Patient FLC and FLC Sphereoid miR correlation

Alaa R. Farghli

1/27/2021

```
library(dplyr)
library(ggpubr)
library(tibble)
library(ggrepel)
```

Load CSV

```
flc.counts <- read.csv("../NML-PRI-MET_noDups_avg-normalized-counts-by-group.csv")
flc.sphereoids <- read.csv("FLC-PDX_DESeq_miRs.csv")
flc.pri <- flc.counts %>%
  select(miR, avg_PRI)
flc.mets <- flc.counts %>%
  select(miR, avg_MET)
```

Workflow with primary patient FLC primary tumors

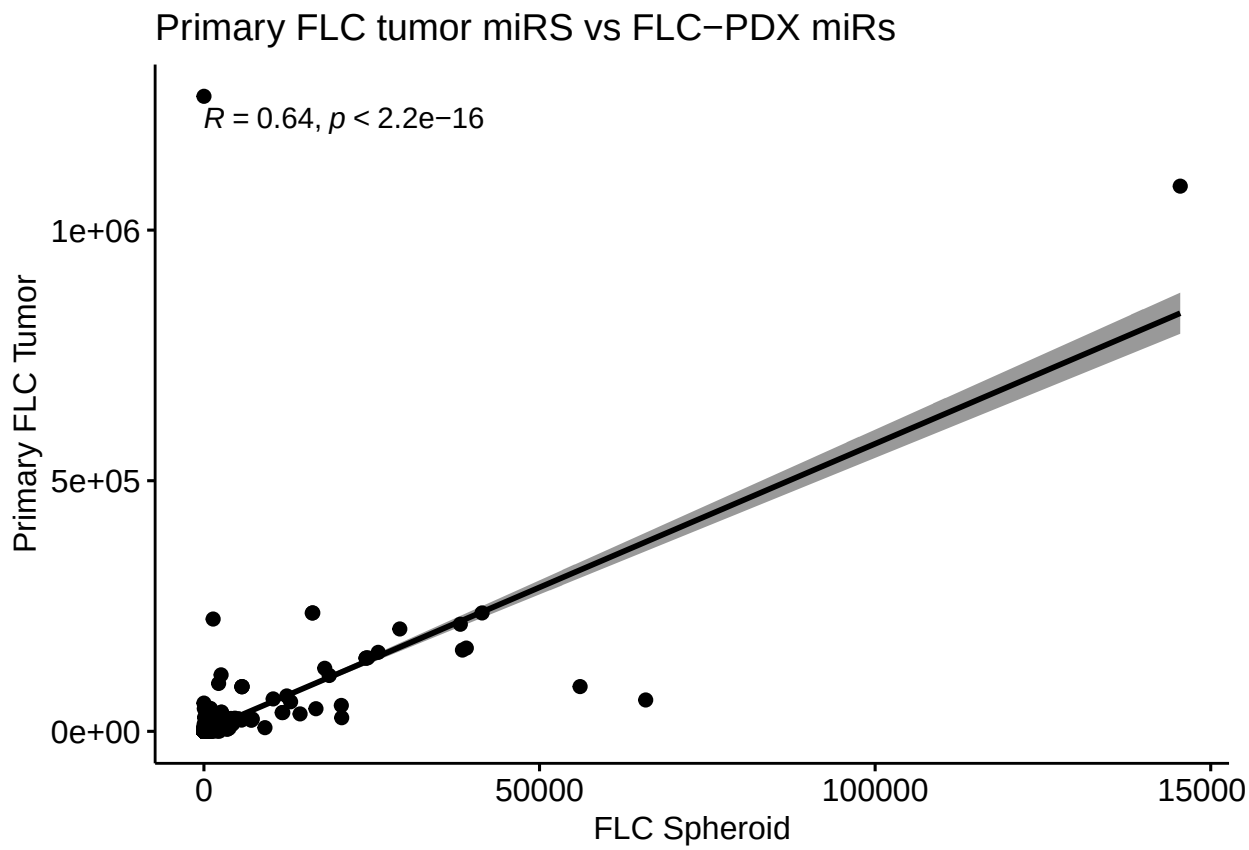
```
#Join intracellular sphereoids miRs with primary tumor miRs
flc.combined.deseq <- full_join(flc.sphereoids, flc.pri, by = "miR")
#change is.na rows to 0
flc.combined.deseq[is.na(flc.combined.deseq)] <- 0
#Move the column to rowname. this is necessary for the final clean up step of removing rows that contai
flc.combined.deseq <- column_to_rownames(flc.combined.deseq, var = "miR")
flc.clean.100 <- flc.combined.deseq %>%
  filter_all(any_vars(. > 100))
#this is done again with reads at least 50 reads in either column
flc.clean.50 <- flc.combined.deseq %>%
  filter_all(any_vars(. > 50))
```

```
most.dev.pri.genes <- flc.clean.100 %>%
  rownames_to_column(var = "miR") %>%
  filter(miR %in% c("hsa-mir-122-5p", "hsa-mir-21-5p", "hsa-mir-320a", "hsa-mir-192-5p"))
```

