Assignment 3

Alessandro Giulivo, Francesca Rinaldi, Martina Monciotti

11/15/2021

```
library(ggplot2)
library(ggfortify)
library(factoextra)
```

```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

```
library(Rtsne)
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
```

Question 1

```
#Read data
setwd("C:/Users/aless/Desktop/Bioinformatics/Erasmus BCN/Data Visualisation/homeassignment")
P3<- read.delim("P3.tsv.gz", row.names=1)
P4<- read.delim("P4.tsv.gz", row.names=1)
P5<- read.delim("P5.tsv.gz", row.names=1)

#Check data
head(P3, 5)
```

```
##                    P2769_N701.S502 P2769_N701.S503 P2769_N701.S505
## ENSG00000000003.14               0               0               0
## ENSG00000000005.5                0               0               0
## ENSG00000000419.12              39              68              16
## ENSG00000000457.13               0               8              17
## ENSG00000000460.16               0              46               4
##                    P2769_N701.S506 P2769_N701.S507 P2769_N701.S508
## ENSG00000000003.14               0               0               0
```

```
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12              0               0               0
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              0               0               0
##                    P2769_N701.S510 P2769_N701.S511 P2769_N702.S502
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             23               2              24
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              0              11               7
##                    P2769_N702.S503 P2769_N702.S505 P2769_N702.S506
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             62              76               0
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              7               0               0
##                    P2769_N702.S507 P2769_N702.S508 P2769_N702.S510
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12              0               0              39
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              0               0             170
##                    P2769_N702.S511 P2769_N703.S502 P2769_N703.S503
## ENSG00000000003.14              0               0            0.00
## ENSG00000000005.5               0               0            0.00
## ENSG00000000419.12             88              29           92.00
## ENSG00000000457.13              0               0           28.79
## ENSG00000000460.16              0               0          187.21
##                    P2769_N703.S505 P2769_N703.S506 P2769_N703.S507
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             39              69               0
## ENSG00000000457.13             15               0               0
## ENSG00000000460.16              0               0               0
##                    P2769_N703.S508 P2769_N703.S510 P2769_N703.S511
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12              0              59              79
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              0               0               3
##                    P2769_N704.S502 P2769_N704.S503 P2769_N704.S505
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             19             111               9
## ENSG00000000457.13              0               1               0
## ENSG00000000460.16             45              30               0
##                    P2769_N704.S506 P2769_N704.S507 P2769_N704.S508
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12            162               0               0
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              0               0               0
##                    P2769_N704.S510 P2769_N704.S511 P2769_N705.S502
## ENSG00000000003.14              0               0               0
```

```
## ENSG00000000005.5                0                0                0
## ENSG00000000419.12              25                0              110
## ENSG00000000457.13               0                0               33
## ENSG00000000460.16               0              110               79
##                     P2769_N705.S503 P2769_N705.S505 P2769_N705.S506
## ENSG00000000003.14               0                0                0
## ENSG00000000005.5                0                0                0
## ENSG00000000419.12              74                9               16
## ENSG00000000457.13               0                0                0
## ENSG00000000460.16             185                0              240
##                     P2769_N705.S507 P2769_N705.S508 P2769_N705.S510
## ENSG00000000003.14               0                0                0
## ENSG00000000005.5                0                0                0
## ENSG00000000419.12              79               45                0
## ENSG00000000457.13               0                6                0
## ENSG00000000460.16               0                0                1
##                     P2769_N705.S511 P2769_N706.S502 P2769_N706.S503
## ENSG00000000003.14               0                0                0
## ENSG00000000005.5                0                0                0
## ENSG00000000419.12              17               53              182
## ENSG00000000457.13               0                0                0
## ENSG00000000460.16               0                0                0
##                     P2769_N706.S505 P2769_N706.S506 P2769_N706.S507
## ENSG00000000003.14               0                0                0
## ENSG00000000005.5                0                0                0
## ENSG00000000419.12              49              125                0
## ENSG00000000457.13               0                0                0
## ENSG00000000460.16               0                0                0
##                     P2769_N706.S508 P2769_N706.S510 P2769_N706.S511
## ENSG00000000003.14               0             0.00                0
## ENSG00000000005.5                0             0.00                0
## ENSG00000000419.12             175           110.00               29
## ENSG00000000457.13               0           185.82                0
## ENSG00000000460.16               0            63.18                0
##                     P2769_N707.S502 P2769_N707.S503 P2769_N707.S505
## ENSG00000000003.14               0                0                0
## ENSG00000000005.5                0                0                0
## ENSG00000000419.12             157               73               47
## ENSG00000000457.13               0                0                0
## ENSG00000000460.16              47              186                0
##                     P2769_N707.S506 P2769_N707.S507 P2769_N707.S508
## ENSG00000000003.14               0                0                0
## ENSG00000000005.5                0                0                0
## ENSG00000000419.12             207              114               84
## ENSG00000000457.13               0                0                0
## ENSG00000000460.16               0              170              152
##                     P2769_N707.S510 P2769_N707.S511 P2769_N710.S502
## ENSG00000000003.14               0                0                0
## ENSG00000000005.5                0                0                0
## ENSG00000000419.12              27               58              152
## ENSG00000000457.13               0               91                0
## ENSG00000000460.16              94               26                0
##                     P2769_N710.S503 P2769_N710.S505 P2769_N710.S506
## ENSG00000000003.14               0                0                0
```

```
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12            107               0              50
## ENSG00000000457.13             21               0               0
## ENSG00000000460.16              0               0               0
##                    P2769_N710.S507 P2769_N710.S508 P2769_N710.S510
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             82             176              68
## ENSG00000000457.13            134               0               0
## ENSG00000000460.16             55               0              87
##                    P2769_N710.S511 P2769_N711.S502 P2769_N711.S503
## ENSG00000000003.14              0            0.00               0
## ENSG00000000005.5               0            0.00               0
## ENSG00000000419.12              0          139.00              69
## ENSG00000000457.13              0           28.19               0
## ENSG00000000460.16              8           99.81               0
##                    P2769_N711.S505 P2769_N711.S506 P2769_N711.S507
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             73              81               0
## ENSG00000000457.13             79               0               8
## ENSG00000000460.16             17               0             147
##                    P2769_N711.S508 P2769_N711.S510 P2769_N711.S511
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12            285              84              60
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16             62              91              32
##                    P2769_N712.S502 P2769_N712.S503 P2769_N712.S505
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12            240              32              33
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16            118             100               0
##                    P2769_N712.S506 P2769_N712.S507 P2769_N712.S508
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             53             178               0
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              0               0               0
##                    P2769_N712.S510 P2769_N712.S511 P2769_N714.S502
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             48               0             103
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              0               0             123
##                    P2769_N714.S503 P2769_N714.S505 P2769_N714.S506
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             58              35               0
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              0             259               0
##                    P2769_N714.S507 P2769_N714.S508 P2769_N714.S510
## ENSG00000000003.14              0               0               0
```

```
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             86             114               0
## ENSG00000000457.13              0               0              10
## ENSG00000000460.16              0               0               0
##                    P2769_N714.S511 P2769_N715.S502 P2769_N715.S503
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             44             351              83
## ENSG00000000457.13             16               0               0
## ENSG00000000460.16              0               0             325
##                    P2769_N715.S505 P2769_N715.S506 P2769_N715.S507
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12            292             110             110
## ENSG00000000457.13              0               0             136
## ENSG00000000460.16              0               0               2
##                    P2769_N715.S508 P2769_N715.S510 P2769_N715.S511
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12              0              56               0
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              0               0              14
```

```
head(P4, 5)
```

```
##                    P2770_N701.S513 P2770_N701.S515 P2770_N701.S516
## ENSG00000000003.14              0            0.00               0
## ENSG00000000005.5               0            0.00               0
## ENSG00000000419.12             49           79.00             144
## ENSG00000000457.13              0           68.42               0
## ENSG00000000460.16              0          127.58               0
##                    P2770_N701.S517 P2770_N701.S518 P2770_N701.S520
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             21               0              87
## ENSG00000000457.13              2               0               0
## ENSG00000000460.16             13               0               0
##                    P2770_N701.S521 P2770_N701.S522 P2770_N702.S513
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12              0               0              75
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16            202               0              82
##                    P2770_N702.S515 P2770_N702.S516 P2770_N702.S517
## ENSG00000000003.14            0.00            0.00               0
## ENSG00000000005.5             0.00            0.00               0
## ENSG00000000419.12          151.00           28.00              83
## ENSG00000000457.13           16.55           30.68               0
## ENSG00000000460.16          114.45           12.32               0
##                    P2770_N702.S518 P2770_N702.S520 P2770_N702.S521
## ENSG00000000003.14              0               0            0.00
## ENSG00000000005.5               0               0            0.00
## ENSG00000000419.12             33             185           95.00
## ENSG00000000457.13              0               0           11.64
```

```
## ENSG00000000460.16              0             105          208.36
##                     P2770_N702.S522 P2770_N703.S513 P2770_N703.S515
## ENSG00000000003.14               0               0               0
## ENSG00000000005.5                0               0               0
## ENSG00000000419.12               1              88              39
## ENSG00000000457.13               0               0               9
## ENSG00000000460.16               0              10              51
##                     P2770_N703.S516 P2770_N703.S517 P2770_N703.S518
## ENSG00000000003.14               0               0               0
## ENSG00000000005.5                0               0               0
## ENSG00000000419.12              84              60             137
## ENSG00000000457.13               0               0               0
## ENSG00000000460.16               0              46              70
##                     P2770_N703.S520 P2770_N703.S521 P2770_N703.S522
## ENSG00000000003.14               0               0               0
## ENSG00000000005.5                0               0               0
## ENSG00000000419.12               0             103              34
## ENSG00000000457.13               0               0               0
## ENSG00000000460.16               0              72               0
##                     P2770_N704.S513 P2770_N704.S515 P2770_N704.S516
## ENSG00000000003.14               0               0               0
## ENSG00000000005.5                0               0               0
## ENSG00000000419.12              53             116              88
## ENSG00000000457.13               0               0               0
## ENSG00000000460.16              17              42               0
##                     P2770_N704.S517 P2770_N704.S518 P2770_N704.S520
## ENSG00000000003.14               0               0            0.00
## ENSG00000000005.5                0               0            0.00
## ENSG00000000419.12              59               0          138.00
## ENSG00000000457.13              33               0           35.02
## ENSG00000000460.16               0               0           78.98
##                     P2770_N704.S521 P2770_N704.S522 P2770_N705.S513
## ENSG00000000003.14               0               0               0
## ENSG00000000005.5                0               0               0
## ENSG00000000419.12               0               0              59
## ENSG00000000457.13               0               0             112
## ENSG00000000460.16              14               0               0
##                     P2770_N705.S515 P2770_N705.S516 P2770_N705.S517
## ENSG00000000003.14               0               0               0
## ENSG00000000005.5                0               0               0
## ENSG00000000419.12              95              89              29
## ENSG00000000457.13               0               0               0
## ENSG00000000460.16              56              41              87
##                     P2770_N705.S518 P2770_N705.S520 P2770_N705.S521
## ENSG00000000003.14               0               0               0
## ENSG00000000005.5                0               0               0
## ENSG00000000419.12              22             134              83
## ENSG00000000457.13               0               0               0
## ENSG00000000460.16              23               0               0
##                     P2770_N705.S522 P2770_N706.S513 P2770_N706.S515
## ENSG00000000003.14               0               0            0.00
## ENSG00000000005.5                0               0            0.00
## ENSG00000000419.12              28              44           64.00
## ENSG00000000457.13               0               0           22.64
```

```
## ENSG00000000460.16              0               0           68.36
##                     P2770_N706.S516 P2770_N706.S517 P2770_N706.S518
## ENSG00000000003.14              0            0.00               0
## ENSG00000000005.5               0            0.00               0
## ENSG00000000419.12             71           64.00               0
## ENSG00000000457.13              0          117.38               0
## ENSG00000000460.16              0          227.62               0
##                     P2770_N706.S520 P2770_N706.S521 P2770_N706.S522
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             96              59              46
## ENSG00000000457.13              0              29               5
## ENSG00000000460.16              0               0               0
##                     P2770_N707.S513 P2770_N707.S515 P2770_N707.S516
## ENSG00000000003.14              0               0            0.00
## ENSG00000000005.5               0               0            0.00
## ENSG00000000419.12             92              99           33.00
## ENSG00000000457.13              1              99            1.02
## ENSG00000000460.16            141               0          202.98
##                     P2770_N707.S517 P2770_N707.S518 P2770_N707.S520
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12            103              37              87
## ENSG00000000457.13              0             125               0
## ENSG00000000460.16              0               0              20
##                     P2770_N707.S521 P2770_N707.S522 P2770_N710.S513
## ENSG00000000003.14              0               0             0.0
## ENSG00000000005.5               0               0             0.0
## ENSG00000000419.12            104              97            81.0
## ENSG00000000457.13              0               0           120.8
## ENSG00000000460.16              0              87            64.2
##                     P2770_N710.S515 P2770_N710.S516 P2770_N710.S517
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             52              18              13
## ENSG00000000457.13              0               2               0
## ENSG00000000460.16              0               0               6
##                     P2770_N710.S518 P2770_N710.S520 P2770_N710.S521
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             23              58              79
## ENSG00000000457.13              0               0               1
## ENSG00000000460.16             51               0               0
##                     P2770_N710.S522 P2770_N711.S513 P2770_N711.S515
## ENSG00000000003.14              0            0.00               0
## ENSG00000000005.5               0            0.00               0
## ENSG00000000419.12            111           75.00               5
## ENSG00000000457.13              0           44.33               0
## ENSG00000000460.16              0          183.67               0
##                     P2770_N711.S516 P2770_N711.S517 P2770_N711.S518
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             41              44               0
## ENSG00000000457.13              0               0               0
```

