Analysis on WT and RT by Seurat

Li Hongmin

Outline

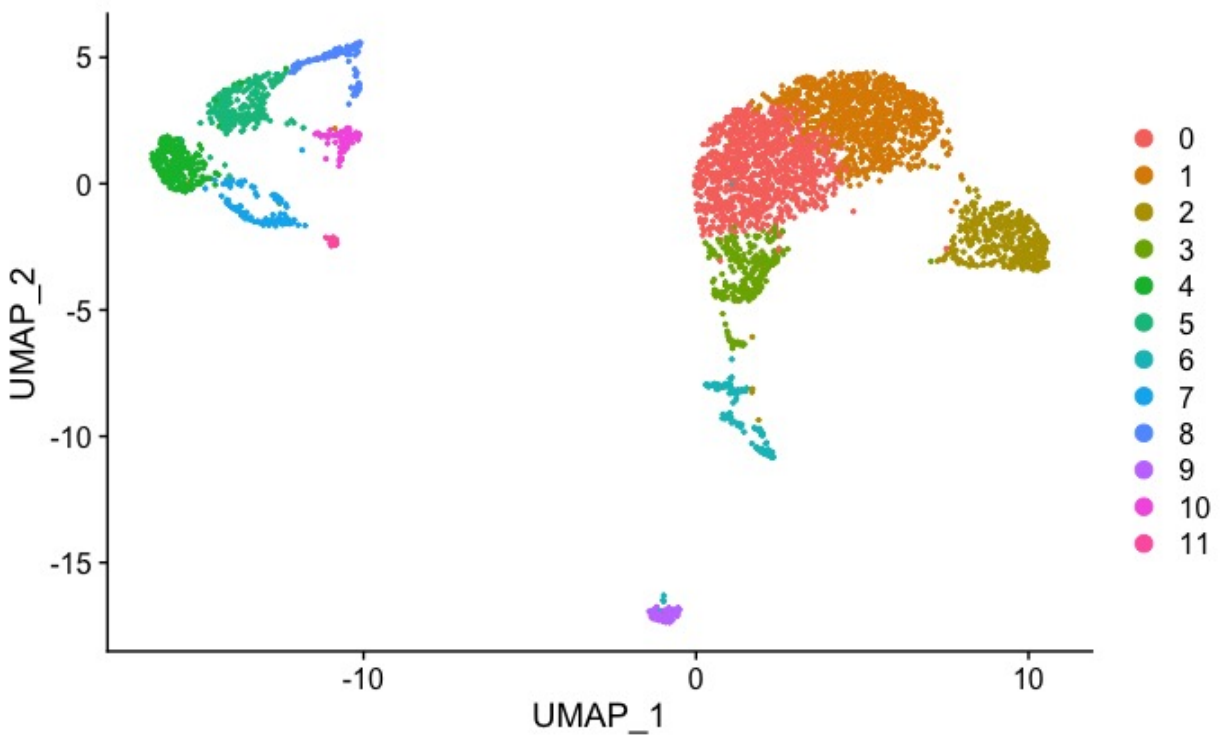
- Analysis on WT by Seurat
- Analysis on RT by Seurat
- Compare the results on WT and RT
- Summary

Analysis on WT by Seurat

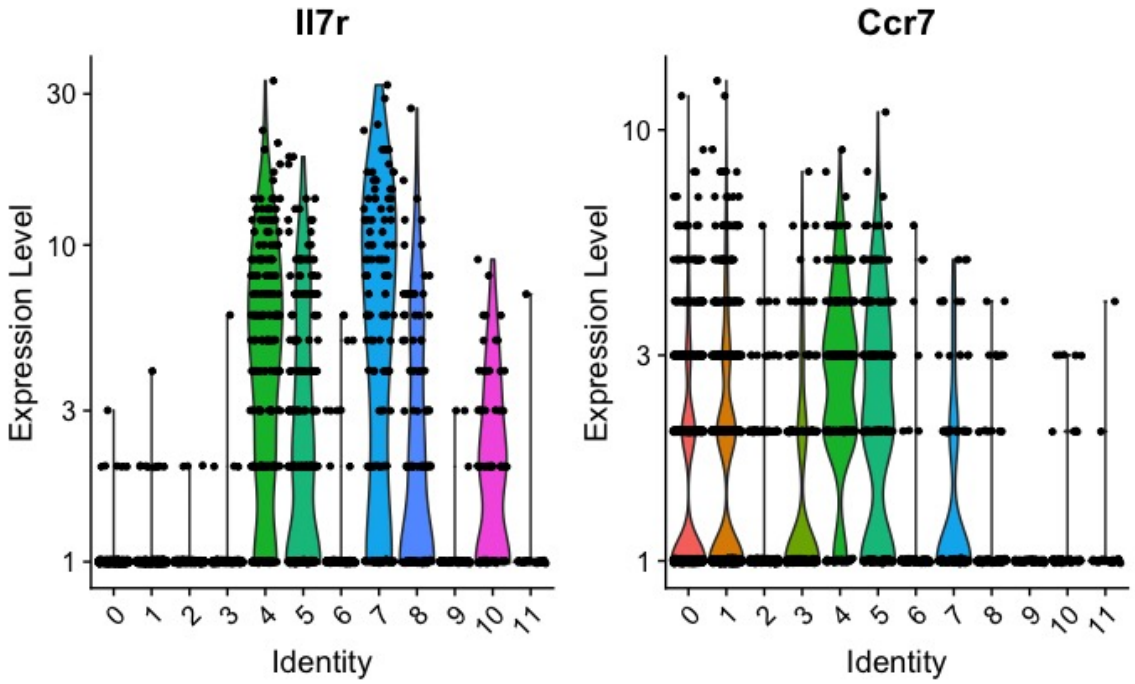
- Standard Seurat workflow
  - **QC** and selecting cells
  - Normalizing the data
  - Identification of highly variable features (feature selection)
  - Scaling the data
  - Perform linear dimensional reduction (**PCA**)
  - Determine the ‘dimensionality’ of the dataset
  - **Cluster** the cells
  - Run non-linear dimensional reduction (**UMAP**)
  - Finding differentially expressed features (**cluster biomarkers**)
  - Shows expression probability distributions across clusters
  - Assign cell type

Analysis on WT by Seurat

Though the standard Seurat workflow, we get a low dimensional spcae of umap.