Plots for primary FLC primary tumors with and without log10 transformation

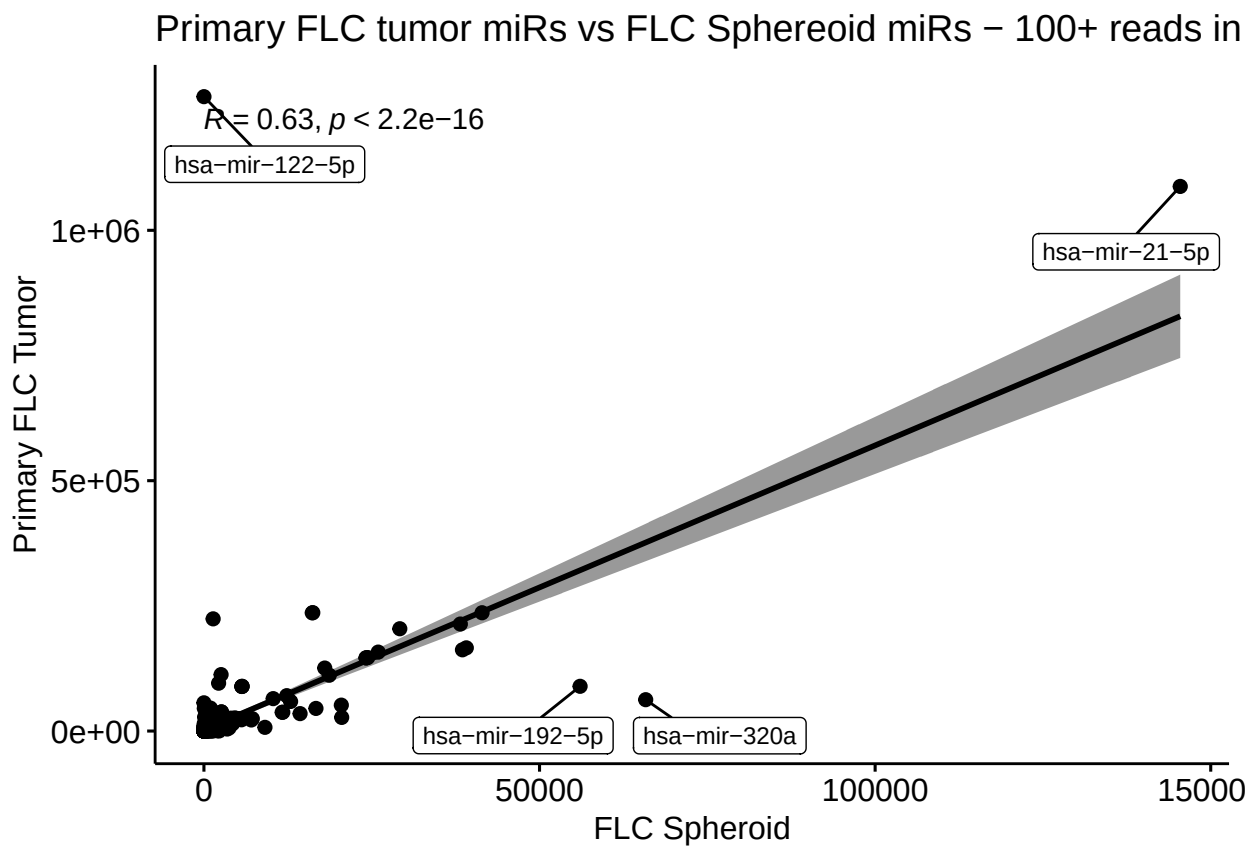
```r
set.seed(1)
p.pri.noclean <- ggscatter(flc.combined.deseq, x = "avg_Intra_FLC", y = "avg_PRI",
          add = "reg.line", conf.int = TRUE,
          cor.coef = TRUE, cor.method = "pearson",
          xlab = "FLC Spheroid", ylab = "Primary FLC Tumor", title = "Primary FLC tumor miRS vs FLC-PDX

p.pri.100 <- ggscatter(flc.clean.100, x = "avg_Intra_FLC", y = "avg_PRI",
          add = "reg.line", conf.int = TRUE,
          cor.coef = TRUE, cor.method = "pearson",
          xlab = "FLC Spheroid", ylab = "Primary FLC Tumor", title = "Primary FLC tumor miRs vs FLC Sph
  geom_label_repel(data = most.dev.pri.genes,
                   aes(label = miR),
                   size = 3,
                   force = 100,
                   nudge_y = -30000)

p.pri.100.log10 <- ggscatter(flc.clean.100, x = "avg_Intra_FLC", y = "avg_PRI",
          add = "reg.line", conf.int = TRUE,
          cor.coef = TRUE, cor.method = "pearson",
          xlab = "log10 FLC Spheroid", ylab = "log 10 Primary FLC", title = "log10(Primary FLC tumor mi
  scale_x_log10() +
  scale_y_log10() +
  geom_label_repel(data = most.dev.pri.genes,
                   aes(label = miR),
                   size = 3) +
  geom_label_repel(data = most.dev.pri.genes,
                   aes(label = miR),
                   size = 3,
                   nudge_y = -1)

p.pri.50 <- ggscatter(flc.clean.50, x = "avg_Intra_FLC", y = "avg_PRI",
          add = "reg.line", conf.int = TRUE,
          cor.coef = TRUE, cor.method = "pearson",
          xlab = "FLC Spheroid", ylab = "Primary FLC Tumor", title = "Primary FLC tumor miRs vs FLC Sph
  geom_label_repel(data = most.dev.pri.genes,
                   aes(label = miR),
                   size = 3,
                   force = 100,
                   nudge_y = -30000)

p.pri.50.log10 <- ggscatter(flc.clean.50, x = "avg_Intra_FLC", y = "avg_PRI",
          add = "reg.line", conf.int = TRUE,
          cor.coef = TRUE, cor.method = "pearson",
          xlab = "log10 FLC Spheroid", ylab = "log 10 Primary FLC Tumor", title = "log10(Primary FLC tu
  scale_x_log10() +
  scale_y_log10() +
  geom_label_repel(data = most.dev.pri.genes,
                   aes(label = miR),
                   size = 3,
                   nudge_y = -1)

p.pri.noclean
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
p.pri.100
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
p.pri.100.log10
```

```
## Warning: Transformation introduced infinite values in continuous x-axis
```

```
## Warning: Transformation introduced infinite values in continuous y-axis
```

```
## Warning: Transformation introduced infinite values in continuous x-axis
```

```
## Warning: Transformation introduced infinite values in continuous y-axis
```

