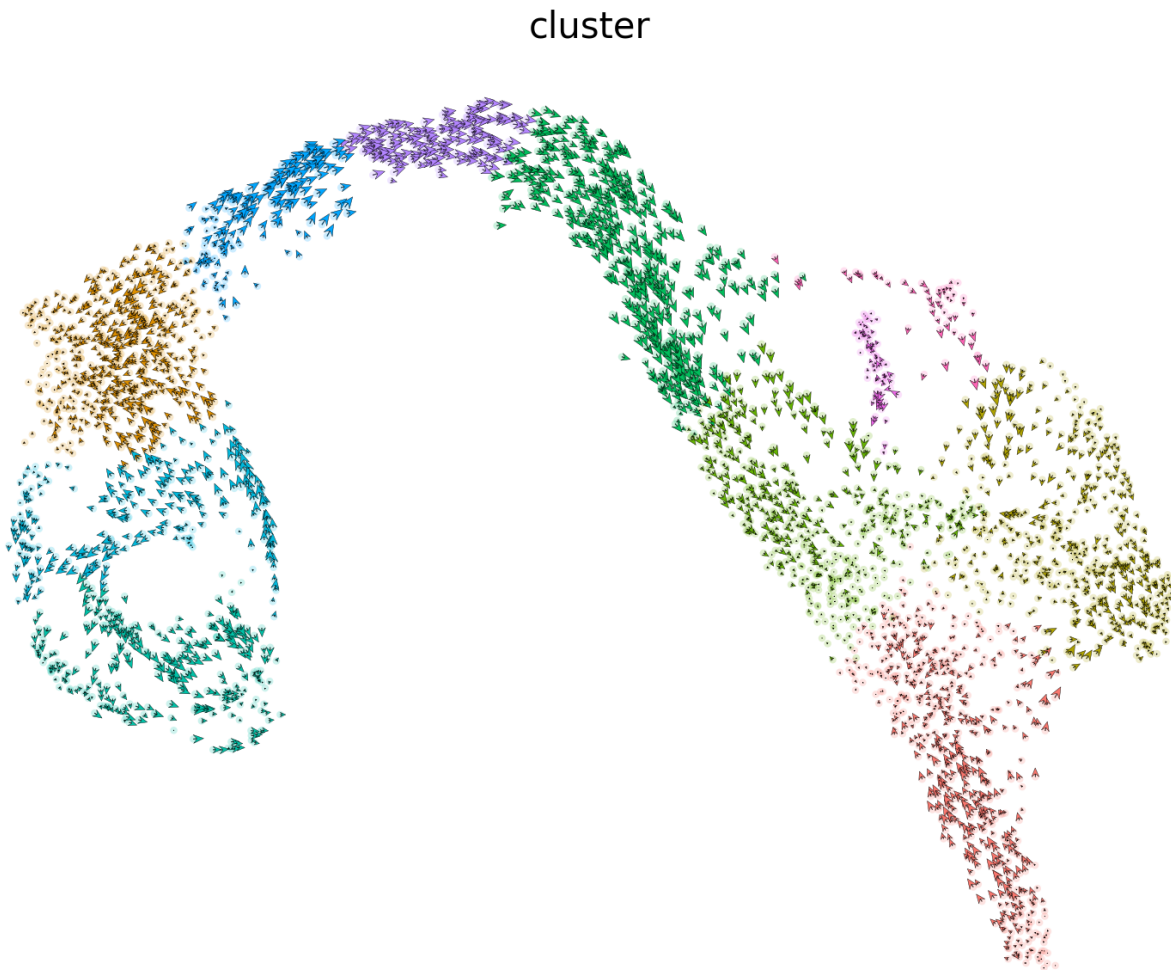


Figure 32: Velocity plot on umap

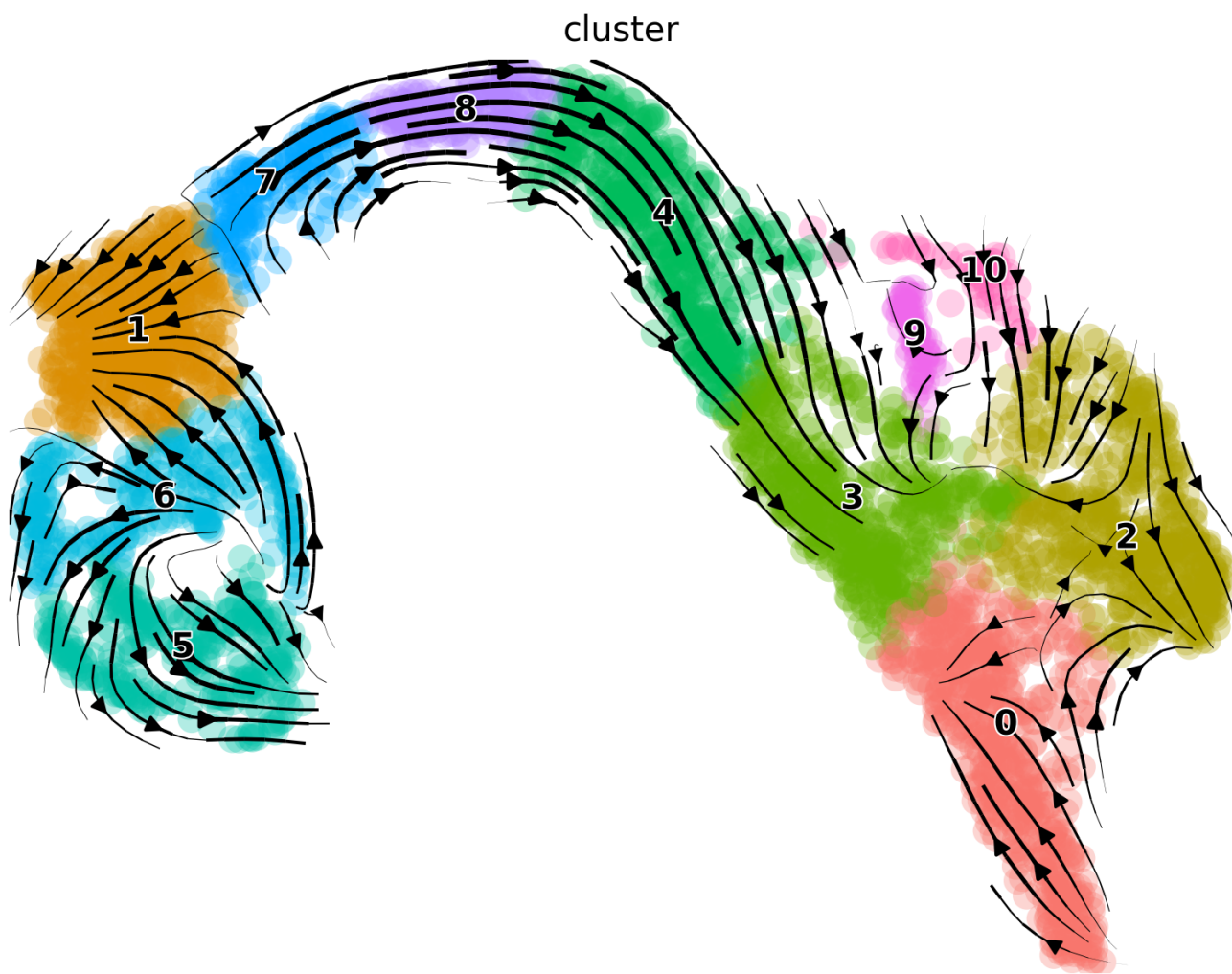


Figure 33: Velocity stream plot on umap

cluster



Figure 34: Velocity grid plot on umap

15.2 PAGA initialised FA2 graphs

cluster

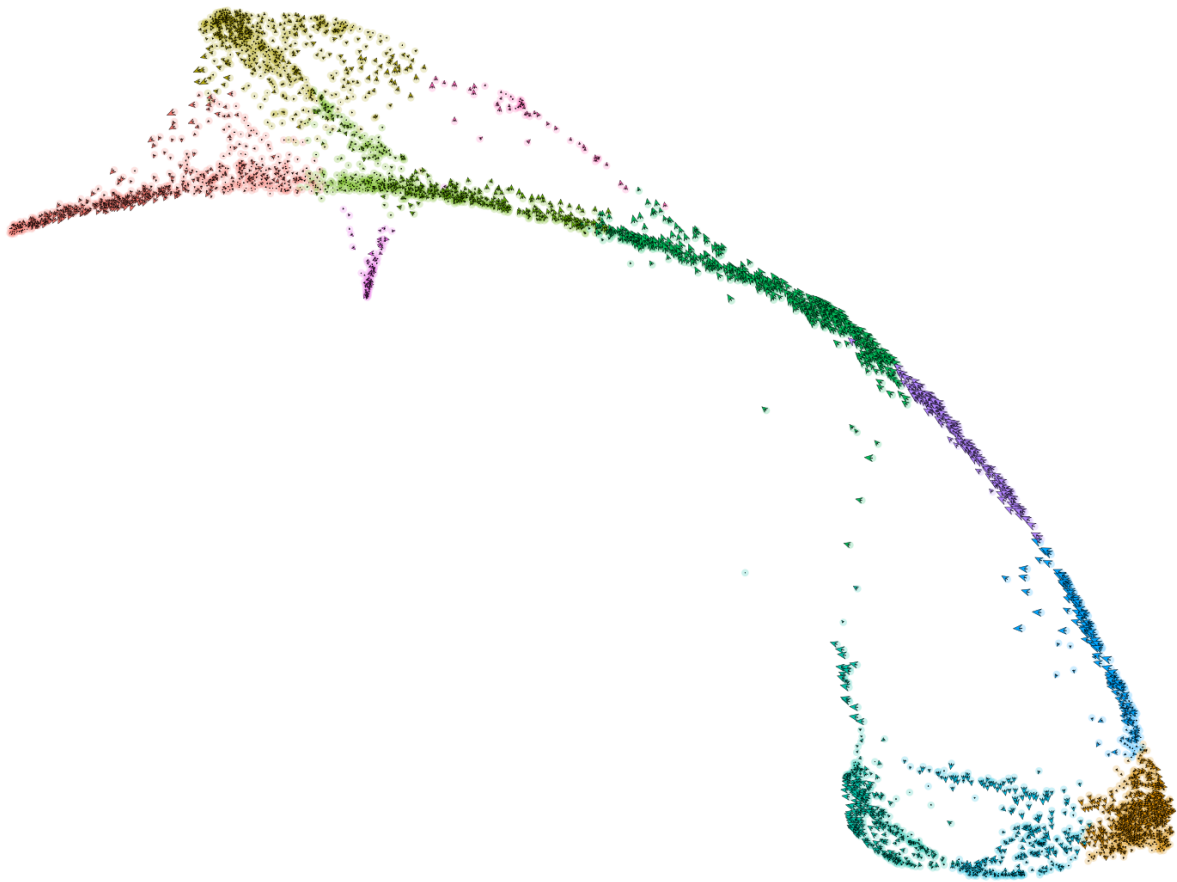


Figure 35: Velocity plot on PAGA initialised force atlas 2 graph

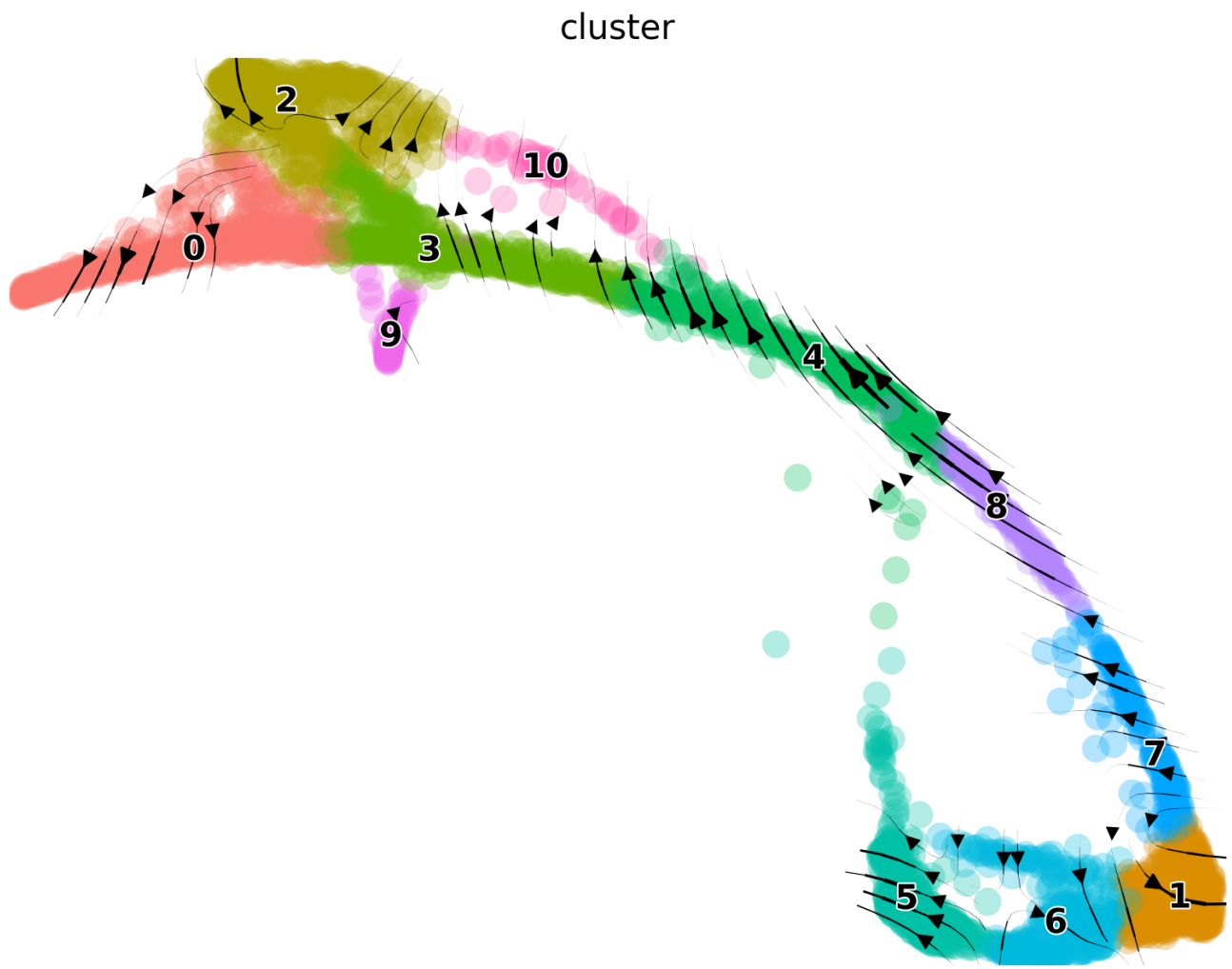


Figure 36: Velocity stream plot on PAGA initialised force atlas 2 graph

cluster

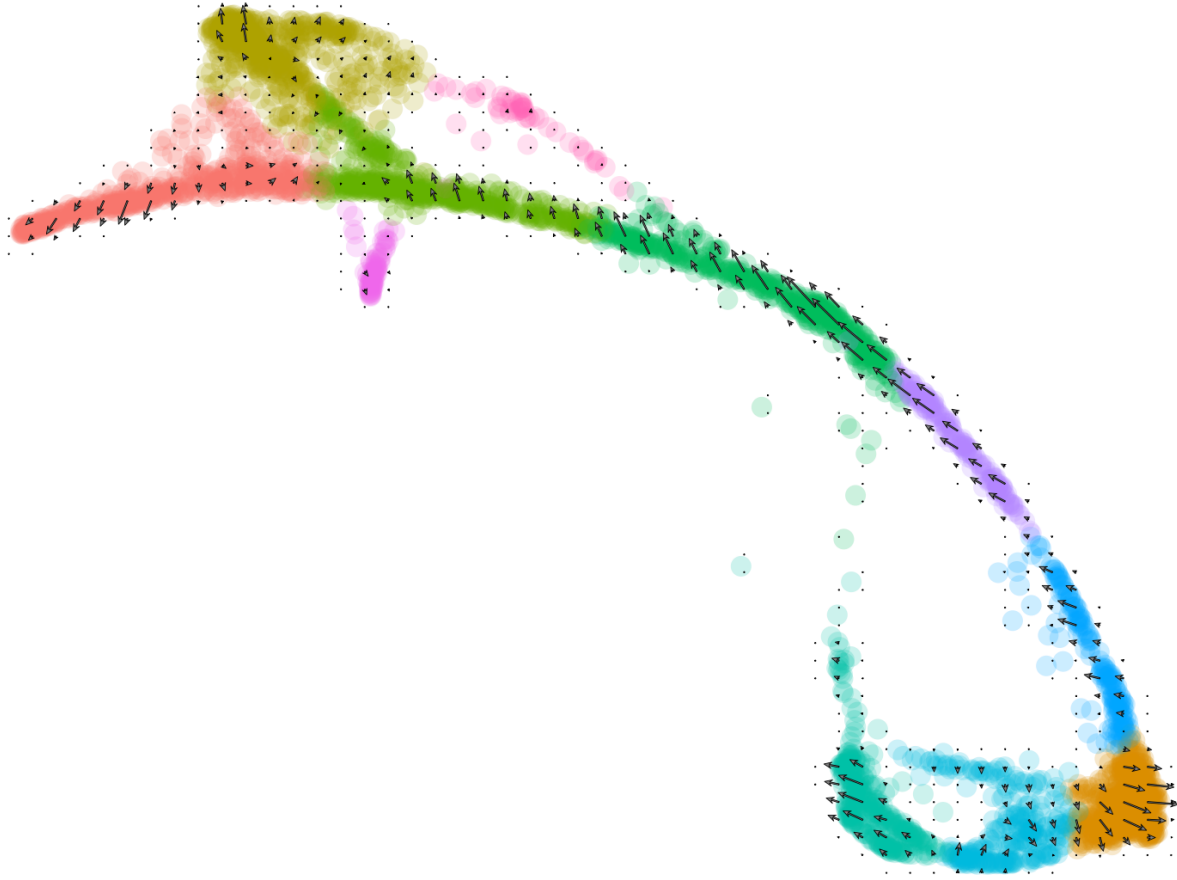


Figure 37: Velocity grid plot on PAGA initialised force atlas 2 graph

15.3 PHATE graphs

cluster

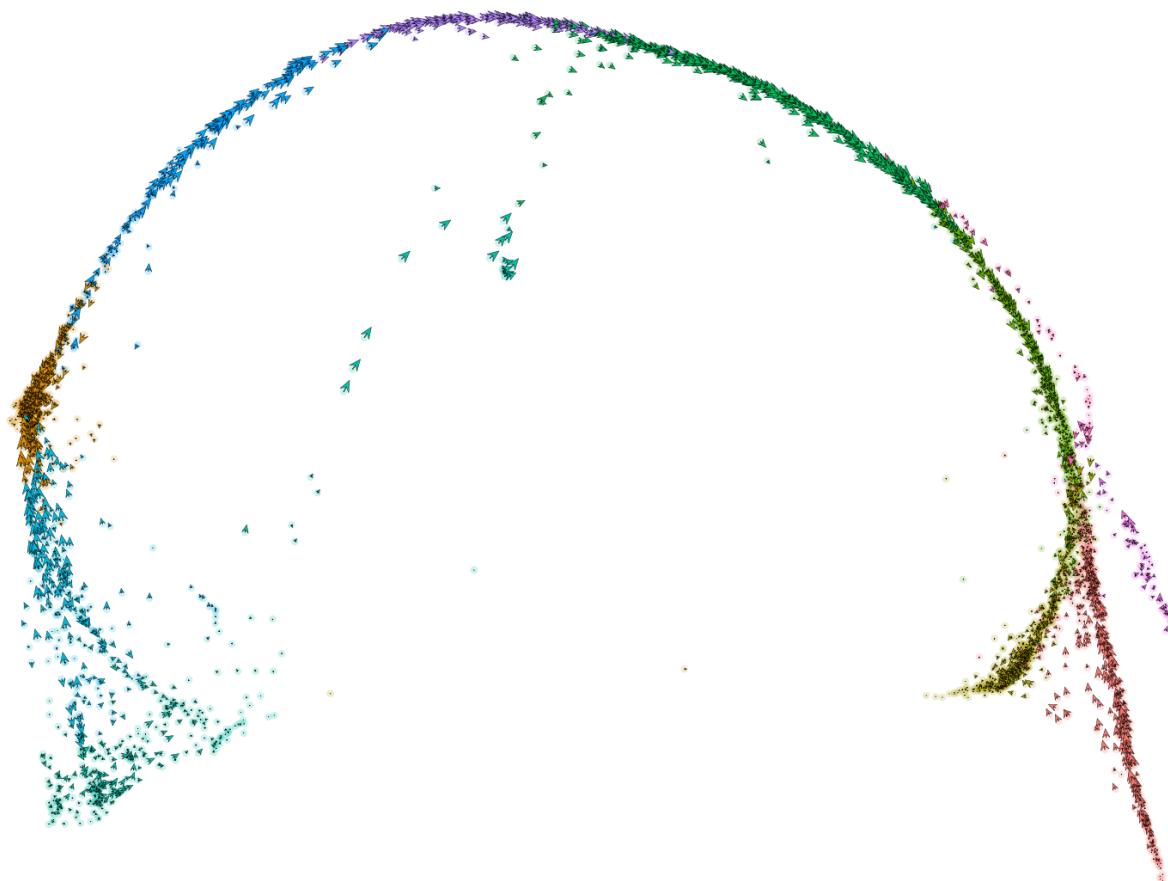


Figure 38: Velocity plot on PHATE graph

cluster

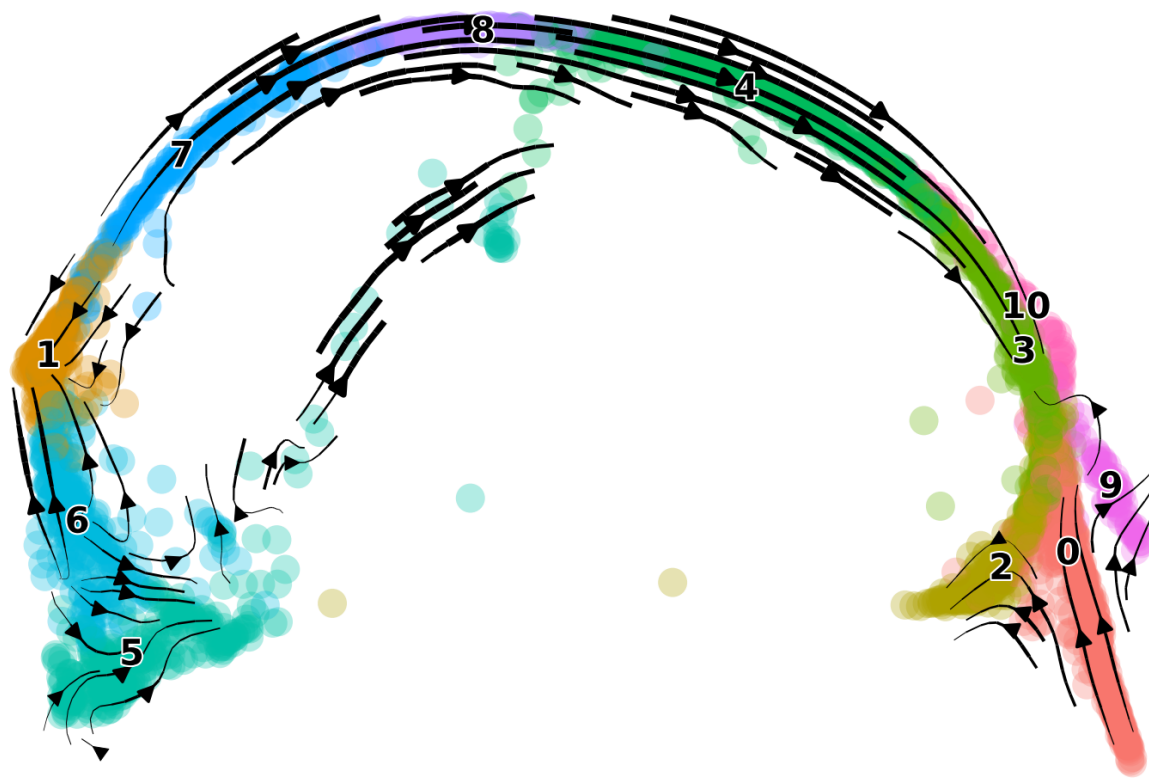


Figure 39: Velocity stream plot on PHATE graph

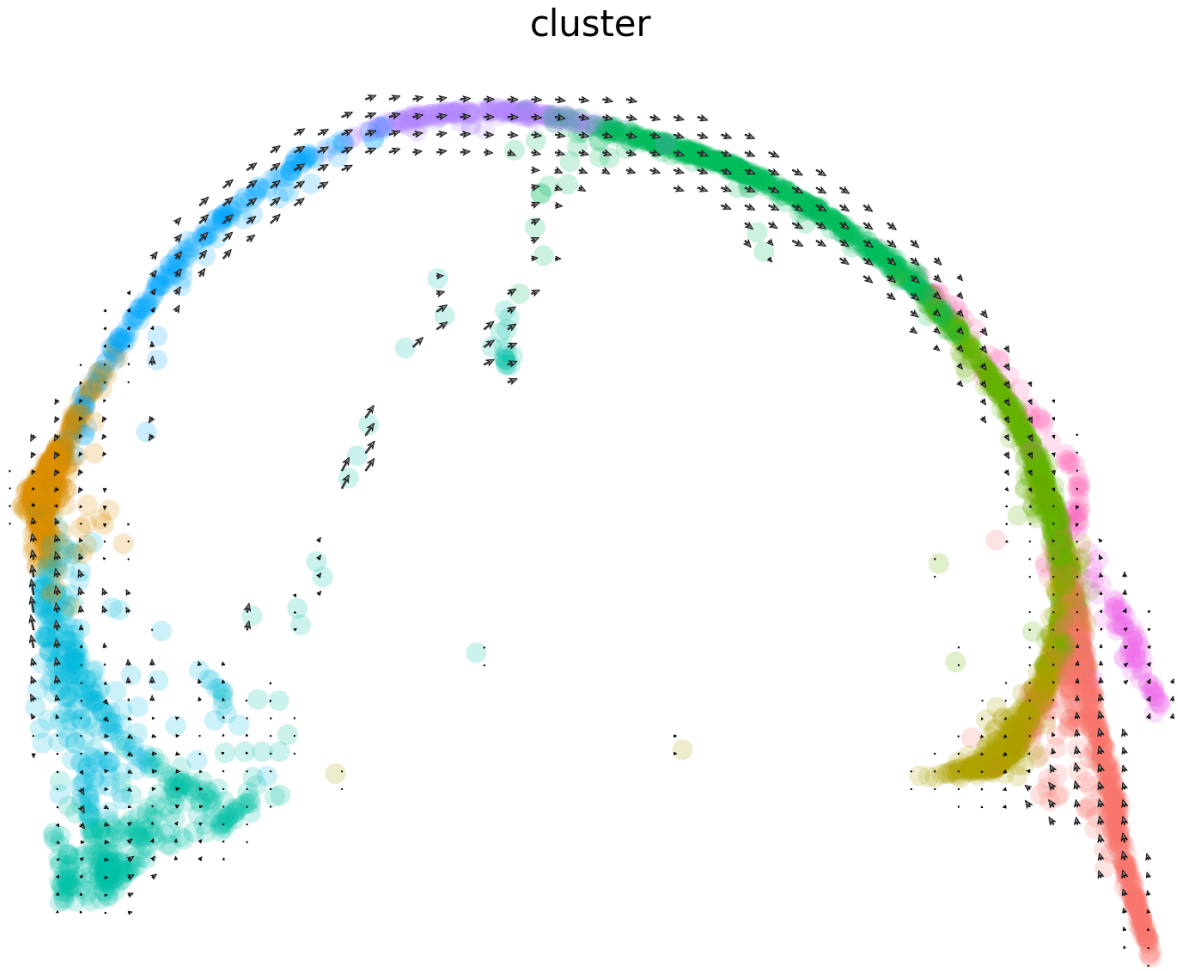


Figure 40: Velocity grid plot on PHATE graph

16 Identification of cluster marker genes

Cluster marker genes were identified using the Seurat FindMarkers routine and the wilcox test. A summary table containing all of the significant markers for all of the clusters (based on BH adjusted p value) is available separately.

Key parameters are:

- Differential expression methods: wilcox
- Testing limited to genes with a log fold change of > 0.25
- Testing limited to genes detected in a minimum fraction of 0.1 of cells
- Conservation factor applied: None

18 Marker genes by cluster

18.0.1 Summary of numbers of DE genes per-cluster

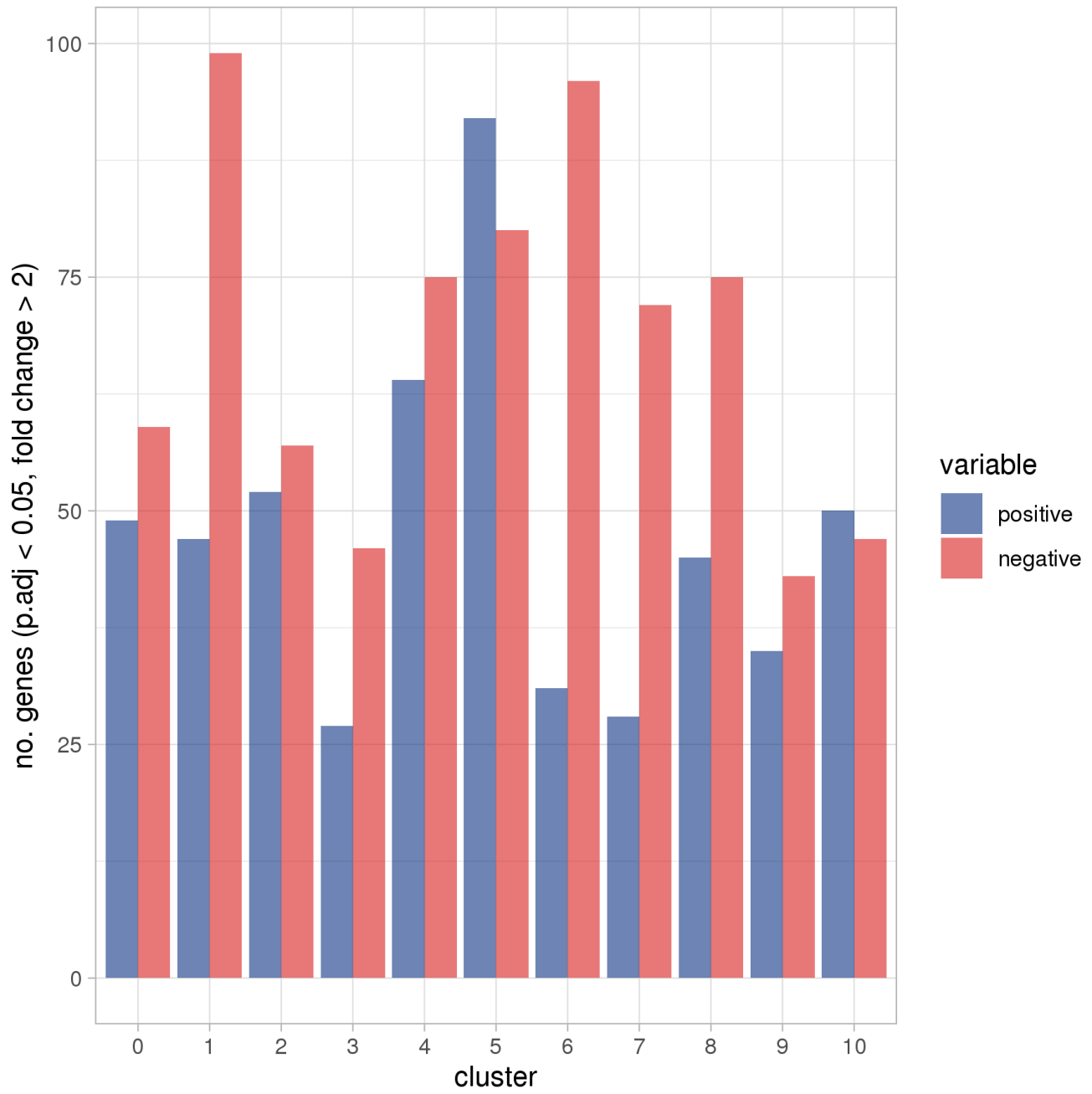


Figure 42: Numbers of differentially expressed genes (adjusted p-value < 0.05, fold change > 2) per cluster

18.1 Cluster 0: summary plots

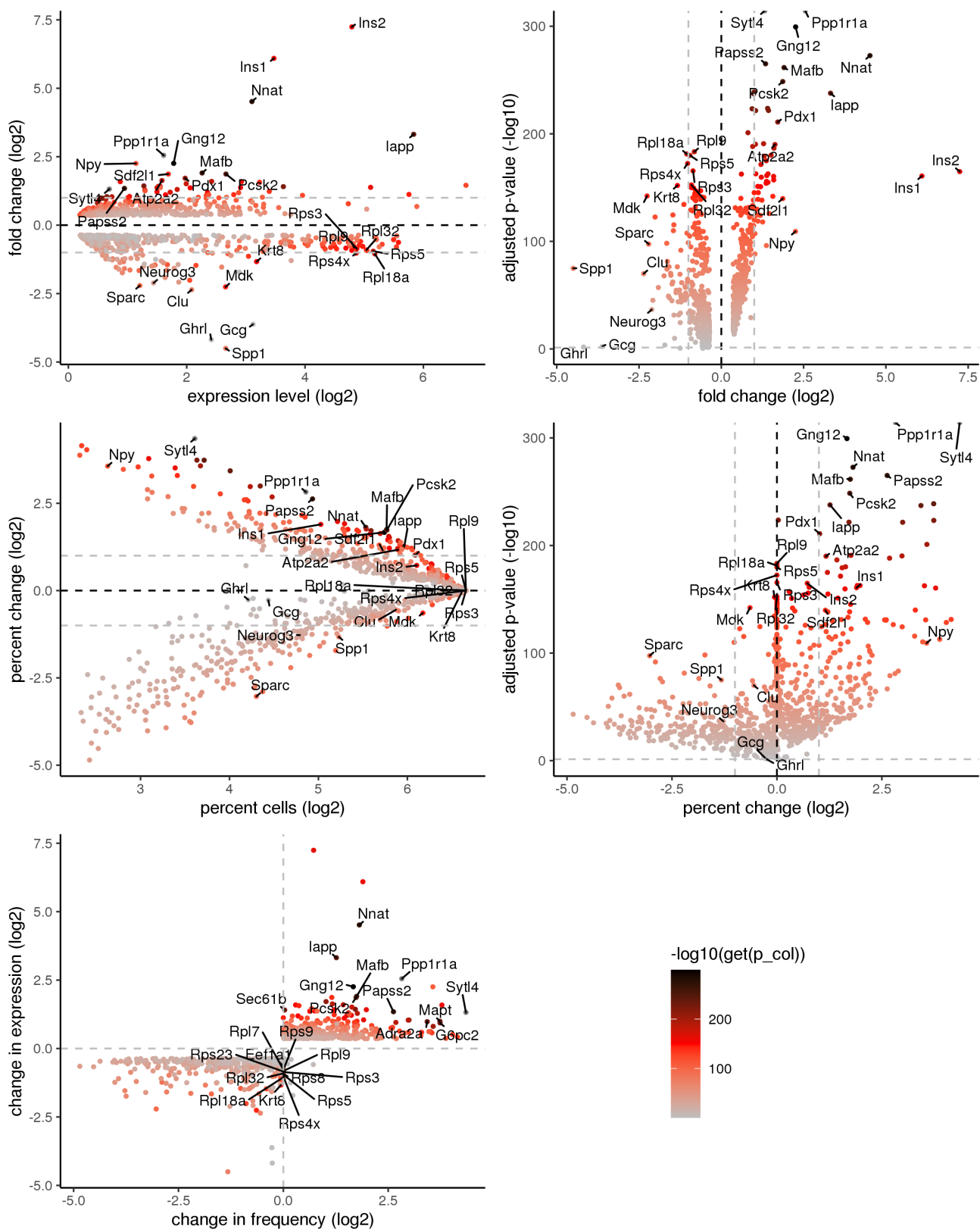
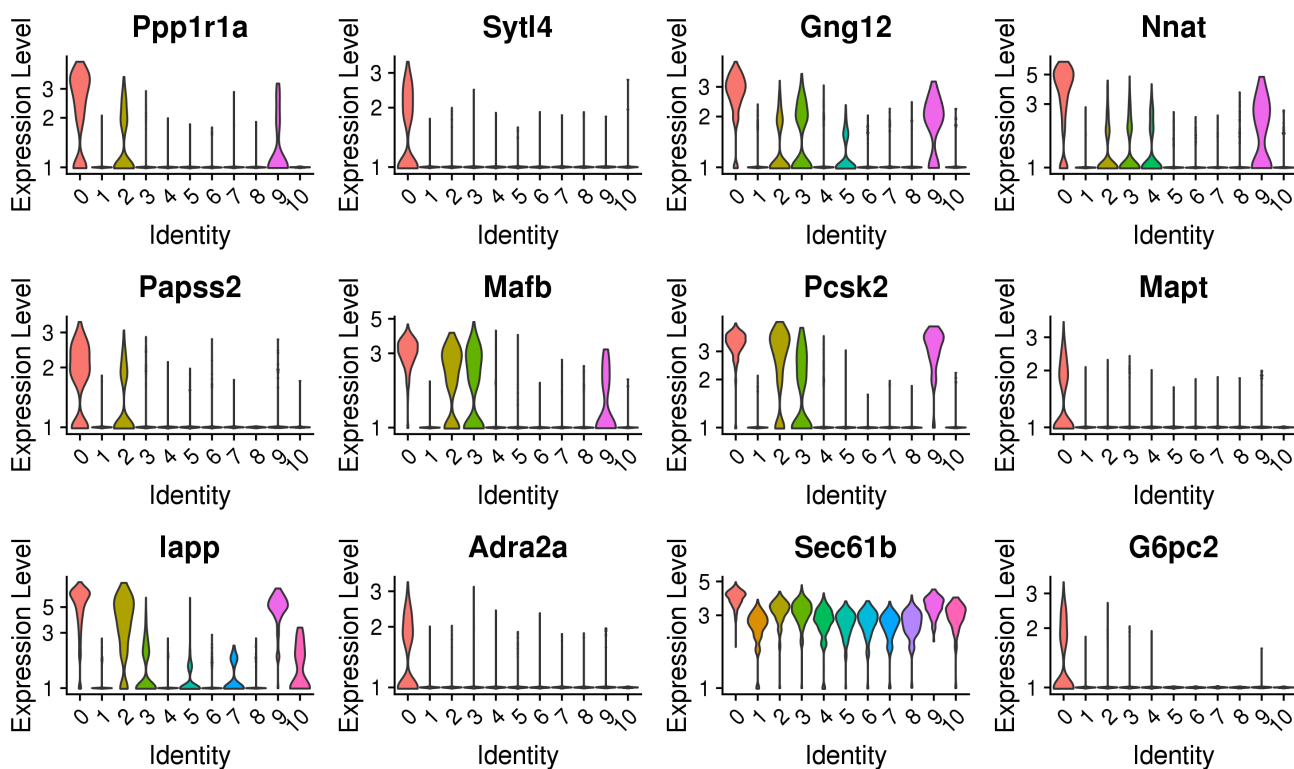
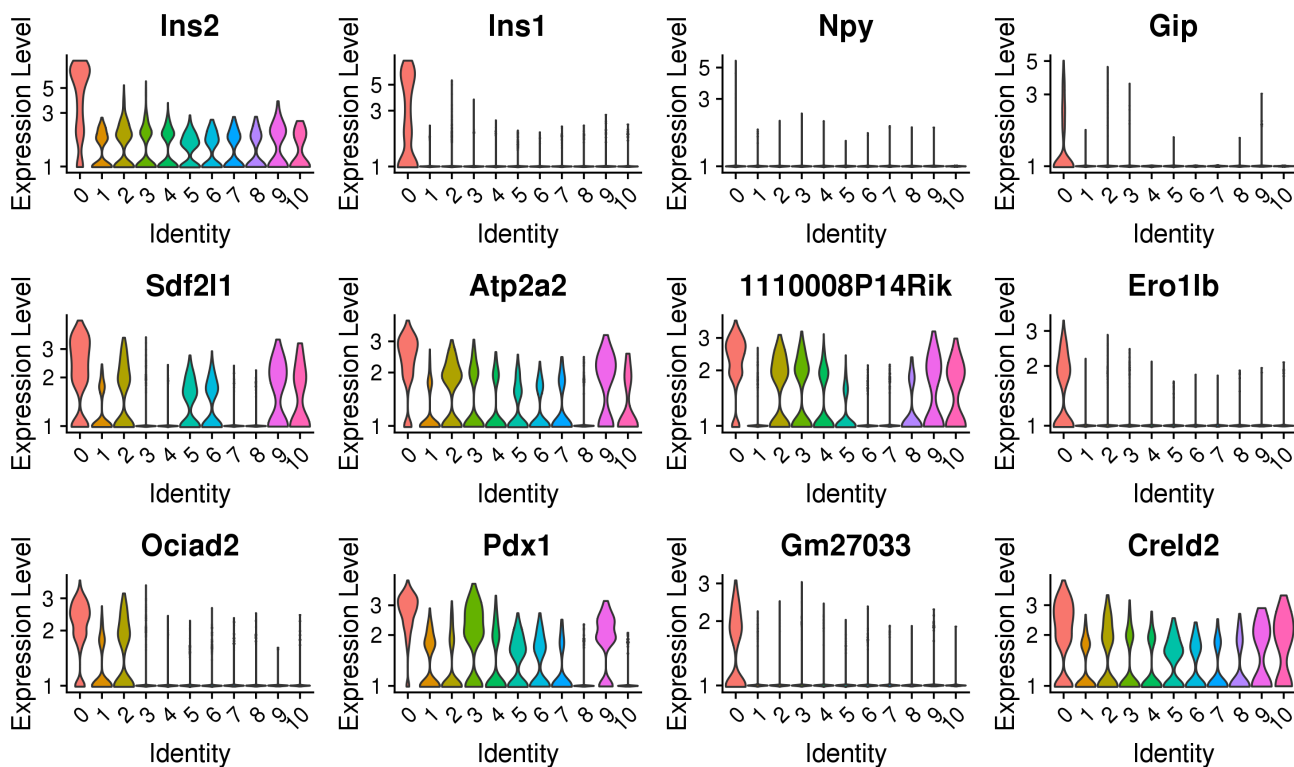


