scAnalyzeR: a comprehensive software package with graphical user interface for single-cell RNA sequencing analysis

User Manual

December 19, 2019

Contents   **Page**

1. Introduction

scAnalyzeR is a comprehensive platform for analysing and visualizing scRNA-seq data with an interactive graphical user interface.

2. How to setup

Firstly, you need to download and install the following softwares(install R then RStudio):

Download and install R and RStudio on your machine,

i. Download and install R (v-3.6.1 or above): https://cran.r-project.org/

ii.Download and install RStudio: RStudio (v-1.1.456 or above): https://rstudio.com/products/rstudio/download/

After installed the R and RStudio on your machine , then you also need to install the shiny package. Please run the following code on RStudio to install the shiny package:

install.packages("shiny")

Now, your machine is ready for running the scAnalyzeR app.

There are many ways to run the scAnalyzeR:

Easiest way is to use runGitHub

Run the following code of line on RStudio.

shiny::runGitHub("scAnalyzeR", "sarwar-chy")

or

Run a tar or zip file directly

shiny::runUrl("https://github.com/sarwar-chy/scAnalyzeR/archive/master.tar.gz")

shiny::runUrl("https://github.com/sarwar-chy/scAnalyzeR/archive/master.zip")

If you want to use it locally, you can simply clone this(https://github.com/sarwar-chy/scAnalyzeR/archive/master.zip) repository as well as unzip it and run the app using RStudio.

To run the app on your local computer, use RStudio to run the following code and replace ~ with the location of your scAnalyzeR-master folder.

shiny::runApp('~/scAnalyzeR-master/')

3. Upload dataset

To upload the data, firstly you choose species then data format. For cellranger output data uploading, select the three files barcodes, genes, and matrix respectively.
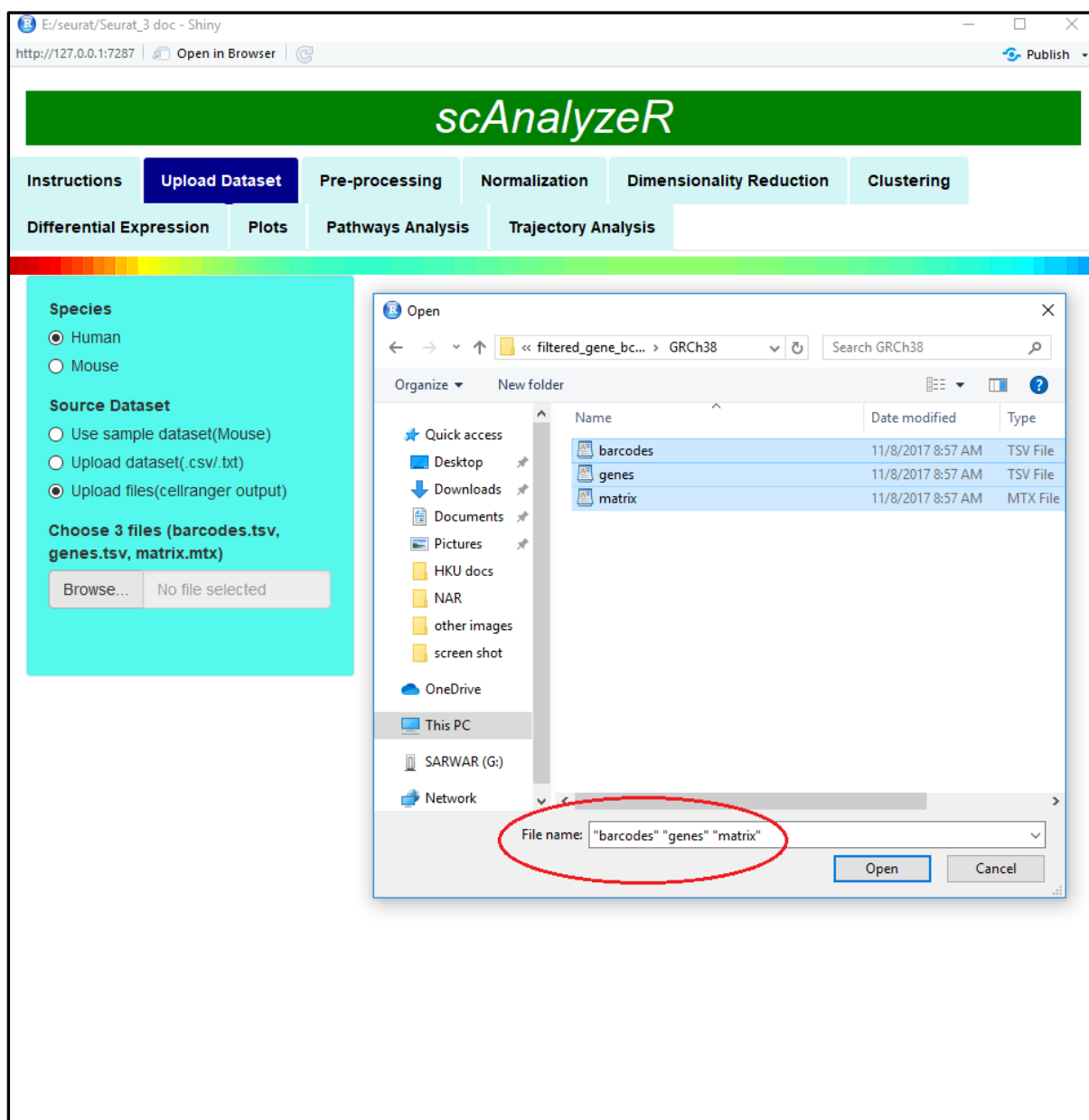
4. **Pre-processing**

scAnalyzeR

Instructions	Upload Dataset	Pre-processing	Normalization	Dimensionality Reduction	Clustering

Differential Expression	Plots	Pathways Analysis	Trajectory Analysis

Create Dataset

Keep all genes expressed in >= MC cells and Keep all cells with at least MFs detected genes

Minimum no. of cells(MC):

> 3

Minimum no. of features(MFs):

> 200

Submit

Display Genes

- ⦿ Mitochondrial
- ◯ None

☑ Plot Meta Data

☐ Download Mitochondrial Genes(as a .csv file)

☑ Plot Gene Summary

☑ *Filtering*

filter cells that have unique feature counts over nFeature_RNA(>) and less than nFeature_RNA(<)

Dataset:
Total number of features(genes): 18340
Total number of samples(cells): 8381

Filtered Dataset:
Total number of features(genes): 18340
Total number of samples(cells): 7802

Show 10 ▾ entries   Search: [   ]

	Total_Count
MT-ND1	59842
MT-ND2	93653
MT-CO1	256609
MT-CO2	157706
MT-ATP8	520
MT-ATP6	78550
MT-CO3	180752
MT-ND3	88809
MT-ND4L	10457
MT-ND4	116821

MT-ND3   88809

MT-ND4L   10457

MT-ND4   116821

Showing 1 to 10 of 13 entries   Previous   **1**   2   Next

Filtering

filter cells that have unique feature counts over nFeature_RNA(>) and less than nFeature_RNA(<)

nFeature_RNA(>):

200

nFeature_RNA(<):

4500

MT gene % :

5

Submit

nFeature_RNA   nCount_RNA   percent.mt
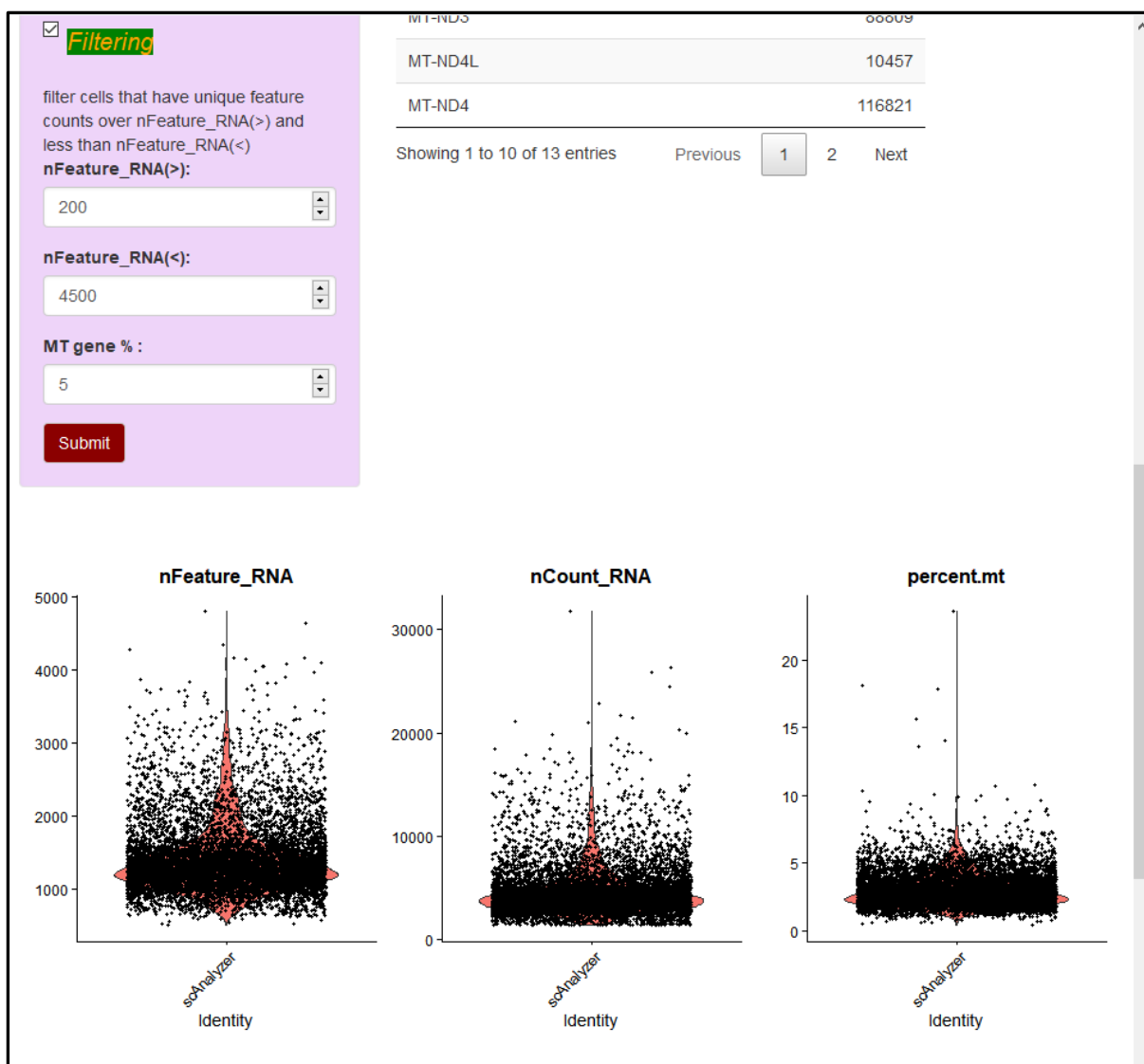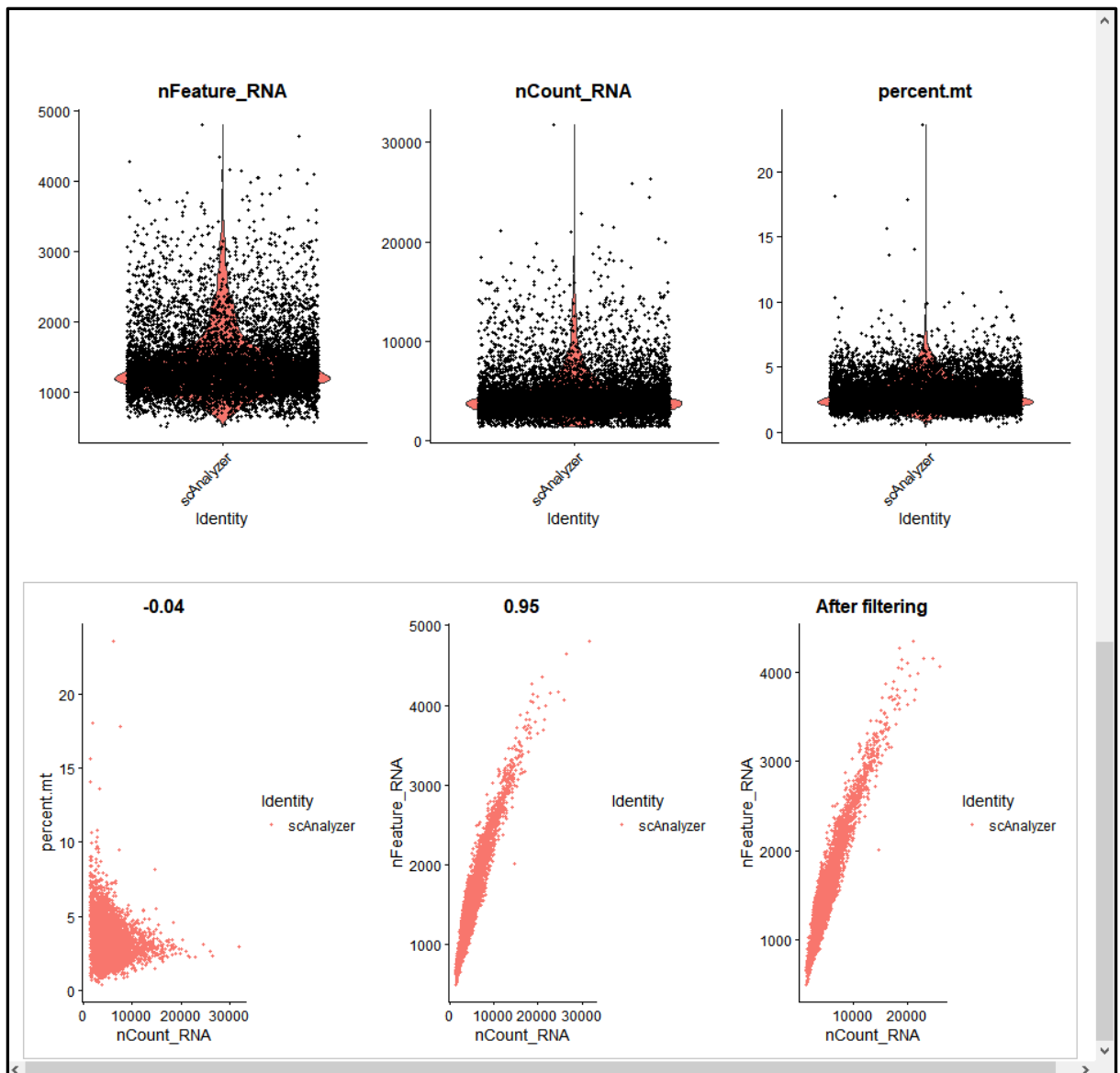
5. **Normalization**

