RdatAI190418 AI190506

```
# Loading packages
library(ggplot2)
library(reshape2)
library(ggpubr)
```

```
## Loading required package: magrittr
```

```
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(RColorBrewer)
library(plotly)
```

```
##
## Attaching package: 'plotly'
```

```
## The following object is masked from 'package:ggplot2':
##
##     last_plot
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

```
## The following object is masked from 'package:graphics':
##
##     layout
```

```
library(rstatix)
```

```
##
## Attaching package: 'rstatix'
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

```
library(RATest)
```

```
## Loading required package: gridExtra
```

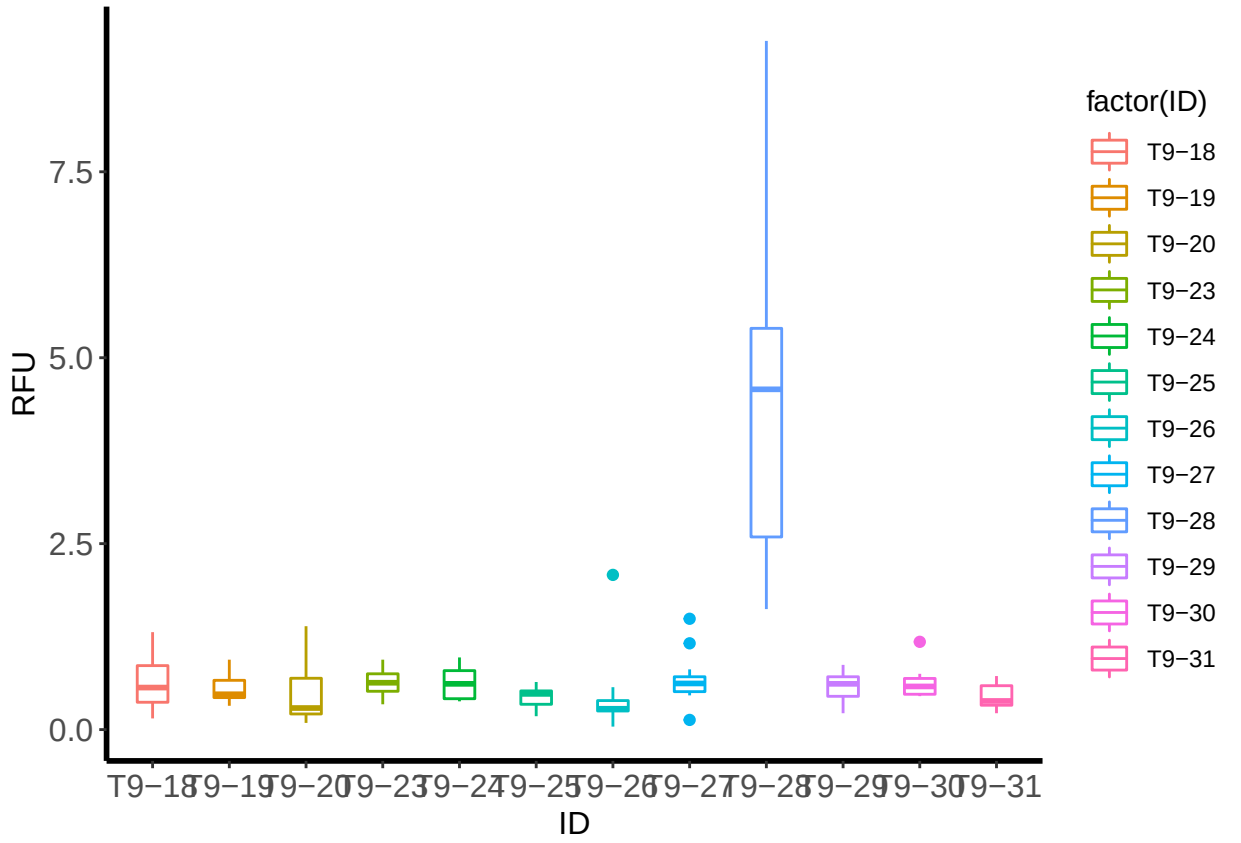
```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##     combine
```

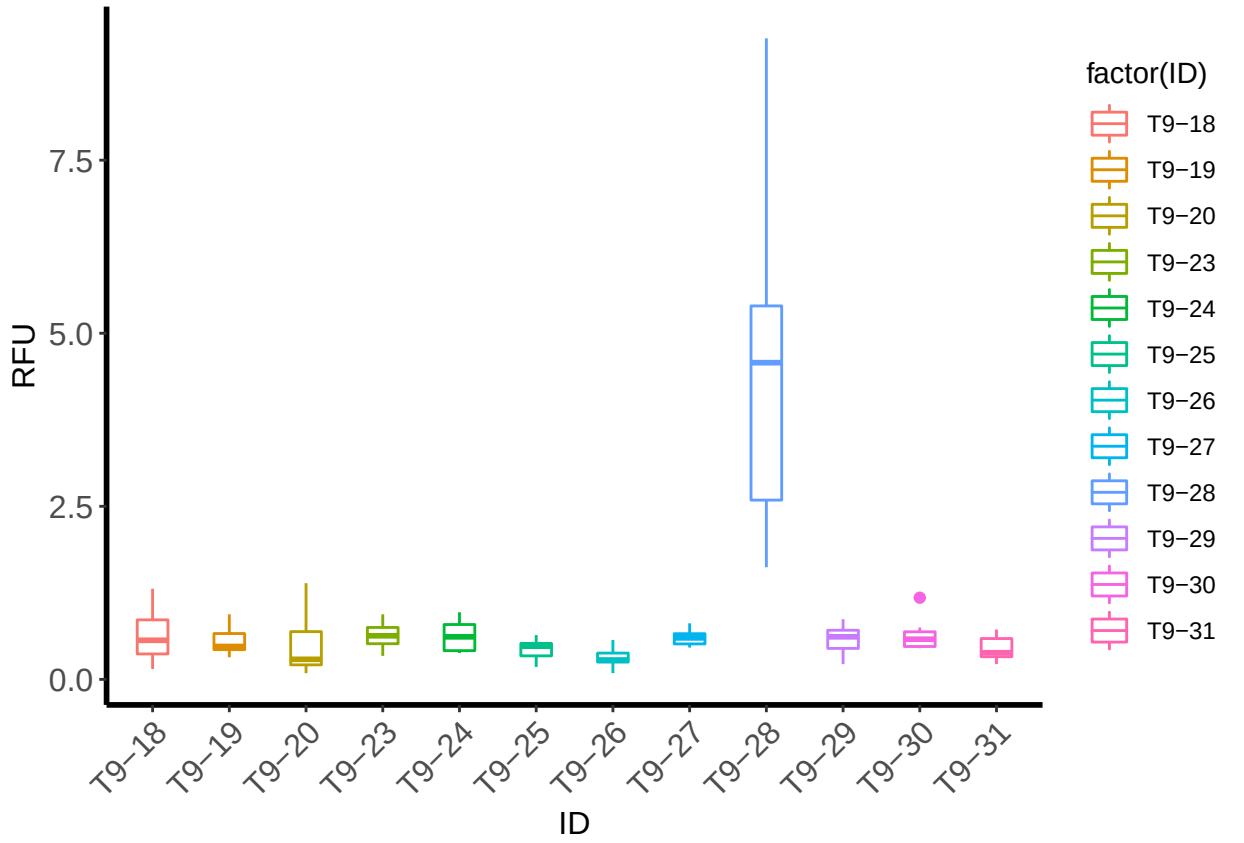
```
AIsumm_df <- read.table("AIsumm.csv", header=TRUE,
                        sep=",") # merged fluorecence data of CRISPR assay 190418 and 190506
attach(AIsumm_df)


# changing the df to long format

AIsumm_df_long <- melt(AIsumm_df,
                       id.vars=c("ID"),
                       measure.vars=c("GFP", "RFP", "GFP.RFP" ),
                       variable.name="Channel",
                       value.name="RFU")
```

Boxplots:

```
# Creating df with only GFP/RFP ratios in long format

AIsumm_df_long_red2 <- AIsumm_df_long[-c(1:282), ]

pall <- ggplot(data=AIsumm_df_long_red2, aes(x=ID,y=RFU, colour=factor(ID))) +
  geom_boxplot(position=position_dodge2(width=0.4), width=0.4) +
  #scale_y_continuous(breaks = 1) +
  #scale_color_grey() +
  theme_classic() +
  #coord_fixed(ratio = 3, xlim = NULL, ylim = c(0,1), expand = TRUE, clip = "on") +
  theme(axis.line = element_line(colour = "black", size = 1, linetype = "solid"),
        axis.title.x = element_text(size = 12),
        axis.title.y = element_text(size = 12),
        axis.text.x= element_text(size=12),
        axis.text.y = element_text(size=12)
        )

pall
```




```
# removing outliers in 9-26 (1) and 9-27 (3)
AIsumm_df_long_red2_mod <-AIsumm_df_long_red2[-c(105,116,124,126,127), ]

pall_mod <- ggplot(data=AIsumm_df_long_red2_mod, aes(x=ID,y=RFU, colour=factor(ID))) +
  geom_boxplot(position=position_dodge2(width=0.4), width=0.4) +
  #scale_y_continuous(breaks = 1) +
  #scale_color_grey() +
  #colorRampPalette(brewer.pal(n_palette, "palette_name"))(n_plot),
  theme_classic() +
  #coord_fixed(ratio = 3, xlim = NULL, ylim = c(0,1), expand = TRUE, clip = "on") +
  theme(axis.line = element_line(colour = "black", size = 1, linetype = "solid"),
        axis.title.x = element_text(size = 12),
        axis.title.y = element_text(size = 12),
        axis.text.x= element_text(size=12, angle = 45, hjust = 1),
        axis.text.y = element_text(size=12)
        )

pall_mod
```



Separating data according to Cas-type:

```
AIsumm1_df <- read.table("AIsumm1.csv", header=TRUE,
                         sep=",") # merged fluorecence data of CRISPR assay 190418 and 190506
AIsumm2_df <- read.table("AIsumm2.csv", header=TRUE,
                         sep=",") # merged fluorecence data of CRISPR assay 190418 and 190506
AIsumm3_df <- read.table("AIsumm3.csv", header=TRUE,
                         sep=",") # merged fluorecence data of CRISPR assay 190418 and 190506

attach(AIsumm1_df)
```

```
## The following objects are masked from AIsumm_df:
##
##     GFP, GFP.RFP, ID, RFP
```

```
attach(AIsumm2_df)
```

```
## The following objects are masked from AIsumm1_df:
##
##     GFP, GFP.RFP, ID, RFP
##
## The following objects are masked from AIsumm_df:
##
##     GFP, GFP.RFP, ID, RFP
```

```
attach(AIsumm3_df)
```

```
## The following objects are masked from AIsumm2_df:
```

```
##
##     GFP, GFP.RFP, ID, RFP

## The following objects are masked from AIsumm1_df:
##
##     GFP, GFP.RFP, ID, RFP

## The following objects are masked from AIsumm_df:
##
##     GFP, GFP.RFP, ID, RFP
```

```
AIsumm1_df_long <- melt(AIsumm1_df,
                        id.vars=c("ID"),
                        measure.vars=c("GFP", "RFP", "GFP.RFP" ),
                        variable.name="Channel",
                        value.name="RFU")

AIsumm2_df_long <- melt(AIsumm2_df,
                        id.vars=c("ID"),
                        measure.vars=c("GFP", "RFP", "GFP.RFP" ),
                        variable.name="Channel",
                        value.name="RFU")

AIsumm3_df_long <- melt(AIsumm3_df,
                        id.vars=c("ID"),
                        measure.vars=c("GFP", "RFP", "GFP.RFP" ),
                        variable.name="Channel",
                        value.name="RFU")

AIsumm1_df_long_red <- AIsumm1_df_long[-c(87:129), ]
AIsumm1_df_long_red2 <- AIsumm1_df_long[-c(1:86), ]

AIsumm2_df_long_red <- AIsumm2_df_long[-c(77:114), ]
AIsumm2_df_long_red2 <- AIsumm2_df_long[-c(1:76), ]

AIsumm3_df_mod <-  AIsumm3_df[-c(24, 35, 43, 45, 46), ]
AIsumm3_df_long_red <- AIsumm3_df_long[-c(121:180), ]
AIsumm3_df_long_red2 <- AIsumm3_df_long[-c(1:120), ]

#removing outliers from 9-26 (2) and 9-27 (3)
AIsumm3_df_long_red2_mod <- AIsumm3_df_long_red2[-c(24, 35, 43, 45, 46), ]


# summarize GFP or RFP between groups
group_by(AIsumm1_df, ID) %>%
  summarise(
    count = n(),
    #medianG = median(GFP, na.rm = TRUE),
    meanG = mean(GFP, na.rm = TRUE),
    sdevG = sd(GFP, na.rm = TRUE),
    #medianR = median(RFP, na.rm = TRUE),
    meanR = mean(RFP, na.rm = TRUE),
    sdevR = sd(RFP, na.rm = TRUE)
    #IQR = IQR(GFP, na.rm = TRUE)
  )
```

```
## # A tibble: 4 x 6
##   ID    count meanG sdevG meanR sdevR
##   <fct> <int> <dbl> <dbl> <dbl> <dbl>
## 1 T9-18    10  2.06 1.28   3.28 0.820
## 2 T9-19     8  1.79 0.578  3.37 0.837
## 3 T9-20    13  1.47 1.13   3.22 0.611
## 4 T9-29    12  2.81 1.53   4.85 2.08
```

```
group_by(AIsumm2_df, ID) %>%
  summarise(
    count = n(),
    #medianG = median(GFP, na.rm = TRUE),
    meanG = mean(GFP, na.rm = TRUE),
    sdevG = sd(GFP, na.rm = TRUE),
    #medianR = median(RFP, na.rm = TRUE),
    meanR = mean(RFP, na.rm = TRUE),
    sdevR = sd(RFP, na.rm = TRUE)
    #IQR = IQR(GFP, na.rm = TRUE)
  )
```

```
## # A tibble: 4 x 6
##   ID    count meanG sdevG meanR sdevR
##   <fct> <int> <dbl> <dbl> <dbl> <dbl>
## 1 T9-23    11  2.14 0.826  3.40 0.943
## 2 T9-24     8  2.39 0.767  3.92 0.875
## 3 T9-25     9  1.50 0.389  3.71 0.801
## 4 T9-30    10  2.87 0.864  4.75 1.49
```

```
group_by(AIsumm3_df, ID) %>%
  summarise(
    count = n(),
    #medianG = median(GFP, na.rm = TRUE),
    meanG = mean(GFP, na.rm = TRUE),
    sdevG = sd(GFP, na.rm = TRUE),
    #medianR = median(RFP, na.rm = TRUE),
    meanR = mean(RFP, na.rm = TRUE),
    sdevR = sd(RFP, na.rm = TRUE)
    #IQR = IQR(GFP, na.rm = TRUE)
  )
```

```
## # A tibble: 4 x 6
##   ID    count meanG sdevG meanR sdevR
##   <fct> <int> <dbl> <dbl> <dbl> <dbl>
## 1 T9-26    21  2.49 2.70   6.22  2.75
## 2 T9-27    13  3.36 2.99   4.54  1.71
## 3 T9-28    12 19.4  8.46   4.61  1.87
## 4 T9-31    14  2.28 0.877  5.41  2.18
```

```
# summarize GFP/RFP ratios between groups
group_by(AIsumm1_df_long_red2, ID) %>%
  summarise(
    count = n(),
    mean = mean(RFU, na.rm = TRUE),
    sdev = sd(RFU, na.rm = TRUE)
  )
```

```
## # A tibble: 4 x 4
##   ID    count  mean  sdev
##   <fct> <int> <dbl> <dbl>
## 1 T9-18    10 0.643 0.389
## 2 T9-19     8 0.555 0.216
## 3 T9-20    13 0.465 0.372
## 4 T9-29    12 0.578 0.190
```

```
group_by(AIsumm2_df_long_red2, ID) %>%
  summarise(
    count = n(),
    mean = mean(RFU, na.rm = TRUE),
    sdev = sd(RFU, na.rm = TRUE)
  )
```

```
## # A tibble: 4 x 4
##   ID    count  mean  sdev
##   <fct> <int> <dbl> <dbl>
## 1 T9-23    11 0.640 0.176
## 2 T9-24     8 0.631 0.241
## 3 T9-25     9 0.427 0.145
## 4 T9-30    10 0.632 0.219
```

