

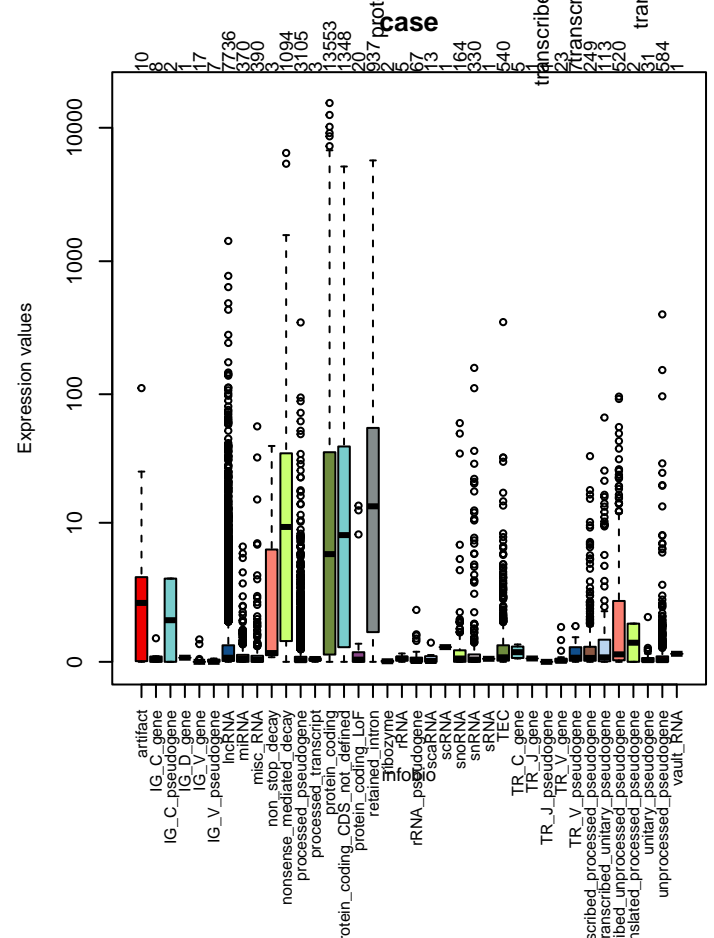
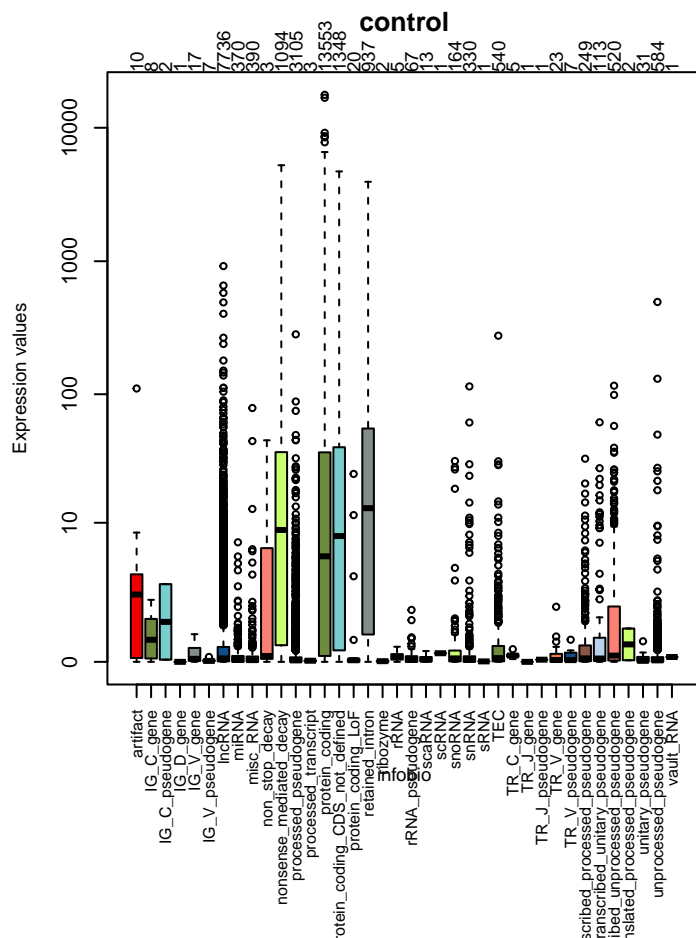