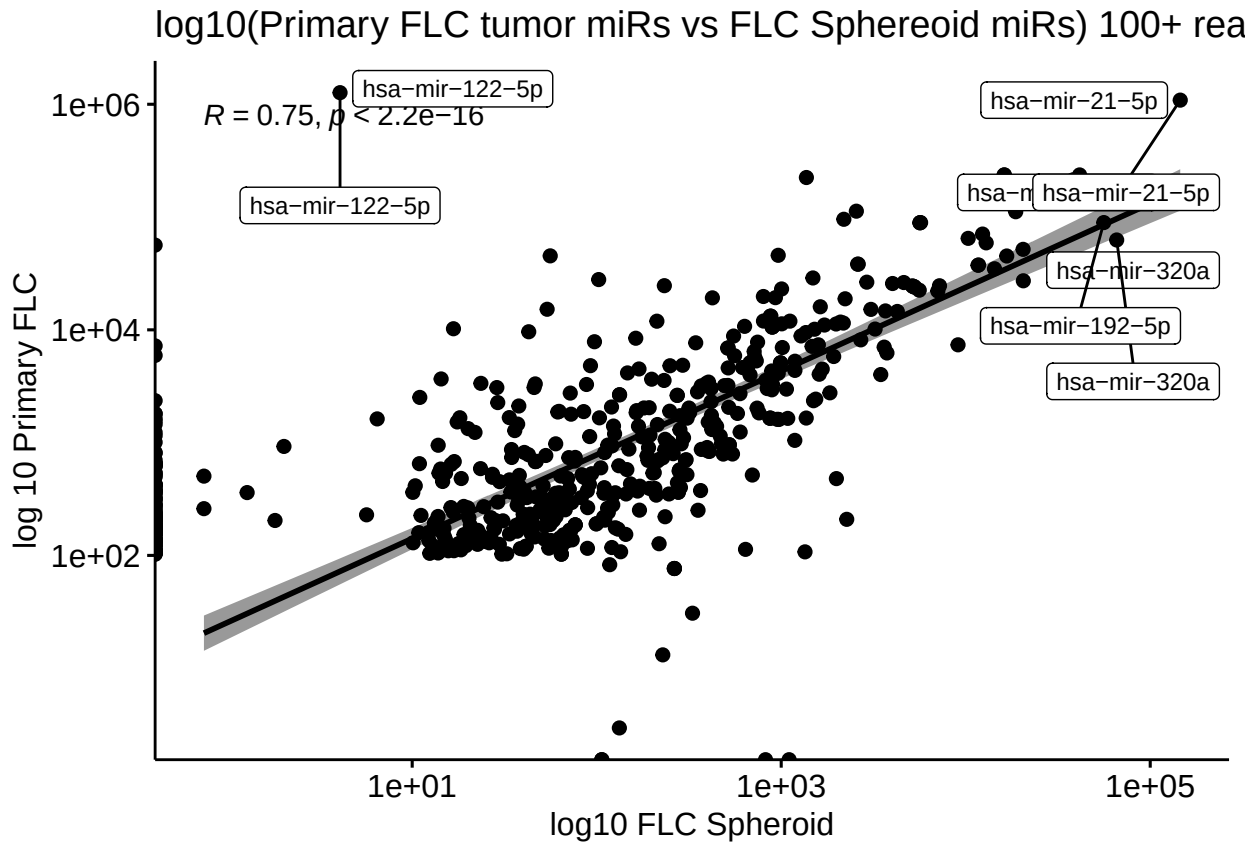
```
## Warning: Transformation introduced infinite values in continuous x-axis
```

```
## Warning: Transformation introduced infinite values in continuous y-axis
```

```
## `geom_smooth()` using formula 'y ~ x'
```

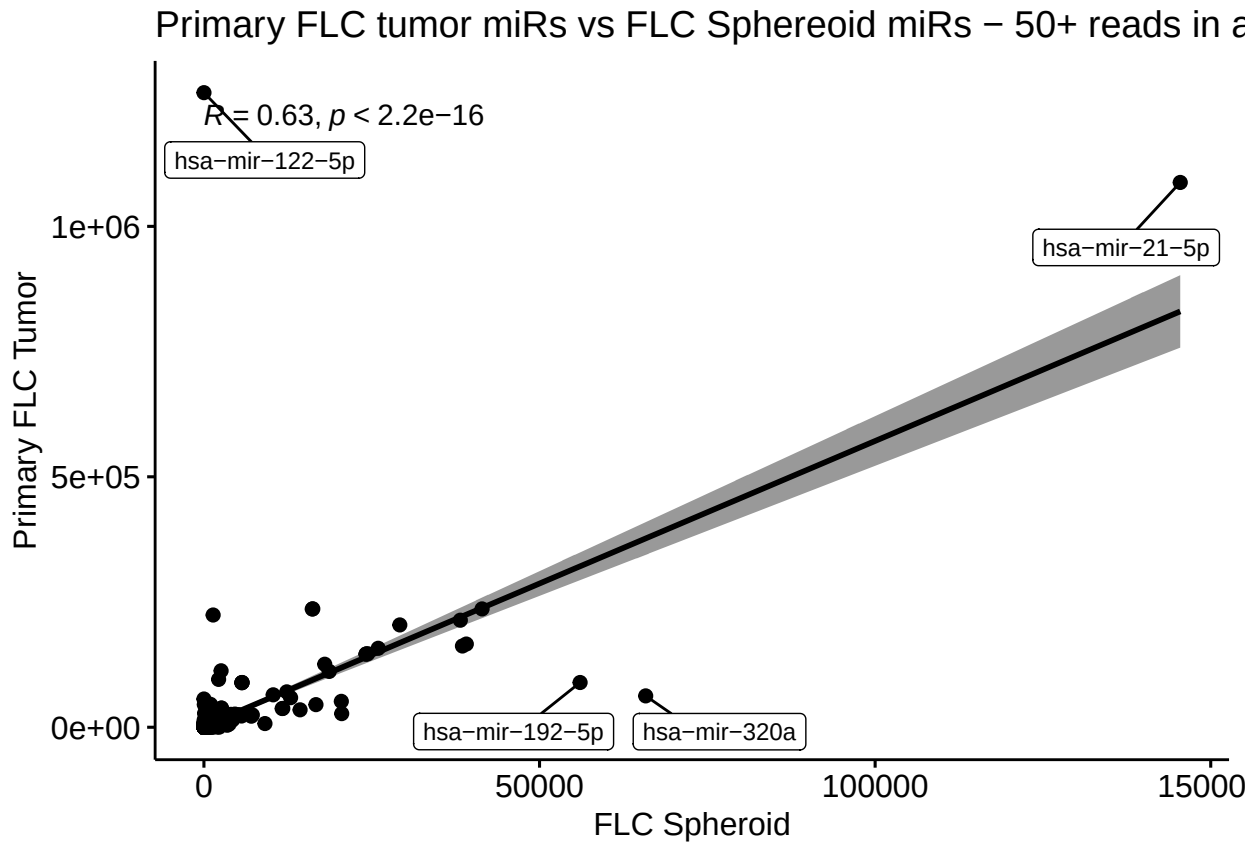
```
## Warning: Removed 119 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 119 rows containing non-finite values (stat_cor).
```