```
## ENSG00000000460.16              0               0              28
##                     P2770_N711.S520 P2770_N711.S521 P2770_N711.S522
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             10               0              19
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              1               3               0
##                     P2770_N712.S513 P2770_N712.S515 P2770_N712.S516
## ENSG00000000003.14            0.00            0.00               0
## ENSG00000000005.5             0.00            0.00               0
## ENSG00000000419.12           78.00           65.00              17
## ENSG00000000457.13           14.69           30.62               0
## ENSG00000000460.16           77.31            2.38               0
##                     P2770_N712.S517 P2770_N712.S518 P2770_N712.S520
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12              0              37              41
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              0              92               0
##                     P2770_N712.S521 P2770_N712.S522 P2770_N714.S513
## ENSG00000000003.14            0.00               0               0
## ENSG00000000005.5             0.00               0               0
## ENSG00000000419.12           24.00              50              60
## ENSG00000000457.13           15.88               0              18
## ENSG00000000460.16            8.12               0             110
##                     P2770_N714.S515 P2770_N714.S516 P2770_N714.S517
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             20              25               0
## ENSG00000000457.13             22               4               0
## ENSG00000000460.16              0               0              84
##                     P2770_N714.S518 P2770_N714.S520 P2770_N714.S521
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12            161              13              78
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16             95              70               0
##                     P2770_N714.S522 P2770_N715.S513 P2770_N715.S515
## ENSG00000000003.14            0.00               0               0
## ENSG00000000005.5             0.00               0               0
## ENSG00000000419.12           21.00              92              50
## ENSG00000000457.13           22.92              40              20
## ENSG00000000460.16          127.08               0               0
##                     P2770_N715.S516 P2770_N715.S517 P2770_N715.S518
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             61              43               3
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              0             118              92
##                     P2770_N715.S520 P2770_N715.S521 P2770_N715.S522
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             32              75              46
## ENSG00000000457.13              0               0               0
```

```
## ENSG00000000460.16               5               0               0
```

```
head(P5, 5)
```

```
##                    P2771_N701.S502 P2771_N701.S503 P2771_N701.S505
## ENSG00000000003.14               0               0               0
## ENSG00000000005.5                0               0               0
## ENSG00000000419.12              41             127             113
## ENSG00000000457.13              60               2              16
## ENSG00000000460.16               0              59              48
##                    P2771_N701.S506 P2771_N701.S507 P2771_N701.S508
## ENSG00000000003.14               0               0               0
## ENSG00000000005.5                0               0               0
## ENSG00000000419.12              60               0             197
## ENSG00000000457.13               0               0               0
## ENSG00000000460.16               0               0               0
##                    P2771_N701.S510 P2771_N701.S511 P2771_N702.S502
## ENSG00000000003.14               0               0               0
## ENSG00000000005.5                0               0               0
## ENSG00000000419.12              14              35              84
## ENSG00000000457.13              55               0               9
## ENSG00000000460.16               0               0              65
##                    P2771_N702.S503 P2771_N702.S505 P2771_N702.S506
## ENSG00000000003.14            0.00            0.00               0
## ENSG00000000005.5             0.00            0.00               0
## ENSG00000000419.12          134.00           12.00              13
## ENSG00000000457.13           34.04           46.47               0
## ENSG00000000460.16           66.96           46.53               0
##                    P2771_N702.S507 P2771_N702.S508 P2771_N702.S510
## ENSG00000000003.14               0               0               0
## ENSG00000000005.5                0               0               0
## ENSG00000000419.12               0             151              57
## ENSG00000000457.13               0               0               0
## ENSG00000000460.16               0               0             102
##                    P2771_N702.S511 P2771_N703.S502 P2771_N703.S503
## ENSG00000000003.14               0               0            0.00
## ENSG00000000005.5                0               0            0.00
## ENSG00000000419.12               0              71           35.00
## ENSG00000000457.13               0               0           31.61
## ENSG00000000460.16               0               0           51.39
##                    P2771_N703.S505 P2771_N703.S506 P2771_N703.S507
## ENSG00000000003.14               0               0               0
## ENSG00000000005.5                0               0               0
## ENSG00000000419.12              24              11             129
## ENSG00000000457.13               0               0               0
## ENSG00000000460.16               6              24               0
##                    P2771_N703.S508 P2771_N703.S510 P2771_N703.S511
## ENSG00000000003.14               0               0               0
## ENSG00000000005.5                0               0               0
## ENSG00000000419.12              61               0              24
## ENSG00000000457.13               0               0               0
## ENSG00000000460.16              55               0               0
##                    P2771_N704.S502 P2771_N704.S503 P2771_N704.S505
## ENSG00000000003.14               0            0.00               0
```

```
## ENSG00000000005.5               0            0.00               0
## ENSG00000000419.12             74          151.00              51
## ENSG00000000457.13              0           26.41               0
## ENSG00000000460.16              0           33.59               0
##                    P2771_N704.S506 P2771_N704.S507 P2771_N704.S508
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12              0              50              60
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              0              46               0
##                    P2771_N704.S510 P2771_N704.S511 P2771_N705.S502
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             21              13              41
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              0              40              89
##                    P2771_N705.S503 P2771_N705.S505 P2771_N705.S506
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12            127              31              22
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16             51               0             169
##                    P2771_N705.S507 P2771_N705.S508 P2771_N705.S510
## ENSG00000000003.14              0            0.00            0.00
## ENSG00000000005.5               0            0.00            0.00
## ENSG00000000419.12             96            3.00           53.00
## ENSG00000000457.13            111          133.87          112.67
## ENSG00000000460.16              0            4.13          195.33
##                    P2771_N705.S511 P2771_N706.S502 P2771_N706.S503
## ENSG00000000003.14              0               0            0.00
## ENSG00000000005.5               0               0            0.00
## ENSG00000000419.12             51              25           37.00
## ENSG00000000457.13              0               0           21.93
## ENSG00000000460.16             90               0           20.07
##                    P2771_N706.S505 P2771_N706.S506 P2771_N706.S507
## ENSG00000000003.14            0.00               0               0
## ENSG00000000005.5             0.00               0               0
## ENSG00000000419.12           67.00              34               7
## ENSG00000000457.13           26.23               0               0
## ENSG00000000460.16           97.77               0               0
##                    P2771_N706.S508 P2771_N706.S510 P2771_N706.S511
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12              0              13              19
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              0              11               0
##                    P2771_N707.S502 P2771_N707.S503 P2771_N707.S505
## ENSG00000000003.14            0.00               0               0
## ENSG00000000005.5             0.00               0               0
## ENSG00000000419.12          114.00              69              82
## ENSG00000000457.13           57.74              42               0
## ENSG00000000460.16           46.26               0              31
##                    P2771_N707.S506 P2771_N707.S507 P2771_N707.S508
## ENSG00000000003.14              0               0               0
```

```
## ENSG00000000005.5                0                0                0
## ENSG00000000419.12              28              177               52
## ENSG00000000457.13               0                0                0
## ENSG00000000460.16               0                4               93
##                    P2771_N707.S510 P2771_N707.S511 P2771_N710.S502
## ENSG00000000003.14               0                0             0.00
## ENSG00000000005.5                0                0             0.00
## ENSG00000000419.12              90               98           122.00
## ENSG00000000457.13               0                0            15.76
## ENSG00000000460.16              22               25            44.24
##                    P2771_N710.S503 P2771_N710.S505 P2771_N710.S506
## ENSG00000000003.14               0                0                0
## ENSG00000000005.5                0                0                0
## ENSG00000000419.12              10                0                0
## ENSG00000000457.13               0               18                0
## ENSG00000000460.16               0               64                0
##                    P2771_N710.S507 P2771_N710.S508 P2771_N710.S510
## ENSG00000000003.14               0                0                0
## ENSG00000000005.5                0                0                0
## ENSG00000000419.12              76               88              113
## ENSG00000000457.13               0                0                0
## ENSG00000000460.16               1                0                9
##                    P2771_N710.S511 P2771_N711.S502 P2771_N711.S503
## ENSG00000000003.14               0                0                0
## ENSG00000000005.5                0                0                0
## ENSG00000000419.12              74               98               61
## ENSG00000000457.13               0               12               55
## ENSG00000000460.16              33               10                0
##                    P2771_N711.S505 P2771_N711.S506 P2771_N711.S507
## ENSG00000000003.14               0                0                0
## ENSG00000000005.5                0                0                0
## ENSG00000000419.12              73              101              165
## ENSG00000000457.13               0                0                0
## ENSG00000000460.16             106               56               40
##                    P2771_N711.S508 P2771_N711.S510 P2771_N711.S511
## ENSG00000000003.14               0                0                0
## ENSG00000000005.5                0                0                0
## ENSG00000000419.12               0               48               27
## ENSG00000000457.13               0               59               10
## ENSG00000000460.16               0                0               65
##                    P2771_N712.S502 P2771_N712.S503 P2771_N712.S505
## ENSG00000000003.14               0                0                0
## ENSG00000000005.5                0                0                0
## ENSG00000000419.12              44               42                0
## ENSG00000000457.13              24                0                0
## ENSG00000000460.16              68                0                0
##                    P2771_N712.S506 P2771_N712.S507 P2771_N712.S508
## ENSG00000000003.14               0                0                0
## ENSG00000000005.5                0                0                0
## ENSG00000000419.12              98              136               85
## ENSG00000000457.13               0                0                0
## ENSG00000000460.16               0              150               54
##                    P2771_N712.S510 P2771_N712.S511 P2771_N714.S502
## ENSG00000000003.14               0                0                0
```

```
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             23              12             108
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16            135               9              28
##                     P2771_N714.S503 P2771_N714.S505 P2771_N714.S506
## ENSG00000000003.14            0.00               0               0
## ENSG00000000005.5             0.00               0               0
## ENSG00000000419.12           60.00              93              75
## ENSG00000000457.13            8.12               0               0
## ENSG00000000460.16           94.88               0               0
##                     P2771_N714.S507 P2771_N714.S508 P2771_N714.S510
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             85              39               0
## ENSG00000000457.13              0               0              43
## ENSG00000000460.16             37               0               0
##                     P2771_N714.S511 P2771_N715.S502 P2771_N715.S503
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             77               0              66
## ENSG00000000457.13              0               3               0
## ENSG00000000460.16              5              41               3
##                     P2771_N715.S505 P2771_N715.S506 P2771_N715.S507
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             32             101             235
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16             57              15              51
##                     P2771_N715.S508 P2771_N715.S510 P2771_N715.S511
## ENSG00000000003.14              0               0               0
## ENSG00000000005.5               0               0               0
## ENSG00000000419.12             59              59              27
## ENSG00000000457.13              0               0               0
## ENSG00000000460.16              0               0               0
```