Figure 43: Differential expression summary plots for cluster 0

18.2 Cluster 0 violin plots: positive marker genes

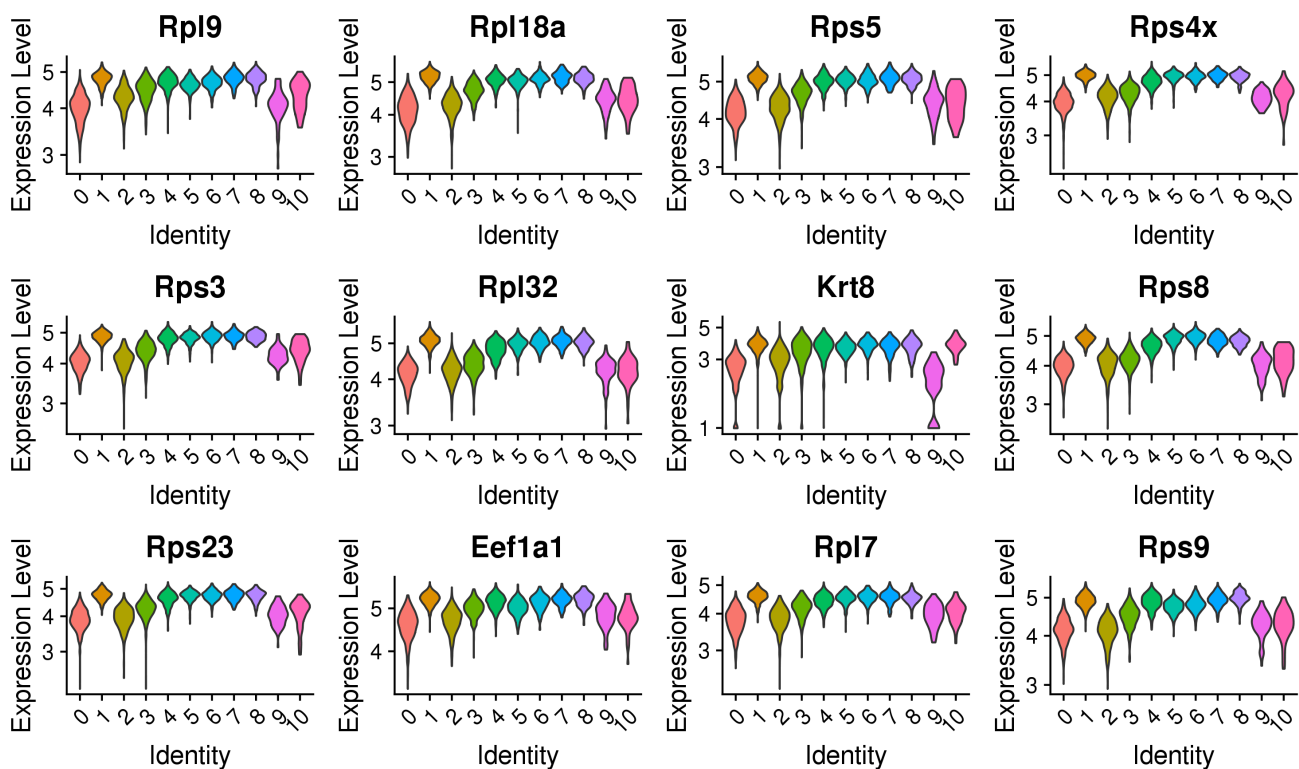


(a) Top positive marker genes ordered by p-value, cluster: 0

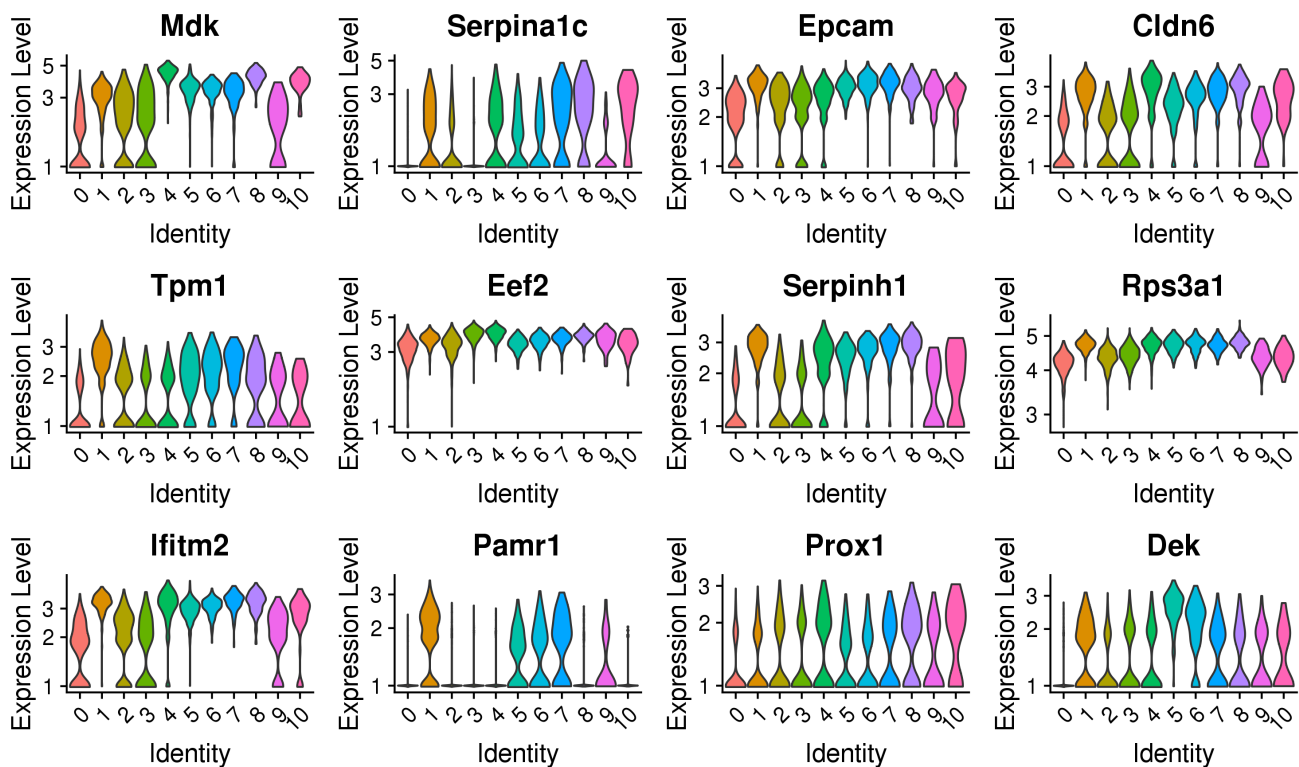


(b) Additional positive marker genes ordered by minimum fold-change vs all other clusters, cluster: 0

18.3 Cluster 0 violin plots: negative marker genes

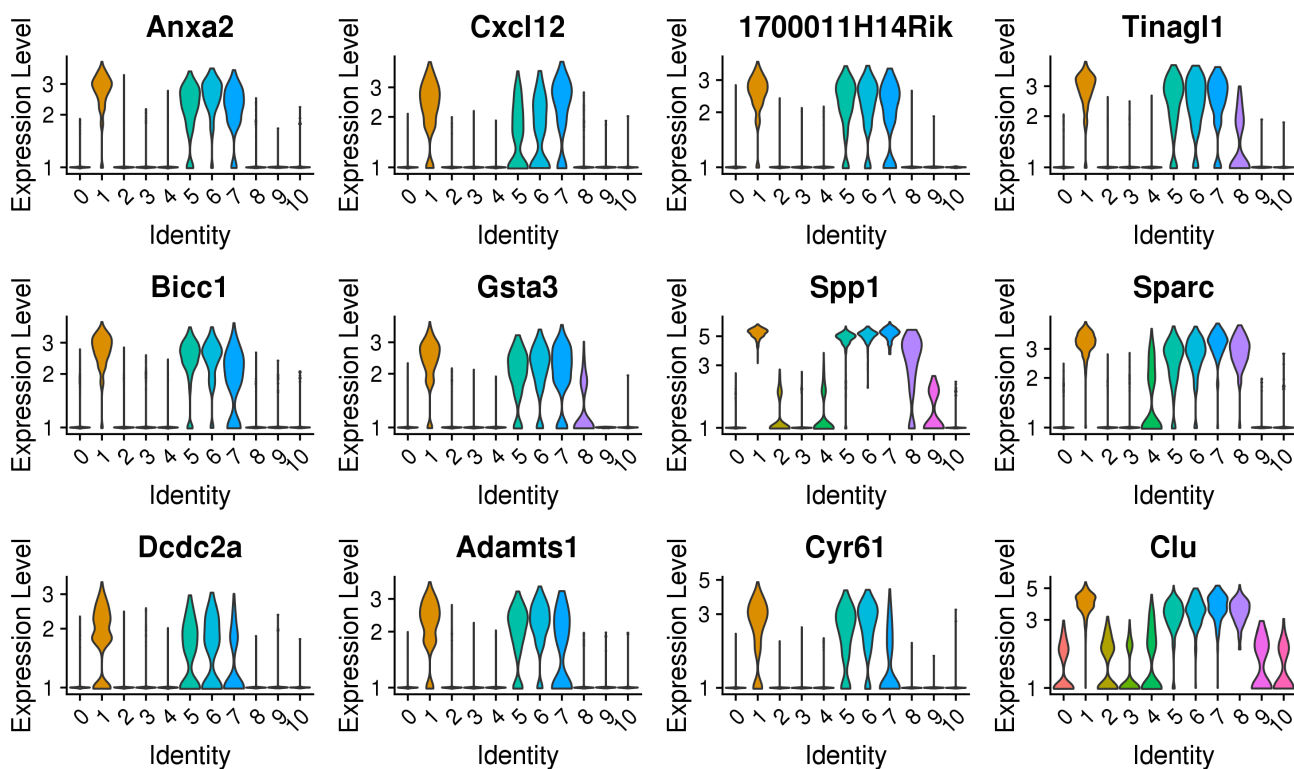


(a) Top negative marker genes ordered by p-value, cluster: 0

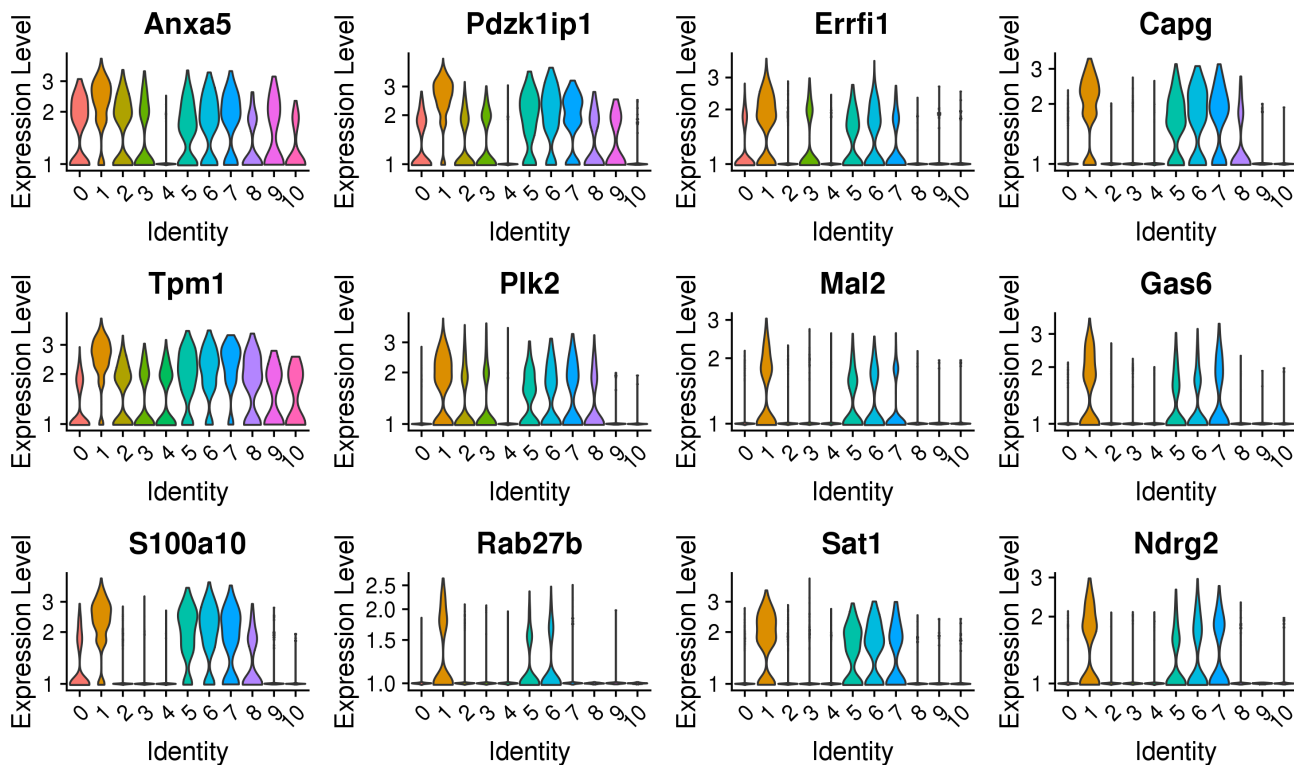


(b) Additional negative marker genes ordered by minimum fold-change vs all other clusters, cluster: 0

18.5 Cluster 1 violin plots: positive marker genes

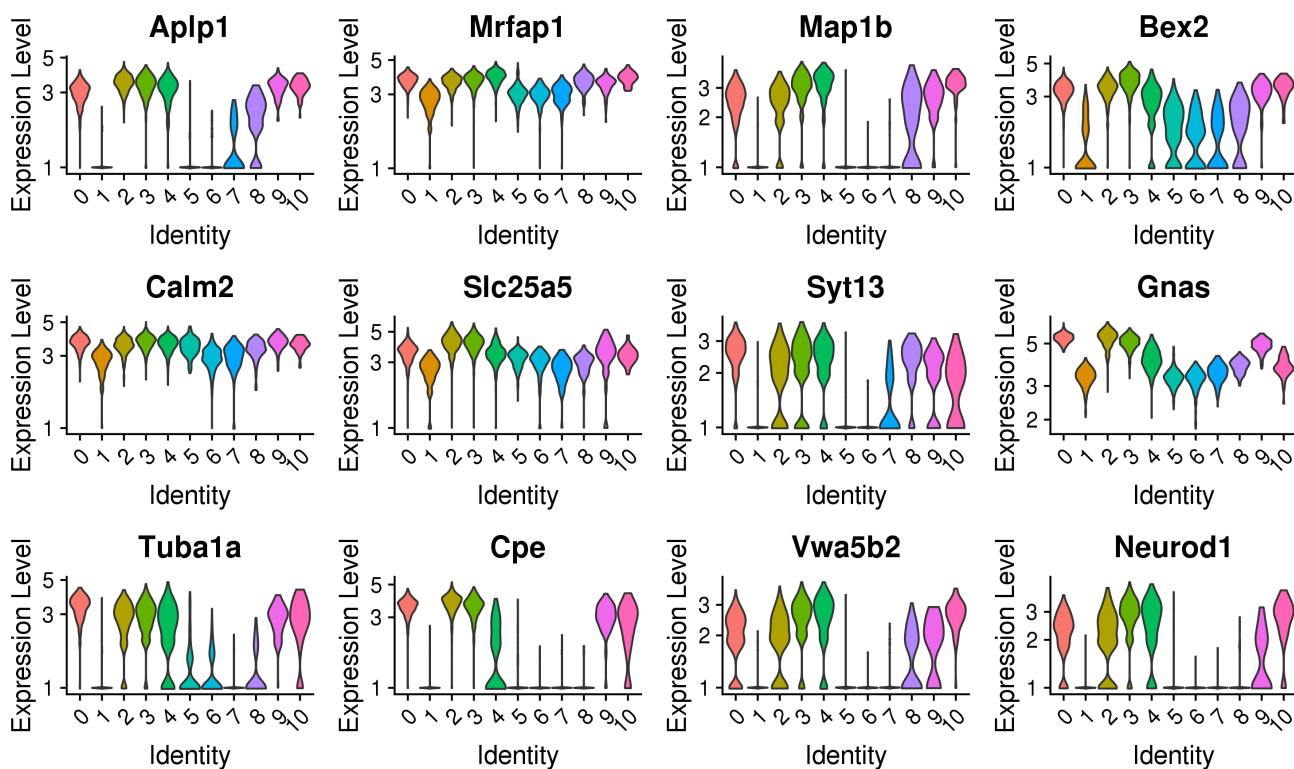


(a) Top positive marker genes ordered by p-value, cluster: 1

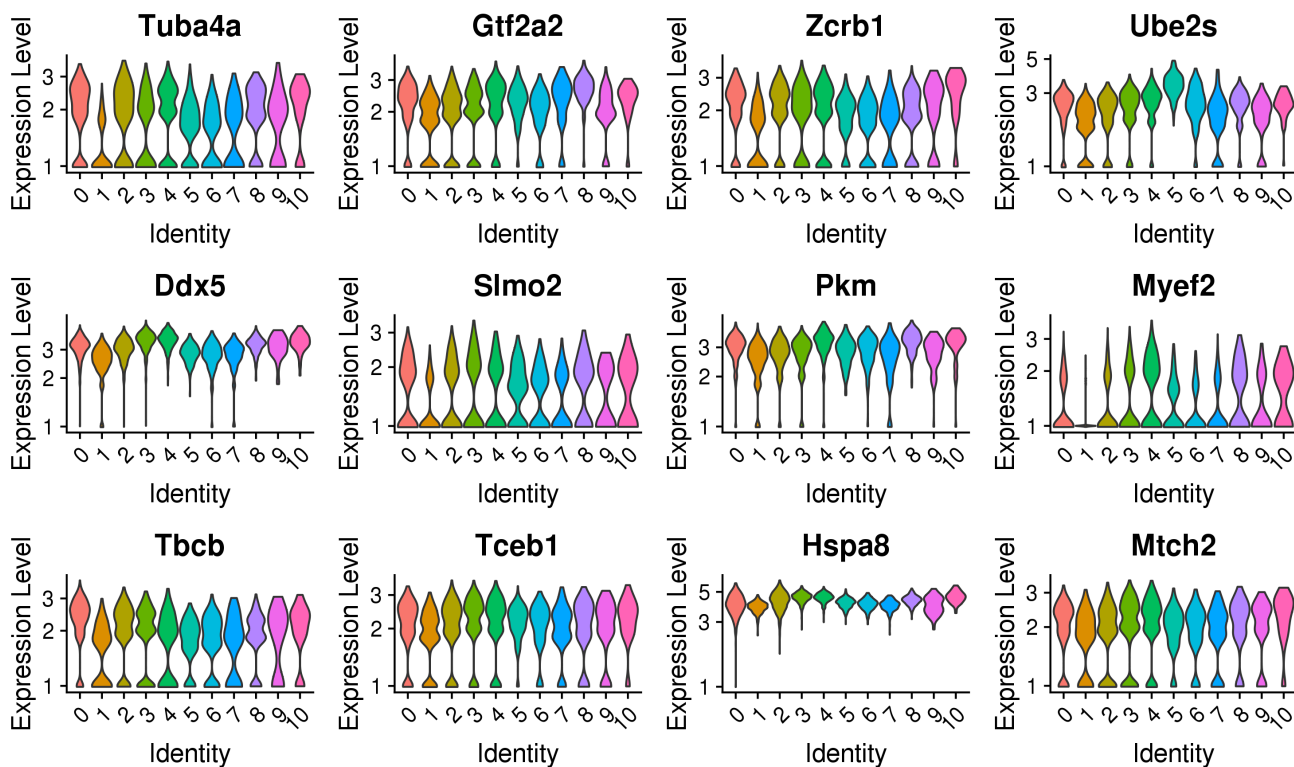


(b) Additional positive marker genes ordered by minimum fold-change vs all other clusters, cluster: 1

18.6 Cluster 1 violin plots: negative marker genes



(a) Top negative marker genes ordered by p-value, cluster: 1



(b) Additional negative marker genes ordered by minimum fold-change vs all other clusters, cluster: 1

18.7 Cluster 2: summary plots

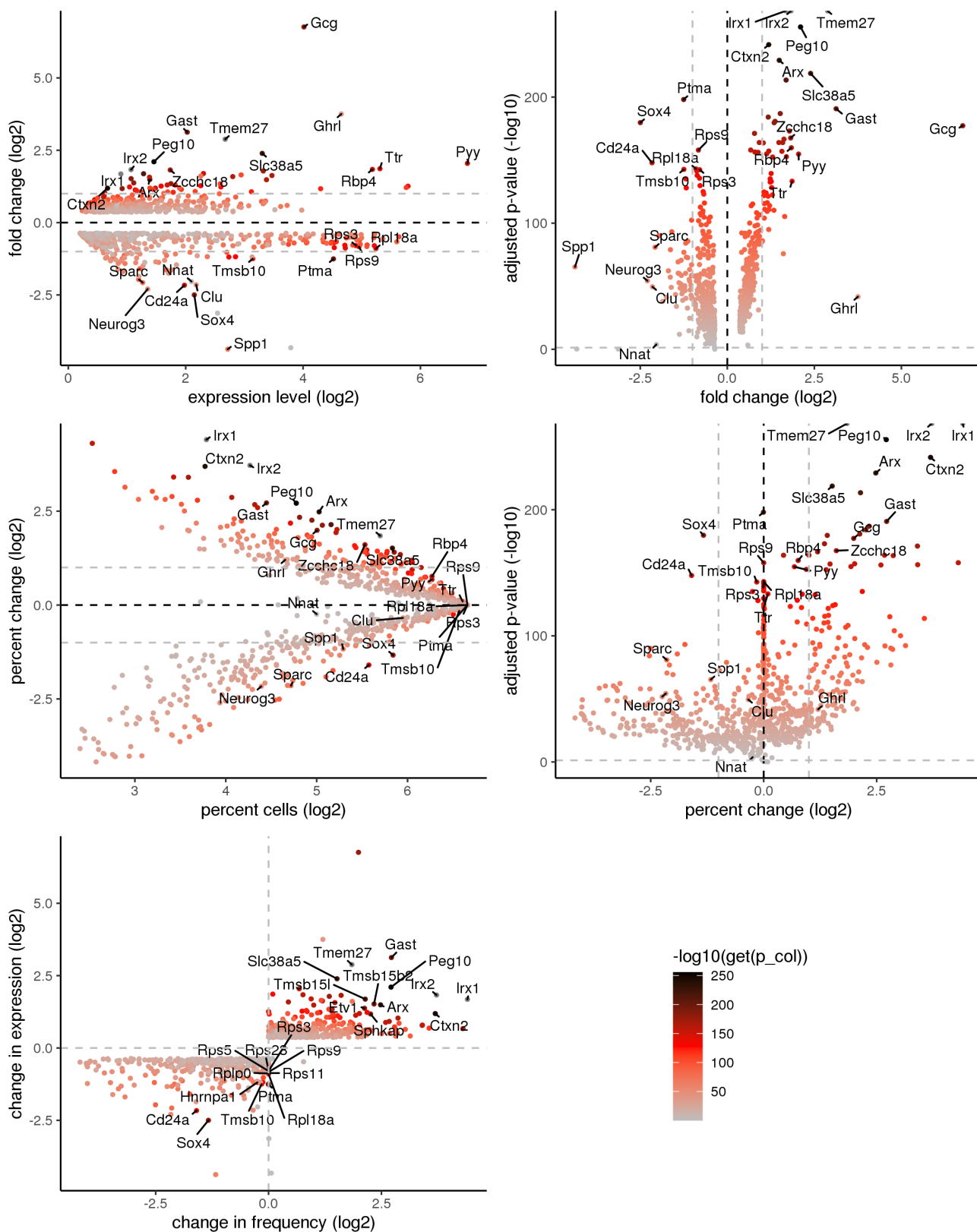
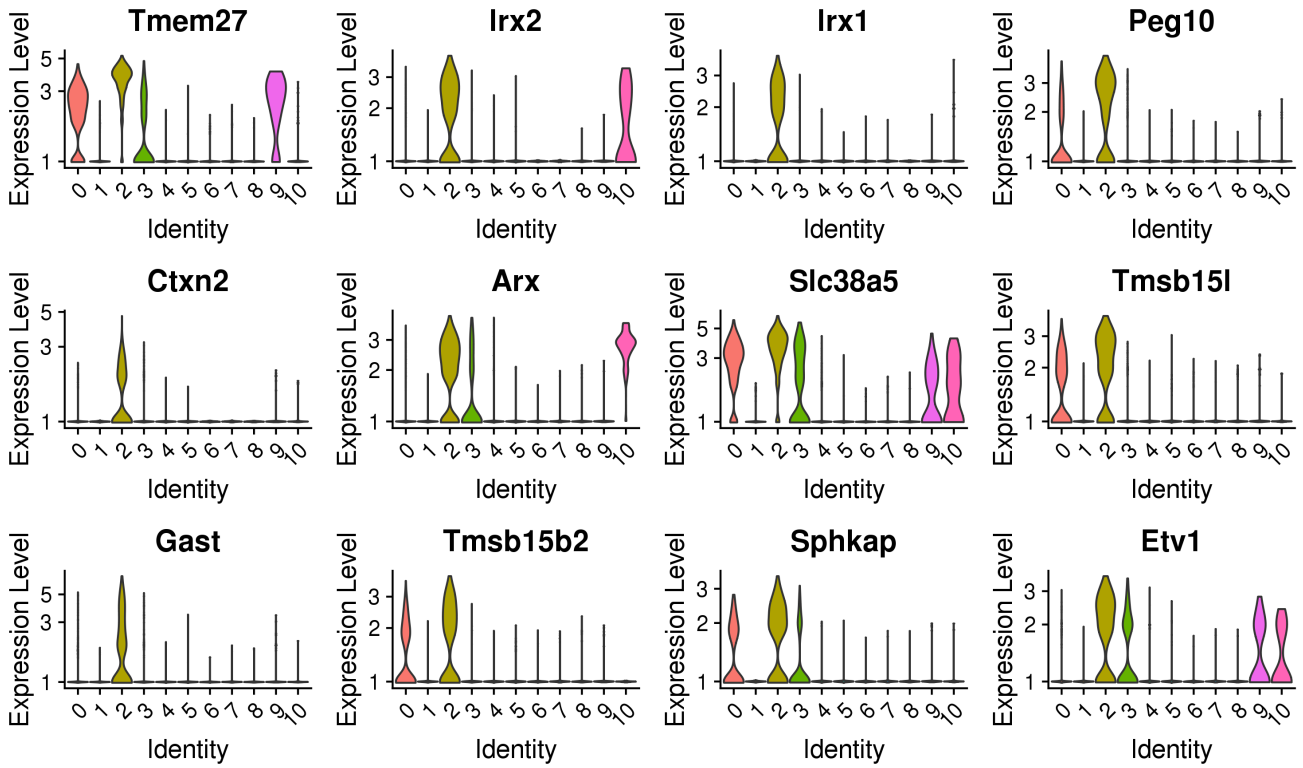
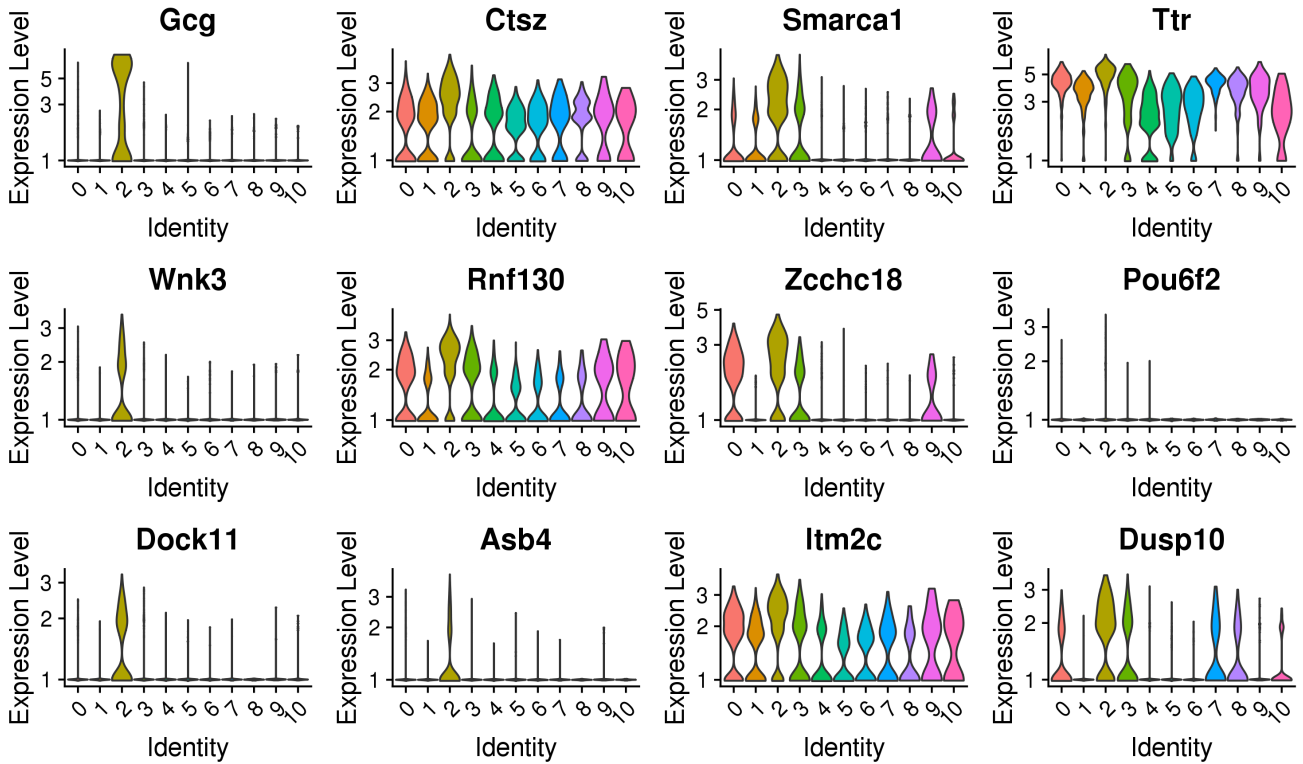


Figure 49: Differential expression summary plots for cluster 2

18.8 Cluster 2 violin plots: positive marker genes

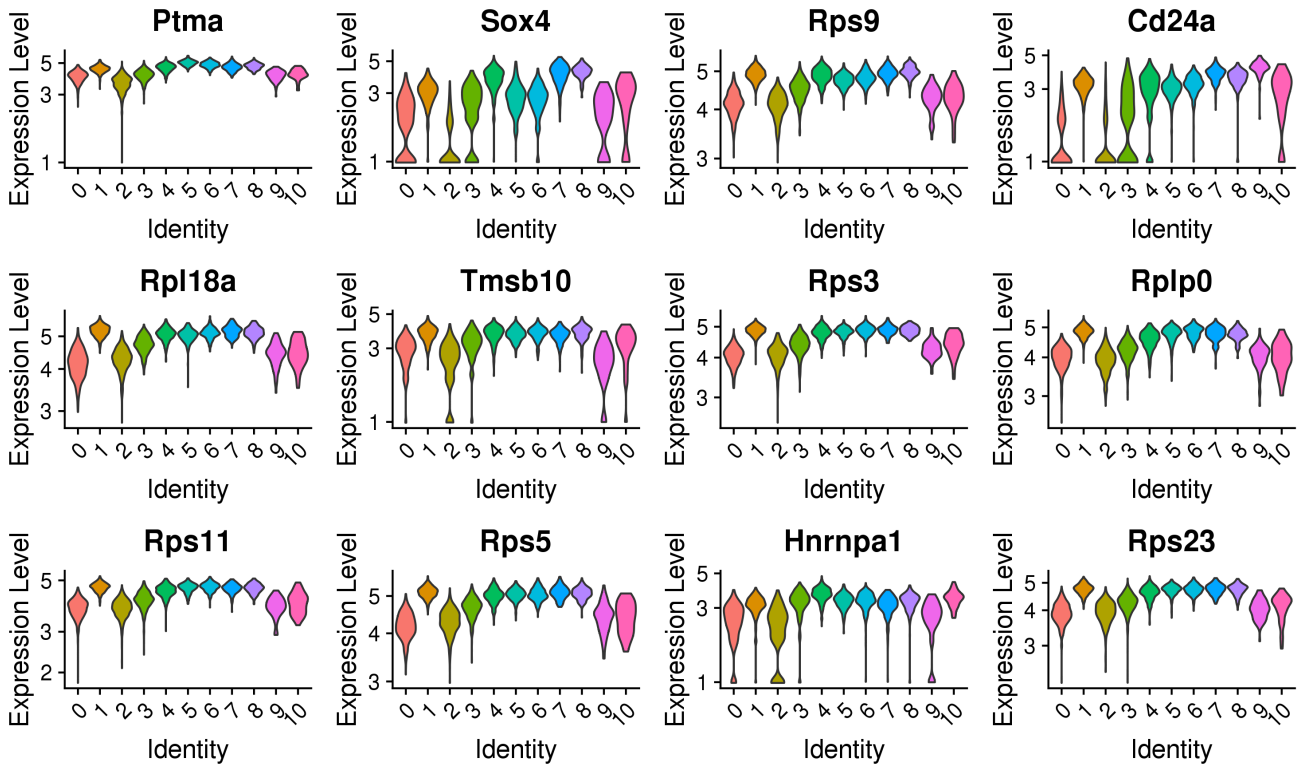


(a) Top positive marker genes ordered by p-value, cluster: 2

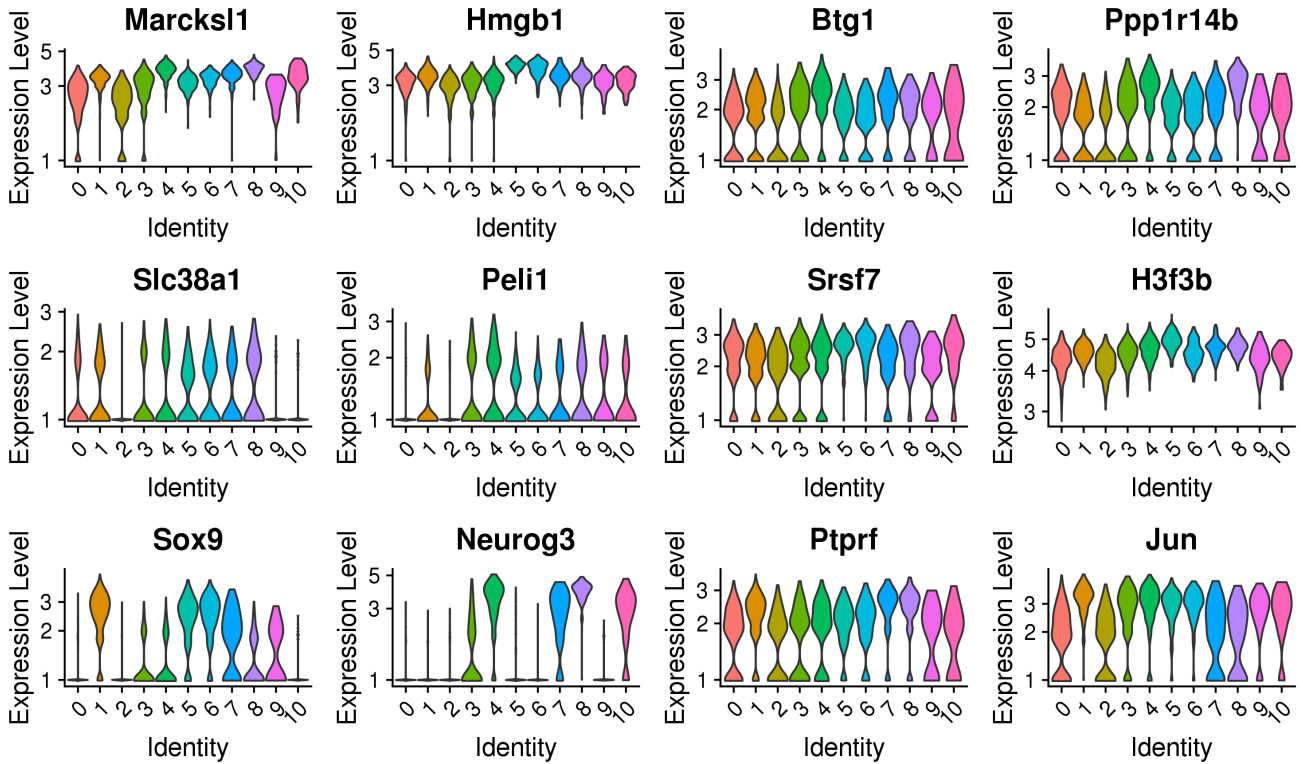


(b) Additional positive marker genes ordered by minimum fold-change vs all other clusters, cluster: 2