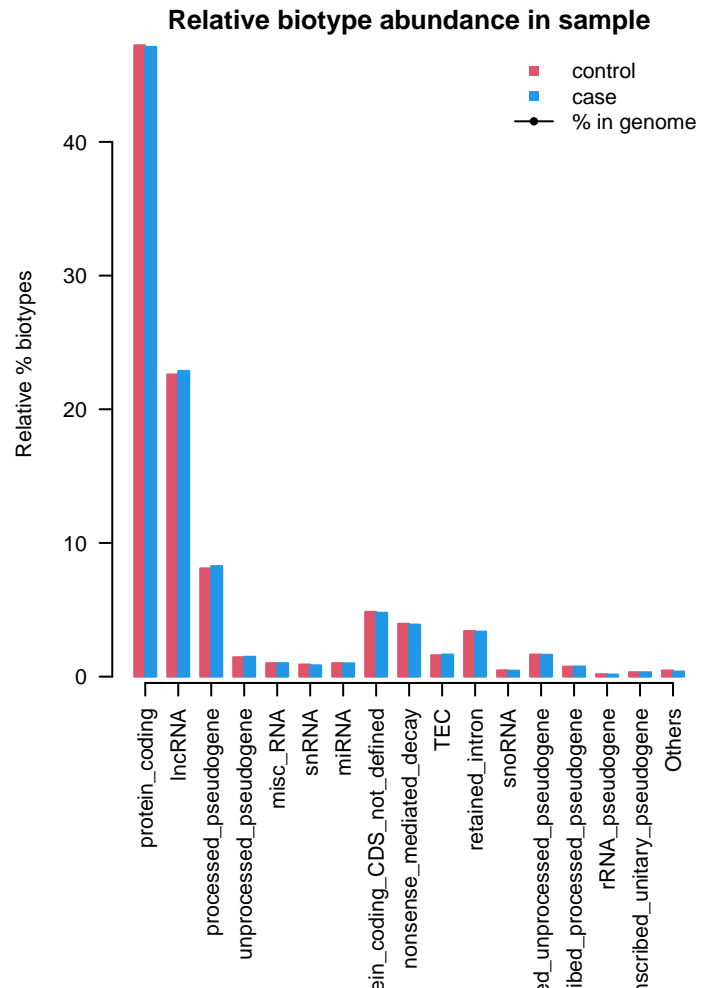
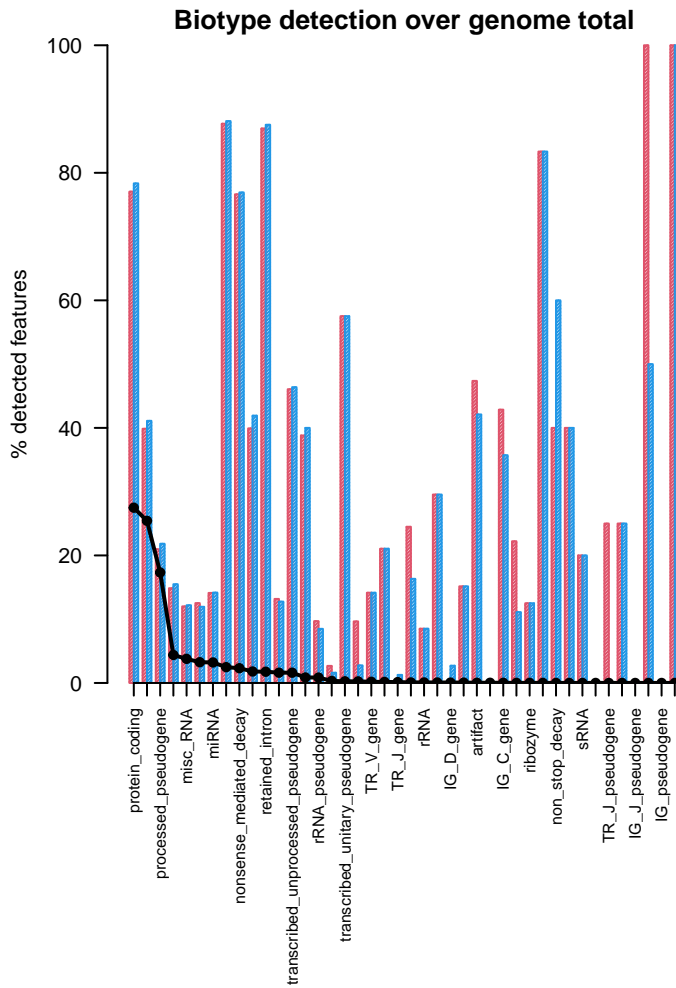
# Quality Control of Expression Data