```
#Alternatives
#P3<- read.table(gzfile("P3.tsv.gz", "rt"), header= TRUE)
#P4<- read.table(gzfile("P4.tsv.gz", "rt"), header= TRUE)
#P5<- read.table(gzfile("P5.tsv.gz", "rt"), header= TRUE)
```

Question 2

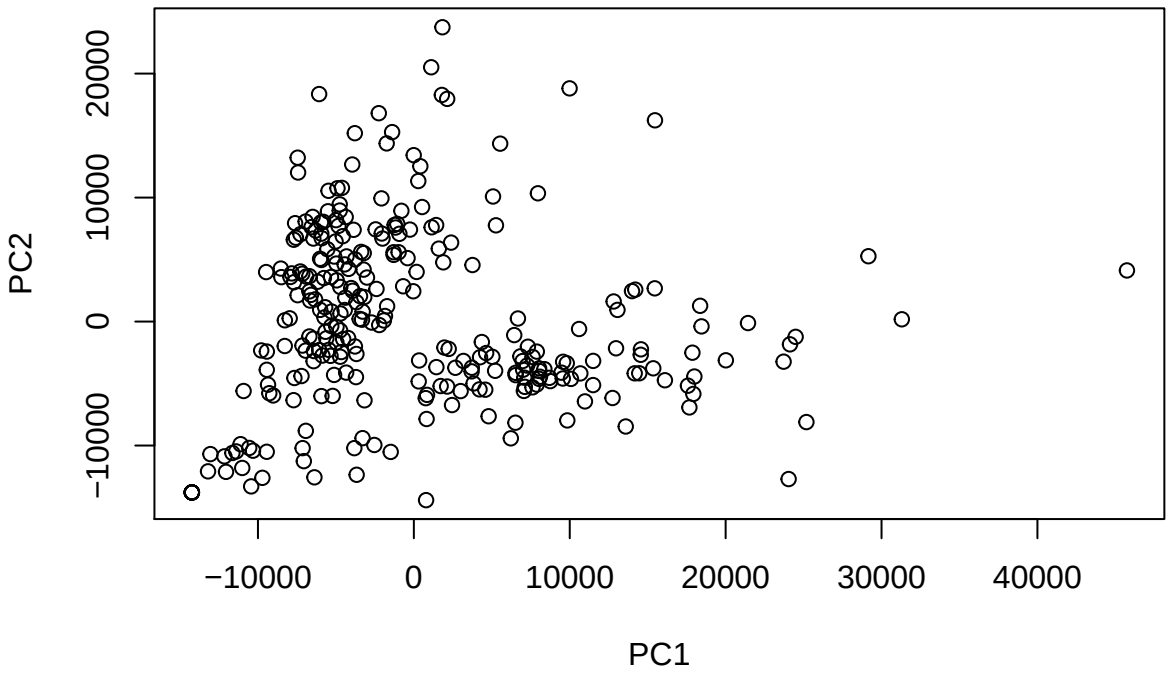
```
treat<- read.delim("samplesheet.tsv", row.names=1)
#treat <- as.factor(readLines("samplesheet.tsv"))
#Alternatives
#treat<- read.table(("samplesheet.tsv", "rt"), header= TRUE))

table(treat)
```

```
## treat
## J-LatA2+DMSO  J-LatA2+PMA J-LatA2+SAHA  Jurkat+DMSO   Jurkat+PMA  Jurkat+SAHA
##           36           81           81           18           36           36
```

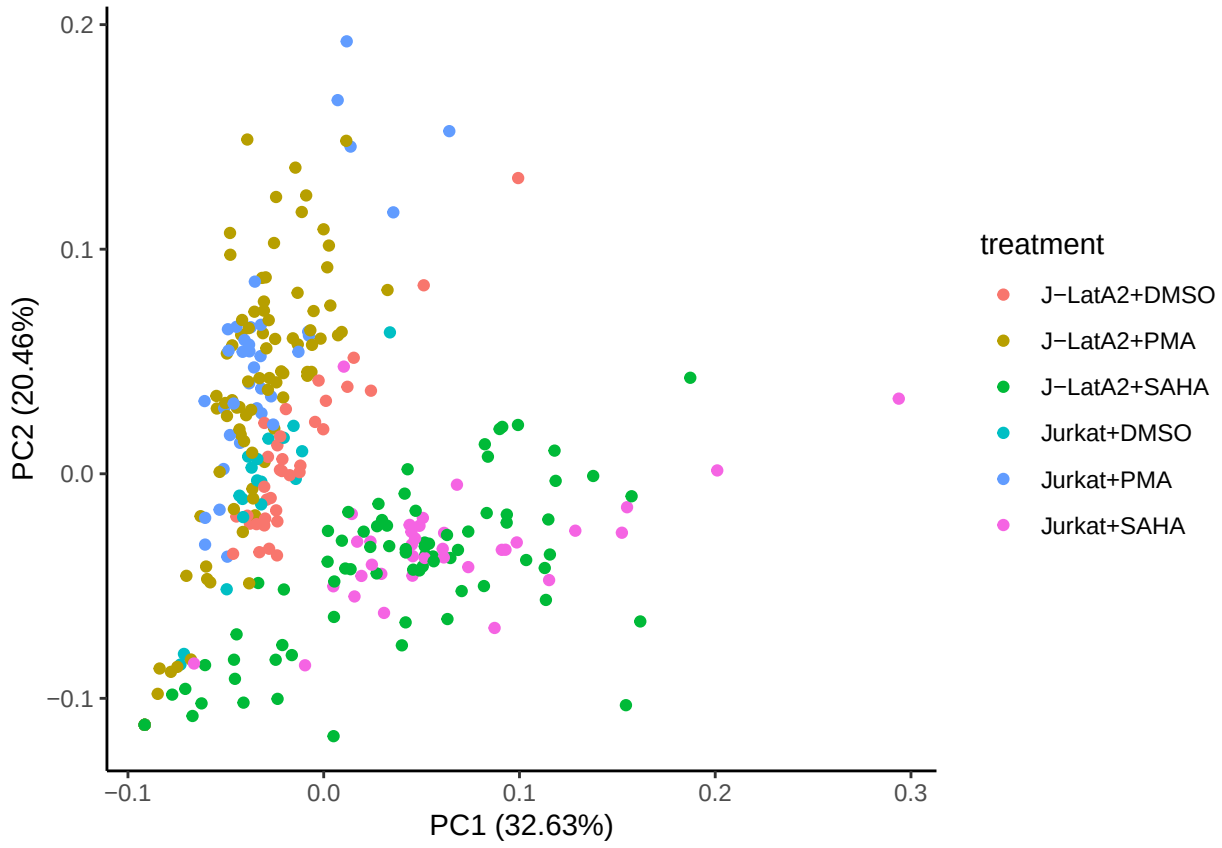
There are 2 kinds of cells (one infected with HIV and one with a latent virus), each one is treated with DMSO, SAHA and PMA. This creates 6 different combinations, the table above shows the number of cells in each condition.

Question 3

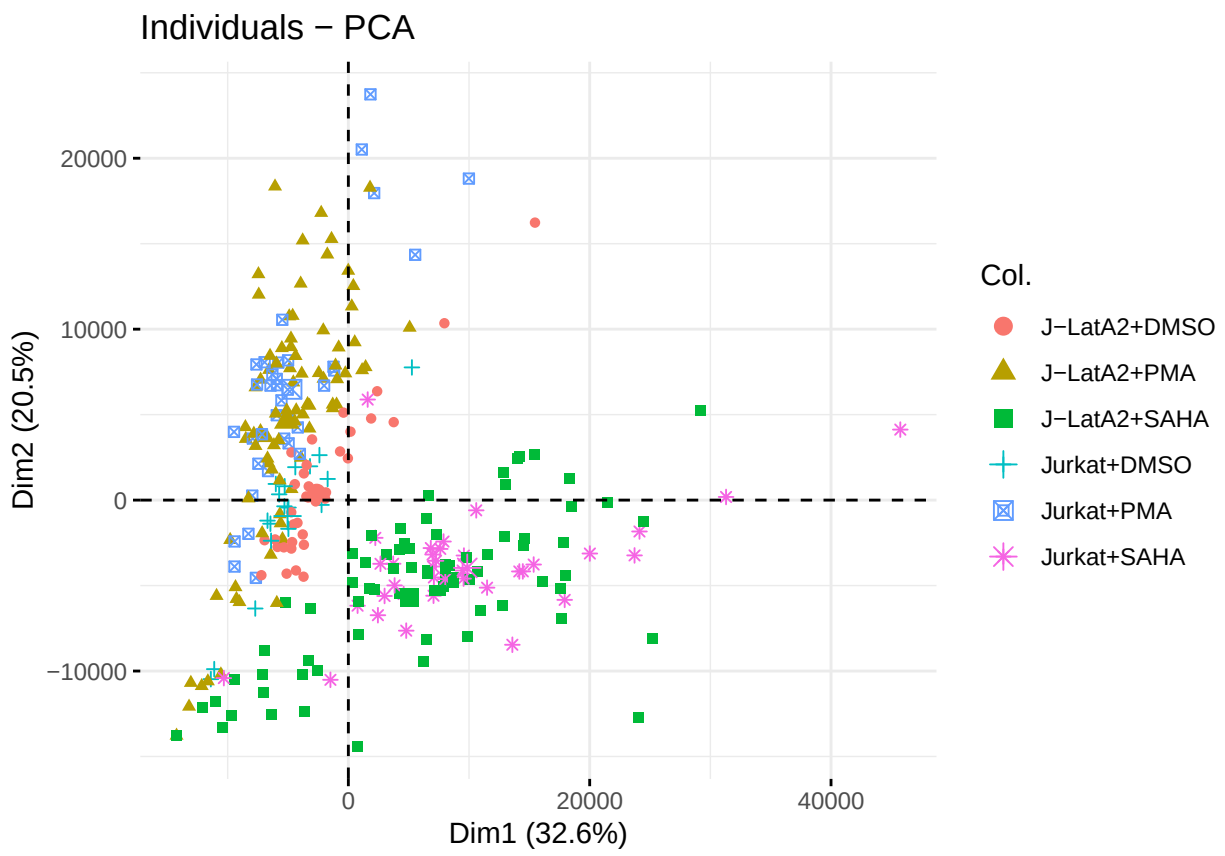
```
X3 <- t(P3)
X4 <- t(P4)
X5 <- t(P5)

X1 = rbind(X3, X4, X5)

data <- data.frame(row.names= rownames(X1), experiment= as.factor(rep(c("A", "B", "C"), each=96)), trea

colnames(data)<- c("experiment", "treatment")

#PCA
PCA = stats::prcomp(X1)
plot(PCA$x)
```



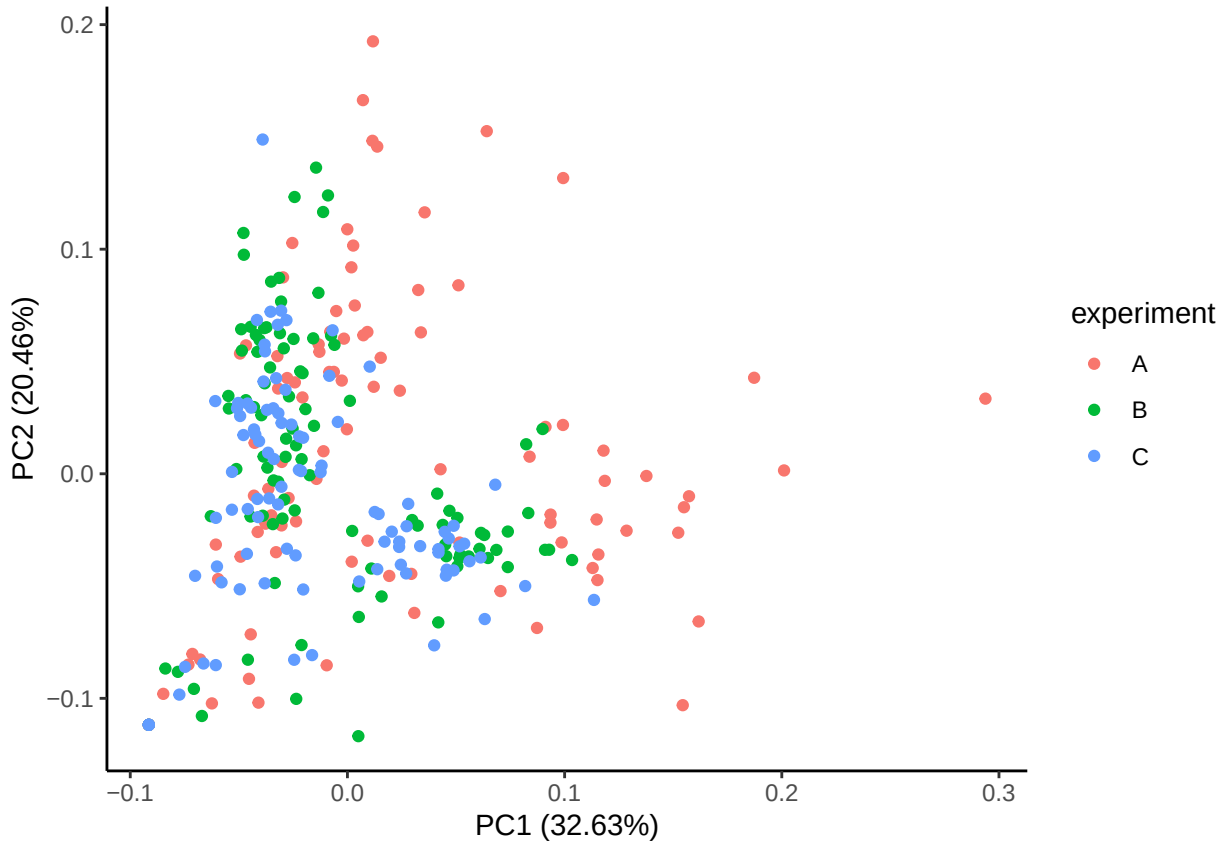