18.9 Cluster 2 violin plots: negative marker genes



(a) Top negative marker genes ordered by p-value, cluster: 2



(b) Additional negative marker genes ordered by minimum fold-change vs all other clusters, cluster: 2

18.10 Cluster 3: summary plots

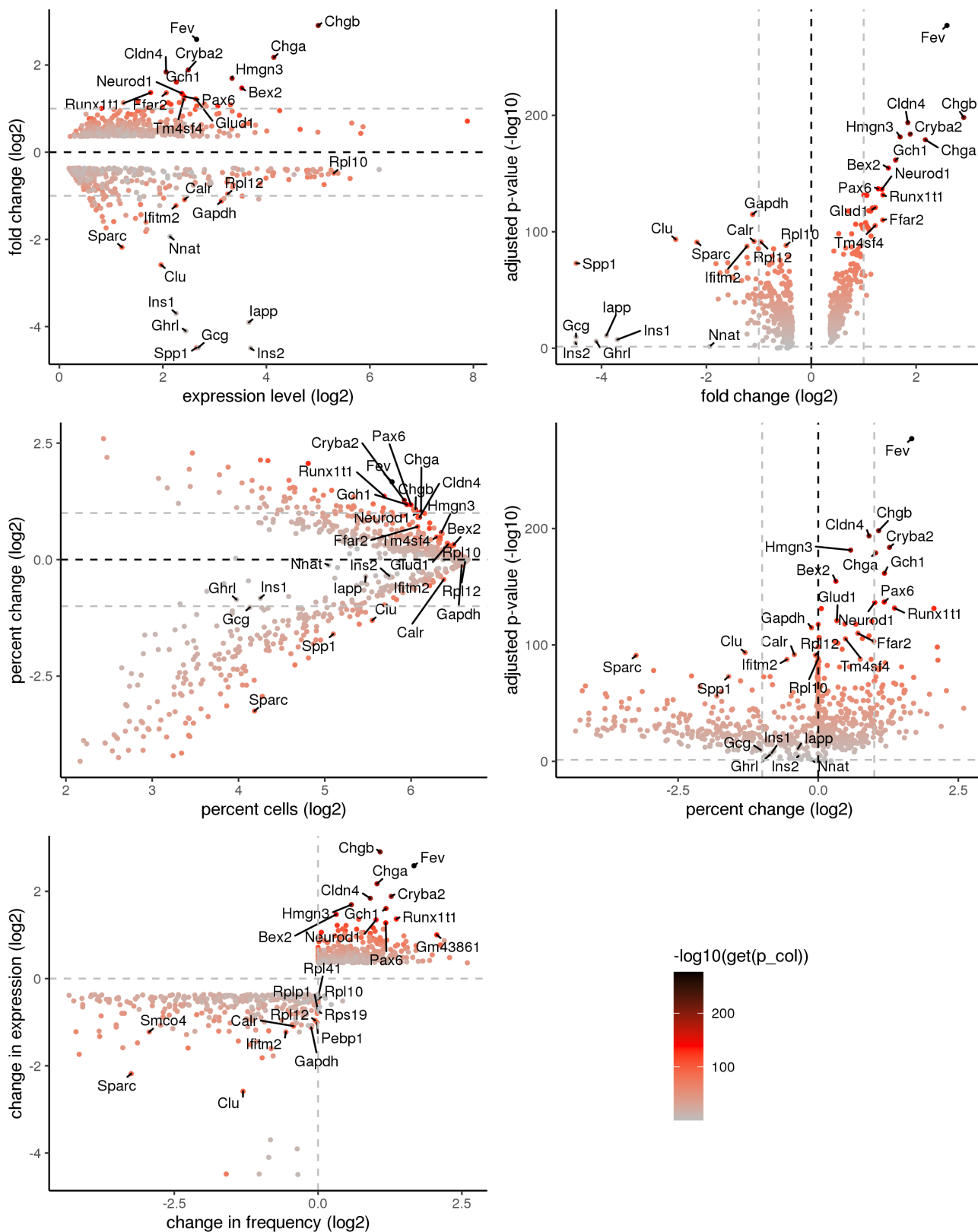
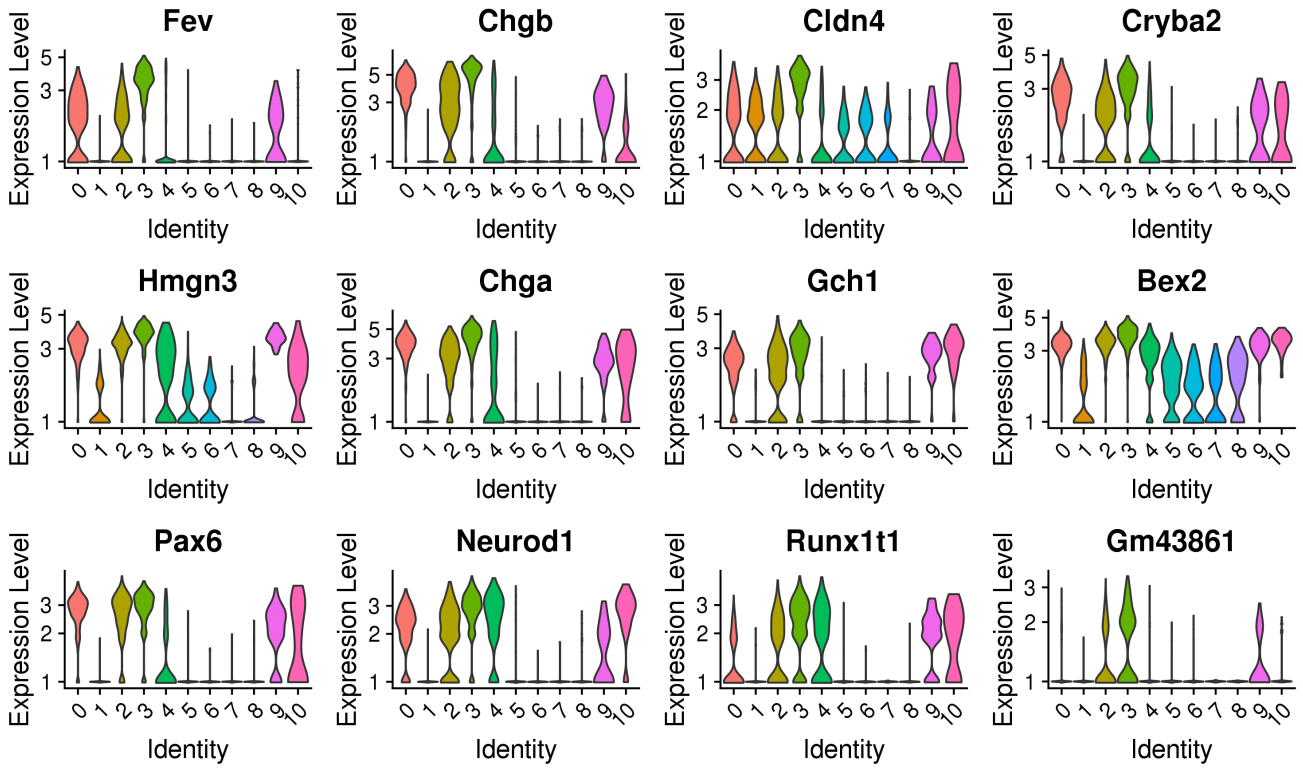
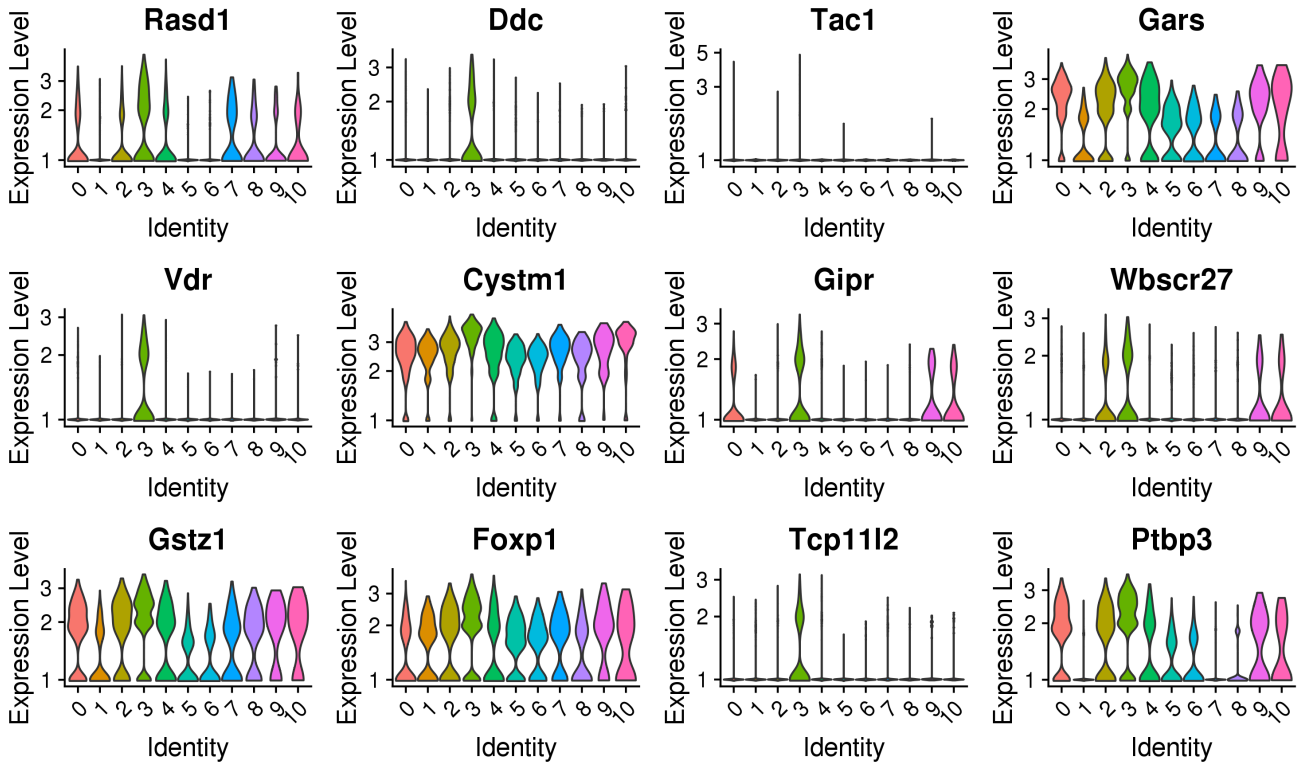


Figure 52: Differential expression summary plots for cluster 3

18.11 Cluster 3 violin plots: positive marker genes

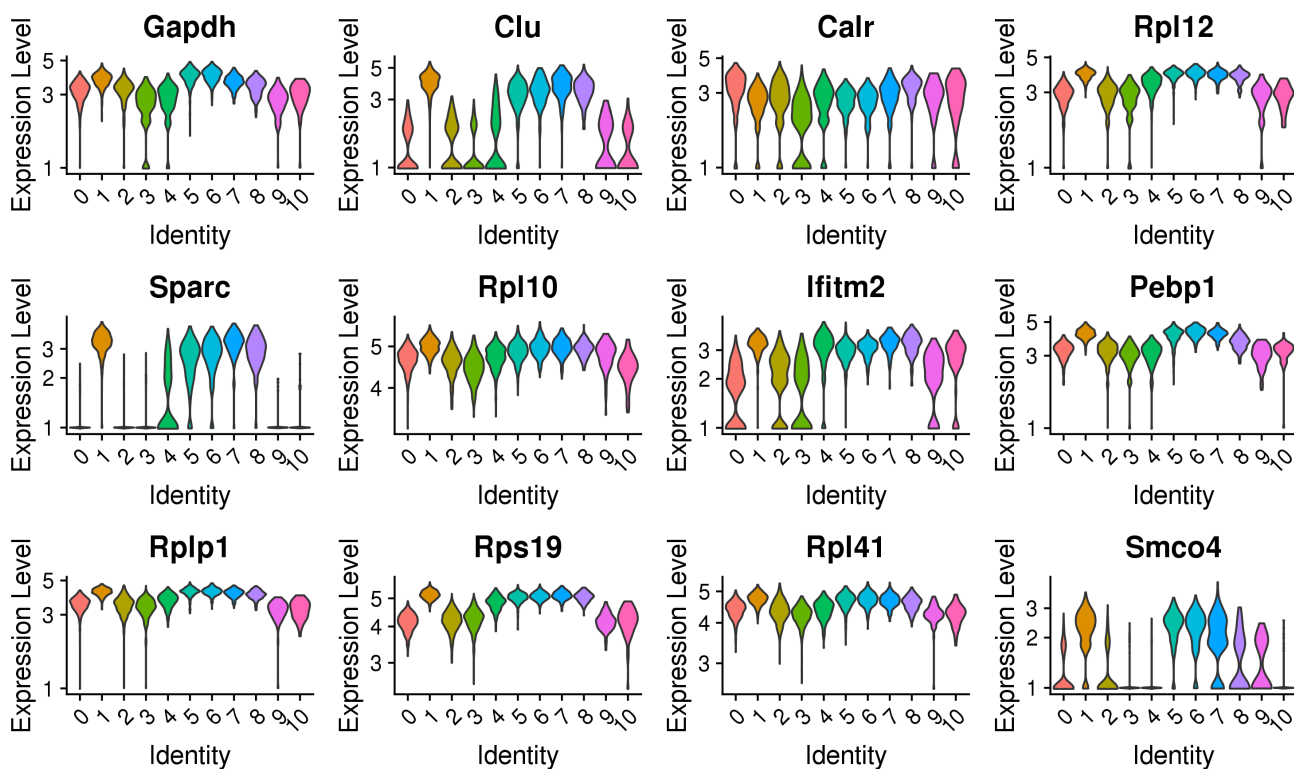


(a) Top positive marker genes ordered by p-value, cluster: 3

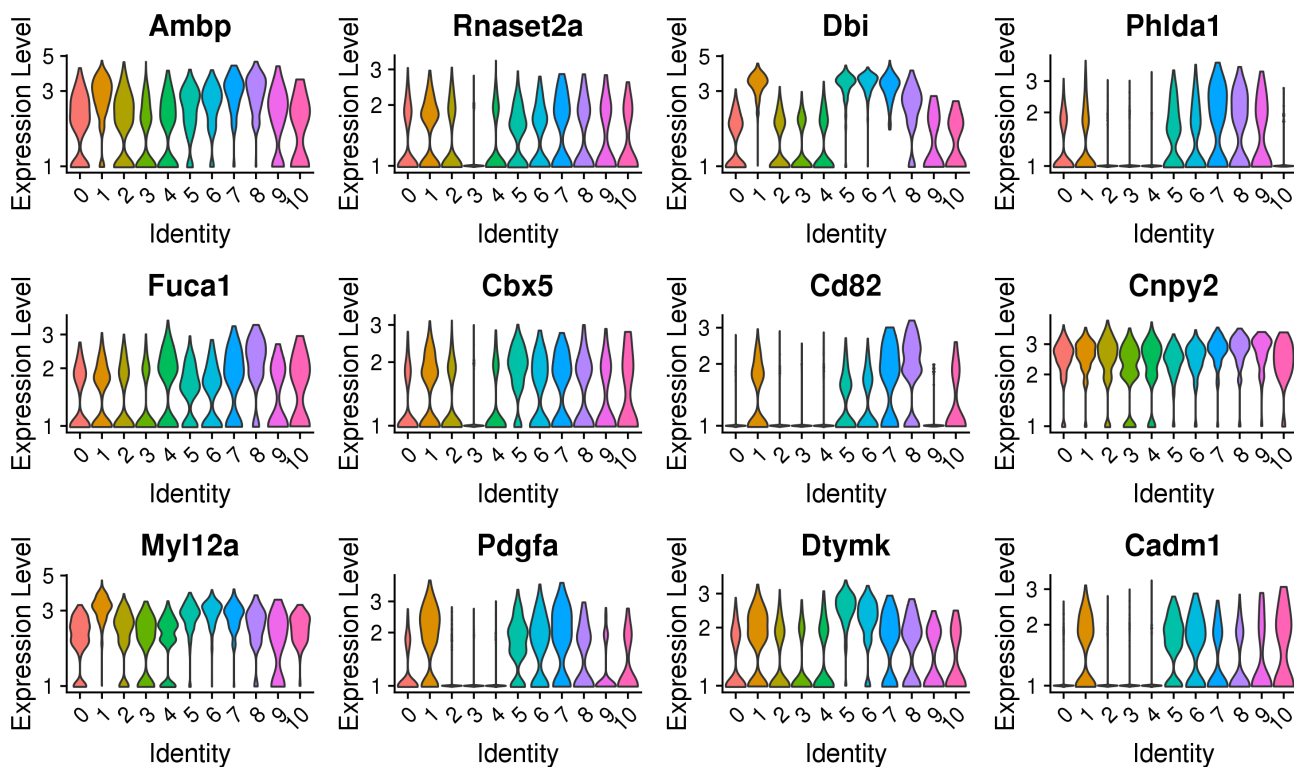


(b) Additional positive marker genes ordered by minimum fold-change vs all other clusters, cluster: 3

18.12 Cluster 3 violin plots: negative marker genes



(a) Top negative marker genes ordered by p-value, cluster: 3



(b) Additional negative marker genes ordered by minimum fold-change vs all other clusters, cluster: 3

18.13 Cluster 4: summary plots

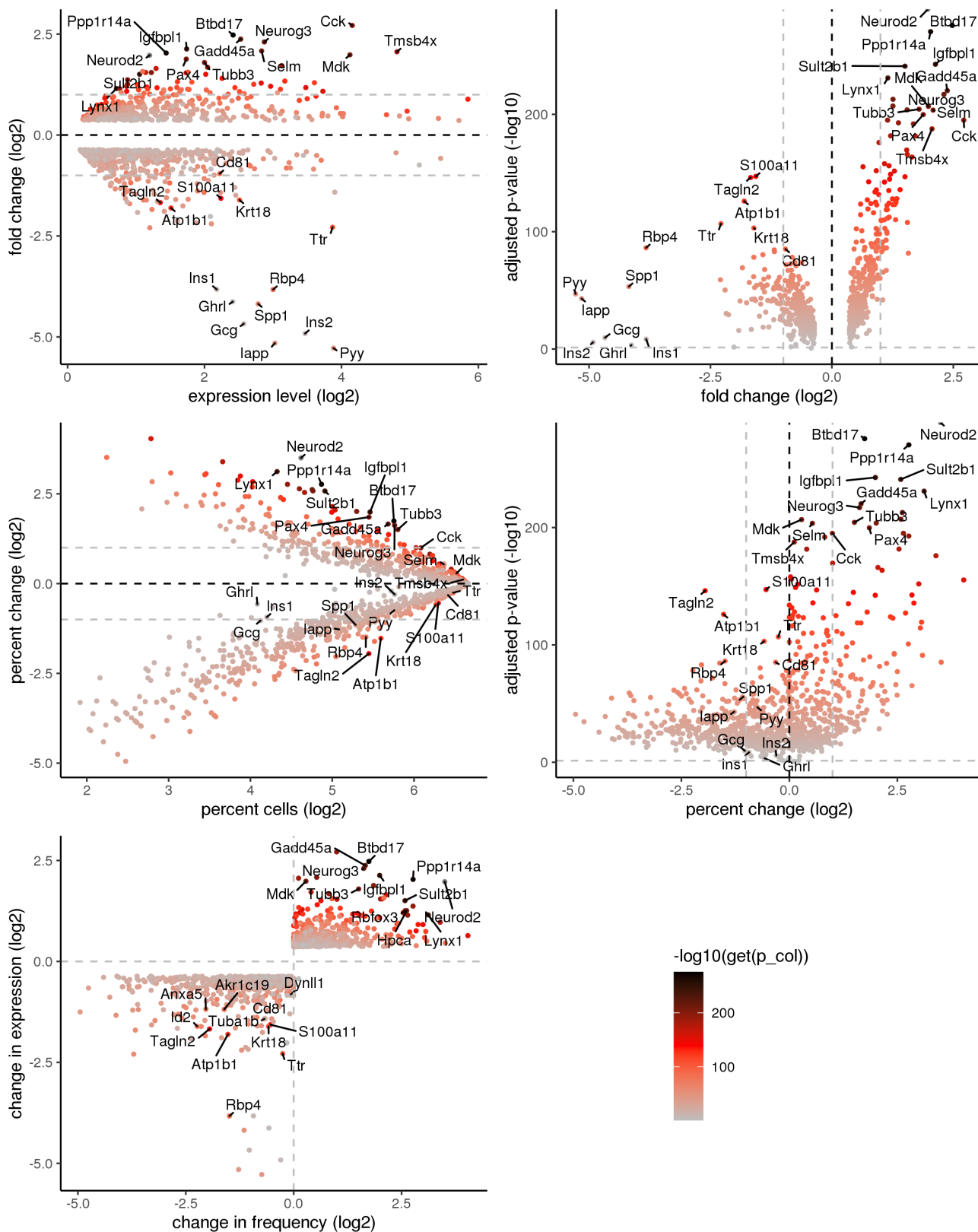
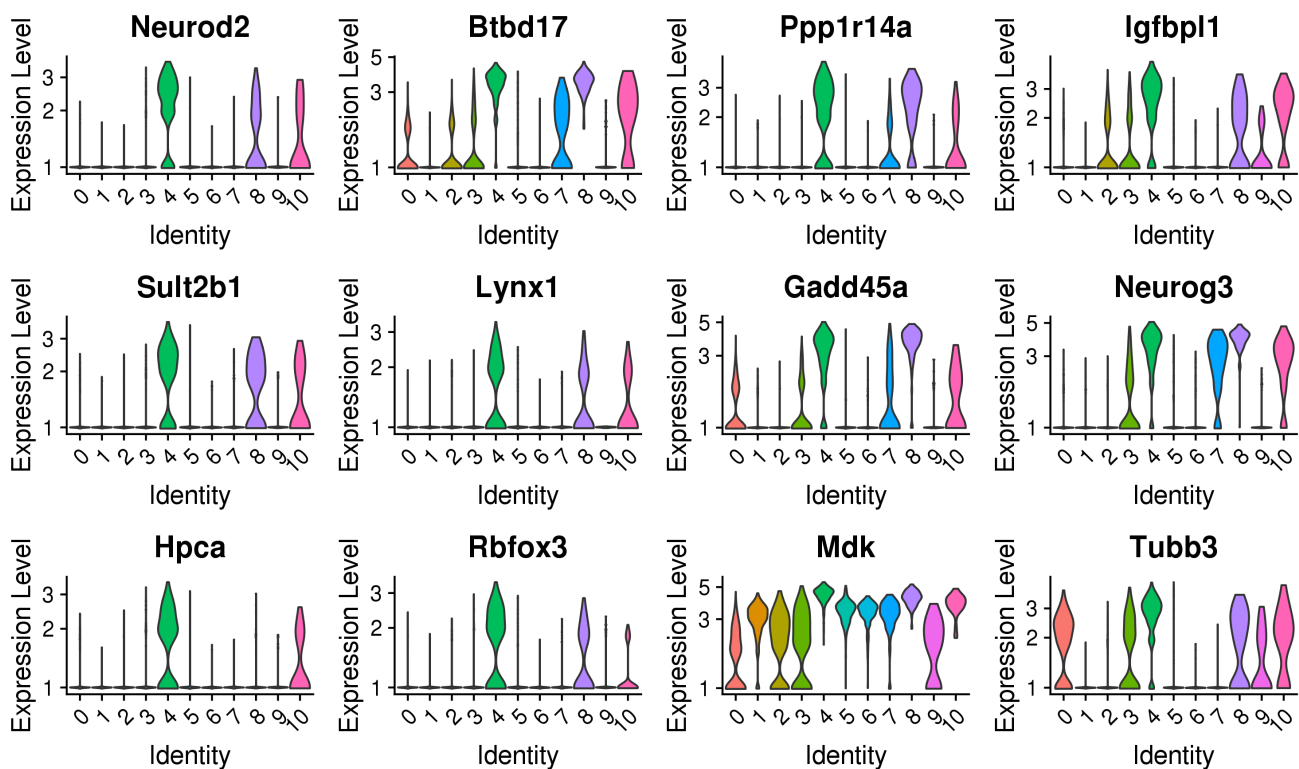
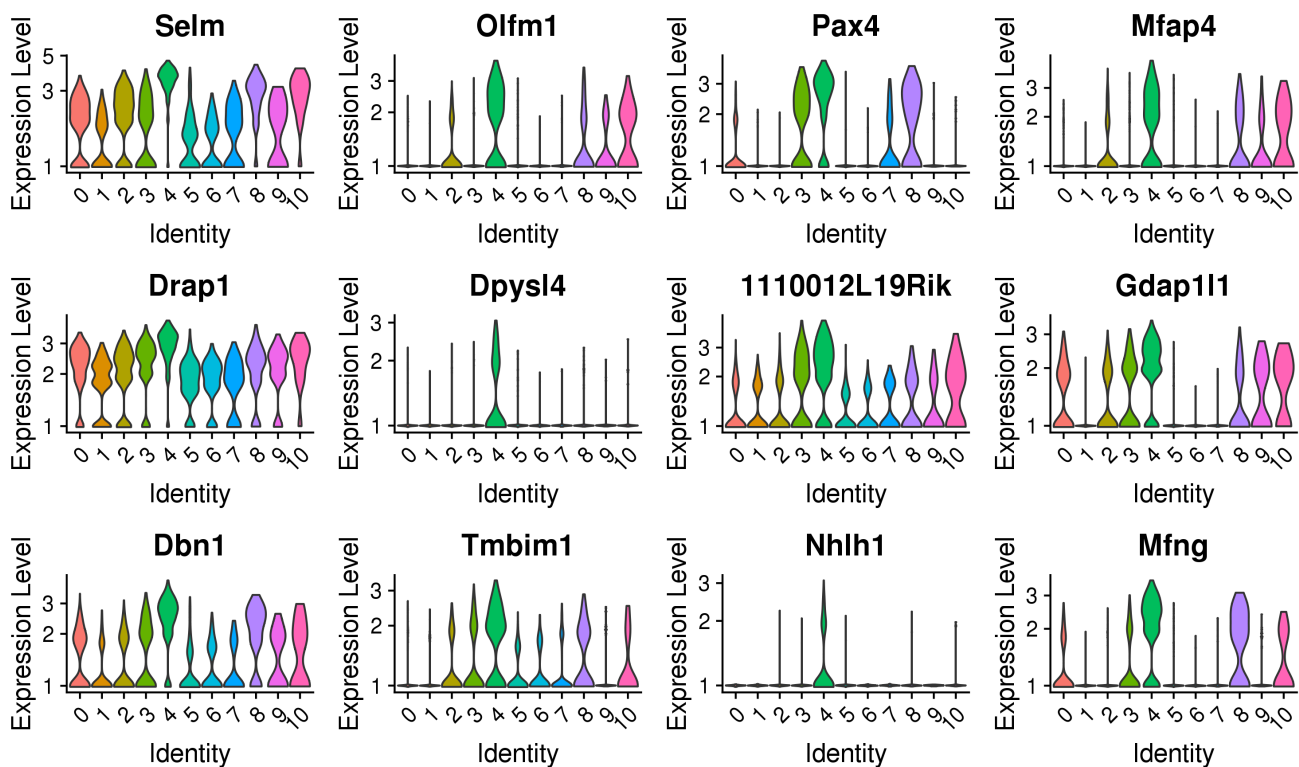


Figure 55: Differential expression summary plots for cluster 4

18.14 Cluster 4 violin plots: positive marker genes

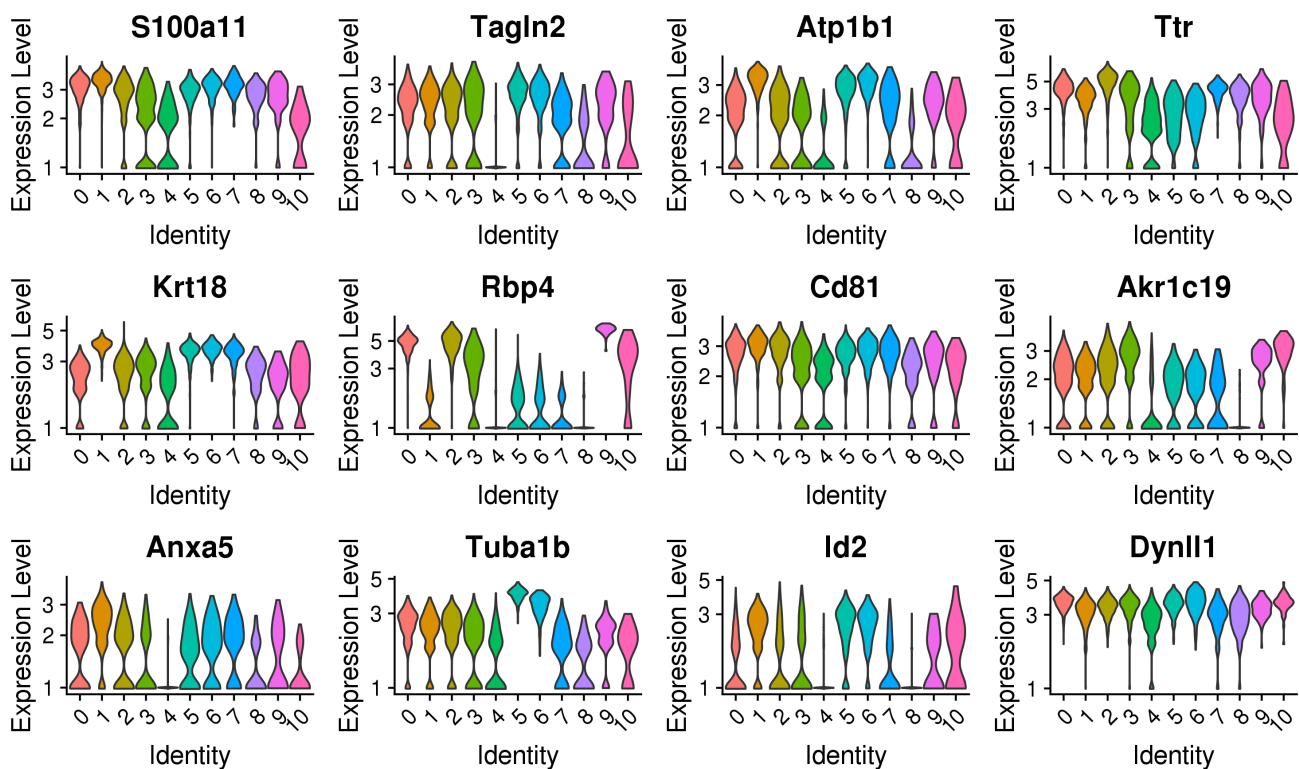


(a) Top positive marker genes ordered by p-value, cluster: 4

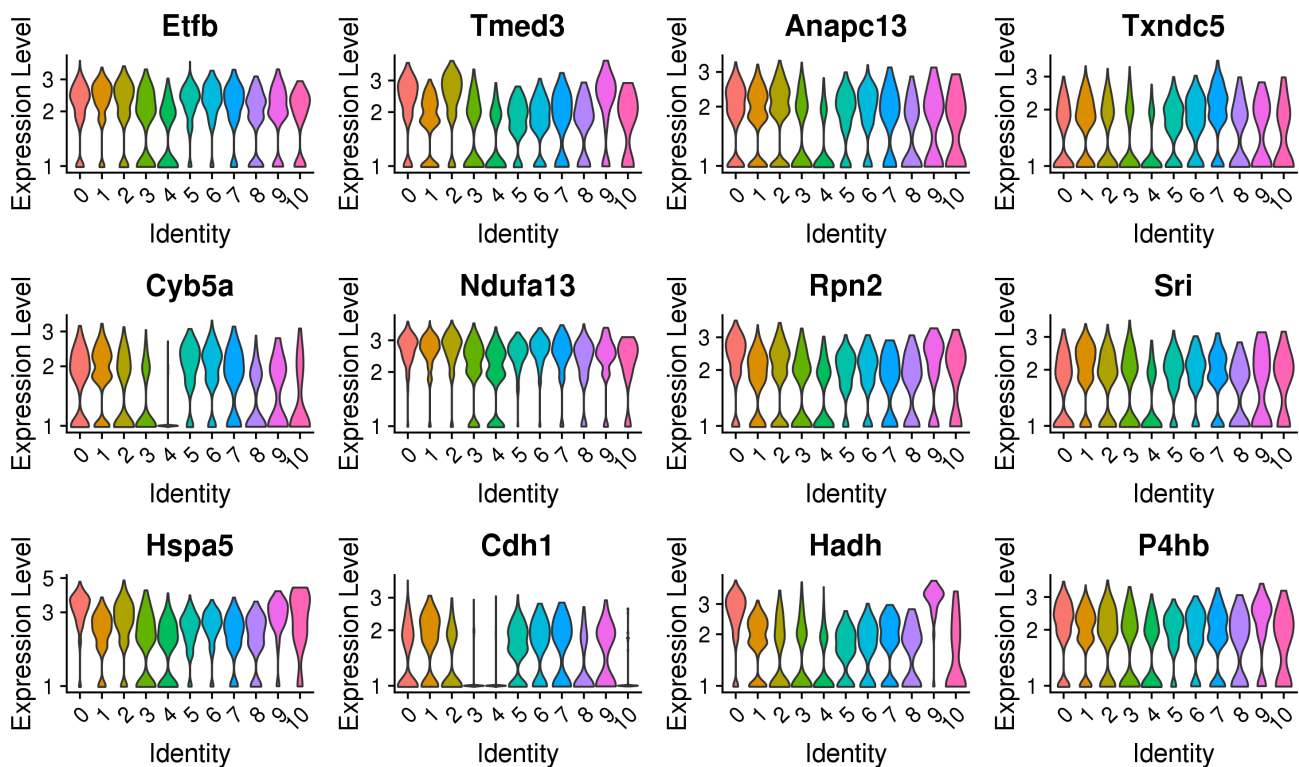


(b) Additional positive marker genes ordered by minimum fold-change vs all other clusters, cluster: 4

18.15 Cluster 4 violin plots: negative marker genes



(a) Top negative marker genes ordered by p-value, cluster: 4



(b) Additional negative marker genes ordered by minimum fold-change vs all other clusters, cluster: 4