Generated by NOISeq on 28 Jan 2024, 15:25:29

## Content

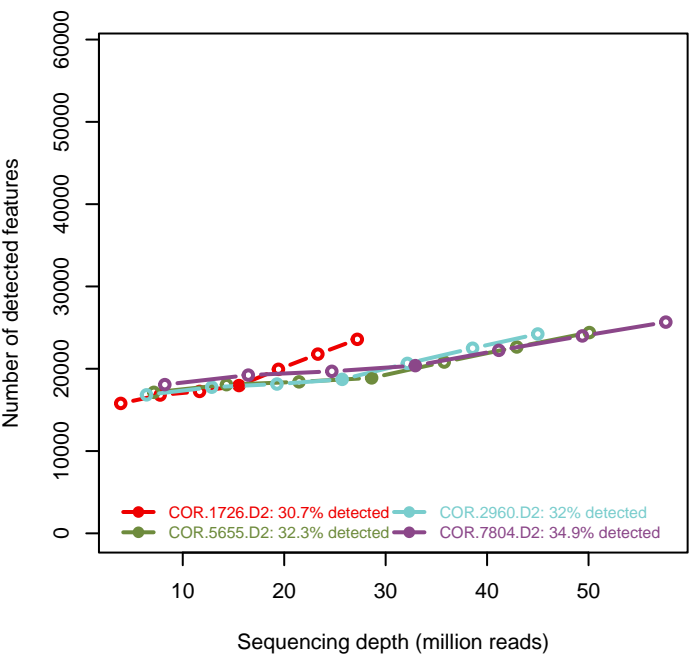
Plot	Description
<b>Biotype detection</b>	Biotype abundance in the genome with %genes detected (counts > 0) in the sample/condition. Biotype abundance within the sample/condition.
<b>Biotype expression</b>	Distribution of gene counts per million per biotype in sample/condition (only genes with counts > 0).
<b>Saturation</b>	Number of detected genes (counts > 0) per sample across different sequencing depths
<b>Expression boxplot</b>	Distribution of gene counts per million (all biotypes) in each sample/condition
<b>Expression barplot</b>	Percentage of genes with >0, >1, >2, >5 or >10 counts per million in each sample/condition.
<b>Length bias</b>	Mean gene expression per each length bin. Fitted curve and diagnostic test.
<b>GC content bias</b>	Mean gene expression per each GC content bin. Fitted curve and diagnostic test.
<b>RNA composition bias</b>	Density plots of log fold changes (M) between pairs of samples. Confidence intervals for the median of M values.
<b>Exploratory PCA</b>	Principal Component Analysis score plots for PC1 vs PC2, and PC1 vs PC3.