```
#coloring by treatment
#using autoplot
ggplot2::autoplot(PCA, data = data, colour = "treatment") + theme_classic()
```



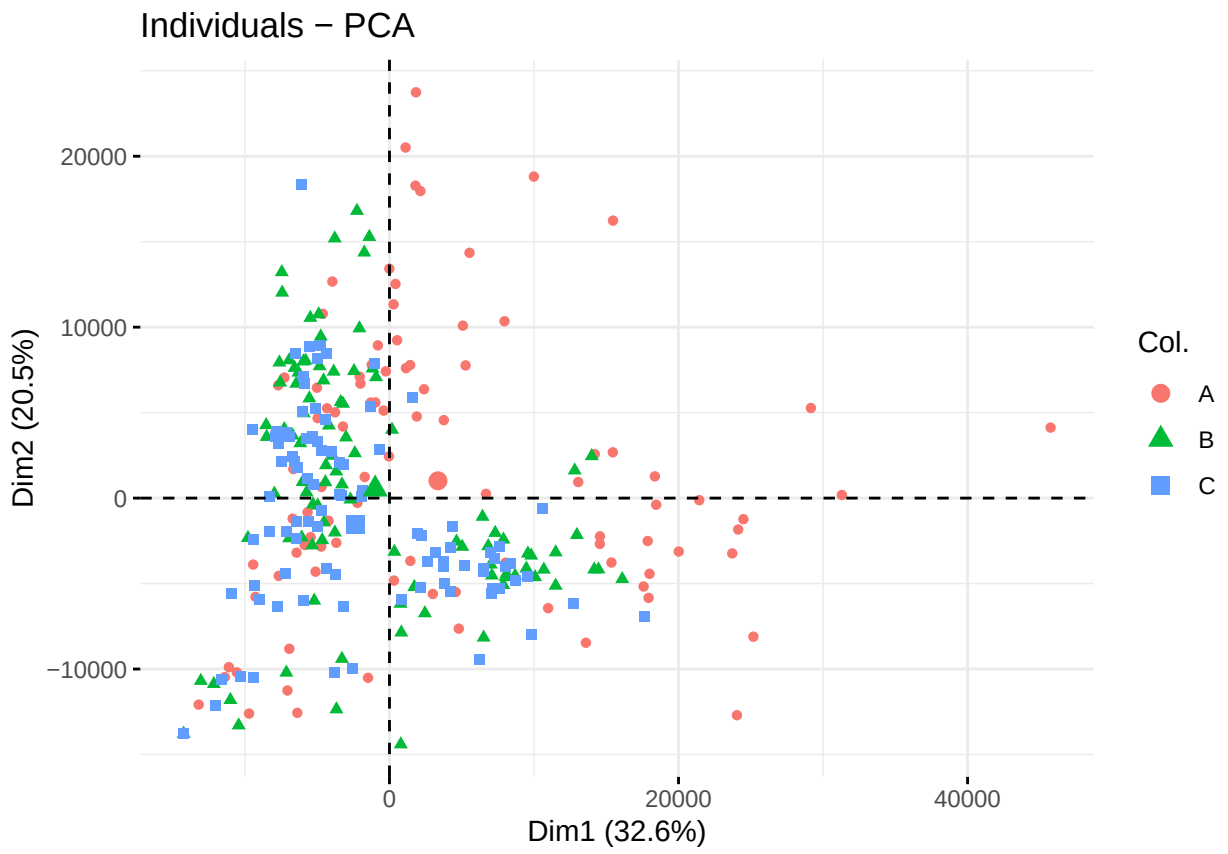

```
#using factoextra
factoextra::fviz_pca_ind(PCA, geom = "point", col.ind = data$treatment)
```




```
#coloring by experiment
#using autoplot
ggplot2::autoplot(PCA, data = data, colour = "experiment") + theme_classic()
```




```
#using factoextra
factoextra::fviz_pca_ind(PCA, geom = "point", col.ind = data$experiment)
```



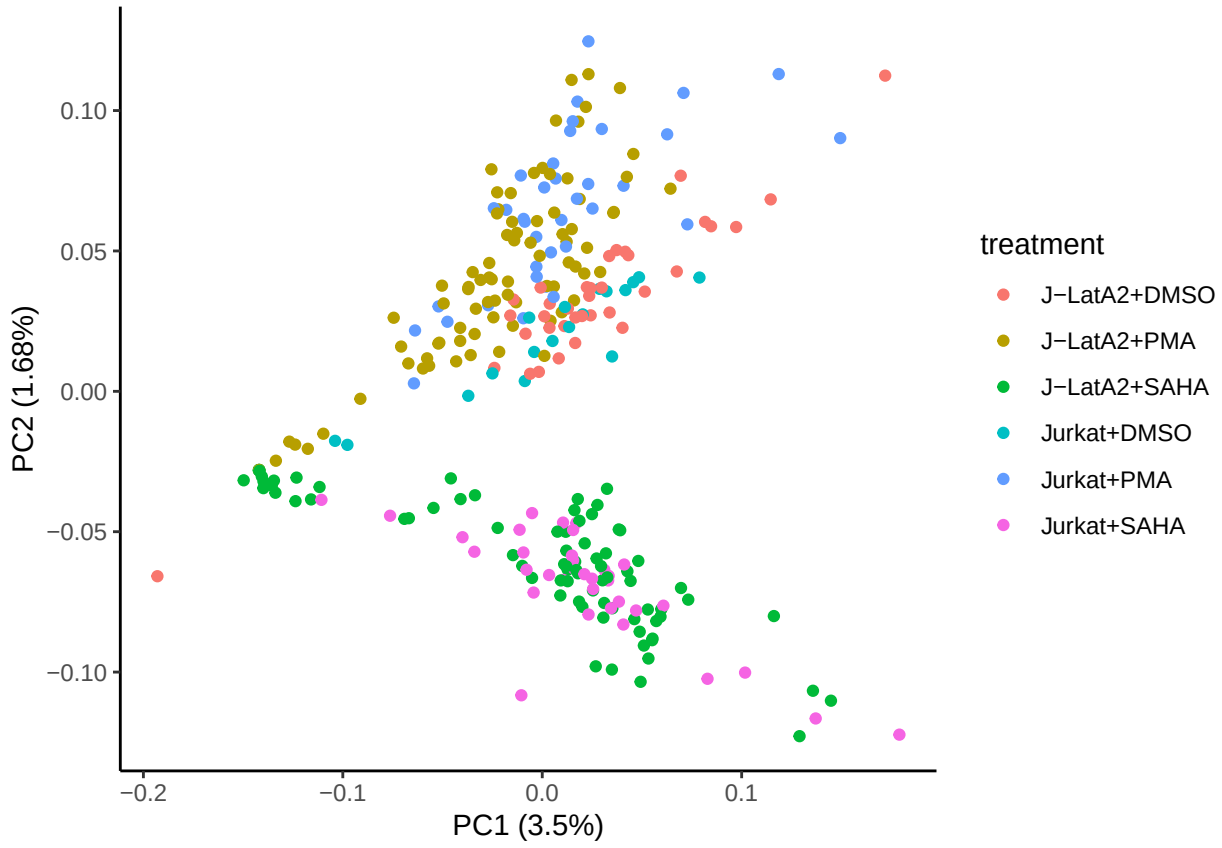

```
#creating scaled PCA
print(dim(X1))
```

```
## [1]   288 58780
```

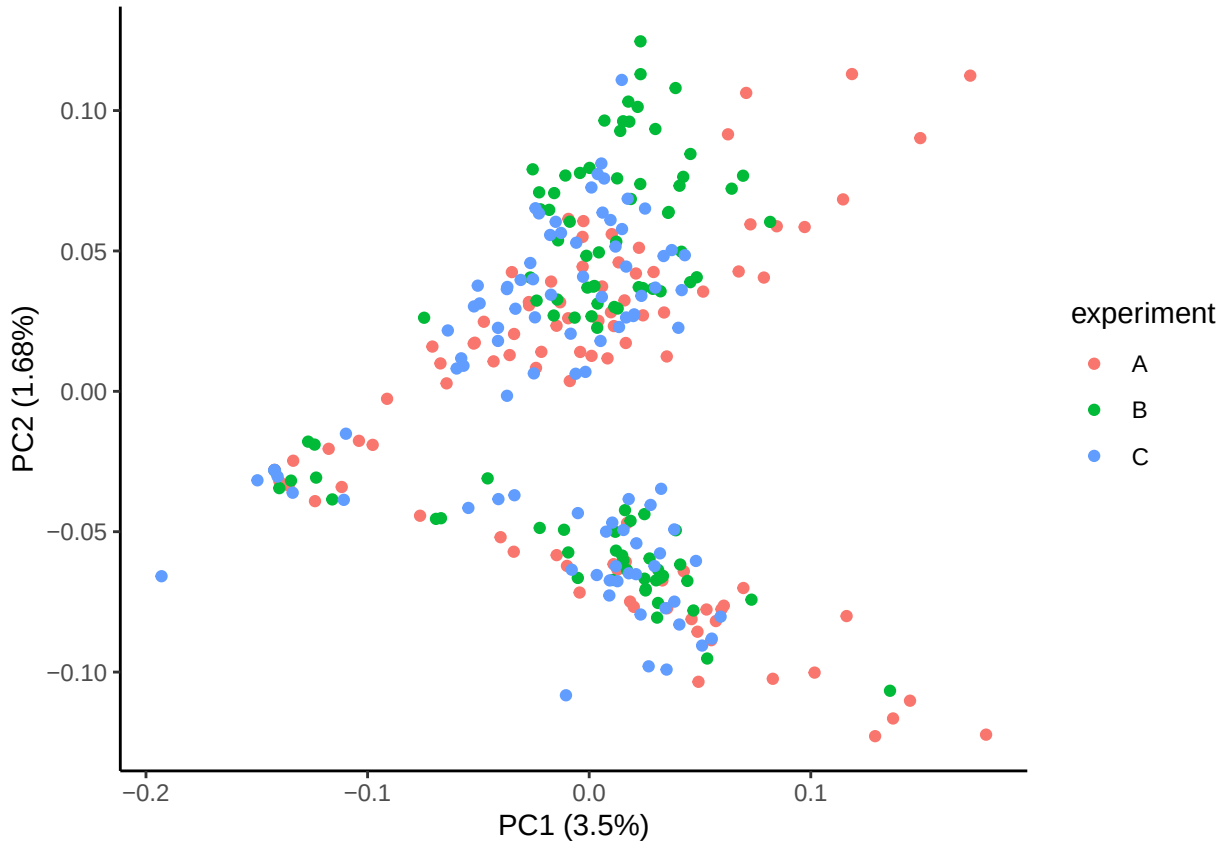
```
X1=X1[, colSums(X1)>0] ## tip: rowSums > 0
print(dim(X1))
```