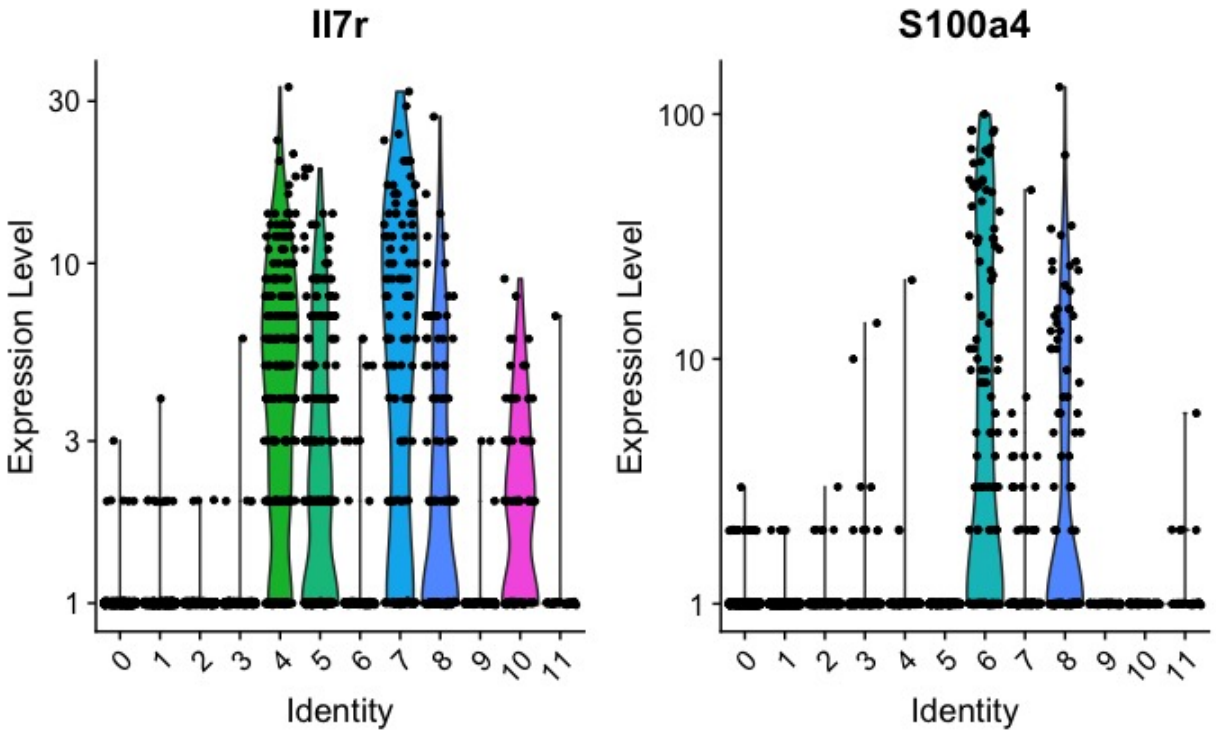


Use canonical markers to match the cell types



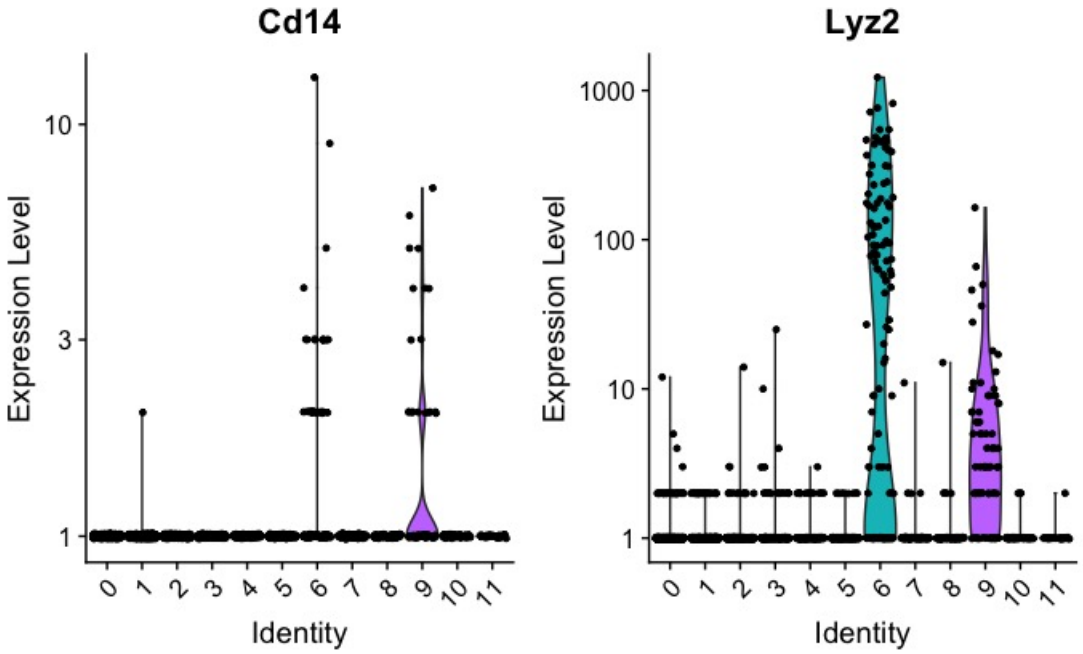
Cluster ID	Makers	Cell types
4,5,7	Il7r, Ccr7	Naive CD4+ T

Use canonical markers to match the cell types



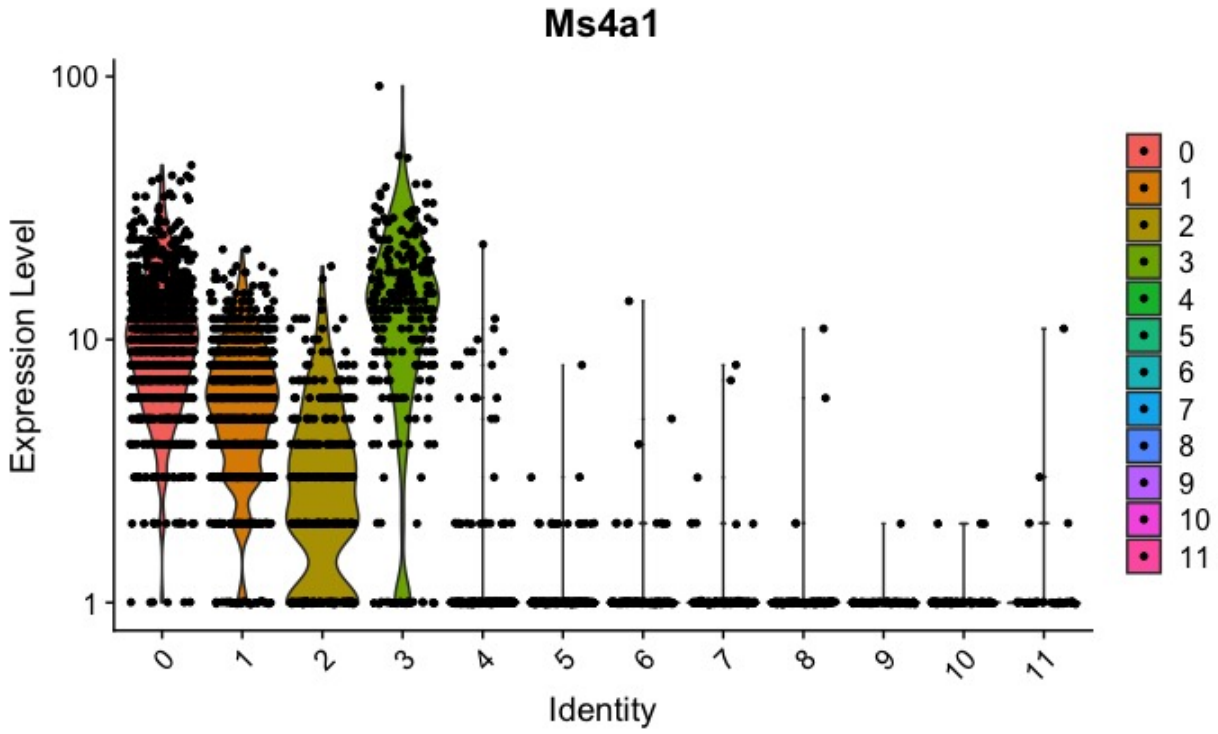
Cluster ID	Makers	Cell types
4,5,7	Il7r, Ccr7	Naive CD4+ T
8	Il7r,S100a4	Memory CD4+ T

Use canonical markers to match the cell types



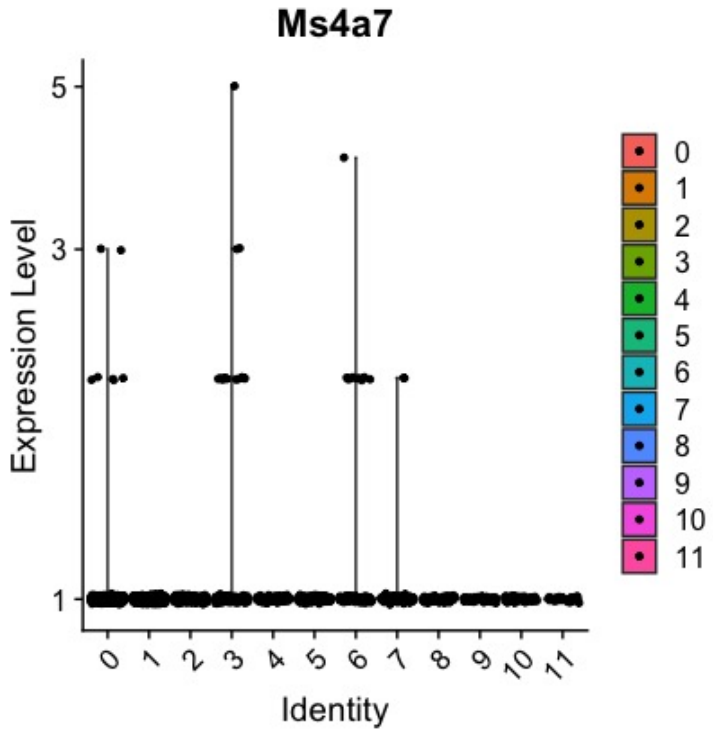
Cluster ID	Makers	Cell types
4,5,7	Il7r, Ccr7	Naive CD4+ T
8	Il7r,S100a4	Memory CD4+ T
6,9	Cd14,Lyz2	CD14+ Mono

Use canonical markers to match the cell types



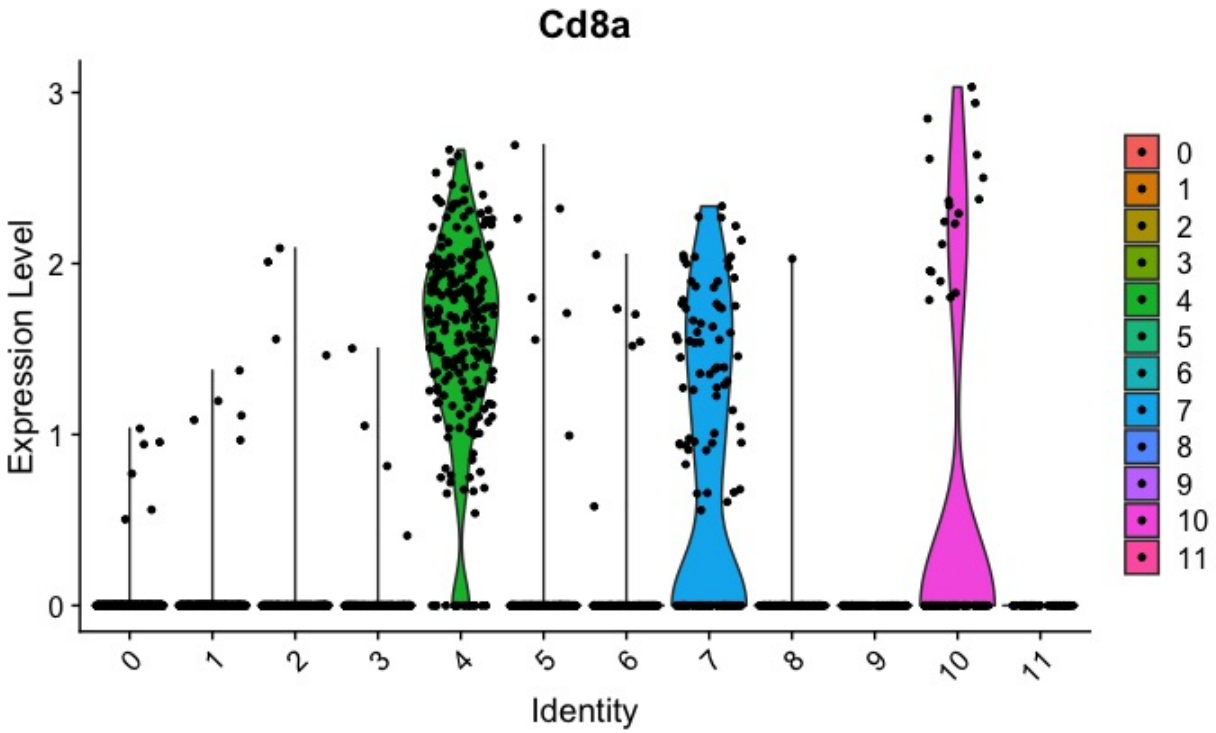
Cluster ID	Makers	Cell types
4,5,7	Il7r, Ccr7	Naive CD4+ T
8	Il7r,S100a4	Memory CD4+ T
6,9	Cd14,Lyz2	CD14+ Mono
0,1,2,3	Ms4a1	B

