Class10R

Katelyn Brown

2/17/2022

Examine 1000 Genome Data

Q5. What proportion of the Mexican Ancestry in Los Angeles sample population (MXL) are homozygous for the asthma associated SNP (G|G)?

```
# Read genotype file ofrom Ensembl
mxl <- read.csv("373531-SampleGenotypes-Homo_sapiens_Variation_Sample_rs8067378.csv")
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      -
## 7                   NA19657 (F)                       A|G ALL, AMR, MXL      -
## 8                   NA19658 (M)                       A|A ALL, AMR, MXL      -
## 9                   NA19661 (M)                       A|G ALL, AMR, MXL      -
## 10                  NA19663 (F)                       A|A ALL, AMR, MXL      -
## 11                  NA19664 (M)                       G|A ALL, AMR, MXL      -
## 12                  NA19669 (F)                       A|A ALL, AMR, MXL      -
## 13                  NA19670 (M)                       A|A ALL, AMR, MXL      -
## 14                  NA19676 (M)                       G|G ALL, AMR, MXL      -
## 15                  NA19678 (F)                       A|A ALL, AMR, MXL      -
## 16                  NA19679 (M)                       A|G ALL, AMR, MXL      -
## 17                  NA19681 (F)                       A|G ALL, AMR, MXL      -
## 18                  NA19682 (M)                       A|G ALL, AMR, MXL      -
## 19                  NA19684 (F)                       A|G ALL, AMR, MXL      -
## 20                  NA19716 (F)                       G|A ALL, AMR, MXL      -
## 21                  NA19717 (M)                       A|G ALL, AMR, MXL      -
## 22                  NA19719 (F)                       G|G ALL, AMR, MXL      -
## 23                  NA19720 (M)                       G|G ALL, AMR, MXL      -
## 24                  NA19722 (F)                       G|A ALL, AMR, MXL      -
## 25                  NA19723 (M)                       G|G ALL, AMR, MXL      -
## 26                  NA19725 (F)                       A|G ALL, AMR, MXL      -
## 27                  NA19726 (M)                       A|A ALL, AMR, MXL      -
## 28                  NA19728 (F)                       A|A ALL, AMR, MXL      -
## 29                  NA19729 (M)                       A|G ALL, AMR, MXL      -
## 30                  NA19731 (F)                       A|A ALL, AMR, MXL      -
## 31                  NA19732 (M)                       A|G ALL, AMR, MXL      -
```

```
## 32                    NA19734 (F)                        G|A ALL, AMR, MXL       -
## 33                    NA19735 (M)                        G|G ALL, AMR, MXL       -
## 34                    NA19740 (F)                        A|A ALL, AMR, MXL       -
## 35                    NA19741 (M)                        A|A ALL, AMR, MXL       -
## 36                    NA19746 (F)                        A|A ALL, AMR, MXL       -
## 37                    NA19747 (M)                        G|A ALL, AMR, MXL       -
## 38                    NA19749 (F)                        A|G ALL, AMR, MXL       -
## 39                    NA19750 (M)                        A|G ALL, AMR, MXL       -
## 40                    NA19752 (F)                        A|G ALL, AMR, MXL       -
## 41                    NA19755 (F)                        A|A ALL, AMR, MXL       -
## 42                    NA19756 (M)                        G|A ALL, AMR, MXL       -
## 43                    NA19758 (F)                        A|G ALL, AMR, MXL       -
## 44                    NA19759 (M)                        G|A ALL, AMR, MXL       -
## 45                    NA19761 (F)                        G|A ALL, AMR, MXL       -
## 46                    NA19762 (M)                        A|A ALL, AMR, MXL       -
## 47                    NA19764 (F)                        A|A ALL, AMR, MXL       -
## 48                    NA19770 (F)                        A|G ALL, AMR, MXL       -
## 49                    NA19771 (M)                        A|A ALL, AMR, MXL       -
## 50                    NA19773 (F)                        A|A ALL, AMR, MXL       -
## 51                    NA19774 (M)                        A|G ALL, AMR, MXL       -
## 52                    NA19776 (F)                        A|G ALL, AMR, MXL       -
## 53                    NA19777 (M)                        A|A ALL, AMR, MXL       -
## 54                    NA19779 (F)                        G|A ALL, AMR, MXL       -
## 55                    NA19780 (M)                        A|A ALL, AMR, MXL       -
## 56                    NA19782 (F)                        G|A ALL, AMR, MXL       -
## 57                    NA19783 (M)                        A|G ALL, AMR, MXL       -
## 58                    NA19785 (F)                        A|A ALL, AMR, MXL       -
## 59                    NA19786 (M)                        G|A ALL, AMR, MXL       -
## 60                    NA19788 (F)                        A|G ALL, AMR, MXL       -
## 61                    NA19789 (M)                        G|G ALL, AMR, MXL       -
## 62                    NA19792 (M)                        A|A ALL, AMR, MXL       -
## 63                    NA19794 (F)                        G|A ALL, AMR, MXL       -
## 64                    NA19795 (M)                        A|G ALL, AMR, MXL       -
##     Mother
## 1        -
## 2        -
## 3        -
## 4        -
## 5        -
## 6        -
## 7        -
## 8        -
## 9        -
## 10       -
## 11       -
## 12       -
## 13       -
## 14       -
## 15       -
## 16       -
## 17       -
## 18       -
## 19       -
## 20       -
```

```
## 21          -
## 22          -
## 23          -
## 24          -
## 25          -
## 26          -
## 27          -
## 28          -
## 29          -
## 30          -
## 31          -
## 32          -
## 33          -
## 34          -
## 35          -
## 36          -
## 37          -
## 38          -
## 39          -
## 40          -
## 41          -
## 42          -
## 43          -
## 44          -
## 45          -
## 46          -
## 47          -
## 48          -
## 49          -
## 50          -
## 51          -
## 52          -
## 53          -
## 54          -
## 55          -
## 56          -
## 57          -
## 58          -
## 59          -
## 60          -
## 61          -
## 62          -
## 63          -
## 64          -
```

```r
# Use table() to answer question
mxl$Genotype..forward.strand.
```

```
##  [1] "A|A" "G|G" "A|A" "G|G" "G|G" "A|G" "A|G" "A|A" "A|G" "A|A" "G|A" "A|A"
## [13] "A|A" "G|G" "A|A" "A|G" "A|G" "A|G" "A|G" "G|A" "A|G" "G|G" "G|G" "G|A"
## [25] "G|G" "A|G" "A|A" "A|A" "A|G" "A|A" "A|G" "G|A" "G|G" "A|A" "A|A" "A|A"
## [37] "G|A" "A|G" "A|G" "A|G" "A|A" "G|A" "A|G" "G|A" "G|A" "A|A" "A|A" "A|G"
## [49] "A|A" "A|A" "A|G" "A|G" "A|A" "G|A" "A|A" "G|A" "A|G" "A|A" "G|A" "A|G"
## [61] "G|G" "A|A" "G|A" "A|G"
```