## Biotype detection

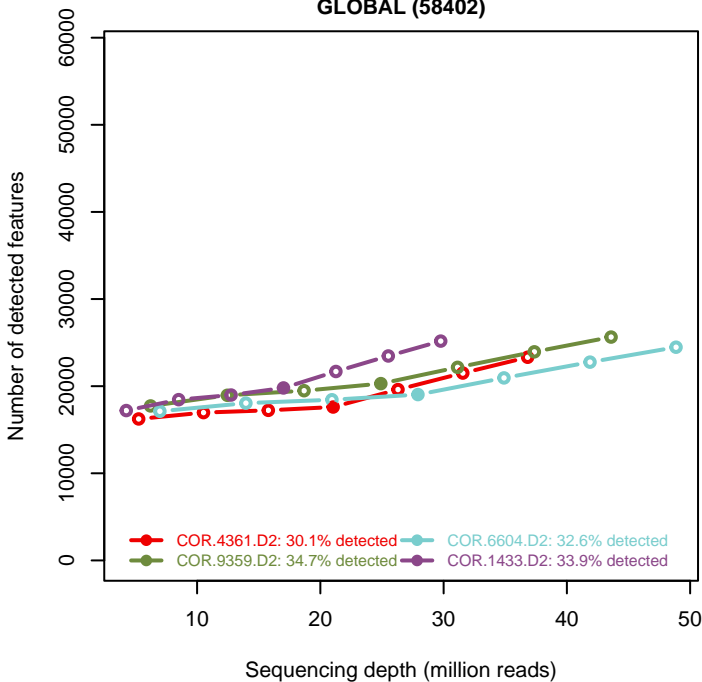


# Sequencing depth & Expression quantification

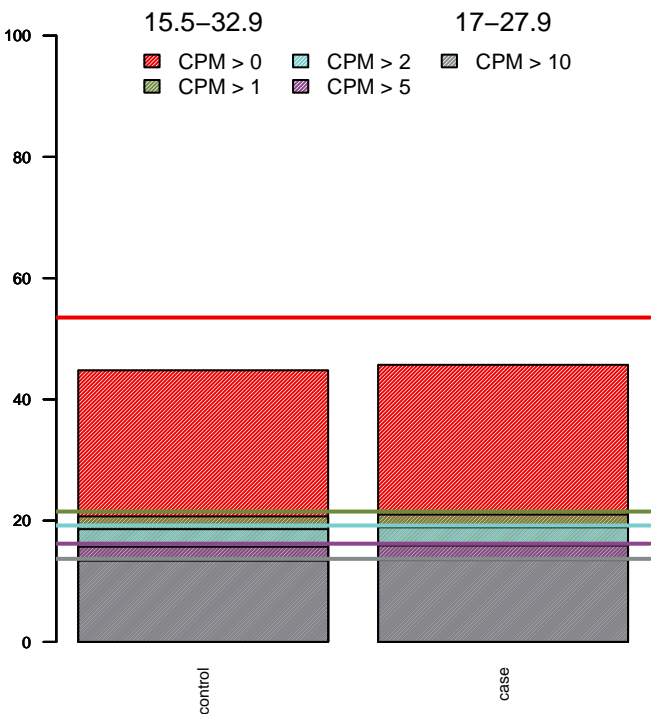
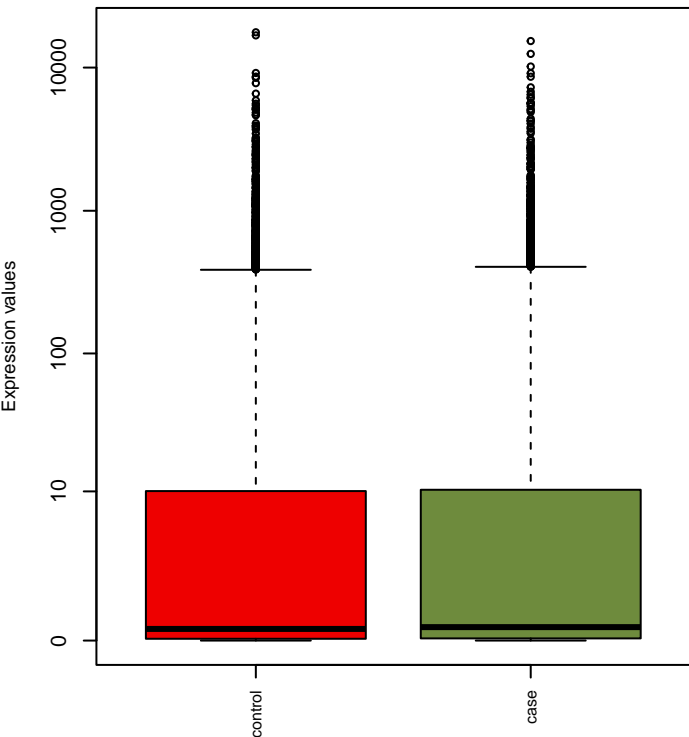
GLOBAL (58402)