18.16 Cluster 5: summary plots

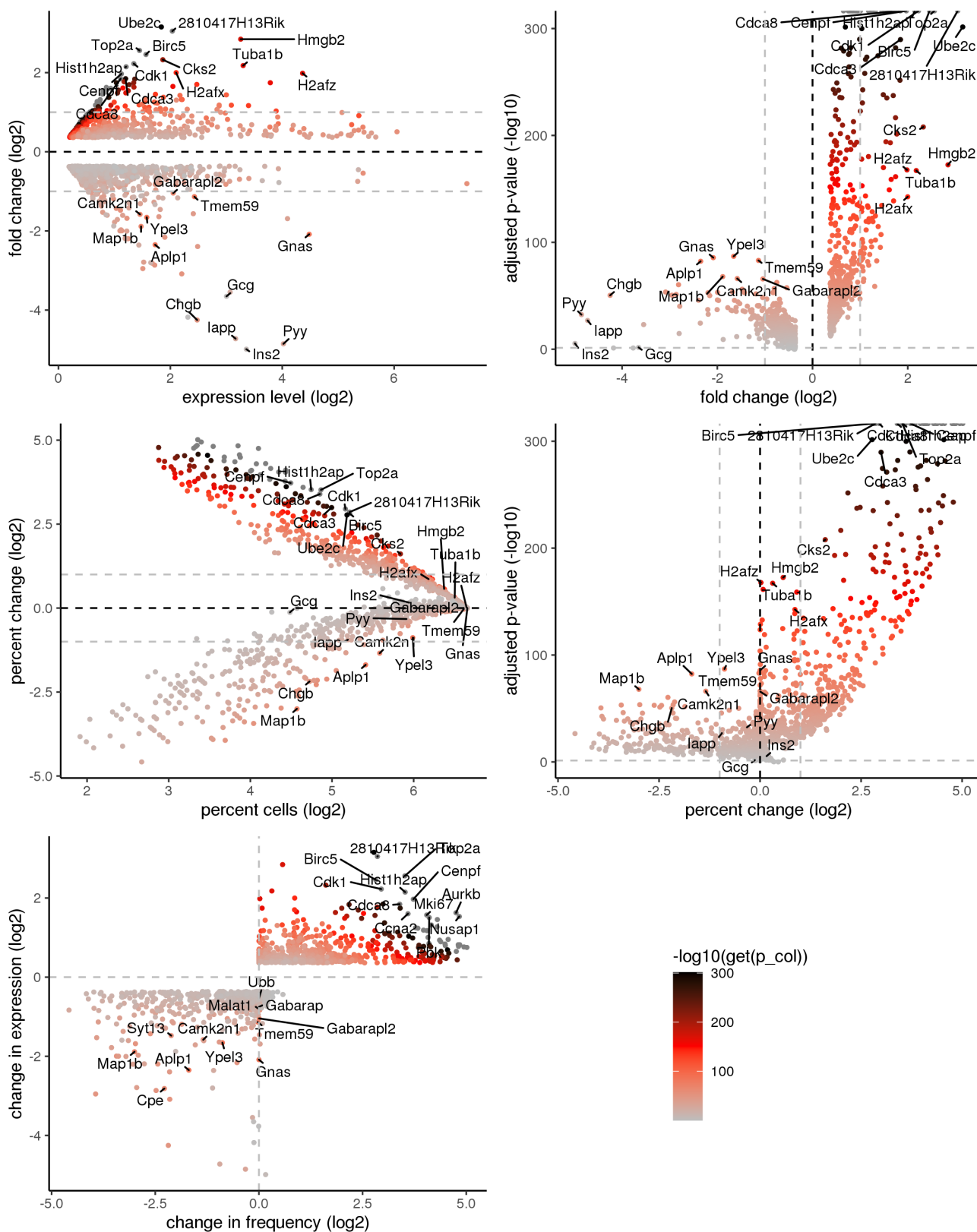
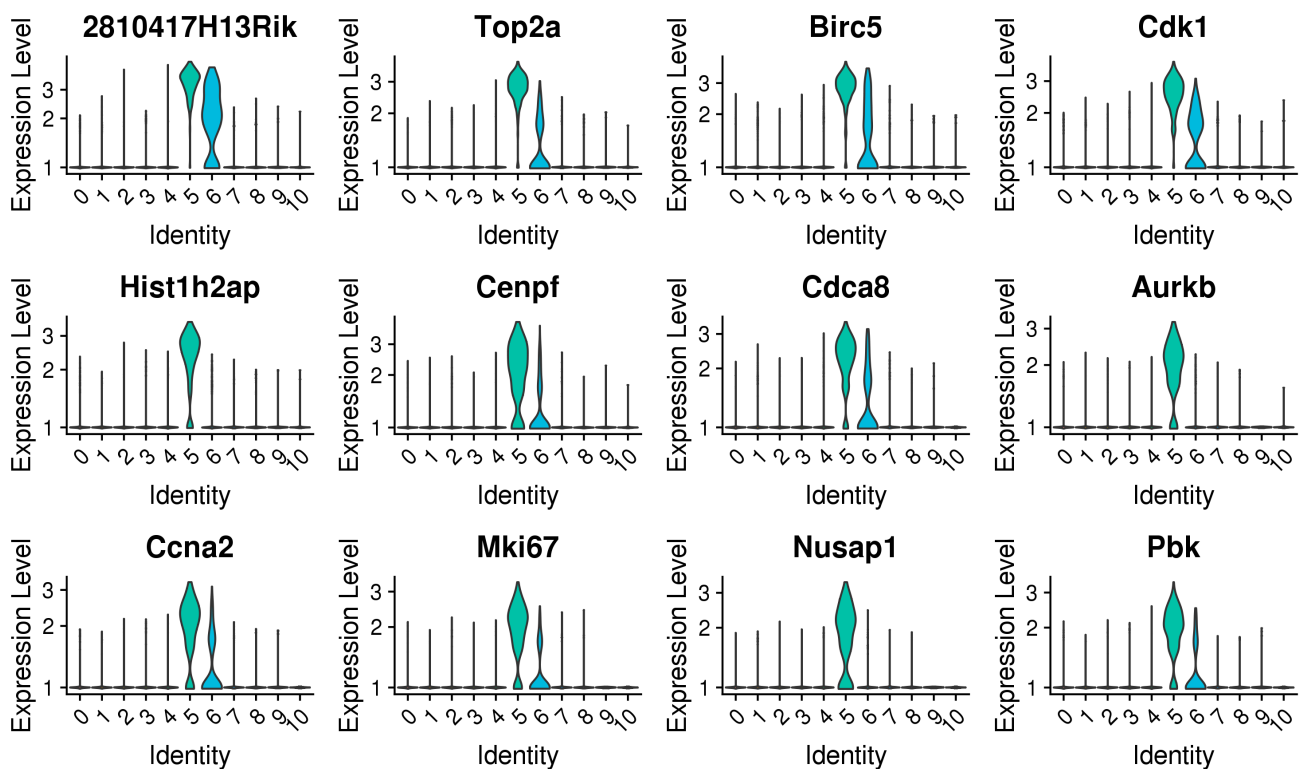
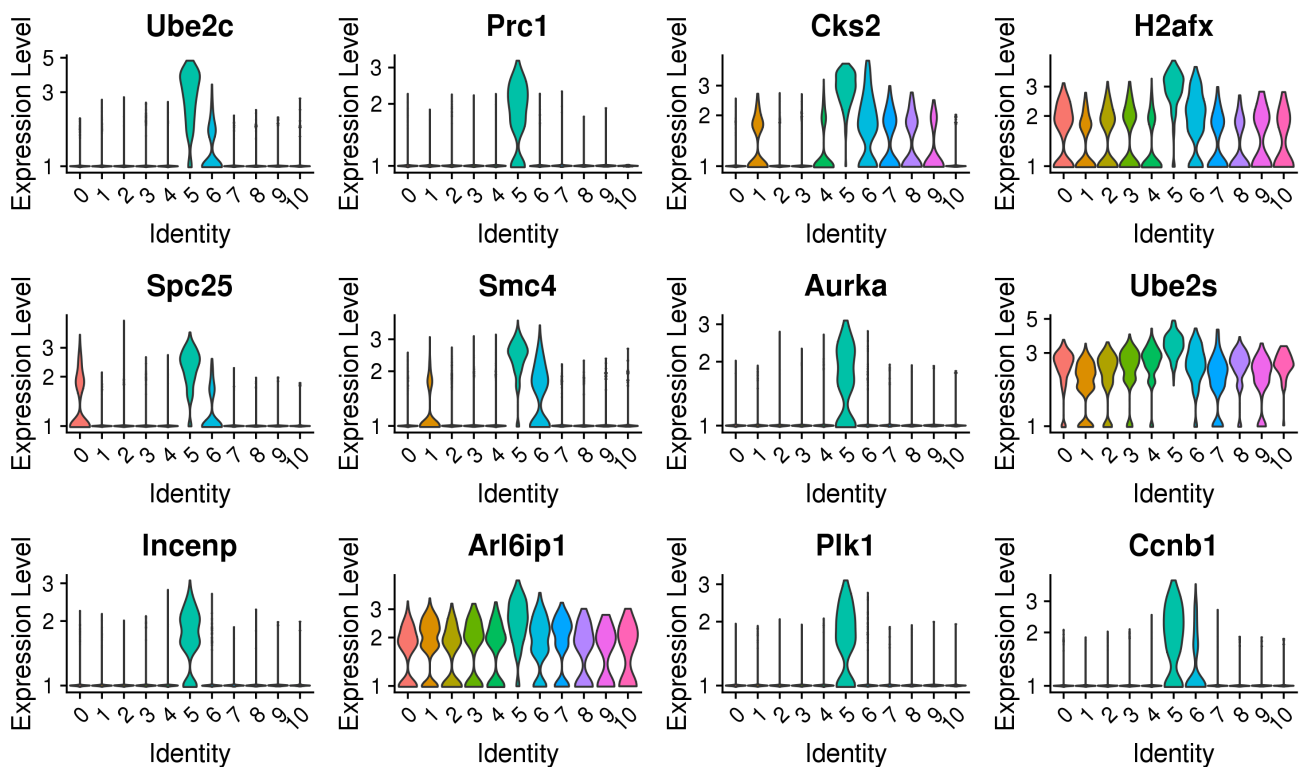


Figure 58: Differential expression summary plots for cluster 5

18.17 Cluster 5 violin plots: positive marker genes

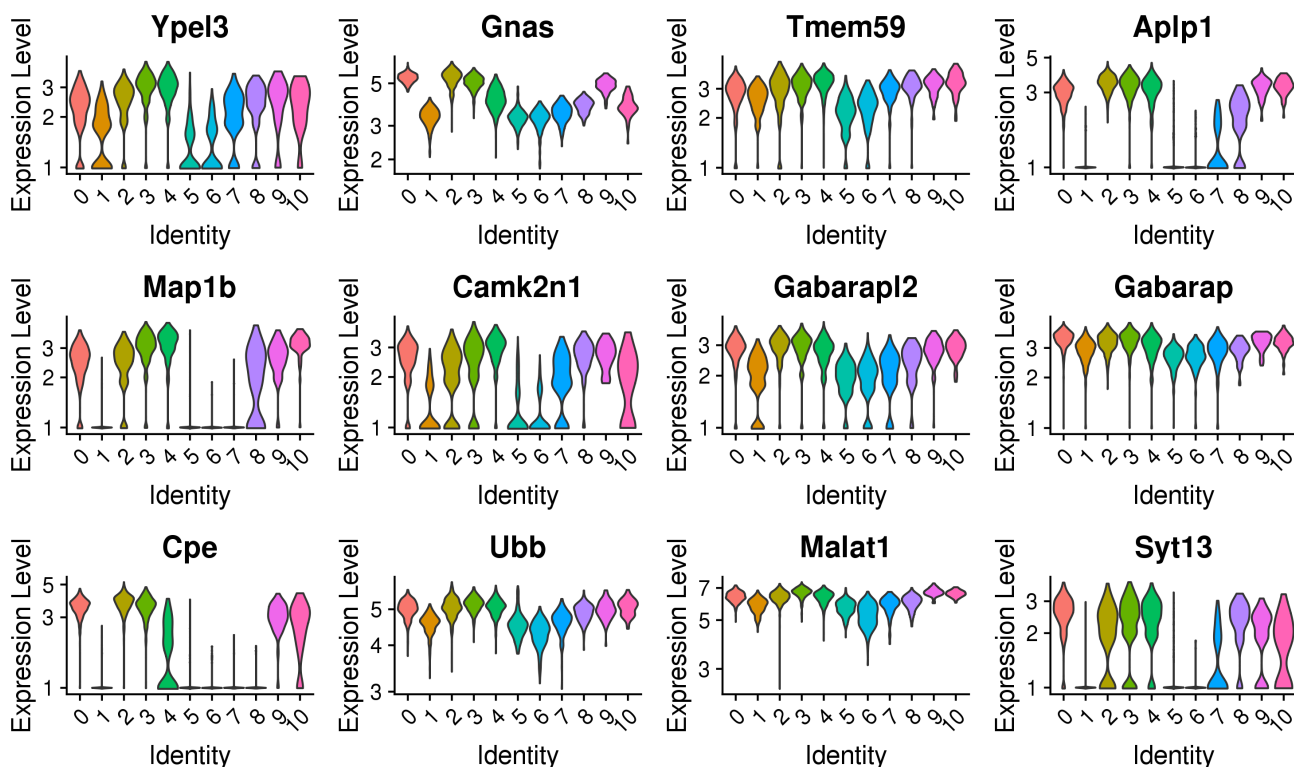


(a) Top positive marker genes ordered by p-value, cluster: 5

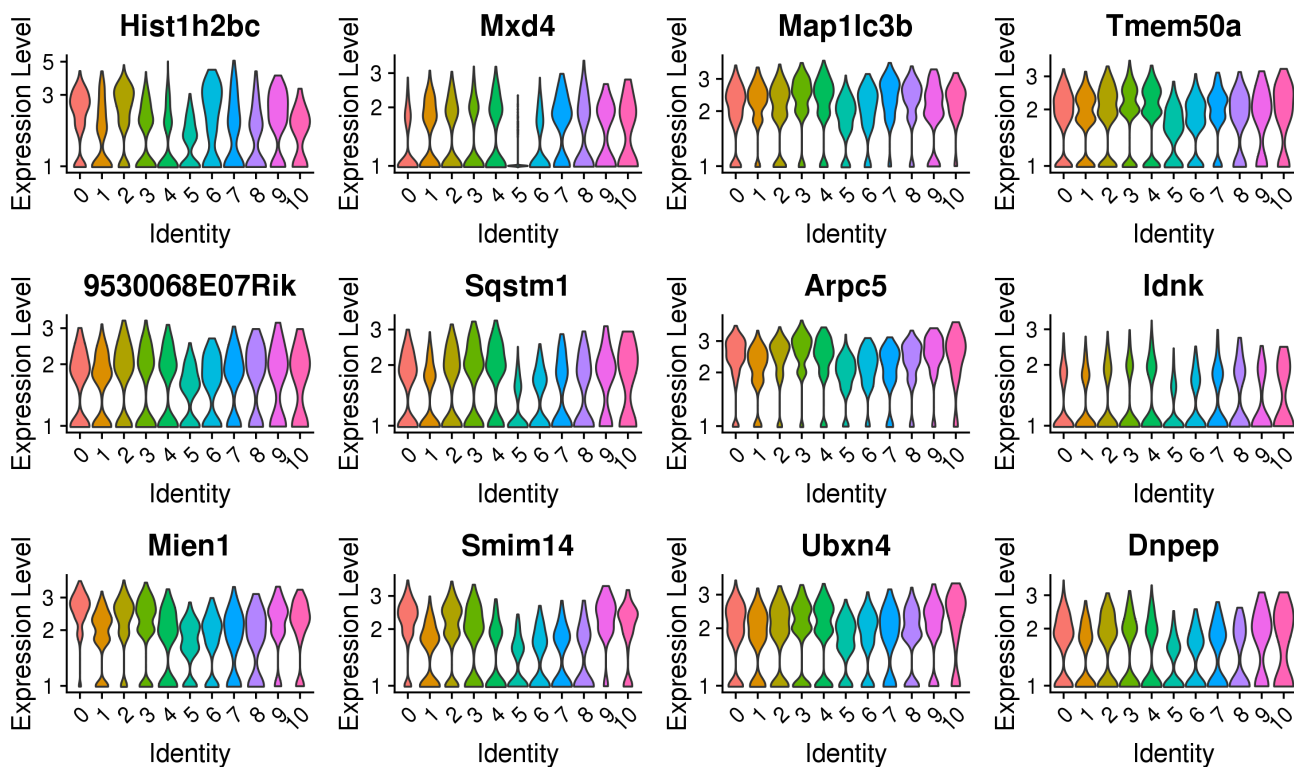


(b) Additional positive marker genes ordered by minimum fold-change vs all other clusters, cluster: 5

18.18 Cluster 5 violin plots: negative marker genes



(a) Top negative marker genes ordered by p-value, cluster: 5



(b) Additional negative marker genes ordered by minimum fold-change vs all other clusters, cluster: 5

18.19 Cluster 6: summary plots

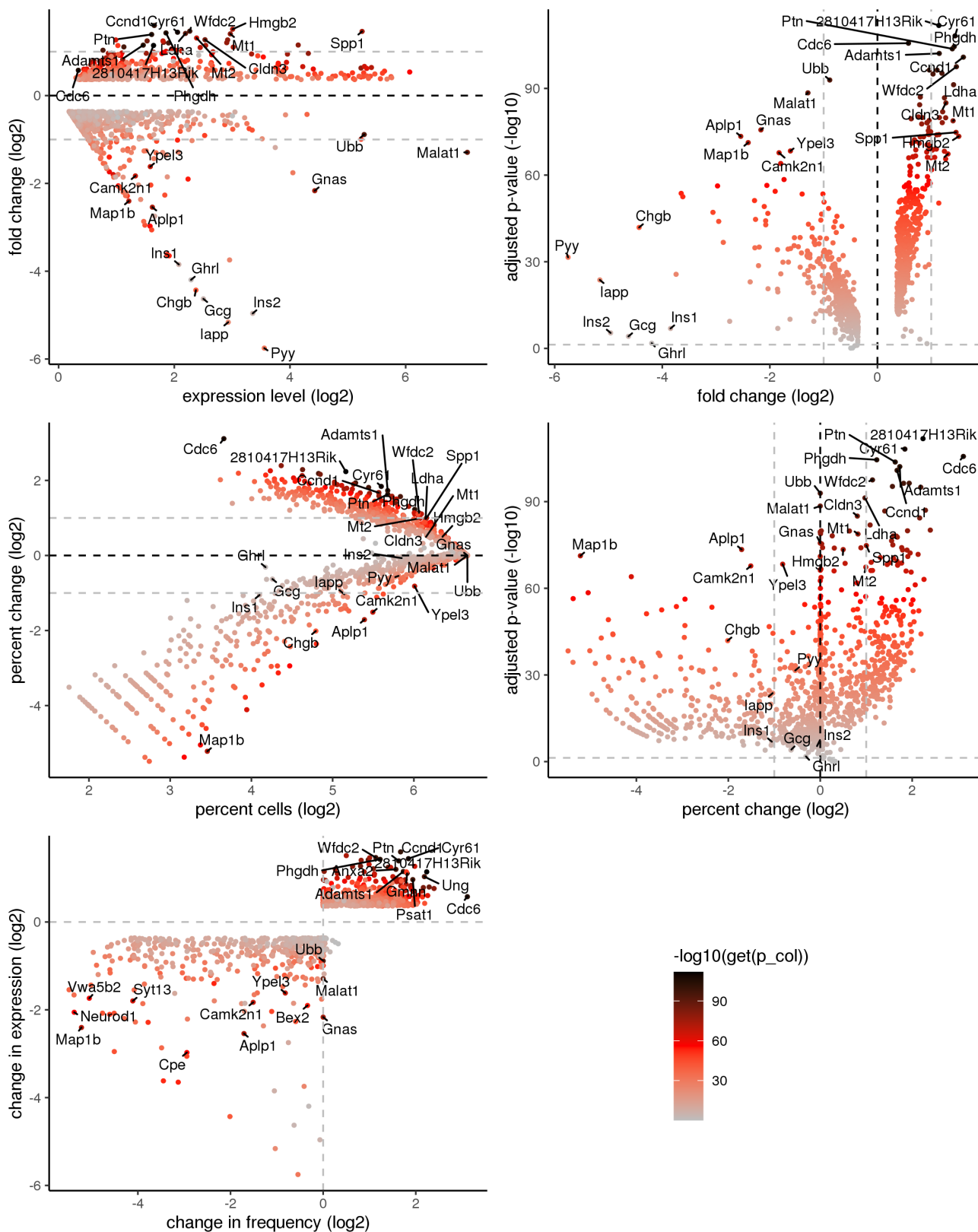
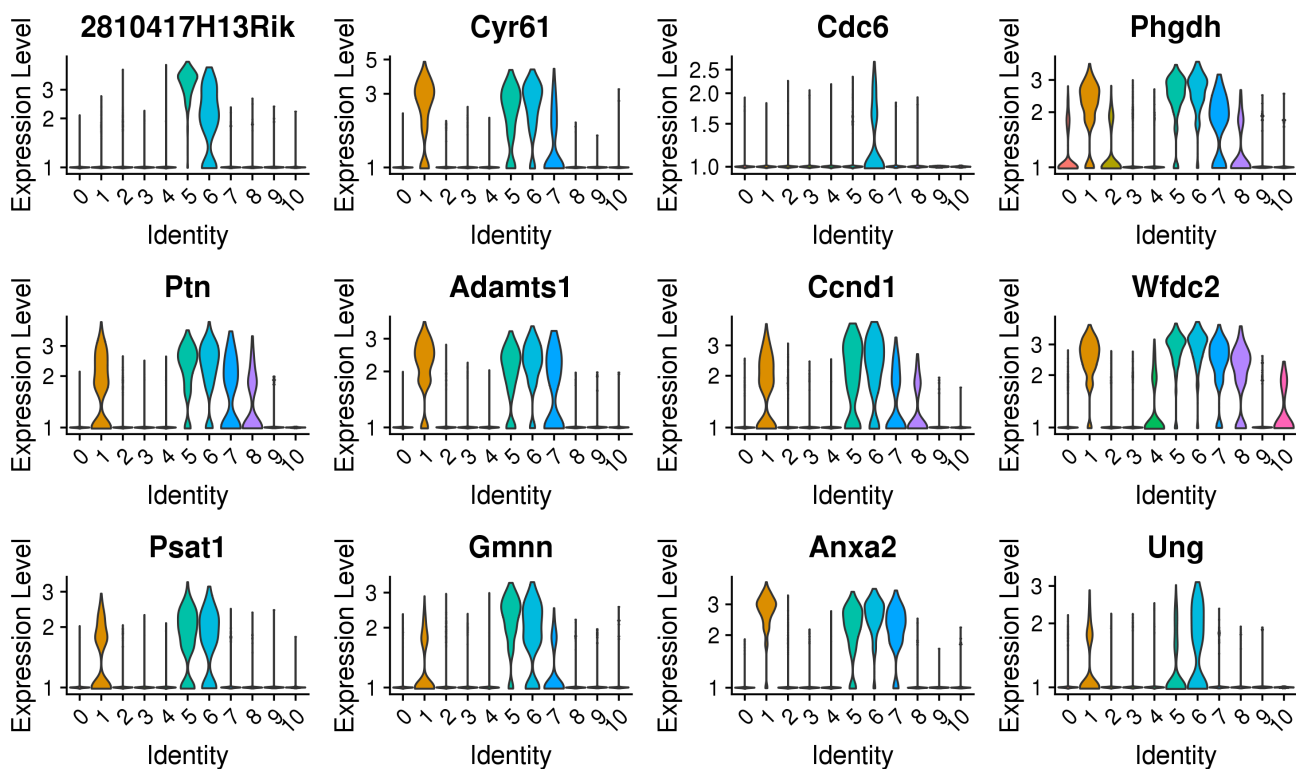
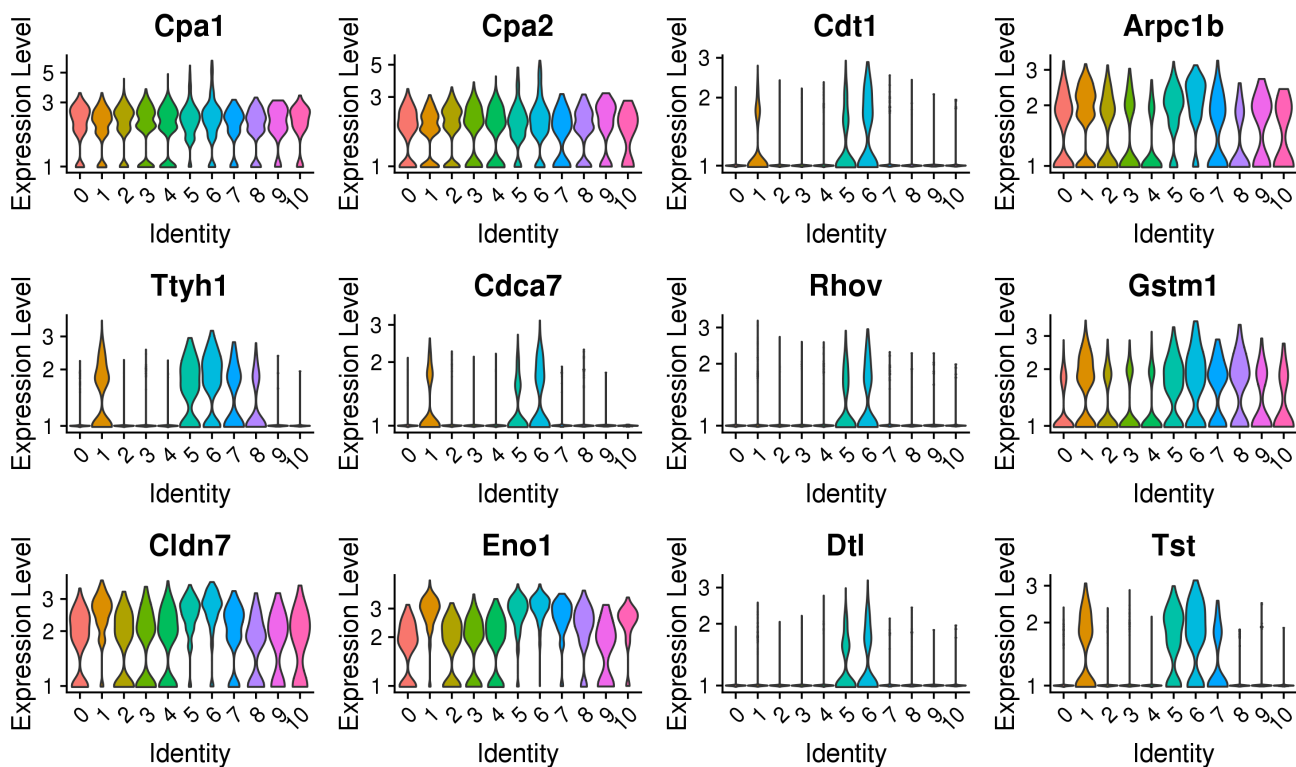


Figure 61: Differential expression summary plots for cluster 6

18.20 Cluster 6 violin plots: positive marker genes

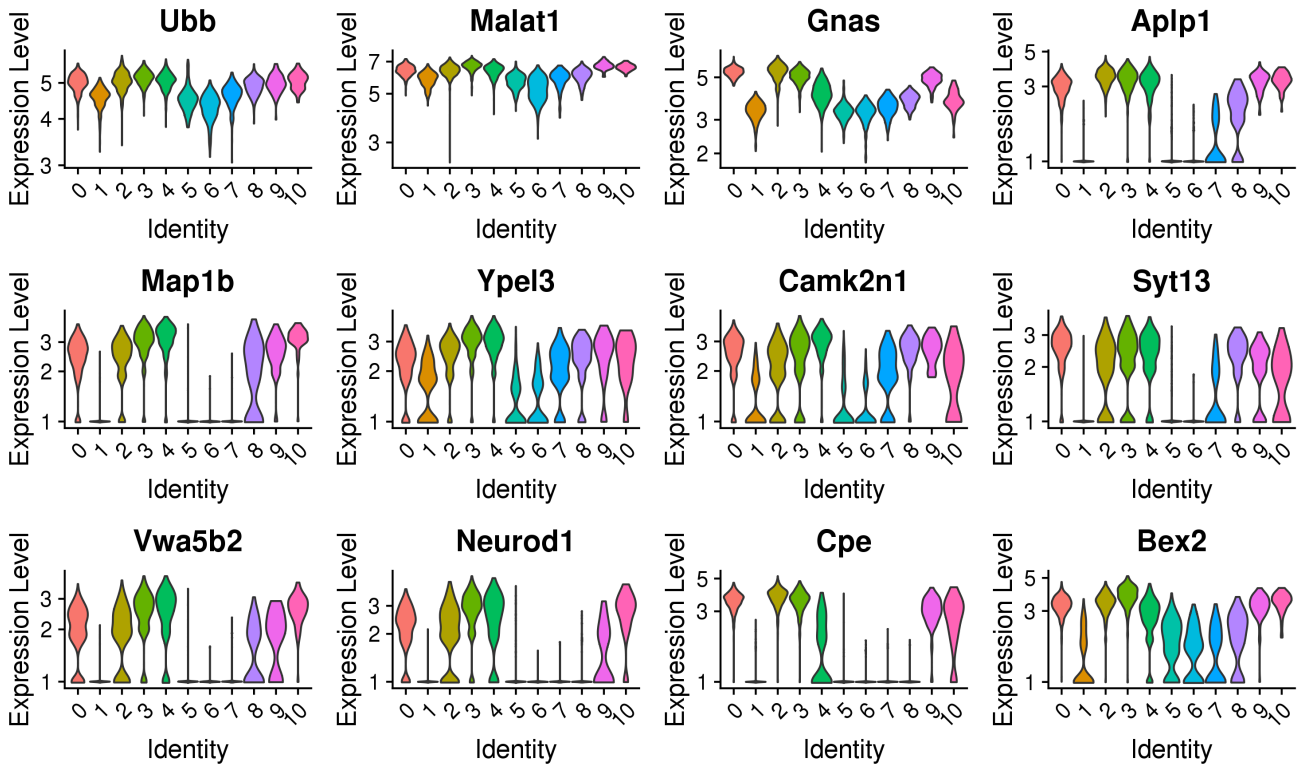


(a) Top positive marker genes ordered by p-value, cluster: 6

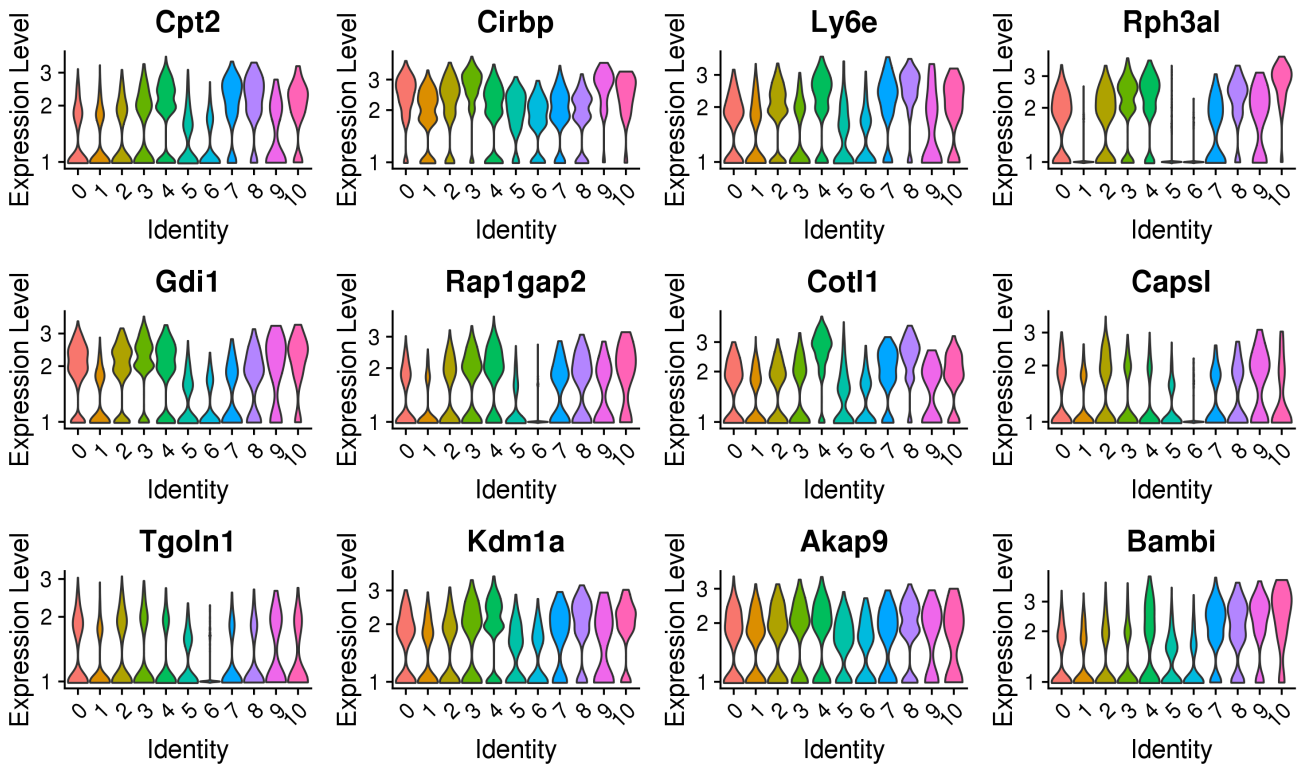


(b) Additional positive marker genes ordered by minimum fold-change vs all other clusters, cluster: 6

18.21 Cluster 6 violin plots: negative marker genes



(a) Top negative marker genes ordered by p-value, cluster: 6



(b) Additional negative marker genes ordered by minimum fold-change vs all other clusters, cluster: 6

18.22 Cluster 7: summary plots

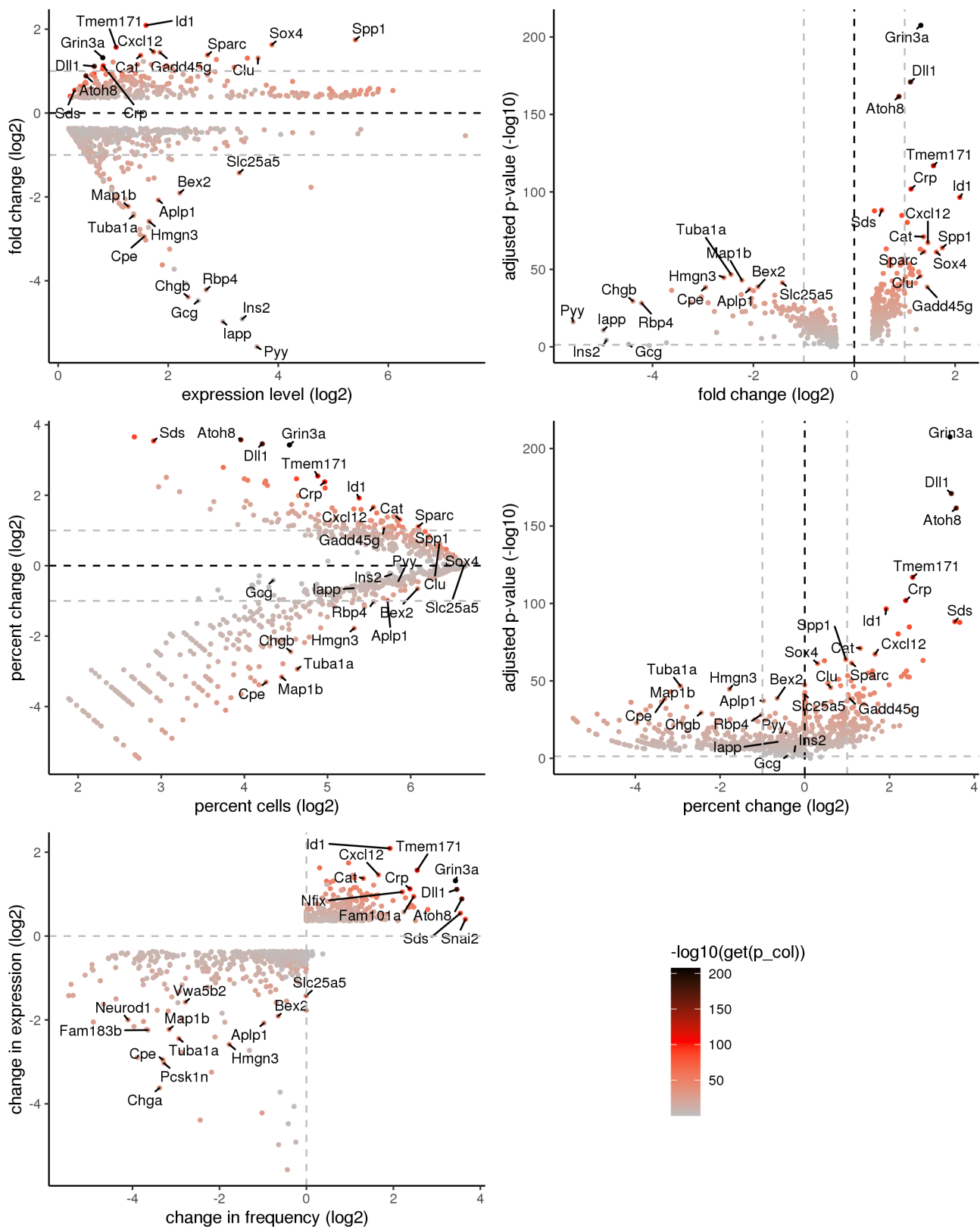
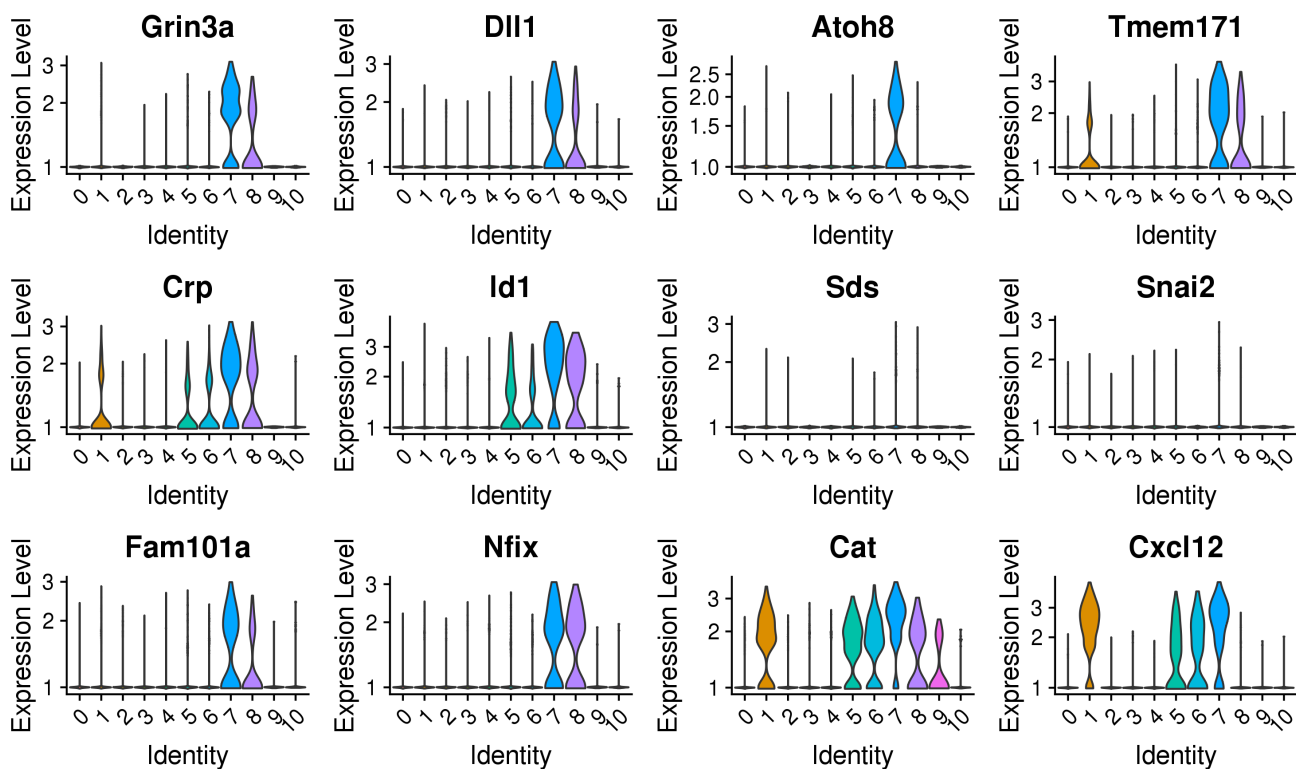
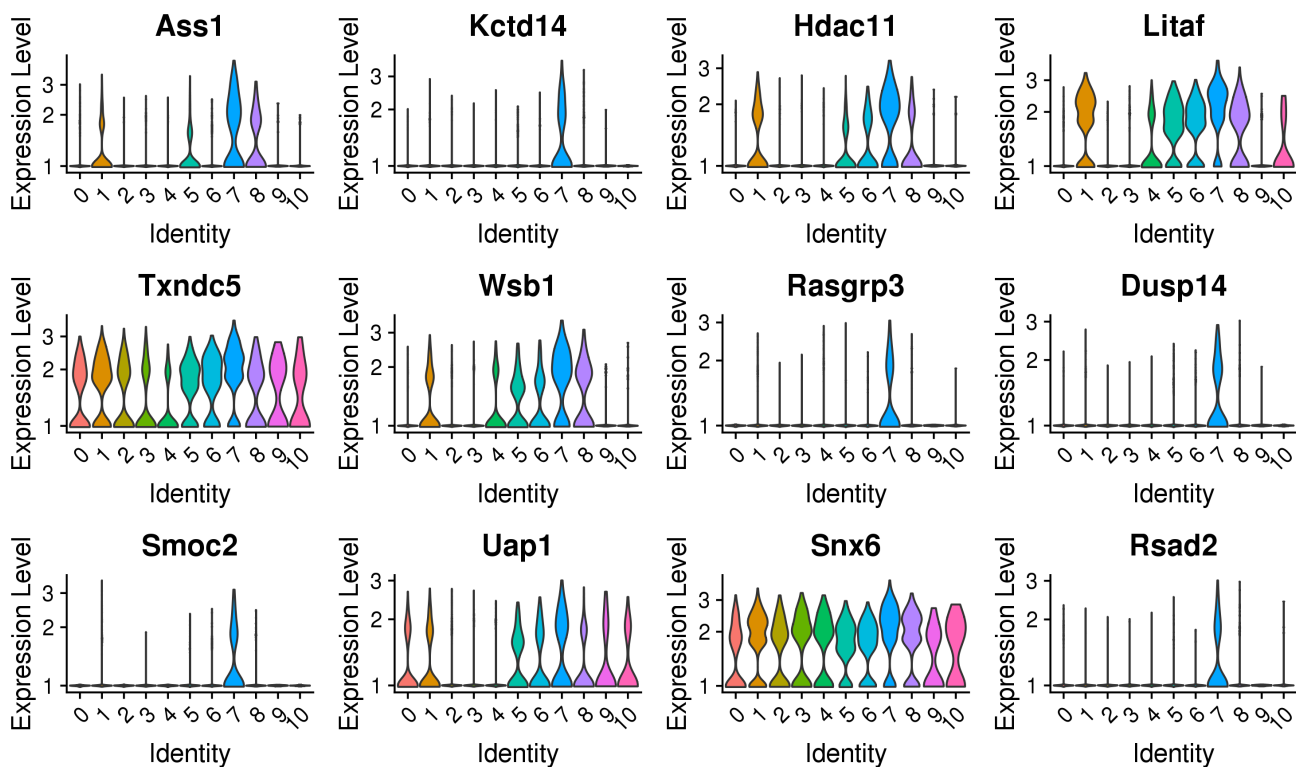