Use canonical markers to match the cell types



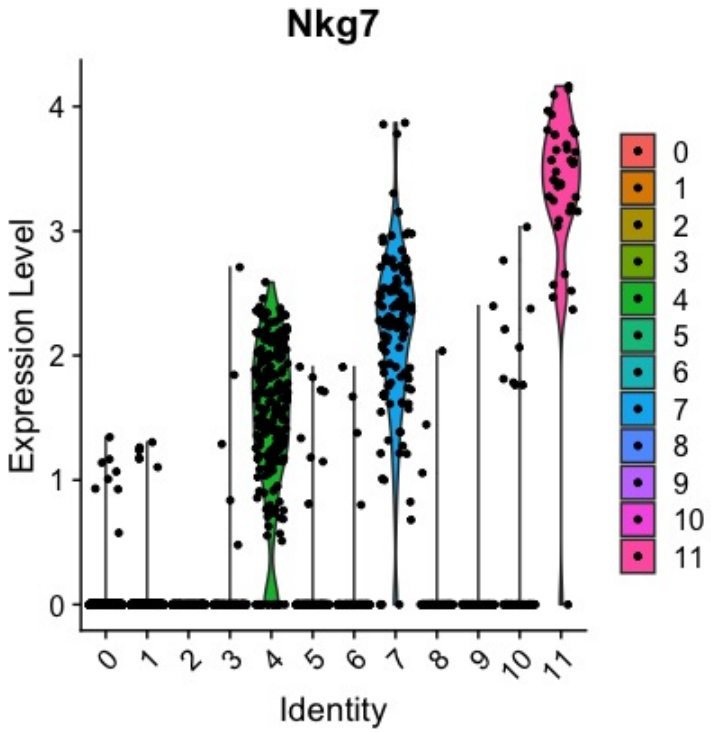
Cluster ID	Makers	Cell types
4,5,7	Il7r, Ccr7	Naive CD4+ T
8	Il7r,S100a4	Memory CD4+ T
6,9	Cd14,Lyz2	CD14+ Mono
0,1,2,3	Ms4a1	B
None	Ms4a7	FCGR3A+ Mono

Use canonical markers to match the cell types



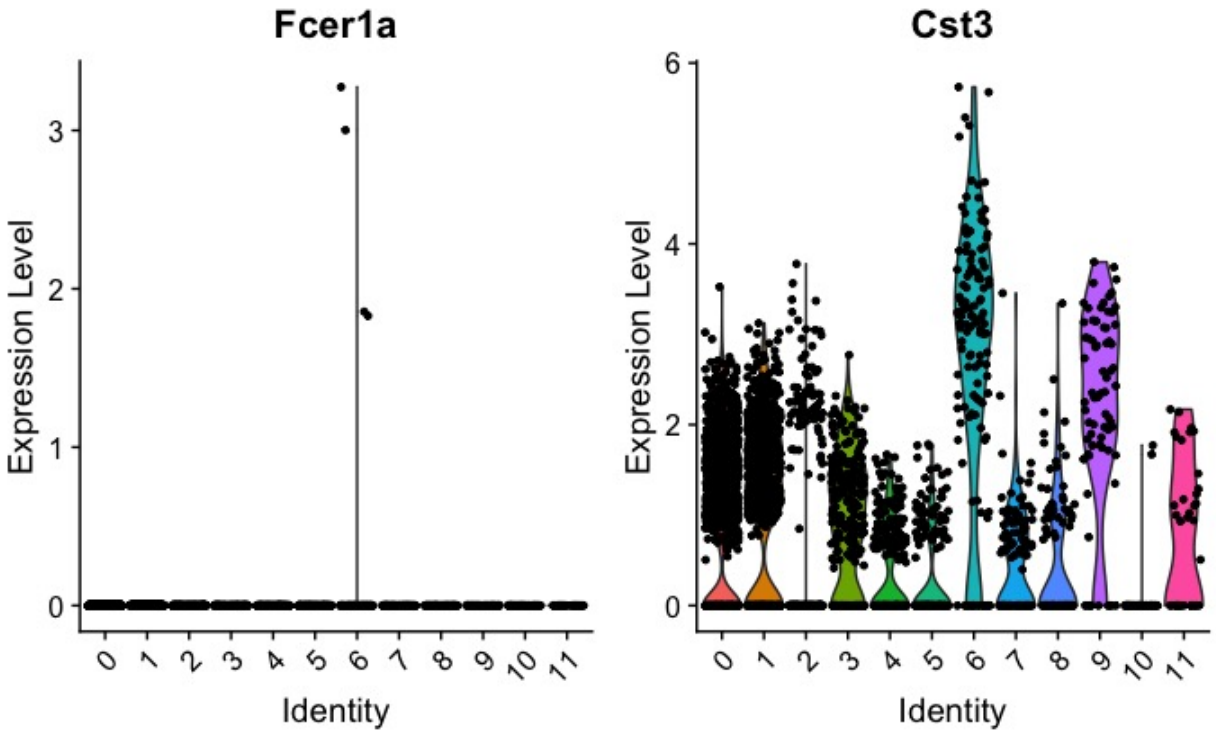
Cluster ID	Makers	Cell types
4,5,7	Il7r, Ccr7	Naive CD4+ T
8	Il7r,S100a4	Memory CD4+ T
6,9	Cd14,Lyz2	CD14+ Mono
0,1,2,3	Ms4a1	B
None	Ms4a7	FCGR3A+ Mono
4, 7, 10	Cd8a	CD8+ T

Use canonical markers to match the cell types



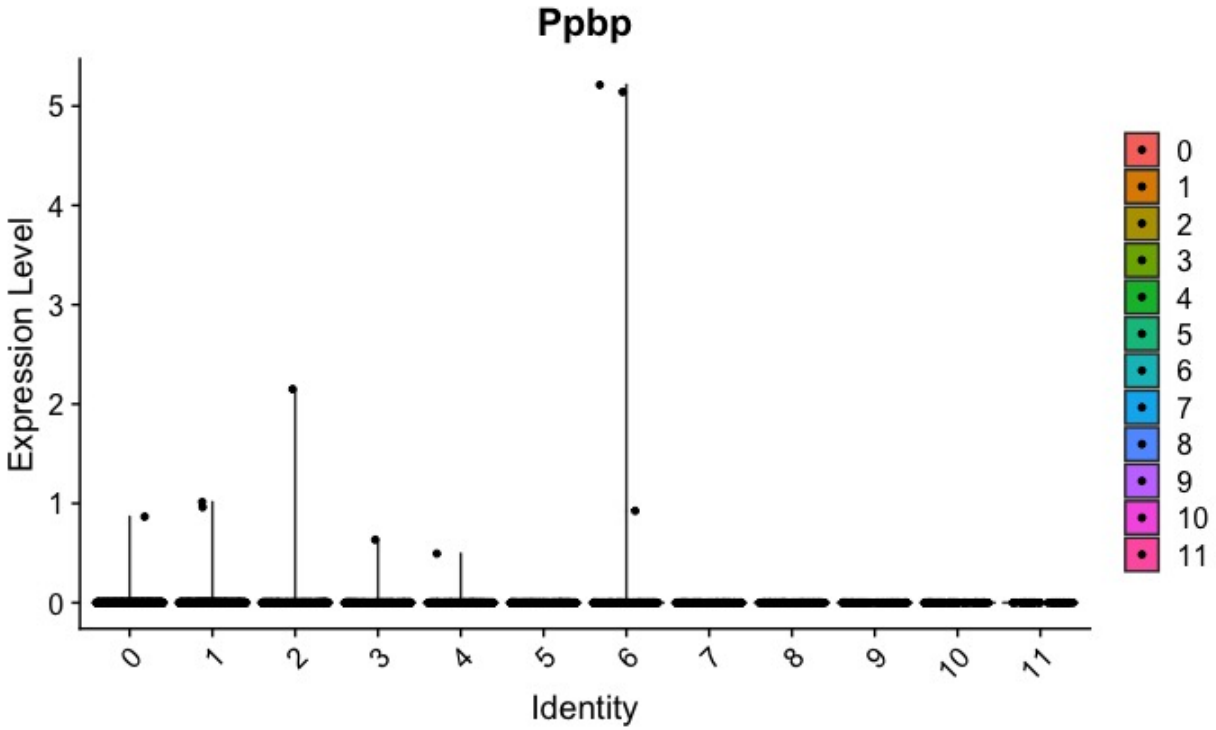
Cluster ID	Makers	Cell types
4,5,7	Il7r, Ccr7	Naive CD4+ T
8	Il7r,S100a4	Memory CD4+ T
6,9	Cd14,Lyz2	CD14+ Mono
0,1,2,3	Ms4a1	B
None	Ms4a7	FCGR3A+ Mono
4, 7, 10	Cd8a	CD8+ T
4, 7, 11	Nkg7	NK