GLOBAL (58402)



GLOBAL (31264)



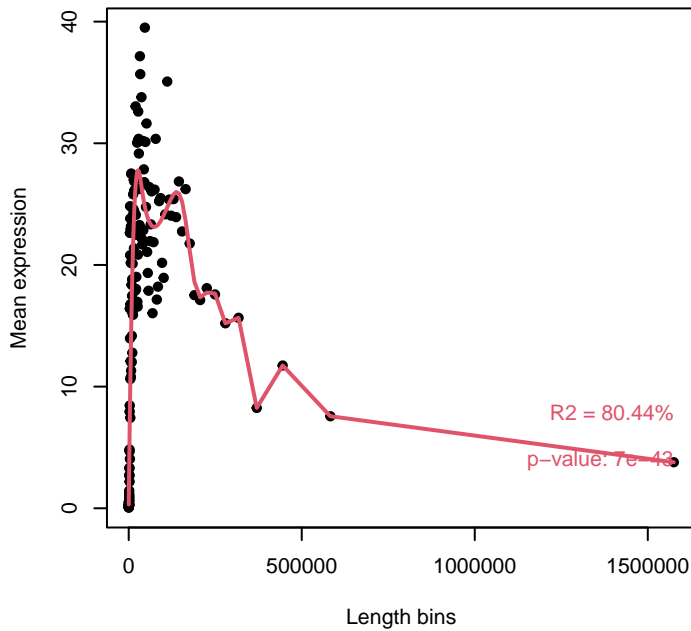
# Sequencing bias detection

## Diagnostic plot for feature length bias

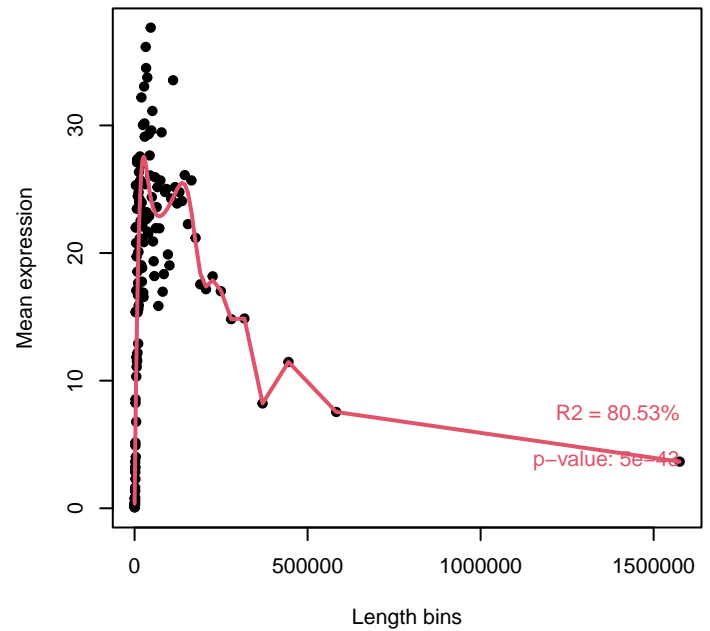
FAILED. At least one of the model p-values was lower than 0.05 and  $R^2 > 70\%$ .