```
## [1]   288 23104
```

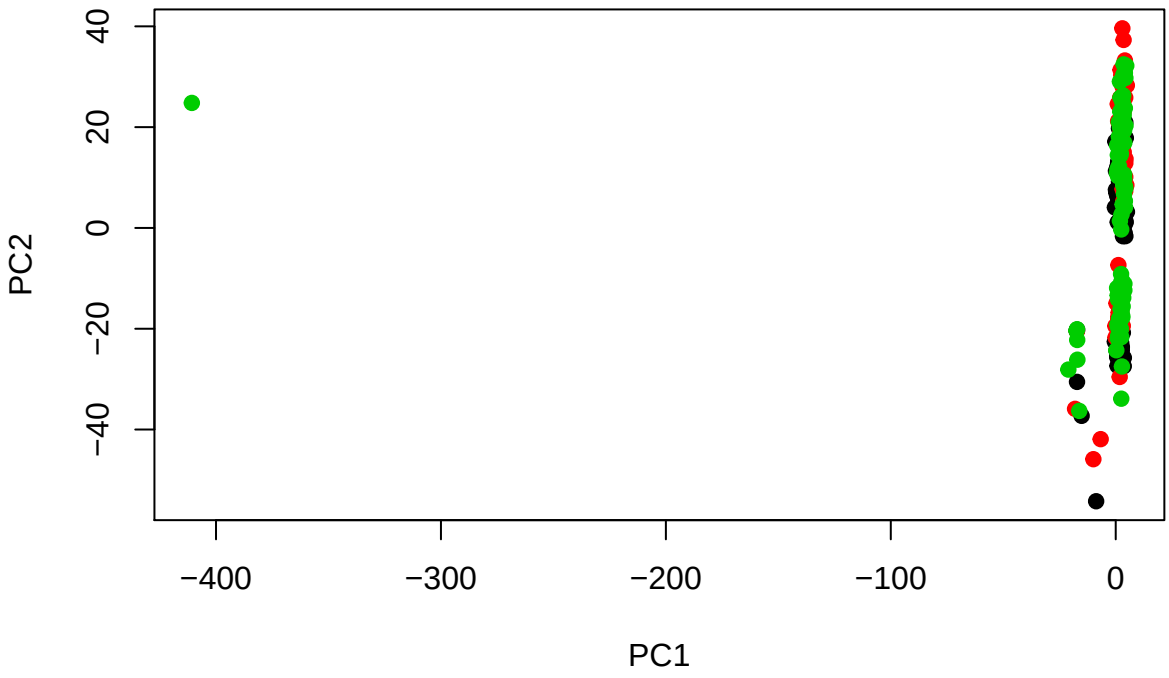
```
pca_res_scaled <- stats::prcomp(X1, scale=TRUE)
#colored by treatment
ggplot2::autoplot(pca_res_scaled, data = data, colour = "treatment") + theme_classic()
```



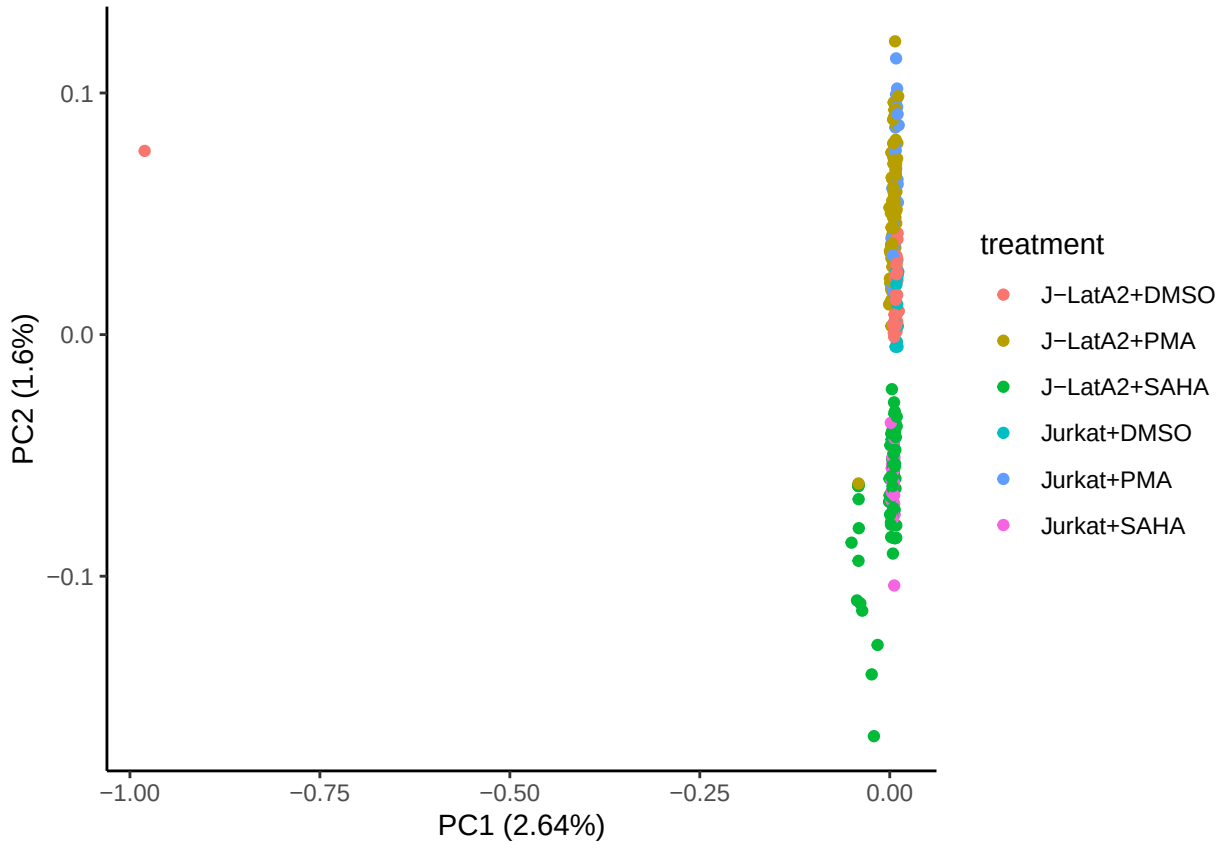

```
#colored by experiment
ggplot2::autoplot(pca_res_scaled, data = data, colour = "experiment") + theme_classic()
```




```
#creating normalized PCA
Y1 = X1/rowSums(X1)
normpca <- stats::prcomp(Y1, scale=TRUE)
#colored by experiment
plot(normpca$x, pch=19, col=data$experiment)
```




```
#colored by treatment, using autoplot
ggplot2::autoplot(normpca, data=data, colour="treatment") + theme_classic()
```



DMSO is the control and it's very similar to PMA, meaning that PMA doesn't have a significant effect. On the other hand, SAHA is very different from both of them, meaning that it has a significant effect. There is no batch effect, because in the graph colored by experiment the samples are not clustered based on the experiment that they're part of. However, we scaled the PCA and the scaled PCAs are more precise

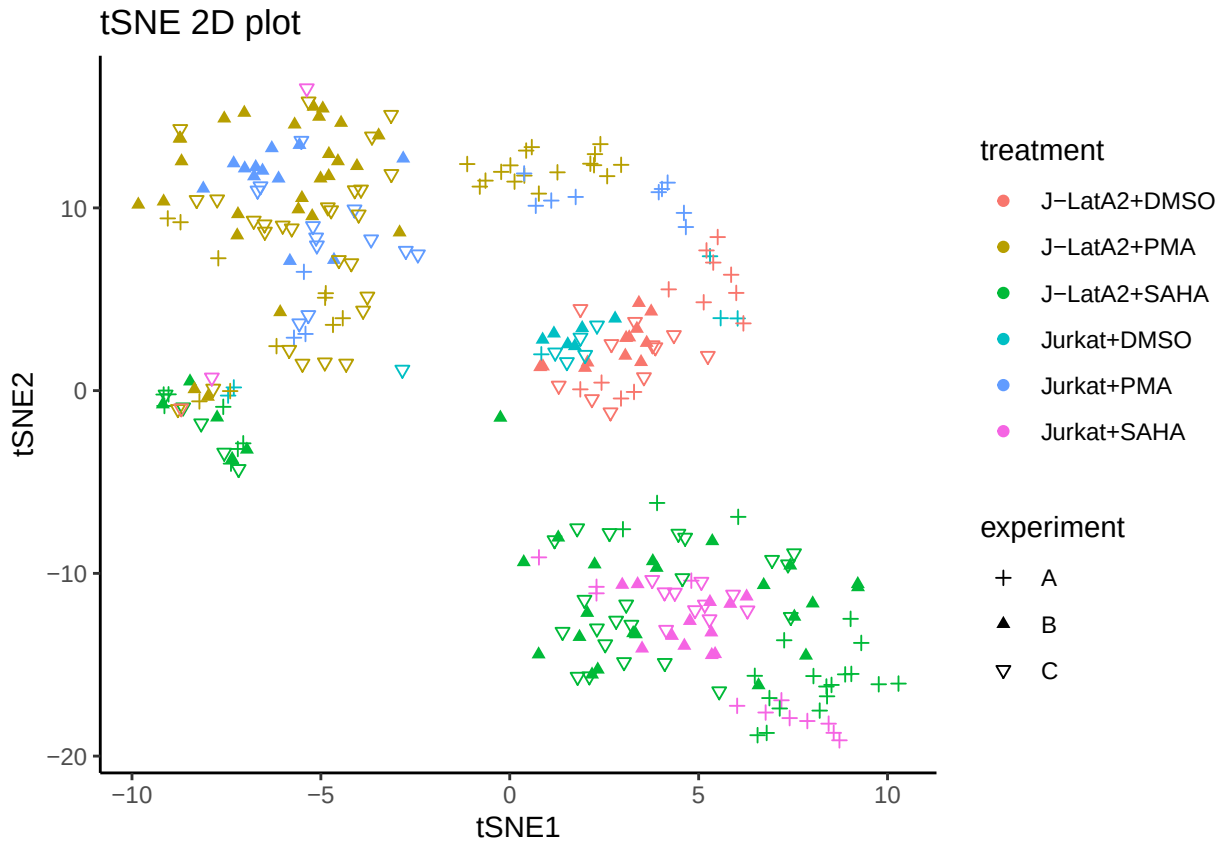
```
#tSNE
tsne1 <- Rtsne(X1)

## Check results generated
names(tsne1)
```

```
##  [1] "N"                 "Y"                 "costs"
##  [4] "itercosts"         "origD"             "perplexity"
##  [7] "theta"             "max_iter"          "stop_lying_iter"
## [10] "mom_switch_iter"   "momentum"          "final_momentum"
## [13] "eta"               "exaggeration_factor"
```

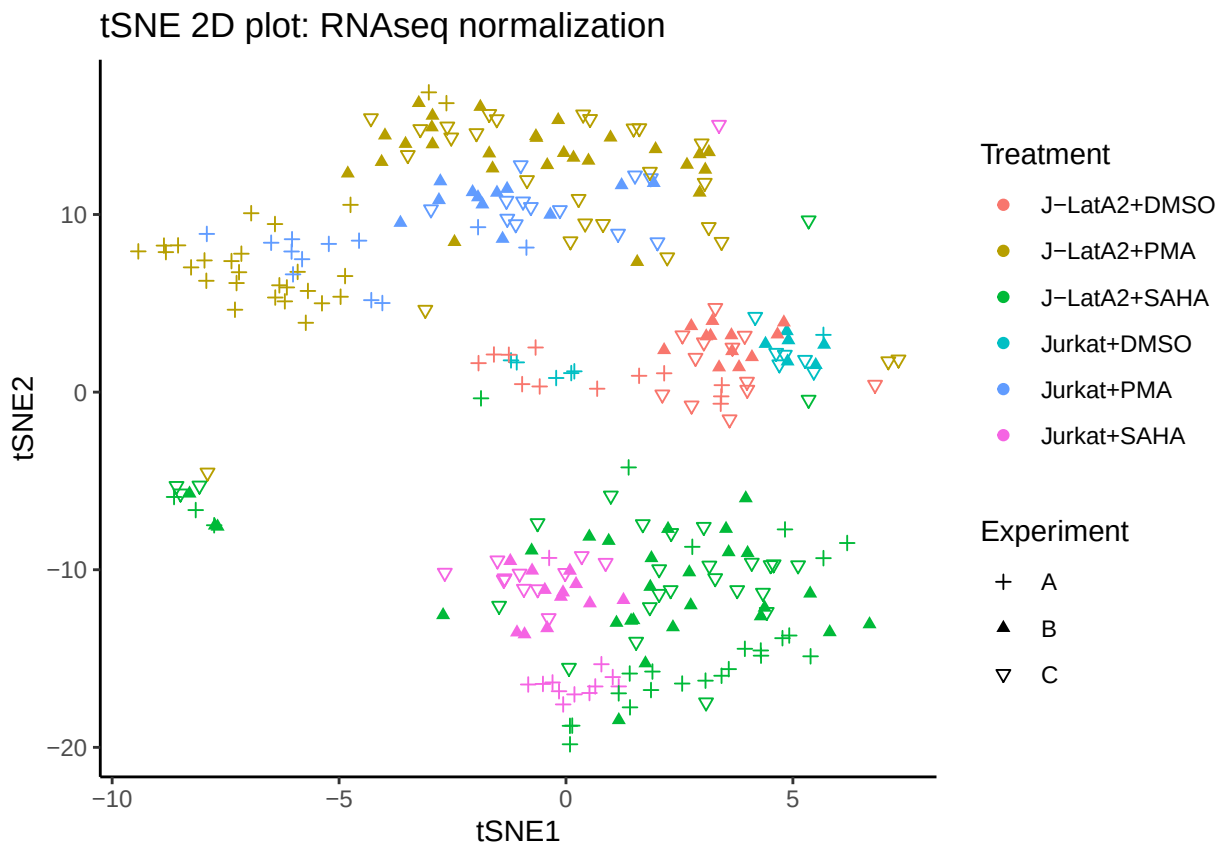
```
head(tsne1$Y)
```

```
##           [,1]         [,2]
## [1,] -7.304639   0.17519535
## [2,]  2.295887 -11.09239025
## [3,] -5.711962   2.89508730
## [4,]  1.867787   0.07082732
## [5,] -7.055208  -2.88328641
## [6,] -7.583680  -0.88662495
```

```
tsne_data <- as.data.frame(tsne1$Y)
tsne_data$experiment <- data$experiment
tsne_data$treatment <- data$treatment

