Class12: Homework

Rachel Lamm (A18518313)

2026-02-16

Q13: Read this file into R and determine the sample size for each genotype and their corresponding median expression levels for each of these genotypes.

```
data <- read.table("rs8067378_ENSG00000172057.6.txt", header=TRUE)
```

```
table(data$geno)
```

```
##
## A/A A/G G/G
## 108 233 121
```

```
summary(data$exp[data$geno=="A/A"])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   11.40   27.02   31.25   31.82   35.92   51.52
```

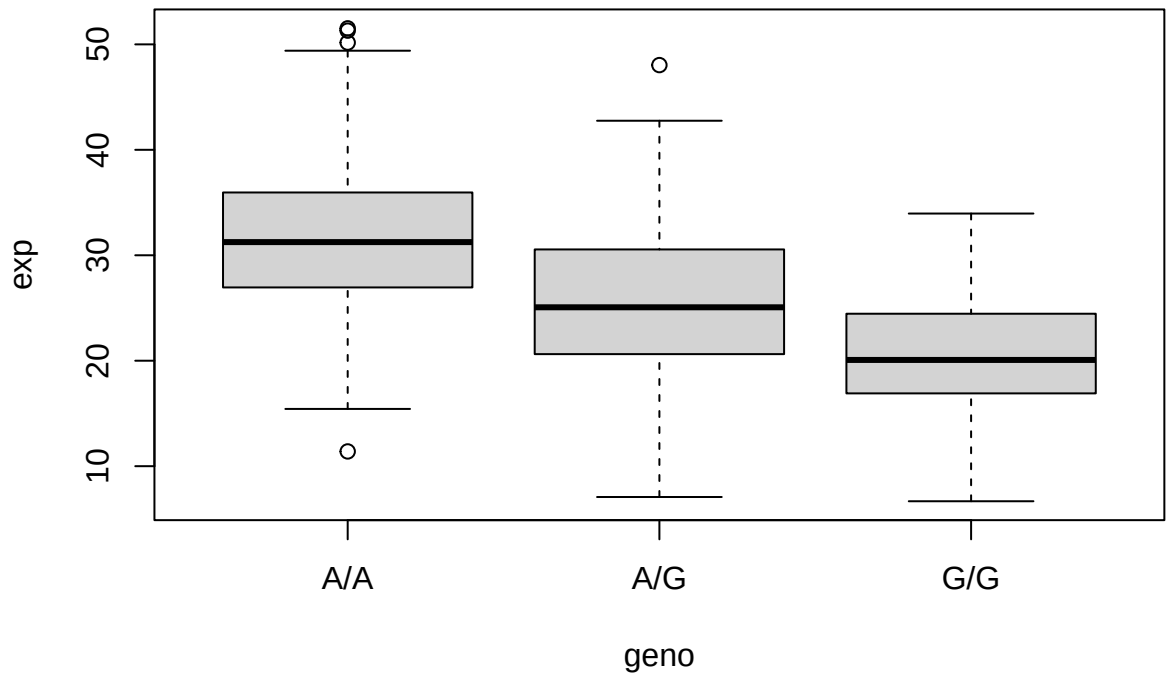
```
summary(data$exp[data$geno=="A/G"])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   7.075  20.626  25.065  25.397  30.552  48.034
```

```
summary(data$exp[data$geno=="G/G"])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   6.675  16.903  20.074  20.594  24.457  33.956
```

```
bp <- boxplot(exp ~ geno, data=data)
```



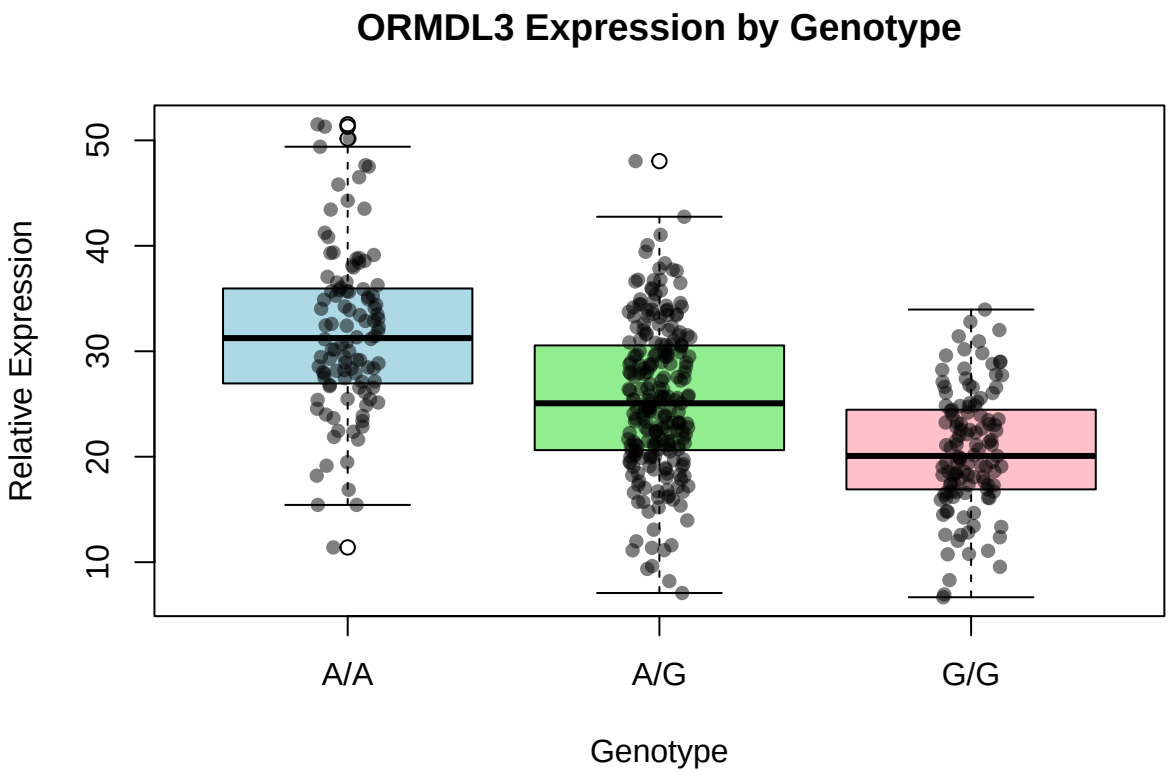

```
bp$stats[3,]
```

```
## [1] 31.24847 25.06486 20.07363
```

Q14: Generate a boxplot with a box per genotype, what could you infer from the relative expression value between A/A and G/G displayed in this plot? Does the SNP effect the expression of ORMDL3?

```
boxplot(exp ~ geno, data=data,
        main="ORMDL3 Expression by Genotype",
        xlab="Genotype",
        ylab="Relative Expression",
        col=c("lightblue", "lightgreen", "pink"))

stripchart(exp ~ geno,
           data=data,
           method="jitter",
           pch=16,
           col=rgb(0,0,0,0.5),
           vertical=TRUE,
           add=TRUE)
```


ORMDL3 Expression by Genotype
Relative Expression
10
20
30
40
50
A/A
A/G
G/G
Genotype