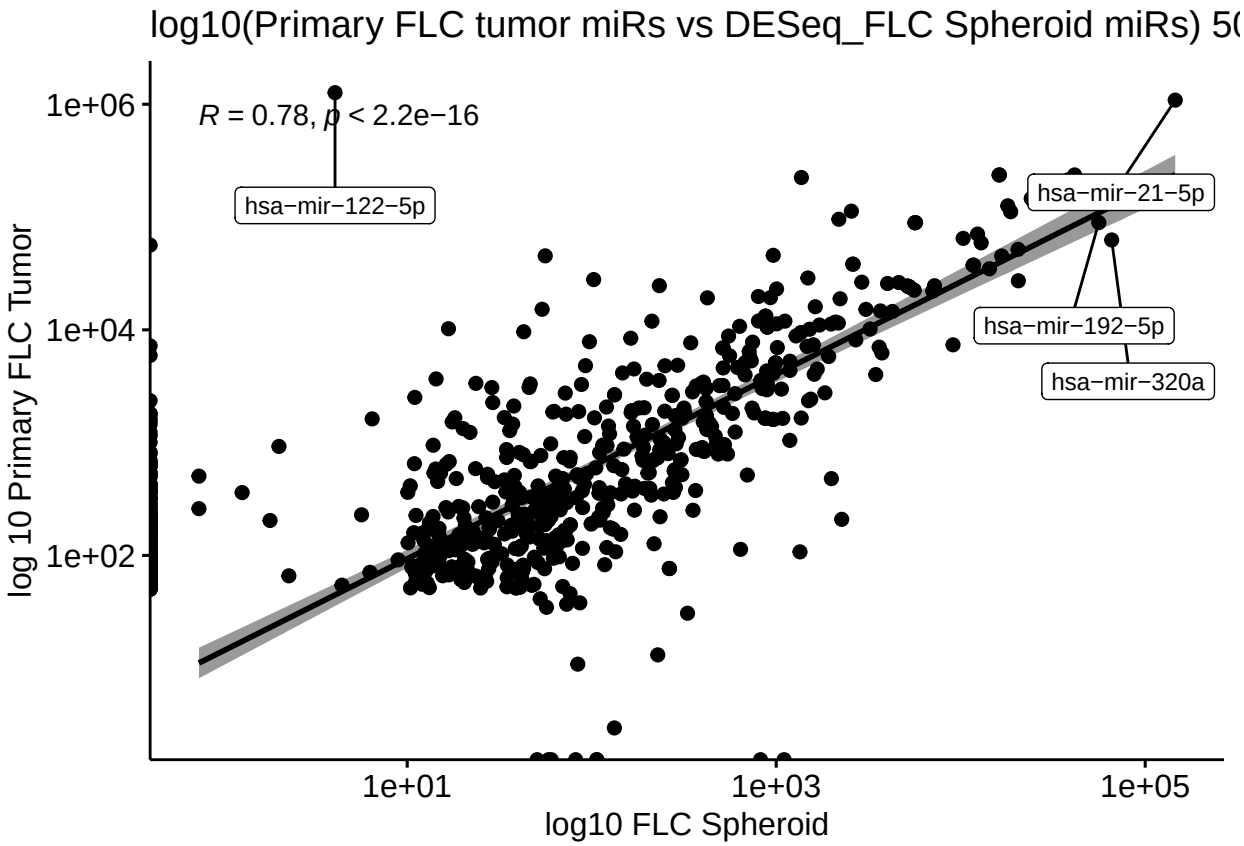

```
p.pri.50
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
p.pri.50.log10
```

```
## Warning: Transformation introduced infinite values in continuous x-axis

## Warning: Transformation introduced infinite values in continuous y-axis

## Warning: Transformation introduced infinite values in continuous x-axis

## Warning: Transformation introduced infinite values in continuous y-axis

## Warning: Transformation introduced infinite values in continuous x-axis

## Warning: Transformation introduced infinite values in continuous y-axis

## `geom_smooth()` using formula 'y ~ x'

## Warning: Removed 213 rows containing non-finite values (stat_smooth).

## Warning: Removed 213 rows containing non-finite values (stat_cor).
```