Use canonical markers to match the cell types



Cluster ID	Makers	Cell types
4,5,7	Il7r, Ccr7	Naive CD4+ T
8	Il7r,S100a4	Memory CD4+ T
6,9	Cd14,Lyz2	CD14+ Mono
0,1,2,3	Ms4a1	B
None	Ms4a7	FCGR3A+ Mono
4, 7, 10	Cd8a	CD8+ T
4,7,11	Nkg7	NK
None	Fcer1a, Cst3	DC

Use canonical markers to match the cell types



Cluster ID	Makers	Cell types
4,5,7	Il7r, Ccr7	Naive CD4+ T
8	Il7r,S100a4	Memory CD4+ T
6,9	Cd14,Lyz2	CD14+ Mono
0,1,2,3	Ms4a1	B
None	Ms4a7	FCGR3A+ Mono
4, 7, 10	Cd8a	CD8+ T
4,7,11	Nkg7	NK
None	Fcer1a, Cst3	DC
None	Ppbp	Platelet

Use canonical markers to match the cell types

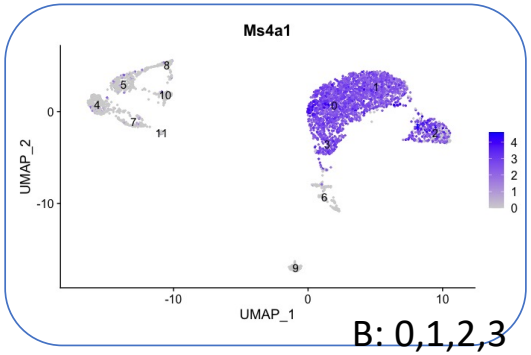
Cluster ID	Makers	Cell types
4, 5, 7	Il7r, Ccr7	Naive CD4+ T
8	Il7r,S100a4	Memory CD4+ T
6,9	Cd14,Lyz2	CD14+ Mono
0,1,2,3	Ms4a1	B
None	Ms4a7	FCGR3A+ Mono
4, 7, 10	Cd8a	CD8+ T
4,7,11	Nkg7	NK
None	Fcer1a, Cst3	DC
None	Ppbp	Platelet

Arrangement

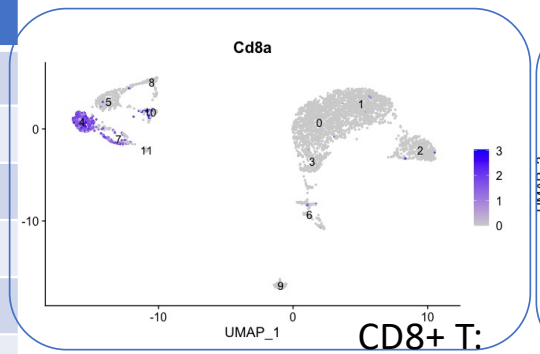
Cluster ID	Makers	Cell types
0,1,2,3	Ms4a1	B
5	Il7r, Ccr7	Naive CD4+ T
6	Cd14,Lyz2	CD14+ Mono
4,7	Cd8a	CD8+ T
8	Il7r,S100a4	Memory CD4+ T
9	Cd14,Lyz2	CD14+ Mono
10	Cd8a	CD8+ T
11	Nkg7	NK

Check FeaturePlot

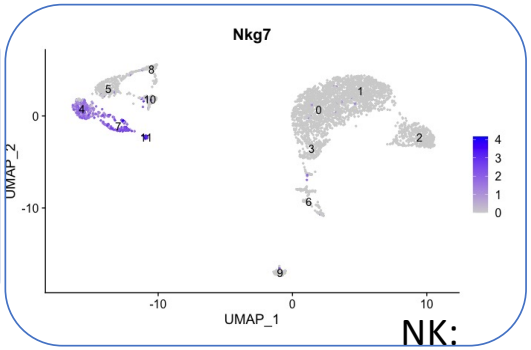
Cluster ID	Makers	Cell types
0,1,2,3	Ms4a1	B
5	Il7r, Ccr7	Naive CD4+ T
6	Cd14,Lyz2	CD14+ Mono
4,7	Cd8a	CD8+ T
8	Il7r,S100a4	Memory CD4+ T
9	Cd14,Lyz2	CD14+ Mono
10	Cd8a	CD8+ T
11	Nkg7	NK