scAnalyzeR

Instructions	Upload Dataset	Pre-processing	Normalization	Dimensionality Reduction	Clustering
Differential Expression	Plots	Pathways Analysis	Trajectory Analysis		

Set Normalization Parameters

Normalization method

Log Normalization ▼

Scaling factor

10000

Submit

Find Highly Variable Features

Features

1000

Submit

☑ Show Highly Variable Genes

☑ Variable Features Plot

Top Features:

10

☐ Download Highly Variable

Normalized Dataset:
Total number of features(genes): 18340
Total number of samples(cells): 7802

1000

Show 10 ▼ entries   Search: [　　　　]

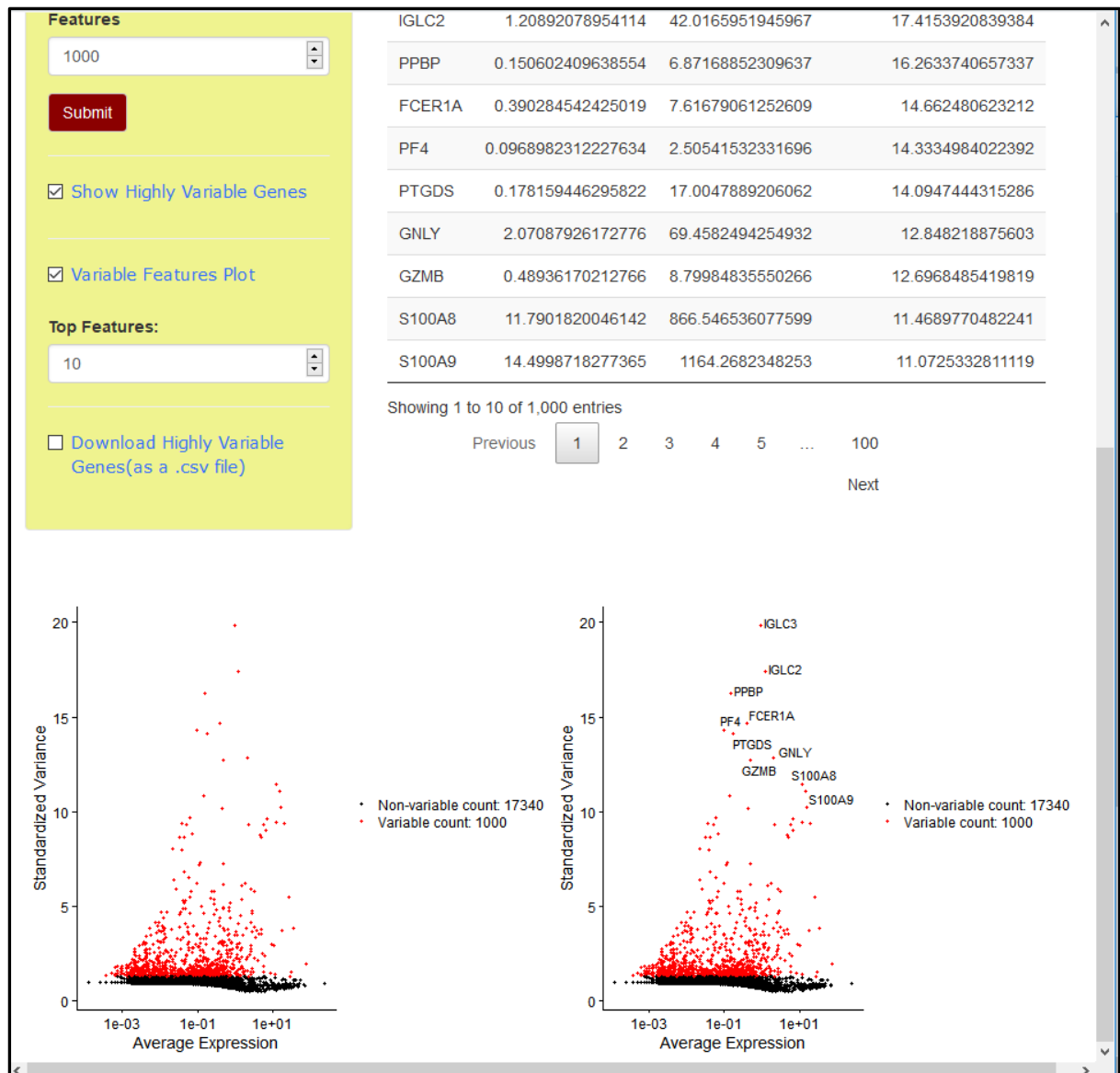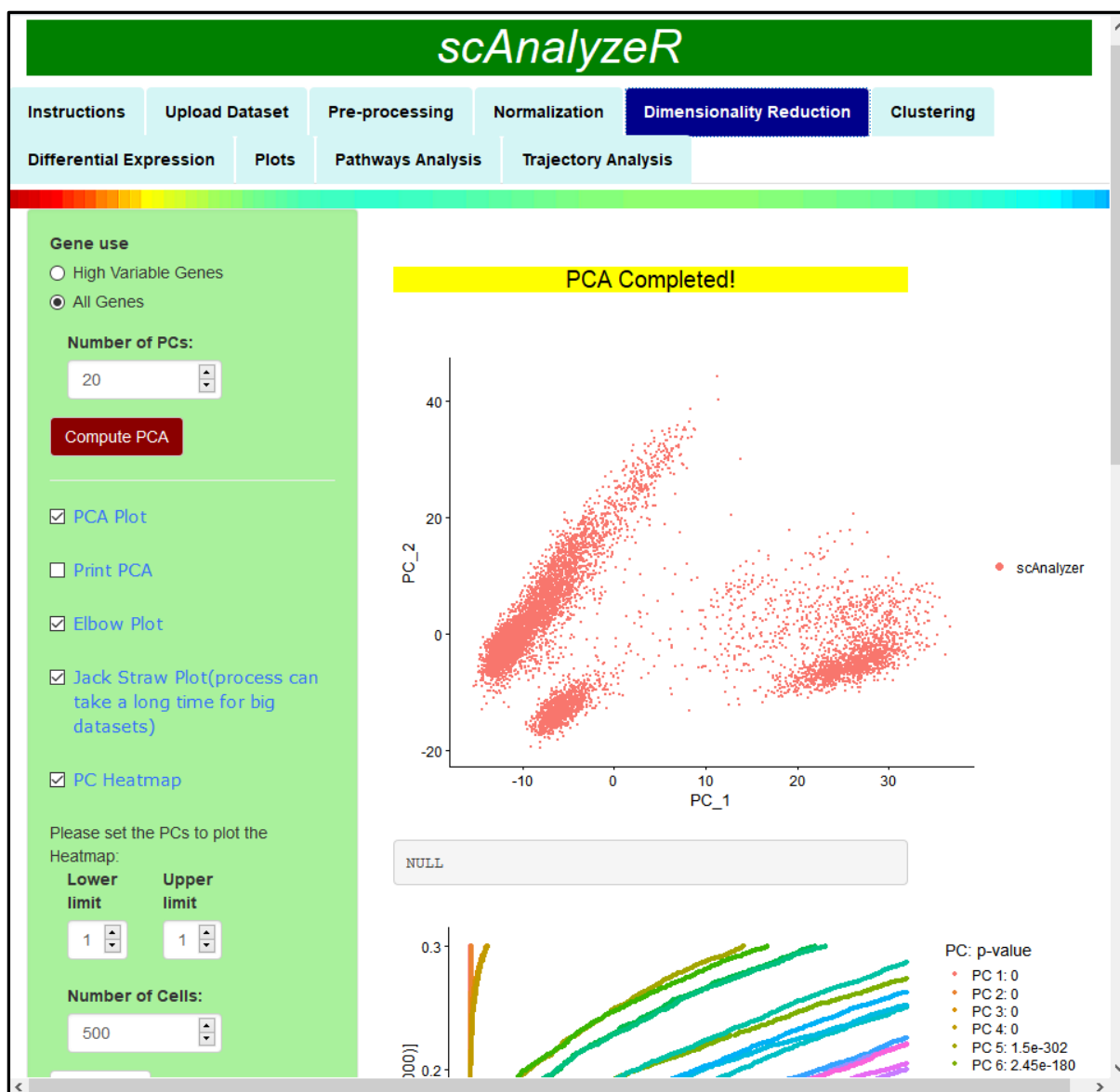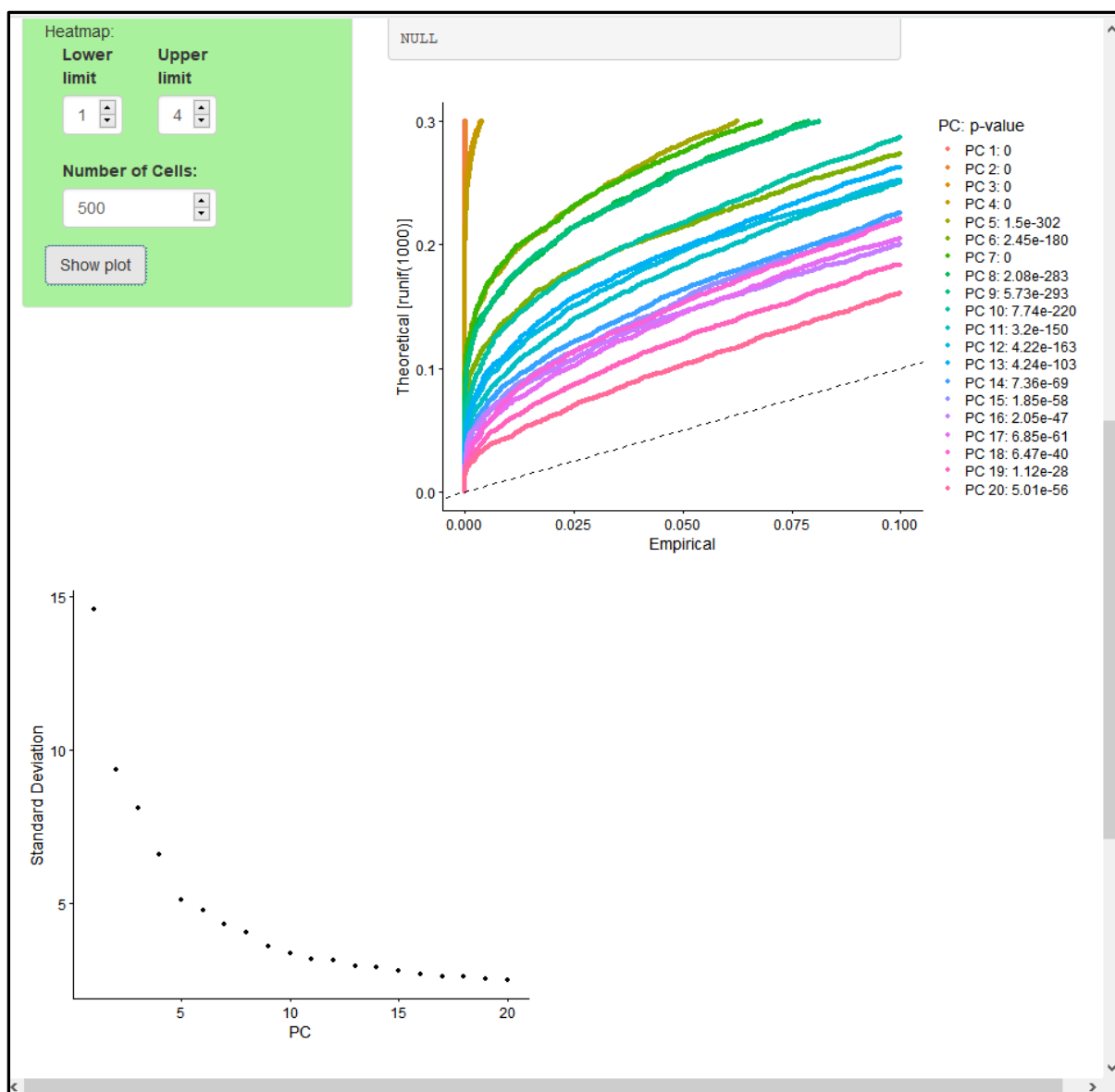|	mean ⇕	variance ⇕	Standardized.variance ⇕
IGLC3	0.941297103306844	44.4757228654423	19.824290188568
IGLC2	1.20892078954114	42.0165951945967	17.4153920839384
PPBP	0.150602409638554	6.87168852309637	16.2633740657337
FCER1A	0.390284542425019	7.61679061252609	14.662480623212
PF4	0.0968982312227634	2.50541532331696	14.3334984022392
PTGDS	0.178159446295822	17.0047889206062	14.0947444315286
GNLY	2.07087926172776	69.4582494254932	12.848218875603
GZMB	0.48936170212766	8.79984835550266	12.6968485419819
S100A8	11.7901820046142	866.546536077599	11.4689770482241
S100A9	14.4998718277365	1164.2682348253	11.0725332811119