Workflow with primary patient FLC metastatic tumors

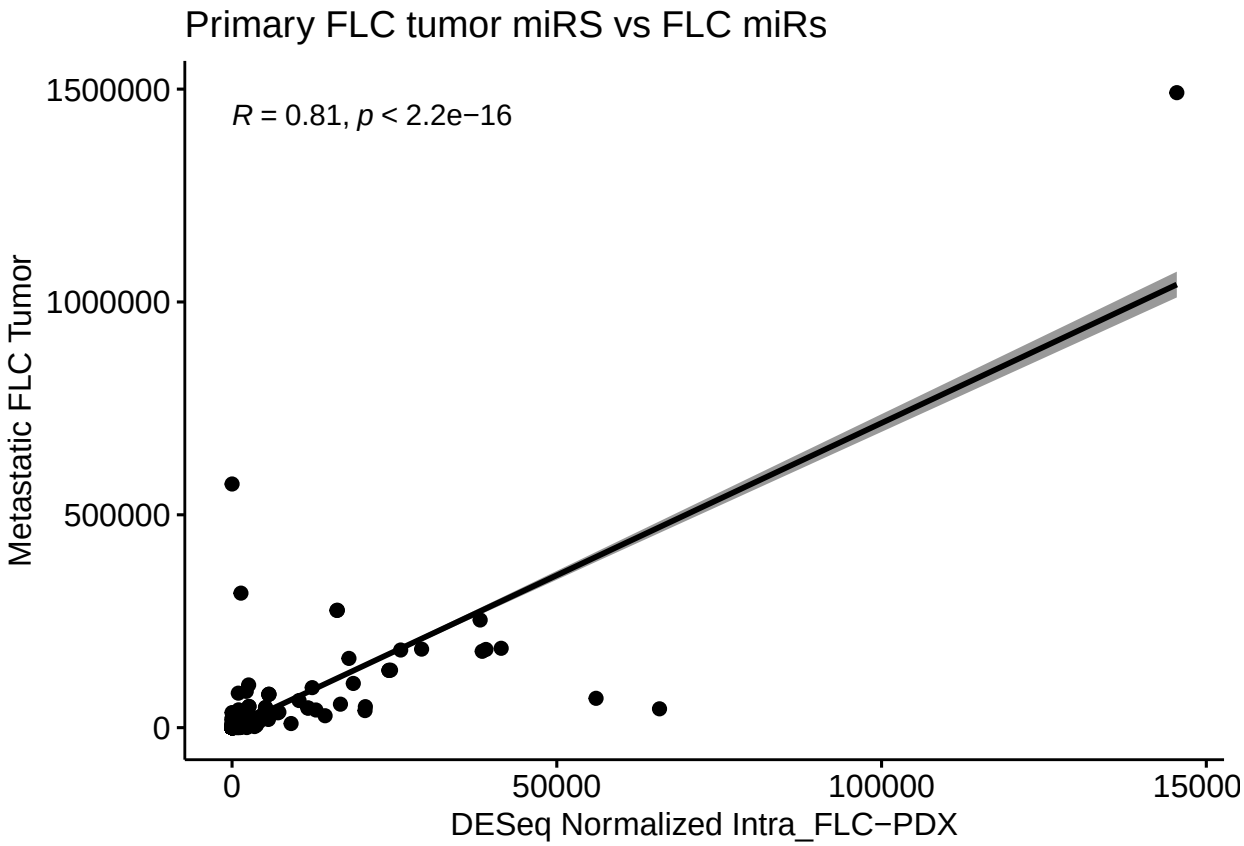
```
#The steps here are similar to that seen in code chunk 4
flc.combined.mets <- full_join(flc.mets, flc.sphereoids, by = "miR")
flc.combined.mets[is.na(flc.combined.mets)] <- 0
flc.combined.mets <- column_to_rownames(flc.combined.mets, var = "miR")
flc.clean.mets.100 <- flc.combined.mets %>%
  filter_all(any_vars(. > 100))
flc.clean.mets.50 <- flc.combined.mets %>%
  filter_all(any_vars(. > 50))
```

```
most.dev.met.genes <- flc.clean.mets.100 %>%
  rownames_to_column(var = "miR") %>%
  filter(miR %in% c("hsa-mir-21-5p", "hsa-mir-122-5p", "hsa-mir-143-3p","hsa-mir-26a-1-5p" , "hsa-mir-3
```

```
set.seed(1)
p.met.noclean <- ggscatter(flc.combined.mets, x = "avg_Intra_FLC", y = "avg_MET",
          add = "reg.line", conf.int = TRUE,
          cor.coef = TRUE, cor.method = "pearson",
          xlab = "DESeq Normalized Intra_FLC-PDX", ylab = "Metastatic FLC Tumor", title = "Primary FLC

p.met.100 <- ggscatter(flc.clean.mets.100, x = "avg_Intra_FLC", y = "avg_MET",
          add = "reg.line", conf.int = TRUE,
```

