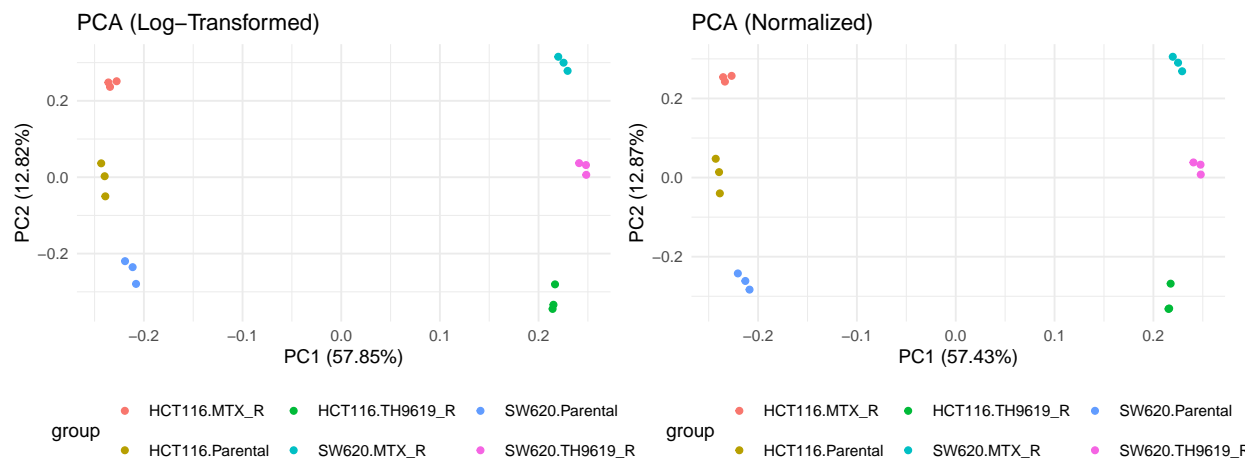


# CRC\_proteomics

EZ

2025-06-05



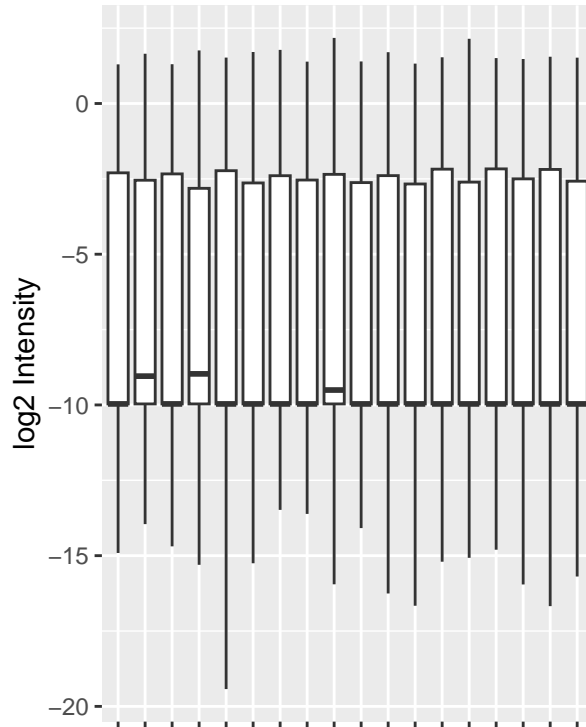
```
# Melt for boxplot comparison
melted_log <- melt(log_mat)
melted_norm <- melt(norm_mat)

p1 <- ggplot(melted_log, aes(x = Var2, y = value)) +
  geom_boxplot(outlier.size = 0.2) +
  theme(axis.text.x = element_blank()) +
  labs(title = "Boxplot: Log-Transformed Intensities", x = "", y = "log2 Intensity")

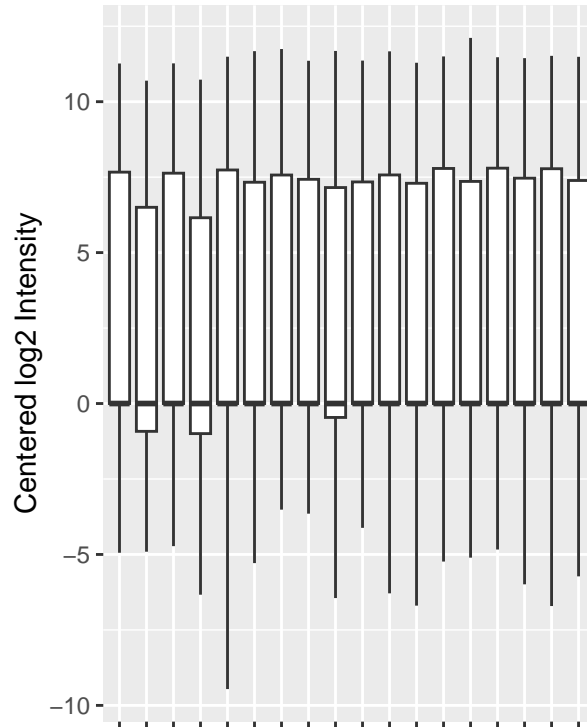
p2 <- ggplot(melted_norm, aes(x = Var2, y = value)) +
  geom_boxplot(outlier.size = 0.2) +
  theme(axis.text.x = element_blank()) +
  labs(title = "Boxplot: Normalized Intensities", x = "", y = "Centered log2 Intensity")

library(patchwork)
p1 + p2
```