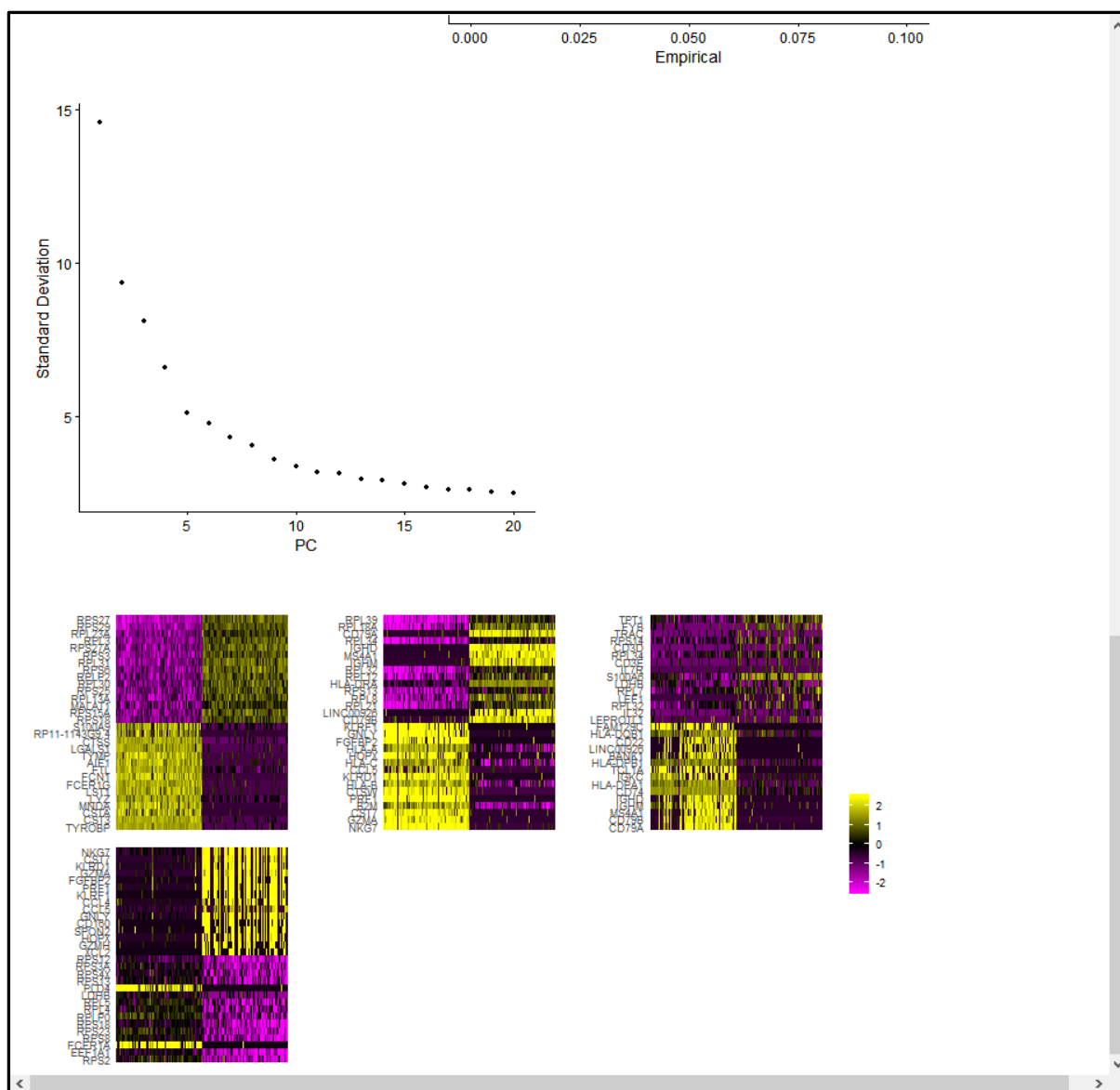Showing 1 to 10 of 1,000 entries

Previous 1 2 3 4 5 … 100

Features

1000

Submit

☑ Show Highly Variable Genes

☑ Variable Features Plot

Top Features:

10

☐ Download Highly Variable Genes(as a .csv file)

IGLC2	1.20892078954114	42.0165951945967	17.4153920839384
PPBP	0.150602409638554	6.87168852309637	16.2633740657337
FCER1A	0.390284542425019	7.61679061252609	14.662480623212
PF4	0.0968982312227634	2.50541532331696	14.3334984022392
PTGDS	0.178159446295822	17.0047889206062	14.0947444315286
GNLY	2.07087926172776	69.4582494254932	12.848218875603
GZMB	0.48936170212766	8.79984835550266	12.6968485419819
S100A8	11.7901820046142	866.546536077599	11.4689770482241
S100A9	14.4998718277365	1164.2682348253	11.0725332811119