Figure 64: Differential expression summary plots for cluster 7

18.23 Cluster 7 violin plots: positive marker genes

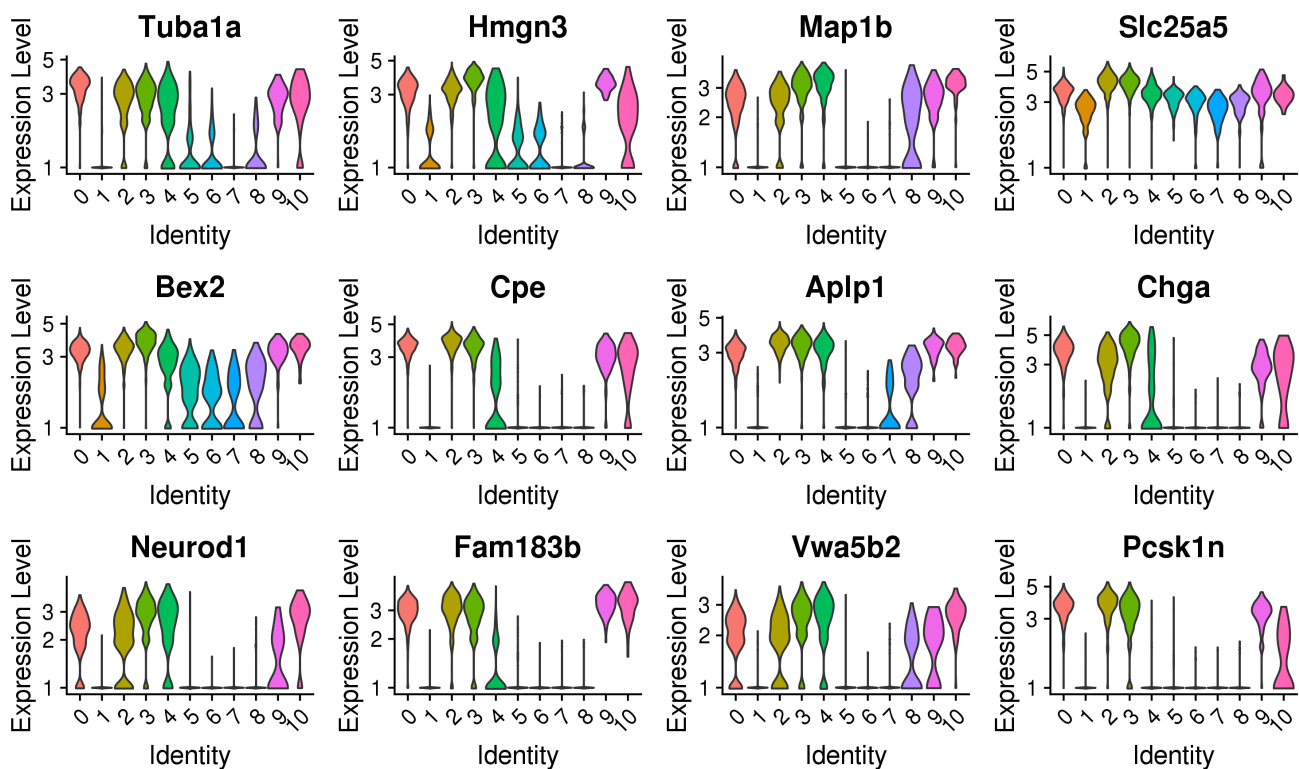


(a) Top positive marker genes ordered by p-value, cluster: 7

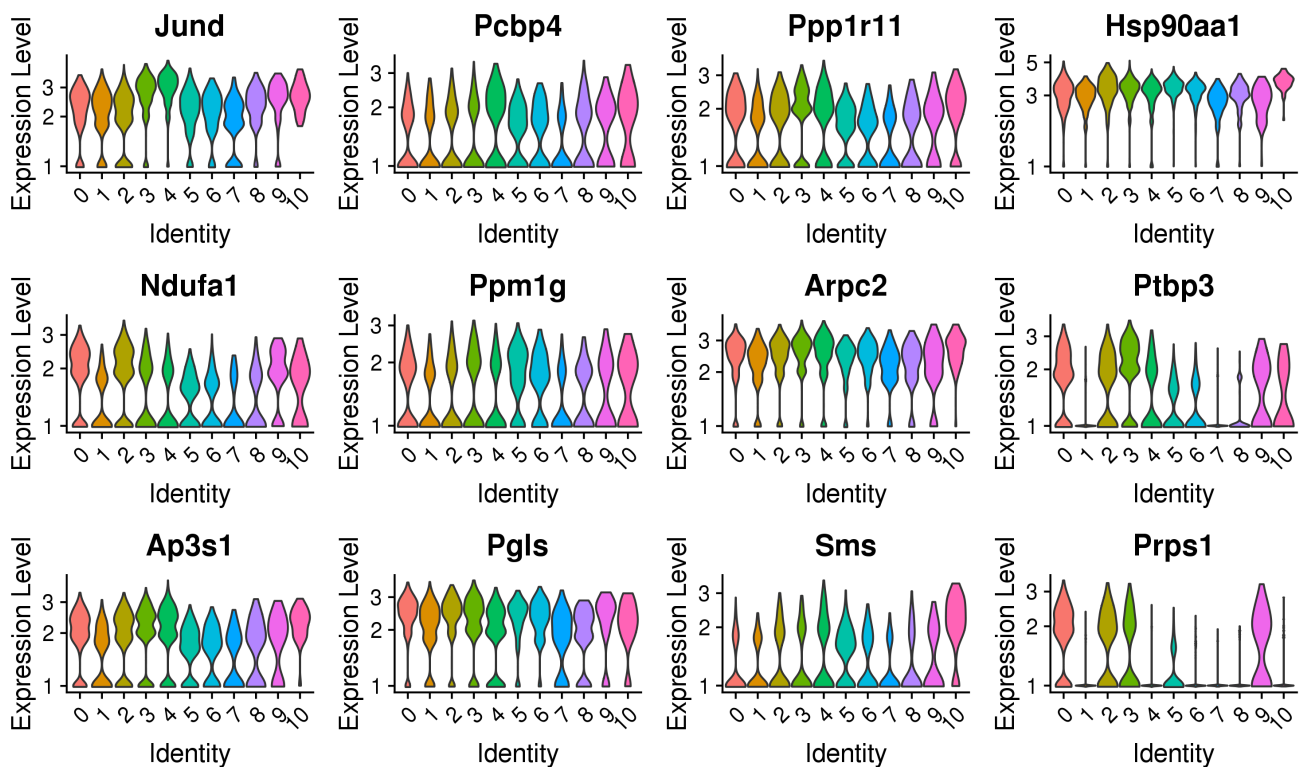


(b) Additional positive marker genes ordered by minimum fold-change vs all other clusters, cluster: 7

18.24 Cluster 7 violin plots: negative marker genes



(a) Top negative marker genes ordered by p-value, cluster: 7



(b) Additional negative marker genes ordered by minimum fold-change vs all other clusters, cluster: 7

18.25 Cluster 8: summary plots

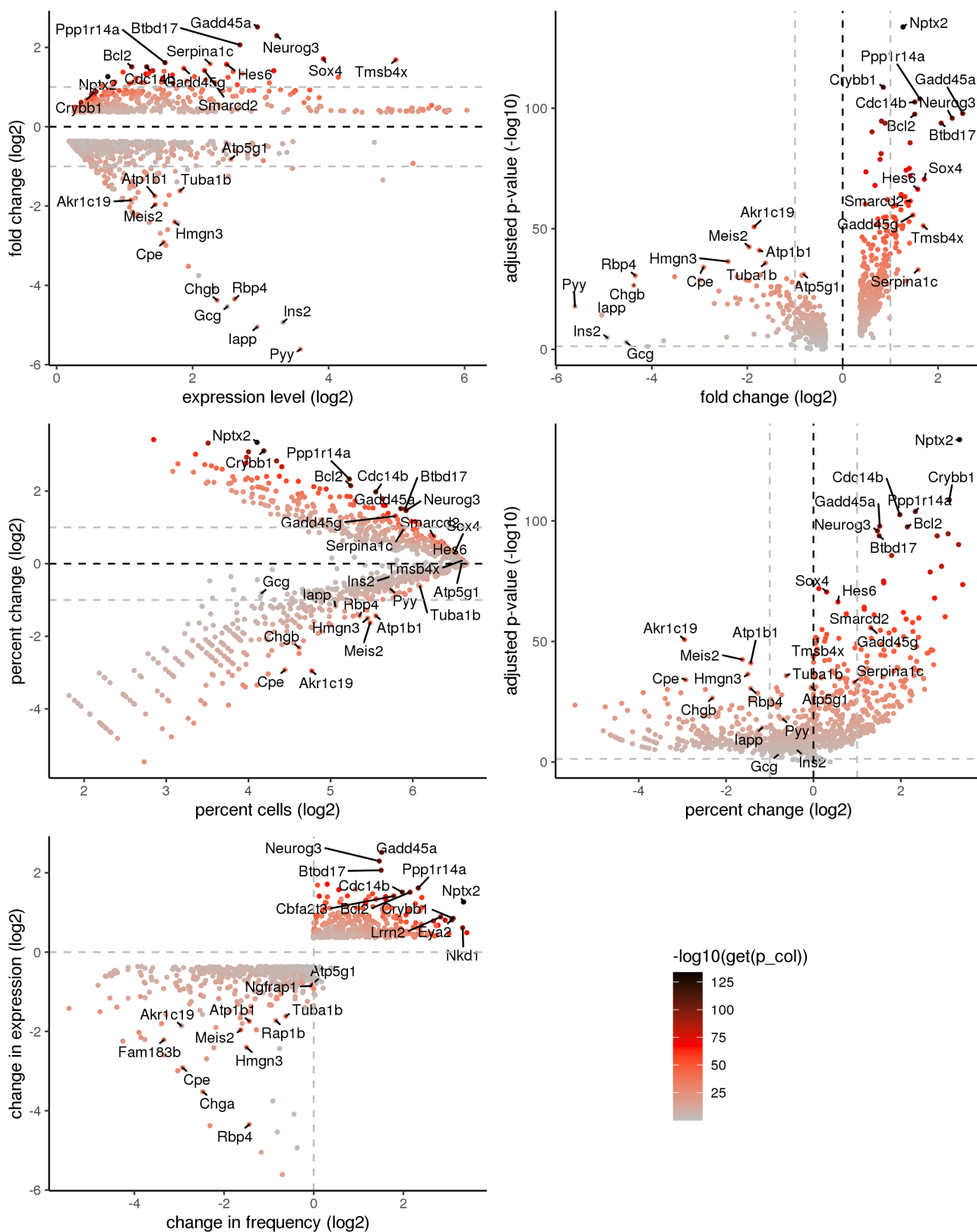
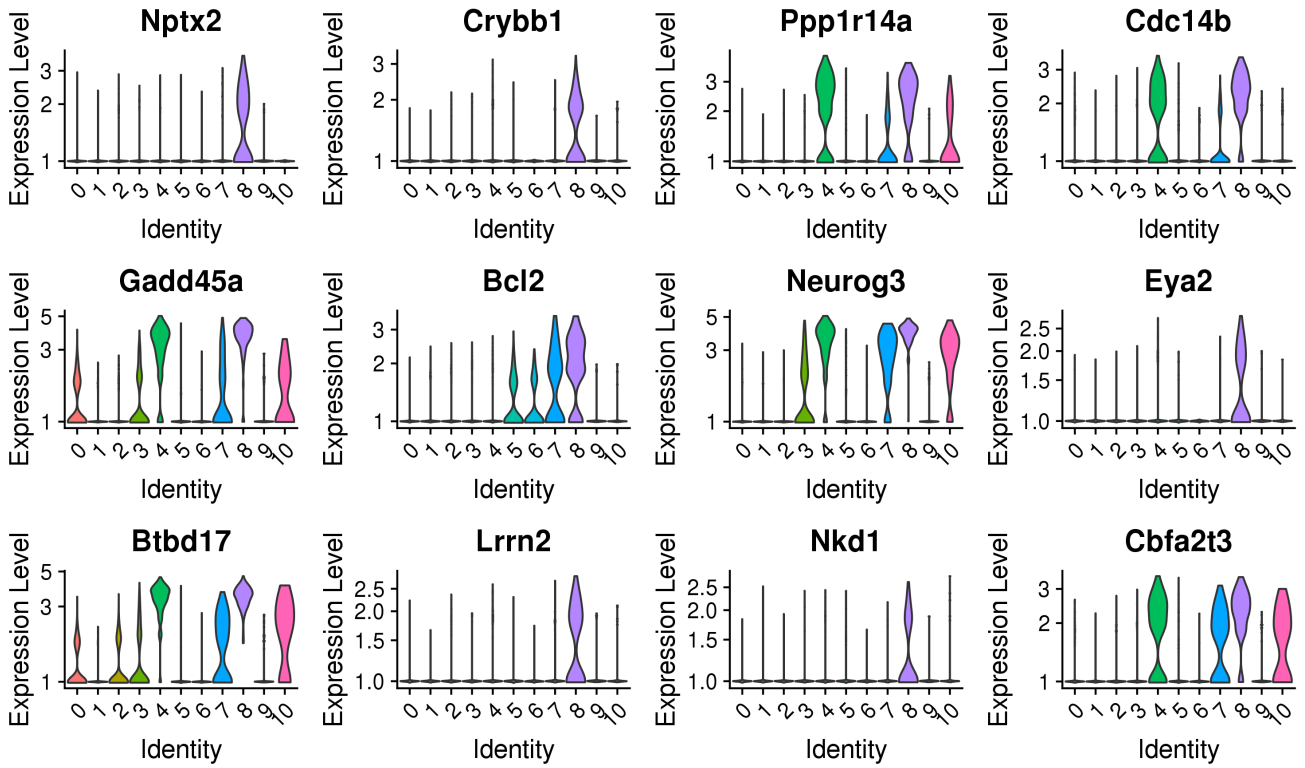
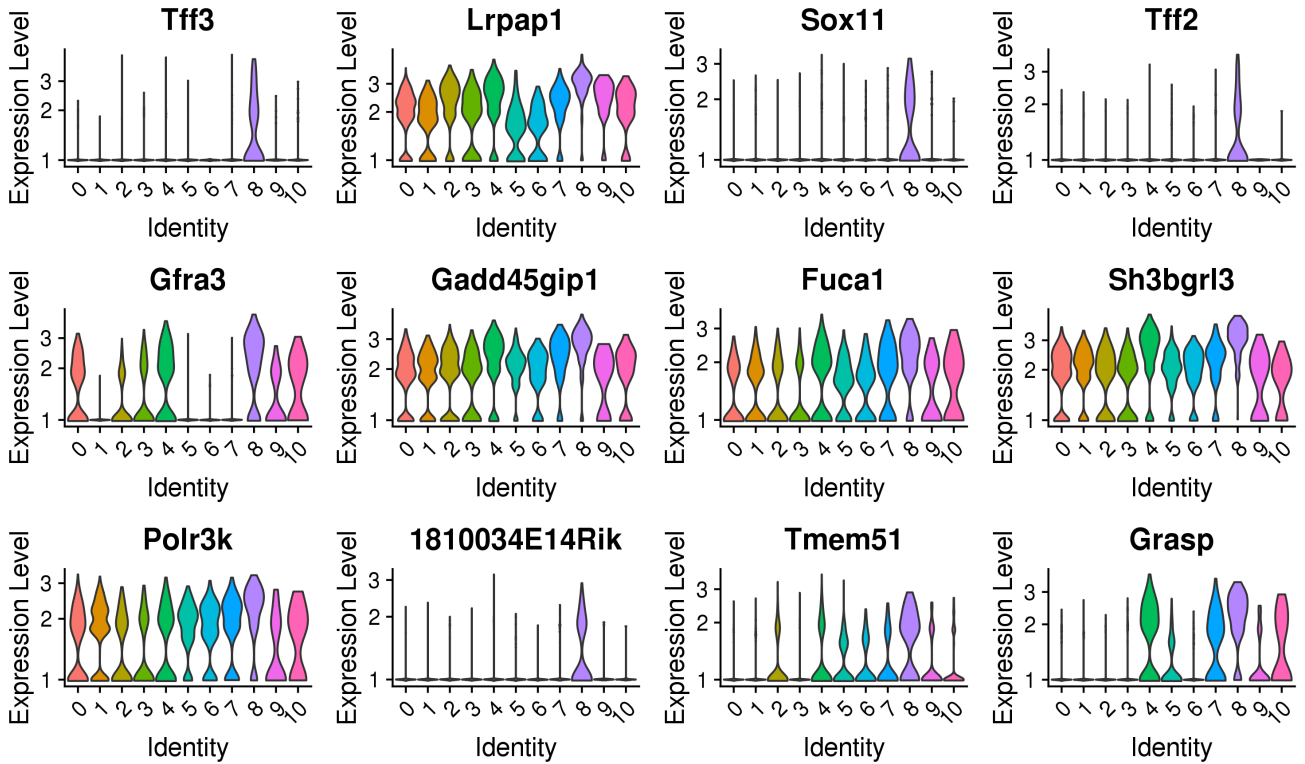


Figure 67: Differential expression summary plots for cluster 8

18.26 Cluster 8 violin plots: positive marker genes

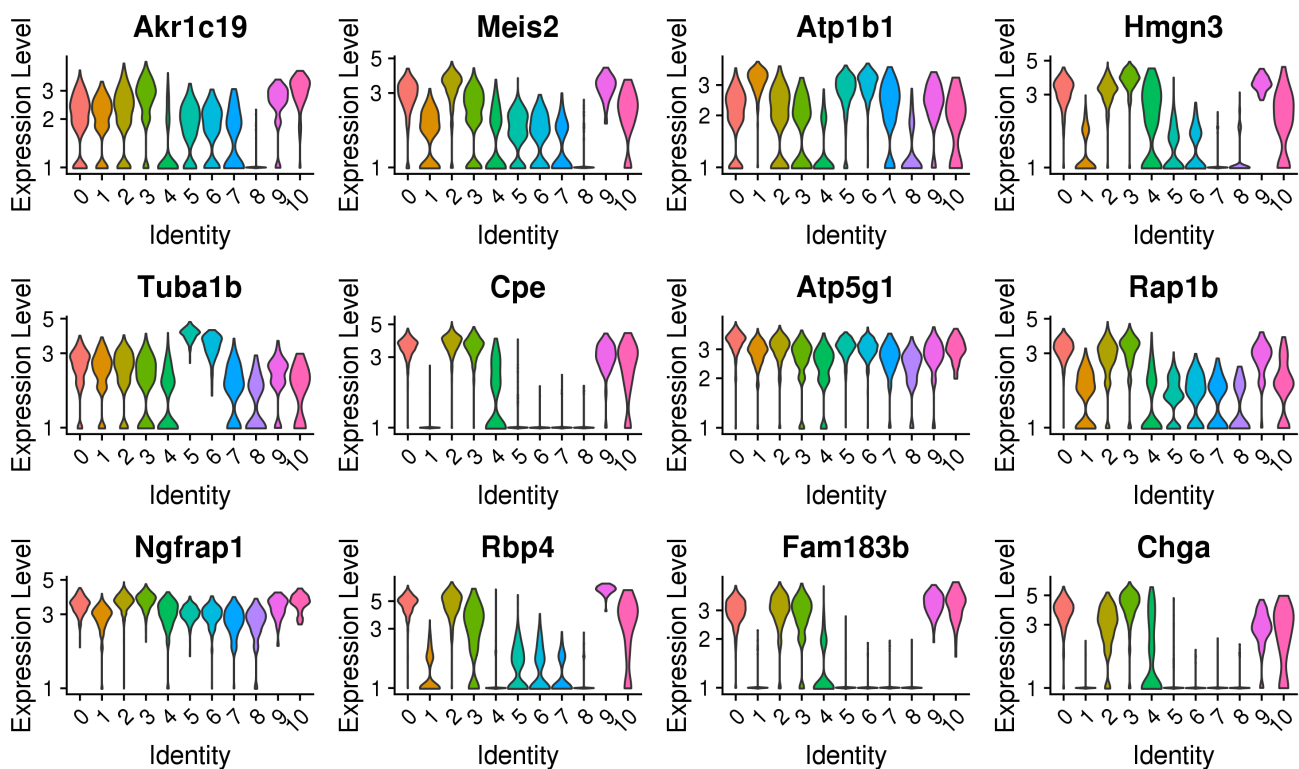


(a) Top positive marker genes ordered by p-value, cluster: 8

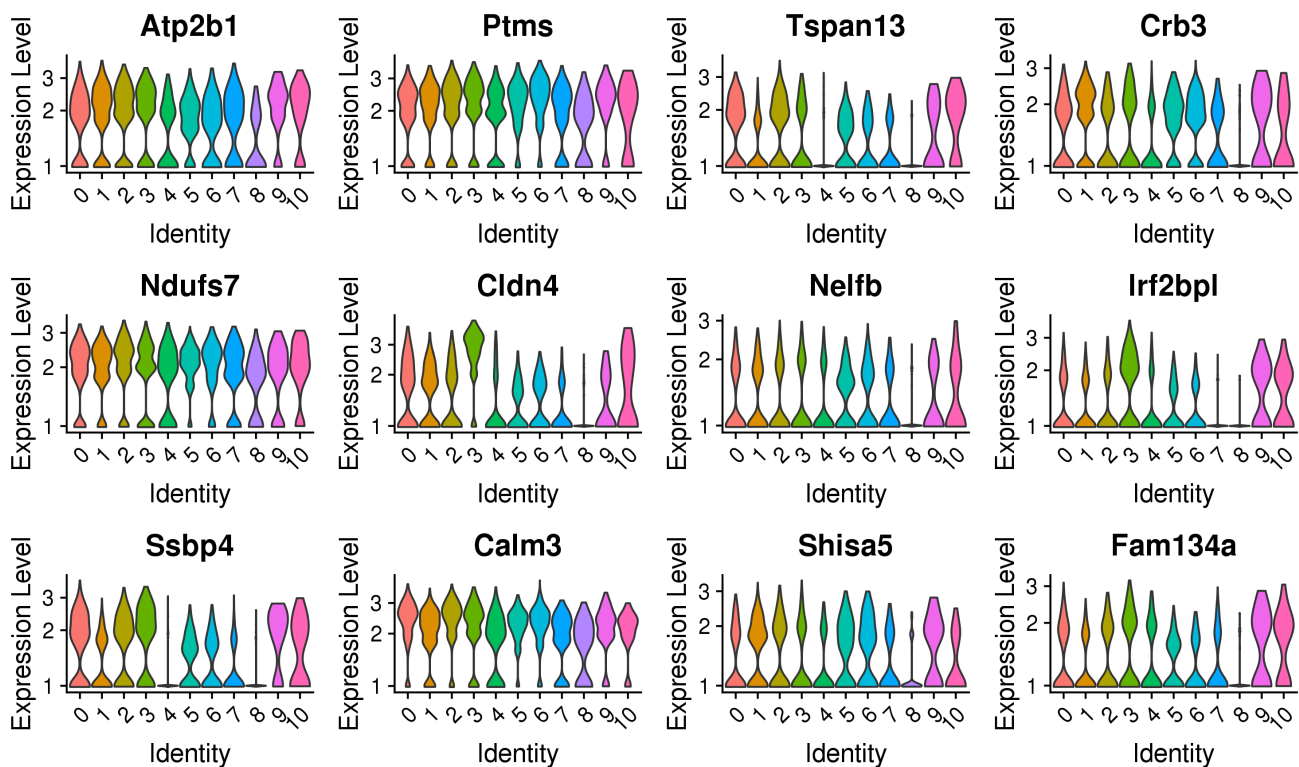


(b) Additional positive marker genes ordered by minimum fold-change vs all other clusters, cluster: 8

18.27 Cluster 8 violin plots: negative marker genes



(a) Top negative marker genes ordered by p-value, cluster: 8



(b) Additional negative marker genes ordered by minimum fold-change vs all other clusters, cluster: 8

18.28 Cluster 9: summary plots

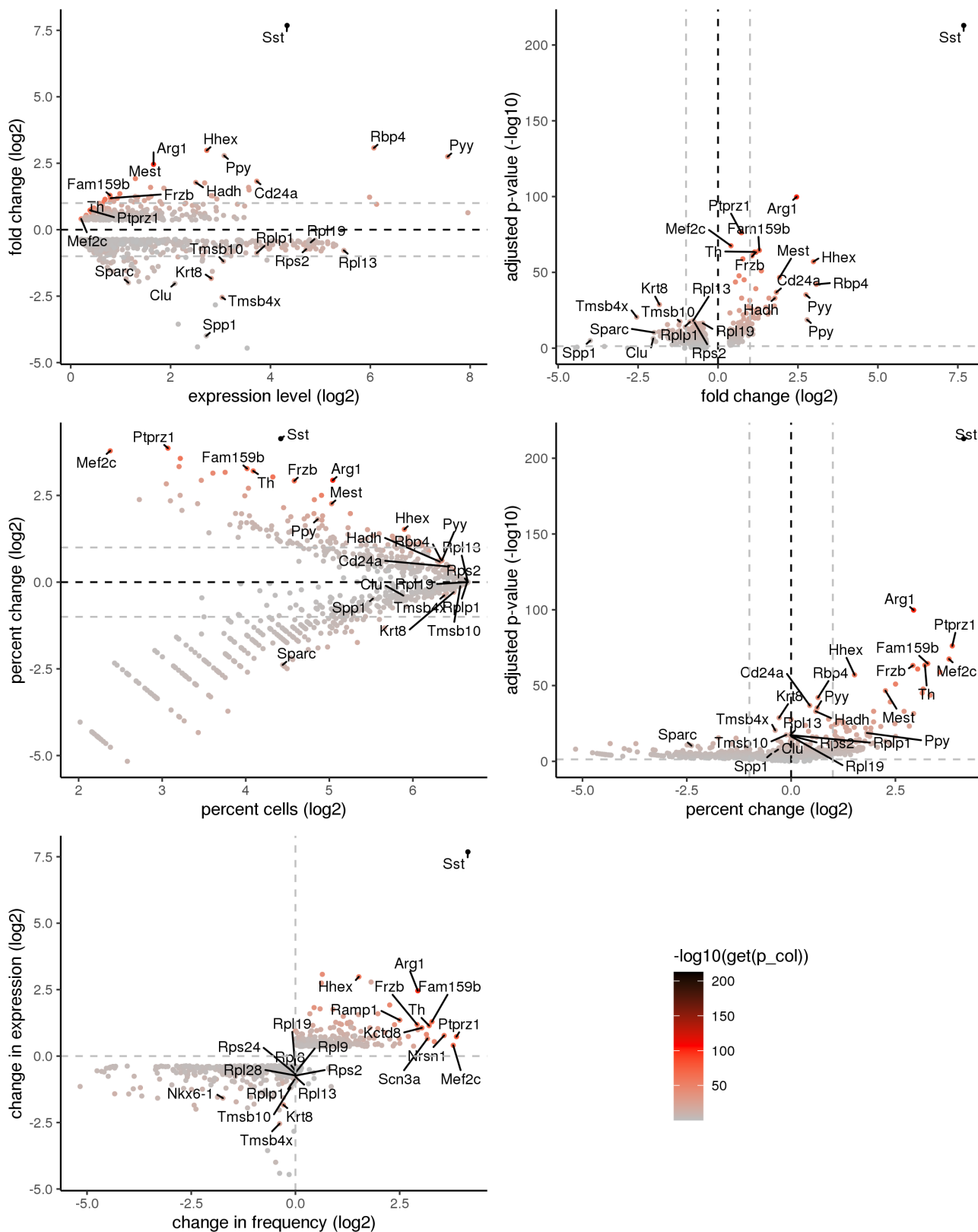
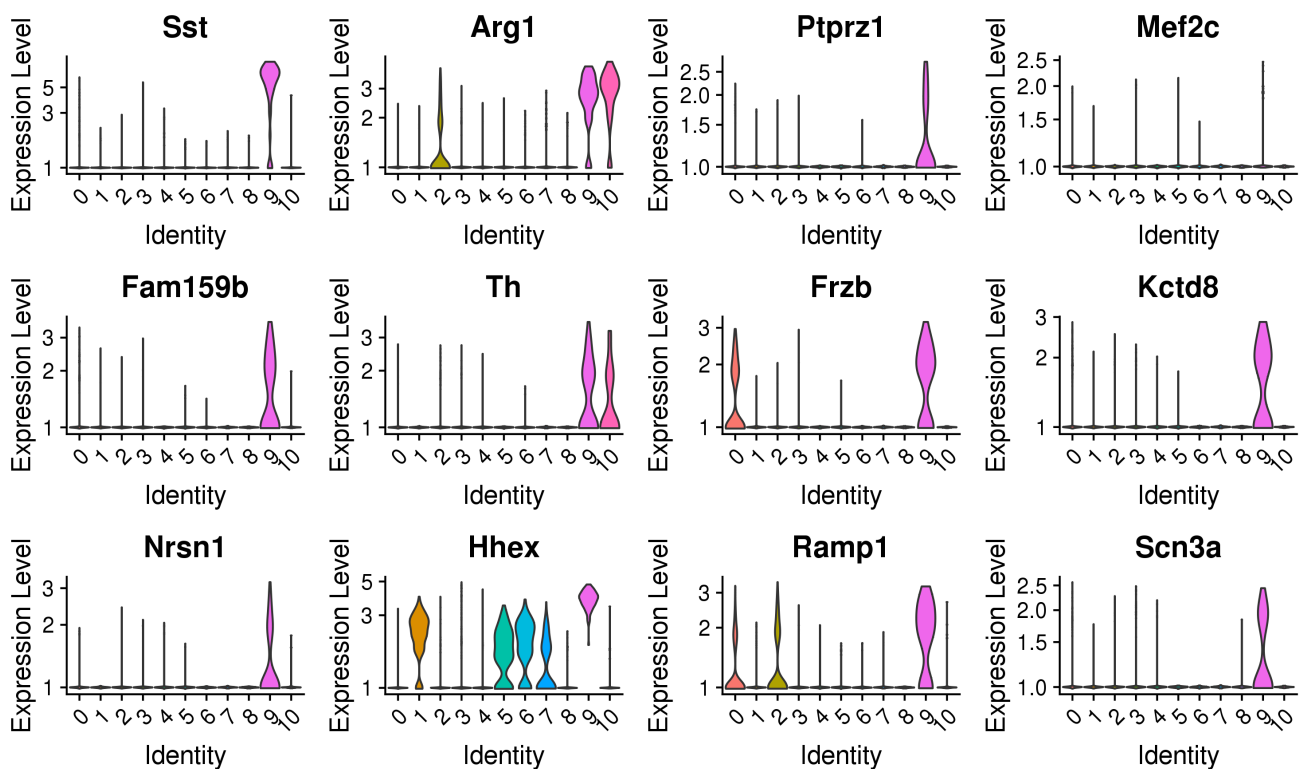
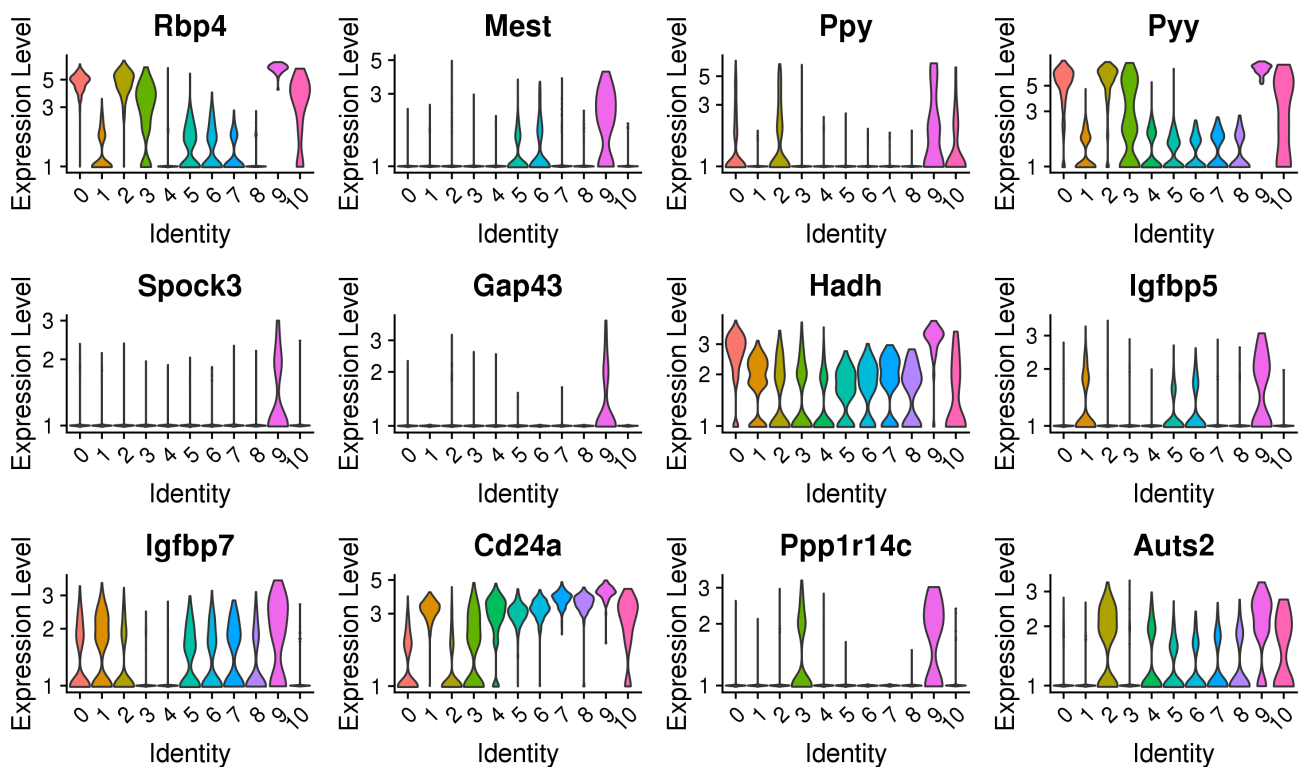


Figure 70: Differential expression summary plots for cluster 9

18.29 Cluster 9 violin plots: positive marker genes

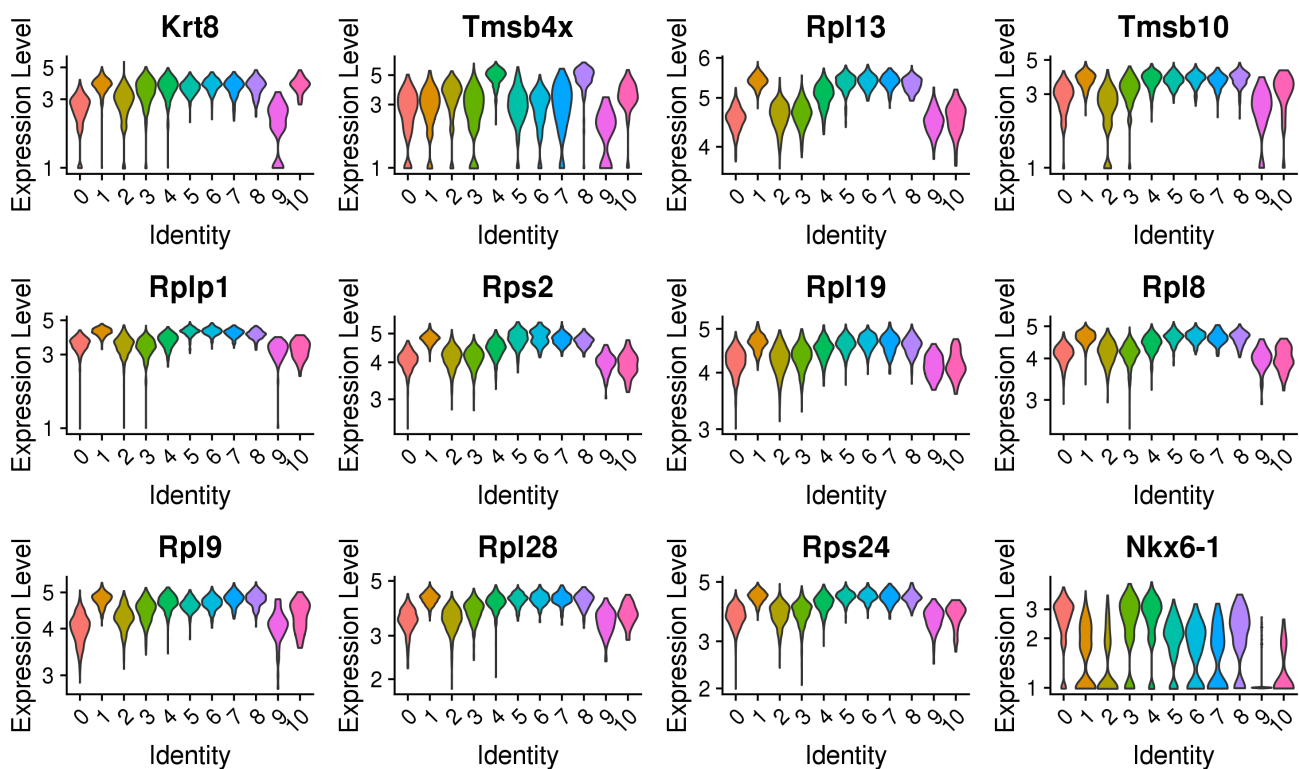


(a) Top positive marker genes ordered by p-value, cluster: 9

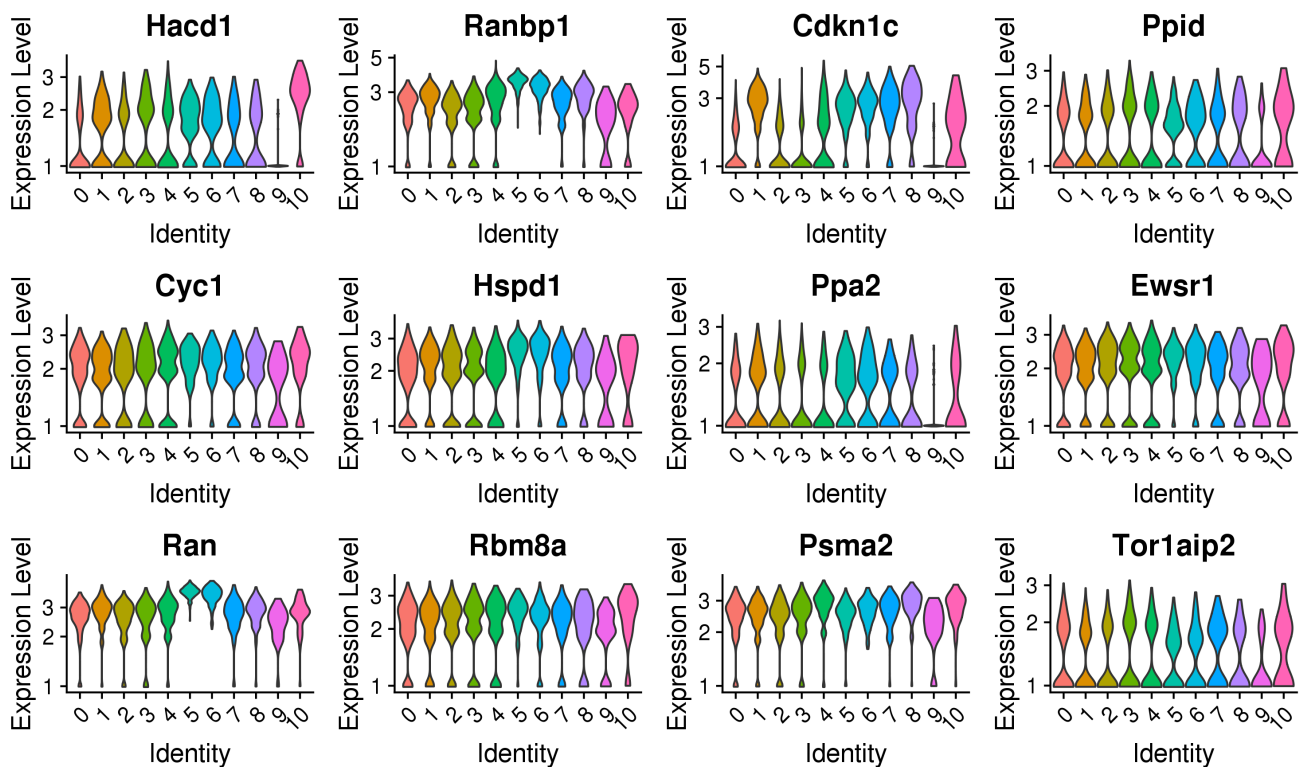


(b) Additional positive marker genes ordered by minimum fold-change vs all other clusters, cluster: 9