Boxplot: Log-Transformed Intensities



Boxplot: Normalized Intensities



```
sum(is.na(norm_mat))
```

```
## [1] 0
```

```
qf <- filterNA(qf, i = "norm_proteins", pNA = 0.5)
na_percent <- rowMeans(is.na(norm_mat)) * 100

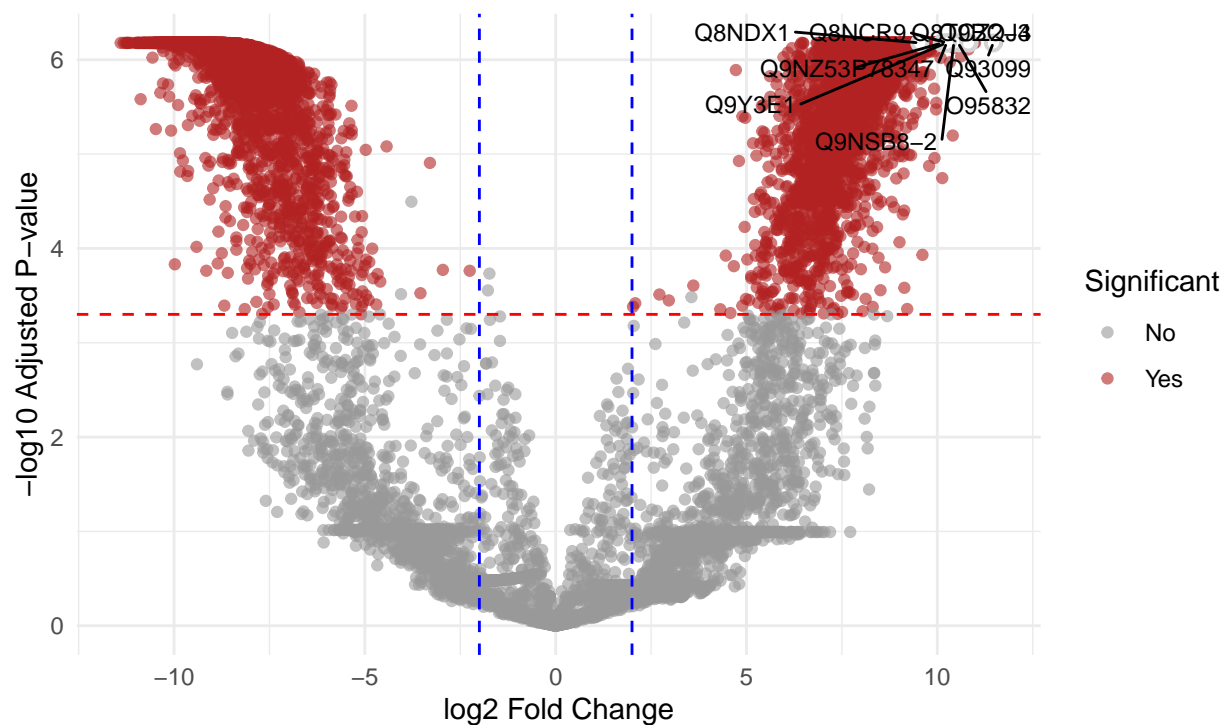
if (any(na_percent > 0)) {
  hist(na_percent,
       breaks = 30,
       main = "Missing Value Distribution (Normalized Data)",
       xlab = "% Missing Across Samples",
       col = "#87ceeb")
} else {
  plot.new()
  title(main = "No Missing Values Detected in Normalized Data")
}
```

