

class12_ec

Cameron Finch (A16734770)

2024-02-24

Population Scale Analysis (Questions 13 & 14)

How many samples do we have?

```
expr <- read.table("rs806-ag.txt")  
# submission note: the link to the file from the class 12 assignment pdf would not work. I instead had  
head(expr)
```

```
##      sample geno      exp  
## 1 HG00367  A/G 28.96038  
## 2 NA20768  A/G 20.24449  
## 3 HG00361  A/A 31.32628  
## 4 HG00135  A/A 34.11169  
## 5 NA18870  G/G 18.25141  
## 6 NA11993  A/A 32.89721
```

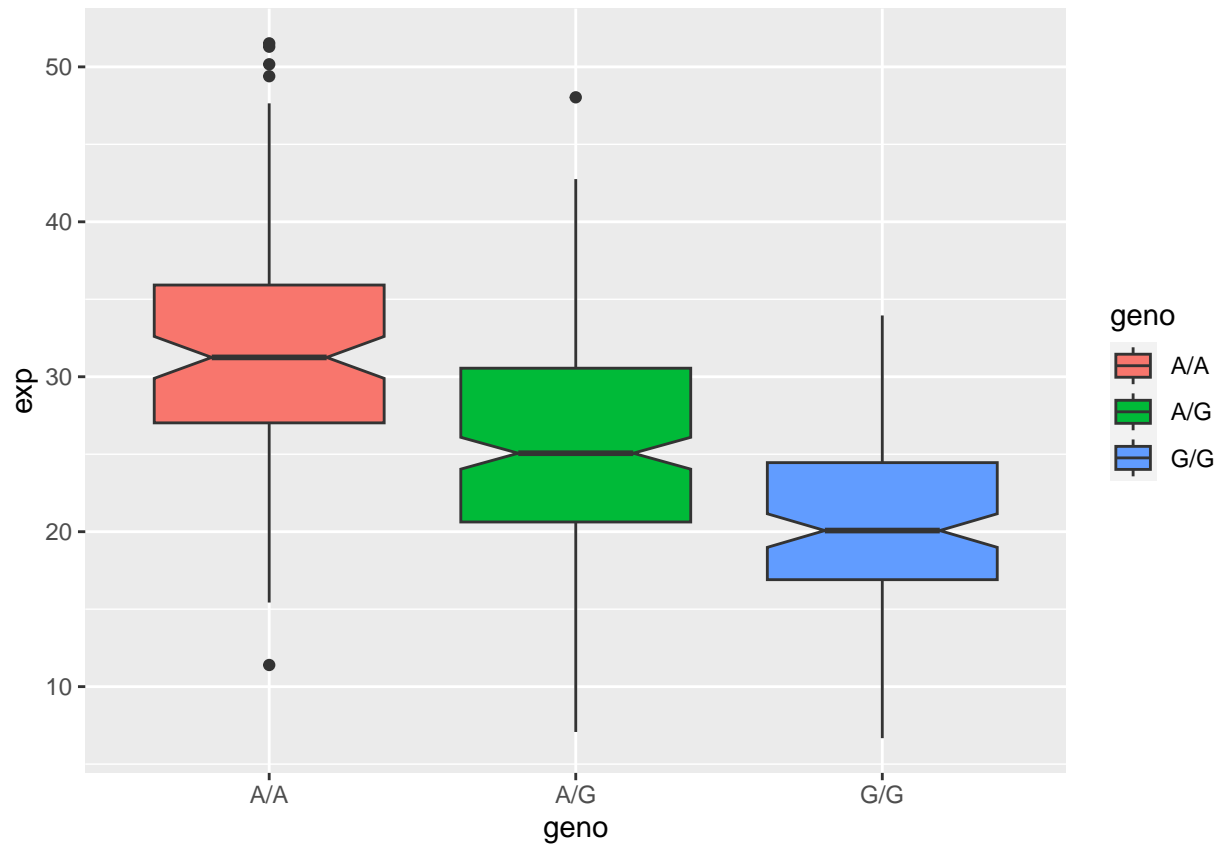
```
nrow(expr)
```

```
## [1] 462
```

```
table(expr$geno)
```

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

```
library(ggplot2)  
ggplot(expr) + aes(x=geno, y=exp, fill=geno) +  
  geom_boxplot(notch=T)
```



> A14. Higher A content seems to increase gene expression (A|A>A|G>G|G). Therefore, the G|G SNP seems to have a clear association with downregulation of ORMDL3 expression.