18.30 Cluster 9 violin plots: negative marker genes



(a) Top negative marker genes ordered by p-value, cluster: 9



(b) Additional negative marker genes ordered by minimum fold-change vs all other clusters, cluster: 9

18.31 Cluster 10: summary plots

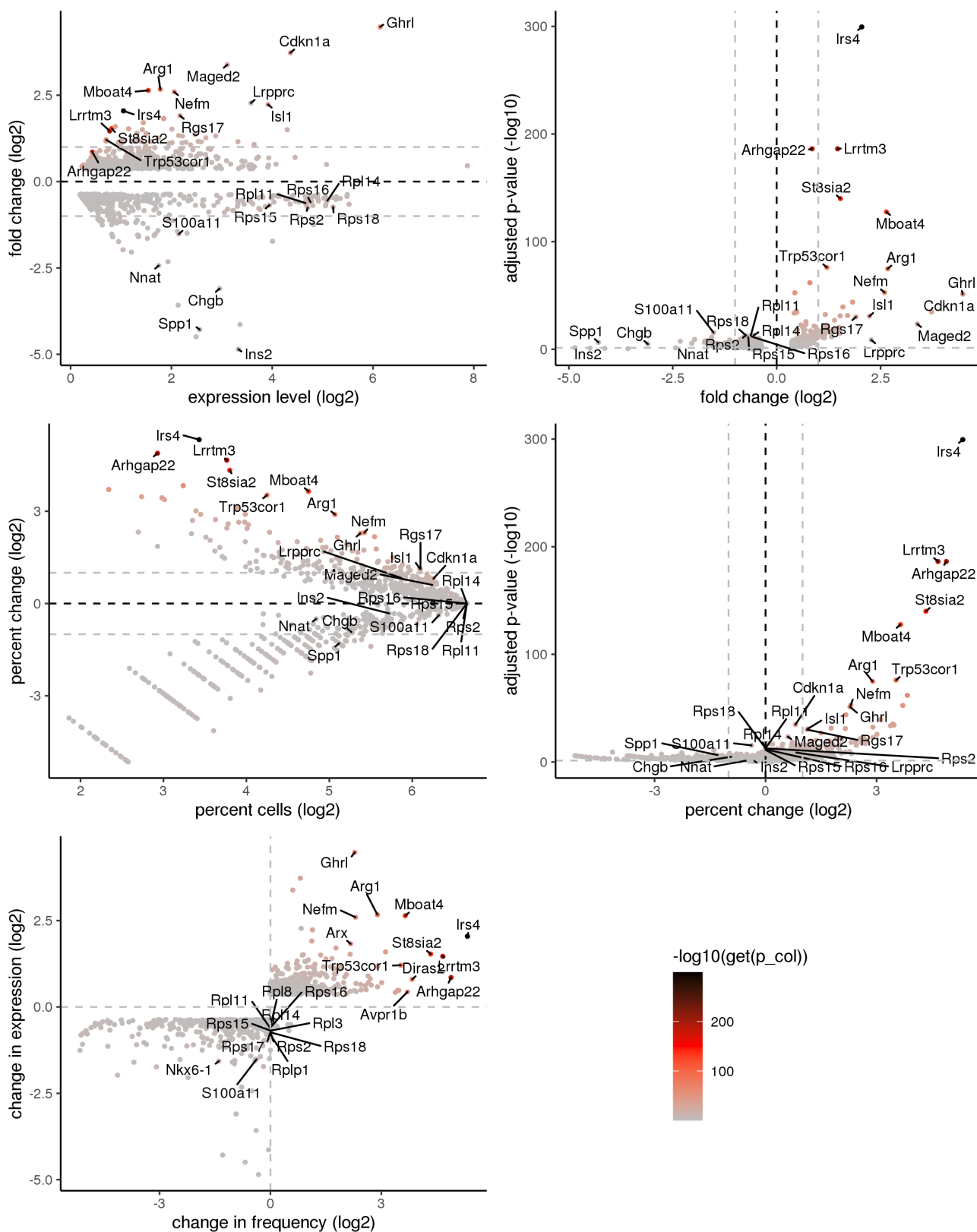
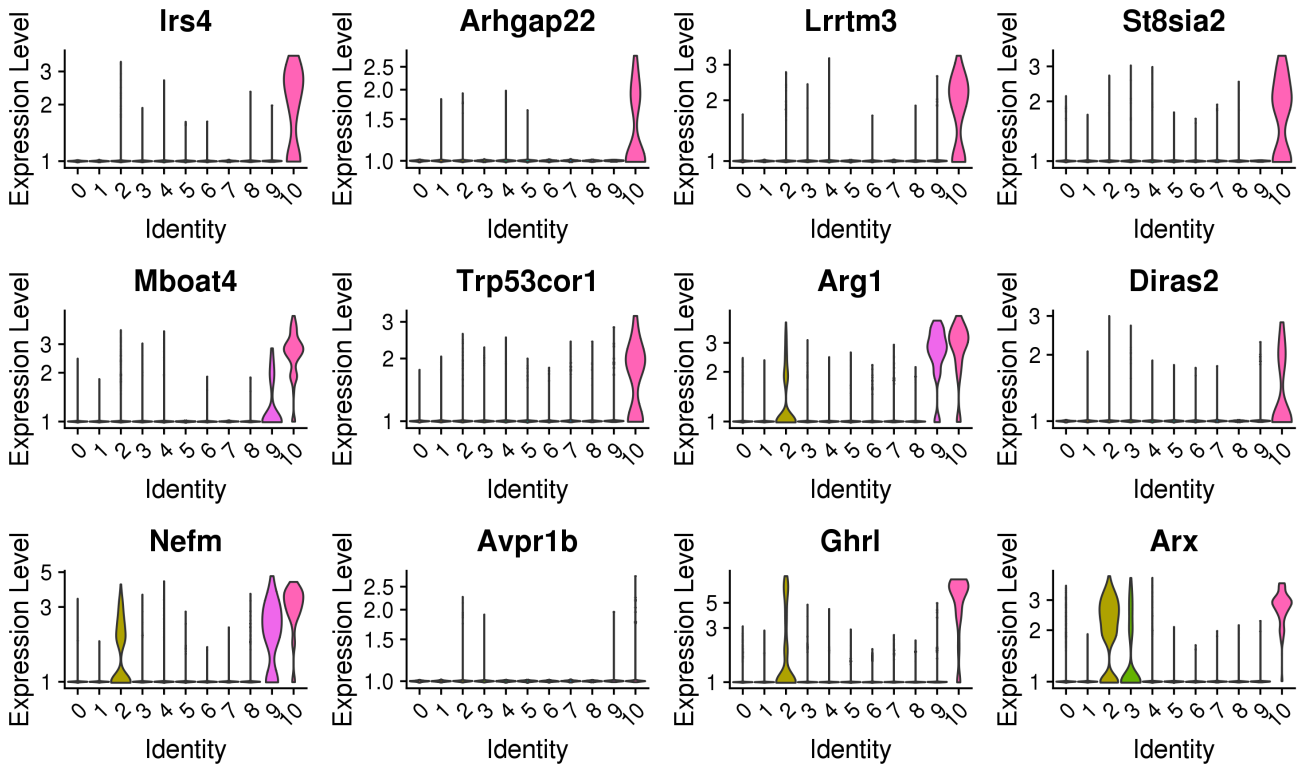
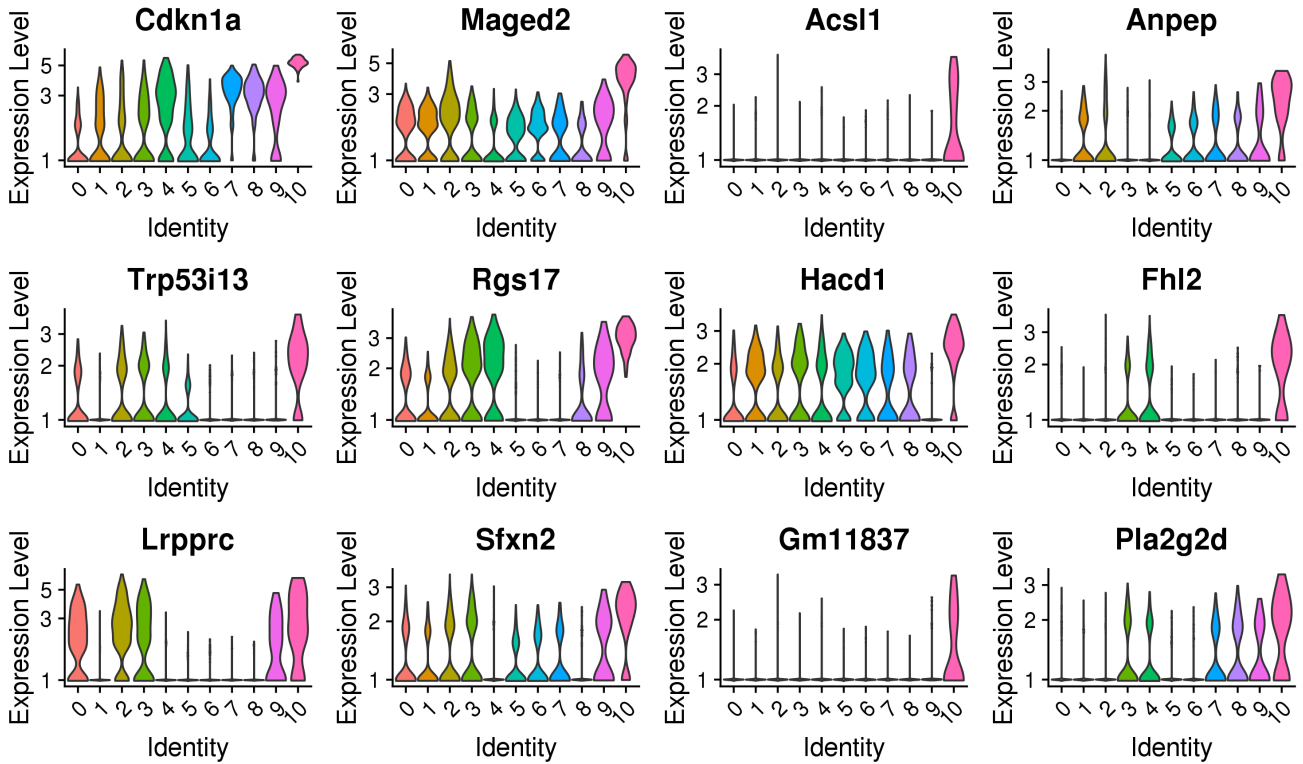


Figure 73: Differential expression summary plots for cluster 10

18.32 Cluster 10 violin plots: positive marker genes

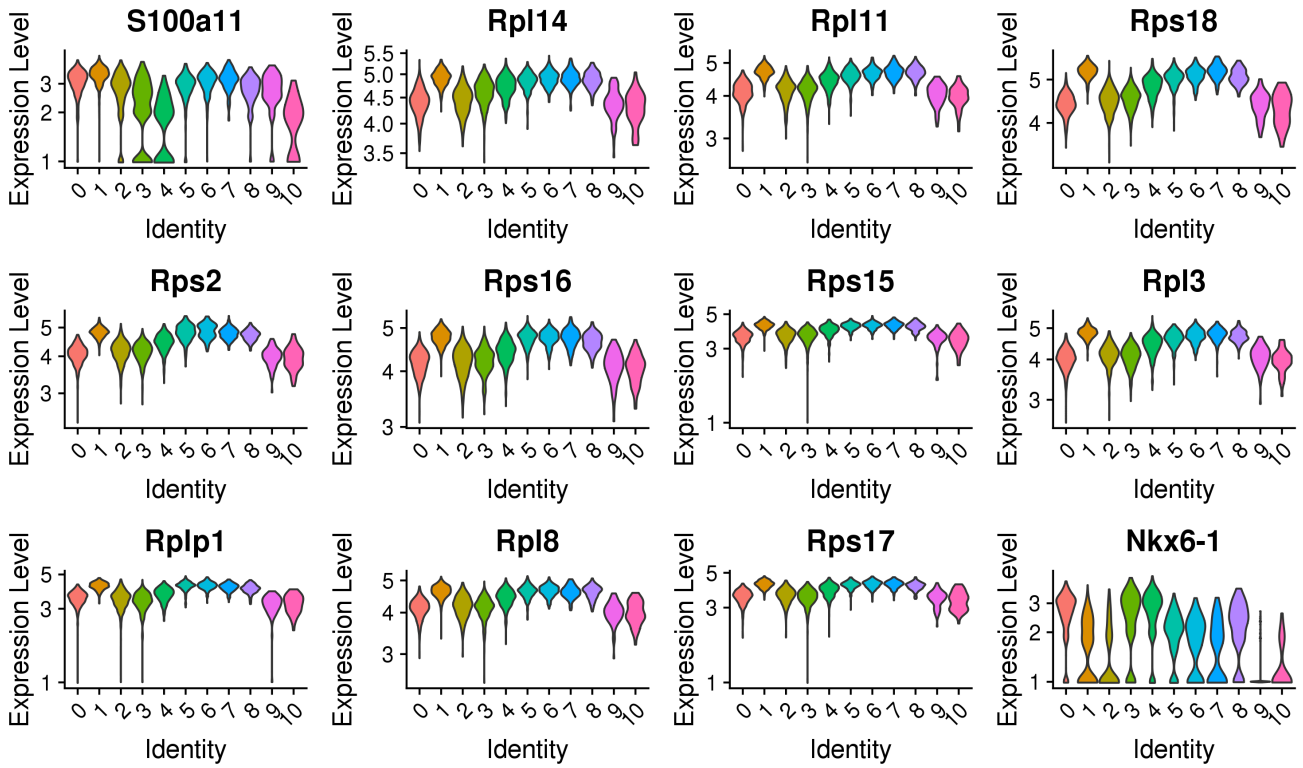


(a) Top positive marker genes ordered by p-value, cluster: 10

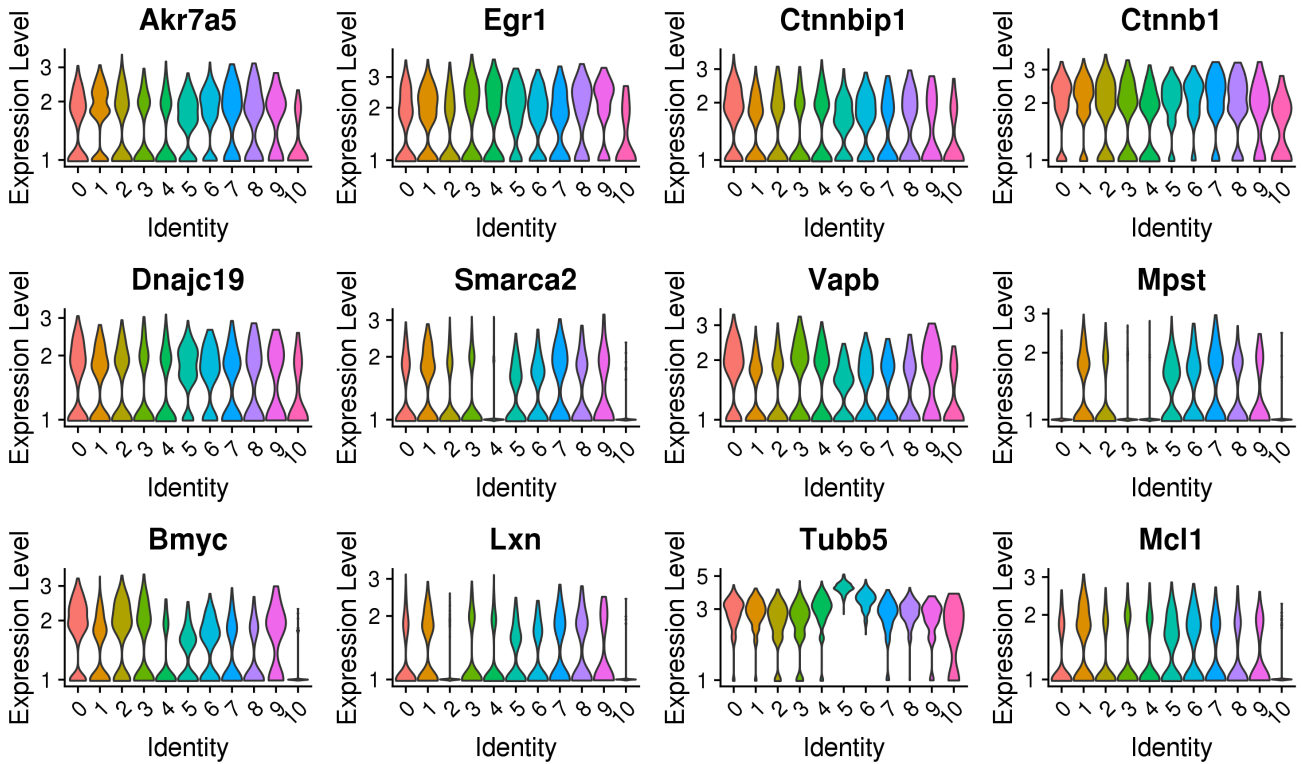


(b) Additional positive marker genes ordered by minimum fold-change vs all other clusters, cluster: 10

18.33 Cluster 10 violin plots: negative marker genes



(a) Top negative marker genes ordered by p-value, cluster: 10



(b) Additional negative marker genes ordered by minimum fold-change vs all other clusters, cluster: 10

19 Geneset analysis

A hypergeometric test is used to test for the enrichment of GO, KEGG and msigdb genesets amongst the positive marker genes for each cluster. The full results are available as a separate xlsx document.

19.1 GO.BP

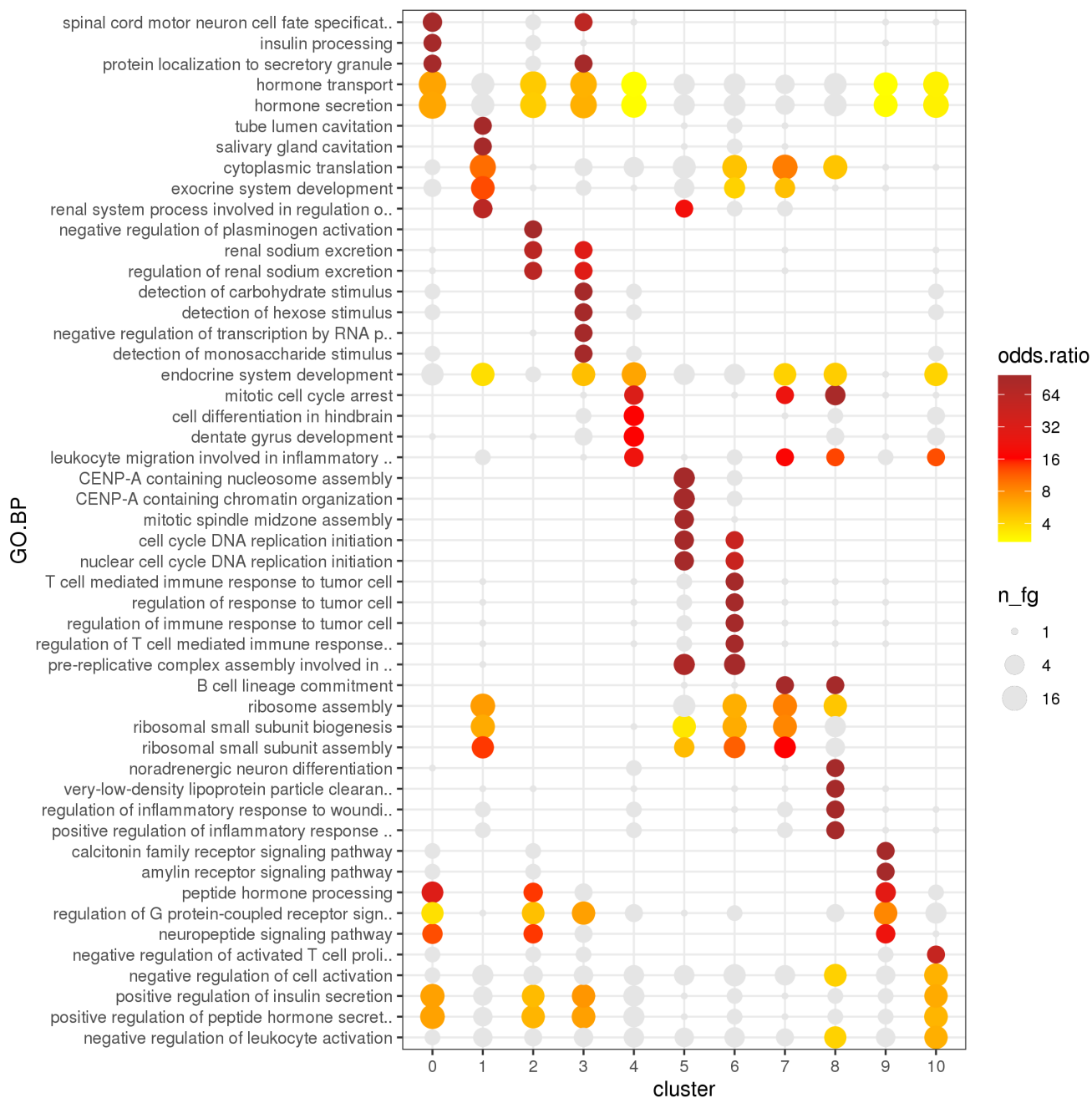


Figure 76: Heatmap of the top GO.BP genesets

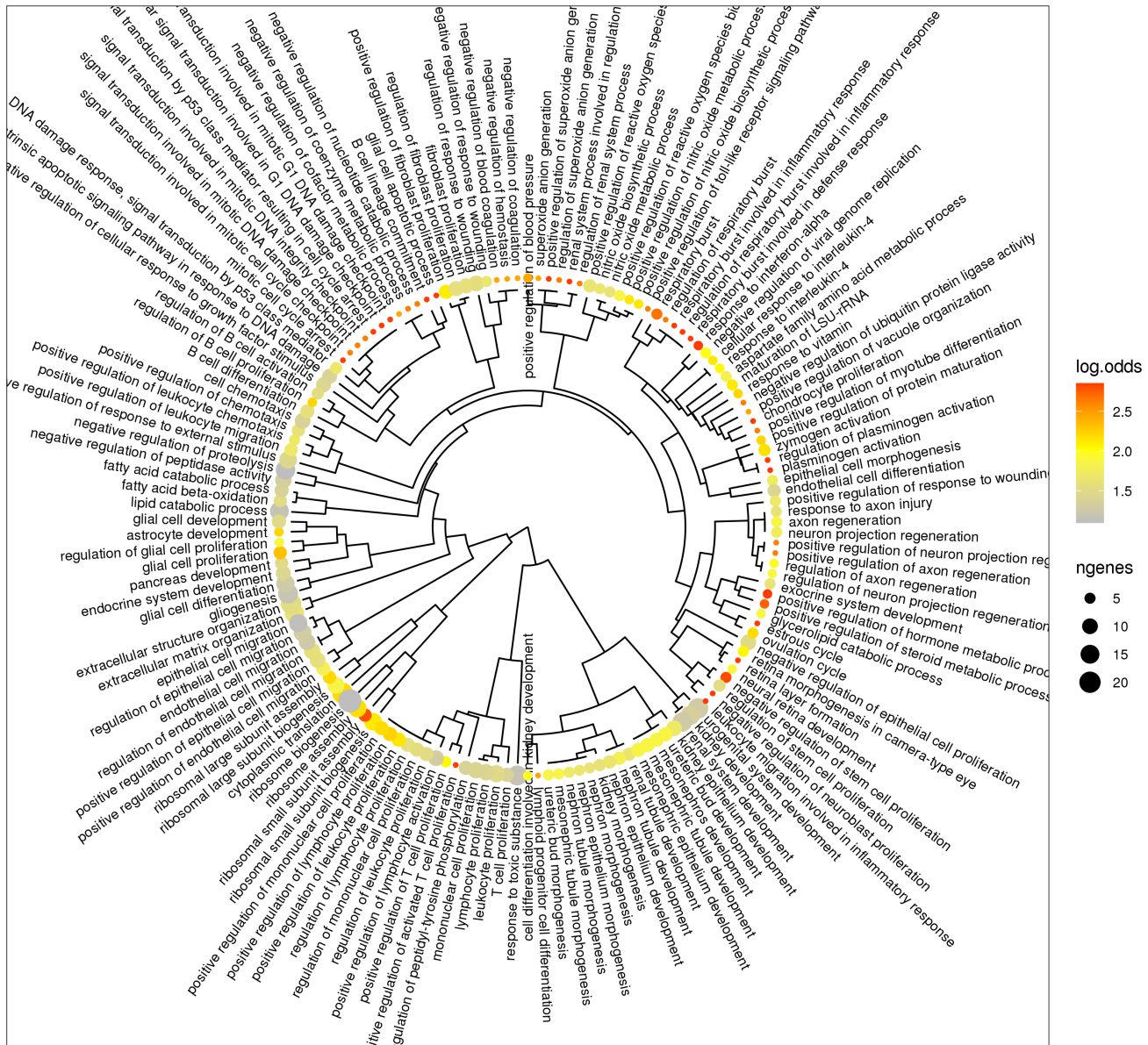


Figure 84: Cluster 7 GO.BP genesets clustered by similarity between over-represented genes.

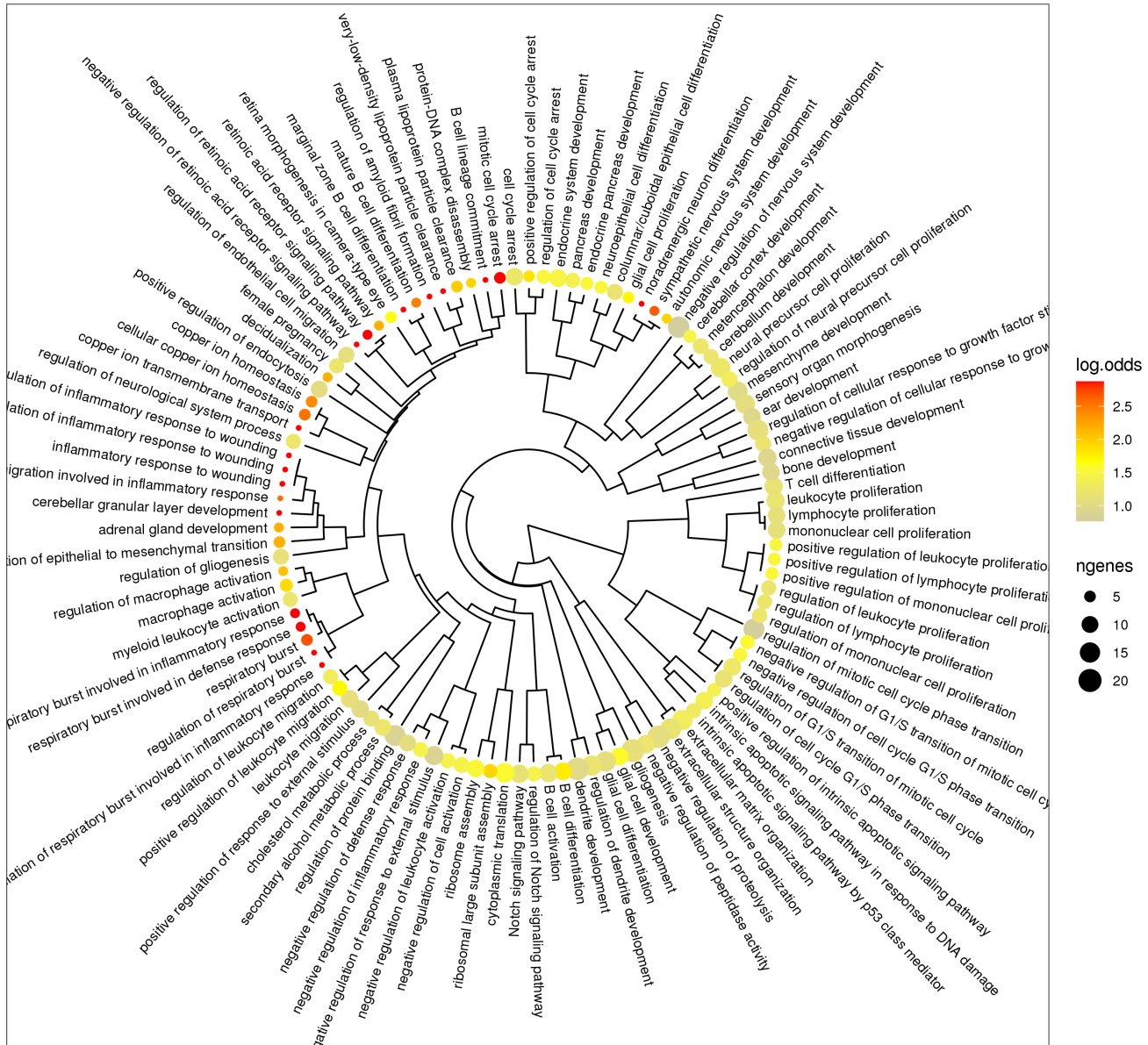


Figure 85: Cluster 8 GO.BP genesets clustered by similarity between over-represented genes.

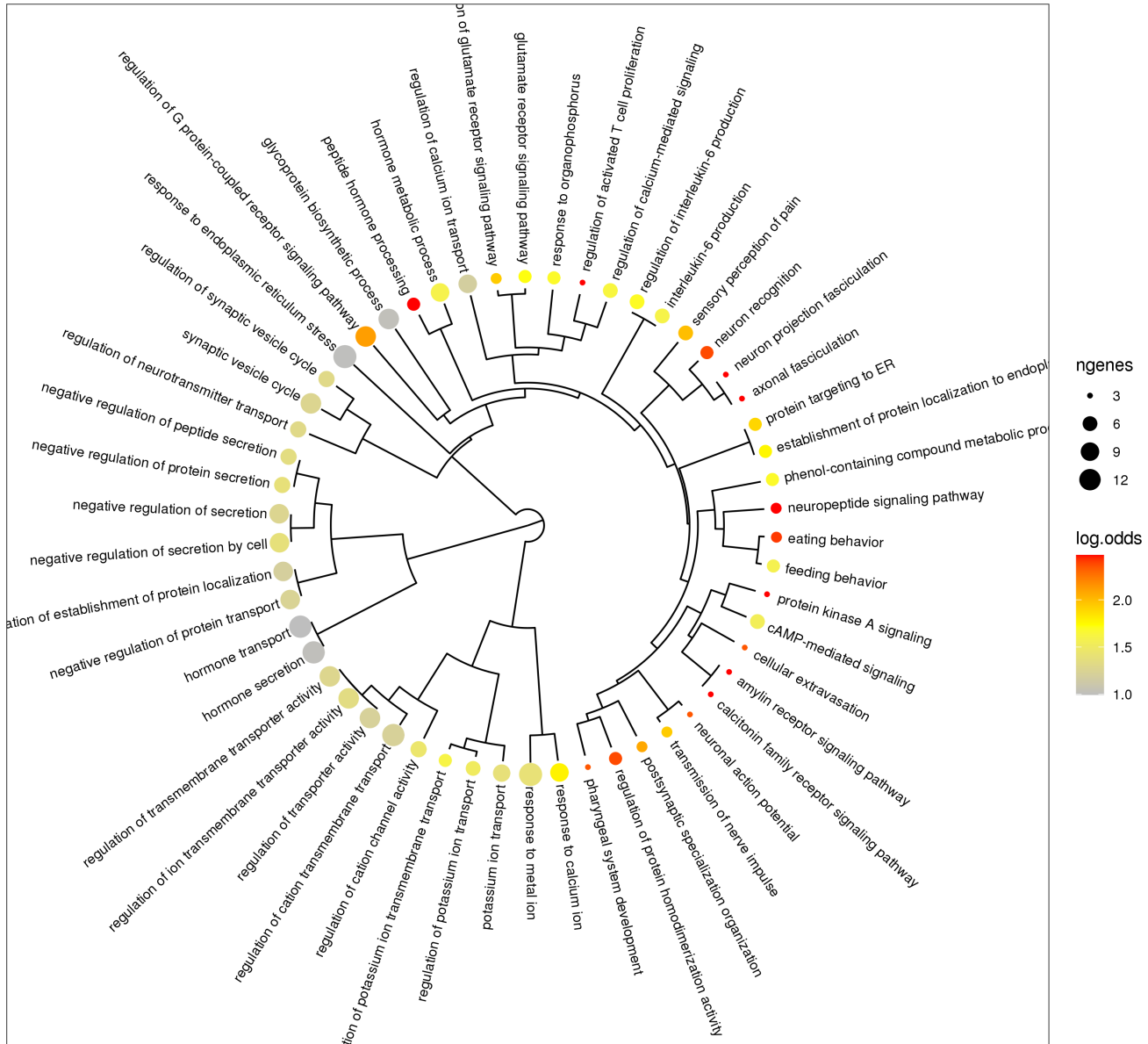


Figure 86: Cluster 9 GO.BP genesets clustered by similarity between over-represented genes.

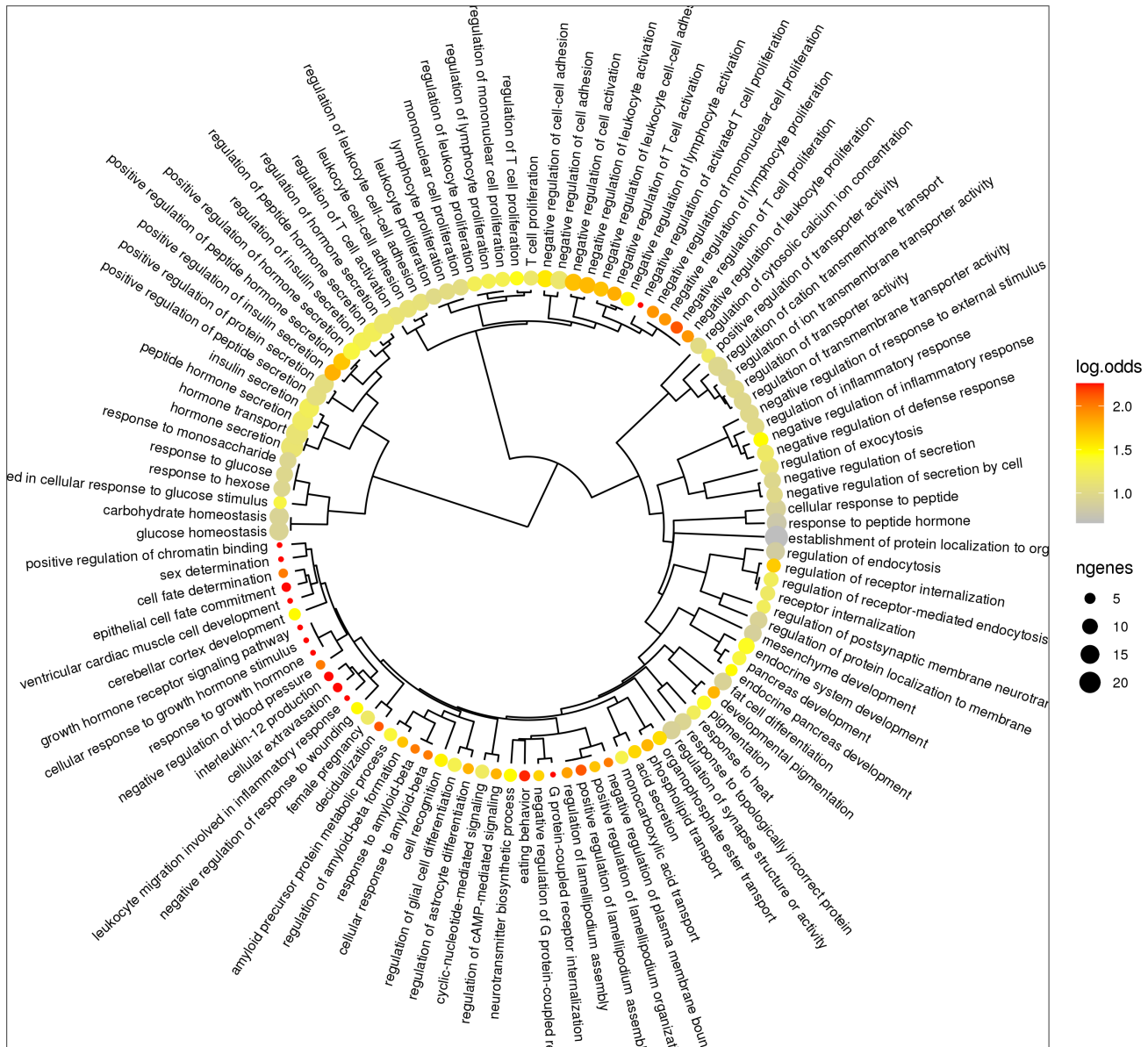


Figure 87: Cluster 10 GO.BP genesets clustered by similarity between over-represented genes.