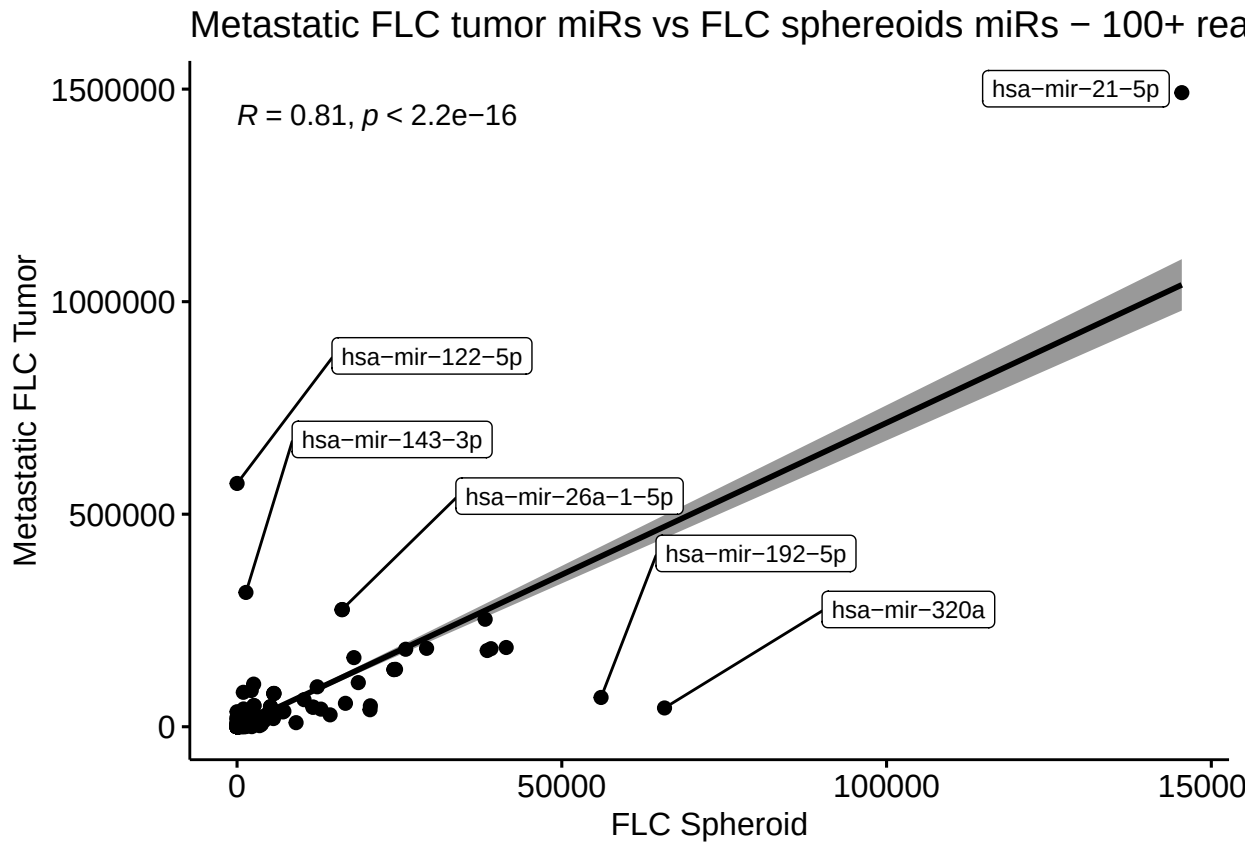
```
          cor.coef = TRUE, cor.method = "pearson",
          xlab = "FLC Spheroid", ylab = "Metastatic FLC Tumor", title = "Metastatic FLC tumor miRs vs Fl
  geom_label_repel(data = most.dev.met.genes,
                   aes(label = miR),
                   size = 3,
                   force = 100,
                   nudge_x = 30000,
                   nudge_y = 300000)

p.met.100.log10 <- ggscatter(flc.clean.mets.100, x = "avg_Intra_FLC", y = "avg_MET",
          add = "reg.line", conf.int = TRUE,
          cor.coef = TRUE, cor.method = "pearson",
          xlab = "log10 FLC SpheroidC", ylab = "log10 Metastatic FLC Tumor", title = "log10(Metastatic
  scale_x_log10() +
  scale_y_log10() +
  geom_text_repel(data = most.dev.met.genes,
                   aes(label = miR),
                   size = 3,
                   force = 100,
                   nudge_y = .5)

p.met.50 <- ggscatter(flc.clean.mets.50, x = "avg_Intra_FLC", y = "avg_MET",
          add = "reg.line", conf.int = TRUE,
          cor.coef = TRUE, cor.method = "pearson",
          xlab = "FLC Spheroid", ylab = "Metastatic FLC Tumor", title = "Metastatic FLC tumor miRs vs Fl
  geom_label_repel(data = most.dev.met.genes,
                   aes(label = miR),
                   size = 3,
                   force = 100,
                   nudge_x = 30000,
                   nudge_y = 300000)

p.met.50.log10 <- ggscatter(flc.clean.mets.50, x = "avg_Intra_FLC", y = "avg_MET",
          add = "reg.line", conf.int = TRUE,
          cor.coef = TRUE, cor.method = "pearson",
          xlab = "log10 FLC SpheroidC", ylab = "log10 Metastatic FLC Tumor", title = "log10(Metastatic
  scale_x_log10() +
  scale_y_log10() +
  geom_text_repel(data = most.dev.met.genes,
                   aes(label = miR),
                   size = 3,
                   force = 100,
                  nudge_y = .5)

p.met.noclean
```

```
## `geom_smooth()` using formula 'y ~ x'
```