Showing 1 to 10 of 1,000 entries

Previous 1 2 3 4 5 ... 100

Next

6. **Dimensionality Reduction**

7. **Clustering**
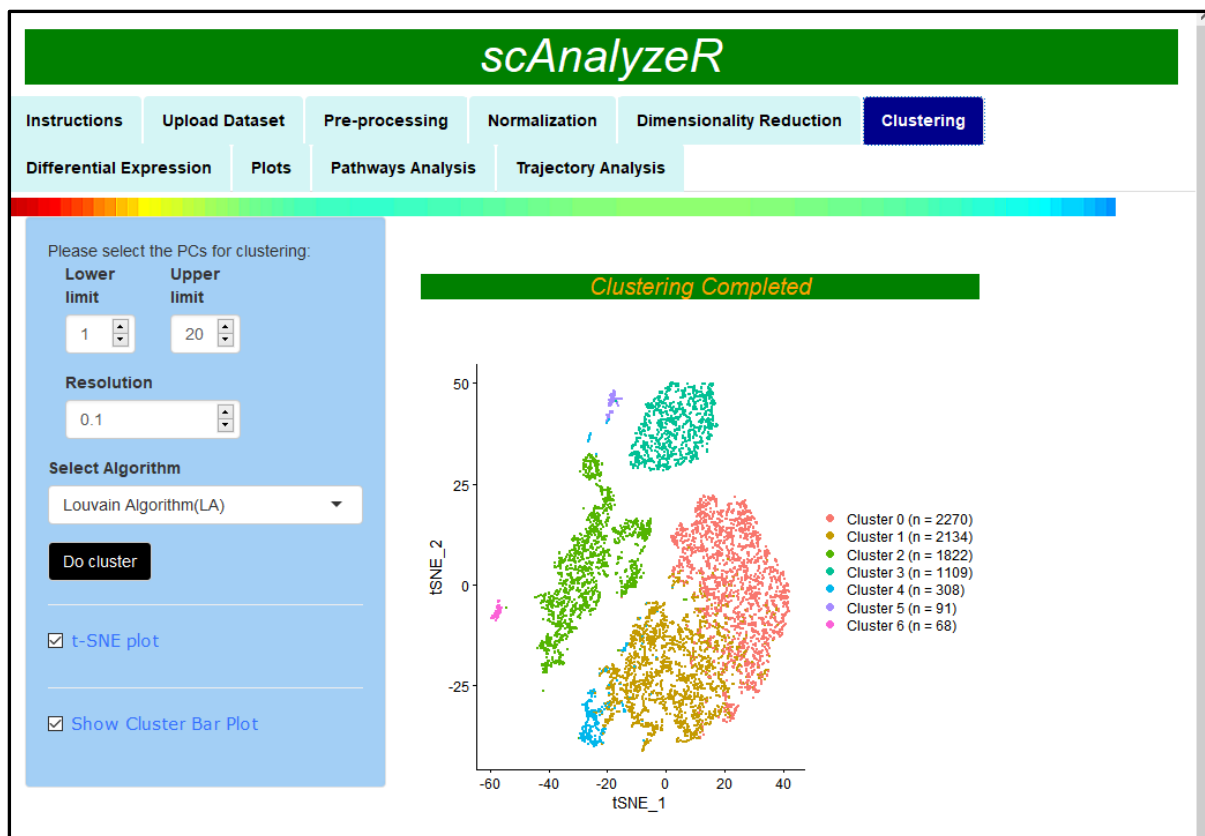
8. Differential Expression Analysis

Markers Selection

◉ Positive only
○ Negative only

☑ Show list

☑ Show Top genes

Top DE genes:

[5 ▲▼]

⬇ Save All Top genes as a text file

⬇ Save +ve Top genes as a text file

⬇ Save -ve Top genes as a text file

☑ Show Heatmap(top genes)

Markers Barplot

◉ All DE genes
○ Filtered DE genes
○ Hide Bar Plot

RPS6	0	0.500593689952276	1	0.999	0	0
RPS25	0	0.453386193921223	1	0.998	0	0
RPL32	0	0.440376612990909	1	0.999	0	0
RPS12	0	0.43606273251681	1	0.999	0	0

Showing 1 to 10 of 40,042 entries

Previous   1   2   3   4   5   …   4005   Next

Show 10 entries

Search: [   ]

	p_val	avg_logFC	pct.1	pct.2	p_val_adj	cluster	gene
1	0	0.989961705166956	0.668	0.165	0	0	CCR7
2	0	0.98267857469619	0.676	0.172	0	0	LEF1
3	0	0.748540684134712	0.964	0.671	0	0	LDHB
4	0	0.676551435415744	0.959	0.535	0	0	TRAC
5	0	0.545532929247101	0.988	0.909	0	0	NPM1
6	0	0.534109740650198	1	0.998	0	0	RPL31
7	0	0.500593689952276	1	0.999	0	0	RPS6
8	0	0.453386193921223	1	0.998	0	0	RPS25
9	0	0.440376612990909	1	0.999	0	0	RPL32
10	0	0.43606273251681	1	0.999	0	0	RPS12

Showing 1 to 10 of 3,104 entries

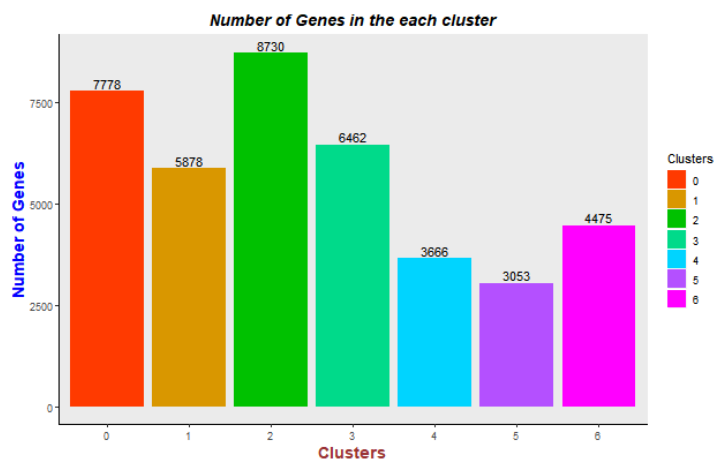Previous   1   2   3   4   5   …   311   Next

☑ Show Heatmap(top genes)

Markers Barplot
◉ All DE genes
○ Filtered DE genes
○ Hide Bar Plot

1	0	0.989961705166956	0.668	0.165	0	0	CCR7
2	0	0.98267857469619	0.676	0.172	0	0	LEF1
3	0	0.748540684134712	0.964	0.671	0	0	LDHB
4	0	0.676551435415744	0.959	0.535	0	0	TRAC
5	0	0.545532929247101	0.988	0.909	0	0	NPM1
6	0	0.534109740650198	1	0.998	0	0	RPL31
7	0	0.500593689952276	1	0.999	0	0	RPS6
8	0	0.453386193921223	1	0.998	0	0	RPS25
9	0	0.440376612990909	1	0.999	0	0	RPL32
10	0	0.43606273251681	1	0.999	0	0	RPS12

Showing 1 to 10 of 3,104 entries

Previous　1　2　3　4　5

… 311 Next

Number of Genes in the each cluster

Show 10 ▾ entries

Search:

	p_val	avg_logFC	pct.1	pct.2	p_val_adj	cluster	gene
1	0	0.989961705166956	0.668	0.165	0	0	CCR7
2	0	0.98267857469619	0.676	0.172	0	0	LEF1
3	2.7448333093562e-160	0.828978424285934	0.426	0.164	5.03402428935927e-156	0	CD8B
4	0	0.748540684134712	0.964	0.671	0	0	LDHB
5	1.12398544628874e-261	0.73633778246992	0.631	0.262	2.06138930849355e-257	0	PIK3IP1
6	0	2.00730129700082	0.679	0.225	0	1	CCL5
7	0	1.59522703490683	0.445	0.066	0	1	GZMK
8	6.57650779039408e-266	1.46084617522336	0.5	0.146	1.20613152875827e-261	1	KLRB1
9	0	1.45048095012378	0.985	0.552	0	1	IL32
10	0	1.22811538548724	0.403	0.05	0	1	TRGC2

Showing 1 to 10 of 35 entries

Previous 1 2 3 4

Next

2	0	0.98267857469619	0.676	0.172	0	0	LEF1
3	2.7448333093562e-160	0.828978424285934	0.426	0.164	5.03402428935927e-156	0	CD8B
4	0	0.748540684134712	0.964	0.671	0	0	LDHB
5	1.12398544628874e-261	0.73633778246992	0.631	0.262	2.06138930849355e-257	0	PIK3IP1
6	0	2.00730129700082	0.679	0.225	0	1	CCL5
7	0	1.59522703490683	0.445	0.066	0	1	GZMK
8	6.57650779039408e-266	1.46084617522336	0.5	0.146	1.20613152875827e-261	1	KLRB1
9	0	1.45048095012378	0.985	0.552	0	1	IL32
10	0	1.22811538548724	0.403	0.05	0	1	TRGC2

Showing 1 to 10 of 35 entries

Previous 1 2 3 4

Next

Differential Expression | Plots | Pathways Analysis | Trajectory Analysis

Find all markers | **Find markers by cluster**

Find markers by clusters vs other clusters

Please select parameters:

Select cluster   **logfc threshold**

2 ▼   0.25

Min % (min.pct)

0.1

Select Test Method

Wilcox ▼

Find Markers

⬇ Save All markers as a csv

☑ Show All Markers(both positive and negative)

☑ Filtering Markes

Adjust p-value

0.05

Markers Selection
◉ Positive only
○ Negative only

☐ Show Heatmap

☐ Show Filtered Markers

Show 10 entries

Search: [   ]

	p_val	avg_logFC	pct.1	pct.2	p_val_adj	Gene
S100A8	0	3.63373580376454	0.984	0.522	0	S100A8
S100A9	0	3.58020227547237	0.992	0.628	0	S100A9
LYZ	0	3.5748764937931	0.997	0.579	0	LYZ
CST3	0	2.79494532990639	0.996	0.368	0	CST3
RP11-1143G9.4	0	2.56344980712131	0.909	0.077	0	RP11-1143G9.4
S100A12	0	2.51260868550754	0.796	0.076	0	S100A12
TYROBP	0	2.50221969256362	0.992	0.323	0	TYROBP
FCN1	0	2.45458451160608	0.944	0.138	0	FCN1
LST1	0	2.43445781564075	0.987	0.243	0	LST1
AIF1	0	2.41099354528212	0.985	0.234	0	AIF1

Showing 1 to 10 of 1,259 entries

Previous 1 2 3 4 5 … 126 Next

S100A12	0	2.51260868550754	0.796	0.076	0	S100A12
TYROBP	0	2.50221969256362	0.992	0.323	0	TYROBP
FCN1	0	2.45458451160608	0.944	0.138	0	FCN1
LST1	0	2.43445781564075	0.987	0.243	0	LST1
AIF1	0	2.41099354528212	0.985	0.234	0	AIF1

Showing 1 to 10 of 1,259 entries

Previous　1　2　3　4　5

...　126　Next

☑ Save All markers as a csv

☑ Show All Markers(both positive and negative)

☑ Filtering Markes

Adjust p-value

0.05

Markers Selection
◉ Positive only
○ Negative only

☑ Show Heatmap

Top Genes:

10

☑ Zoom the cluster

☑ Show Filtered Markers

Show 10 ▾ entries

Search:

	Gene	p_val	avg_logFC	pct.1	pct.2	p_val_adj
1	S100A8	0	3.63373580376454	0.984	0.522	0
2	S100A9	0	3.58020227547237	0.992	0.628	0
3	LYZ	0	3.5748764937931	0.997	0.579	0
4	CST3	0	2.79494532990639	0.996	0.368	0
5	RP11-1143G9.4	0	2.56344980712131	0.909	0.077	0
6	S100A12	0	2.51260868550754	0.796	0.076	0
7	TYROBP	0	2.50221969256362	0.992	0.323	0
8	FCN1	0	2.45458451160608	0.944	0.138	0
9	LST1	0	2.43445781564075	0.987	0.243	0
10	AIF1	0	2.41099354528212	0.985	0.234	0