## No Missing Values Detected in Normalized Data

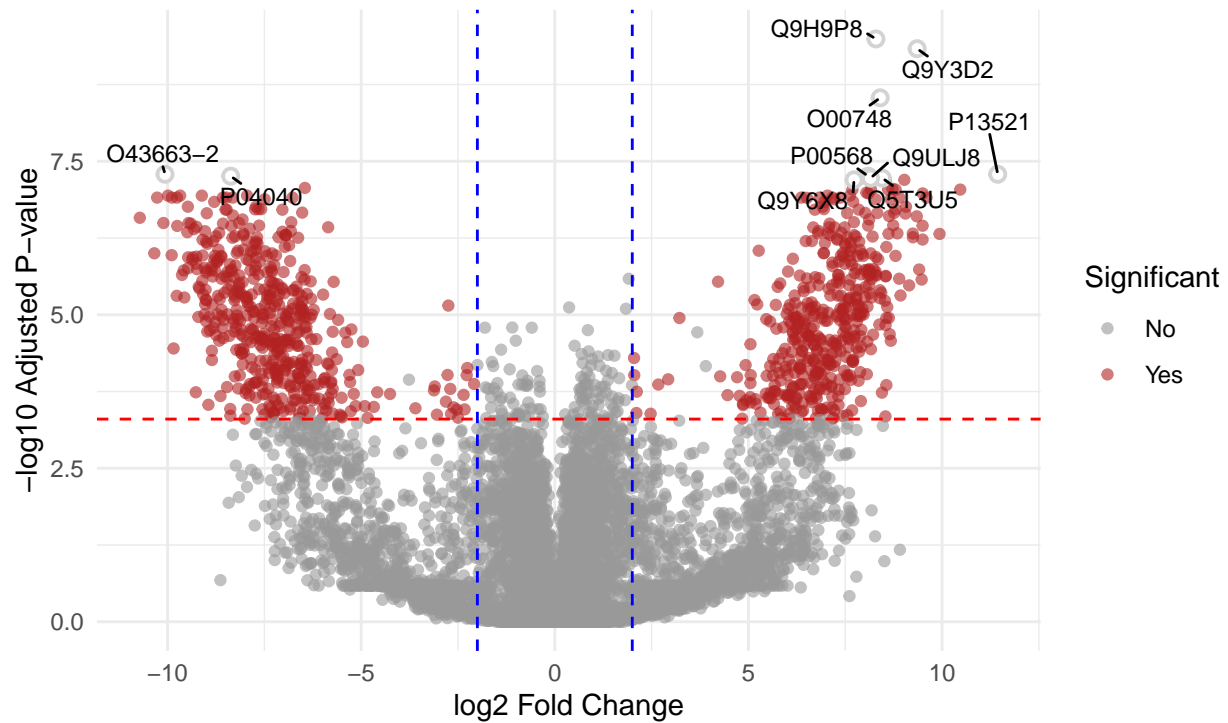
Table 1: This table summarizes how many proteins were significantly differentially expressed ( $\text{FDR} < 0.0005$ ,  $|\log_2\text{FC}| > 2$ ,  $\text{Ave-Expr} > 2$ ) in each comparison

Comparison		DEG_Count
HCT116_TH9619_R_vs_Parental	HCT116_TH9619_R_vs_Parental	6208
HCT116_MTX_R_vs_Parental	HCT116_MTX_R_vs_Parental	943
SW620_TH9619_R_vs_Parental	SW620_TH9619_R_vs_Parental	6670
SW620_MTX_R_vs_Parental	SW620_MTX_R_vs_Parental	6716