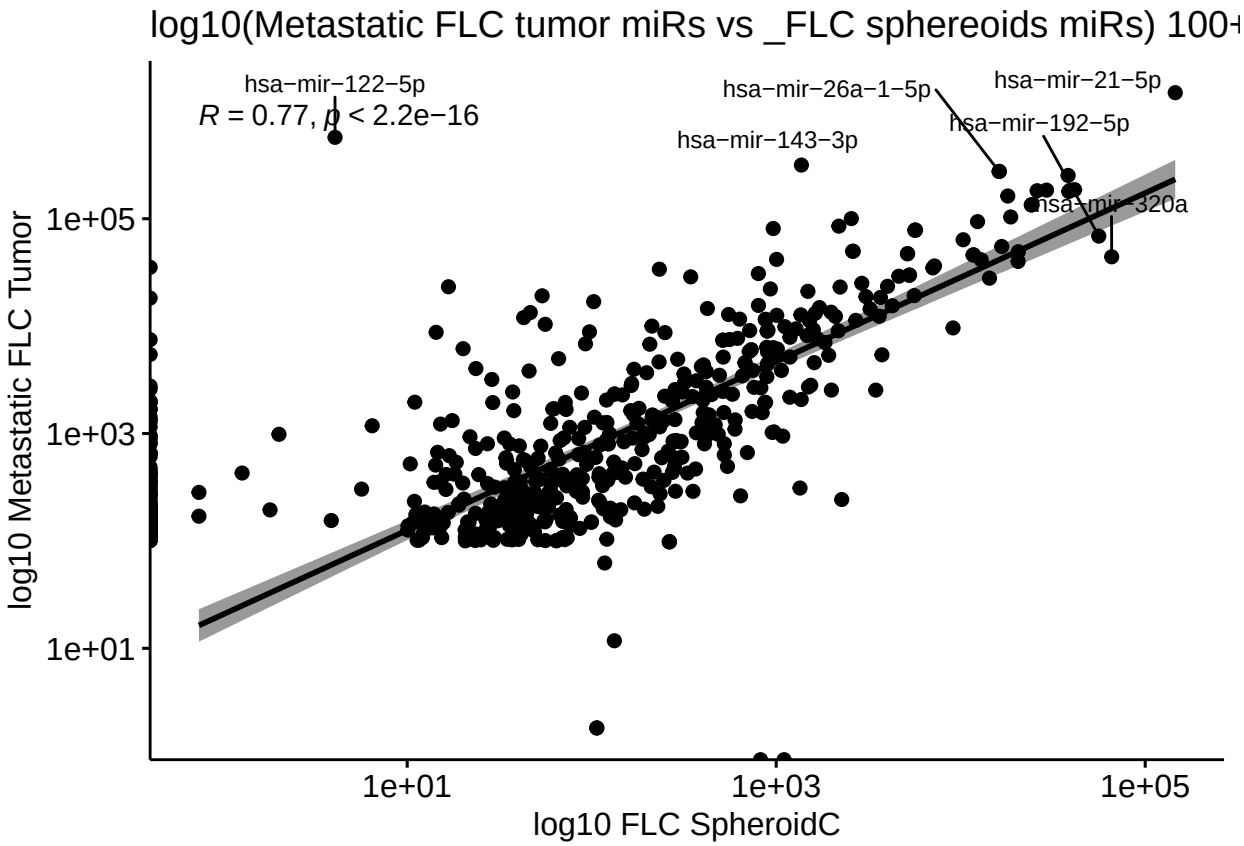

```
p.met.100
```

```
## `geom_smooth()` using formula 'y ~ x'
```



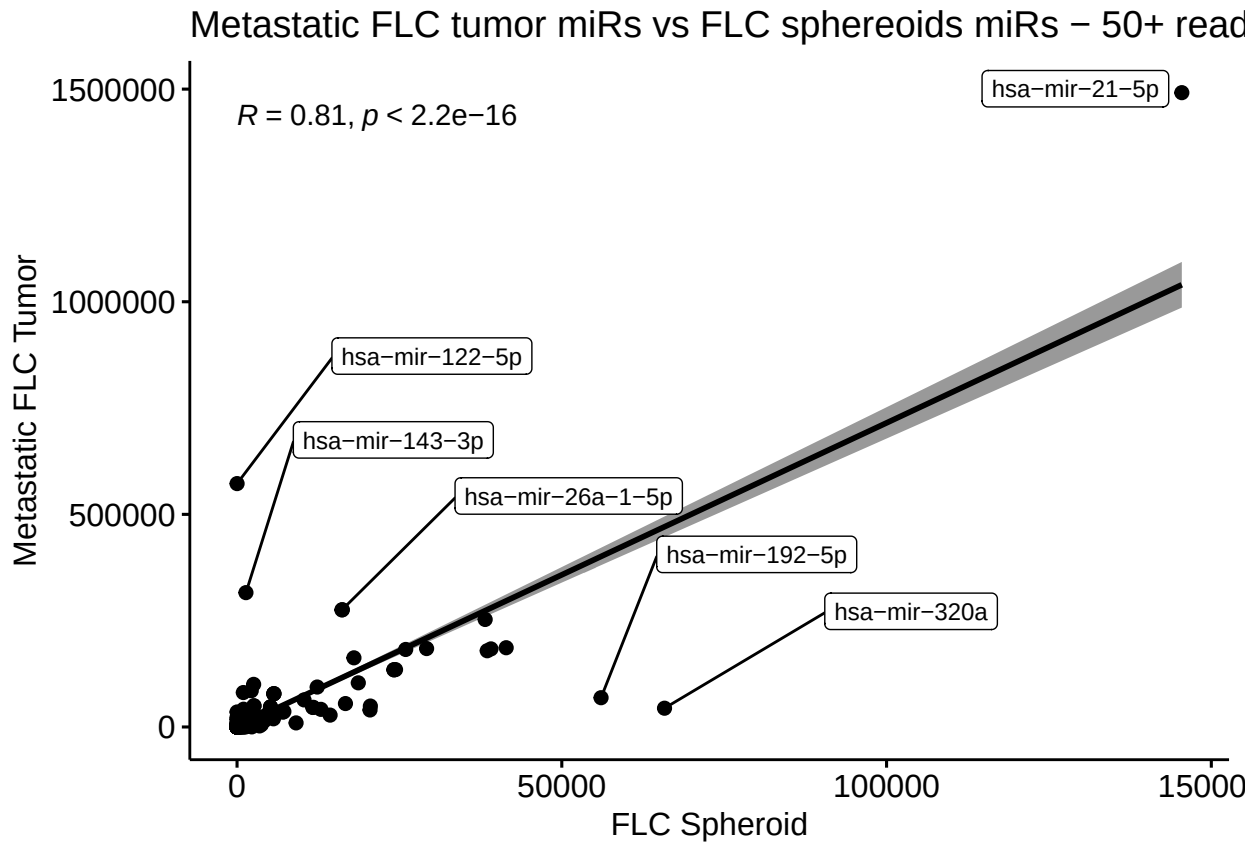

```
p.met.100.log10
```

```
## Warning: Transformation introduced infinite values in continuous x-axis

## Warning: Transformation introduced infinite values in continuous y-axis

## Warning: Transformation introduced infinite values in continuous x-axis

## Warning: Transformation introduced infinite values in continuous y-axis

## Warning: Transformation introduced infinite values in continuous x-axis

## Warning: Transformation introduced infinite values in continuous y-axis

## `geom_smooth()` using formula 'y ~ x'

## Warning: Removed 120 rows containing non-finite values (stat_smooth).

## Warning: Removed 120 rows containing non-finite values (stat_cor).
```



