

# Quality Control of Expression Data

Generated by NOISEq on 01 Jul 2015, 14:55:28

## Content

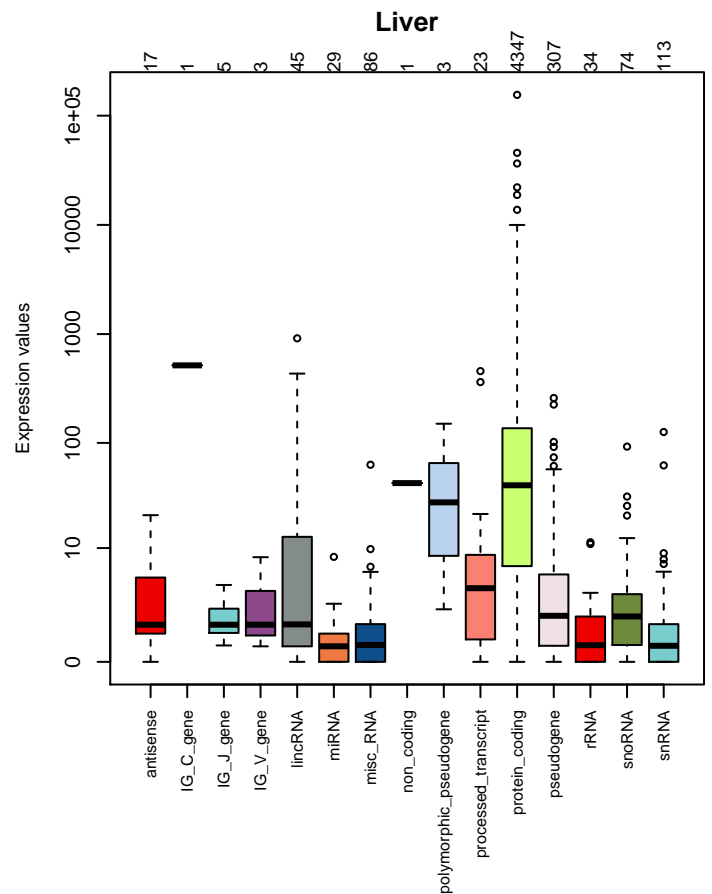