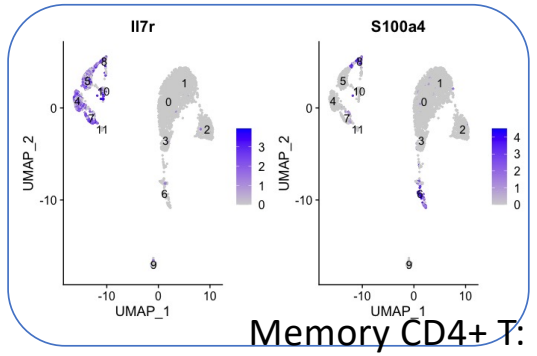
B: 0,1,2,3



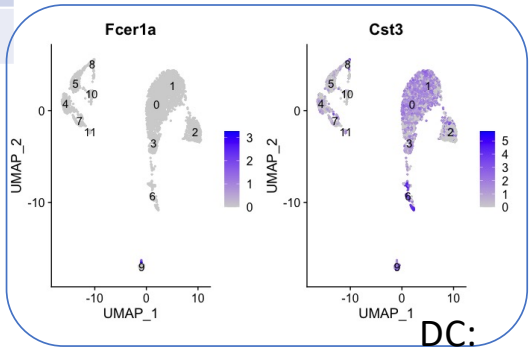
CD8+ T: 4,7



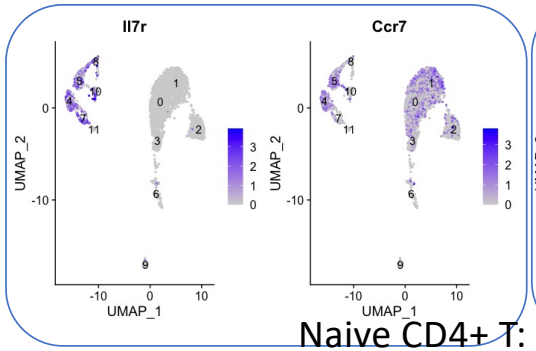
NK: 11



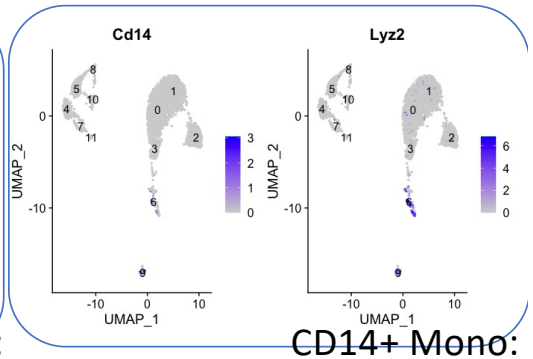
Memory CD4+ T: 8



DC: None

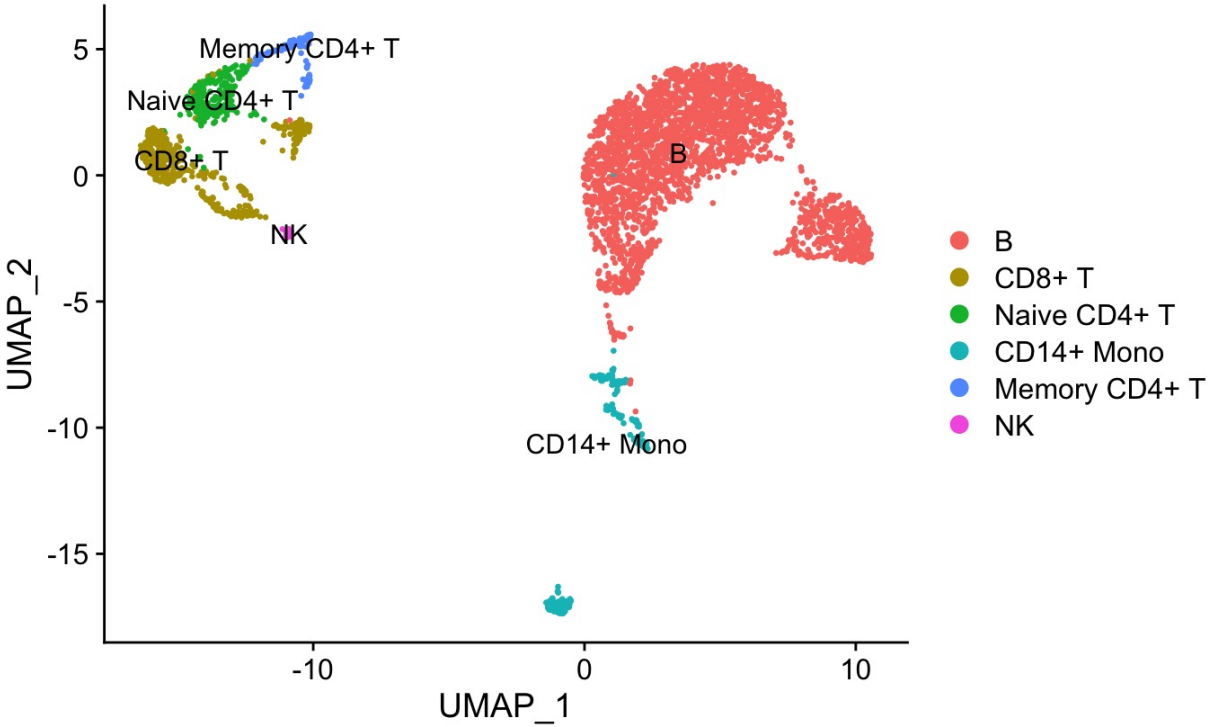


Naive CD4+ T: 4,5



CD14+ Mono: 6, 9

Final label on WT by Seurat

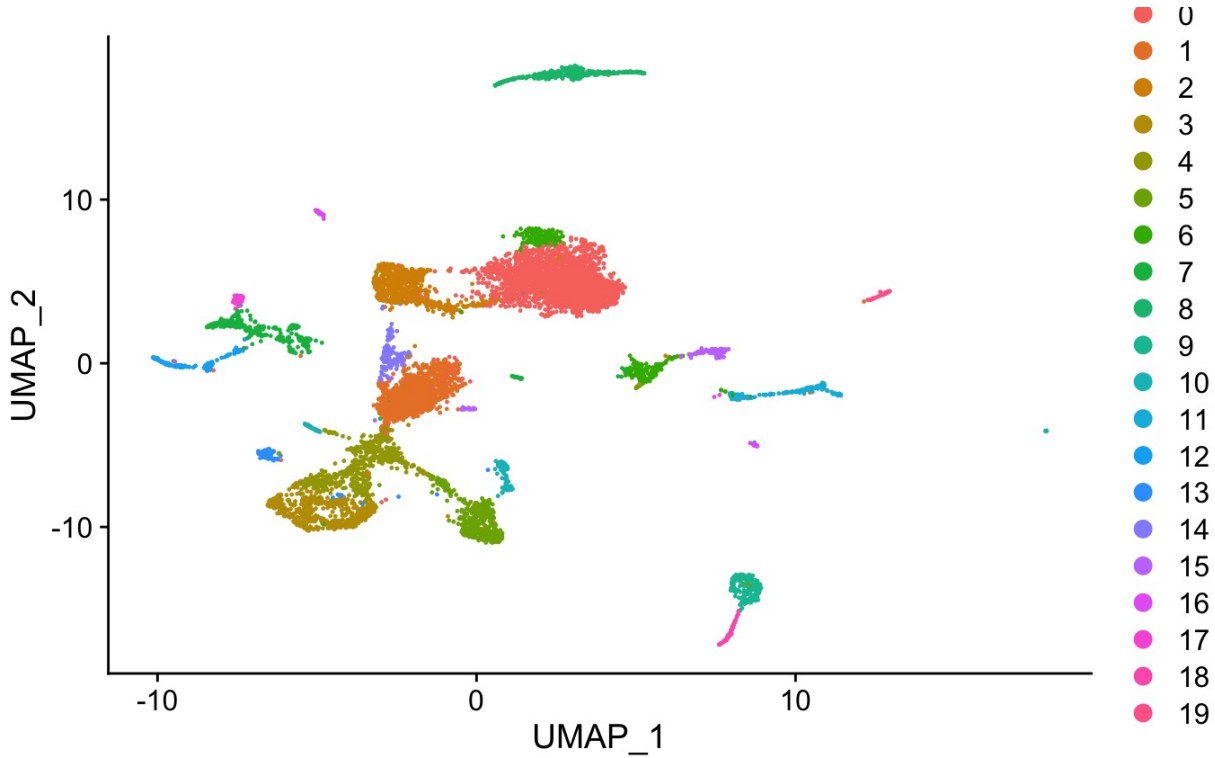


Augment:

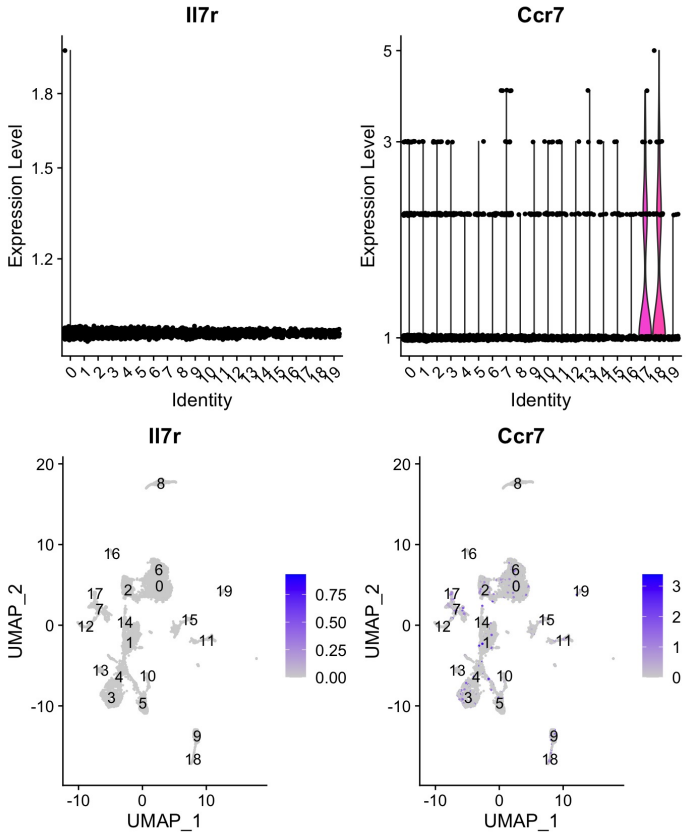
- If we compare WT and RT in the same space, we will assume that the structs of two data set are similar.
- If we make WT and RT overlap in a certain space, the overlapped parts should share similar structs, i.e., B cell of WT and B cell of RT should overlap.
- However, I found B and T cells may not exist in RT, which rejects the original assumption and indicate that RT set may be full of errors.

Analysis on raw data of RT

- Though the standard Seurat workflow, we get a low dimensional spcae of umap.

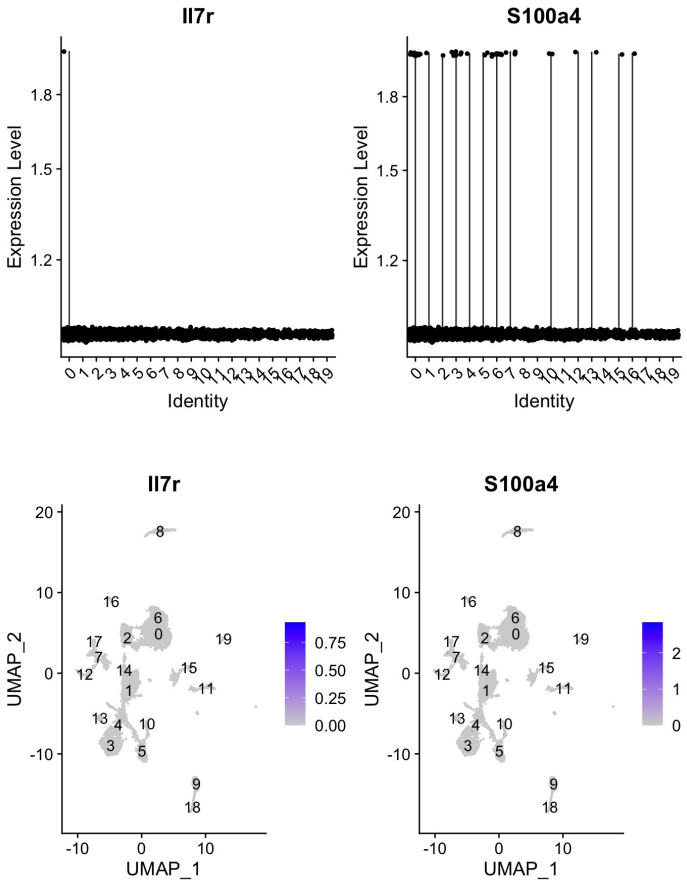


Use canonical markers to match the cell types For RT



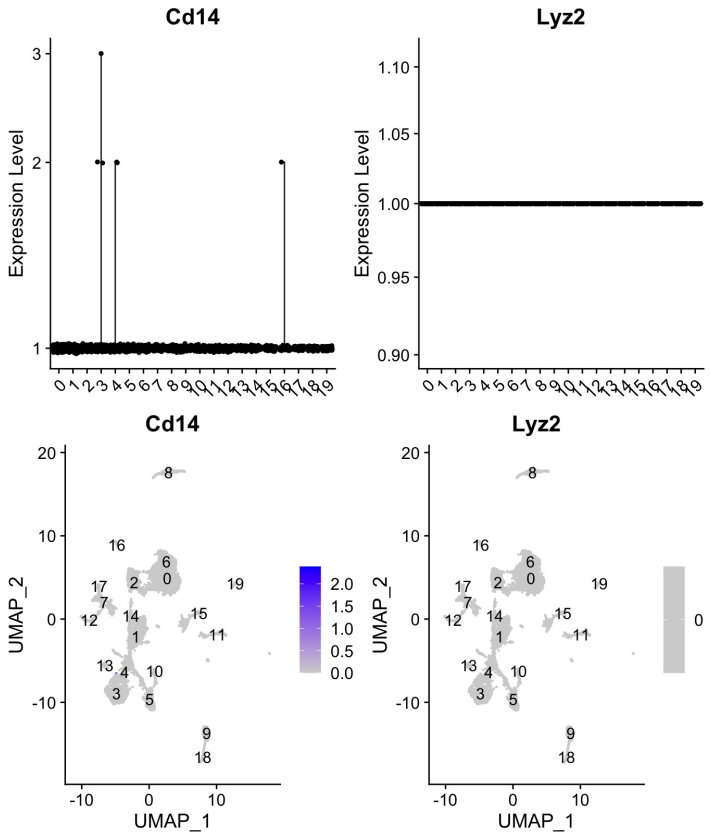
Cluster ID	Makers	Cell types
None	Il7r, Ccr7	Naive CD4+ T

Use canonical markers to match the cell types For RT



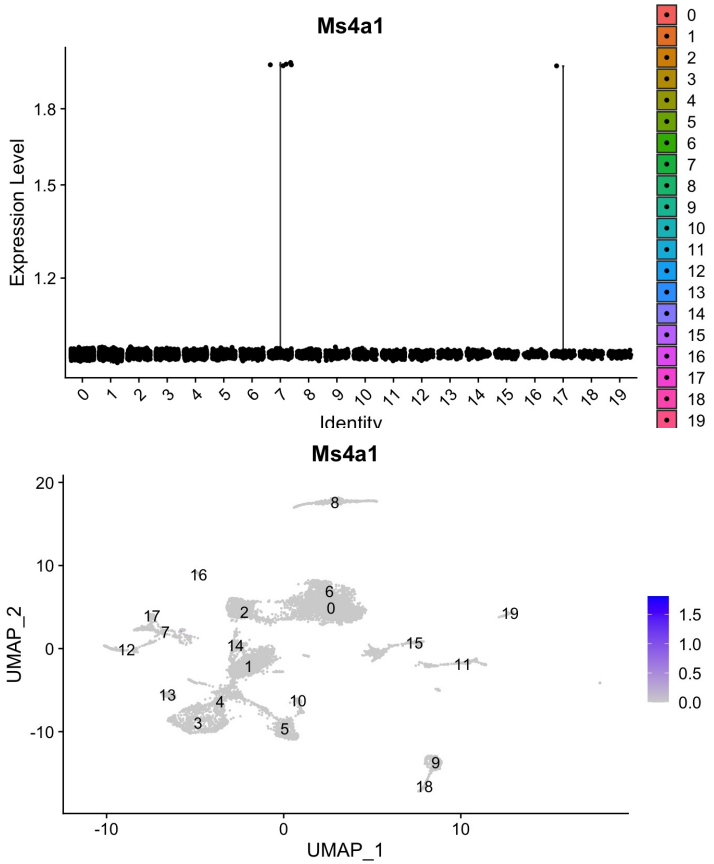
Cluster ID	Makers	Cell types
None	Il7r, Ccr7	Naive CD4+ T
None	Il7r,S100a4	Memory CD4+ T

Use canonical markers to match the cell types For RT



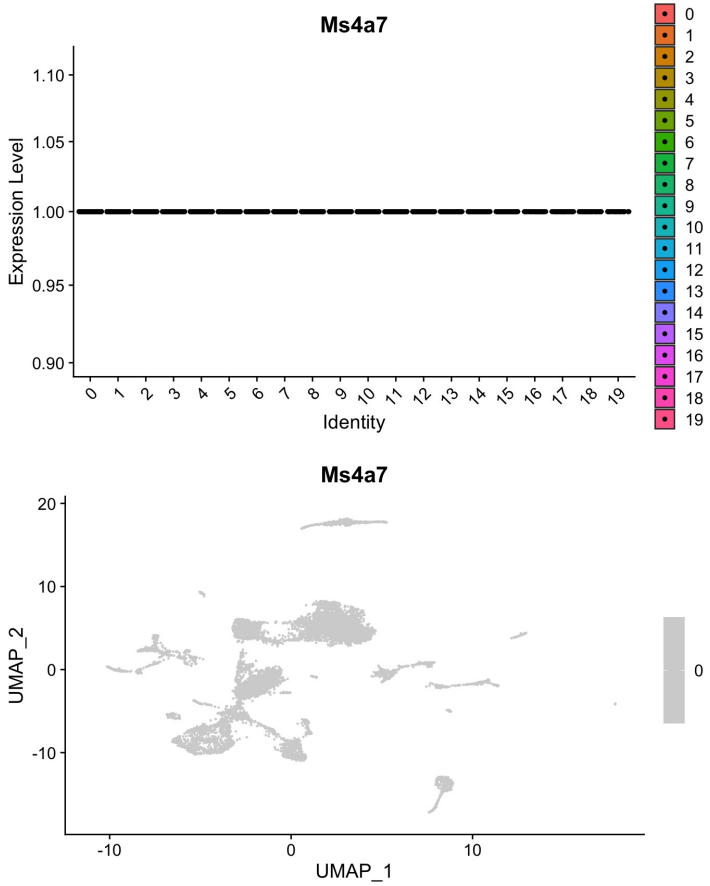
Cluster ID	Makers	Cell types
None	Il7r, Ccr7	Naive CD4+ T
None	Il7r,S100a4	Memory CD4+ T
None	Cd14,Lyz2	CD14+ Mono

Use canonical markers to match the cell types For RT



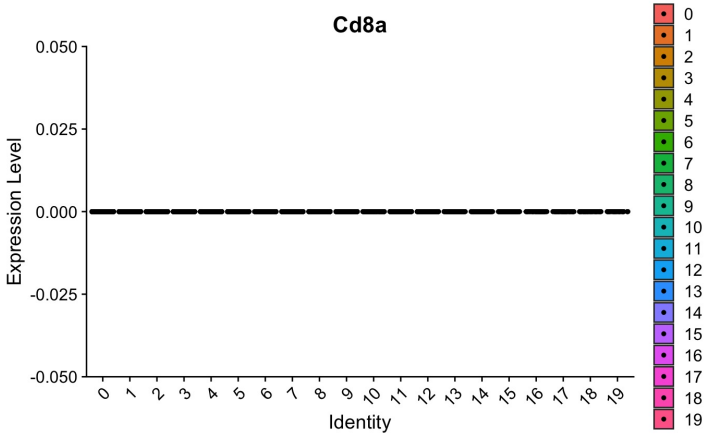
Cluster ID	Makers	Cell types
None	Il7r, Ccr7	Naive CD4+ T
None	Il7r,S100a4	Memory CD4+ T
None	Cd14,Lyz2	CD14+ Mono
None	Ms4a1	B

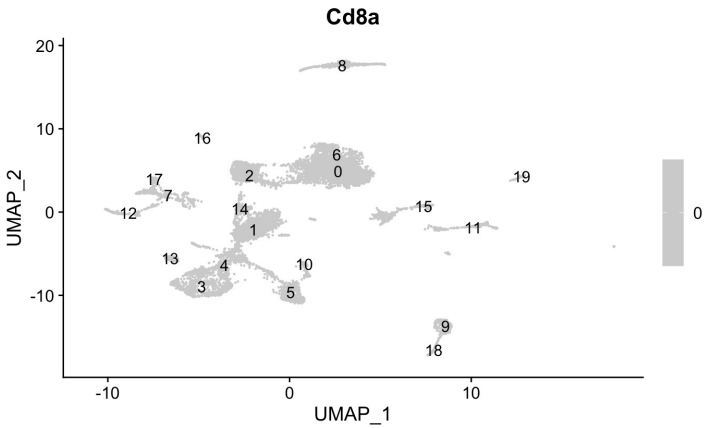
Use canonical markers to match the cell types For RT



Cluster ID	Makers	Cell types
None	Il7r, Ccr7	Naive CD4+ T
None	Il7r,S100a4	Memory CD4+ T
None	Cd14,Lyz2	CD14+ Mono
None	Ms4a1	B
None	Ms4a7	FCGR3A+ Mono

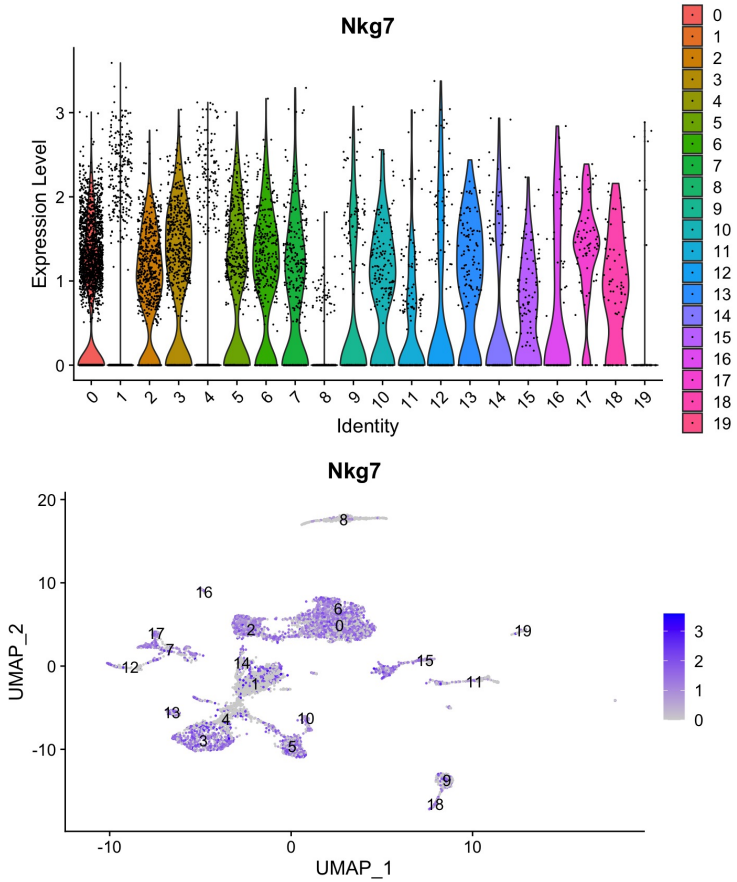
Use canonical markers to match the cell types For RT





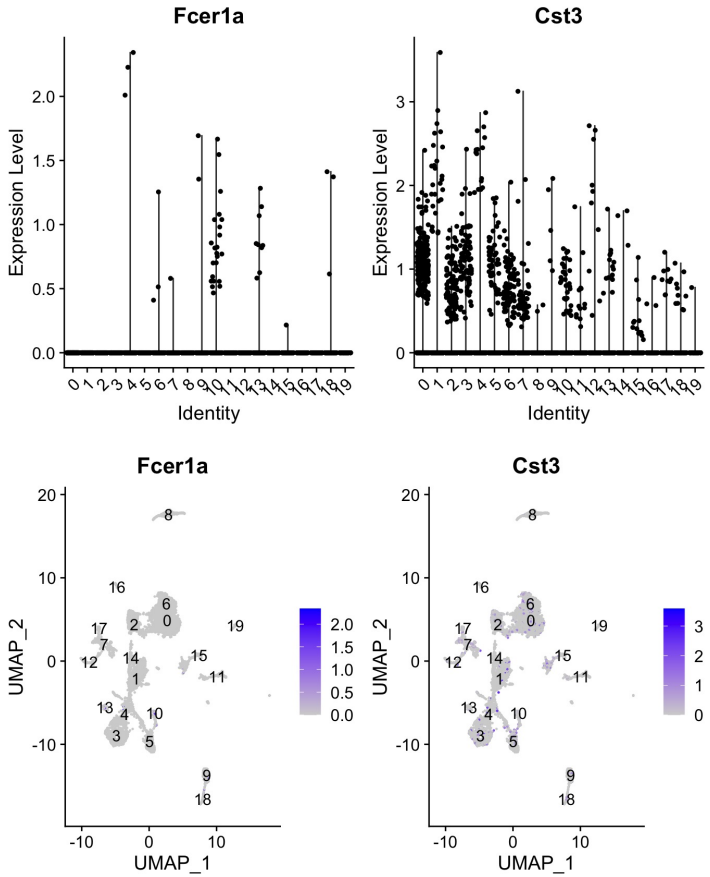
Cluster ID	Makers	Cell types
None	Il7r, Ccr7	Naive CD4+ T
None	Il7r,S100a4	Memory CD4+ T
None	Cd14,Lyz2	CD14+ Mono
None	Ms4a1	B
None	Ms4a7	FCGR3A+ Mono
None	Cd8a	CD8+ T

Use canonical markers to match the cell types For RT



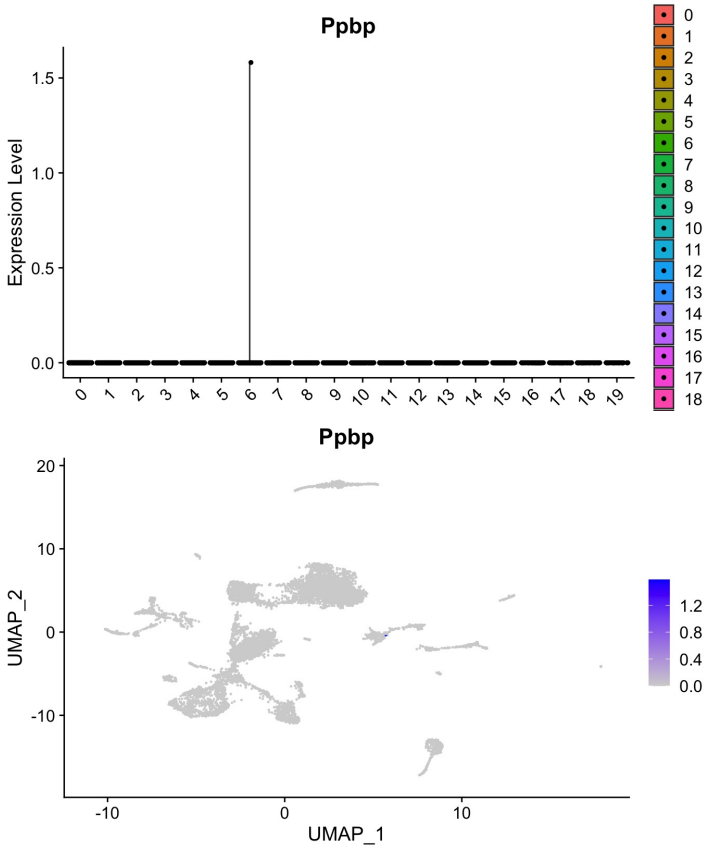
Cluster ID	Makers	Cell types
None	Il7r, Ccr7	Naive CD4+ T
None	Il7r,S100a4	Memory CD4+ T
None	Cd14,Lyz2	CD14+ Mono
None	Ms4a1	B
None	Ms4a7	FCGR3A+ Mono
None	Cd8a	CD8+ T
0,2,3,5,6, 7,9,10,11, 12,13,14, 15,16,17, 18	Nkg7	NK

Use canonical markers to match the cell types For RT



Cluster ID	Makers	Cell types
None	Il7r, Ccr7	Naive CD4+ T
None	Il7r,S100a4	Memory CD4+ T
None	Cd14,Lyz2	CD14+ Mono
None	Ms4a1	B
None	Ms4a7	FCGR3A+ Mono
None	Cd8a	CD8+ T
0,2,3,5,6, 7,9,10,11, 12,13,14, 15,16,17, 18	Nkg7	NK
None	Fcer1a, Cst3	DC

Use canonical markers to match the cell types For RT



Cluster ID	Makers	Cell types
None	Il7r, Ccr7	Naive CD4+ T
None	Il7r,S100a4	Memory CD4+ T
None	Cd14,Lyz2	CD14+ Mono
None	Ms4a1	B
None	Ms4a7	FCGR3A+ Mono
None	Cd8a	CD8+ T
0,2,3,5,6, 7,9,10,11, 12,13,14, 15,16,17, 18	Nkg7	NK
None	Fcer1a, Cst3	DC
None	Ppbp	Platelet

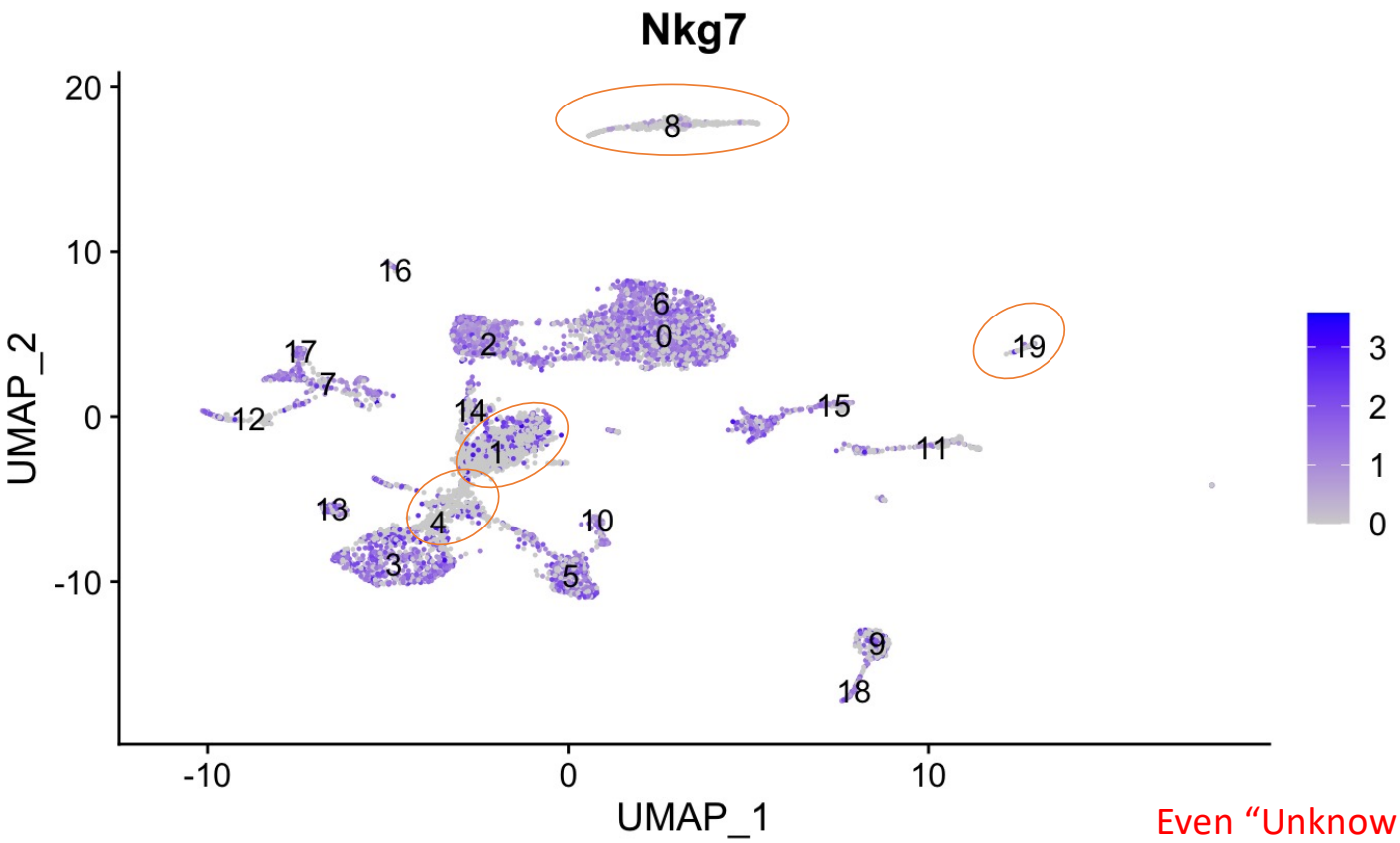
Use canonical markers to match the cell types

Cluster ID	Makers	Cell types
None	Il7r, Ccr7	Naive CD4+ T
None	Il7r,S100a4	Memory CD4+ T
None	Cd14,Lyz2	CD14+ Mono
None	Ms4a1	B
None	Ms4a7	FCGR3A+ Mono
None	Cd8a	CD8+ T
0,2,3,5,6,7,9,10,11,12,13,14,15,16,17,18	Nkg7	NK
None	Fcer1a, Cst3	DC
None	Ppbp	Platelet

Arrangement

Cluster ID	Makers	Cell types
0,2,3,5,6,7,9,10,11,12,13,14,15,16,17,18	Nkg7	NK
1,4,8,19	Unknown	Unknown

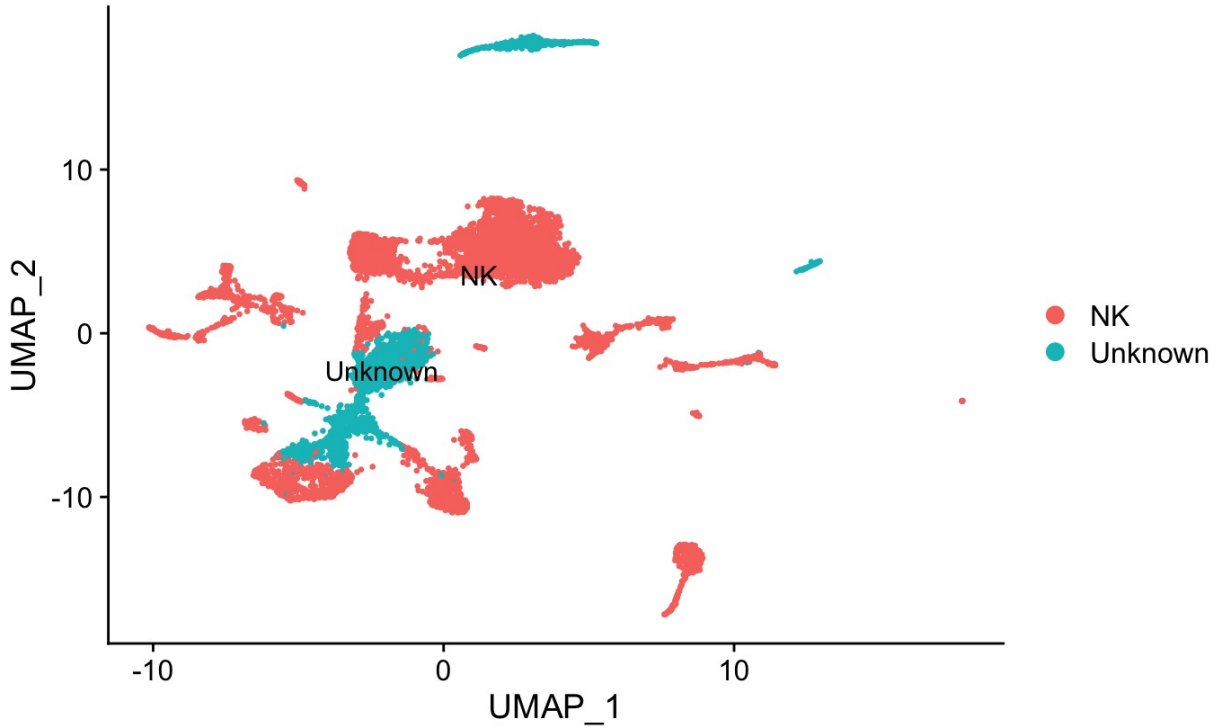
Check the FeaturePlot on RT



Cluster ID	Makers	Cell types
0,2,3,5,6,7,9, 10,11,12,13, 14,15,16,17, 18	Nkg7	NK
1,4,8,19	Unknown	Unknown

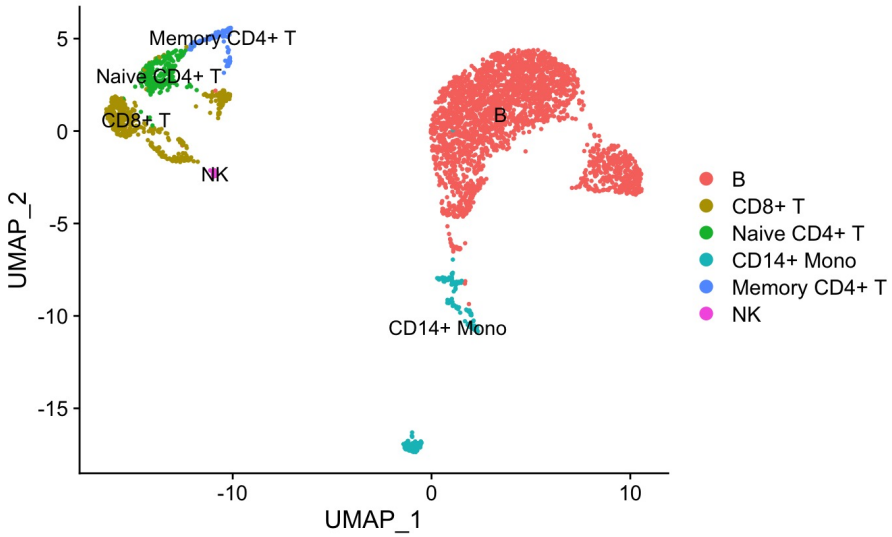
Even “Unknown” clusters contain some NK cells.

Final label on WT by Seurat

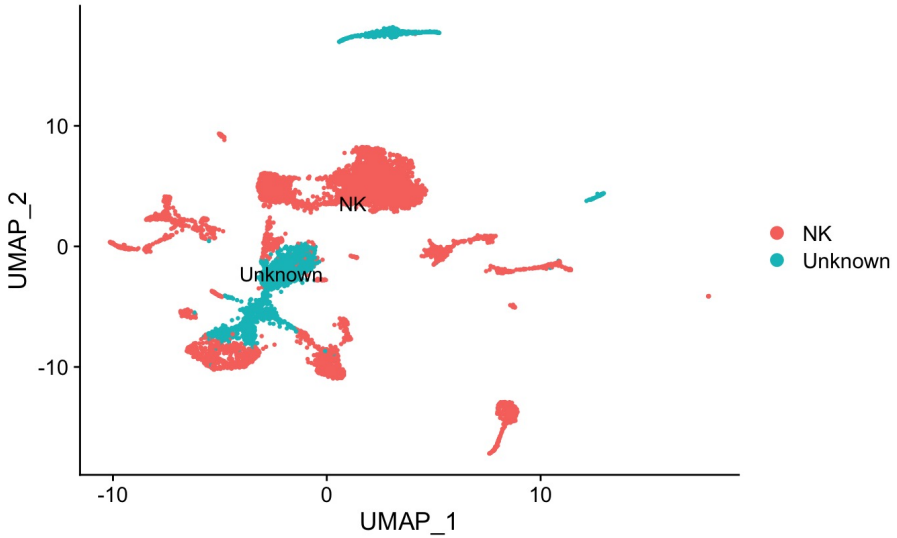


Compare the results between WT and RT

WT



RT



- The cell types in WT and RT seem so different that we may not plot them in the same space correctly.
- This result also explain why the two sets can not overlap in integrating experiments.

Summary

- We conduct several analysis on WT and RT datasets by Seurat.
- While WT includes multiple cell types, RT only contains one cell types (NK).
- Consider cells in RT are tumour cell, we can suspect there is certain "relation" between NK and tumor. Fortunately, the relation between NK cell activity and suppression of tumor occurrence has been documented in several studies (see Shimasaki et al. 2020).
- The different distubtion between WT and RT may destroy the base assumption of making WT and RT overlap in the same space and inspire new thinking.

Shimasaki, N., Jain, A. & Campana, D. NK cells for cancer immunotherapy. *Nat Rev Drug Discov* **19,** 200–218 (2020). https://doi.org/10.1038/s41573-019-0052-1