Normalization for correcting length bias is recommended.

control



case



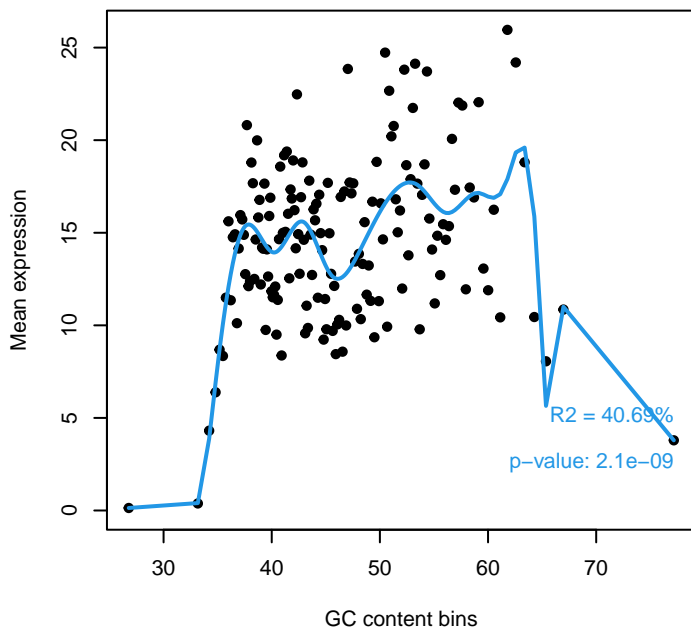
## Diagnostic plot for GC content bias

WARNING. At least one of the model p-values was lower than 0.05, but  $R^2 < 70\%$  for at least one condition.

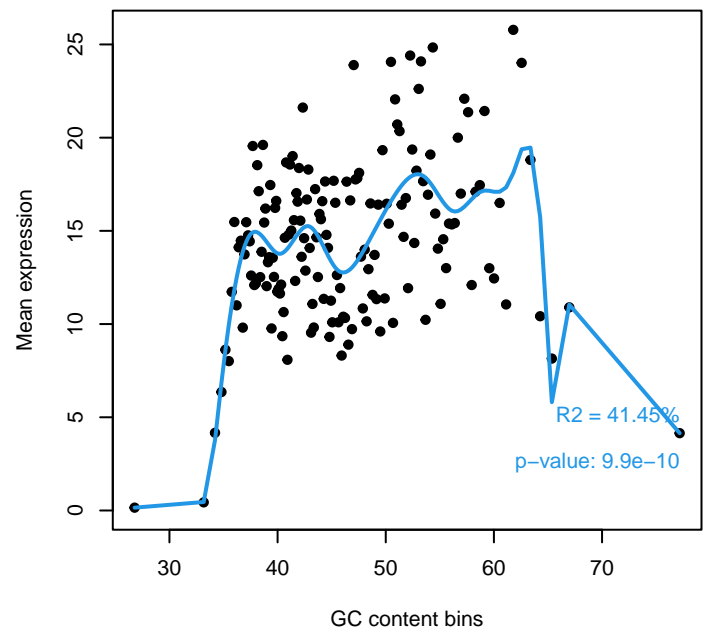
Normalization for correcting GC content bias could be advisable.

Please check in the plots below the strength of the relationship between GC content and expression.

control



case

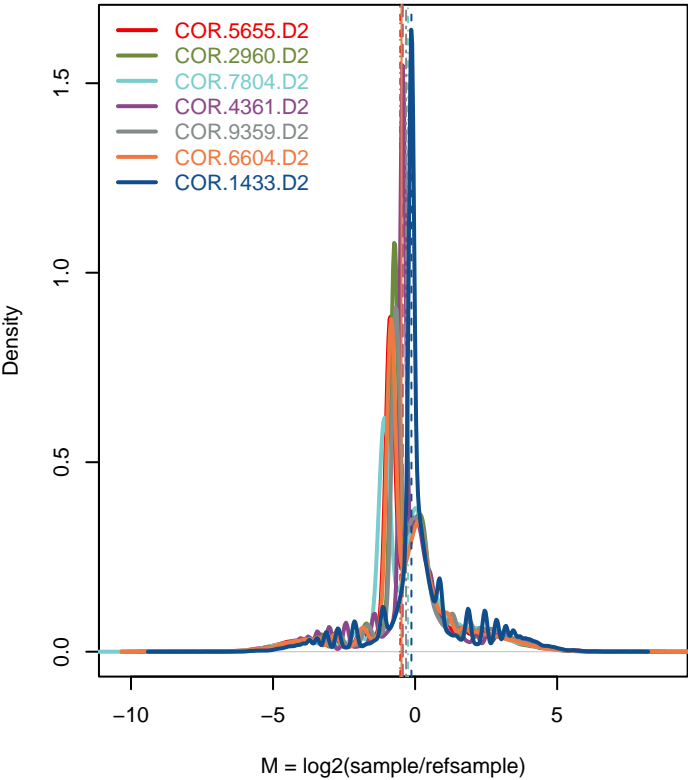


Diagnostic plot for differences in RNA composition

FAILED. There is a pair of samples with significantly different RNA composition

Normalization for correcting this bias is required.

Reference sample: COR.1726.D2



Confidence intervals for median of M values

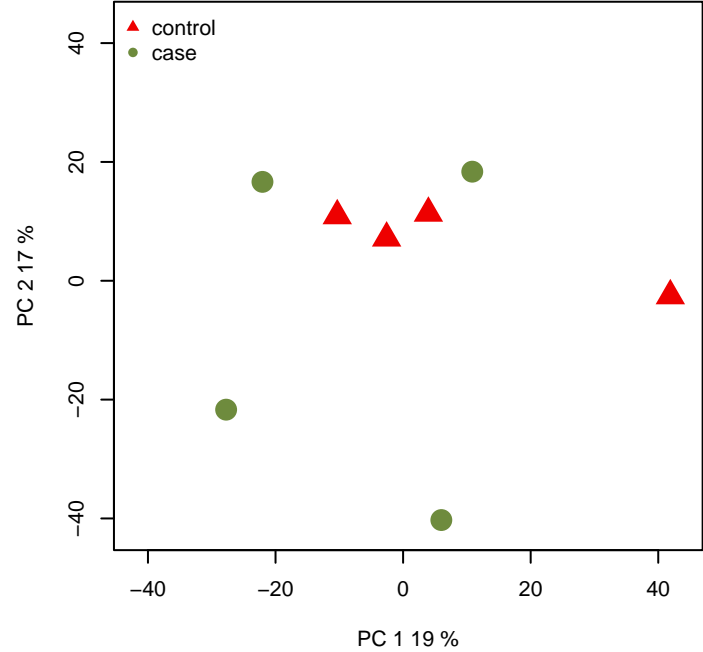
Sample	0.36%	99.64%	Diagnostic Test
COR.5655.D2	-0.5553	-0.4695	FAILED
COR.2960.D2	-0.574	-0.4728	FAILED
COR.7804.D2	-0.2766	-0.2192	FAILED
COR.4361.D2	-0.4365	-0.4365	FAILED
COR.9359.D2	-0.3413	-0.2914	FAILED
COR.6604.D2	-0.517	-0.4356	FAILED
COR.1433.D2	-0.1297	-0.1297	FAILED

# Exploratory PCA

Use this plot to see if samples are clustered according to the experimental design.

Use ARSyNseq function to correct potential batch effects.

Scores



Scores