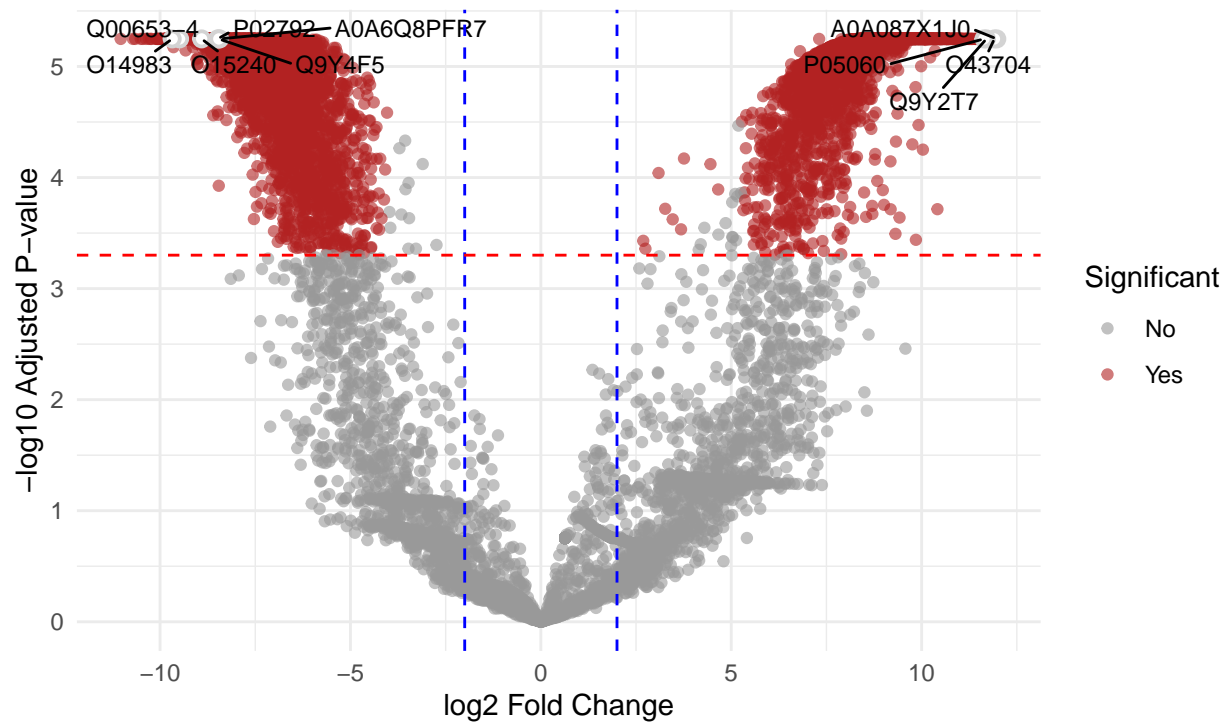
The volcano plots visualize each comparison, showing log2 fold change (x-axis) vs.  $-\log_{10}$  adjusted p-value (y-axis). Top 10 most significant genes are labeled HCT116\_TH9619\_R\_vs\_Parental