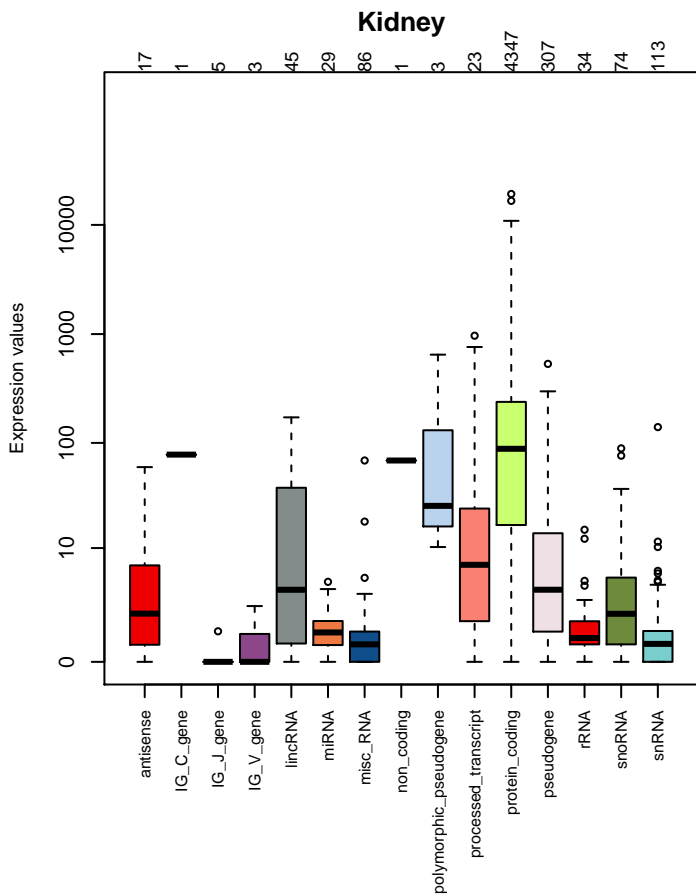
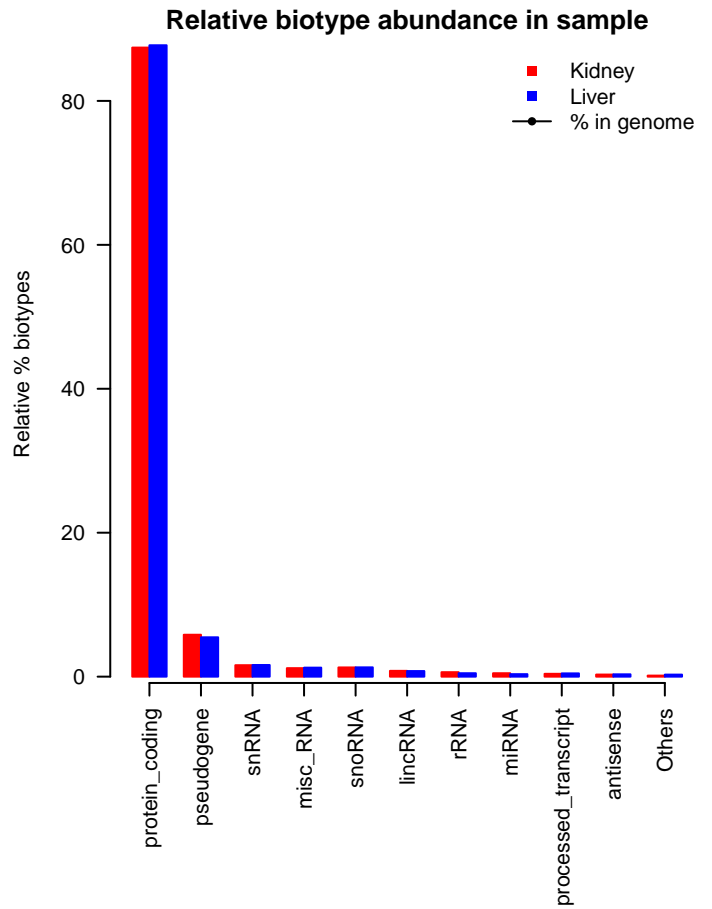
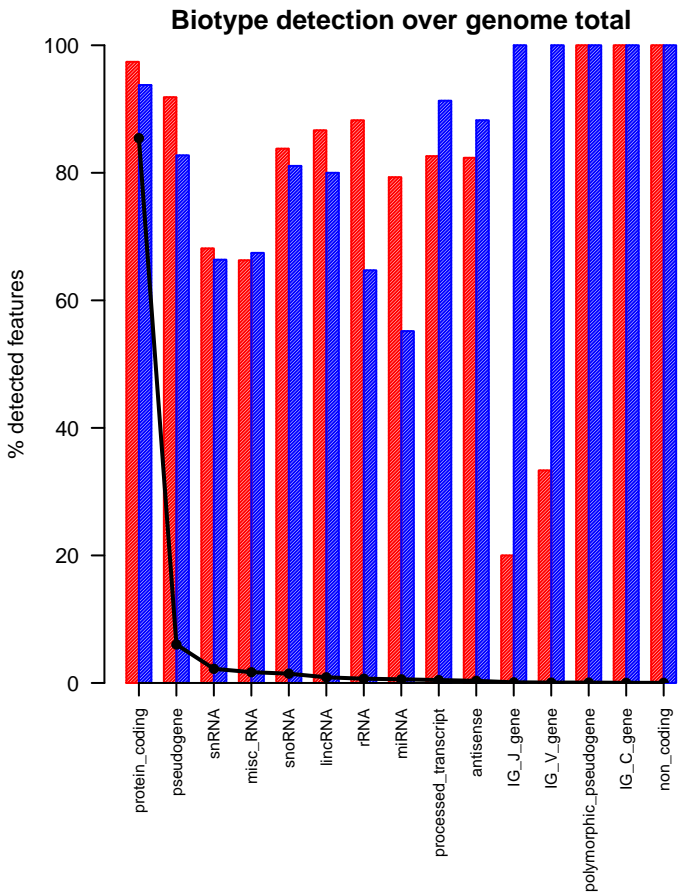
*Plot*