19.2 GO.MF

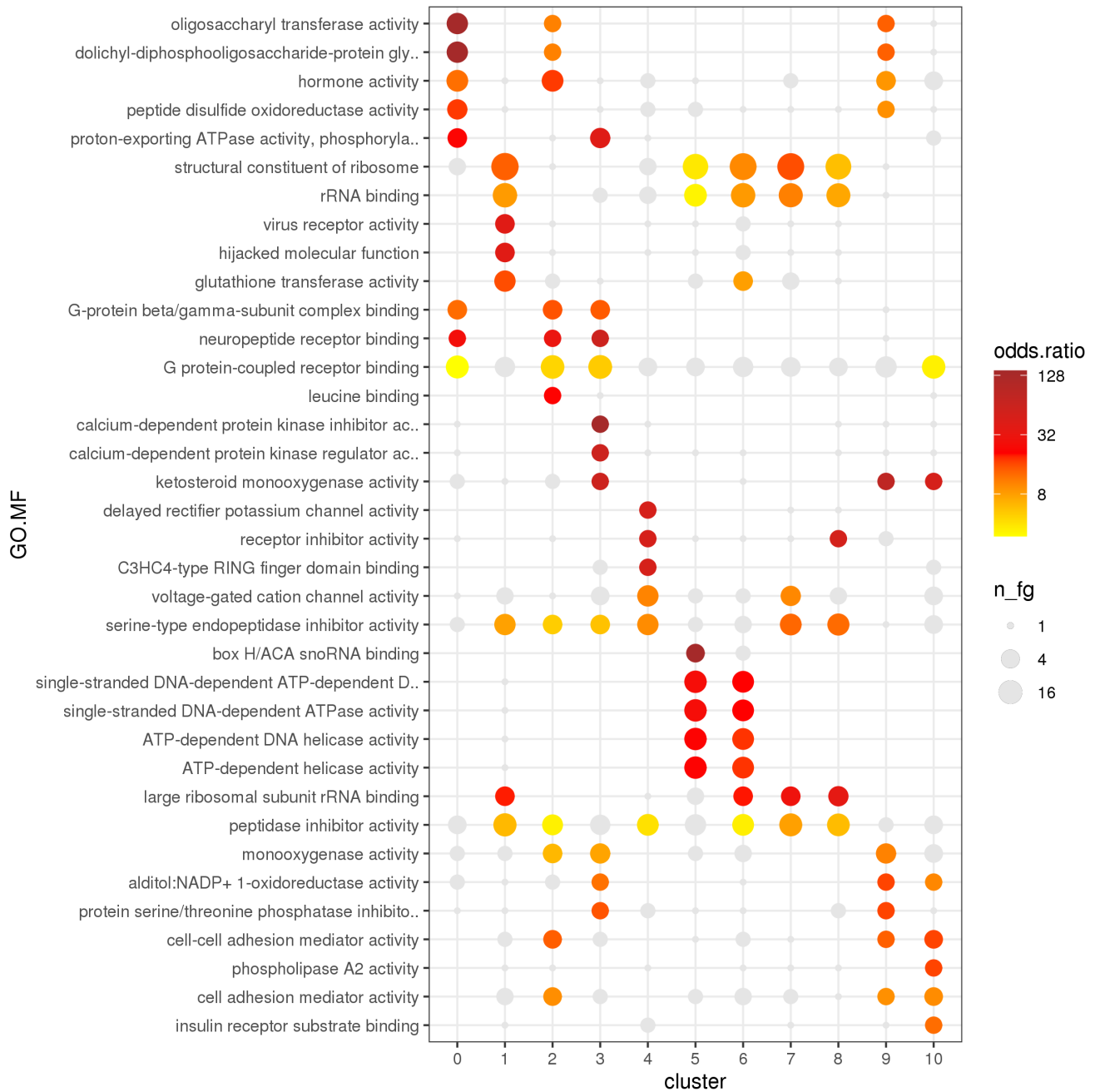


Figure 88: Heatmap of the top GO.MF genesets

19.3 GO.CC

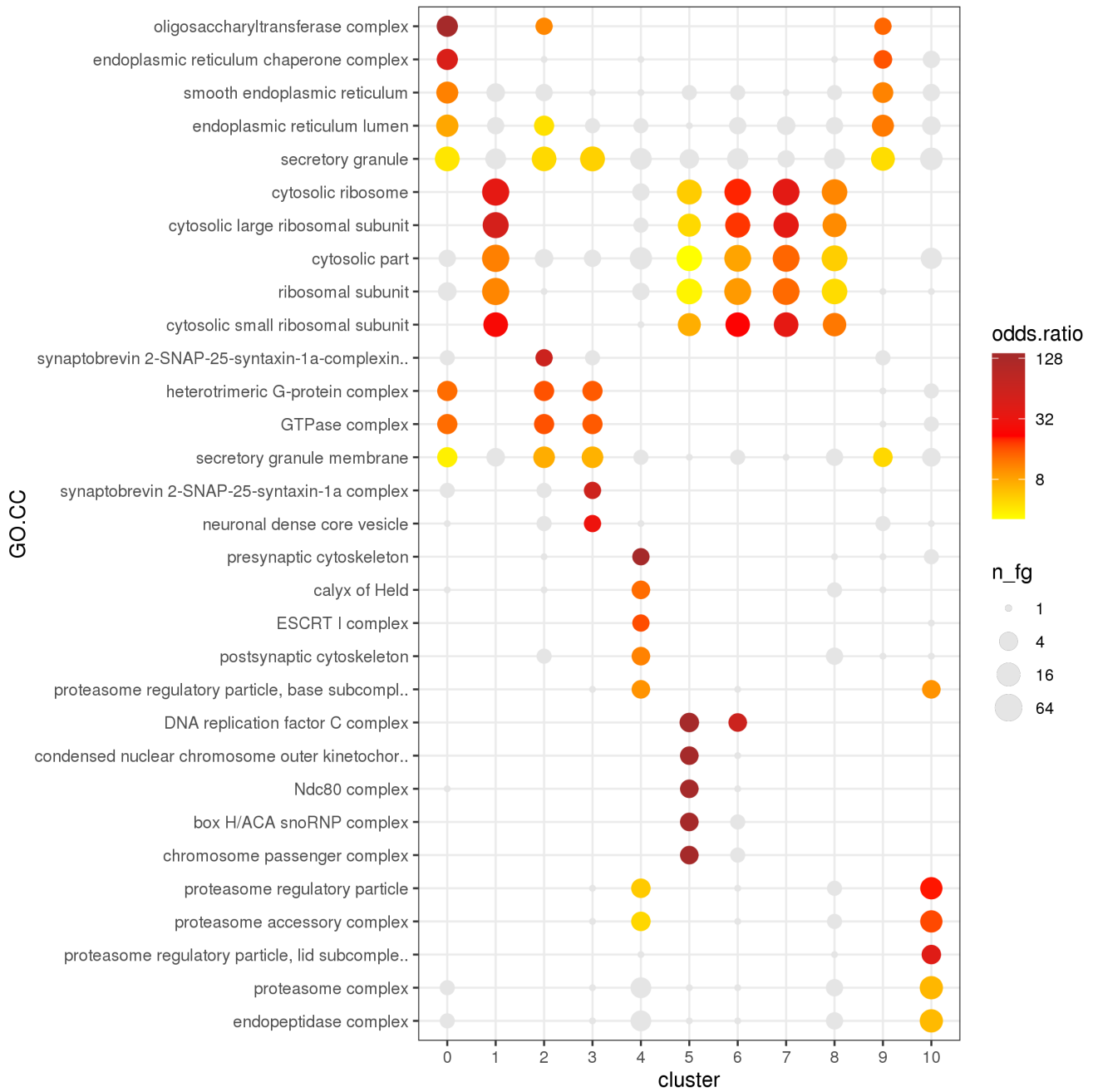


Figure 89: Heatmap of the top GO.CC genesets

19.4 KEGG

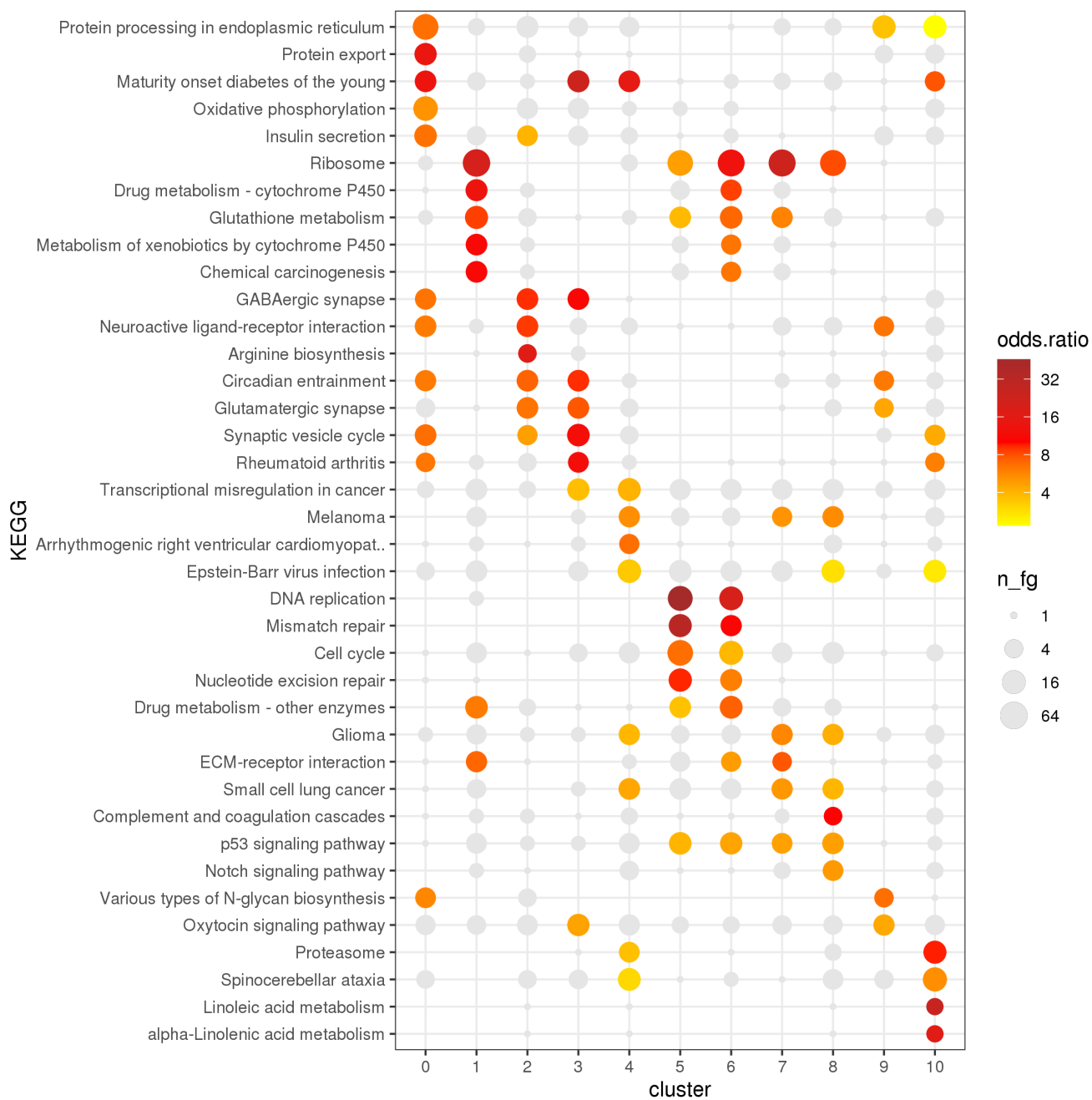


Figure 90: Heatmap of the top KEGG genesets

19.5 msigdb_biocarta

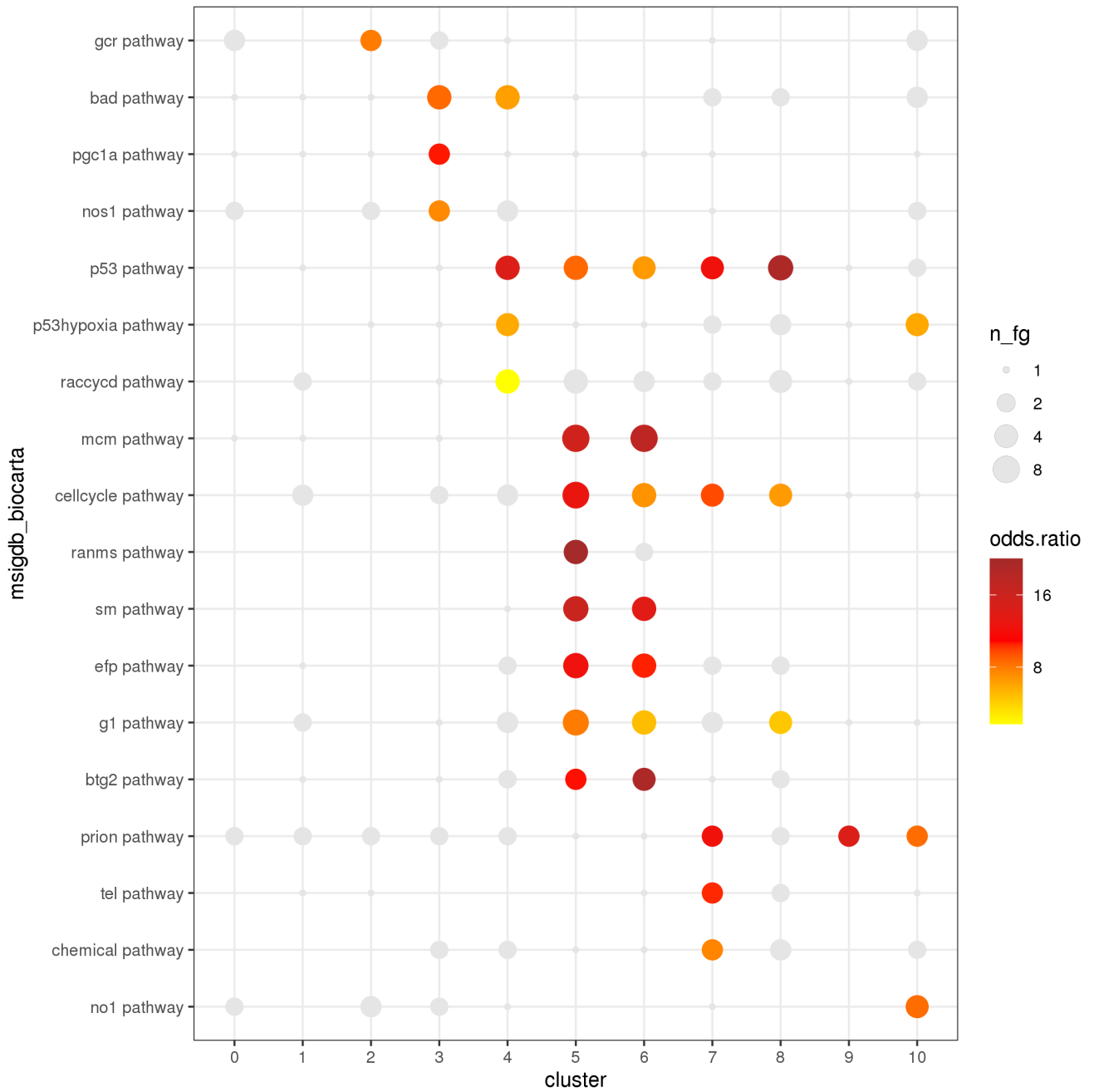


Figure 91: Heatmap of the top msigdb_biocarta genesets

19.6 msigdb_reactome

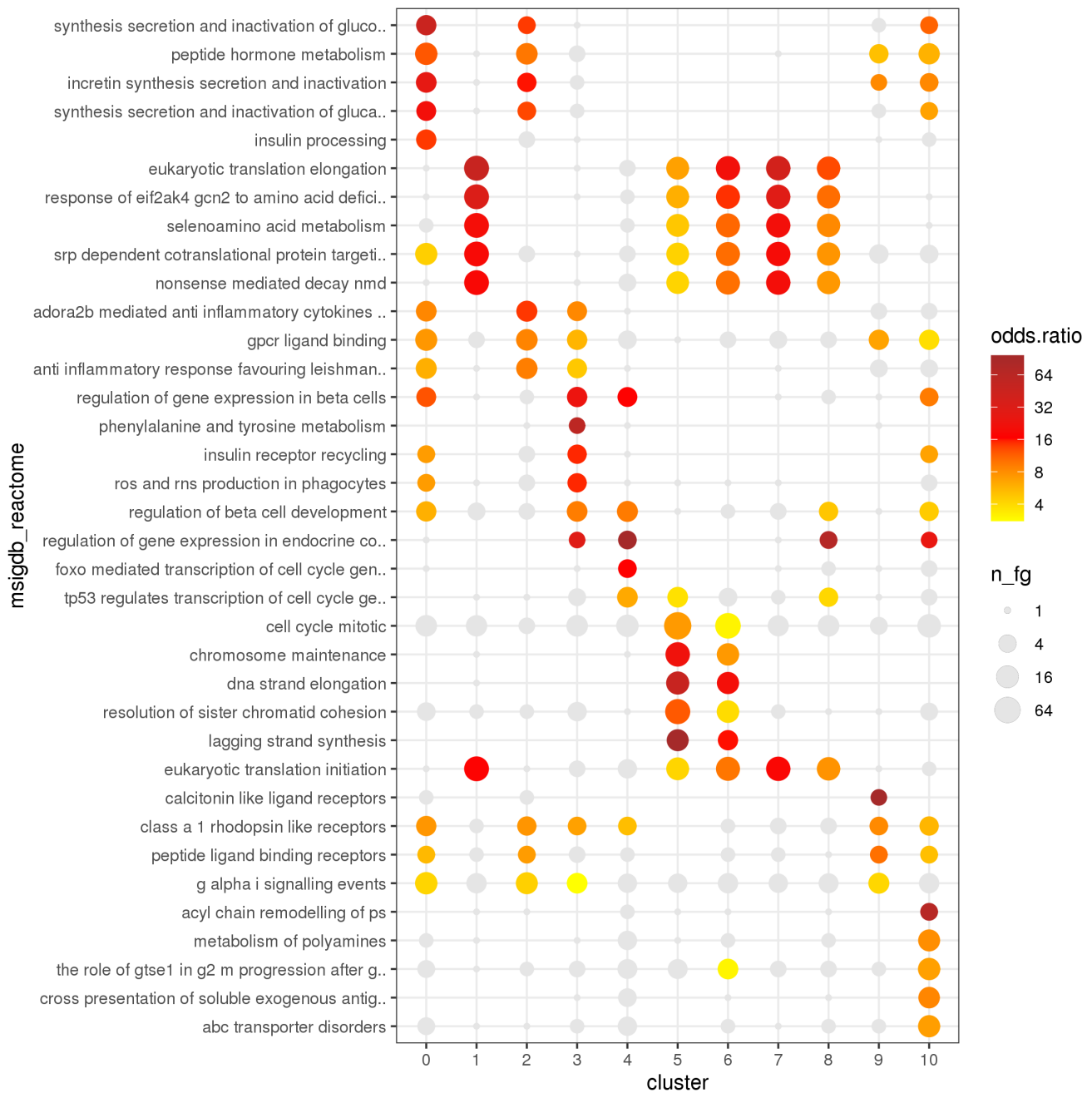


Figure 92: Heatmap of the top msigdb_reactome genesets

19.7 msigdb_canonical_pathways

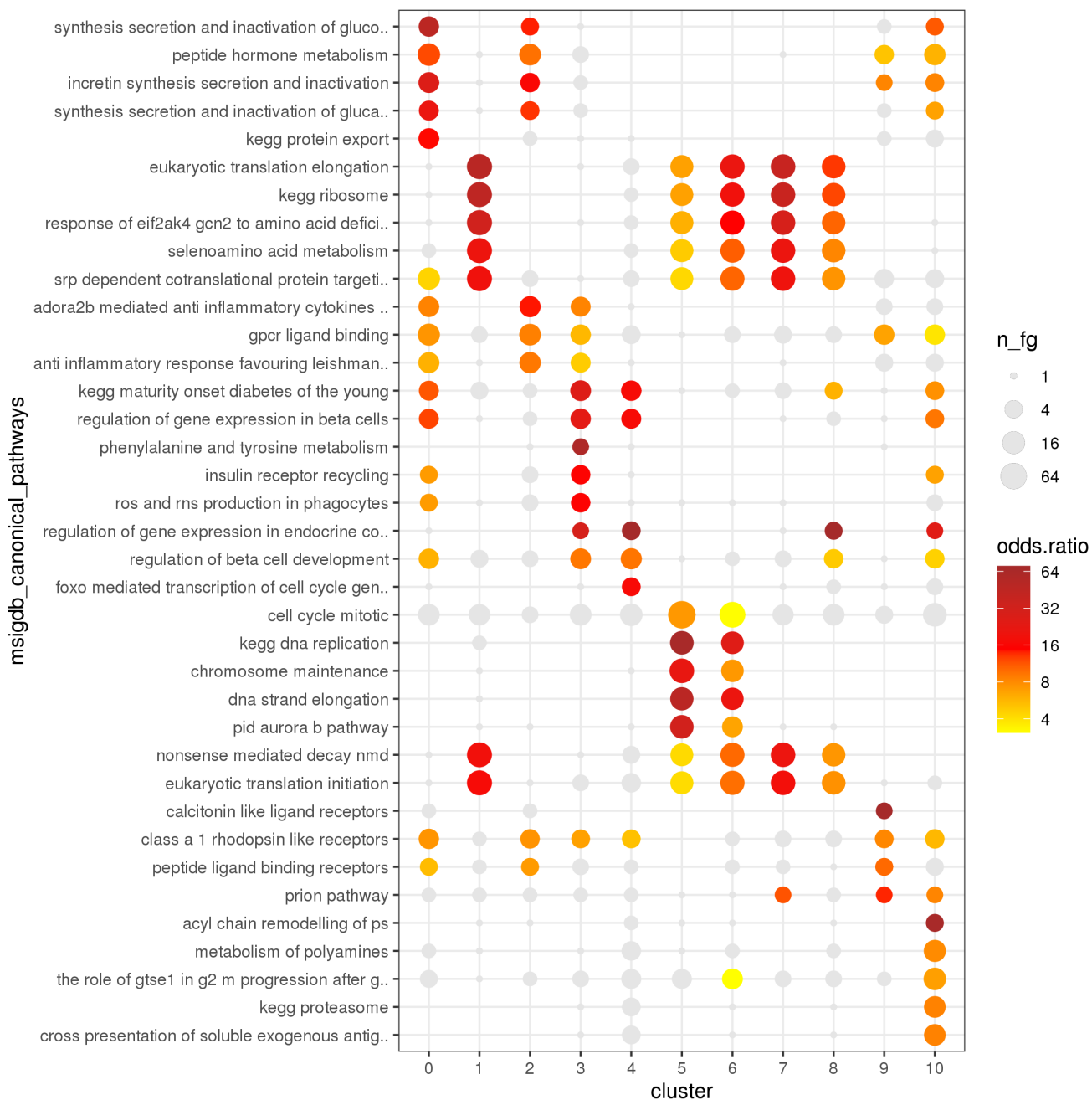


Figure 93: Heatmap of the top msigdb_canonical_pathways genesets

19.8 msigdb_tf_motifs

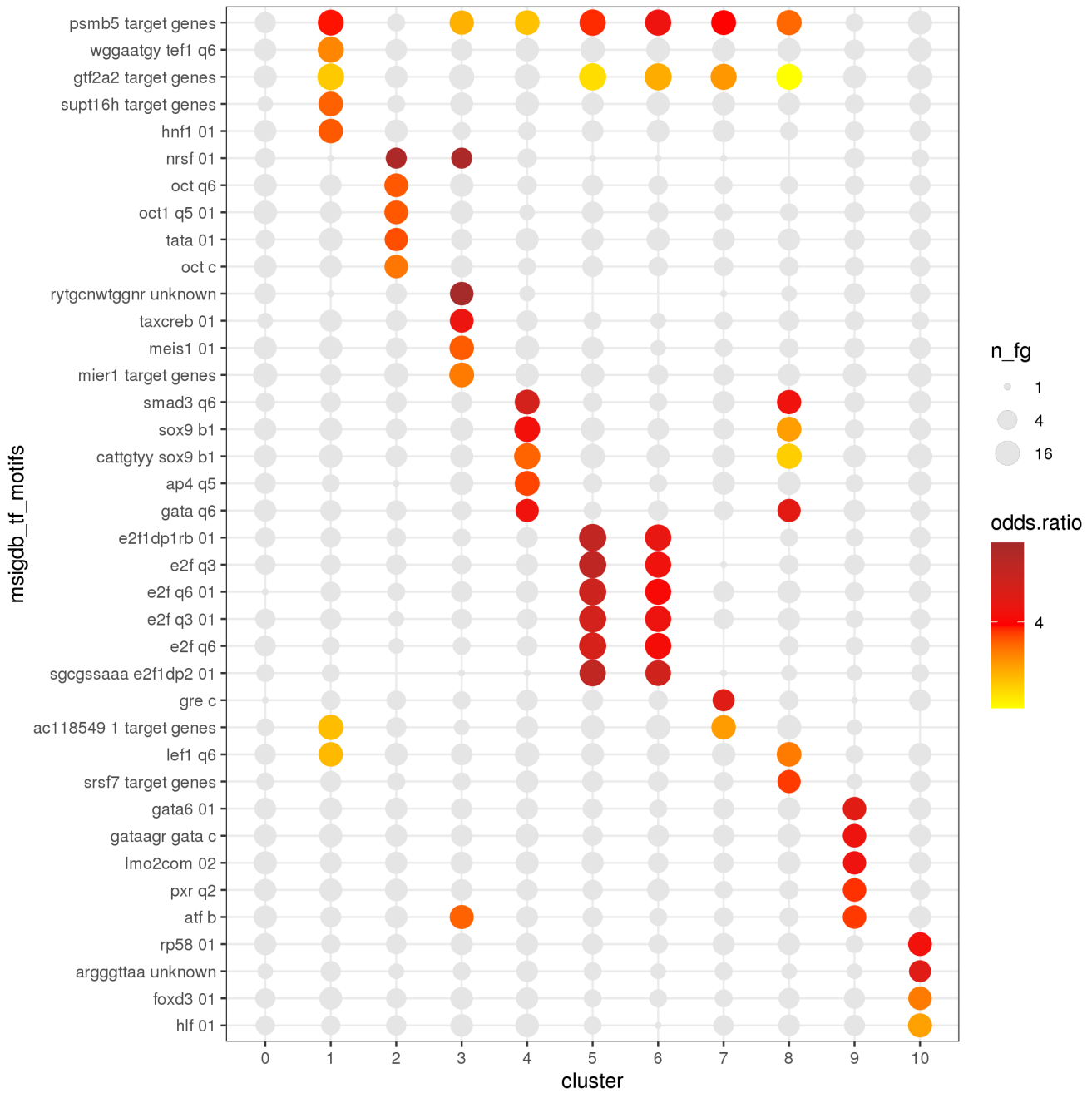


Figure 94: Heatmap of the top msigdb_tf_motifs genesets

19.9 msigdb_immunological_signatures

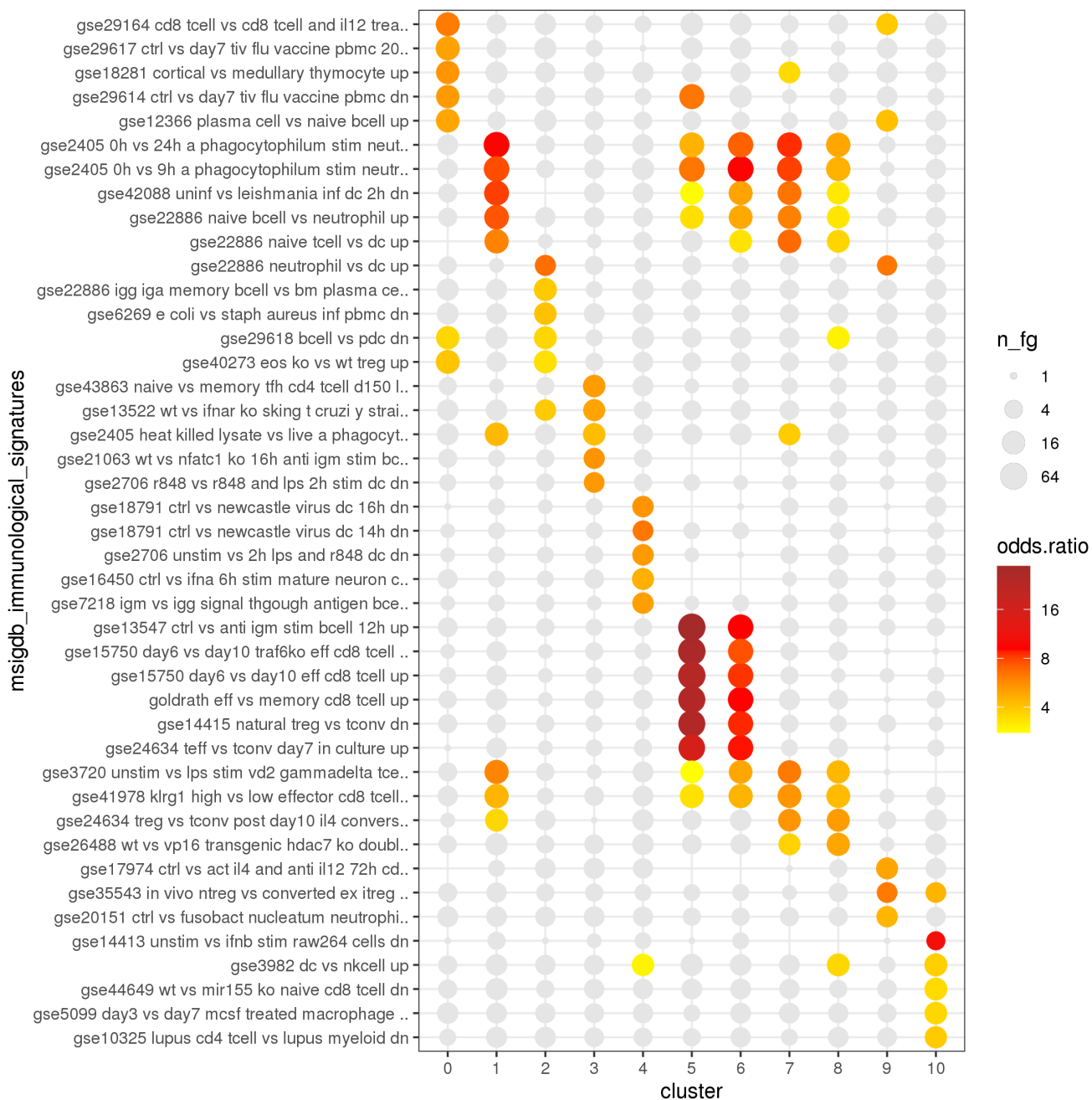


Figure 95: Heatmap of the top msigdb_immunological_signatures genesets

19.10 Top genesets by cluster

The table lists the top (filtered by nominal p value) genesets by cluster.

cluster	type	description	p.val	p.adj	n_fg	odds.ratio	n.clust
0	GO.BP	hormone transport	2.16e-17	3.8e-14	41	6.44	6
0	GO.BP	hormone secretion	8.67e-17	1.38e-13	40	6.32	6
0	GO.BP	regulation of hormone secretion	9.15e-16	1.32e-12	35	6.82	5
0	GO.BP	peptide hormone secretion	1.95e-15	2.39e-12	35	6.62	5
0	GO.BP	regulation of peptide hormone secretion	8.97e-15	8.37e-12	31	7.27	5
0	GO.MF	oligosaccharyl transferase activity	5.76e-10	1.15e-07	8	143	3
0	GO.MF	dolichyl-diphosphooligosaccharide-protein gly..	5.76e-10	1.15e-07	8	143	3
0	GO.MF	hormone activity	9.4e-07	5.38e-05	9	12.4	3
0	GO.MF	peptide disulfide oxidoreductase activity	1.66e-05	0.000497	6	17.8	2
0	GO.MF	proton-exporting ATPase activity	3.61e-05	0.000872	7	10.4	2
0	GO.CC	oligosaccharyltransferase complex	5.76e-10	2.53e-08	8	143	3
0	GO.CC	endoplasmic reticulum chaperone complex	9.59e-09	3.31e-07	8	47.5	2
0	GO.CC	secretory granule	4.26e-08	1.45e-06	25	4.18	4
0	GO.CC	smooth endoplasmic reticulum	4.56e-07	1.2e-05	10	11.2	2
0	GO.CC	endoplasmic reticulum lumen	1.74e-06	4.1e-05	11	7.9	3
0	KEGG	Protein processing in endoplasmic reticulum	1.95e-15	4.53e-13	35	6.62	3
0	KEGG	Oxidative phosphorylation	4.41e-09	5.11e-07	23	5.22	1
0	KEGG	Protein export	2.21e-08	1.92e-06	11	14.1	1
0	KEGG	Parkinson disease	7.36e-08	5.39e-06	30	3.49	2
0	KEGG	Maturity onset diabetes of the young	5.83e-07	3.38e-05	9	13.4	4
0	msigdb_reactome	peptide hormone metabolism	7.32e-11	2.23e-09	16	12.1	4
0	msigdb_reactome	asparagine n linked glycosylation	5.18e-09	1.28e-07	33	3.7	1
0	msigdb_reactome	synthesis secretion and inactivation of gluco..	9.59e-09	2.28e-07	8	47.5	3
0	msigdb_reactome	incretin synthesis secretion and inactivation	1.33e-08	3.09e-07	9	26.8	4
0	msigdb_reactome	gpcr ligand binding	1.26e-07	2.42e-06	14	7.44	5
0	msigdb_canonical_pathways	peptide hormone metabolism	7.32e-11	2.83e-09	16	12.1	4
0	msigdb_canonical_pathways	asparagine n linked glycosylation	5.18e-09	1.64e-07	33	3.7	1
0	msigdb_canonical_pathways	synthesis secretion and inactivation of gluco..	9.59e-09	2.9e-07	8	47.5	3
0	msigdb_canonical_pathways	incretin synthesis secretion and inactivation	1.33e-08	3.91e-07	9	26.8	4
0	msigdb_canonical_pathways	gpcr ligand binding	1.26e-07	3.11e-06	14	7.44	5
0	msigdb_immunological_signatures	gse29164 cd8 tcell vs cd8 tcell and il12 trea..	2.6e-10	5.25e-08	23	6.22	2
0	msigdb_immunological_signatures	gse29617 ctrl vs day7 tiv flu vaccine pbmc 20..	1.98e-09	3.58e-07	25	5.01	1
0	msigdb_immunological_signatures	gse12366 plasma cell vs naive bcell up	5.52e-08	7.67e-06	21	4.86	2
0	msigdb_immunological_signatures	gse18281 cortical vs medullary thymocyte up	5.55e-08	7.69e-06	19	5.41	2
0	msigdb_immunological_signatures	gse22886 naive bcell vs bm plasma cell dn	5.57e-08	7.7e-06	23	4.44	1
1	GO.BP	cytoplasmic translation	3.79e-15	4.01e-12	26	10.2	4
1	GO.BP	extracellular structure organization	1.45e-09	3.55e-07	26	4.95	5
1	GO.BP	regulation of epithelial cell proliferation	2.29e-09	5.27e-07	31	4.08	4
1	GO.BP	exocrine system development	1.48e-08	2.83e-06	12	12.8	3
1	GO.BP	extracellular matrix organization	1.52e-08	2.83e-06	22	5.11	5
1	GO.MF	structural constituent of ribosome	1.2e-39	2.4e-36	63	13.9	5
1	GO.MF	rRNA binding	2.72e-11	7.78e-09	21	8.33	5
1	GO.MF	cell adhesion molecule binding	3.79e-07	2.79e-05	19	4.72	3
1	GO.MF	sulfur compound binding	4.48e-07	2.99e-05	20	4.43	3
1	GO.MF	extracellular matrix structural constituent	6.47e-07	4.05e-05	12	7.99	4
1	GO.CC	cytosolic ribosome	2.37e-55	4.58e-52	63	36.6	5
1	GO.CC	cytosolic large ribosomal subunit	1.17e-38	7.52e-36	40	52.5	5
1	GO.CC	cytosolic part	6.88e-38	3.33e-35	67	11.3	5
1	GO.CC	ribosomal subunit	2.61e-35	6.31e-33	64	10.7	5
1	GO.CC	large ribosomal subunit	4.16e-23	5.37e-21	41	10.4	5
1	KEGG	Ribosome	2.22e-50	3.08e-47	69	20.3	5
1	KEGG	Glutathione metabolism	3.65e-08	2.82e-06	14	8.64	4
1	KEGG	Drug metabolism - cytochrome P450	1.89e-07	1.32e-05	10	13.3	2
1	KEGG	Tight junction	7.4e-07	3.96e-05	22	3.92	2
1	KEGG	Cell adhesion molecules	1.25e-06	6.01e-05	12	7.37	1
1	msigdb_reactome	eukaryotic translation elongation	6.19e-41	7.34e-38	43	49.2	5
1	msigdb_reactome	response of eif2ak4 gn2 to amino acid defici..	9.06e-36	7.17e-33	41	33.4	5
1	msigdb_reactome	selenoamino acid metabolism	5.94e-30	2.35e-27	40	20	5
1	msigdb_reactome	srp dependent cotranslational protein targeti..	2.4e-29	8.15e-27	40	18.9	6
1	msigdb_reactome	nonsense mediated decay nmd	4.73e-29	1.4e-26	40	18.4	5
1	msigdb_canonical_pathways	eukaryotic translation elongation	6.19e-41	1.07e-37	43	49.2	5
1	msigdb_canonical_pathways	kegg ribosome	1.48e-37	1.69e-34	40	45.5	5
1	msigdb_canonical_pathways	response of eif2ak4 gn2 to amino acid defici..	9.06e-36	7.79e-33	41	33.4	5
1	msigdb_canonical_pathways	selenoamino acid metabolism	5.94e-30	2.56e-27	40	20	5
1	msigdb_canonical_pathways	srp dependent cotranslational protein targeti..	2.4e-29	9.19e-27	40	18.9	6
1	msigdb_tf_motifs	psmb5 target genes	1.66e-07	2.86e-05	25	3.88	7
1	msigdb_tf_motifs	wggaatgy tef1 q6	1.99e-06	0.000298	27	3.14	1
1	msigdb_tf_motifs	gtf2a2 target genes	6.03e-06	0.000828	34	2.56	5
1	msigdb_tf_motifs	supt16h target genes	4.91e-05	0.00561	17	3.45	1
1	msigdb_tf_motifs	hnf1 01	6.75e-05	0.007	16	3.51	1