Showing 1 to 10 of 791 entries

Previous　1　2　3　4　5

...　80　Next
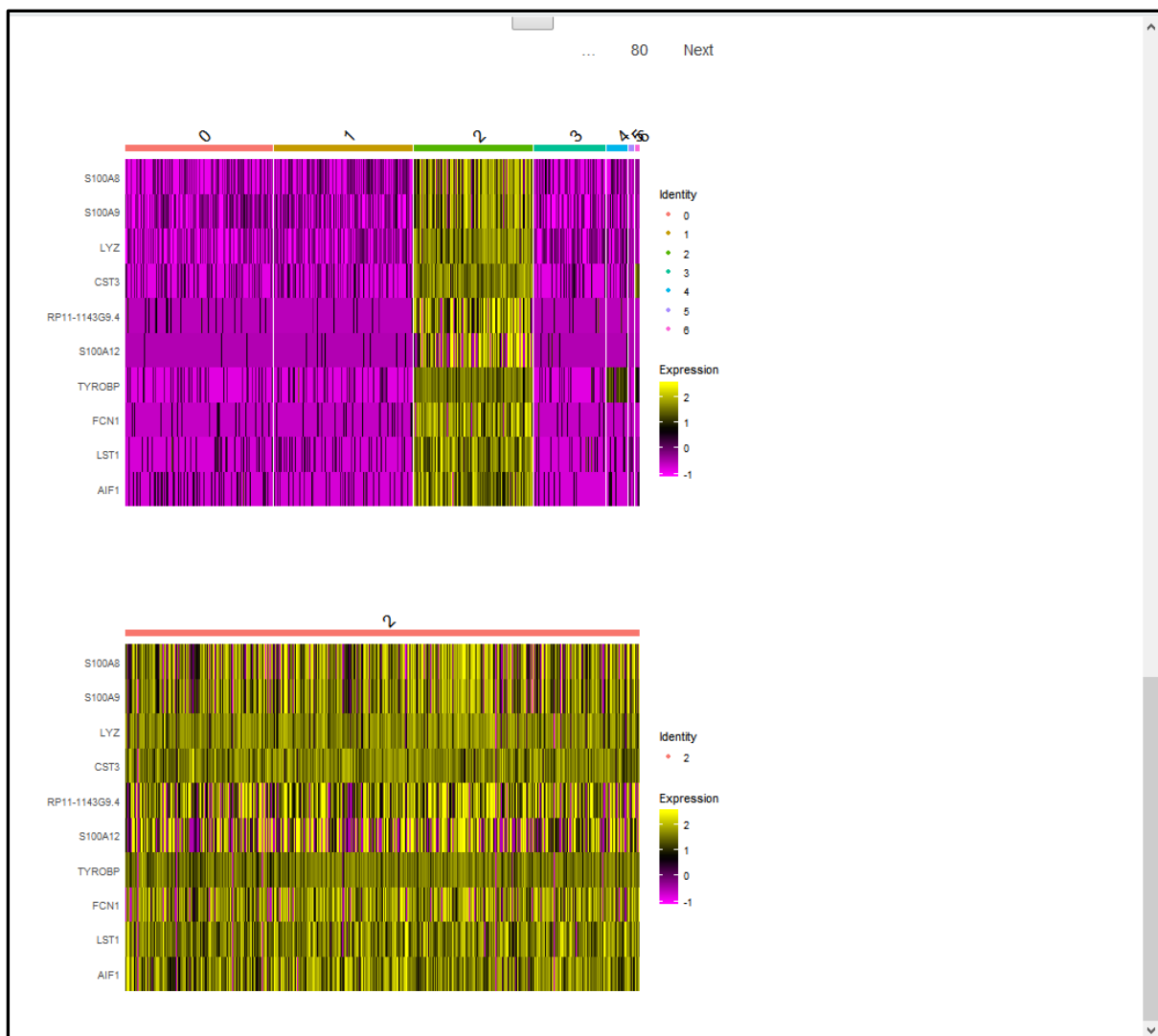
| Instructions | Upload Dataset | Pre-processing | Normalization | Dimensionality Reduction | Clustering |

Differential Expression | Plots | Pathways Analysis | Trajectory Analysis

Find all markers   Find markers by cluster

Find markers by clusters vs other clusters

Please select parameters:

Select Cluster/s

The cluster number must be separated by a commma

Select complementary cluster/s

The cluster number must be separated by a commma

logfc threshold   **Min % (min.pct)**

0.25   0.1

Select Test Method

Wilcox

Find markers

☐ Show All Markers(Both Positive and Negative)

☐ Filtering Markers

Find markers by clusters vs other clusters

Please select parameters:

Select Cluster/s

0,3

Select complementary cluster/s

1,5

logfc threshold	Min % (min.pct)
0.25	0.1

Select Test Method

Wilcox

Find markers

☑ Show All Markers(Both Positive and Negative)

☐ Filtering Markers

Show 10 entries

Search:

	p_val	avg_logFC	pct.1	pct.2	p_val_adj	Gene
RPS23	0	0.341986516375793	1	1	0	RPS23
HLA-B	0	-0.477157353001287	1	1	0	HLA-B
B2M	0	-0.487664881545657	1	1	0	B2M
HLA-C	0	-0.492715744713098	0.997	1	0	HLA-C
HLA-A	0	-0.556744523186942	0.994	1	0	HLA-A
CD99	0	-0.946486407637166	0.323	0.799	0	CD99
IL32	0	-1.16441083863475	0.671	0.983	0	IL32
LYAR	0	-1.19140458046004	0.097	0.577	0	LYAR
ANXA1	0	-1.37177218345366	0.193	0.76	0	ANXA1
CST7	0	-1.44789736555727	0.035	0.504	0	CST7

Showing 1 to 10 of 371 entries

Previous 1 2 3 4

5 ... 38 Next

☑ Filtering Markers

Adjust p-value

0.05

Markers Selection
◉ Positive only
◯ Negative only

⬇ Save Filtered Markers as csv

☑ Show Heatmap

Top Genes:

10

☑ Show Filtered Markers

Previous **1** 2 3 4 5
... 38 Next

Show 10 ▾ entries

Search: [   ]

|	Gene ⇕	p_val ⇕	avg_logFC ⇕	pct.1 ⇕	pct.2 ⇕	p_val_adj ⇕
53	IGKC	1.78433677752888e-48	2.16208138492901	0.43	0.304	3.27247364998797e-44
28	HLA-DRA	8.06197781863657e-65	1.69019696990863	0.677	0.592	1.47856673193795e-60
67	IGLC2	4.98271223459222e-40	1.61111724873305	0.179	0.063	9.13829423824214e-36
74	IGLC3	2.78296462573471e-38	1.51707196477959	0.159	0.05	5.10395712359745e-34
16	IGHM	4.75203987257259e-95	1.50525954372878	0.369	0.148	8.71524112629812e-91
11	CD79A	1.85927928246677e-129	1.478624341117	0.392	0.121	3.40991820404405e-125
23	TCL1A	5.22737069896925e-70	1.32109676796572	0.248	0.075	9.5869978619096e-66
12	IGHD	1.07833686044379e-115	1.2850612911637	0.313	0.067	1.97766980205391e-111
21	CD79B	2.71530718596355e-71	1.26566533243811	0.405	0.237	4.97987337905715e-67
14	MS4A1	1.30734235763571e-99	1.21848621636033	0.343	0.119	2.39766588390389e-95

Showing 1 to 10 of 113 entries

Previous **1** 2 3 4 5
... 12 Next

74	IGLC3	2.78296462573471e-38	1.51707196477959	0.159	0.05	5.10395712359743e-34
16	IGHM	4.75203987257259e-95	1.50525954372878	0.369	0.148	8.71524112629812e-91
11	CD79A	1.85927928246677e-129	1.478624341117	0.392	0.121	3.40991820404405e-125
23	TCL1A	5.22737069896925e-70	1.32109676796572	0.248	0.075	9.5869978619096e-66
12	IGHD	1.07833686044379e-115	1.2850612911637	0.313	0.067	1.97766980205391e-111
21	CD79B	2.71530718596355e-71	1.26566533243811	0.405	0.237	4.97987337905715e-67
14	MS4A1	1.30734235763571e-99	1.21848621636033	0.343	0.119	2.39766588390389e-95