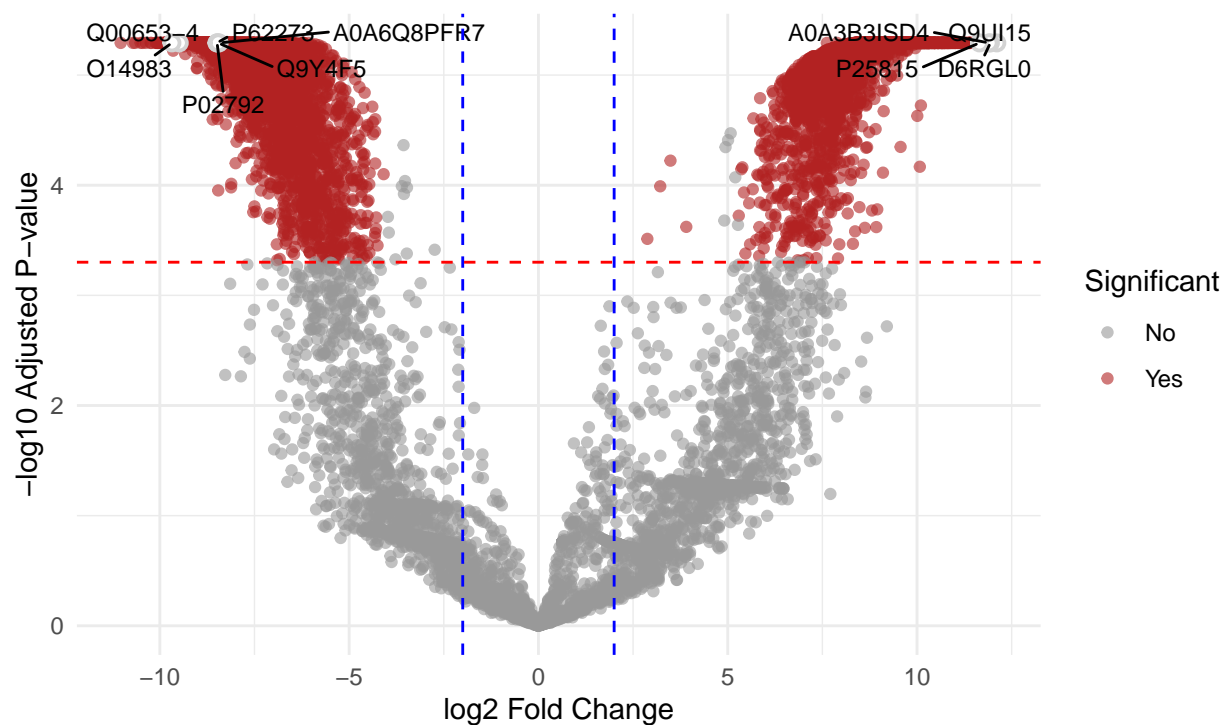
The volcano plots visualize each comparison, showing log2 fold change (x-axis) vs.  $-\log_{10}$  adjusted p-value (y-axis). Top 10 most significant genes are labeled.



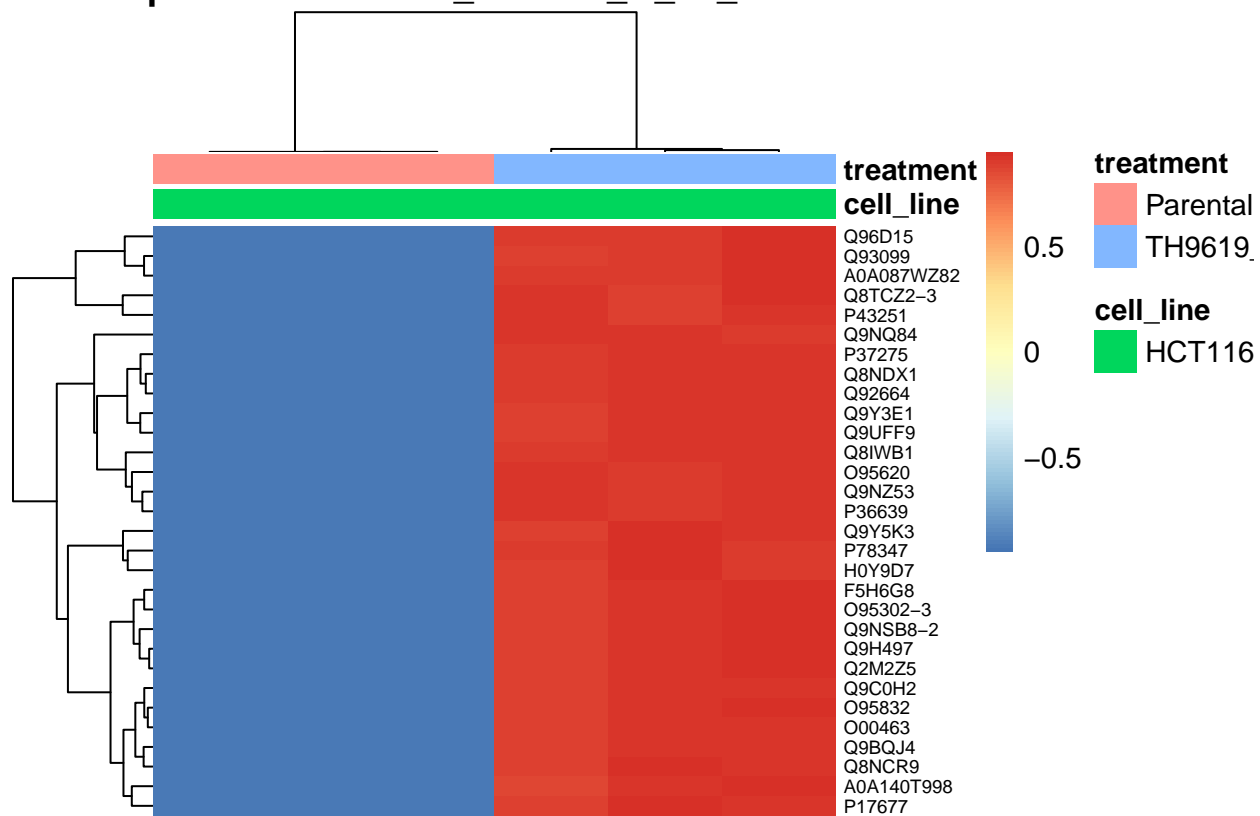
The volcano plots visualize each comparison, showing log2 fold change (x-axis) vs.  $-\log_{10}$  adjusted p-value (y-axis). Top 10 most significant genes are labeled SW620\_TH9619\_R\_vs\_Parental



