

Seurat_ to RISC

Yang Liu, Tao Wang, Deyou Zheng

6/27/2021

R Markdown

We strongly suggest to clean the gene-cell matrix when using RISC, as shown in the vignettes. The RISC is based on the gene expression variance which could be severely distorted by abnormal genes with extremely large number of UMIs.

Here we show how to transfer Seurat object to RISC object without filtering abnormal genes, we get the consistent result. The data can be download from Seurat website:
https://satijalab.org/seurat/articles/pbm3k_tutorial.html

The analysis is completely based on Seurat's parameters and cell-population information, not optimized for RISC.

Seurat v4

```
library(dplyr)

##

## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##   filter, lag

## The following objects are masked from 'package:base':
##   intersect, setdiff, setequal, union

library(Seurat)

## Attaching SeuratObject

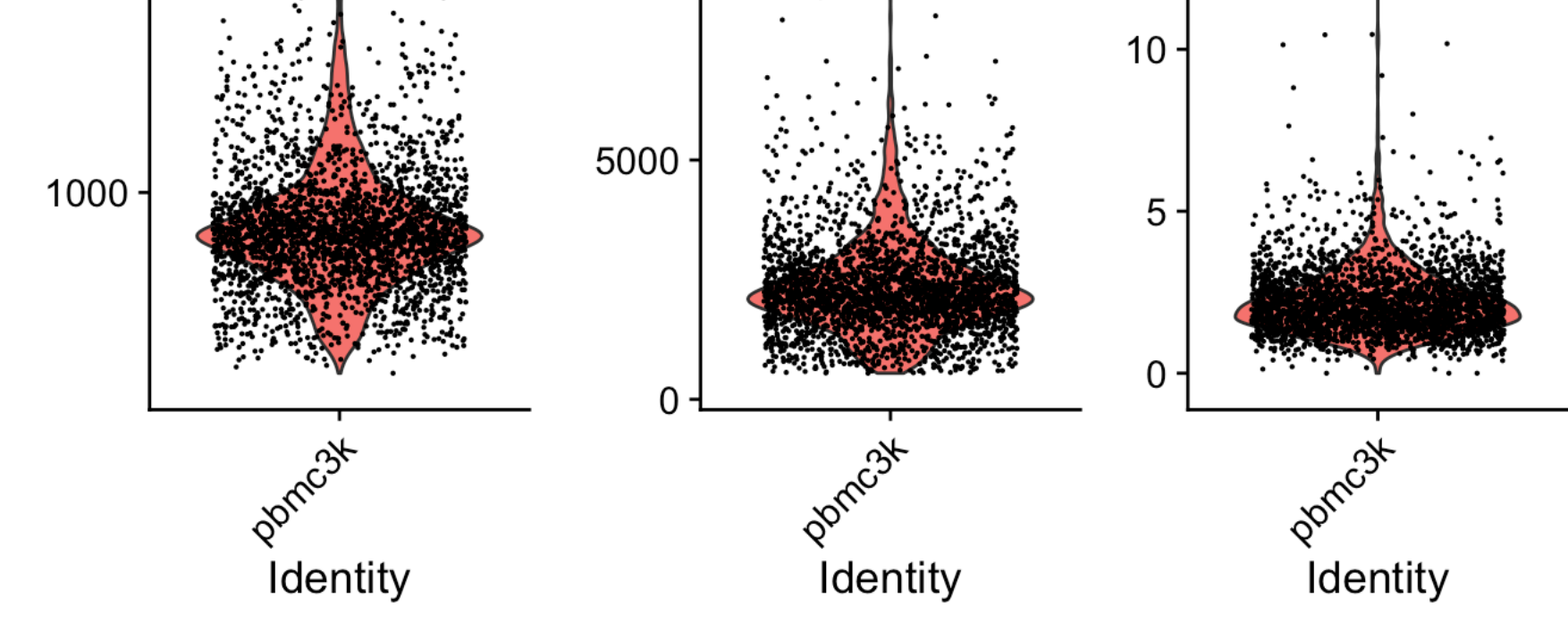
library(patchwork)
library(ggplot2)
library(RColorBrewer)

pbmc.data = Read10X(data.dir = "/Users/yanliu/Desktop/PBM3_1k") # change directory path to your own
pbmc = CreateSeuratObject(counts = pbmc.data, project = "pbmc3k", min.cells = 3, min.features = 200)

## Warning: Feature names cannot have underscores ('_'), replacing with dashes
## ('-')

pbmc[["percent.mt"]] = PercentageFeatureSet(pbmc, pattern = "MT-")

# Plot1
plot1 = FeatureScatter(pbmc, feature1 = "nCount_RNA", feature2 = "percent.mt")
# Plot2
plot2 = FeatureScatter(pbmc, feature1 = "nCount_RNA", feature2 = "nFeature_RNA")
# Plot1 + Plot2
```



```
pbmc = subset(pbmc, subset = nFeature_RNA > 200 & nFeature_RNA < 2500 & percent.mt < 5)
pbmc = NormalizeData(pbmc)
pbmc = FindVariableFeatures(pbmc, selection.method = "vst", nfeatures = 2000)
all.genes = rownames(pbmc)
pbmc = ScaleData(pbmc, features = all.genes)

## Centering and scaling data matrix

pbmc = RunPCA(pbmc, features = VariableFeatures(object = pbmc))