## ggplot
ggplot(tsne_data, aes(x=V1, y=V2)) +
  geom_point(aes(color=treatment, shape=experiment)) +
  scale_shape_manual(values = c(3,17, 6)) +
  ggtitle("tSNE 2D plot") + xlab("tSNE1") + ylab("tSNE2") +
  theme_classic()
```




```
## Normalized tSNE
X1=X1[, colSums(X1)>0]
Y = X1/rowSums(X1)
tSNE_norm <- Rtsne(Y, pca=TRUE)
tsne_data2 <- as.data.frame(tSNE_norm$Y)
tsne_data2$Experiment <- data$experiment
tsne_data2$Treatment <- data$treatment
## ggplot
tSNE2 <- ggplot(tsne_data2, aes(x=V1, y=V2)) +
  geom_point(aes(color=Treatment, shape=Experiment)) +
  scale_shape_manual(values = c(3,17,6)) +
  ggtitle("tSNE 2D plot: RNAseq normalization") + xlab("tSNE1") + ylab("tSNE2") + theme_classic()
print(tSNE2)
```

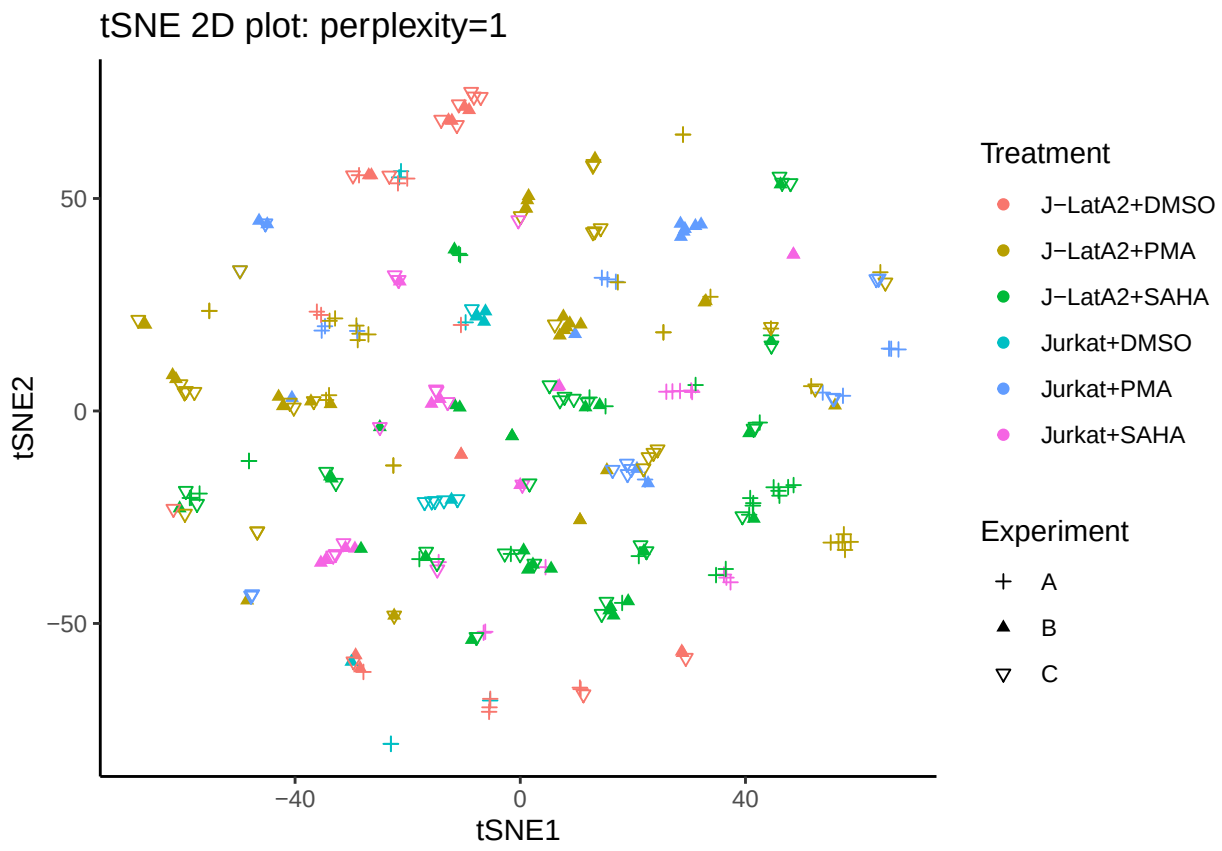


Question 4

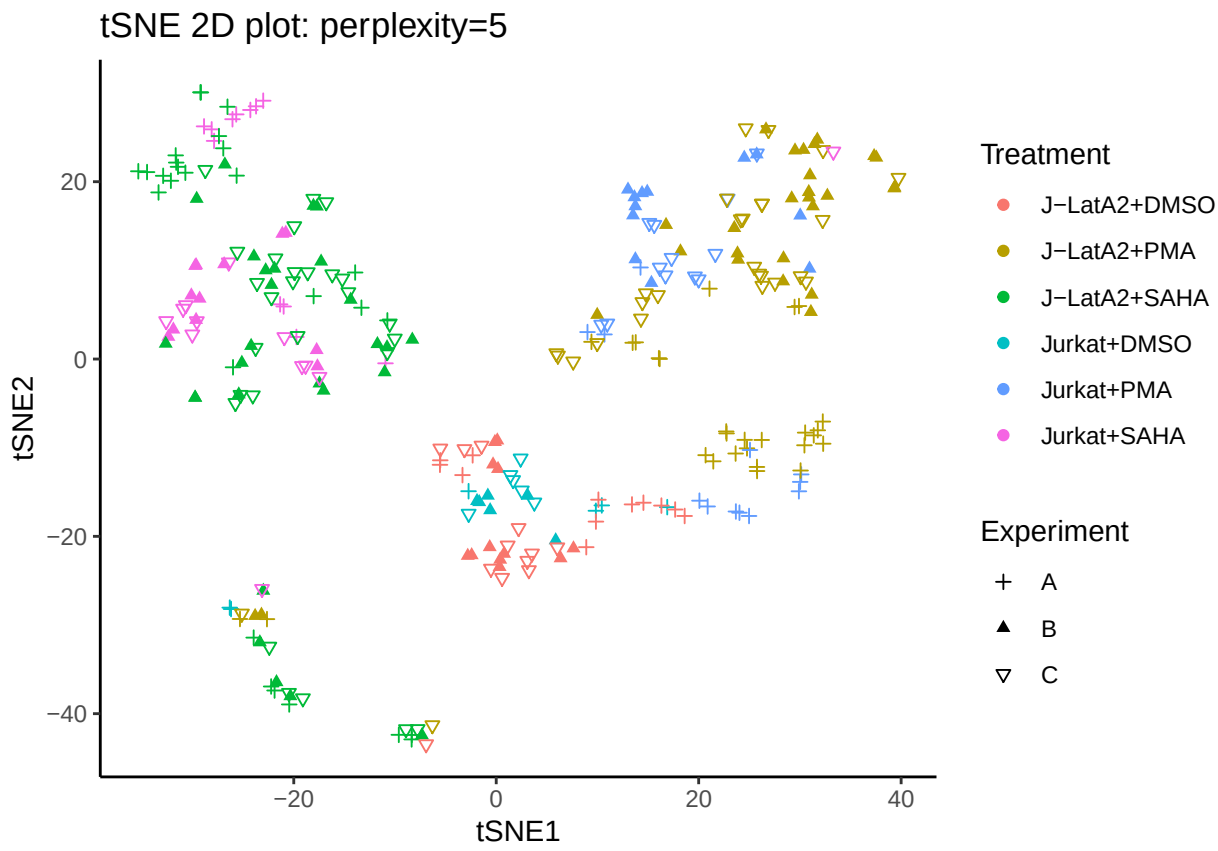
```
#changing perplexity
set.seed(123) ## allows us to reproduce results

list_plots <- list()
for (p_value in c(1, 5, 10, 25, 50)) {
  print(paste0("tSNE perplexity=", p_value))

  ## Default parameters for Rtsne
  tsne_p <- Rtsne(X1, perplexity = p_value)

  tsne_data2 <- as.data.frame(tsne_p$Y)
  tsne_data2$Treatment <- data$treatment
  tsne_data2$Experiment <- data$experiment

  ## ggplot
  plot_p <- ggplot(tsne_data2, aes(x=V1, y=V2)) +
    geom_point(aes(color=Treatment, shape=Experiment)) +
    scale_shape_manual(values = c(3,17,6)) +
    ggtitle(paste0("tSNE 2D plot: perplexity=", p_value)) + xlab("tSNE1") + ylab("tSNE2") +
    theme_classic()

  print(plot_p)

  # alternative: save plots in list_plots
```

```
  list_plots[[paste0("perp_", p_value)]] = plot_p
}
```

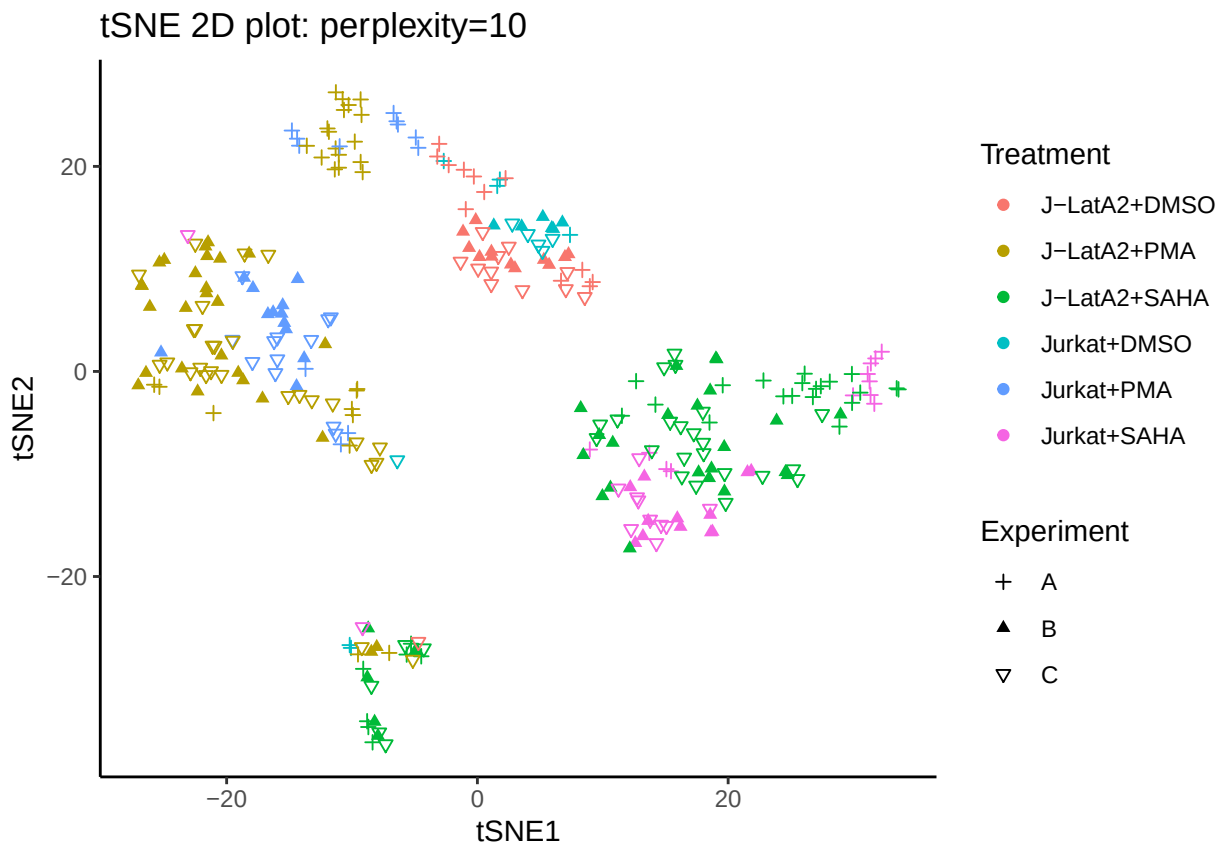
```
## [1] "tSNE perplexity=1"
```



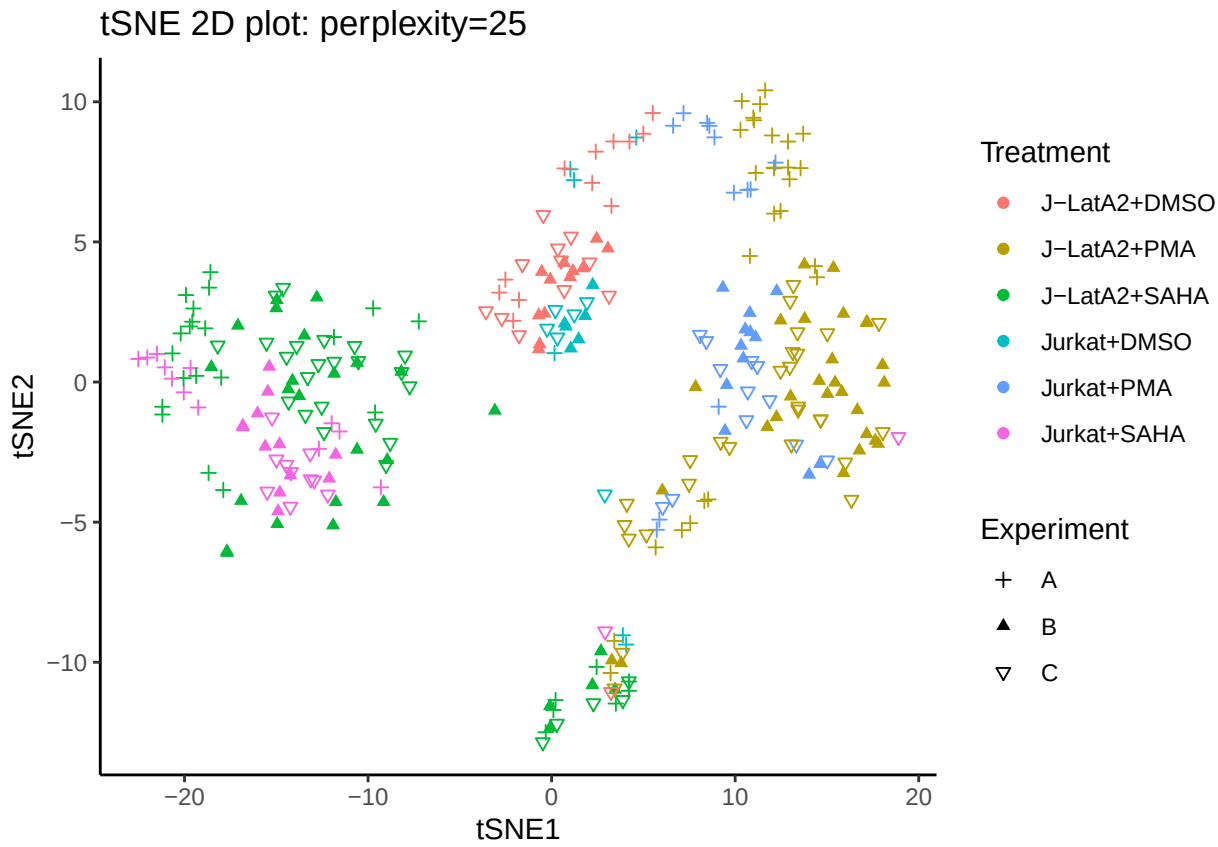