*Description*

---

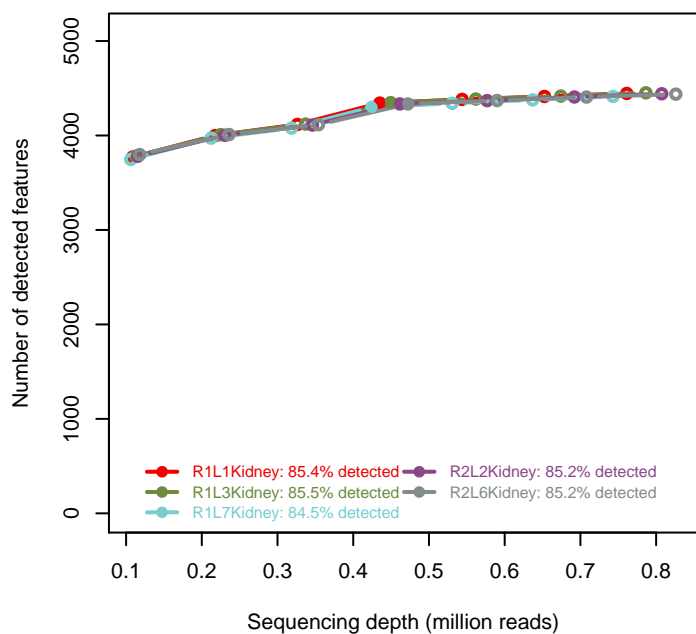
<b>Biotype detection</b>	Biotype abundance in the genome with %genes detected (counts > 0) in the sample/condition. Biotype abundance within the sample/condition.
<b>Biotype expression</b>	Distribution of gene counts per million per biotype in sample/condition (only genes with counts > 0).
<b>Saturation</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

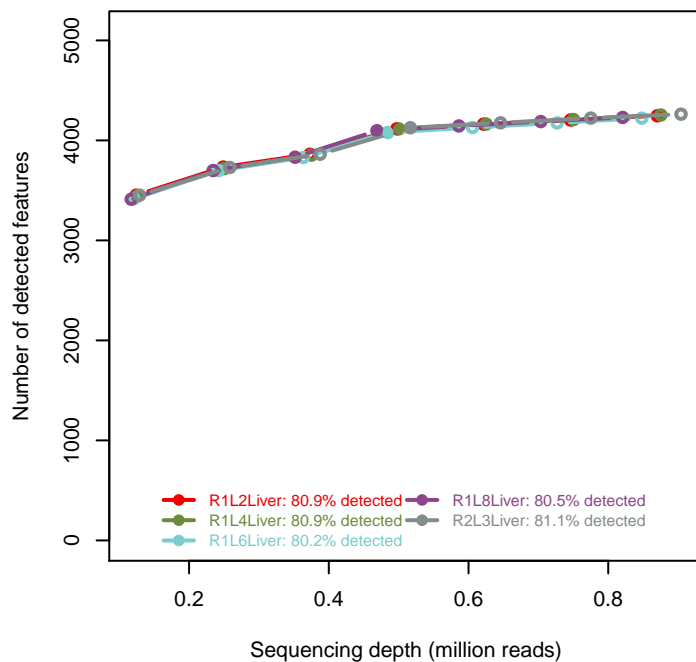


# Sequencing depth & Expression quantification

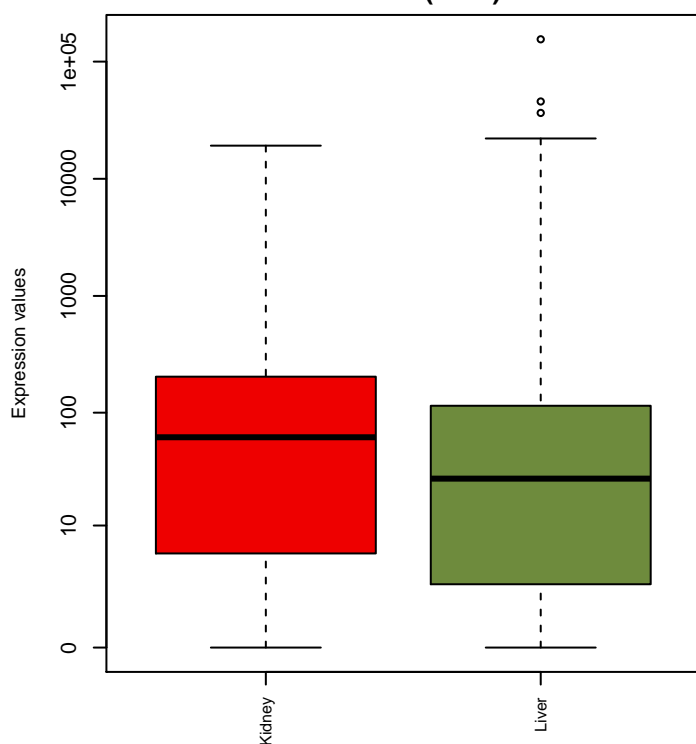
GLOBAL (5088)



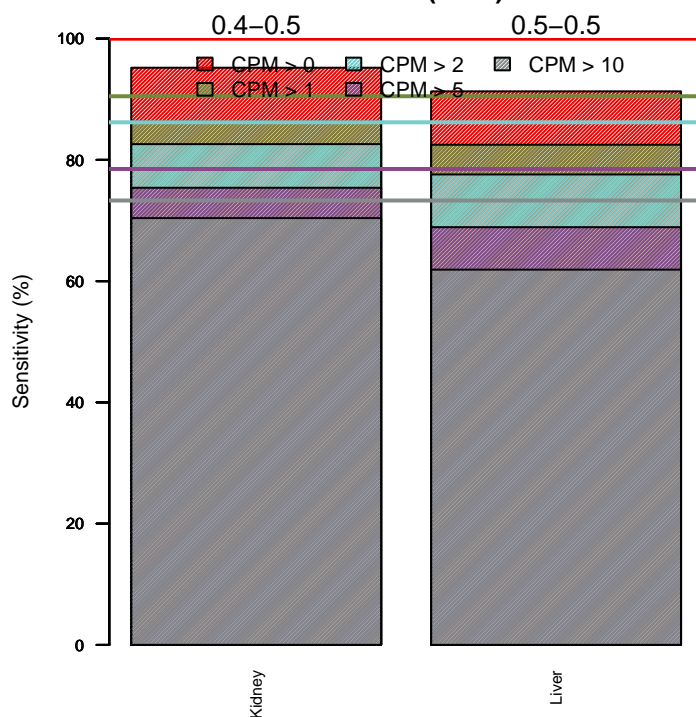
GLOBAL (5088)



GLOBAL (5088)



GLOBAL (5088)



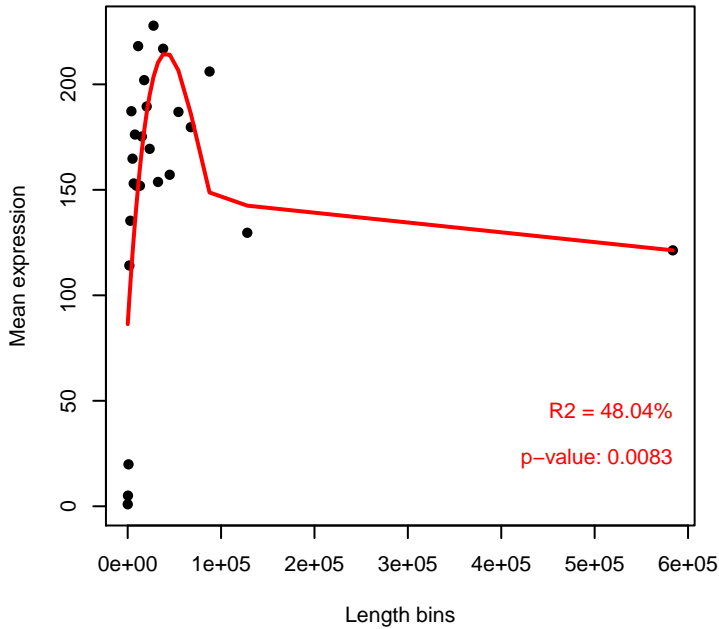
# Sequencing bias detection

## Diagnostic plot for feature length bias

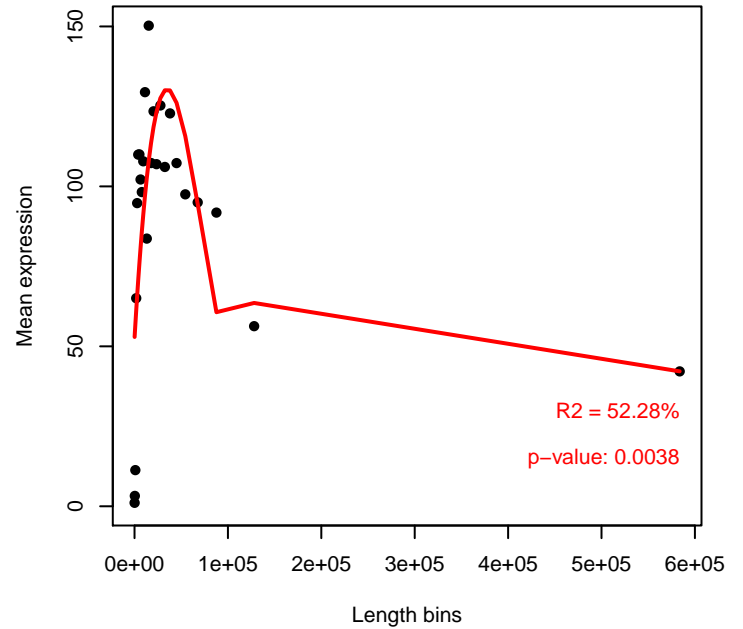
WARNING. At least one of the model p-values was lower than 0.05, but  $R^2 < 70\%$  for at least one condition.

Normalization for correcting length bias could be advisable.  
Please check in the plots below the strength of the relationship between length and expression.

### Kidney



### Liver

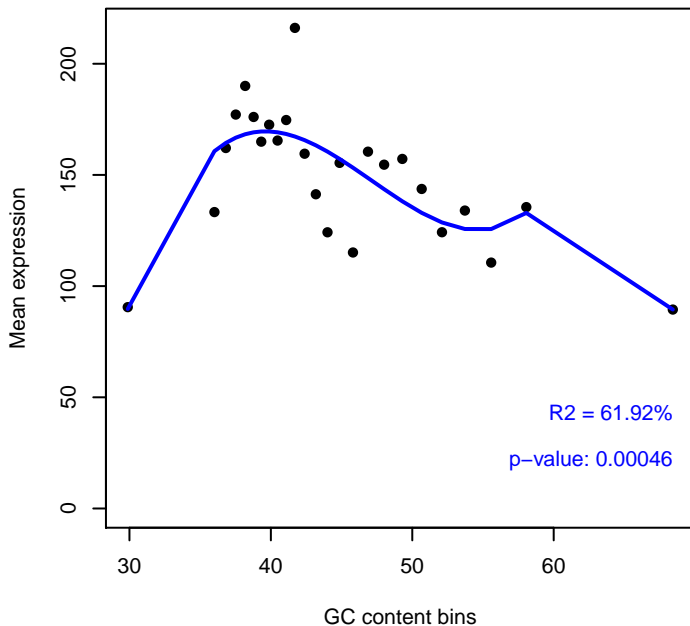


## Diagnostic plot for GC content bias

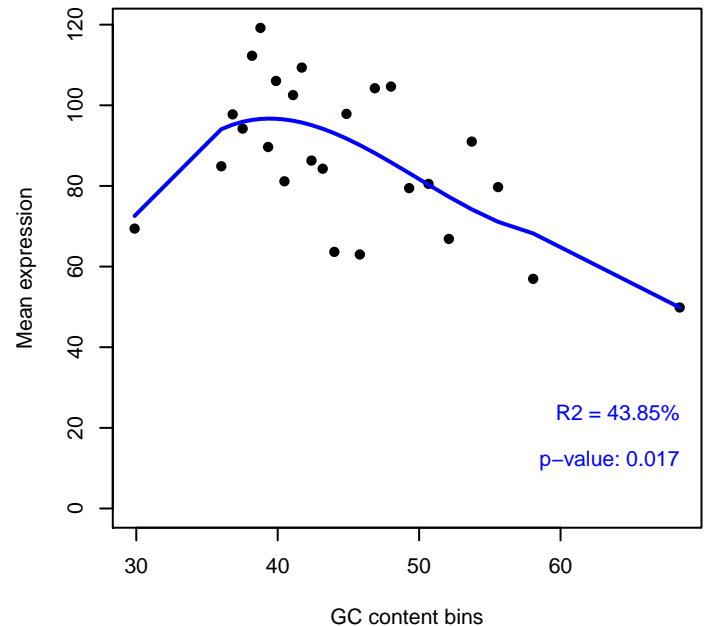
WARNING. At least one of the model p-values was lower than 0.05, but  $R^2 < 70\%$  for at least one condition.

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.

### Kidney



### Liver

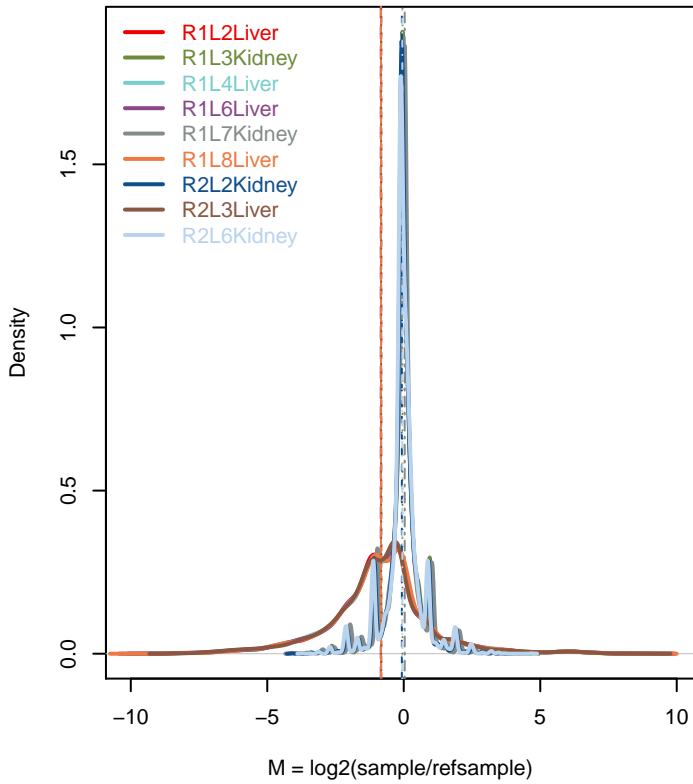


## Diagnostic plot for differences in RNA composition

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

Normalization for correcting this bias is required.

### Reference sample: R1L1Kidney



### Confidence intervals for median of M values

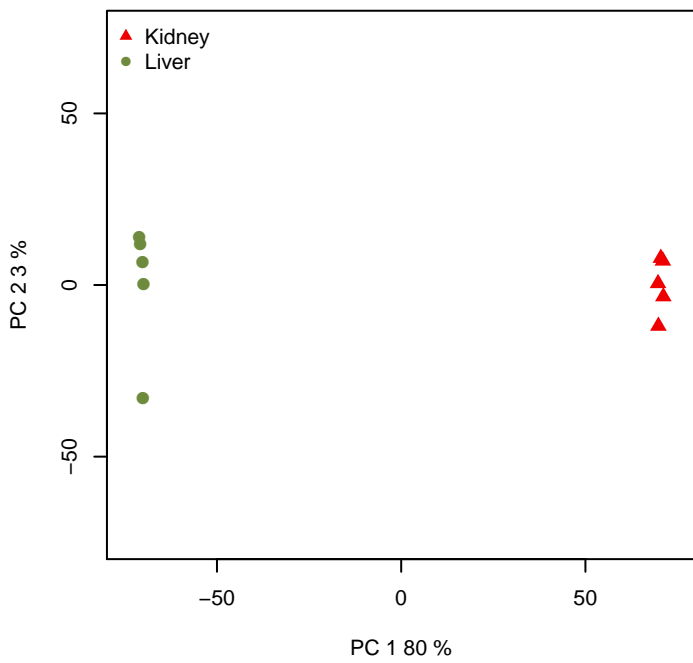
Sample	0.28%	99.72%	Diagnostic Test
R1L2Liver	-0.8895	-0.7676	FAILED
R1L3Kidney	-0.0472	-0.0472	FAILED
R1L4Liver	-0.8796	-0.7572	FAILED
R1L6Liver	-0.9148	-0.7632	FAILED
R1L7Kidney	0.0348	0.0348	FAILED
R1L8Liver	-0.8927	-0.7552	FAILED
R2L2Kidney	-0.085	-0.0454	FAILED
R2L3Liver	-0.8735	-0.7572	FAILED
R2L6Kidney	-0.0696	-0.0361	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