Showing 1 to 10 of 113 entries

Previous | 1 | 2 | 3 | 4 | 5

... | 12 | Next

9. **Plots**
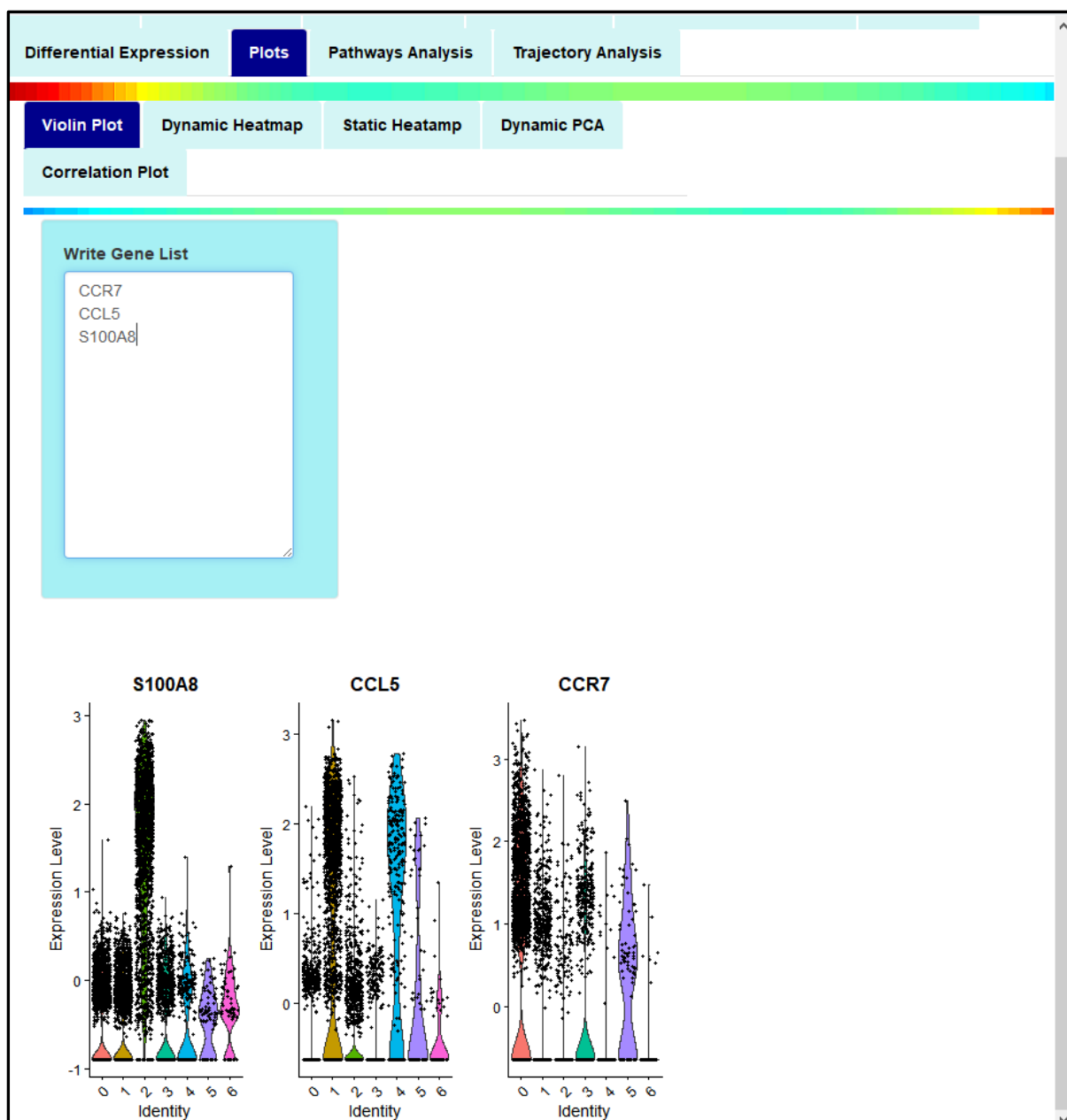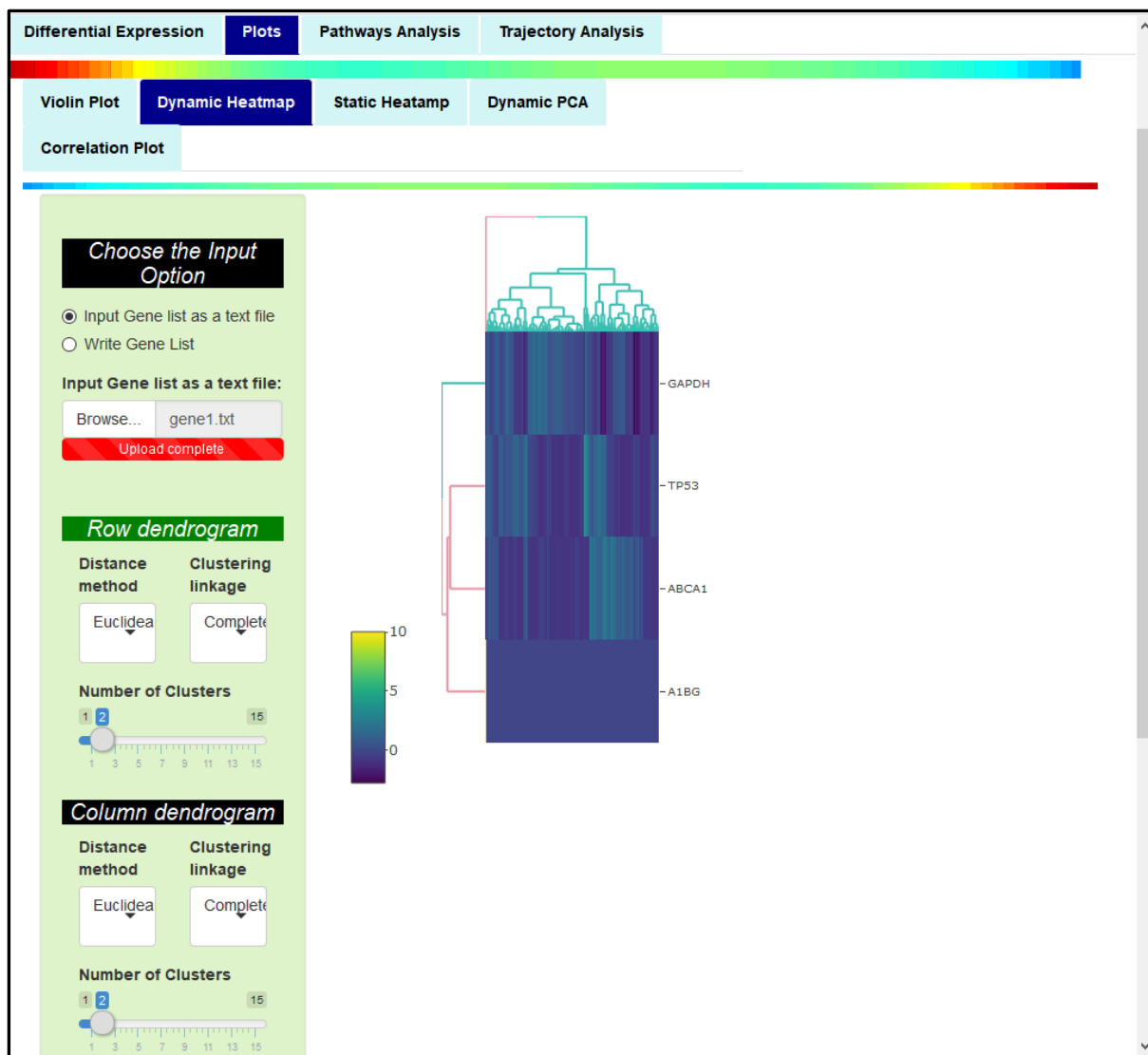
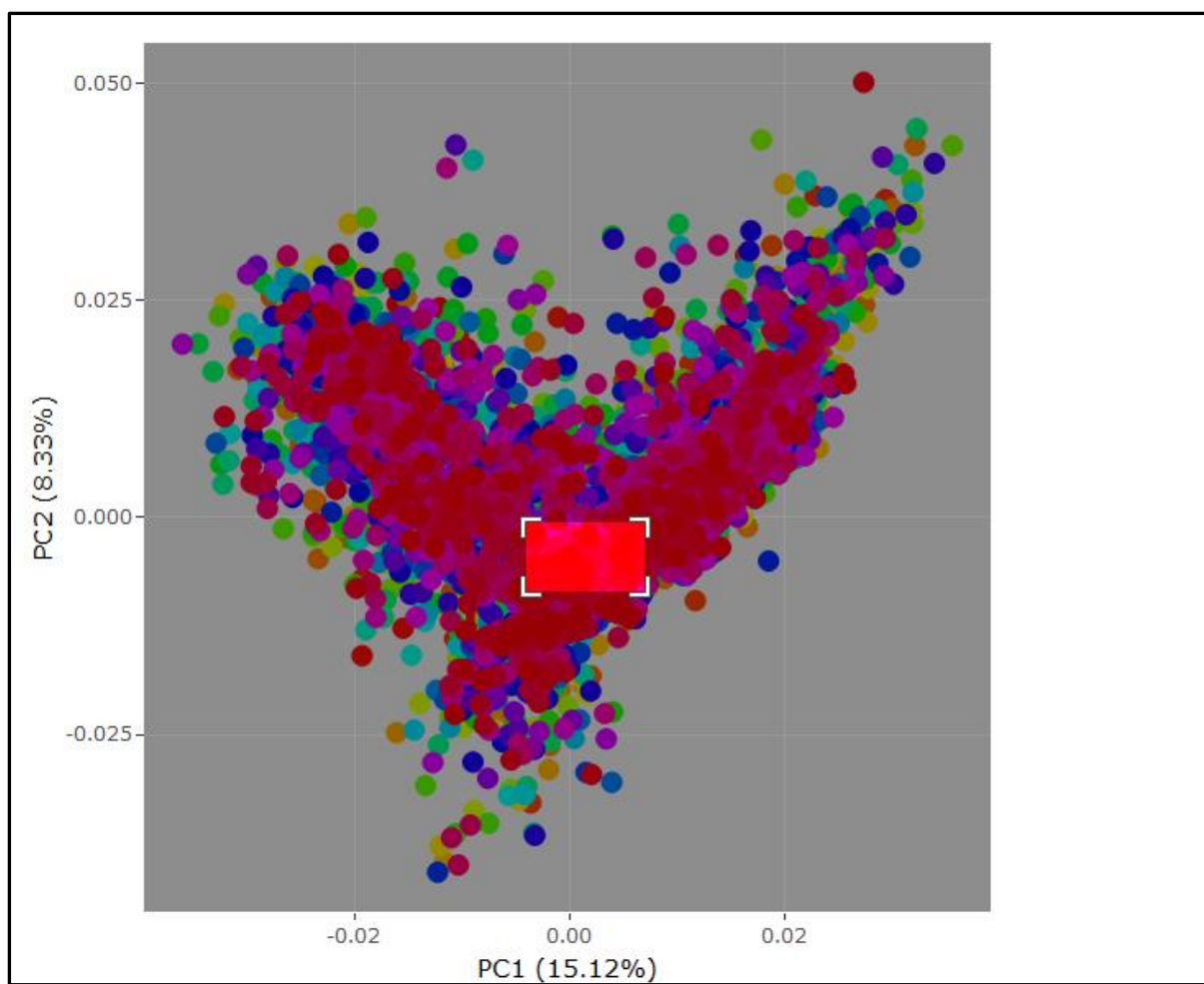**Select Plot**
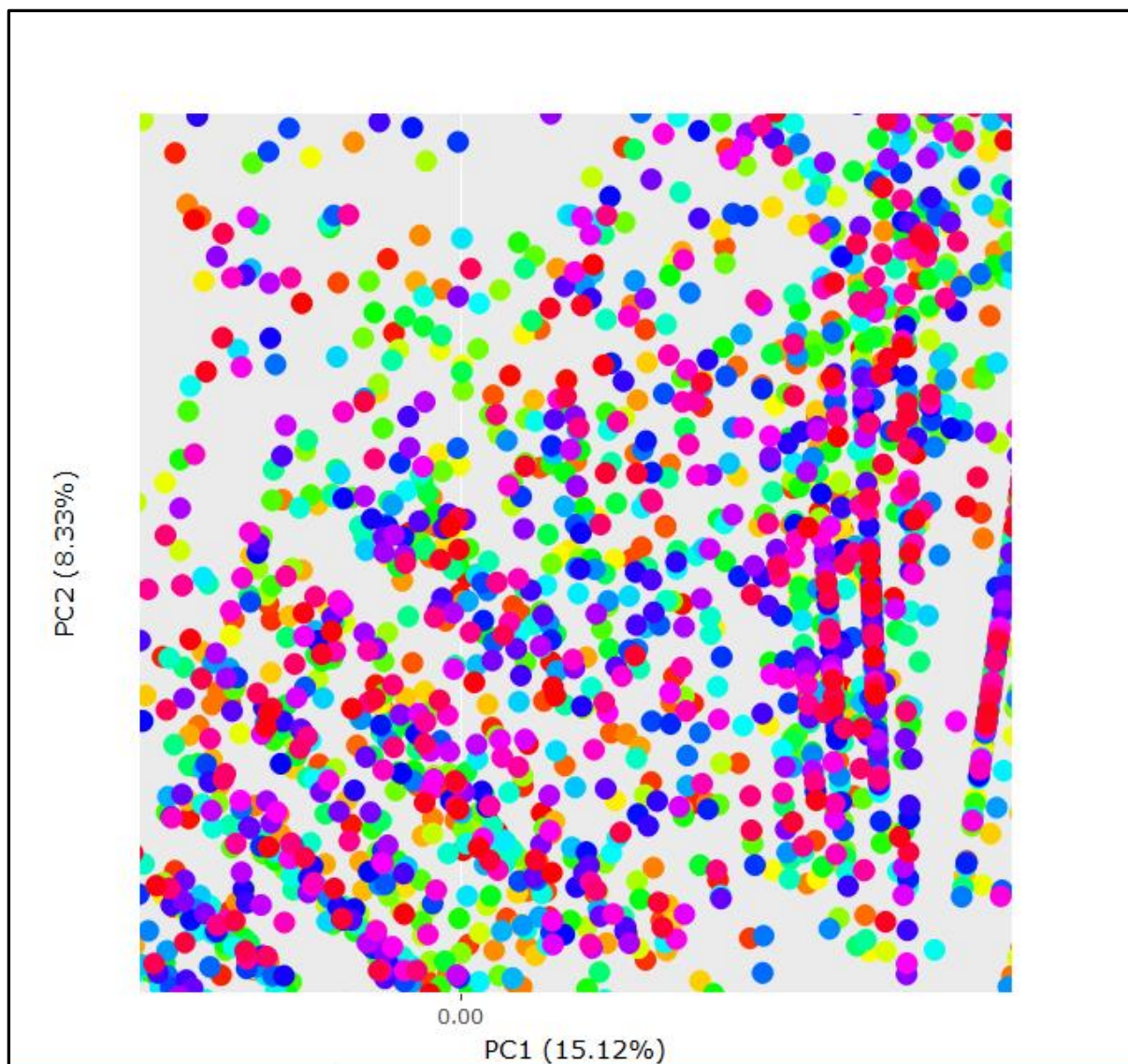
○ Heatmap with clusters