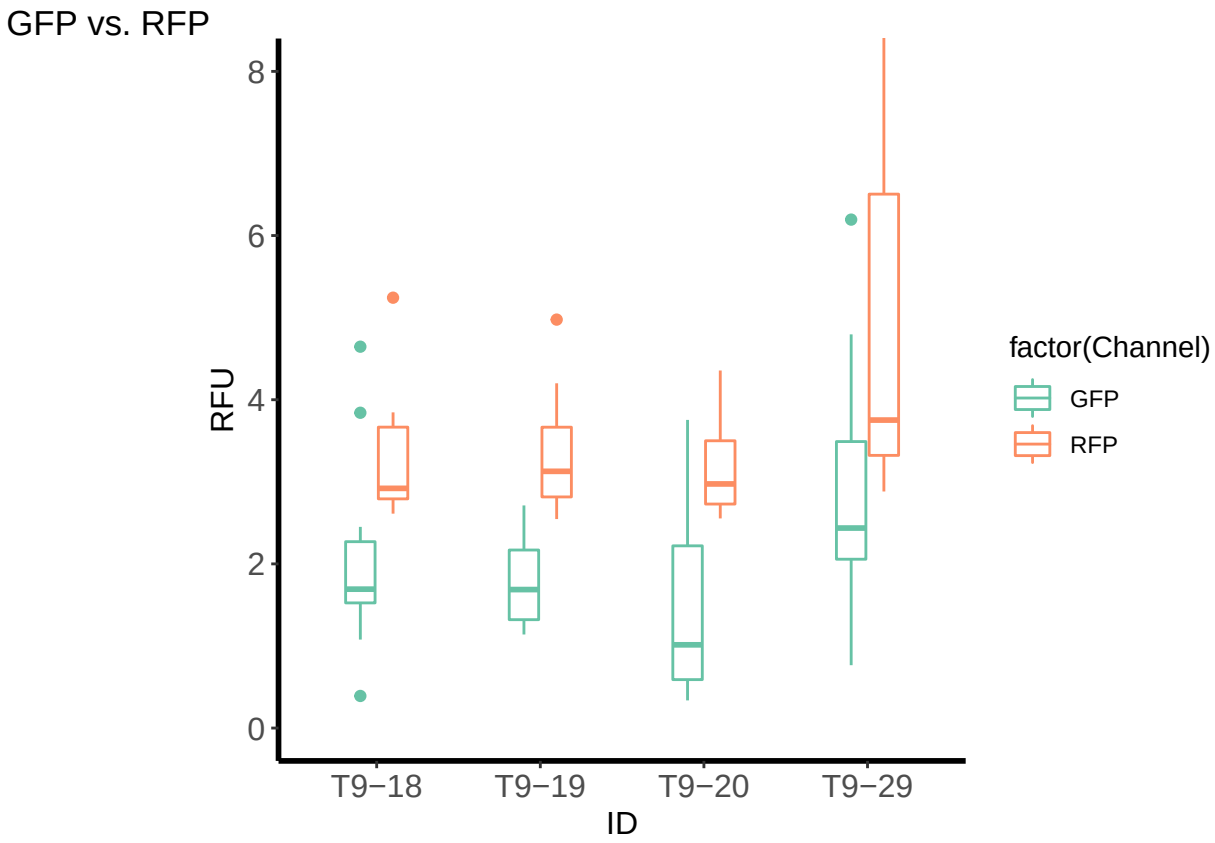
```
group_by(AIsumm3_df_long_red2_mod, ID) %>%
  summarise(
    count = n(),
    mean = mean(RFU, na.rm = TRUE),
    sdev = sd(RFU, na.rm = TRUE)
  )
```

```
## # A tibble: 4 x 4
##   ID    count  mean  sdev
##   <fct> <int> <dbl> <dbl>
## 1 T9-26    19 0.308 0.112
## 2 T9-27    10 0.597 0.111
## 3 T9-28    12 4.57  2.23
## 4 T9-31    14 0.448 0.154
```

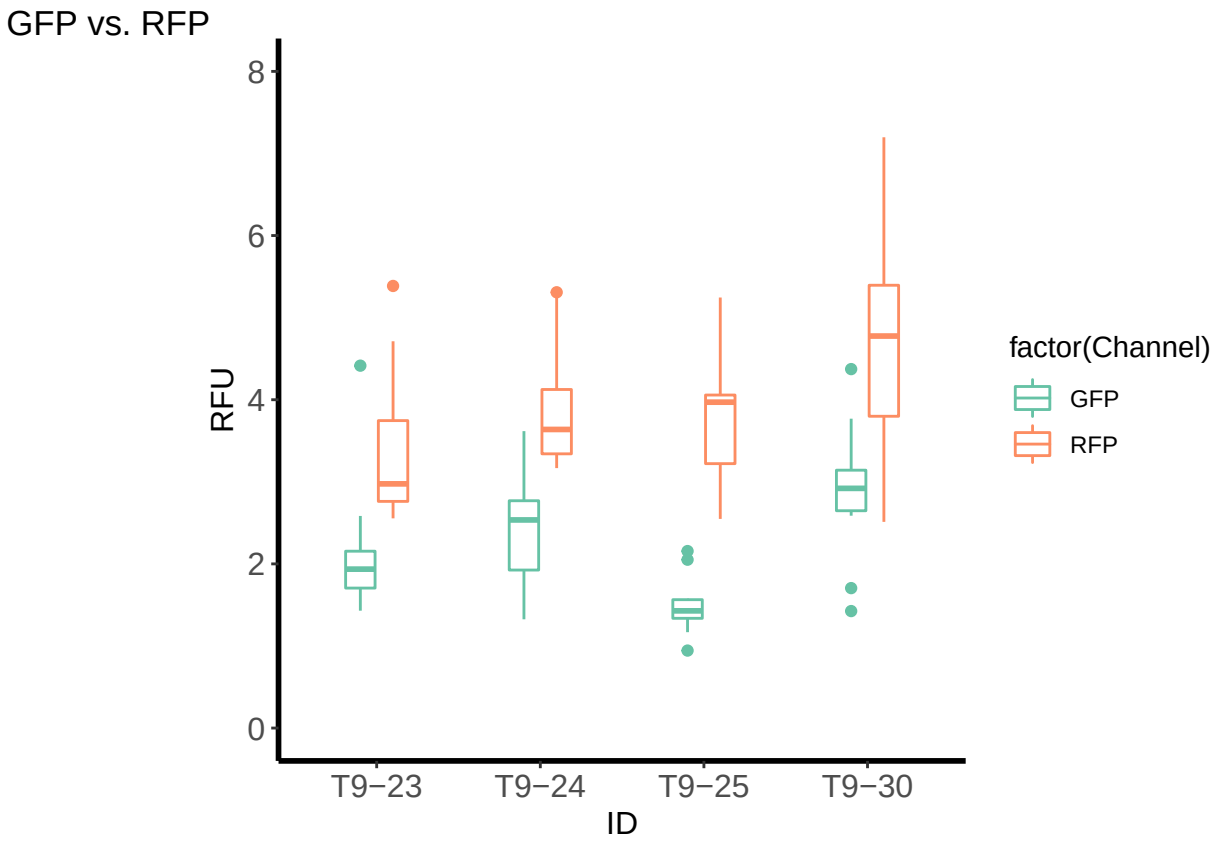
Boxplots GFP vs. RFP

```
p1 <- ggplot(data=AIsumm1_df_long_red, aes(x=ID,y=RFU, colour=factor(Channel))) +
  geom_boxplot(position=position_dodge2(width=0.4), width=0.4) +
  labs(tag = "GFP vs. RFP") +
  scale_color_brewer(palette = "Set2") +
  theme_classic() +
  coord_cartesian(ylim = c(0, 8)) +
  #facet_wrap(~Channel) +
  #coord_fixed(ratio = 0.05, xlim = NULL, ylim = NULL, expand = TRUE, clip = "on") +
  theme(axis.line = element_line(colour = "black", size = 1, linetype = "solid"),
        axis.title.x = element_text(size = 12),
        axis.title.y = element_text(size = 12),
        axis.text.x= element_text(size=12),
        axis.text.y = element_text(size=12)
        )
```

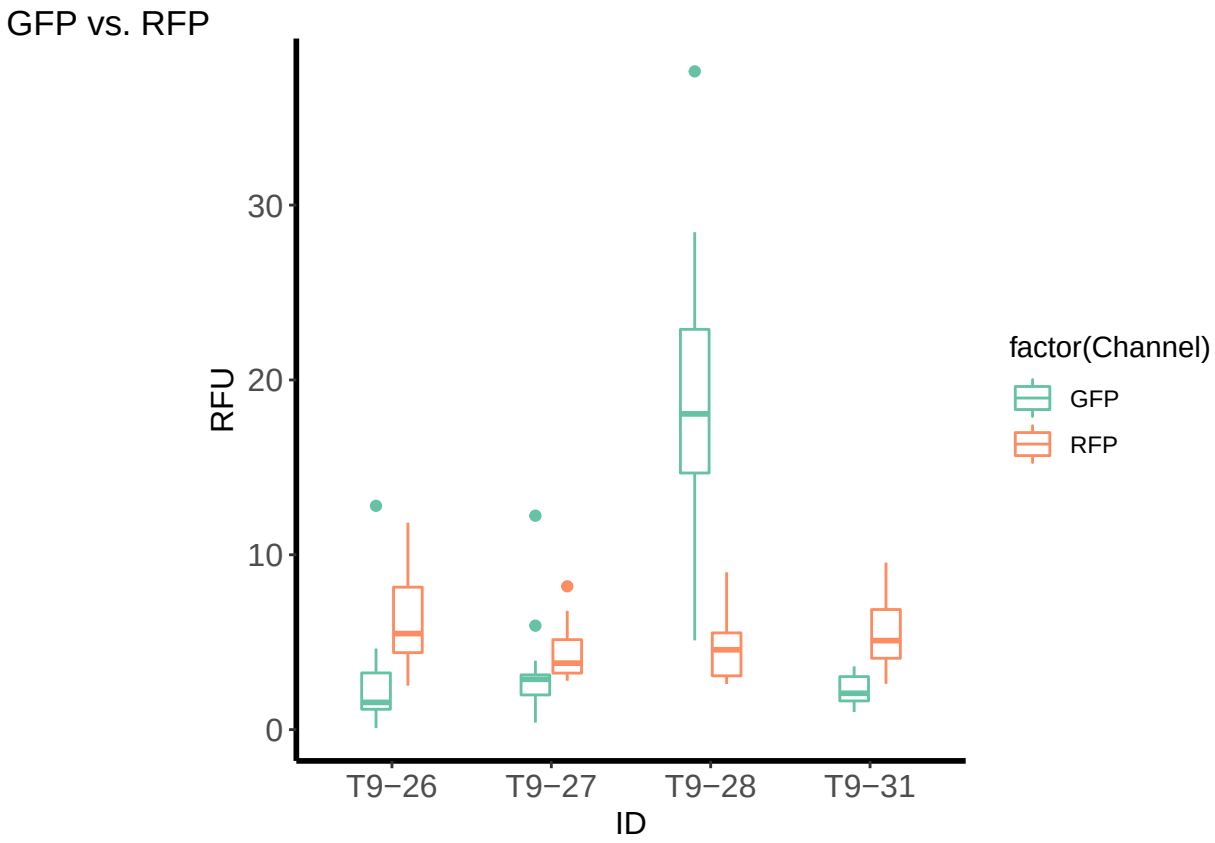
```
p2 <- ggplot(data=AIsumm2_df_long_red, aes(x=ID,y=RFU, colour=factor(Channel))) +
  geom_boxplot(position=position_dodge2(width=0.4), width=0.4) +
  labs(tag = "GFP vs. RFP") +
  scale_color_brewer(palette = "Set2") +
  theme_classic() +
  coord_cartesian(ylim = c(0, 8)) +
  #facet_wrap(~Channel) +
  #coord_fixed(ratio = 0.05, xlim = NULL, ylim = NULL, expand = TRUE, clip = "on") +
  theme(axis.line = element_line(colour = "black", size = 1, linetype = "solid"),
        axis.title.x = element_text(size = 12),
        axis.title.y = element_text(size = 12),
        axis.text.x= element_text(size=12),
        axis.text.y = element_text(size=12)
        )


p3 <- ggplot(data=AIsumm3_df_long_red, aes(x=ID,y=RFU, colour=factor(Channel))) +
  geom_boxplot(position=position_dodge2(width=0.4), width=0.4) +
  labs(tag = "GFP vs. RFP") +
  scale_color_brewer(palette = "Set2") +
  theme_classic() +
  #facet_wrap(~Channel) +
  #coord_fixed(ratio = 0.05, xlim = NULL, ylim = NULL, expand = TRUE, clip = "on") +
  theme(axis.line = element_line(colour = "black", size = 1, linetype = "solid"),
        axis.title.x = element_text(size = 12),
        axis.title.y = element_text(size = 12),
        axis.text.x= element_text(size=12),
        axis.text.y = element_text(size=12)
        )


p3zoom <- ggplot(data=AIsumm3_df_long_red, aes(x=ID,y=RFU, colour=factor(Channel))) +
  geom_boxplot(position=position_dodge2(width=0.4), width=0.4) +
  labs(tag = "GFP vs. RFP") +
  scale_color_brewer(palette = "Set2") +
  theme_classic() +
  coord_cartesian(ylim = c(0, 8)) +
  #facet_wrap(~Channel) +
  #coord_fixed(ratio = 0.05, xlim = NULL, ylim = NULL, expand = TRUE, clip = "on") +
  theme(axis.line = element_line(colour = "black", size = 1, linetype = "solid"),
        axis.title.x = element_text(size = 12),
        axis.title.y = element_text(size = 12),
        axis.text.x= element_text(size=12),
        axis.text.y = element_text(size=12)
        )
p1
```