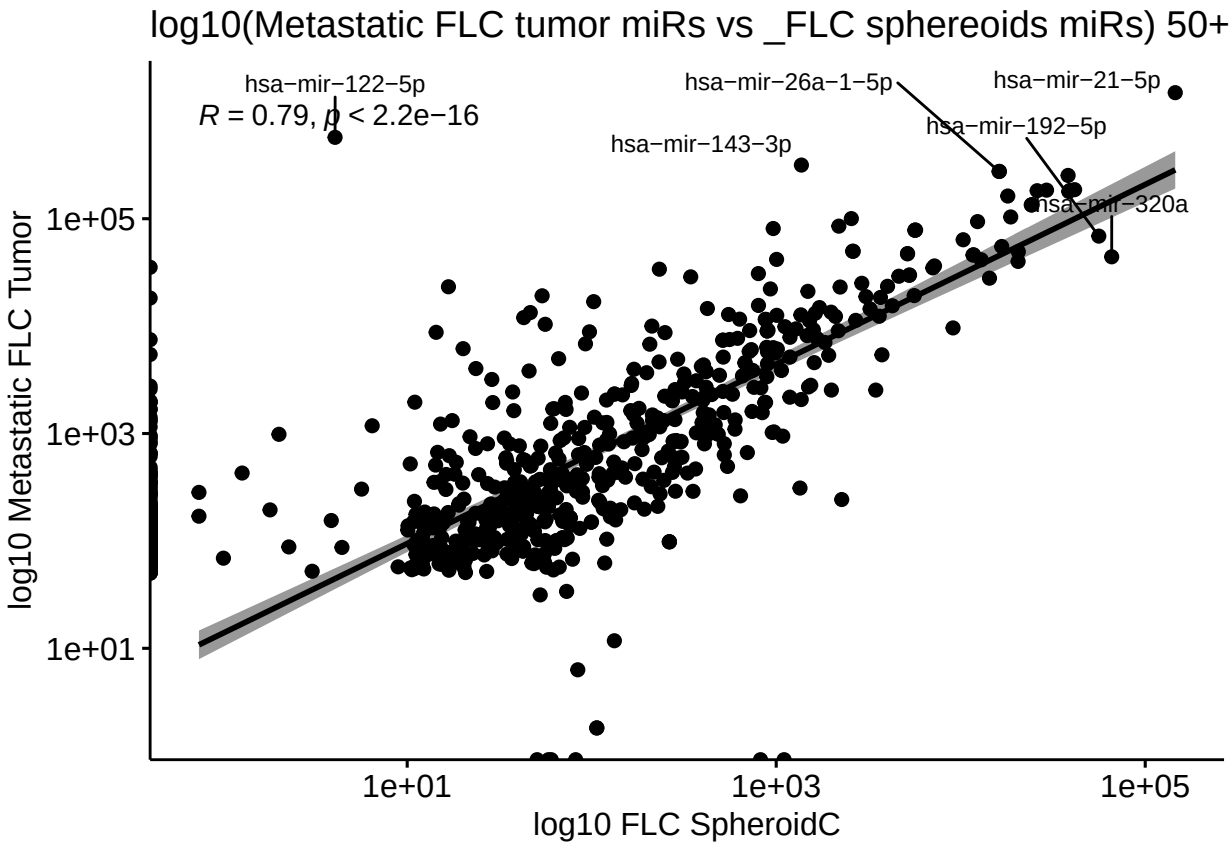

```
p.met.50
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
p.met.50.log10
```

```
## Warning: Transformation introduced infinite values in continuous x-axis

## Warning: Transformation introduced infinite values in continuous y-axis

## Warning: Transformation introduced infinite values in continuous x-axis

## Warning: Transformation introduced infinite values in continuous y-axis

## Warning: Transformation introduced infinite values in continuous x-axis

## Warning: Transformation introduced infinite values in continuous y-axis

## `geom_smooth()` using formula 'y ~ x'

## Warning: Removed 203 rows containing non-finite values (stat_smooth).

## Warning: Removed 203 rows containing non-finite values (stat_cor).
```


log10(Metastatic FLC tumor miRs vs _FLC sphereoids miRs) 50+
hsa−mir−122−5p
R = 0.79, p < 2.2e−16
hsa−mir−26a−1−5p
hsa−mir−21−5p
hsa−mir−192−5p
hsa−mir−143−3p
hsa−mir−320a
log10 Metastatic FLC Tumor
1e+05
1e+03
1e+01
1e+01
1e+03
1e+05
log10 FLC SpheroidC