◉ Heatmap without Clusters

PC2 (8.33%)

0.00

PC1 (15.12%)
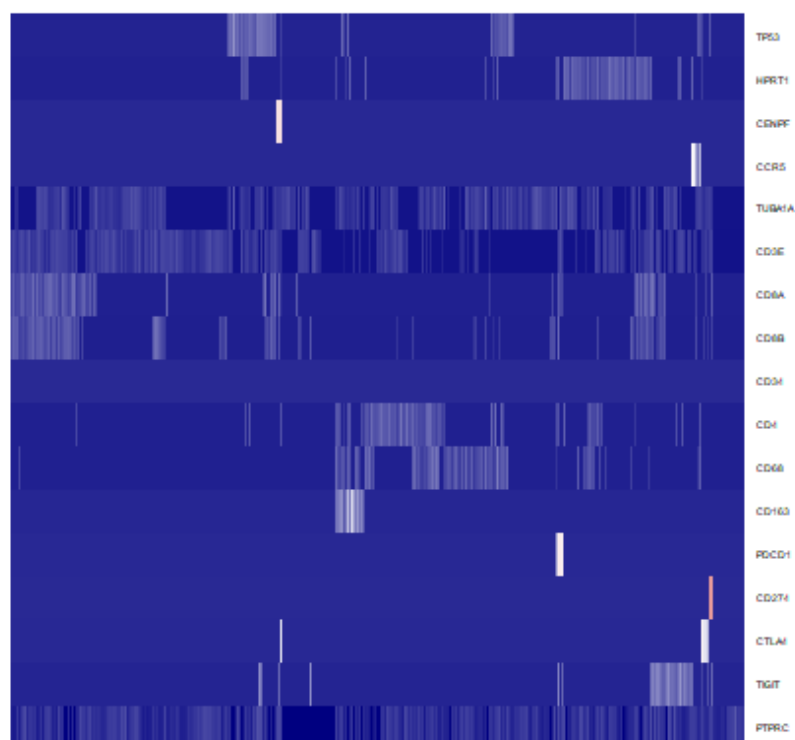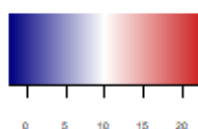
10.Pathway Analysis

scAnalyzeR

Instructions | Upload Dataset | Pre-processing | Normalization | Dimensionality Reduction | Clustering | Differential Expression | Plots | Pathways Analysis | Trajectory Analysis

Select DE Genes

- Cluster Specific DE Genes
- Cluster(s) vs Cluster(s) DE Genes

Select Pathway Source

KEGG Pathways

Compute Pathways

⬇ Download All Pathways as a csv

Select Pathways

- Positive Pathways
- Negative pathways

Cutoff q-value(<= cutoff):

0.5

⬇ Download Filtered Pathways as a csv

☑ Show Pathways(plot)

Top Pathways:

10

☑ Show Pathways list

☐ Show Gene sets:

Show 10 entries   Search:

	pathway	pval	padj	ES	NES	nMoreExtreme	size	Sig_genesets
1	Phagosome	0.00073385518590998	0.0235246808082839	0.475248818678833	1.95081430736081	5	48	20
2	Staphylococcus aureus infection	0.000844475721323012	0.0235246808082839	0.630347646706248	2.05290490688612	5	19	13
3	Asthma	0.00014801657785672	0.0235246808082839	0.714936798069136	2.09784533507546	0	13	9
4	Lysosome	0.000527078666490974	0.0235246808082839	0.5724424449476	2.083779924655	3	26	16
5	Leishmaniasis	0.00181300181300181	0.0441919191919192	0.517509206603927	1.94752070646304	13	32	22
6	Metabolic pathways	0.0118805159538357	0.151515151515152	0.350722417004452	1.60404210797303	104	87	60
7	Osteoclast differentiation	0.0117470632341915	0.151515151515152	0.431723555507705	1.71500143238771	93	41	19
8	Toll-like receptor signaling pathway	0.00834652467981004	0.151515151515152	0.576854825850755	1.79218984220634	57	16	8
9	Rheumatoid arthritis	0.0124320124320124	0.151515151515152	0.463816734146028	1.74546207531489	95	32	19
10	Bladder cancer	0.00939620671654776	0.151515151515152	0.825609976717102	1.66403548235537	53	4	2

Showing 1 to 10 of 29 entries   Previous 1 2 3 Next

NULL

Top Pathways:

10

☑ **Show Pathways list**

☑ Show Gene sets:

Pathway No.:

1

Select Gene Set

- Significant Genes Only
- All Genes

☐ Show Gene set Heatmap:

Showing 1 to 10 of 29 entries   Previous 1 2 3 Next
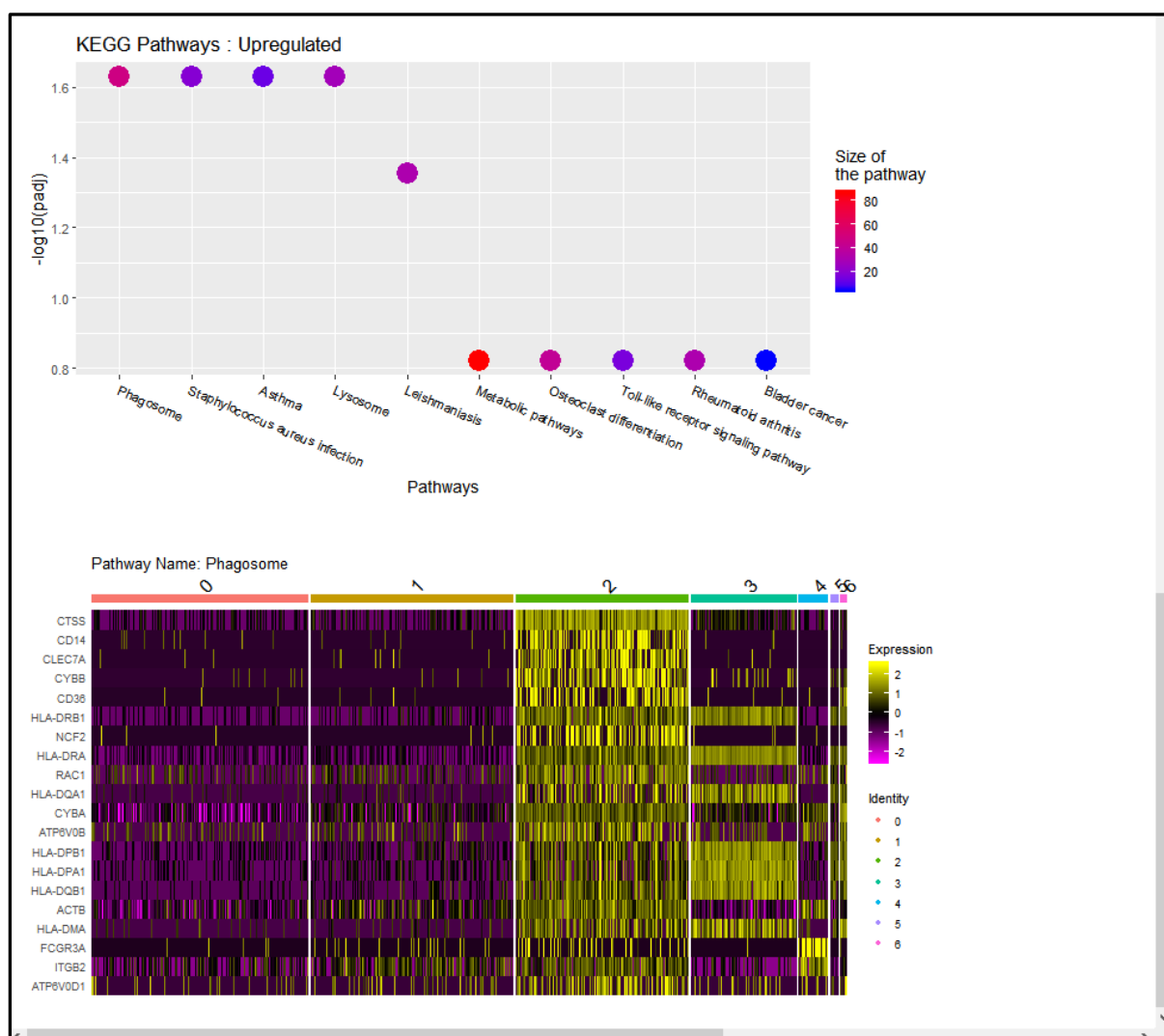
```
 [1] "CTSS"      "CD14"      "CLEC7A"    "CYBB"      "CD36"      "HLA-DRB1"  "NCF
[13] "HLA-DPB1"  "HLA-DPA1"  "HLA-DQB1"  "ACTB"      "HLA-DMA"   "FCGR3A"    "ITG
```
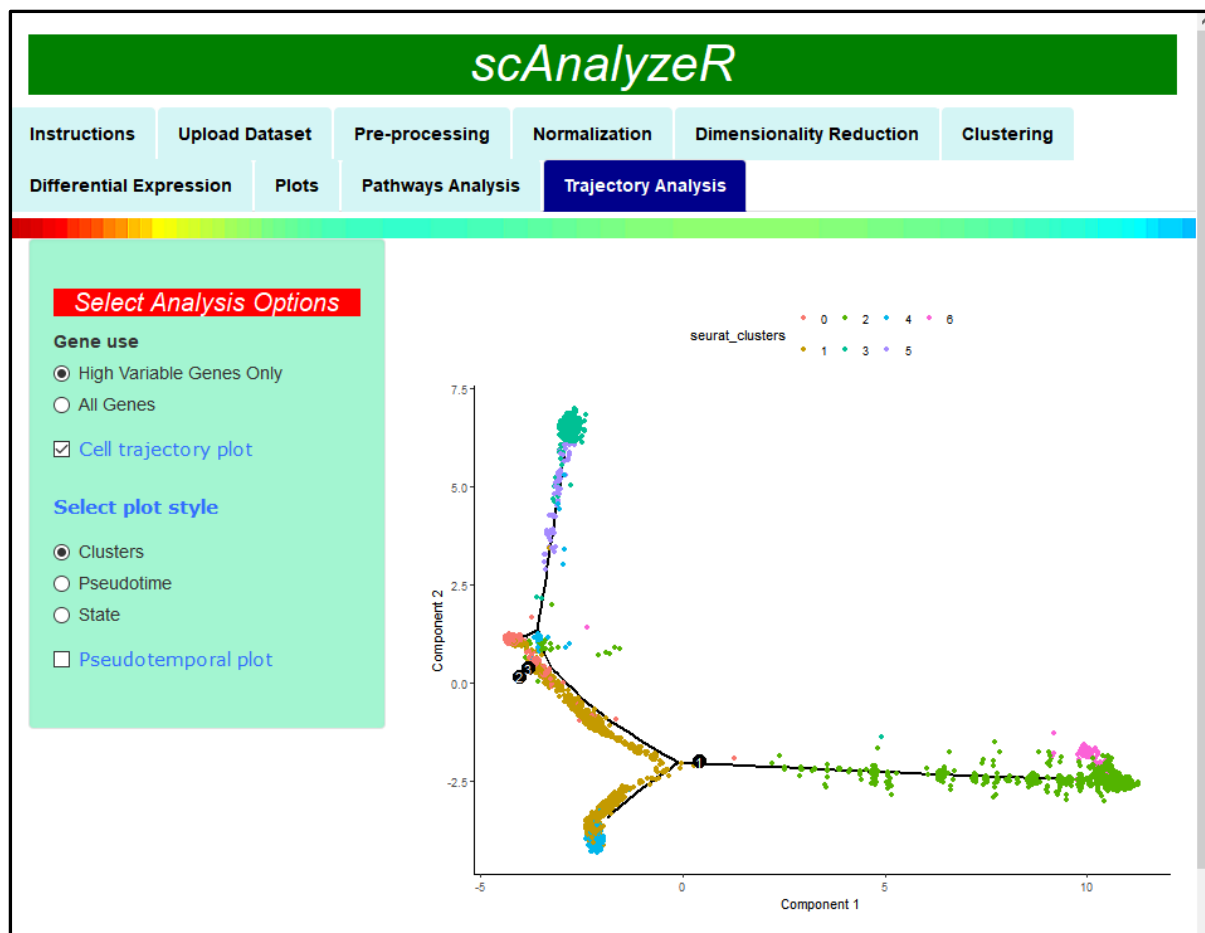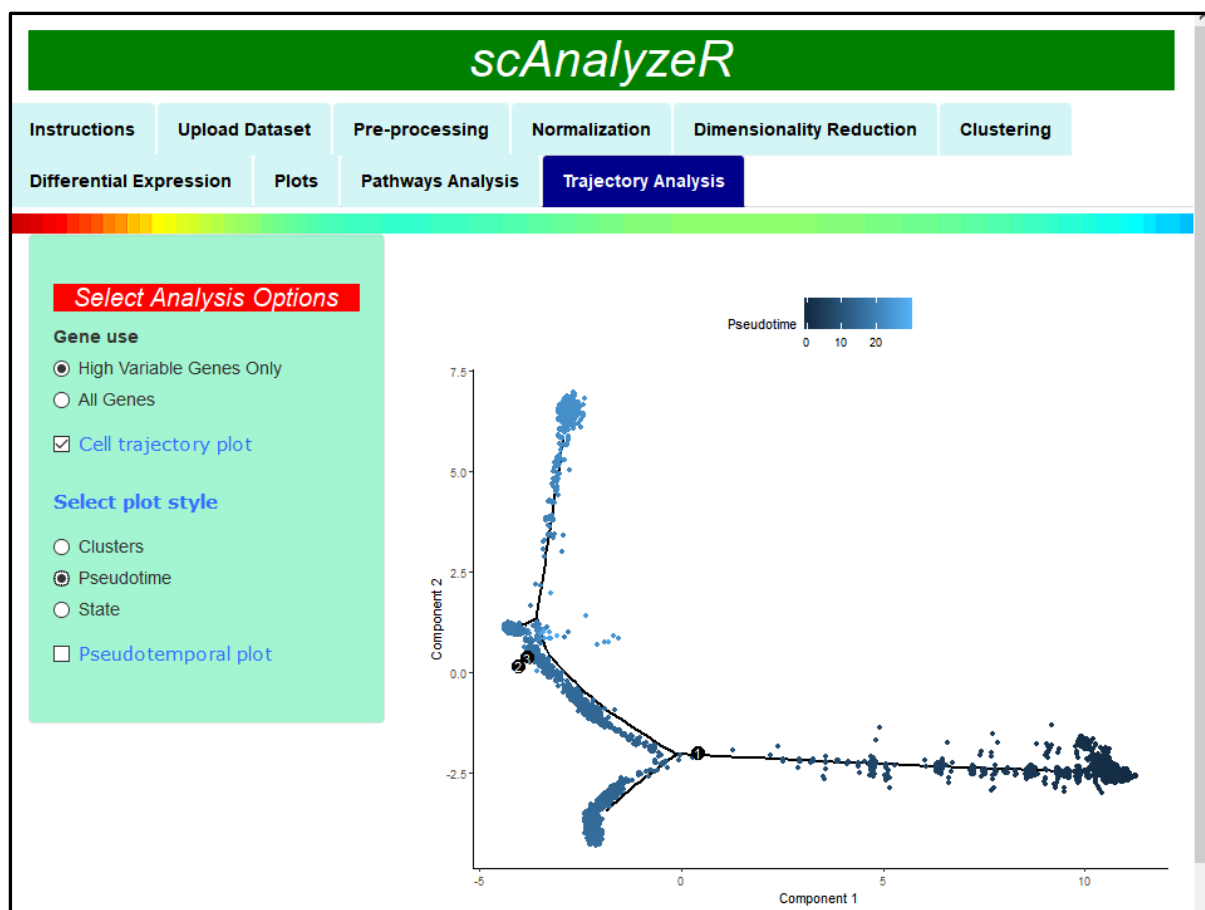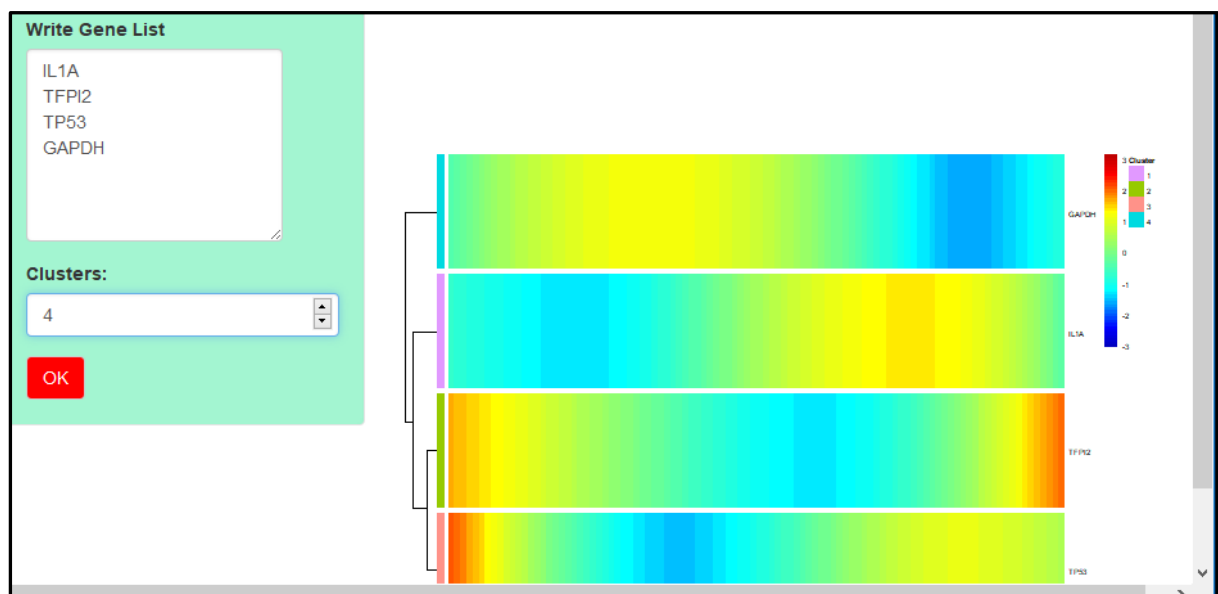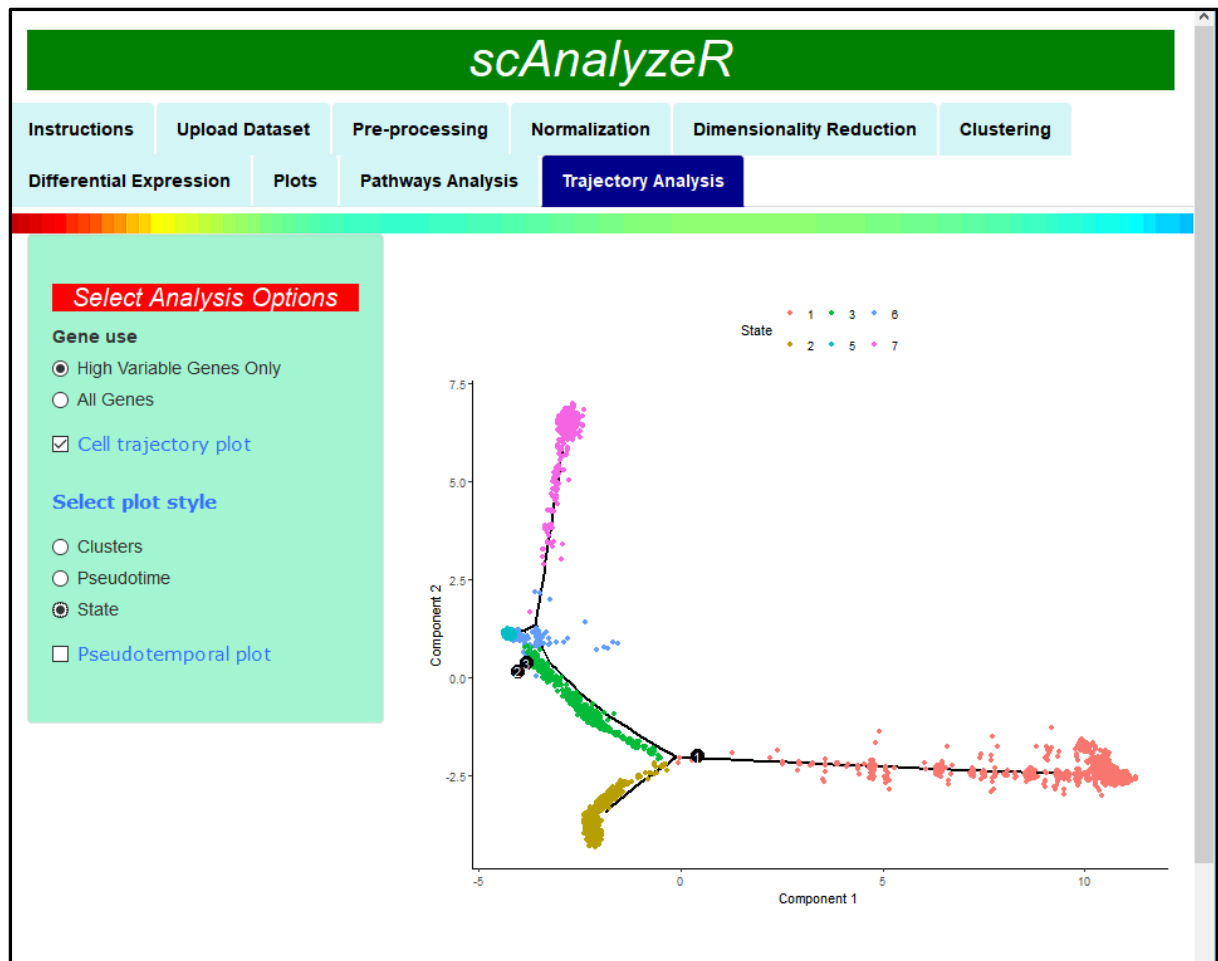

KEGG Pathways : Upregulated

11. <u>Trajectory Analysis</u>

12. References