```
## [1] "tSNE perplexity=5"
```



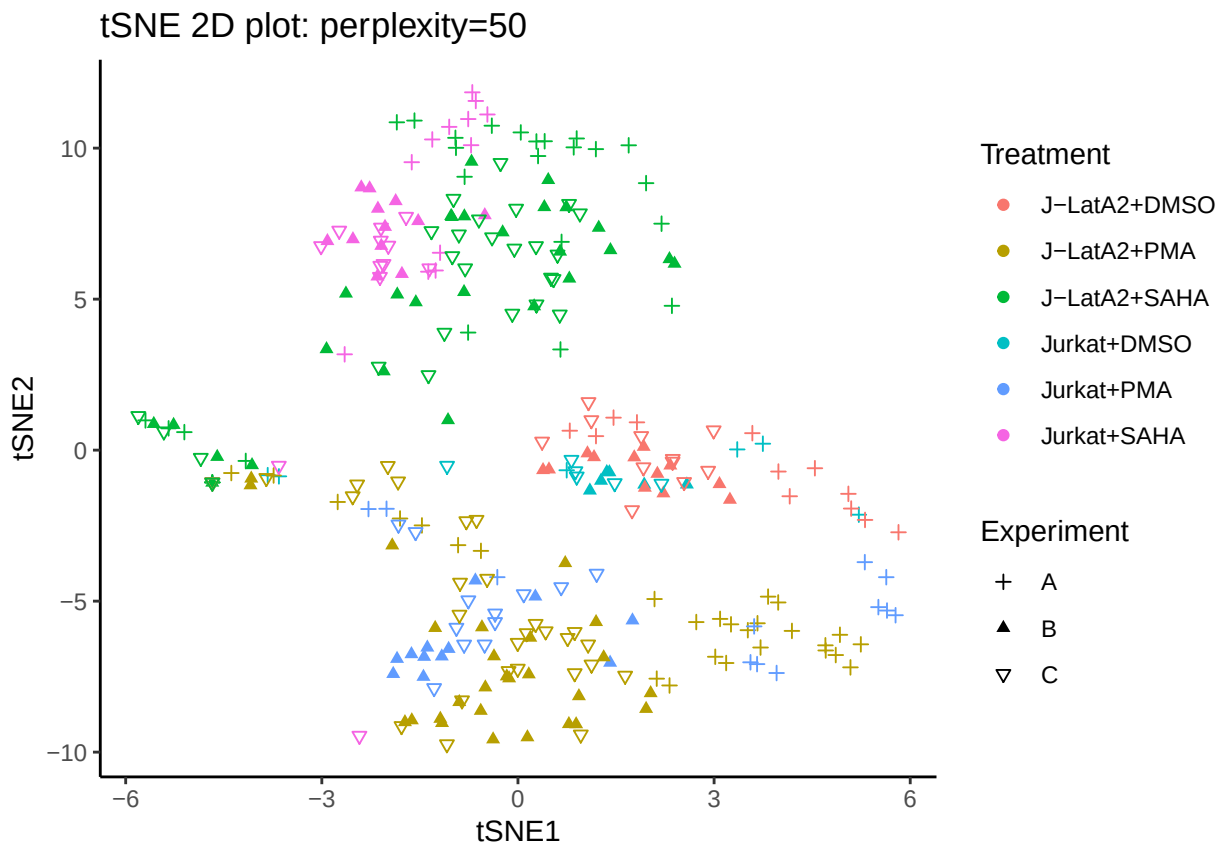

```
## [1] "tSNE perplexity=10"
```



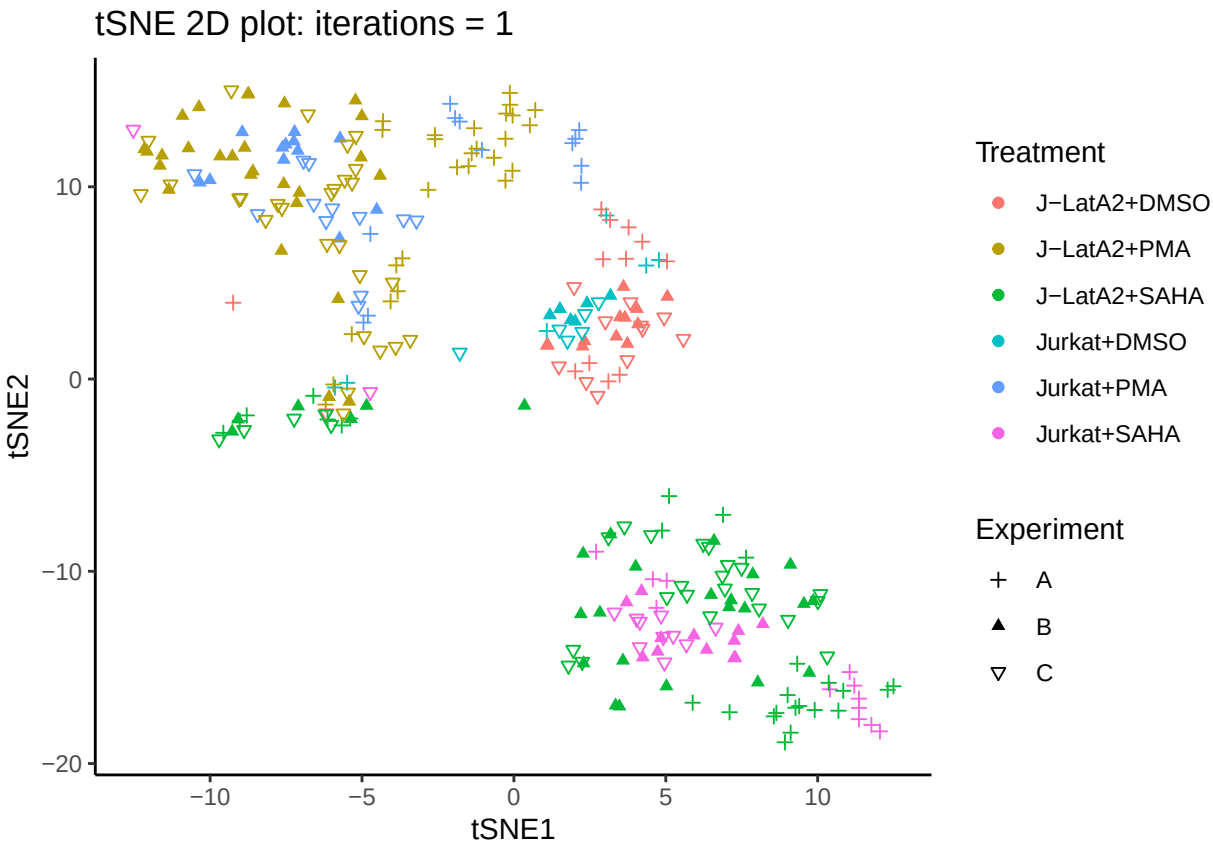

```
## [1] "tSNE perplexity=25"
```

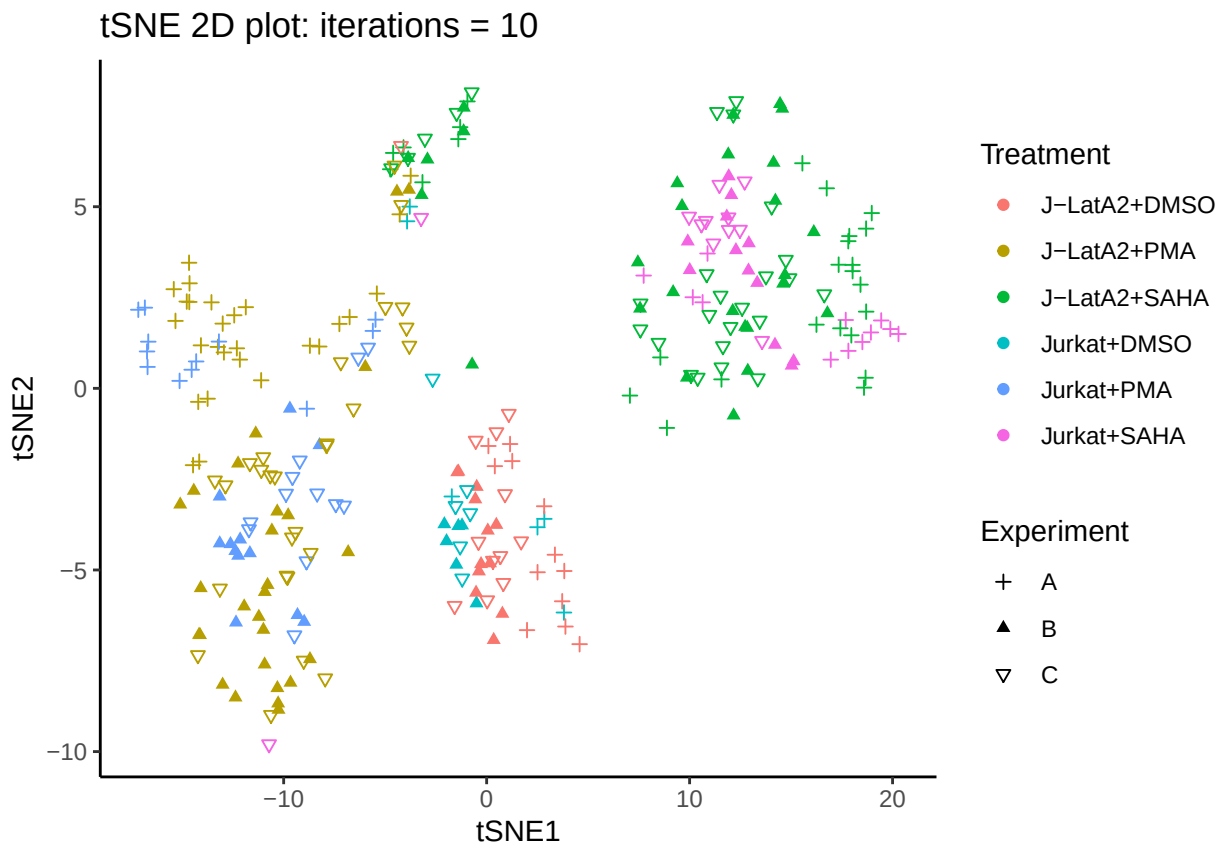


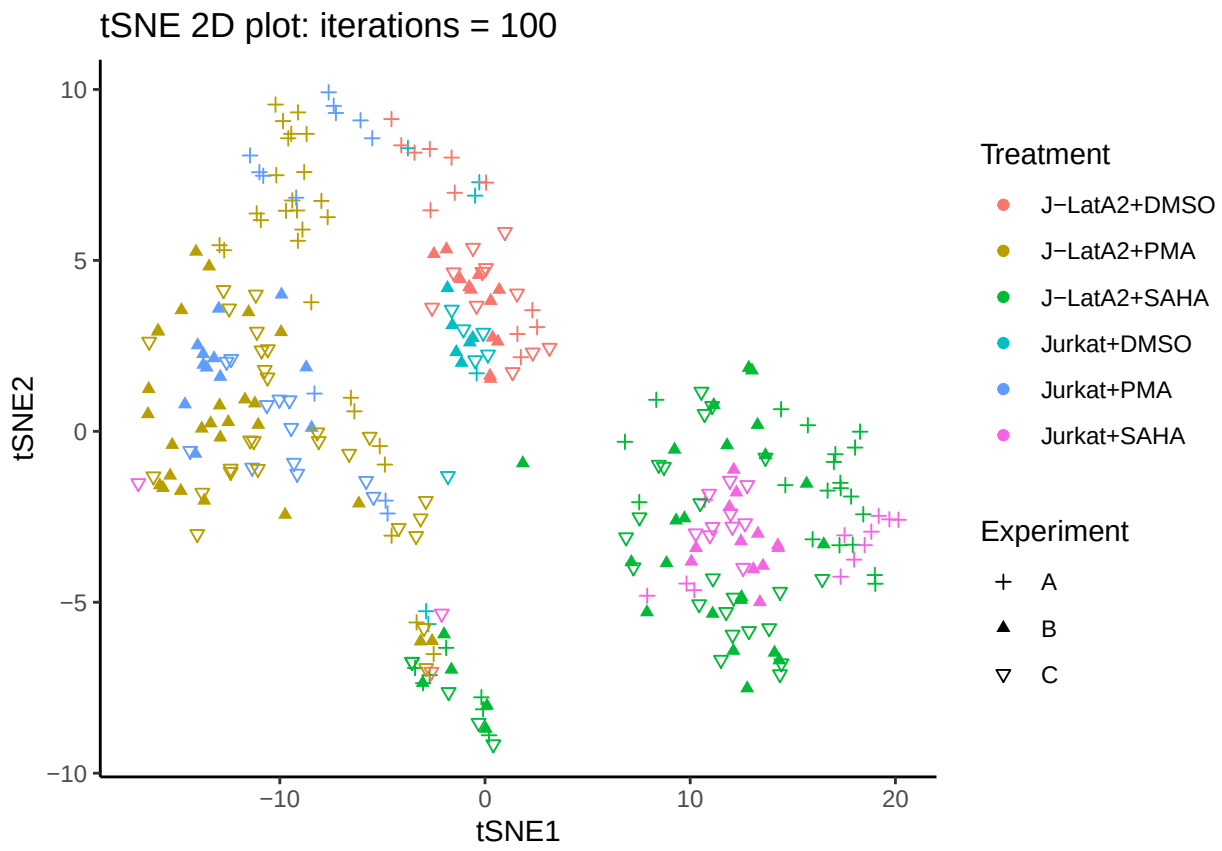