GFP vs. RFP
RFU
8
6
4
2
0
T9−18
T9−19
T9−20
T9−29
ID
factor(Channel)
GFP
RFP

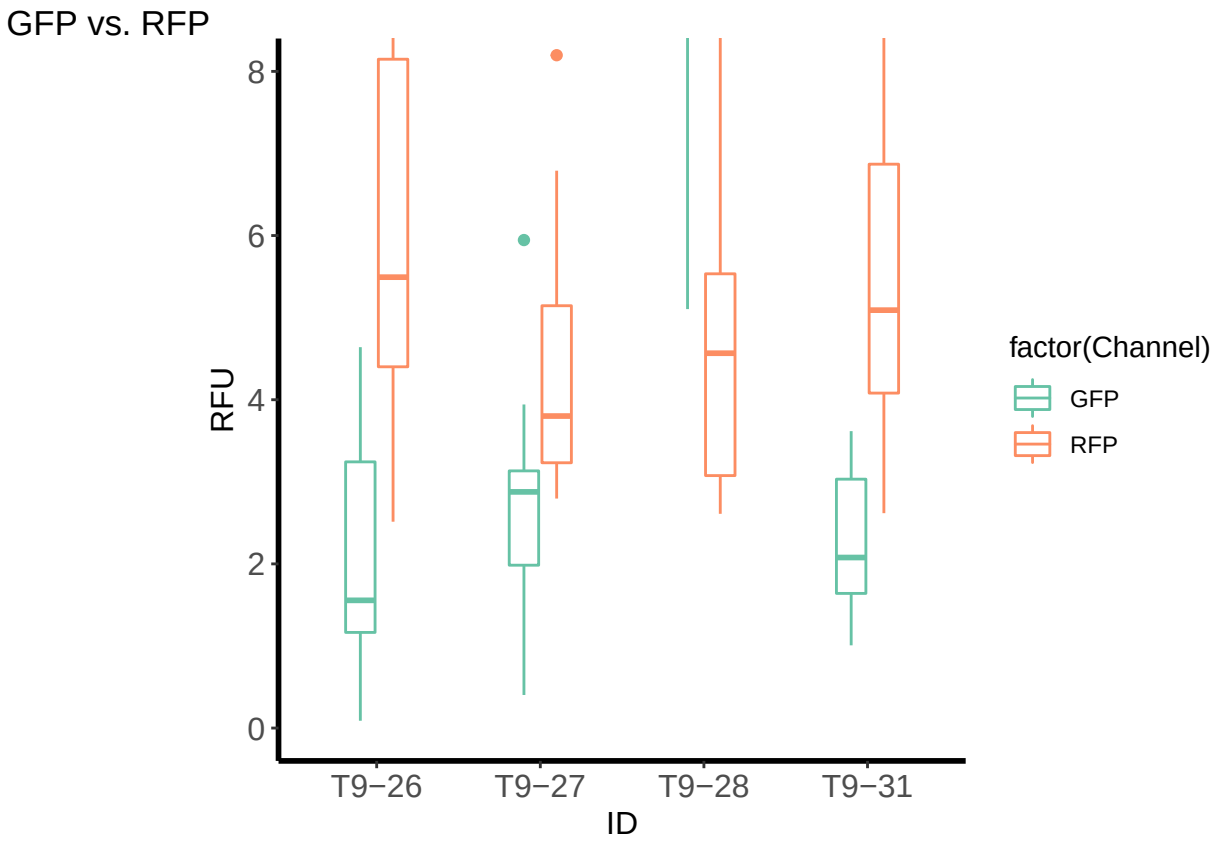
```
p2
```


GFP vs. RFP
8
6
4
2
0
RFU
T9−23
T9−24
T9−25
T9−30
ID
factor(Channel)
GFP
RFP

```
p3
```


GFP vs. RFP
RFU
30
20
10
0
T9−26
T9−27
T9−28
T9−31
ID
factor(Channel)
GFP
RFP

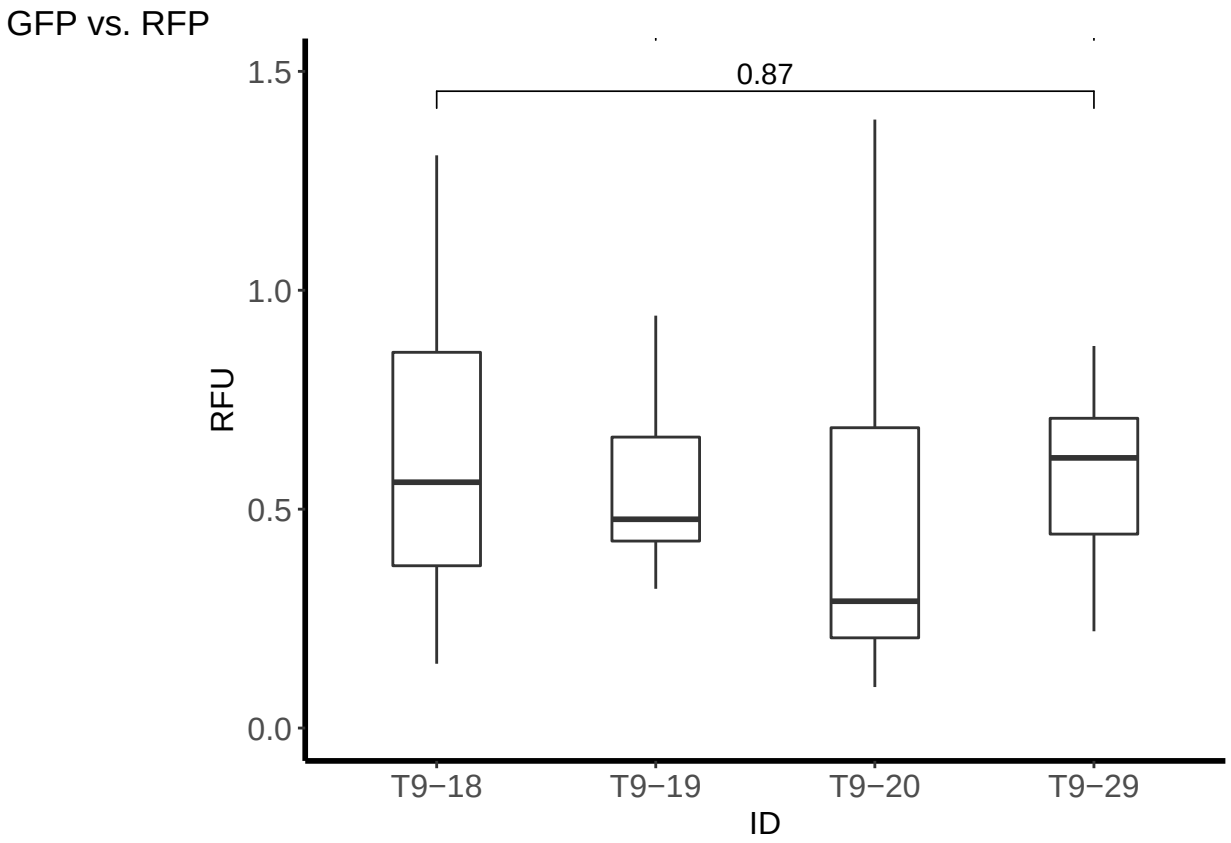
```
p3zoom
```



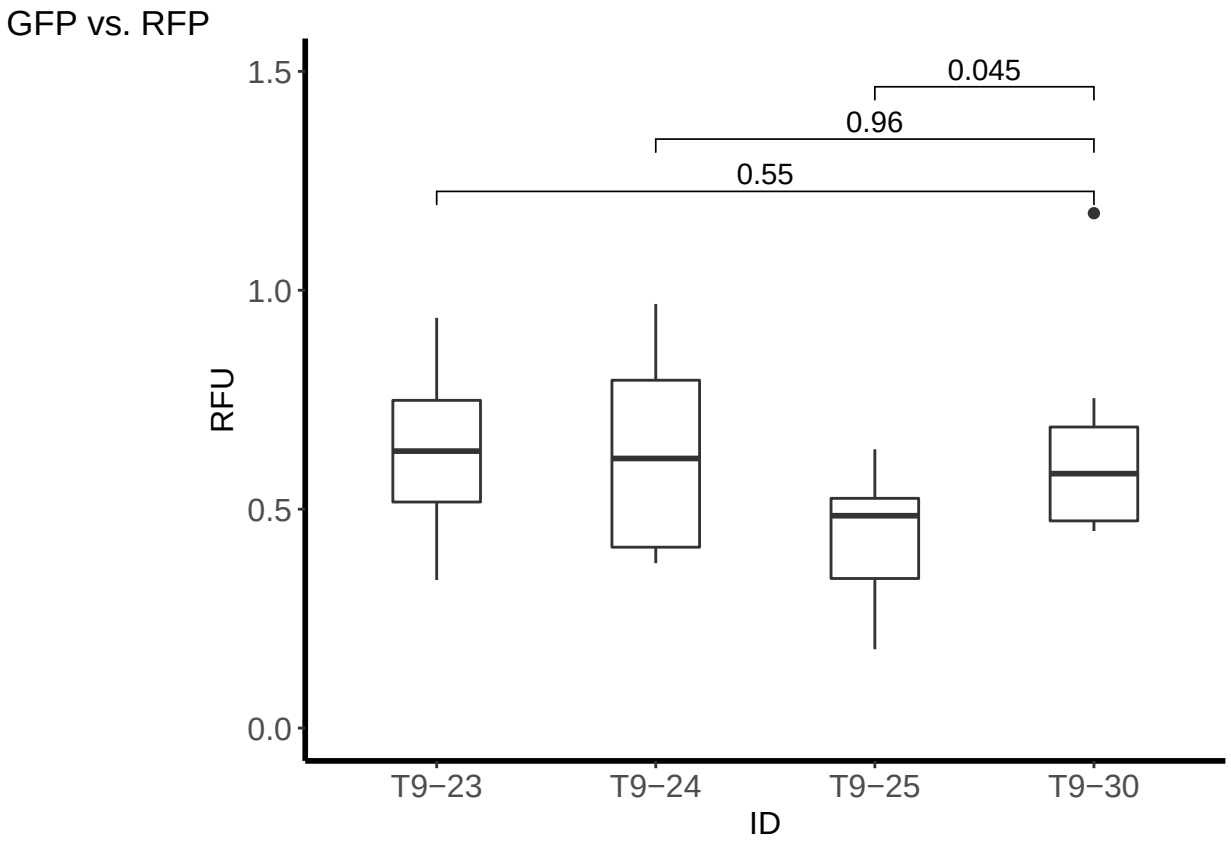
Boxplots ratios:

```
p4 <- ggplot(data=AIsumm1_df_long_red2, aes(x=ID,y=RFU)) +
  geom_boxplot(position=position_dodge2(width=0.4), width=0.4) +
  #geom_dotplot(binaxis='y', stackdir='center', dotsize=0.4, colour='grey', fill='grey') +
  labs(tag = "GFP vs. RFP") +
  scale_color_grey() +
  theme_classic() +
  coord_cartesian(ylim = c(0, 1.5)) +
  #facet_wrap(~Channel) +
  #coord_fixed(ratio = 3.55, xlim = NULL, ylim = NULL, expand = TRUE, clip = "on") +
  theme(axis.line = element_line(colour = "black", size = 1, linetype = "solid"),
        axis.title.x = element_text(size = 12),
        axis.title.y = element_text(size = 12),
        axis.text.x= element_text(size=12),
        axis.text.y = element_text(size=12)
        )

p5 <- ggplot(data=AIsumm2_df_long_red2, aes(x=ID,y=RFU)) +
  geom_boxplot(position=position_dodge2(width=0.4), width=0.4) +
  #geom_dotplot(binaxis='y', stackdir='center', dotsize=0.5, colour='grey', fill='grey') +
  labs(tag = "GFP vs. RFP") +
  scale_color_grey() +
  theme_classic() +
  coord_cartesian(ylim = c(0, 1.5)) +
  #facet_wrap(~Channel) +
  #coord_fixed(ratio = 0.55, xlim = NULL, ylim = NULL, expand = TRUE, clip = "on") +
```

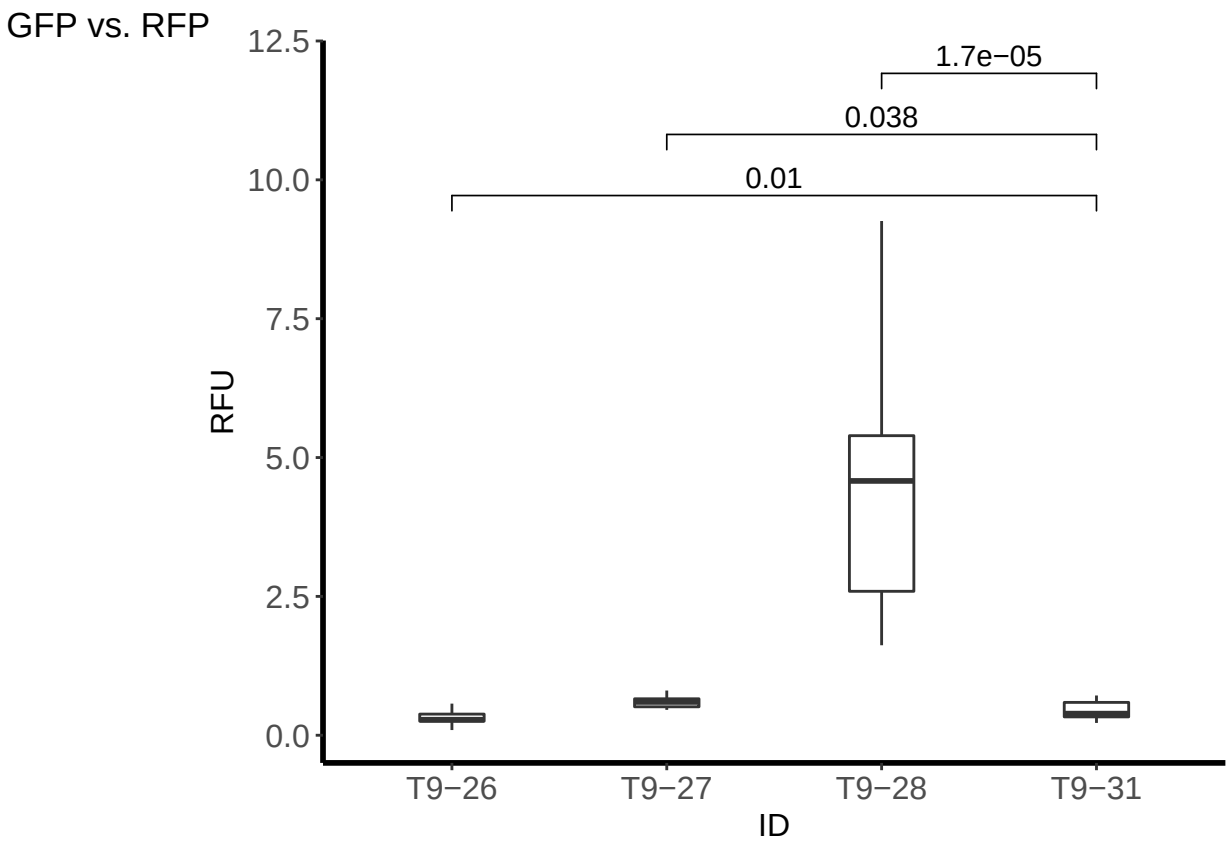
```
  theme(axis.line = element_line(colour = "black", size = 1, linetype = "solid"),
        axis.title.x = element_text(size = 12),
        axis.title.y = element_text(size = 12),
        axis.text.x= element_text(size=12),
        axis.text.y = element_text(size=12)
        )

p6 <- ggplot(data=AIsumm3_df_long_red2_mod, aes(x=ID,y=RFU)) +
  geom_boxplot(position=position_dodge2(width=0.3), width=0.3) +
  #geom_dotplot(binaxis='y', stackdir='center', dotsize=0.5, colour='grey', fill='grey') +
  labs(tag = "GFP vs. RFP") +
  scale_color_grey() +
  theme_classic() +
  #facet_wrap(~Channel) +
  #coord_fixed(ratio = 0.55, xlim = NULL, ylim = NULL, expand = TRUE, clip = "on") +
  theme(axis.line = element_line(colour = "black", size = 1, linetype = "solid"),
        axis.title.x = element_text(size = 12),
        axis.title.y = element_text(size = 12),
        axis.text.x= element_text(size=12),
        axis.text.y = element_text(size=12)
        )



p6zoom <- ggplot(data=AIsumm3_df_long_red2_mod, aes(x=ID,y=RFU)) +
  geom_boxplot(position=position_dodge2(width=0.4), width=0.4) +
  labs(tag = "GFP vs. RFP") +
  scale_color_grey() +
  theme_classic() +
  coord_cartesian(ylim = c(0, 1.5)) +
  #facet_wrap(~Channel) +
  #coord_fixed(ratio = 0.55, xlim = NULL, ylim = NULL, expand = TRUE, clip = "on") +
  theme(axis.line = element_line(colour = "black", size = 1, linetype = "solid"),
        axis.title.x = element_text(size = 12),
        axis.title.y = element_text(size = 12),
        axis.text.x= element_text(size=12),
        axis.text.y = element_text(size=12)
        )


comparisons1 <- list(c("T9-29", "T9-18"), c("T9-29", "T9-19"), c("T9-29", "T9-20"))
comparisons2 <- list(c("T9-30", "T9-23"), c("T9-30", "T9-24"), c("T9-30", "T9-25"))
comparisons3 <- list(c("T9-31", "T9-26"), c("T9-31", "T9-27"), c("T9-31", "T9-28"))

p4p <- p4 + stat_compare_means(method = "wilcox.test", label = "p.format", comparisons = comparisons1)
p5p <- p5 + stat_compare_means(method = "wilcox.test", label = "p.format", comparisons = comparisons2)
p6p <- p6 + stat_compare_means(method = "wilcox.test", label = "p.format", comparisons = comparisons3)

pctrl <- pall_mod + stat_compare_means(method = "wilcox.test", label = "p.format", comparisons = compari


p4p
```



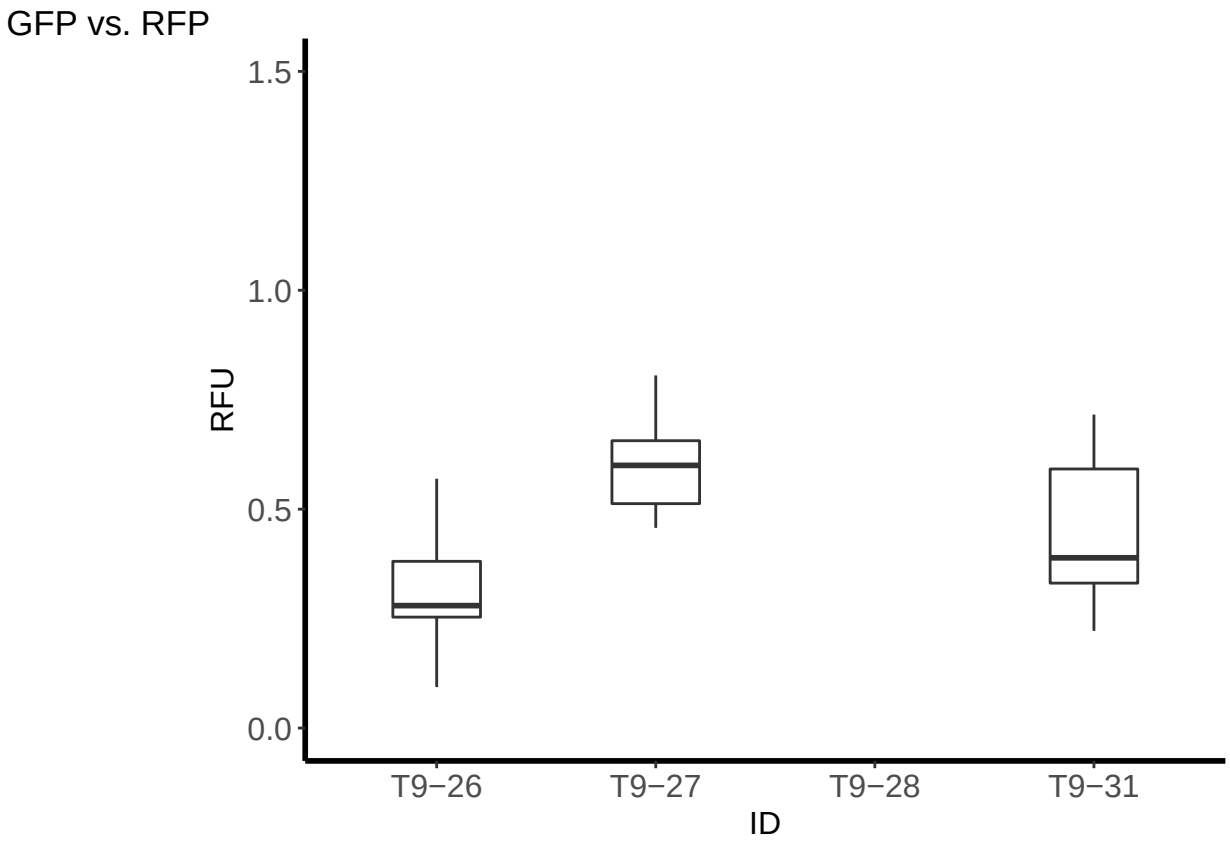

```
p5p
```




```
p6p
```


GFP vs. RFP
12.5
10.0
7.5
5.0
2.5
0.0
RFU
1.7e−05
0.038
0.01
T9−26
T9−27
T9−28
T9−31
ID

```
p6zoom
```



Splitting dataframes:

```
library('dplyr')


###### Only ratios

summary(AIsumm1_df_long_red2)
```

```
##      ID        Channel        RFU
##  T9-18:10   GFP    : 0   Min.   :0.0937
##  T9-19: 8   RFP    : 0   1st Qu.:0.3308
##  T9-20:13   GFP.RFP:43   Median :0.5351
##  T9-29:12                Mean   :0.5546
##                          3rd Qu.:0.7050
##                          Max.   :1.3899
```

```
t918_df <- filter(AIsumm1_df_long_red2, ID == "T9-18")
t919_df <- filter(AIsumm1_df_long_red2, ID == "T9-19")
t920_df <- filter(AIsumm1_df_long_red2, ID == "T9-20")
t929_df <- filter(AIsumm1_df_long_red2, ID == "T9-29")

summary(AIsumm2_df_long_red2)
```

```
##      ID        Channel        RFU
##  T9-23:11   GFP    : 0   Min.   :0.1800
##  T9-24: 8   RFP    : 0   1st Qu.:0.4527
##  T9-25: 9   GFP.RFP:38   Median :0.5303
```

```
##  T9-30:10                Mean   :0.5856
##                          3rd Qu.:0.7143
##                          Max.   :1.1760
```

```
t923_df <- filter(AIsumm2_df_long_red2, ID == "T9-23")
t924_df <- filter(AIsumm2_df_long_red2, ID == "T9-24")
t925_df <- filter(AIsumm2_df_long_red2, ID == "T9-25")
t930_df <- filter(AIsumm2_df_long_red2, ID == "T9-30")

summary(AIsumm3_df_long_red2_mod)
```

```
##      ID        Channel        RFU
##  T9-26:19   GFP    : 0   Min.   :0.0936
##  T9-27:10   RFP    : 0   1st Qu.:0.3165
##  T9-28:12   GFP.RFP:55   Median :0.4730
##  T9-31:14                Mean   :1.3271
##                          3rd Qu.:0.7143
##                          Max.   :9.2588
```

```
t926_df <- filter(AIsumm3_df_long_red2, ID == "T9-26")
t927_df <- filter(AIsumm3_df_long_red2, ID == "T9-27")
t928_df <- filter(AIsumm3_df_long_red2, ID == "T9-28")
t931_df <- filter(AIsumm3_df_long_red2, ID == "T9-31")

#Controls
tctrl_df <- rbind(t929_df, t930_df, t931_df)
summary(tctrl_df)
```

```
##        ID        Channel        RFU
##  T9-31  :14   GFP    : 0   Min.   :0.2211
##  T9-29  :12   RFP    : 0   1st Qu.:0.3926
##  T9-30  :10   GFP.RFP:36   Median :0.5423
##  T9-18  : 0                Mean   :0.5428
##  T9-19  : 0                3rd Qu.:0.6532
##  T9-20  : 0                Max.   :1.1760
##  (Other): 0
```

Summary stats within

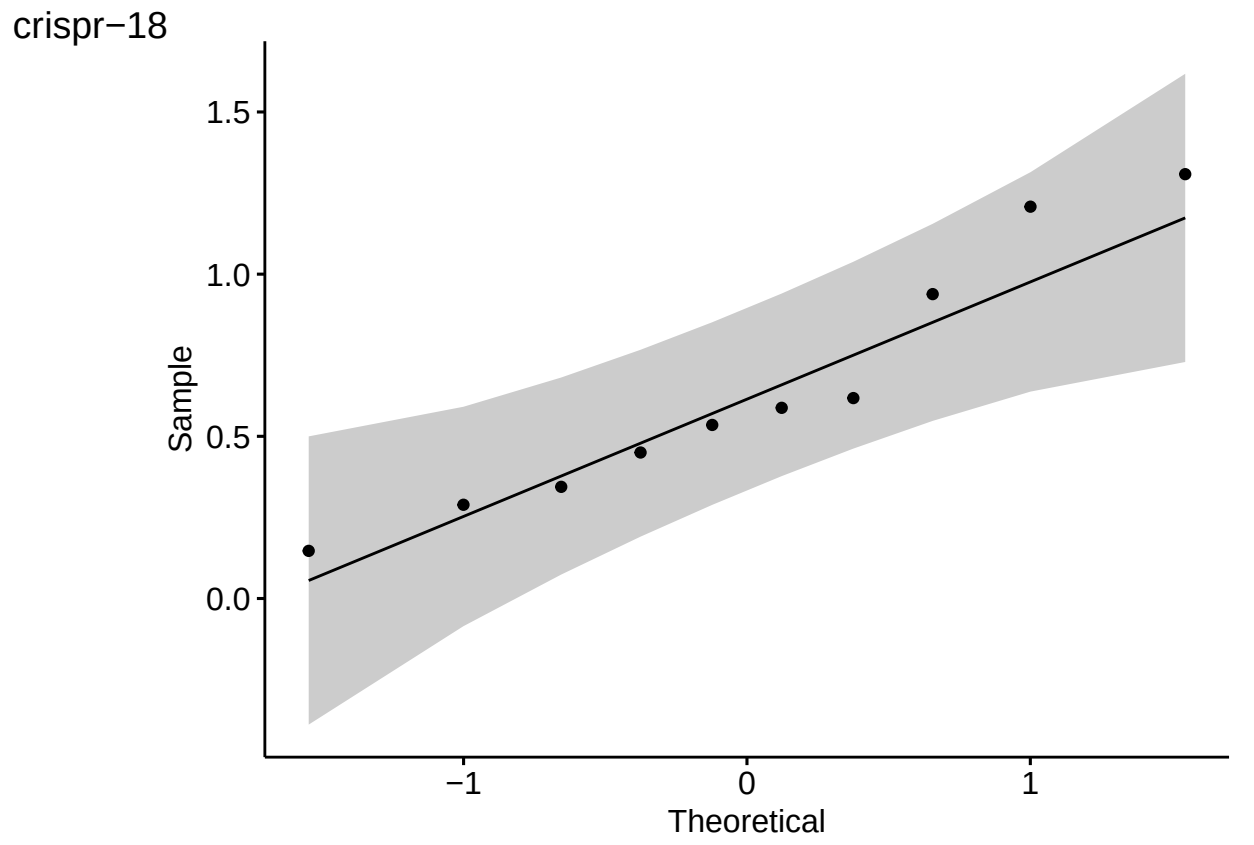
```
summary(t918_df)
```

```
##      ID        Channel        RFU
##  T9-18:10   GFP    : 0   Min.   :0.1469
##  T9-19: 0   RFP    : 0   1st Qu.:0.3708
##  T9-20: 0   GFP.RFP:10   Median :0.5615
##  T9-29: 0                Mean   :0.6426
##                          3rd Qu.:0.8583
##                          Max.   :1.3082
```

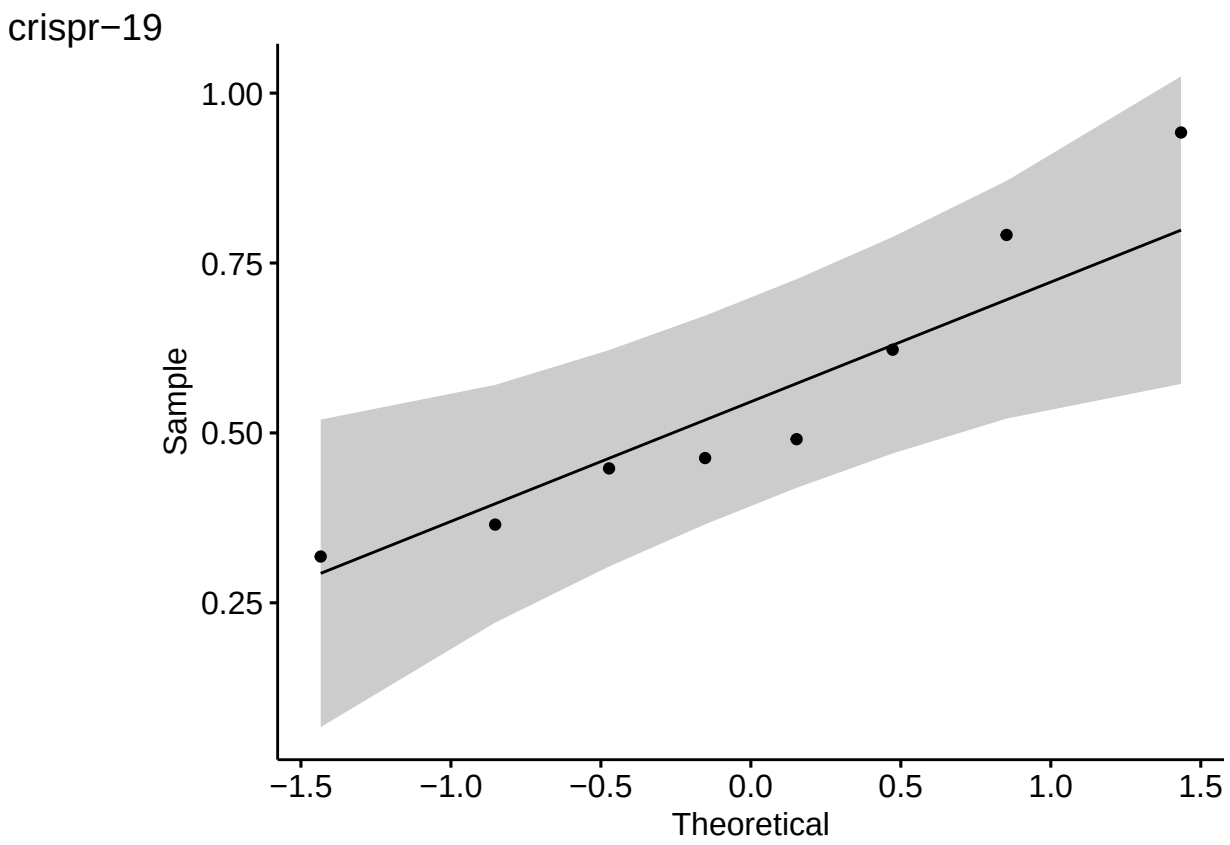
Statistic analysis:

```
### Normality tests:
qq918 <- ggqqplot(data=t918_df$RFU) + labs(tag = "crispr-18")
qq919 <- ggqqplot(data=t919_df$RFU) + labs(tag = "crispr-19")
qq920 <- ggqqplot(data=t920_df$RFU) + labs(tag = "crispr-20")
qq929 <- ggqqplot(data=t929_df$RFU) + labs(tag = "crispr-ctrl")

qq923 <- ggqqplot(data=t923_df$RFU) + labs(tag = "crisprI-23")
```

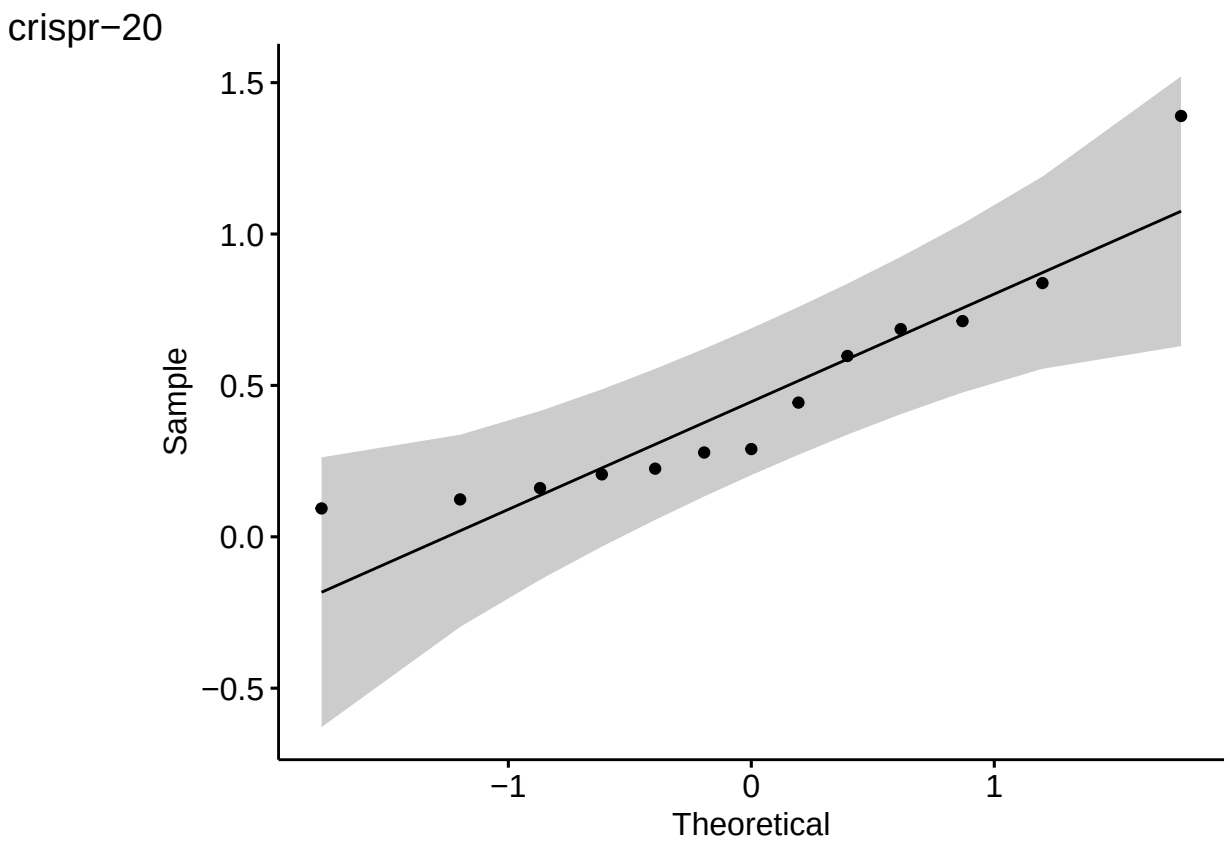
```
qq924 <- ggqqplot(data=t924_df$RFU) + labs(tag = "crisprI-24")
qq925 <- ggqqplot(data=t925_df$RFU) + labs(tag = "crisprI-25")
qq930 <- ggqqplot(data=t930_df$RFU) + labs(tag = "crisprI-ctrl")

qq926 <- ggqqplot(data=t926_df$RFU) + labs(tag = "crisprA-26")
qq927 <- ggqqplot(data=t927_df$RFU) + labs(tag = "crisprA-27")
qq928 <- ggqqplot(data=t928_df$RFU) + labs(tag = "crisprA-28")
qq931 <- ggqqplot(data=t931_df$RFU) + labs(tag = "crisprA-ctrl")

qq918
```



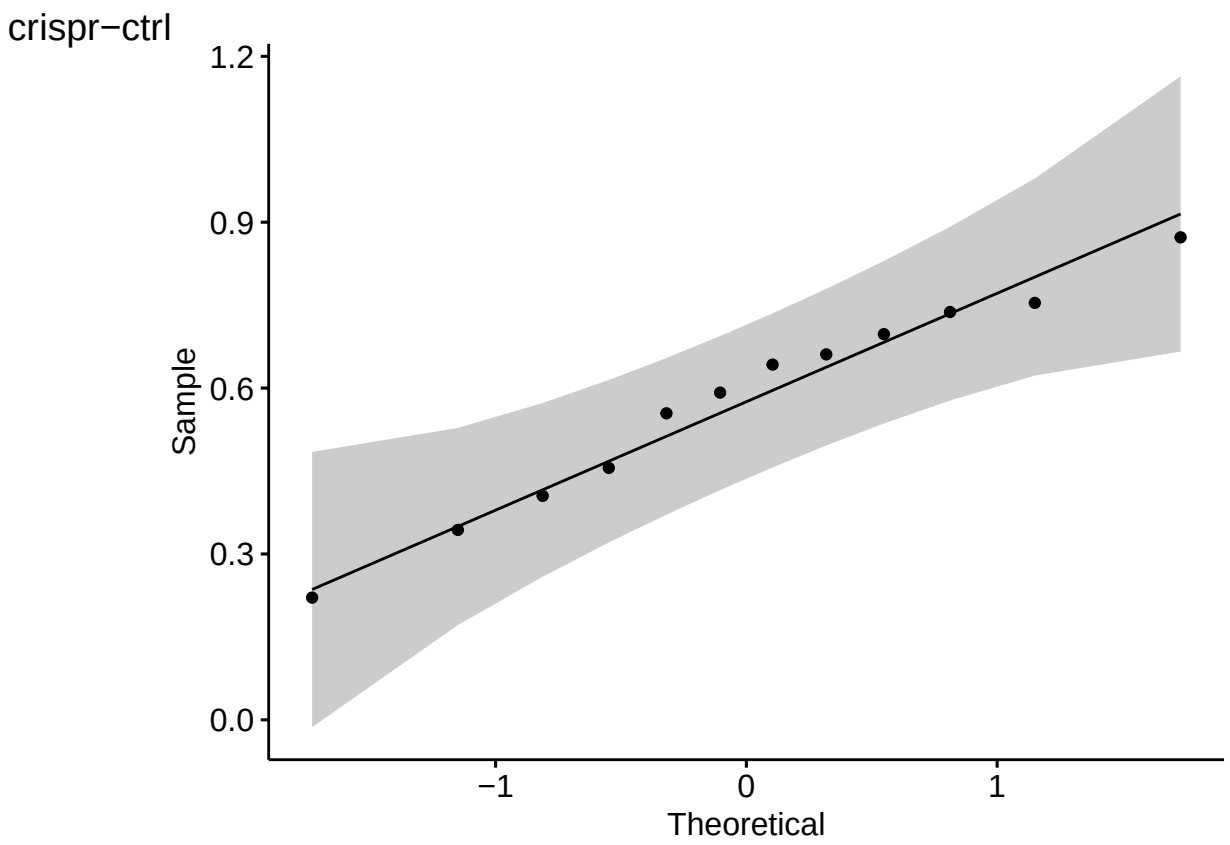

```
qq919
```


crispr−19
1.00
0.75
0.50
0.25
Sample
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
Theoretical

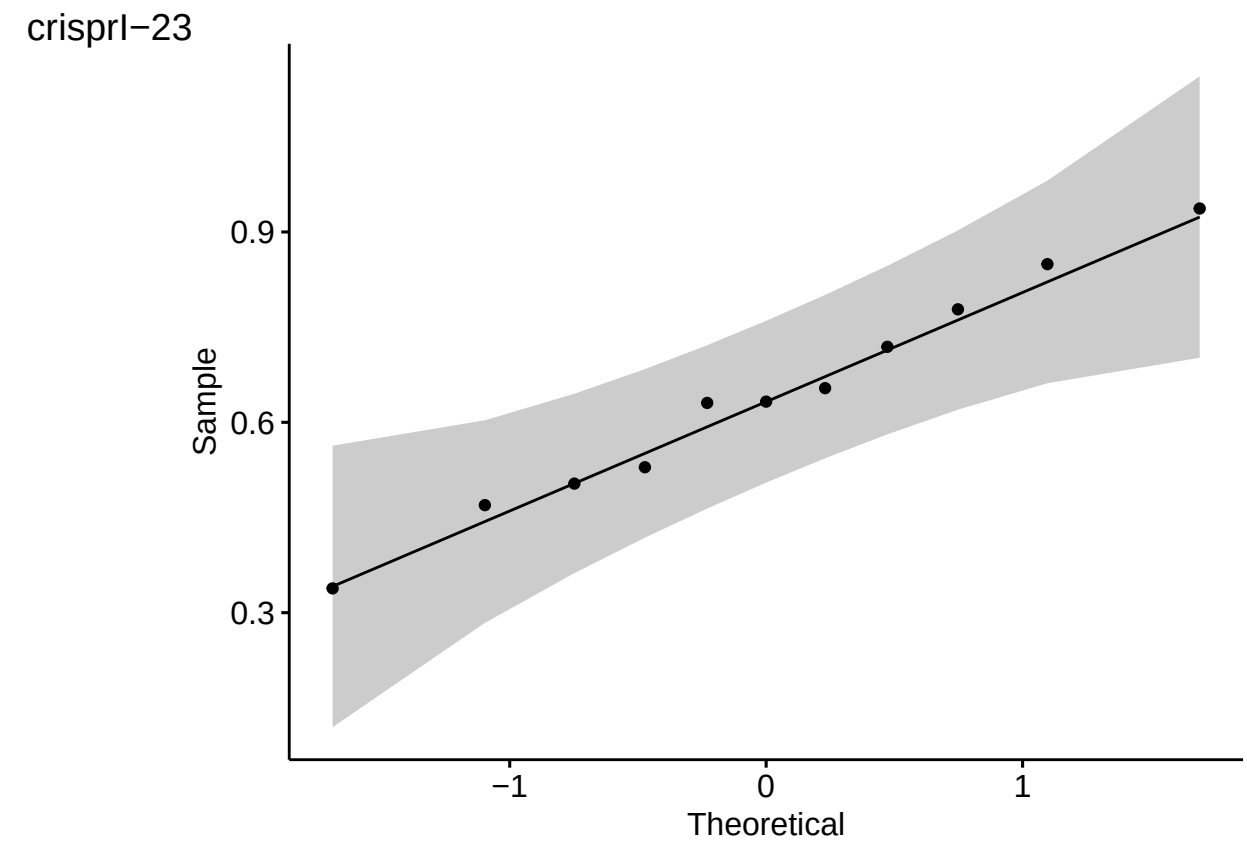
```
qq920
```



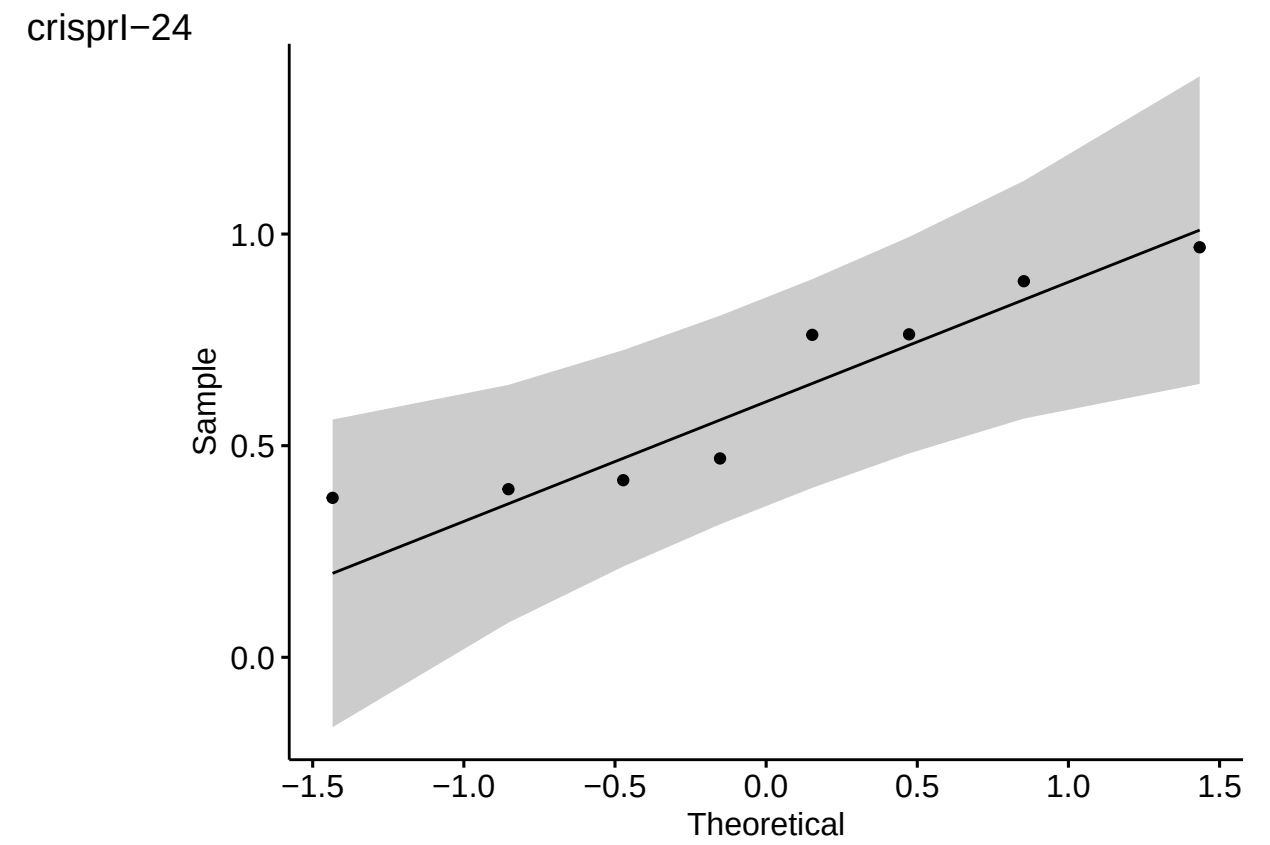

```
qq929
```



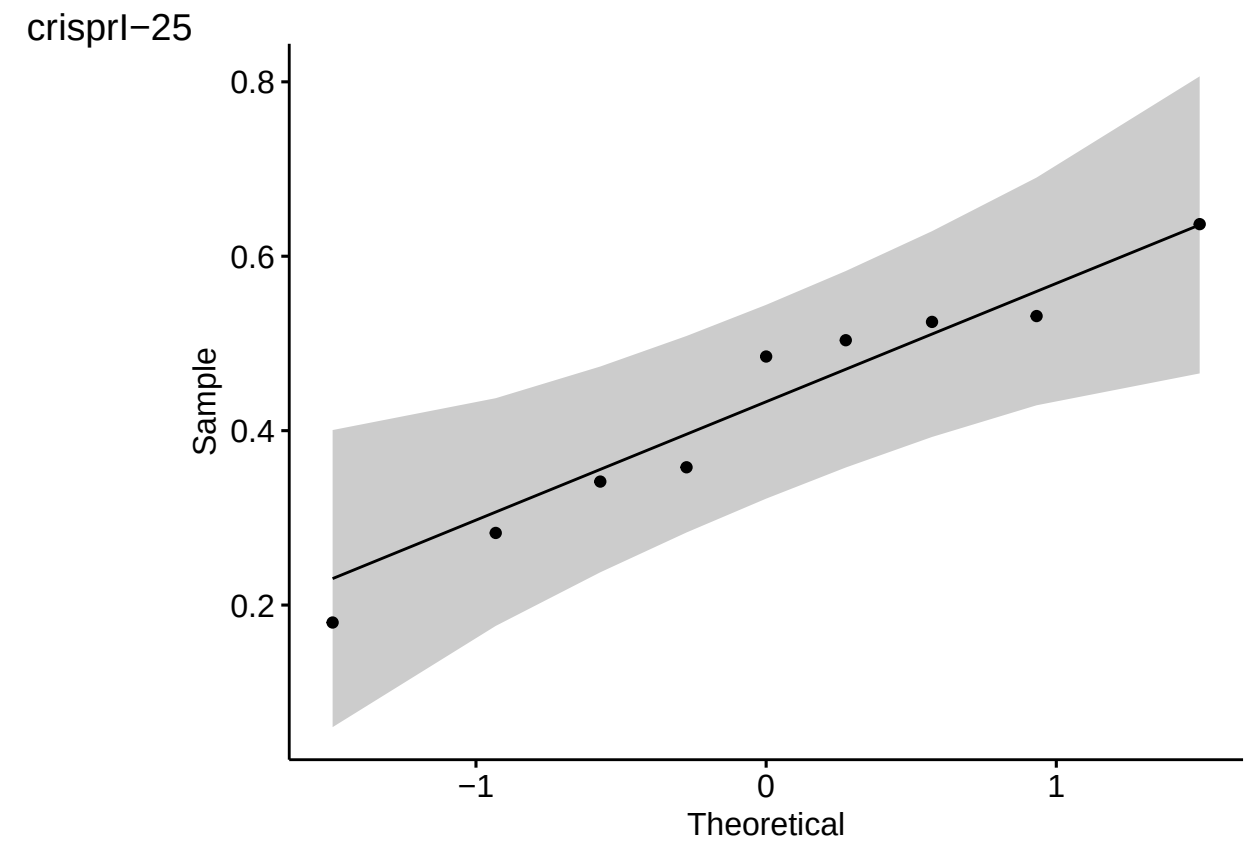

```
qq923
```



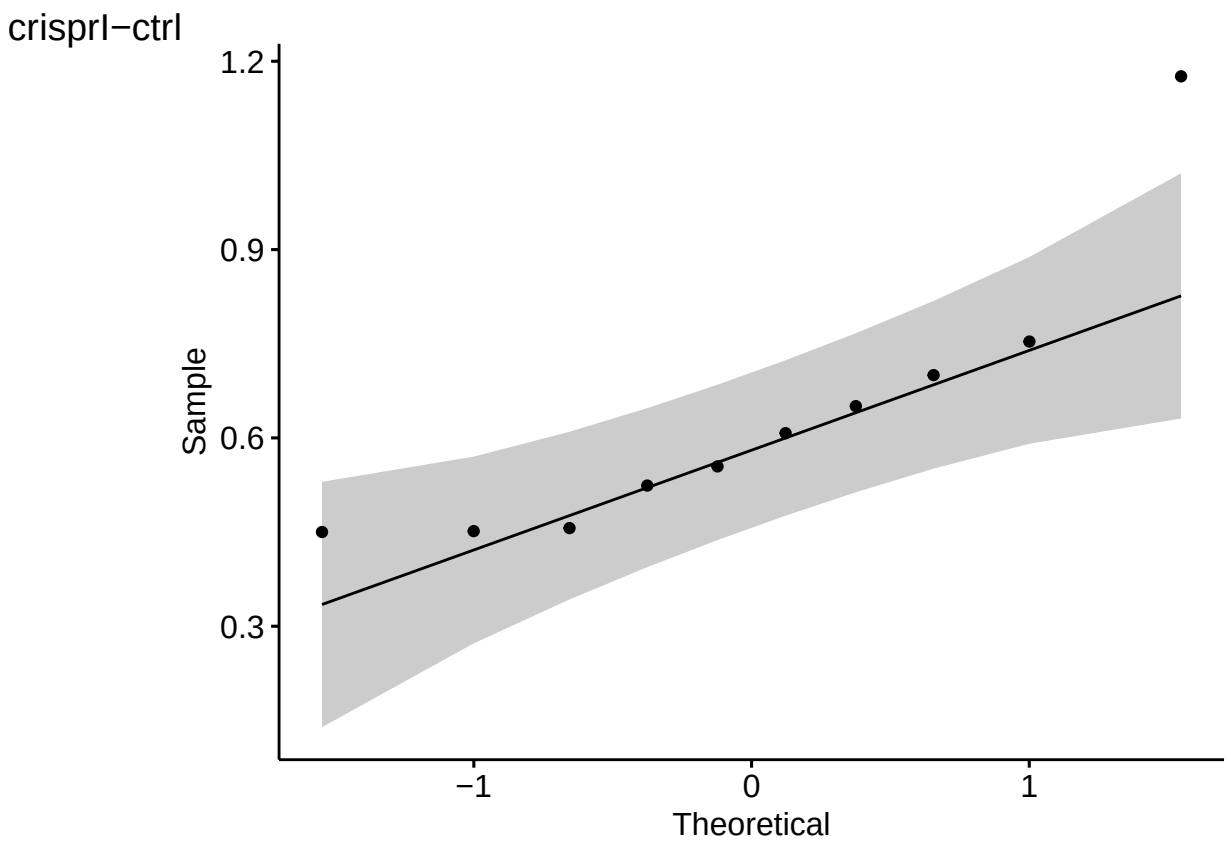

```
qq924
```



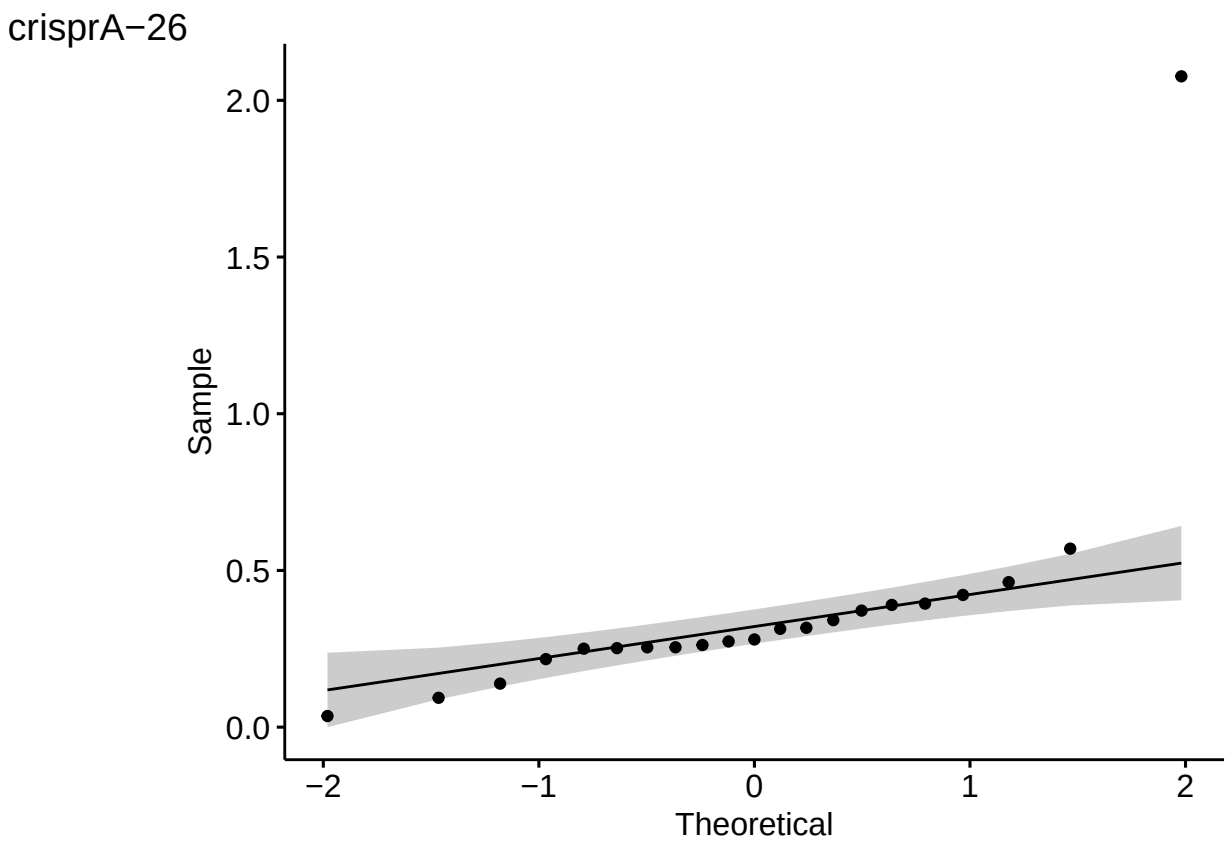

```
qq925
```



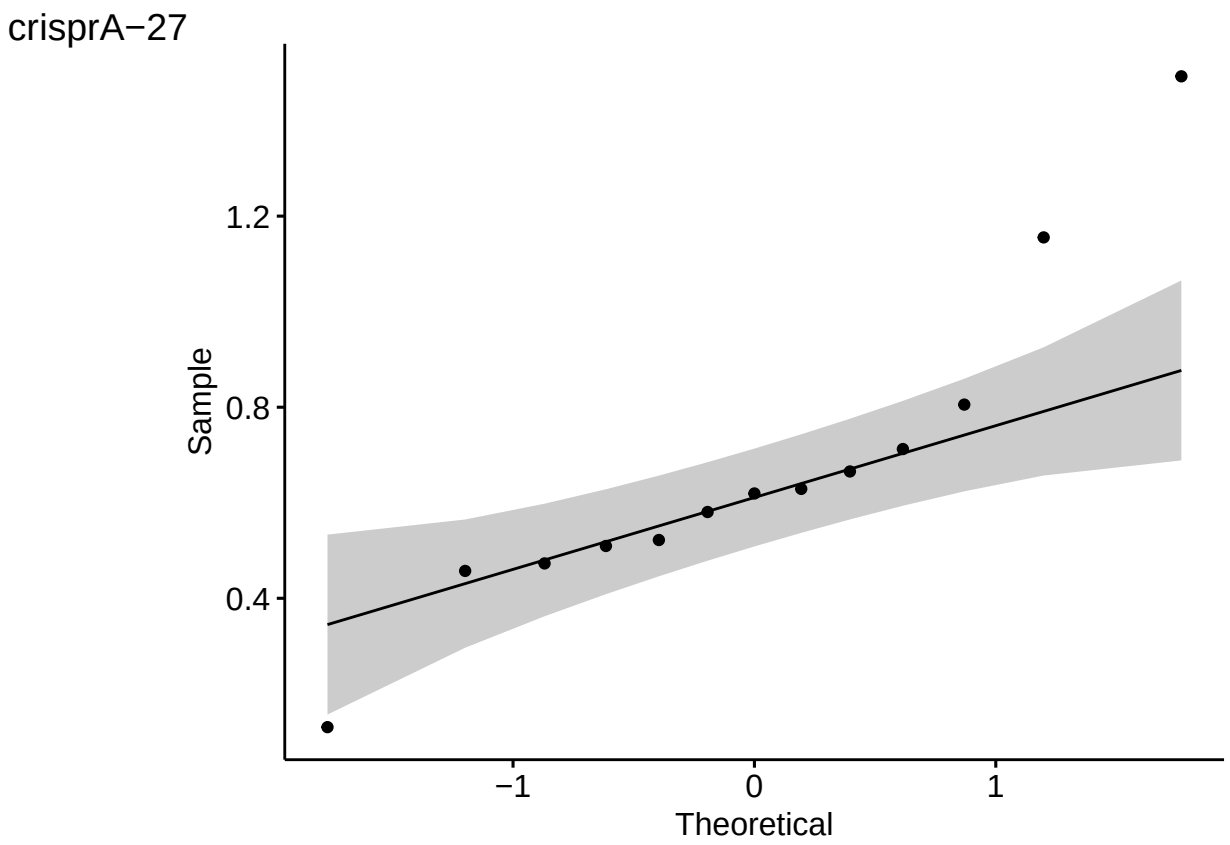

```
qq930
```



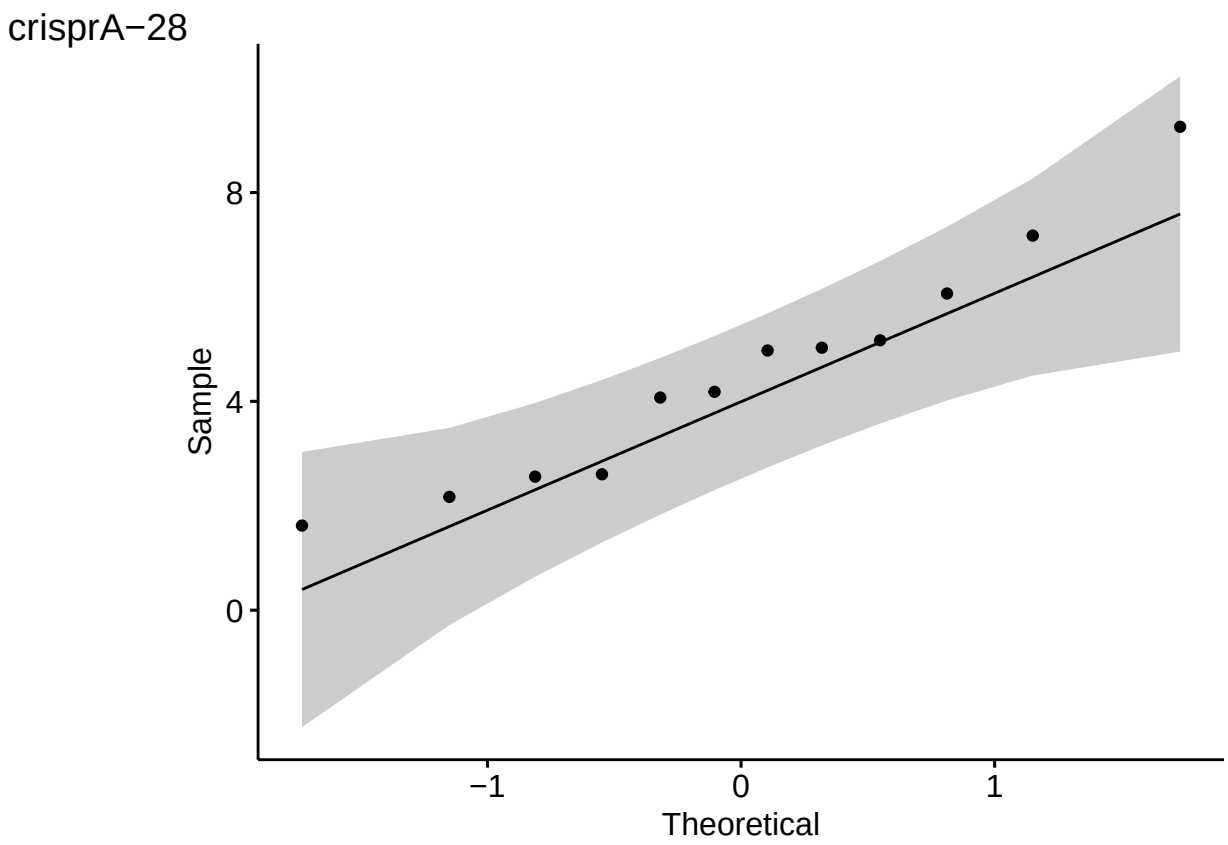

```
qq926
```



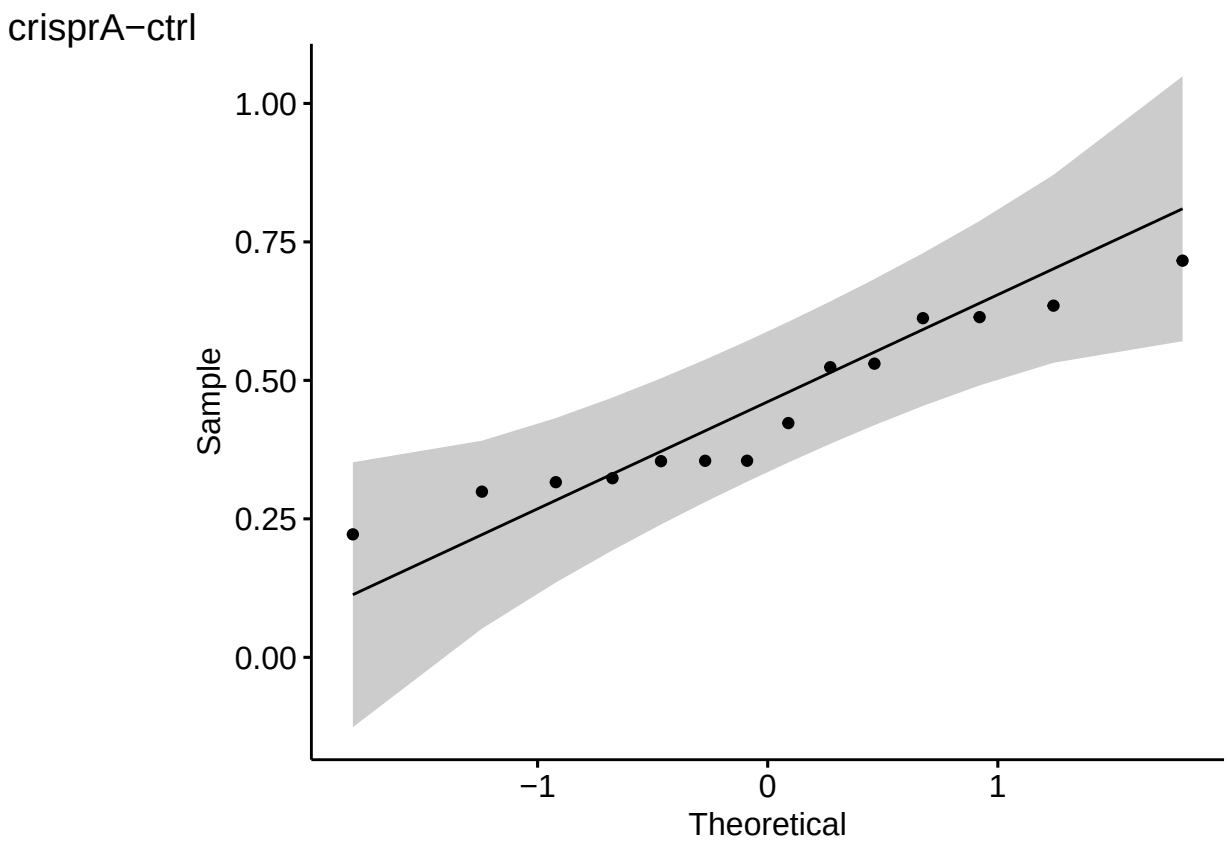

```
qq927
```




```
qq928
```




```
qq931
```




```
#  follow for the most part noramlity criteria

## Shapiro-wilk test

shapiro.test(t918_df$RFU) # p-value = 0.346
```

```
##
##  Shapiro-Wilk normality test
##
## data:  t918_df$RFU
## W = 0.91867, p-value = 0.346
```

```
shapiro.test(t919_df$RFU) # p-value = 0.3411
```

```
##
##  Shapiro-Wilk normality test
##
## data:  t919_df$RFU
## W = 0.90814, p-value = 0.3411
```

```
shapiro.test(t920_df$RFU) # p-value = 0.04243     ..tests not normal
```

```
##
##  Shapiro-Wilk normality test
##
## data:  t920_df$RFU
## W = 0.86318, p-value = 0.04243
```

```r
shapiro.test(t929_df$RFU) # p-value = 0.9347
```

```
##
##  Shapiro-Wilk normality test
##
## data:  t929_df$RFU
## W = 0.97245, p-value = 0.9347
```

```r
shapiro.test(t923_df$RFU) # p-value = 0.9936
```

```
##
##  Shapiro-Wilk normality test
##
## data:  t923_df$RFU
## W = 0.98741, p-value = 0.9936
```

```r
shapiro.test(t924_df$RFU) # p-value = 0.1198
```

```
##
##  Shapiro-Wilk normality test
##
## data:  t924_df$RFU
## W = 0.85991, p-value = 0.1198
```

```r
shapiro.test(t925_df$RFU) # p-value = 0.7255
```

```
##
##  Shapiro-Wilk normality test
##
## data:  t925_df$RFU
## W = 0.95323, p-value = 0.7255
```

```r
shapiro.test(t930_df$RFU) # p-value = 0.01332     ..tests not normal
```

```
##
##  Shapiro-Wilk normality test
##
## data:  t930_df$RFU
## W = 0.79698, p-value = 0.01332
```

```r
shapiro.test(t926_df$RFU) # p-value = 2.379e-07   ..tests not normal
```

```
##
##  Shapiro-Wilk normality test
##
## data:  t926_df$RFU
## W = 0.50735, p-value = 2.379e-07
```

```r
shapiro.test(t927_df$RFU) # p-value = 0.06551     ..tests not normal
```

```
##
##  Shapiro-Wilk normality test
##
## data:  t927_df$RFU
## W = 0.87728, p-value = 0.06551
```

```r
shapiro.test(t928_df$RFU) # p-value = 0.6075
```

```
##
```

```
##  Shapiro-Wilk normality test
##
## data:  t928_df$RFU
## W = 0.94797, p-value = 0.6075
```

```
shapiro.test(t931_df$RFU) # p-value = 0.237
```

```
##
##  Shapiro-Wilk normality test
##
## data:  t931_df$RFU
## W = 0.92227, p-value = 0.237
```

```
# Calculating and adding adjusted p-values to the boxplots
#Source: https://github.com/kassambara/ggpubr/issues/65#issuecomment-407211245
#stat_pvalue_manual function: https://www.rdocumentation.org/packages/ggpubr/versions/0.2/topics/stat_p

# Pairwise wilcoxon rank sum test between groups
stat.test1 <- AIsumm1_df_long_red2 %>%
  group_by("ID") %>%
  wilcox_test(RFU ~ ID) %>%
  adjust_pvalue(method = 'holm') %>%
  mutate(y.position = 2)#, p = signif(p, digits = 4),
  #p.adj = signif(p.adj, digits = 6))

#adding adjusted p-value to boxplot
p4adjust <- p4p + stat_pvalue_manual(stat.test1, label = "p = {p.adj}")
```

```
## Warning: Ignoring unknown aesthetics: xmin, xmax, annotations, y_position
```

```
stat.test2 <- AIsumm2_df_long_red2 %>%
  group_by("ID") %>%
  wilcox_test(RFU ~ ID) %>%
  adjust_pvalue(method = 'holm') %>%
  mutate(y.position = 2)#, p = signif(p, digits = 4),
  #p.adj = signif(p.adj, digits = 6))

p5adjust <- p5p + stat_pvalue_manual(stat.test2, label = "p = {p.adj}")
```

```
## Warning: Ignoring unknown aesthetics: xmin, xmax, annotations, y_position
```

```
stat.test3 <- AIsumm3_df_long_red2_mod %>%
  group_by("ID") %>%
  wilcox_test(RFU ~ ID) %>%
  adjust_pvalue(method = 'holm') %>%
  mutate(y.position = 7.5)#, p = signif(p, digits = 4),
  #p.adj = signif(p.adj, digits = 6))

p6adjust <- p6p + stat_pvalue_manual(stat.test3, label = "p = {p.adj}")
```

```
## Warning: Ignoring unknown aesthetics: xmin, xmax, annotations, y_position
```

```
# comparing means of controls 9-29, 9-30, 9-31
stat.testCtrl <- tctrl_df %>%
  group_by("ID") %>%
  kruskal_test(RFU ~ ID) %>%
  #adjust_pvalue(method = 'holm') %>%
```

```
  mutate(y.position = 7.5)#, p = signif(p, digits = 4),
  #p.adj = signif(p.adj, digits = 6))


# Output: comparing means within CRISPR-groups (wilcoxon)
stat.test1
```

```
## # A tibble: 6 x 9
##   `"ID"` .y.   group1 group2 statistic     p p.adj p.adj.signif y.position
##   <chr>  <chr> <chr>  <chr>      <dbl> <dbl> <dbl> <chr>              <dbl>
## 1 ID     RFU   T9-18  T9-19         42 0.897     1 ns                     2
## 2 ID     RFU   T9-18  T9-20         84 0.257     1 ns                     2
## 3 ID     RFU   T9-18  T9-29         57 0.872     1 ns                     2
## 4 ID     RFU   T9-19  T9-20         70 0.21      1 ns                     2
## 5 ID     RFU   T9-19  T9-29         43 0.734     1 ns                     2
## 6 ID     RFU   T9-20  T9-29         52 0.168     1 ns                     2
```

```
stat.test2
```

```
## # A tibble: 6 x 9
##   `"ID"` .y.   group1 group2 statistic      p p.adj p.adj.signif y.position
##   <chr>  <chr> <chr>  <chr>      <dbl>  <dbl> <dbl> <chr>              <dbl>
## 1 ID     RFU   T9-23  T9-24         46 0.904  1     ns                     2
## 2 ID     RFU   T9-23  T9-25         79 0.0251 0.151 ns                     2
## 3 ID     RFU   T9-23  T9-30         64 0.557  1     ns                     2
## 4 ID     RFU   T9-24  T9-25         52 0.139  0.556 ns                     2
## 5 ID     RFU   T9-24  T9-30         39 0.965  1     ns                     2
## 6 ID     RFU   T9-25  T9-30         20 0.0435 0.217 ns                     2
```

```
stat.test3
```

```
## # A tibble: 6 x 9
##   `"ID"` .y.   group1 group2 statistic       p   p.adj p.adj.signif
##   <chr>  <chr> <chr>  <chr>      <dbl>   <dbl>   <dbl> <chr>
## 1 ID     RFU   T9-26  T9-27          5 1.90e-6 7.60e-6 ****
## 2 ID     RFU   T9-26  T9-28          0 1.42e-8 8.52e-8 ****
## 3 ID     RFU   T9-26  T9-31         62 8.85e-3 1.77e-2 *
## 4 ID     RFU   T9-27  T9-28          0 3.09e-6 9.27e-6 ****
## 5 ID     RFU   T9-27  T9-31        106 3.58e-2 3.58e-2 *
## 6 ID     RFU   T9-28  T9-31        168 2.07e-7 1.03e-6 ****
## # ... with 1 more variable: y.position <dbl>
```

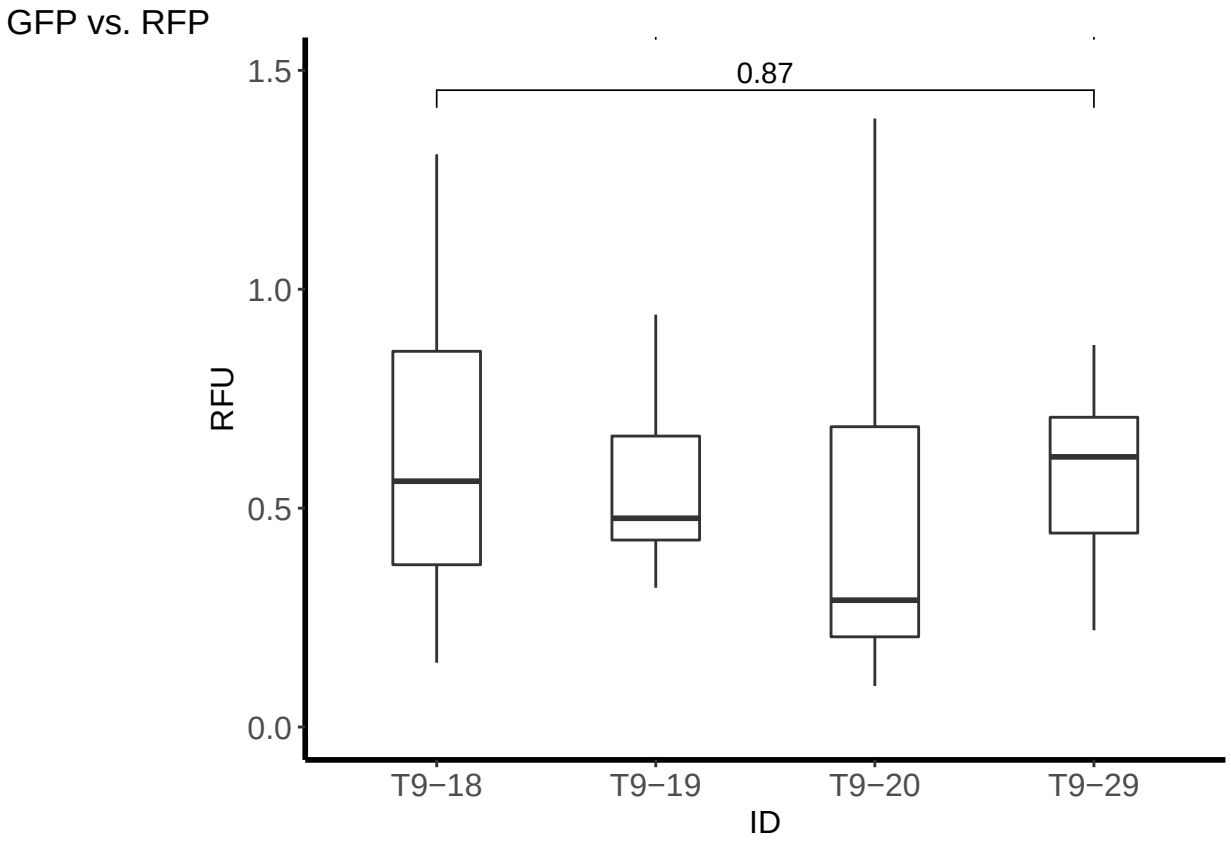
```
# Output: comparing means of controls 9-29, 9-30, 9-31
stat.testCtrl
```

```
## # A tibble: 1 x 7
##   `"ID"` .y.   statistic    df      p method         y.position
##   <chr>  <chr>     <dbl> <int>  <dbl> <chr>               <dbl>
## 1 ID     RFU        5.69     2 0.0582 Kruskal-Wallis        7.5
```

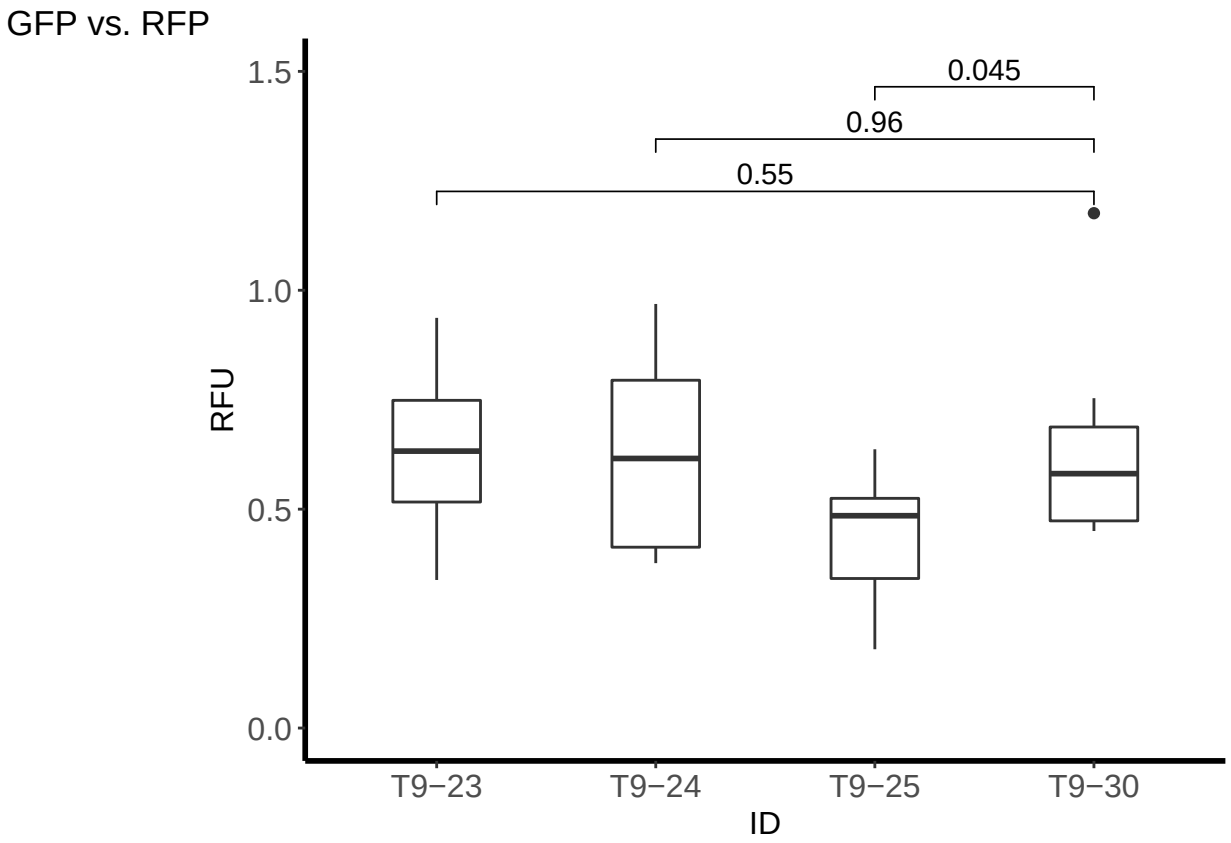
```
kruskal.test(RFU ~ ID, data = tctrl_df)
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  RFU by ID
## Kruskal-Wallis chi-squared = 5.6882, df = 2, p-value = 0.05819
```

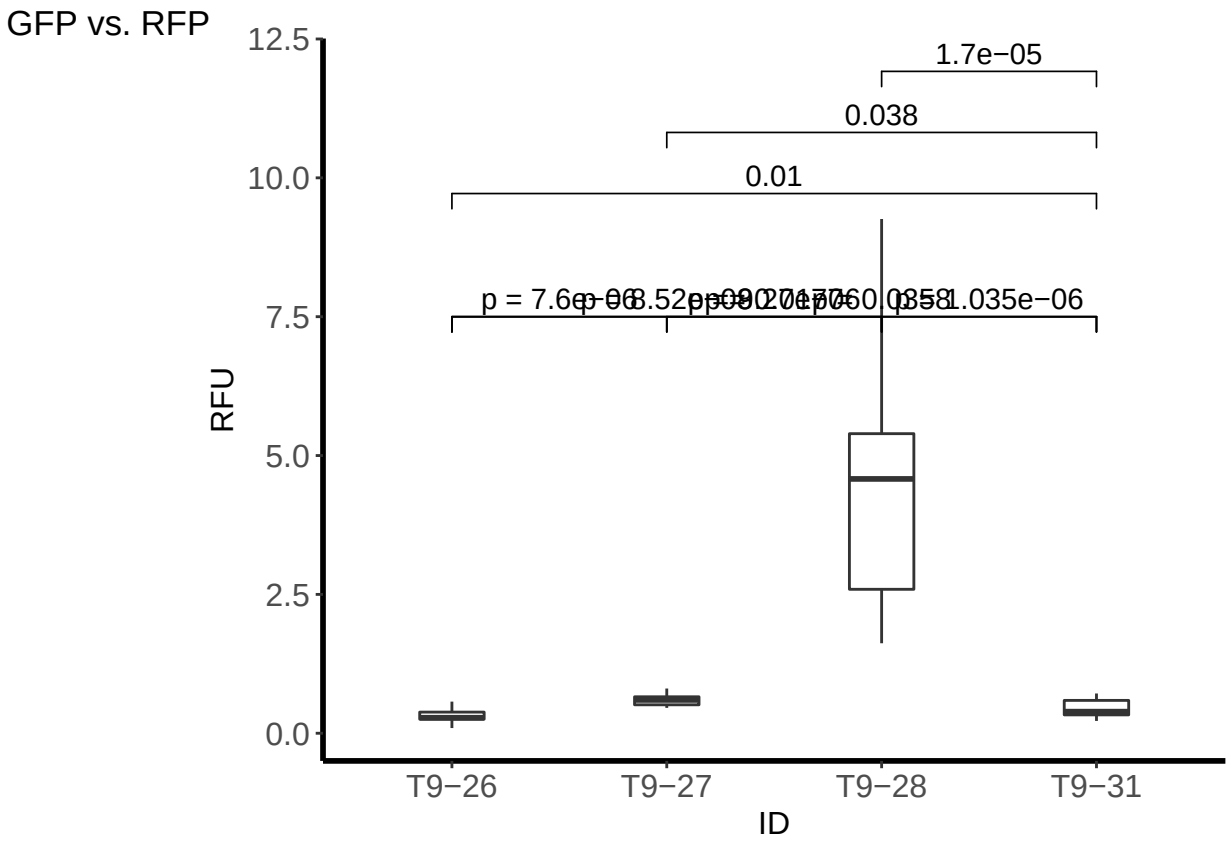
```
p4adjust
```



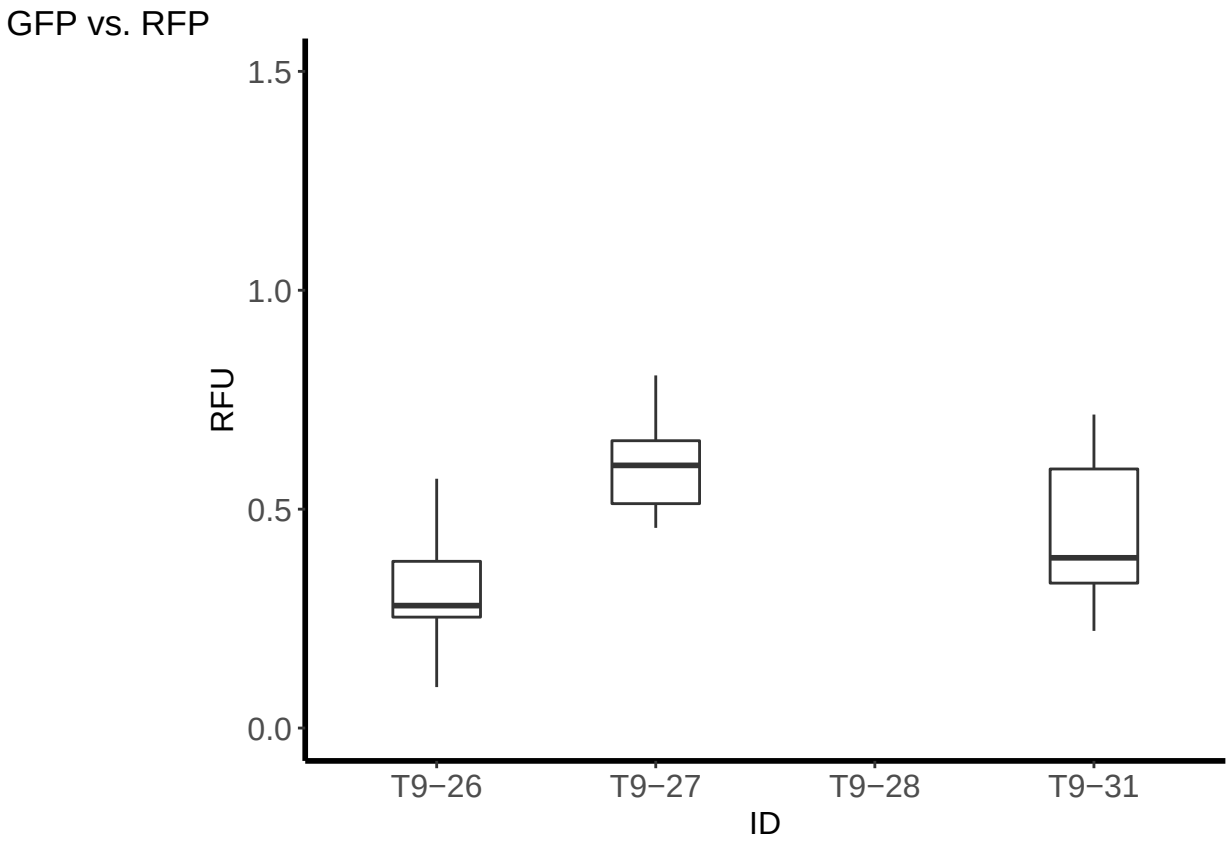

```
p5adjust
```




```
p6adjust
```


GFP vs. RFP
12.5
10.0
7.5
5.0
2.5
0.0
RFU
1.7e−05
0.038
0.01
T9−26
T9−27
T9−28
T9−31
ID

```
p6zoom
```




```
# RA test von Mau
#install.packages("RATest")

#male<-rnorm(50,1,1)
#female<-rnorm(50,1,2)
#dta<-data.frame(group=c(rep(1,50),rep(2,50)),outcome=c(male,female))
#rpt.var<-RPT(dta$outcome~dta$group,test="variances")
#summary(rpt.var)
#rpt.var$pvalue

#rpt.ben <-
```