1	msigdb_immunological_signatures	gse2405 0h vs 24h a phagocytophilum stim neut..	2.94e-25	2.87e-22	48	9.4	5
1	msigdb_immunological_signatures	gse2405 0h vs 9h a phagocytophilum stim neutr..	9.12e-21	5.94e-18	44	7.73	5
1	msigdb_immunological_signatures	gse42088 uninfl vs leishmania inf dc 2h dn	5.54e-17	2.36e-14	34	8.11	5
1	msigdb_immunological_signatures	gse22886 naive bcell vs neutrophil up	1.61e-14	5.23e-12	30	7.52	5
1	msigdb_immunological_signatures	gse22886 naive tcell vs dc up	4.71e-11	1.04e-08	26	6	4
2	GO.BP	hormone transport	4.52e-09	9.83e-07	26	4.55	6
2	GO.BP	hormone secretion	1.57e-08	2.83e-06	25	4.39	6
2	GO.BP	regulation of hormone secretion	5.43e-07	5.98e-05	20	4.28	5
2	GO.BP	peptide hormone secretion	8.08e-07	8.17e-05	20	4.16	5
2	GO.BP	positive regulation of hormone secretion	8.9e-06	0.000591	13	5.13	4
2	GO.MF	hormone activity	7.88e-08	7.52e-06	9	17.6	3
2	GO.MF	G protein-coupled receptor binding	2.62e-06	0.000112	16	4.66	4
2	GO.MF	G-protein beta/gamma-subunit complex binding	0.000113	0.00218	5	15.2	3
2	GO.MF	amyloid-beta binding	0.000436	0.0059	6	7.54	1
2	GO.MF	cell-cell adhesion mediator activity	0.000702	0.00789	4	14.2	3
2	GO.CC	secretory granule	7.34e-09	2.63e-07	24	4.78	4
2	GO.CC	secretory granule membrane	1.81e-05	0.000326	9	7.45	4
2	GO.CC	heterotrimeric G-protein complex	1.84e-05	0.000326	6	16	3
2	GO.CC	GTPase complex	1.84e-05	0.000326	6	16	3
2	GO.CC	perikaryon	0.00011	0.00134	10	5.01	5
2	KEGG	GABAergic synapse	4.48e-06	0.000152	9	9.23	3
2	KEGG	Neuroactive ligand-receptor interaction	6.07e-06	0.000188	9	8.81	3
2	KEGG	Circadian entrainment	2.32e-05	0.000601	9	7.18	4
2	KEGG	Glutamatergic synapse	4.62e-05	0.000974	9	6.46	3
2	KEGG	Lysosome	8.11e-05	0.00147	13	4.02	2
2	msigdb_biocarta	gcr pathway	0.0116	0.0423	3	7.95	1
2	msigdb_reactome	adora2b mediated anti inflammatory cytokines ..	5.74e-08	1.21e-06	10	14.4	3
2	msigdb_reactome	peptide hormone metabolism	8.06e-08	1.65e-06	12	9.66	4
2	msigdb_reactome	gpcr ligand binding	2.02e-07	3.71e-06	12	8.69	5
2	msigdb_reactome	anti inflammatory response favouring leishman..	4.2e-07	7.03e-06	11	9.16	3
2	msigdb_reactome	regulation of insulin secretion	4.89e-06	6.21e-05	10	7.71	3
2	msigdb_canonical_pathways	adora2b mediated anti inflammatory cytokines ..	5.74e-08	1.53e-06	10	14.4	3
2	msigdb_canonical_pathways	peptide hormone metabolism	8.06e-08	2.1e-06	12	9.66	4
2	msigdb_canonical_pathways	gpcr ligand binding	2.02e-07	4.72e-06	12	8.69	5
2	msigdb_canonical_pathways	anti inflammatory response favouring leishman..	4.2e-07	8.88e-06	11	9.16	3
2	msigdb_canonical_pathways	regulation of insulin secretion	4.89e-06	7.66e-05	10	7.71	3
2	msigdb_tf_motifs	oct q6	0.000264	0.0208	13	3.52	1
2	msigdb_tf_motifs	oct1 q5 01	0.000264	0.0208	13	3.52	1
2	msigdb_tf_motifs	tata 01	0.000645	0.0393	11	3.59	1
2	msigdb_tf_motifs	nrsf 01	0.000708	0.0394	6	6.75	2
2	msigdb_tf_motifs	oct c	0.000776	0.0411	12	3.28	1
2	msigdb_immunological_signatures	gse22886 igg iga memory bcell vs bm plasma ce..	1.02e-05	0.000927	17	3.91	1
2	msigdb_immunological_signatures	gse6269 e coli vs staph aureus inf pbmc dn	3.62e-05	0.00291	14	4.11	1
2	msigdb_immunological_signatures	gse22886 neutrophil vs dc up	3.7e-05	0.00296	9	6.68	2
2	msigdb_immunological_signatures	gse29618 bcell vs pdc dn	4.02e-05	0.00316	16	3.63	3
2	msigdb_immunological_signatures	gse40273 eos ko vs wt treg up	4.9e-05	0.00376	17	3.4	2
3	GO.BP	hormone secretion	2.27e-12	1.2e-09	30	5.79	6
3	GO.BP	hormone transport	3.33e-12	1.65e-09	30	5.69	6
3	GO.BP	regulation of hormone secretion	7.09e-11	2.56e-08	25	6	5
3	GO.BP	peptide hormone secretion	7.46e-10	1.94e-07	24	5.49	5
3	GO.BP	regulation of peptide hormone secretion	2.12e-09	4.94e-07	21	6	5
3	GO.MF	G protein-coupled receptor binding	4.19e-07	2.89e-05	17	5.15	4
3	GO.MF	proton-exporting ATPase activity, phosphoryla..	8.13e-07	4.93e-05	6	40.8	2
3	GO.MF	syntaxin binding	5.54e-06	0.000222	11	6.67	2
3	GO.MF	proton-exporting ATPase activity	1.56e-05	0.000482	7	11.9	2
3	GO.MF	tubulin binding	4.89e-05	0.00113	21	2.95	2
3	GO.CC	secretory granule	4.35e-10	1.96e-08	26	5.19	4
3	GO.CC	exocytic vesicle	1.55e-05	0.000297	18	3.61	3
3	GO.CC	perikaryon	2.01e-05	0.000351	11	5.67	5
3	GO.CC	synaptic vesicle membrane	2.28e-05	0.000374	12	5.06	2
3	GO.CC	exocytic vesicle membrane	2.28e-05	0.000374	12	5.06	2
3	KEGG	Maturity onset diabetes of the young	3.2e-09	4.05e-07	10	22.9	4
3	KEGG	Synaptic vesicle cycle	1.55e-08	1.44e-06	12	11.8	4
3	KEGG	GABAergic synapse	1.67e-06	7.51e-05	9	10.9	3
3	KEGG	Circadian entrainment	4.73e-06	0.000155	9	9.24	4
3	KEGG	Glutamatergic synapse	1.54e-05	0.000427	9	7.7	3
3	msigdb_biocarta	bad pathway	0.0009	0.00907	5	8.47	2
3	msigdb_biocarta	pgc1a pathway	0.00718	0.0316	3	10.1	1
3	msigdb_biocarta	nos1 pathway	0.0131	0.0452	3	7.58	1
3	msigdb_reactome	regulation of gene expression in beta cells	1.14e-07	2.21e-06	8	23.4	4
3	msigdb_reactome	regulation of beta cell development	4.73e-06	6.1e-05	9	9.24	5
3	msigdb_reactome	insulin receptor recycling	2.37e-05	0.000231	6	15.3	3
3	msigdb_reactome	ros and rns production in phagocytes	2.37e-05	0.000231	6	15.3	2
3	msigdb_reactome	regulation of insulin secretion	2.59e-05	0.000248	9	7.11	3
3	msigdb_canonical_pathways	kegg maturity onset diabetes of the young	9.36e-09	2.85e-07	9	26.4	5

3	msigdb_canonical_pathways	regulation of gene expression in beta cells	1.14e-07	2.84e-06	8	23.4	4
3	msigdb_canonical_pathways	regulation of beta cell development	4.73e-06	7.5e-05	9	9.24	5
3	msigdb_canonical_pathways	insulin receptor recycling	2.37e-05	0.000288	6	15.3	3
3	msigdb_canonical_pathways	ros and rns production in phagocytes	2.37e-05	0.000288	6	15.3	2
3	msigdb_tf_motifs	rytgenwtggnr unknown	1.25e-06	0.000191	12	7.09	1
3	msigdb_tf_motifs	taxcreb 01	3.72e-05	0.00431	13	4.41	1
3	msigdb_tf_motifs	mier1 target genes	5.63e-05	0.00618	18	3.23	1
3	msigdb_tf_motifs	meis1 01	6.15e-05	0.00658	16	3.5	1
3	msigdb_tf_motifs	efc q6	8.54e-05	0.00872	16	3.39	1
3	msigdb_immunological_signatures	gse43863 naive vs memory tfh cd4 tcell d150 l..	3.86e-06	0.000383	14	5.19	1
3	msigdb_immunological_signatures	gse2405 heat killed lysate vs live a phagocyt..	2.62e-05	0.00219	14	4.26	3
3	msigdb_immunological_signatures	gse13522 wt vs ifnar ko sking t cruzi y strai..	2.71e-05	0.00224	12	4.95	2
3	msigdb_immunological_signatures	gse21063 wt vs nfatc1 ko 16h anti igm stim bc..	0.000148	0.00938	9	5.43	1
3	msigdb_immunological_signatures	gse29618 monocyte vs mdc day7 flu vaccine up	0.000163	0.0102	13	3.73	2
4	GO.BP	endocrine system development	2.46e-07	3.26e-05	15	6.39	6
4	GO.BP	pancreas development	1.21e-05	0.000751	11	6.2	7
4	GO.BP	endocrine pancreas development	1.53e-05	0.000881	9	7.98	5
4	GO.BP	negative regulation of neurogenesis	7.33e-05	0.00289	21	2.87	3
4	GO.BP	negative regulation of nervous system develop..	7.41e-05	0.0029	22	2.79	4
4	GO.MF	voltage-gated cation channel activity	1.1e-05	0.00038	8	10.3	2
4	GO.MF	serine-type endopeptidase inhibitor activity	1.64e-05	0.000497	8	9.61	6
4	GO.MF	endopeptidase regulator activity	0.000301	0.00461	10	4.43	5
4	GO.MF	voltage-gated ion channel activity	0.00051	0.00676	8	5.17	3
4	GO.MF	voltage-gated channel activity	0.00051	0.00676	8	5.17	3
4	GO.CC	presynaptic cytoskeleton	0.000184	0.00201	3	Inf	1
4	GO.CC	cortical cytoskeleton	0.000657	0.00573	11	3.64	1
4	GO.CC	anchored component of membrane	0.000694	0.00594	9	4.32	6
4	GO.CC	calyx of Held	0.00104	0.00801	4	13.3	1
4	GO.CC	intermediate filament	0.00148	0.0105	7	4.89	3
4	KEGG	Maturity onset diabetes of the young	3.24e-07	2.14e-05	9	15.2	4
4	KEGG	Transcriptional misregulation in cancer	4.39e-05	0.000953	14	4.09	2
4	KEGG	Epstein-Barr virus infection	5.52e-05	0.00108	17	3.41	3
4	KEGG	Hepatocellular carcinoma	0.000116	0.00196	15	3.49	3
4	KEGG	Chronic myeloid leukemia	0.000301	0.00428	10	4.43	2
4	msigdb_biocarta	p53 pathway	0.000204	0.0032	5	13.9	5
4	msigdb_biocarta	bad pathway	0.00205	0.0159	5	6.96	2
4	msigdb_biocarta	p53hypoxia pathway	0.0066	0.031	4	6.67	2
4	msigdb_biocarta	raccycd pathway	0.00842	0.036	5	4.64	1
4	msigdb_reactome	regulation of gene expression in endocrine co..	5.89e-07	9.71e-06	5	Inf	4
4	msigdb_reactome	regulation of beta cell development	1.69e-06	2.48e-05	10	9.38	5
4	msigdb_reactome	regulation of gene expression in beta cells	4.51e-06	5.85e-05	7	16.8	4
4	msigdb_reactome	tp53 regulates transcription of cell cycle ge..	6.6e-05	0.000574	9	6.31	3
4	msigdb_reactome	foxo mediated transcription of cell cycle gen..	0.000117	0.000918	5	16.7	1
4	msigdb_canonical_pathways	regulation of gene expression in endocrine co..	5.89e-07	1.21e-05	5	Inf	4
4	msigdb_canonical_pathways	kegg maturity onset diabetes of the young	8.97e-07	1.8e-05	8	16.8	5
4	msigdb_canonical_pathways	regulation of beta cell development	1.69e-06	3.03e-05	10	9.38	5
4	msigdb_canonical_pathways	pid p53 downstream pathway	3.11e-06	5.1e-05	15	5.01	3
4	msigdb_canonical_pathways	regulation of gene expression in beta cells	4.51e-06	7.19e-05	7	16.8	4
4	msigdb_tf_motifs	sox9 b1	6.75e-08	1.23e-05	24	4.24	2
4	msigdb_tf_motifs	smad3 q6	2.03e-07	3.33e-05	18	5.23	2
4	msigdb_tf_motifs	cattgtty sox9 b1	2.86e-07	4.48e-05	28	3.42	2
4	msigdb_tf_motifs	ap4 q5	1.49e-05	0.00183	18	3.66	1
4	msigdb_tf_motifs	pitx2 q2	0.000146	0.0134	14	3.59	1
4	msigdb_immunological_signatures	gse16450 ctrl vs ifna 6h stim mature neuron c..	5.68e-05	0.00423	12	4.61	1
4	msigdb_immunological_signatures	gse26928 centr memory vs cxcr5 pos cd4 tcell ..	7.13e-05	0.00509	13	4.15	1
4	msigdb_immunological_signatures	gse18791 ctrl vs newcastle virus dc 16h dn	7.39e-05	0.00526	10	5.44	1
4	msigdb_immunological_signatures	gse29618 pdc vs mdc up	9.11e-05	0.00631	16	3.4	4
4	msigdb_immunological_signatures	gse2706 unstim vs 2h lps and r848 dc dn	9.22e-05	0.00637	10	5.27	1
5	GO.BP	chromosome segregation	4.21e-35	6.68e-31	77	8.08	2
5	GO.BP	mitotic nuclear division	6.46e-26	5.13e-22	62	6.89	2
5	GO.BP	nuclear chromosome segregation	1.83e-24	9.71e-21	57	7.08	1
5	GO.BP	mitotic sister chromatid segregation	1.73e-23	6.86e-20	47	8.78	2
5	GO.BP	sister chromatid segregation	6.01e-23	1.91e-19	50	7.69	2
5	GO.MF	catalytic activity, acting on DNA	2.9e-10	7.26e-08	32	4.43	2
5	GO.MF	single-stranded DNA-dependent ATP-dependent D..	1.29e-09	2.15e-07	11	25.2	2
5	GO.MF	single-stranded DNA-dependent ATPase activity	1.29e-09	2.15e-07	11	25.2	2
5	GO.MF	ATP-dependent DNA helicase activity	3.1e-09	3.89e-07	11	21.6	2
5	GO.MF	ATP-dependent helicase activity	3.1e-09	3.89e-07	11	21.6	2
5	GO.CC	chromosome, centromeric region	9.82e-31	2.11e-28	59	9.76	2
5	GO.CC	condensed chromosome	5.05e-26	7.52e-24	45	11.5	1
5	GO.CC	kinetochore	1.51e-22	1.82e-20	43	9.43	2
5	GO.CC	condensed chromosome, centromeric region	1.84e-18	1.62e-16	24	21.1	1
5	GO.CC	spindle	6.81e-16	5.27e-14	51	4.63	1
5	KEGG	DNA replication	3.95e-24	1.37e-21	26	45.9	2
5	KEGG	Cell cycle	7.17e-15	1.43e-12	34	6.72	2

5	KEGG	Mismatch repair	1.24e-12	1.91e-10	14	32.3	2
5	KEGG	Ribosome	3.25e-11	4.52e-09	33	4.76	5
5	KEGG	Nucleotide excision repair	7.11e-09	7.61e-07	15	9.43	2
5	msigdb_biocarta	mcm pathway	4.87e-07	3.43e-05	9	15.4	2
5	msigdb_biocarta	cellcycle pathway	6.77e-06	0.000305	8	12.2	4
5	msigdb_biocarta	sm pathway	3.58e-05	0.00101	6	16.4	2
5	msigdb_biocarta	ranms pathway	7.2e-05	0.00169	5	22.7	1
5	msigdb_biocarta	efp pathway	0.000118	0.00238	6	11.7	2
5	msigdb_reactome	cell cycle mitotic	2.09e-49	4.95e-46	122	7.24	2
5	msigdb_reactome	cell cycle checkpoints	2.26e-28	5.35e-26	68	6.96	3
5	msigdb_reactome	mitotic prometaphase	2.56e-27	5.51e-25	58	8.19	2
5	msigdb_reactome	resolution of sister chromatid cohesion	4.28e-27	7.82e-25	46	12	2
5	msigdb_reactome	chromosome maintenance	1.61e-26	2.54e-24	35	21.8	2
5	msigdb_canonical_pathways	cell cycle mitotic	2.09e-49	7.19e-46	122	7.24	2
5	msigdb_canonical_pathways	cell cycle checkpoints	2.26e-28	6.47e-26	68	6.96	3
5	msigdb_canonical_pathways	mitotic prometaphase	2.56e-27	6.77e-25	58	8.19	2
5	msigdb_canonical_pathways	resolution of sister chromatid cohesion	4.28e-27	9.83e-25	46	12	2
5	msigdb_canonical_pathways	chromosome maintenance	1.61e-26	3.25e-24	35	21.8	2
5	msigdb_tf_motifs	e2f1dp1rb 01	2.76e-16	1.86e-12	42	5.86	2
5	msigdb_tf_motifs	e2f q3	8.12e-16	2.73e-12	40	5.98	2
5	msigdb_tf_motifs	e2f q6 01	9.09e-15	2.04e-11	40	5.46	2
5	msigdb_tf_motifs	e2f q6	5.13e-14	7.31e-11	40	5.11	2
5	msigdb_tf_motifs	e2f q3 01	6.32e-14	7.31e-11	39	5.21	2
5	msigdb_immunological_signatures	gse15750 day6 vs day10 traf6ko eff cd8 tcell ..	3.99e-76	1.71e-71	98	27.3	2
5	msigdb_immunological_signatures	gse13547 ctrl vs anti igm stim bcell 12h up	7.36e-75	1.58e-70	94	29.6	2
5	msigdb_immunological_signatures	gse15750 day6 vs day10 eff cd8 tcell up	1.79e-69	2.57e-65	93	24	2
5	msigdb_immunological_signatures	goldrath eff vs memory cd8 tcell up	4e-62	4.3e-58	82	24.7	2
5	msigdb_immunological_signatures	gse14415 natural treg vs tconv dn	3.01e-60	2.59e-56	79	25.1	2
6	GO.BP	DNA replication	2.48e-12	1.27e-09	38	4.57	2
6	GO.BP	protein-DNA complex assembly	1.45e-08	2.81e-06	23	4.88	2
6	GO.BP	ribosome biogenesis	1.63e-08	2.9e-06	37	3.24	4
6	GO.BP	DNA-dependent DNA replication	3.96e-08	6.55e-06	22	4.78	2
6	GO.BP	protein-DNA complex subunit organization	8.95e-08	1.37e-05	25	4.02	2
6	GO.MF	structural constituent of ribosome	5.25e-28	3.51e-25	52	9.87	5
6	GO.MF	rRNA binding	2.03e-11	6.77e-09	21	8.47	5
6	GO.MF	single-stranded DNA-dependent ATP-dependent D..	6.53e-08	6.54e-06	9	20.8	2
6	GO.MF	single-stranded DNA-dependent ATPase activity	6.53e-08	6.54e-06	9	20.8	2
6	GO.MF	ATP-dependent DNA helicase activity	1.31e-07	1.07e-05	9	18.2	2
6	GO.CC	cytosolic ribosome	4e-37	1.29e-34	50	20.1	5
6	GO.CC	ribosomal subunit	1.02e-28	1.79e-26	57	8.91	5
6	GO.CC	cytosolic part	2.87e-26	4.64e-24	55	8.12	5
6	GO.CC	cytosolic large ribosomal subunit	4.37e-21	4.98e-19	28	18.9	5
6	GO.CC	cytosolic small ribosomal subunit	1.19e-18	1.1e-16	23	22.5	5
6	KEGG	Ribosome	8.12e-34	3.76e-31	55	12.9	5
6	KEGG	DNA replication	2.63e-14	4.58e-12	18	19.8	2
6	KEGG	Drug metabolism - other enzymes	4.87e-07	2.94e-05	13	7.29	3
6	KEGG	Glutathione metabolism	1.96e-06	8.27e-05	12	6.96	4
6	KEGG	Cell cycle	3.6e-06	0.000129	19	3.94	2
6	msigdb_biocarta	mcm pathway	1.31e-07	1.85e-05	9	18.2	2
6	msigdb_biocarta	sm pathway	0.000244	0.00344	5	13.4	2
6	msigdb_biocarta	btg2 pathway	0.000368	0.00472	4	21.3	2
6	msigdb_biocarta	efp pathway	0.000616	0.00724	5	10	2
6	msigdb_biocarta	cellcycle pathway	0.0018	0.0149	5	7.29	4
6	msigdb_reactome	eukaryotic translation elongation	9.79e-25	1.29e-22	32	20.9	5
6	msigdb_reactome	response of eif2ak4 gcn2 to amino acid defici..	7.12e-20	6.03e-18	29	14.8	5
6	msigdb_reactome	selenoamino acid metabolism	2.79e-17	1.84e-15	29	10.8	5
6	msigdb_reactome	srp dependent cotranslational protein targeti..	6.58e-17	4.22e-15	29	10.4	6
6	msigdb_reactome	eukaryotic translation initiation	7.51e-17	4.69e-15	30	9.72	5
6	msigdb_canonical_pathways	eukaryotic translation elongation	9.79e-25	1.69e-22	32	20.9	5
6	msigdb_canonical_pathways	kegg ribosome	9.11e-22	1.13e-19	29	18.8	5
6	msigdb_canonical_pathways	response of eif2ak4 gcn2 to amino acid defici..	7.12e-20	7.42e-18	29	14.8	5
6	msigdb_canonical_pathways	selenoamino acid metabolism	2.79e-17	2.34e-15	29	10.8	5
6	msigdb_canonical_pathways	srp dependent cotranslational protein targeti..	6.58e-17	5.4e-15	29	10.4	6
6	msigdb_tf_motifs	e2f1dp1rb 01	5.17e-10	1.94e-07	30	4.52	2
6	msigdb_tf_motifs	sgcgssaaa e2f1dp2 01	1.82e-09	6.13e-07	24	5.3	2
6	msigdb_tf_motifs	e2f q3 01	3.06e-09	9.8e-07	28	4.4	2
6	msigdb_tf_motifs	psmb5 target genes	3.65e-09	1.12e-06	28	4.36	7
6	msigdb_tf_motifs	e2f q6	5.46e-09	1.6e-06	29	4.13	2
6	msigdb_immunological_signatures	gse2405 0h vs 9h a phagocytophilum stim neutr..	2.32e-26	2.56e-23	50	9.45	5
6	msigdb_immunological_signatures	gse13547 ctrl vs anti igm stim bcell 12h up	6.05e-23	4.56e-20	44	9.08	2
6	msigdb_immunological_signatures	goldrath eff vs memory cd8 tcell up	1.86e-22	1.38e-19	43	9.07	2
6	msigdb_immunological_signatures	gse24634 teff vs tconv day7 in culture up	7.64e-21	5.05e-18	40	8.93	2
6	msigdb_immunological_signatures	gse15750 day6 vs day10 eff cd8 tcell up	1.44e-20	9e-18	41	8.4	2
7	GO.BP	cytoplasmic translation	3.33e-11	1.29e-08	19	9.03	4
7	GO.BP	ribosome assembly	1.55e-08	2.83e-06	14	8.8	4

7	GO.BP	ribosomal small subunit biogenesis	8.08e-08	1.26e-05	13	8.36	4
7	GO.BP	extracellular structure organization	7.32e-07	7.49e-05	18	4.6	5
7	GO.BP	extracellular matrix organization	8.86e-07	8.73e-05	16	5.1	5
7	GO.MF	structural constituent of ribosome	5.2e-37	5.21e-34	53	15.7	5
7	GO.MF	rRNA binding	3.04e-12	1.22e-09	19	10.7	5
7	GO.MF	peptidase inhibitor activity	1.34e-07	1.07e-05	13	7.93	6
7	GO.MF	serine-type endopeptidase inhibitor activity	3.91e-07	2.79e-05	9	13.2	6
7	GO.MF	endopeptidase inhibitor activity	5.25e-07	3.39e-05	12	7.69	5
7	GO.CC	cytosolic ribosome	1.9e-51	1.84e-48	54	38.2	5
7	GO.CC	cytosolic part	2.01e-37	7.77e-35	57	13.9	5
7	GO.CC	ribosomal subunit	2.77e-36	7.67e-34	56	13.5	5
7	GO.CC	cytosolic large ribosomal subunit	3.5e-30	6.77e-28	31	37.4	5
7	GO.CC	cytosolic small ribosomal subunit	1.14e-23	1.57e-21	24	37.1	5
7	KEGG	Ribosome	1.24e-47	8.62e-45	59	23	5
7	KEGG	PI3K-Akt signaling pathway	0.000103	0.00177	16	3.31	2
7	KEGG	Glutathione metabolism	0.000202	0.00303	8	5.86	4
7	KEGG	Glioma	0.000242	0.00357	8	5.68	3
7	KEGG	Small cell lung cancer	0.000473	0.00587	8	5.06	3
7	msigdb_biocarta	p53 pathway	0.00113	0.0107	4	11.6	5
7	msigdb_biocarta	cellcycle pathway	0.00214	0.0159	4	9.27	4
7	msigdb_biocarta	prion pathway	0.00503	0.0284	3	11.6	3
7	msigdb_biocarta	tel pathway	0.00696	0.0316	3	9.91	1
7	msigdb_biocarta	chemical pathway	0.012	0.0423	3	7.71	1
7	msigdb_reactome	eukaryotic translation elongation	2.99e-34	1.77e-31	35	39.1	5
7	msigdb_reactome	response of eif2ak4 gcn2 to amino acid defici..	3.84e-30	1.82e-27	33	30.1	5
7	msigdb_reactome	nonsense mediated decay nmd	3.52e-27	6.95e-25	34	20.2	5
7	msigdb_reactome	selenoamino acid metabolism	1.41e-26	2.39e-24	33	20.5	5
7	msigdb_reactome	srp dependent cotranslational protein targeti..	4.12e-26	6.12e-24	33	19.6	6
7	msigdb_canonical_pathways	eukaryotic translation elongation	2.99e-34	2.06e-31	35	39.1	5
7	msigdb_canonical_pathways	kegg ribosome	3.26e-32	1.87e-29	33	38.3	5
7	msigdb_canonical_pathways	response of eif2ak4 gcn2 to amino acid defici..	3.84e-30	1.89e-27	33	30.1	5
7	msigdb_canonical_pathways	nonsense mediated decay nmd	3.52e-27	8.64e-25	34	20.2	5
7	msigdb_canonical_pathways	selenoamino acid metabolism	1.41e-26	3.03e-24	33	20.5	5
7	msigdb_tf_motifs	gtf2a2 target genes	3.83e-06	0.000561	27	2.99	5
7	msigdb_tf_motifs	psmb5 target genes	5.33e-06	0.000747	18	3.92	7
7	msigdb_tf_motifs	acl18549 1 target genes	0.000339	0.0246	16	2.94	2
7	msigdb_tf_motifs	gre c	0.000641	0.0393	8	4.8	1
7	msigdb_immunological_signatures	gse2405 0h vs 24h a phagocytophilum stim neut..	4.82e-18	2.25e-15	34	8.55	5
7	msigdb_immunological_signatures	gse2405 0h vs 9h a phagocytophilum stim neutr..	2.09e-17	9.45e-15	34	8.06	5
7	msigdb_immunological_signatures	gse42088 uninif vs leishmania inf dc 2h dn	2.29e-10	4.65e-08	22	6.47	5
7	msigdb_immunological_signatures	gse22886 naive tcell vs dc up	2.54e-10	5.15e-08	21	6.81	4
7	msigdb_immunological_signatures	gse41978 klr1 high vs low effector cd8 tcell..	3.93e-09	6.81e-07	22	5.41	5
8	GO.BP	mitotic cell cycle arrest	2.54e-06	0.000218	5	88.8	3
8	GO.BP	cytoplasmic translation	9.83e-06	0.000645	14	4.77	4
8	GO.BP	B cell differentiation	1.48e-05	0.000866	11	5.99	2
8	GO.BP	extracellular matrix organization	4.64e-05	0.00205	15	3.81	5
8	GO.BP	gliogenesis	4.93e-05	0.00215	20	3.05	4
8	GO.MF	structural constituent of ribosome	9.5e-14	4.76e-11	34	5.84	5
8	GO.MF	rRNA binding	2.07e-09	3.19e-07	18	7.47	5
8	GO.MF	serine-type endopeptidase inhibitor activity	1.83e-07	1.41e-05	10	12.8	6
8	GO.MF	large ribosomal subunit rRNA binding	1.73e-06	8.23e-05	6	35.6	4
8	GO.MF	peptidase inhibitor activity	6.03e-06	0.000237	12	6	6
8	GO.CC	cytosolic ribosome	4.25e-20	4.32e-18	34	10.8	5
8	GO.CC	cytosolic part	3.75e-14	2.53e-12	38	5.36	5
8	GO.CC	cytosolic large ribosomal subunit	1.13e-11	6.84e-10	19	10.2	5
8	GO.CC	ribosomal subunit	1.75e-11	9.7e-10	34	4.67	5
8	GO.CC	cytosolic small ribosomal subunit	7.16e-11	3.85e-09	16	12.1	5
8	KEGG	Ribosome	1.97e-19	5.49e-17	39	8.14	5
8	KEGG	Hepatocellular carcinoma	0.000292	0.00419	14	3.32	3
8	KEGG	Melanoma	0.000348	0.00472	8	5.49	3
8	KEGG	MicroRNAs in cancer	0.000349	0.00472	13	3.44	5
8	KEGG	p53 signaling pathway	0.000384	0.00504	9	4.73	4
8	msigdb_biocarta	p53 pathway	8.65e-06	0.000305	6	21.3	5
8	msigdb_biocarta	cellcycle pathway	0.0054	0.0284	4	7.08	4
8	msigdb_biocarta	g1 pathway	0.00901	0.0362	4	5.9	3
8	msigdb_reactome	eukaryotic translation elongation	4.29e-16	2.42e-14	24	13.1	5
8	msigdb_reactome	response of eif2ak4 gcn2 to amino acid defici..	2.21e-13	9.91e-12	22	10.4	5
8	msigdb_reactome	selenoamino acid metabolism	2.31e-12	8.43e-11	23	8.34	5
8	msigdb_reactome	eukaryotic translation initiation	3.28e-12	1.14e-10	24	7.67	5
8	msigdb_reactome	srp dependent cotranslational protein targeti..	3.42e-11	1.1e-09	22	7.51	6
8	msigdb_canonical_pathways	eukaryotic translation elongation	4.29e-16	3.14e-14	24	13.1	5
8	msigdb_canonical_pathways	kegg ribosome	1.76e-14	1.16e-12	22	12.3	5
8	msigdb_canonical_pathways	response of eif2ak4 gcn2 to amino acid defici..	2.21e-13	1.23e-11	22	10.4	5
8	msigdb_canonical_pathways	selenoamino acid metabolism	2.31e-12	1.07e-10	23	8.34	5
8	msigdb_canonical_pathways	eukaryotic translation initiation	3.28e-12	1.46e-10	24	7.67	5

8	msigdb_tf_motifs	psmb5 target genes	1.28e-05	0.00162	20	3.4	7
8	msigdb_tf_motifs	smad3 q6	1.38e-05	0.00172	15	4.3	2
8	msigdb_tf_motifs	lef1 q6	9.19e-05	0.00923	17	3.24	2
8	msigdb_tf_motifs	gata q6	9.48e-05	0.00938	11	4.7	2
8	msigdb_tf_motifs	sox9 b1	0.000267	0.0208	17	2.92	2
8	msigdb_immunological_signatures	gse2405 0h vs 24h a phagocytophilum stim neut..	4.13e-10	8.11e-08	28	4.87	5
8	msigdb_immunological_signatures	gse2405 0h vs 9h a phagocytophilum stim neut..	1.41e-09	2.6e-07	28	4.56	5
8	msigdb_immunological_signatures	gse41978 klr1 high vs low effector cd8 tcell..	8.82e-08	1.17e-05	23	4.31	5
8	msigdb_immunological_signatures	gse26488 wt vs vp16 transgenic hdac7 ko doubl..	4.28e-07	5.06e-05	18	4.89	2
8	msigdb_immunological_signatures	gse24634 treg vs tcon post day10 il4 convers..	4.79e-07	5.59e-05	17	5.15	3
9	GO.BP	regulation of G protein-coupled receptor sign..	6.93e-07	7.19e-05	11	8.29	4
9	GO.BP	peptide hormone processing	1.1e-05	0.000695	5	28.1	3
9	GO.BP	response to metal ion	3.56e-05	0.00169	14	4.07	2
9	GO.BP	calcitonin family receptor signaling pathway	4.23e-05	0.00193	3	Inf	1
9	GO.BP	amylin receptor signaling pathway	4.23e-05	0.00193	3	Inf	1
9	GO.MF	monooxygenase activity	8.03e-05	0.00163	6	10.6	3
9	GO.MF	ketosteroid monooxygenase activity	0.000165	0.00278	3	83.3	3
9	GO.MF	hormone activity	0.000646	0.00789	5	8.77	3
9	GO.MF	oxidoreductase activity, acting on paired don..	0.000948	0.00974	8	4.43	1
9	GO.MF	alditol:NADP+ 1-oxidoreductase activity	0.00208	0.0168	3	16.7	3
9	GO.CC	endoplasmic reticulum lumen	1.44e-07	4.49e-06	10	11.9	3
9	GO.CC	secretory granule	6.5e-07	1.66e-05	18	4.61	4
9	GO.CC	smooth endoplasmic reticulum	1.65e-05	0.00031	7	11	2
9	GO.CC	endoplasmic reticulum chaperone complex	0.000397	0.00377	4	16	2
9	GO.CC	perikaryon	0.000665	0.00577	8	4.71	5
9	KEGG	Protein processing in endoplasmic reticulum	6.82e-05	0.00128	15	3.6	3
9	KEGG	Neuroactive ligand-receptor interaction	0.000727	0.00839	6	6.49	3
9	KEGG	Circadian entrainment	0.000863	0.00968	6	6.25	4
9	KEGG	Oxytocin signaling pathway	0.000948	0.0105	8	4.43	2
9	KEGG	Various types of N-glycan biosynthesis	0.00181	0.017	5	6.68	2
9	msigdb_biocarta	prion pathway	0.00304	0.0214	3	13.9	3
9	msigdb_reactome	calcitonin like ligand receptors	4.23e-05	0.000383	3	Inf	1
9	msigdb_reactome	gpcr ligand binding	8.43e-05	0.00071	8	6.66	5
9	msigdb_reactome	g alpha i signalling events	0.000394	0.00276	10	4.12	4
9	msigdb_reactome	class a 1 rhodopsin like receptors	0.000813	0.00512	5	8.25	6
9	msigdb_reactome	peptide ligand binding receptors	0.00147	0.00854	4	10.2	4
9	msigdb_canonical_pathways	calcitonin like ligand receptors	4.23e-05	0.000474	3	Inf	1
9	msigdb_canonical_pathways	gpcr ligand binding	8.43e-05	0.000869	8	6.66	5
9	msigdb_canonical_pathways	g alpha i signalling events	0.000394	0.00331	10	4.12	4
9	msigdb_canonical_pathways	class a 1 rhodopsin like receptors	0.000813	0.00603	5	8.25	6
9	msigdb_canonical_pathways	peptide ligand binding receptors	0.00147	0.00985	4	10.2	3
9	msigdb_tf_motifs	gata6 01	3.11e-05	0.00367	12	4.77	1
9	msigdb_tf_motifs	gataagr gata c	6.26e-05	0.00658	12	4.4	1
9	msigdb_tf_motifs	lmo2com 02	0.000134	0.0127	11	4.36	1
9	msigdb_tf_motifs	pxr q2	0.000233	0.0194	12	3.77	1
9	msigdb_tf_motifs	atf b	0.000256	0.0207	12	3.73	2
9	msigdb_immunological_signatures	gse17974 ctrl vs act il4 and anti il12 72h cd..	2.43e-05	0.00203	12	4.91	1
9	msigdb_immunological_signatures	gse12366 plasma cell vs naive bcell up	9.63e-05	0.00656	12	4.18	2
9	msigdb_immunological_signatures	gse35543 in vivo ntreg vs converted ex itreg ..	0.000119	0.00787	8	6.29	2
9	msigdb_immunological_signatures	gse20151 ctrl vs fusobact nucleatum neutrophi..	0.000229	0.0134	10	4.44	1
9	msigdb_immunological_signatures	gse22886 neutrophil vs dc up	0.000294	0.0163	7	6.37	2
10	GO.BP	negative regulation of cell activation	1.01e-05	0.000657	12	5.68	2
10	GO.BP	positive regulation of peptide hormone secret..	1.28e-05	0.000781	12	5.52	4
10	GO.BP	positive regulation of insulin secretion	1.47e-05	0.000866	11	6.03	4
10	GO.BP	hormone secretion	1.74e-05	0.000976	22	3.11	6
10	GO.BP	negative regulation of leukocyte activation	1.88e-05	0.00103	11	5.85	2
10	GO.MF	G protein-coupled receptor binding	0.000124	0.0023	15	3.46	4
10	GO.MF	GTP binding	0.000543	0.00693	23	2.32	2
10	GO.MF	cell-cell adhesion mediator activity	0.000594	0.00742	4	16.8	3
10	GO.MF	ketosteroid monooxygenase activity	0.000693	0.00789	3	50.2	3
10	GO.MF	purine nucleoside binding	0.000836	0.00886	23	2.24	2
10	GO.CC	proteasome regulatory particle	1.8e-09	7.43e-08	11	20.8	2
10	GO.CC	proteasome accessory complex	6.83e-09	2.49e-07	11	17	2
10	GO.CC	proteasome complex	1.34e-07	4.24e-06	15	6.77	1
10	GO.CC	endopeptidase complex	1.76e-07	5.39e-06	15	6.6	1
10	GO.CC	myelin sheath	1.01e-06	2.45e-05	24	3.56	6
10	KEGG	Proteasome	1.17e-08	1.16e-06	14	9.6	2
10	KEGG	Spinocerebellar ataxia	2.76e-08	2.26e-06	20	5.41	2
10	KEGG	Prion disease	2.77e-06	0.000107	27	3.07	2
10	KEGG	Huntington disease	3.08e-05	0.000764	28	2.58	2
10	KEGG	Parkinson disease	4.48e-05	0.000958	25	2.68	2
10	msigdb_biocarta	no1 pathway	0.0035	0.0215	4	8.39	1
10	msigdb_biocarta	p53hypoxia pathway	0.00647	0.031	4	6.71	2
10	msigdb_biocarta	prion pathway	0.0118	0.0423	3	8.37	3
10	msigdb_reactome	the role of gtse1 in g2 m progression after g..	4.37e-08	9.35e-07	16	6.88	2

10	msigdb_reactome	metabolism of polyamines	6.84e-08	1.41e-06	14	8	1
10	msigdb_reactome	abc transporter disorders	1.01e-07	2e-06	15	6.96	1
10	msigdb_reactome	disorders of transmembrane transporters	1.21e-07	2.34e-06	19	5.13	1
10	msigdb_reactome	defective cftr causes cystic fibrosis	2.33e-07	4.22e-06	14	7.05	1
10	msigdb_canonical_pathways	the role of gtse1 in g2 m progression after g..	4.37e-08	1.18e-06	16	6.88	2
10	msigdb_canonical_pathways	metabolism of polyamines	6.84e-08	1.8e-06	14	8	1
10	msigdb_canonical_pathways	abc transporter disorders	1.01e-07	2.56e-06	15	6.96	1
10	msigdb_canonical_pathways	disorders of transmembrane transporters	1.21e-07	3.01e-06	19	5.13	1
10	msigdb_canonical_pathways	defective cftr causes cystic fibrosis	2.33e-07	5.38e-06	14	7.05	1
10	msigdb_tf_motifs	rp58 01	5.69e-05	0.00618	13	4.26	1
10	msigdb_tf_motifs	argggtaa unknown	0.00074	0.0405	8	4.83	1
10	msigdb_tf_motifs	foxd3 01	0.000945	0.0474	12	3.23	1
10	msigdb_tf_motifs	hlf 01	0.00095	0.0474	14	2.9	1
10	msigdb_immunological_signatures	gse3982 dc vs nkcell up	1.72e-05	0.00149	17	3.79	3
10	msigdb_immunological_signatures	gse17721 ctrl vs gardiquimod 24h bmec dn	5.78e-05	0.00431	18	3.25	1
10	msigdb_immunological_signatures	gse44649 wt vs mir155 ko naive cd8 tcell dn	6.56e-05	0.00476	16	3.51	1
10	msigdb_immunological_signatures	gse5099 day3 vs day7 mcsf treated macrophage ..	8.28e-05	0.00578	15	3.61	1
10	msigdb_immunological_signatures	gse29618 bcell vs mdc day7 flu vaccine dn	9.5e-05	0.00649	15	3.56	3

Table 3: The top (lowest p-value) genesets found (uniquely) in each cluster