## PC_1
## Positive: CST3, TYROBP, LST1, AIF1, FTL, FTH1, LY2, FCN1, S100A9, TYMP
## FCRIG, CFD, LGALS1, S100A8, CTSS, LGALS2, SERPINA1, IFITM3, SPI1, CFP
## PSAP, IFI30, SAT1, COTL1, S100A11, NPC2, GRN, LGALS3, GSTP1, PTCAR0
## Negative: MALAT1, LTB, IL12, IL7R, CD2, B2M, ACAP1, CD27, STK17A, CTSSW
## CD247, GIMAP5, AQP3, CCL5, SELL, TRAF3IP3, GZMA, MAL, CST7, ITM2A
## MYC, GIMAP7, HOPX, BE2, LDLRAP1, GZMK, ETS1, ZAP70, TNFAIP8, R1C3
## PC_2
## Positive: CD79A, MS4A1, TCL1A, HLA-DQA1, HLA-DQB1, HLA-DMA, LINC00926, CD79B, HLA-DRB1, CD74
## HLA-DMA, HLA-DRB1, HLA-DQA2, CD37, HLA-DRB5, HLA-DNB, HLA-DPA1, FCRLA, HVGN1, LTB
## BLNK, P2RX5, IGLL5, IRF8, SHAP7, ARHGAP24, FCR2B, SHIM4, PPP1R14A, C16orf74
## Negative: NKG7, PRF1, CST7, GZMB, GZMA, PGFBP2, CTSSW, GNLY, B2M, SPON2
## CCL4, GZMH, FCGR3A, CCL5, CD247, XCL2, CLIC3, AKR1C3, SRGN, HOPX
## TYC38, APMAP, CTSC, S100A4, IGFBP7, ANXA1, ID2, IL32, XCL1, RHOC
## PC_3
## Positive: HLA-DQA1, CD79A, CD79B, HLA-DQB1, HLA-DPA1, CD74, MS4A1, HLA-DRB1, HLA-DRA
## HLA-DRB5, HLA-DQA2, TCL1A, LINC00926, HLA-DMA, CD37, HVGN1, FCRLA, IRF8
## PLAC8, BLNK, MALAT1, SMIM14, PLD4, LAT2, IGLL5, P2RX5, SHAP70, FCGR2B
## Negative: PSBP, PF4, SDRB, SPARC, GNG11, NRGN, GPR, RGS18, TUBB1, CLU
## HIST1H2AC, AP01189.4, ITGA2B, CD9, TMEM40, PTCRA, CA2, ACRRP, MMD, TREML1
## NGRAP1, F13A1, SEPT5, RUFY1, TSC2D1, MPP1, CMT5, RP11-367G6.3, MYL9, GF18A
## PC_4
## Positive: HLA-DQA1, CD79B, CD79A, MS4A1, HLA-DQB1, CD74, HLA-DPA1, HIST1H2AC, PF4, TCL1A
## SDRB, HLA-DPA1, HLA-DRB1, HLA-DQA2, HLA-DRA, PPPB, LINC00926, GNG11, HLA-DRB5, SPARC
## GPR, AP01189.4, CA2, PTCRA, CD9, NRGN, RGS18, GZMB, CLU, TUBB1
## Negative: VIM, IL7R, S100A6, IL32, S100A8, S100A4, GIMAP7, S100A10, S100A9, MAL
## AQP3, CD2, CD14, FIB, LGALS2, GIMAP4, ANXA1, CD27, FCN1, RSP7
## LY2, S100A11, GIMAP5, MS4A6A, S100A12, FOLR3, TRAB2A, AIF1, IL8, IFI6
## PC_5
## Positive: GZMB, NKG7, S100A8, PGFBP2, GNLY, CCL4, CST7, PRF1, GZMA, SPON2
## GZMH, S100A9, LGALS2, CCL3, CTSSW, XCL2, CD14, CLIC3, S100A12, CCL5
## RSP7, MS4A6A, GSTP1, FOLR3, IGFBP7, TYROBP, TYC38, AKR1C3, XCL1, HOPX
## Negative: LTB, IL7R, CD9, VIM, MS4A7, AQP3, CYTIP, RP11-290P20.3, SIGLEC10, HMOX1
## PTGES3, LILRB2, MAL, CD27, HNI, CD2, GD12, ANXA5, CORO1B, TUBA1B
## FAMIL10A, ATP1A1, TRAOB, PPA1, CCDC109B, ABRACL, CTD-2006K23.1, HARS, VMO1, FIB
```

```
VisDimLoadings(pbmc, dims = 1:2, reduction = "pca")
DimPlot(pbmc, reduction = "pca")
DimHeatmap(pbmc, dims = 1, cells = 500, balanced = TRUE)
```