```
## [1] "tSNE perplexity=50"
```

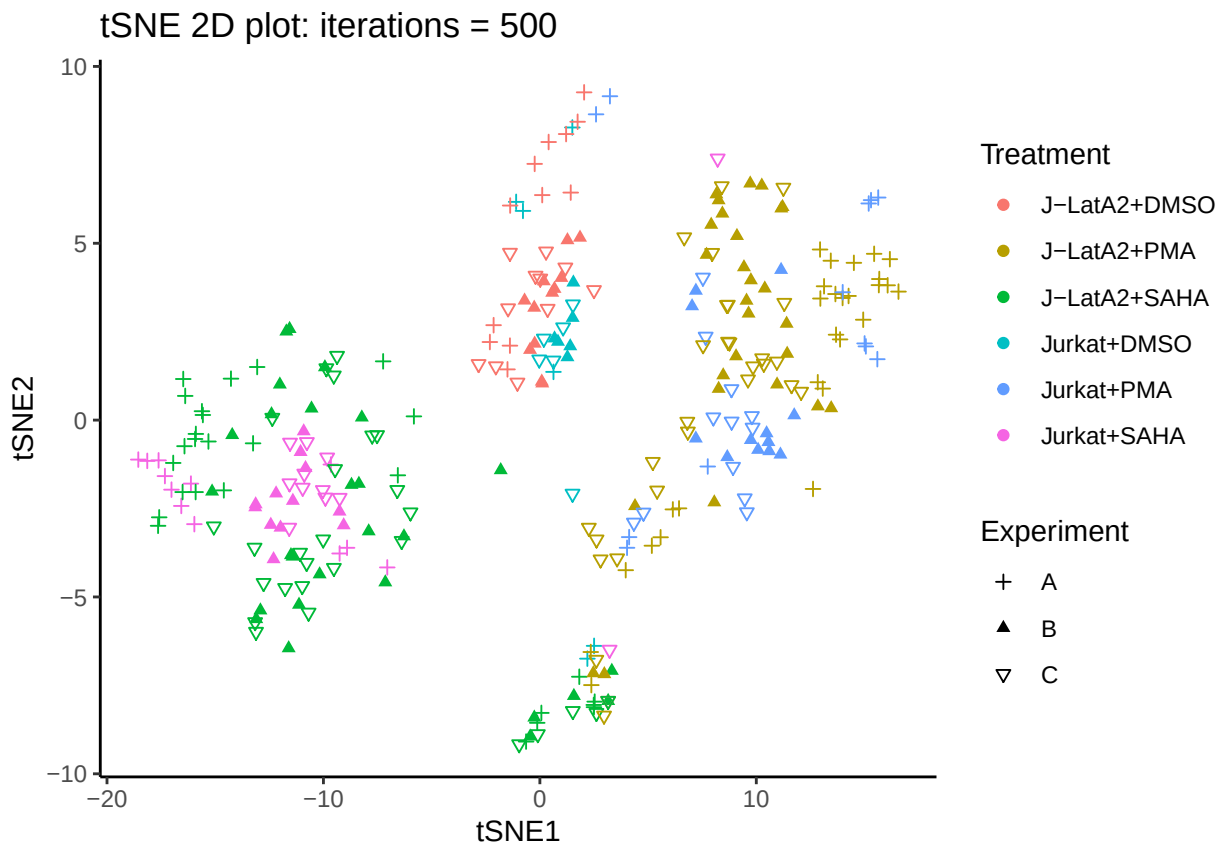


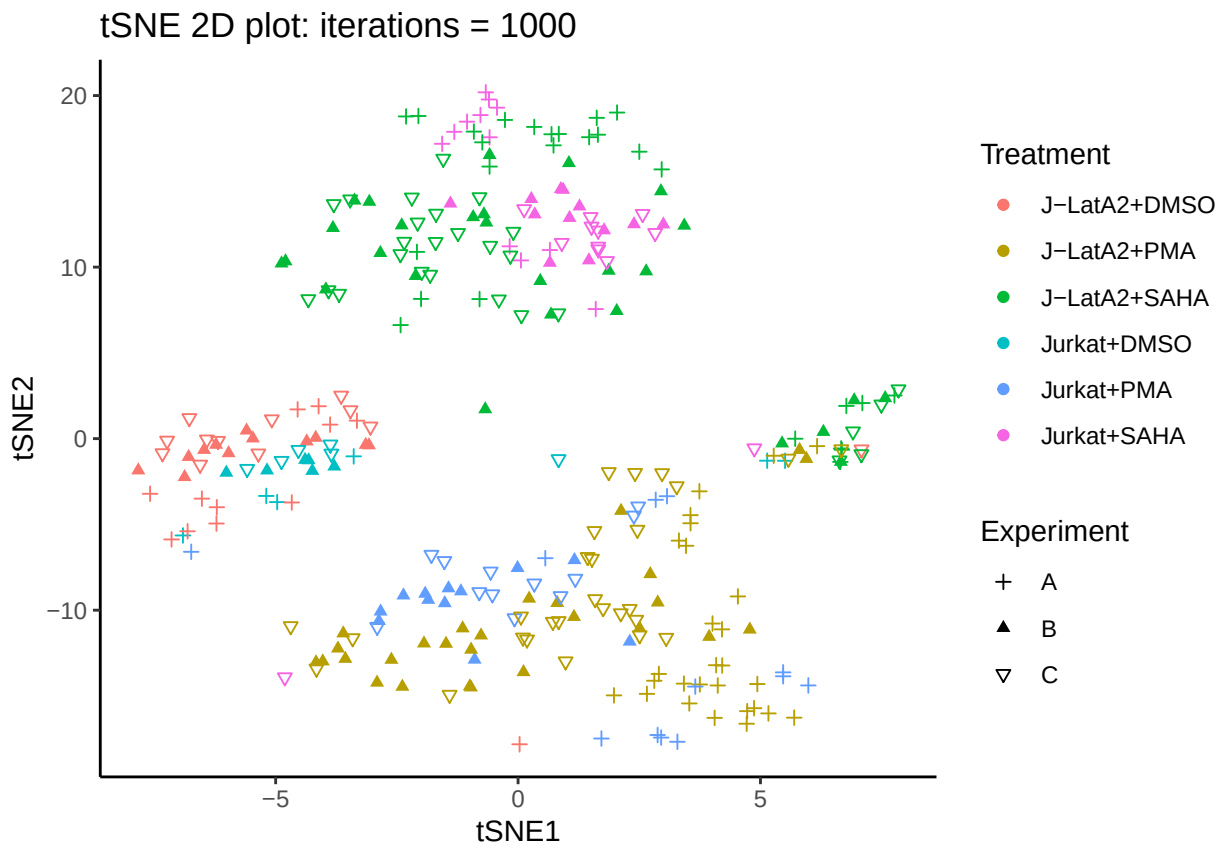

```
#changing max_iterations
set.seed(123) ## allows us to reproduce results

for (max_iter in c(1, 10, 100, 500, 1000)) {
  tsne <- Rtsne(X1)

  tsne_data <- as.data.frame(tsne$Y)
  tsne_data$Experiment <- data$experiment
  tsne_data$Treatment <- data$treatment

  library(ggplot2)
  p <- ggplot(tsne_data, aes(x=V1, y=V2)) +
   geom_point(aes(color=Treatment, shape=Experiment)) +
   scale_shape_manual(values = c(3,17,6)) +
   ggtitle(paste("tSNE 2D plot: iterations =", max_iter)) + xlab("tSNE1") + ylab("tSNE2") +
   theme_classic()
  print(p)
}
```


tSNE 2D plot: iterations = 1
tSNE2
tSNE1
10
0
−10
−20
−10
−5
0
5
10
Treatment
J−LatA2+DMSO
J−LatA2+PMA
J−LatA2+SAHA
Jurkat+DMSO
Jurkat+PMA
Jurkat+SAHA
Experiment
A
B
C


tSNE 2D plot: iterations = 10
tSNE2
tSNE1
5
0
−5
−10
−10
0
10
20
Treatment
J−LatA2+DMSO
J−LatA2+PMA
J−LatA2+SAHA
Jurkat+DMSO
Jurkat+PMA
Jurkat+SAHA
Experiment
A
B
C


tSNE 2D plot: iterations = 100
tSNE2
tSNE1
10
5
0
−5
−10
−10
0
10
20
Treatment
J−LatA2+DMSO
J−LatA2+PMA
J−LatA2+SAHA
Jurkat+DMSO
Jurkat+PMA
Jurkat+SAHA
Experiment
A
B
C


tSNE 2D plot: iterations = 500
10
5
0
−5
−10
tSNE2
−20
−10
0
10
tSNE1
Treatment
J−LatA2+DMSO
J−LatA2+PMA
J−LatA2+SAHA
Jurkat+DMSO
Jurkat+PMA
Jurkat+SAHA
Experiment
A
B
C



In the previous plots we show how increasing the perplexity or the iterations numbers for a tSNE can generate more dense and defined and more distinct clusters, respectively.

QUESTION 5

No, it is pretty clear from both the PCA and the tSNE graphs that SAHA and PMA have a different effect on the cells as the points associated to each drug are grouped very distinctly in two separate clusters. In particular, from the PCA plot we see that PMA treated cells are observed to be closer in similarity to DMSO (control) cells, so we might assume that PMA has a smaller effect than SAHA.

QUESTION 6

The PC1 separates dead from alive cells, meanwhile PC2 separates differently tretated cells. On the right we have cells with low expression of genes, on the left with high expression.

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this: