class11

Q13

```r
file<-read.table("SNP.txt")
GG<-file[file$geno=="G/G",]
summary(GG)
```

```
    sample              geno                  exp
 Length:121         Length:121         Min.   : 6.675
 Class :character   Class :character   1st Qu.:16.903
 Mode  :character   Mode  :character   Median :20.074
                                       Mean   :20.594
                                       3rd Qu.:24.457
                                       Max.   :33.956
```

```r
AG<-file[file$geno=="A/G",]
summary(AG)
```

```
    sample              geno                  exp
 Length:233         Length:233         Min.   : 7.075
 Class :character   Class :character   1st Qu.:20.626
 Mode  :character   Mode  :character   Median :25.065
                                       Mean   :25.397
                                       3rd Qu.:30.552
                                       Max.   :48.034
```
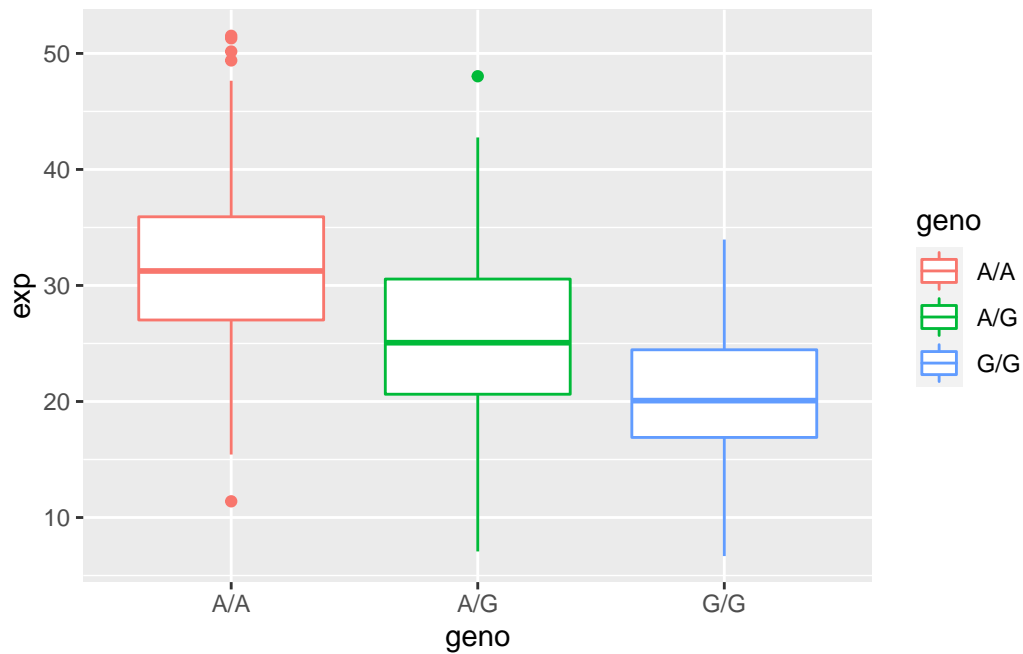
```r
AA<-file[file$geno=="A/A",]
summary(AA)
```

```
    sample              geno                  exp
 Length:108          Length:108          Min.   :11.40
 Class :character    Class :character    1st Qu.:27.02
 Mode  :character    Mode  :character    Median :31.25
                                         Mean   :31.82
                                         3rd Qu.:35.92
                                         Max.   :51.52
```

GG has 121 samples with a median expression level of 20.074.AG has 233 samples with median expression of 25.065. AA has 108 samples with a median expression level of 31.25.

Q14

```
library(ggplot2)
ggplot(file,aes(geno,exp,color=geno))+geom_boxplot()
```


Q14) Based off this boxplot, it appears that the A/A genotype has higher overall expression of ORMDL3 compared to the G/G genotype. The SNP effect appears to affect expression of ORMDL3