The volcano plots visualize each comparison, showing log2 fold change (x-axis) vs.  $-\log_{10}$  adjusted p-value (y-axis). Top 10 most significant genes are labeled SW620\_MTX\_R\_vs\_Parental

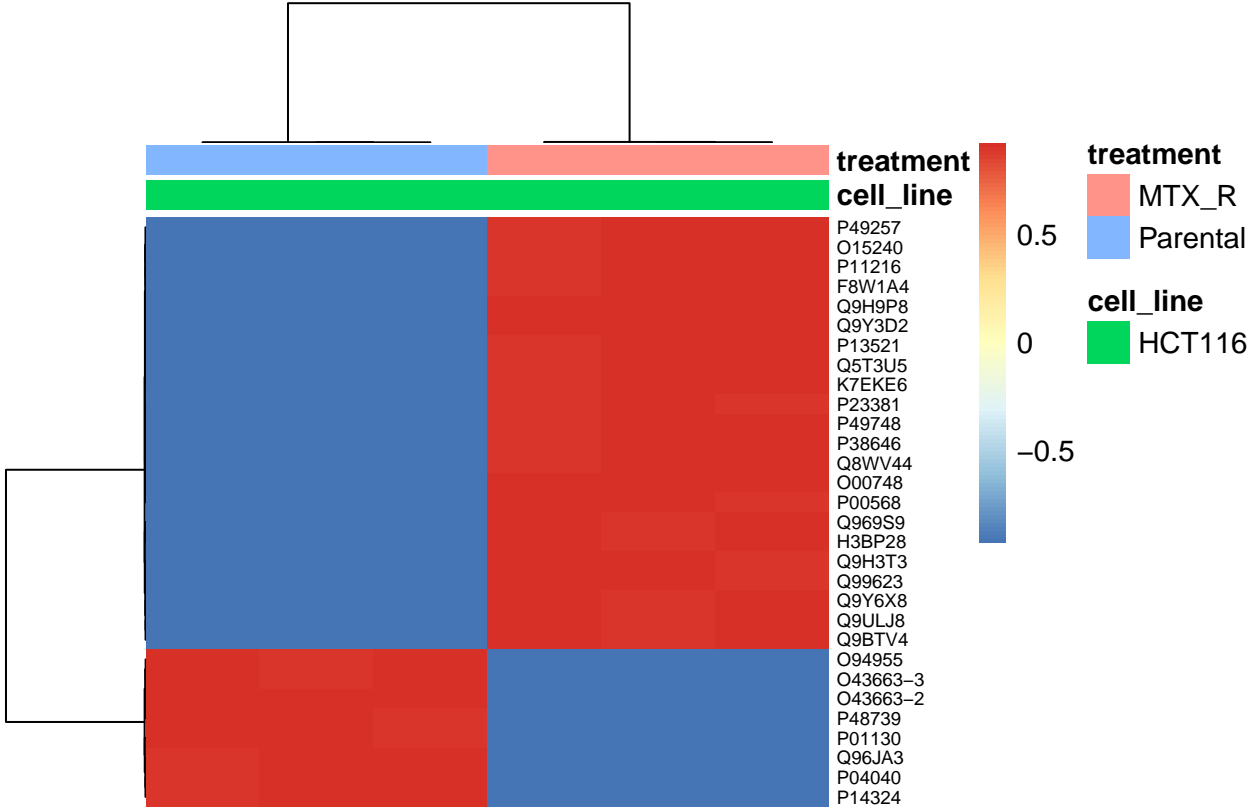


