

Class 11

Cynthia Perez (A16393492)

Section 1. Proportion of G/G in a population

Downloaded the csv file from Ensembl < https://useast.ensembl.org/Homo_sapiens/Variation/Sample?db=core;v=rs8067378;vdb=variation;vf=959672880#373531_tablePanel >

Here we read this CSV file

```
mxl <- read.csv("373531-SampleGenotypes-Homo_sapiens_Variation_Sample_rs8067378.csv")  
  
head(mxl)
```

	Sample..Male.Female.Unknown.	Genotype..forward.strand.	Population.s.	Father
1	NA19648 (F)	A A	ALL, AMR, MXL	-
2	NA19649 (M)	G G	ALL, AMR, MXL	-
3	NA19651 (F)	A A	ALL, AMR, MXL	-
4	NA19652 (M)	G G	ALL, AMR, MXL	-
5	NA19654 (F)	G G	ALL, AMR, MXL	-
6	NA19655 (M)	A G	ALL, AMR, MXL	-
Mother				
1	-			
2	-			
3	-			
4	-			
5	-			
6	-			

We want to find proportion that are homozygous (G|G) for asthma. Use the `table()` function.

```
table(mxl$Genotype..forward.strand.) / nrow(mxl) *100
```

A A	A G	G A	G G
34.3750	32.8125	18.7500	14.0625

Section 4 Population Scale Analysis

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.

```
#need to read in the file first
expr <- read.table("rs8067378_ENSG00000172057.6.txt", row.names=1)
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

Sample size for each genotype

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

A/A	A/G	G/G
108	233	121

Median Expression levels for each genotype

```
medis <- tapply(expr$exp, expr$geno, median)
medis
```

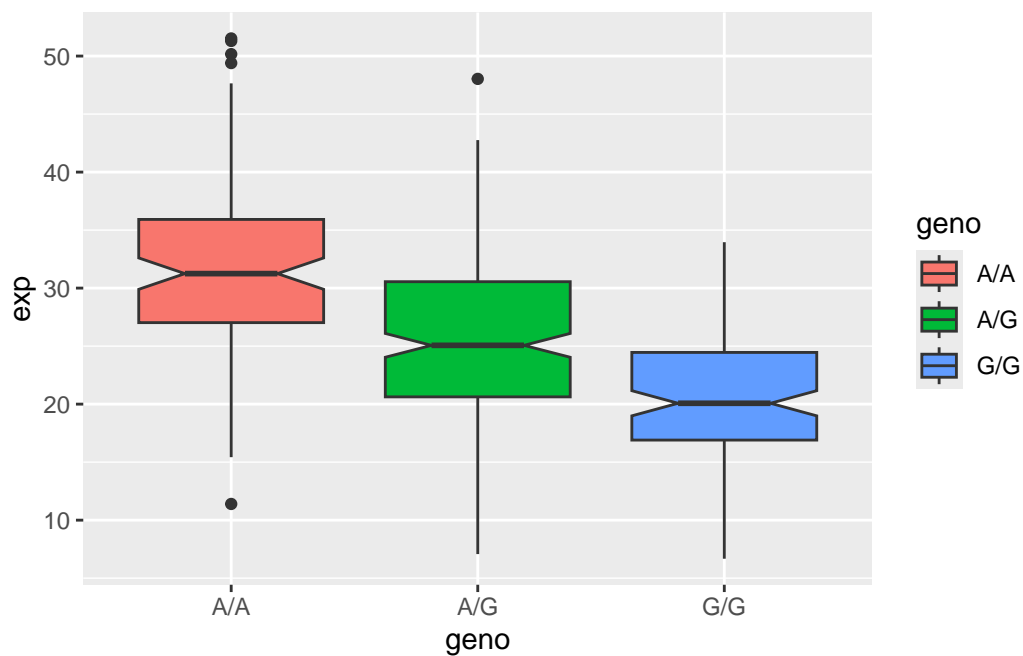
A/A	A/G	G/G
31.24847	25.06486	20.07363

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?

Make a boxplot

```
library (ggplot2)
```

```
genobox <- ggplot(expr) + aes(geno, exp, fill=geno) +  
  geom_boxplot(notch=TRUE)  
genobox
```



G/G has a lower expression level relative to A/A genotype The SNP does effect the expression of ORMDL3.