## Top 30 DEGs – HCT\_TH9619\_R\_vs\_Parental

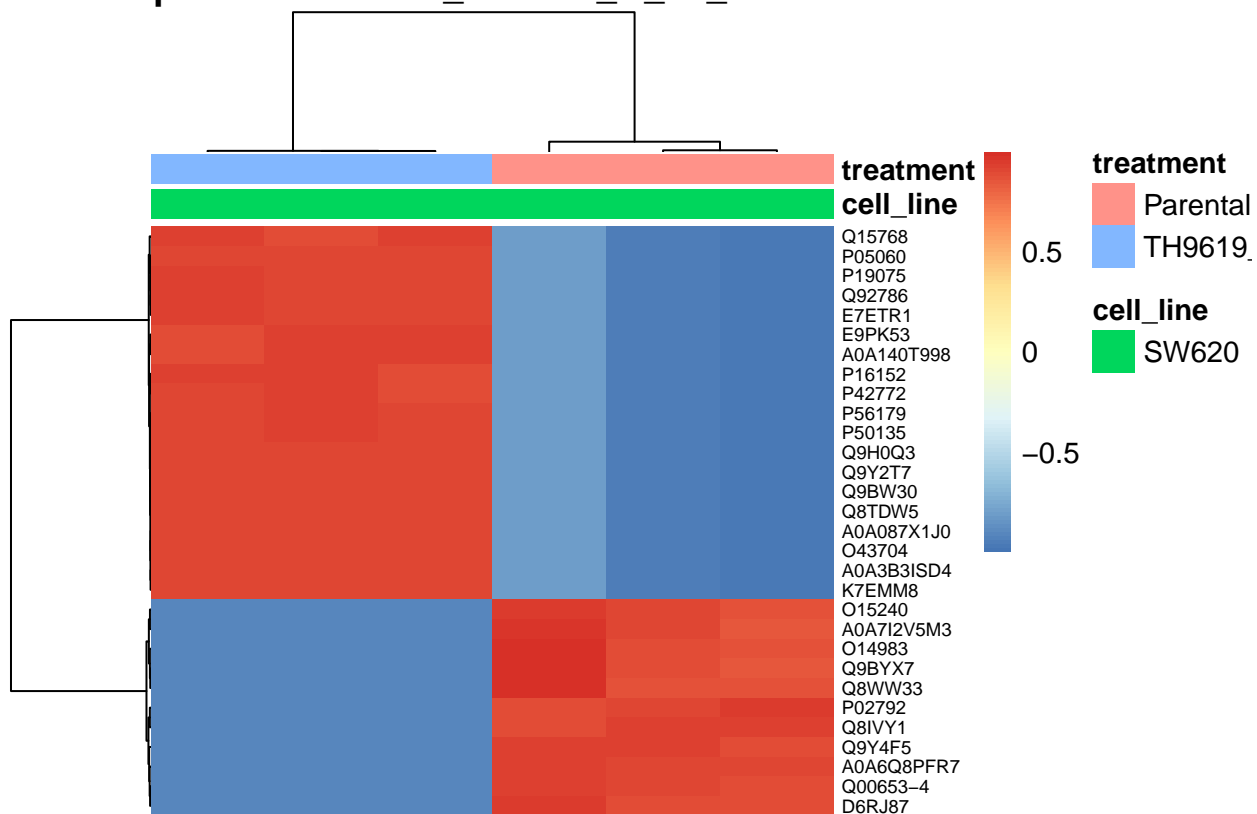




Top 30 DEGs – HCT\_MTX\_R\_vs\_Parental



## Top 30 DEGs – SW\_TH9619\_R\_vs\_Parental



Top 30 DEGs – SW\_MTX\_R\_vs\_Parental