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

```
##
##      A|A      A|G      G|A      G|G
## 0.343750 0.328125 0.187500 0.140625
```

14% of mxl are homozygous for the asthma associated SNP (G|G).

Examine GBR data set for the same SNP to compare populations.

```
gbr <- read.csv("373522-SampleGenotypes-Homo_sapiens_Variation_Sample_rs8067378.csv")
gbr
```

```
##     Sample..Male.Female.Unknown. Genotype..forward.strand. Population.s. Father
## 1                   HG00096 (M)                       A|A ALL, EUR, GBR      -
## 2                   HG00097 (F)                       G|A ALL, EUR, GBR      -
## 3                   HG00099 (F)                       G|G ALL, EUR, GBR      -
## 4                   HG00100 (F)                       A|A ALL, EUR, GBR      -
## 5                   HG00101 (M)                       A|A ALL, EUR, GBR      -
## 6                   HG00102 (F)                       A|A ALL, EUR, GBR      -
## 7                   HG00103 (M)                       A|G ALL, EUR, GBR      -
## 8                   HG00105 (M)                       A|A ALL, EUR, GBR      -
## 9                   HG00106 (F)                       G|A ALL, EUR, GBR      -
## 10                  HG00107 (M)                       G|G ALL, EUR, GBR      -
## 11                  HG00108 (M)                       A|A ALL, EUR, GBR      -
## 12                  HG00109 (M)                       G|G ALL, EUR, GBR      -
## 13                  HG00110 (F)                       A|G ALL, EUR, GBR      -
## 14                  HG00111 (F)                       A|A ALL, EUR, GBR      -
## 15                  HG00112 (M)                       G|G ALL, EUR, GBR      -
## 16                  HG00113 (M)                       G|G ALL, EUR, GBR      -
## 17                  HG00114 (M)                       G|A ALL, EUR, GBR      -
## 18                  HG00115 (M)                       A|G ALL, EUR, GBR      -
## 19                  HG00116 (M)                       G|G ALL, EUR, GBR      -
## 20                  HG00117 (M)                       A|A ALL, EUR, GBR      -
## 21                  HG00118 (F)                       G|G ALL, EUR, GBR      -
## 22                  HG00119 (M)                       G|A ALL, EUR, GBR      -
## 23                  HG00120 (F)                       G|G ALL, EUR, GBR      -
## 24                  HG00121 (F)                       A|G ALL, EUR, GBR      -
## 25                  HG00122 (F)                       G|G ALL, EUR, GBR      -
## 26                  HG00123 (F)                       G|A ALL, EUR, GBR      -
## 27                  HG00125 (F)                       A|G ALL, EUR, GBR      -
## 28                  HG00126 (M)                       G|G ALL, EUR, GBR      -
## 29                  HG00127 (F)                       G|A ALL, EUR, GBR      -
## 30                  HG00128 (F)                       A|G ALL, EUR, GBR      -
## 31                  HG00129 (M)                       G|G ALL, EUR, GBR      -
## 32                  HG00130 (F)                       A|G ALL, EUR, GBR      -
## 33                  HG00131 (M)                       G|G ALL, EUR, GBR      -
## 34                  HG00132 (F)                       A|A ALL, EUR, GBR      -
## 35                  HG00133 (F)                       G|A ALL, EUR, GBR      -
## 36                  HG00136 (M)                       G|G ALL, EUR, GBR      -
## 37                  HG00137 (F)                       G|A ALL, EUR, GBR      -
## 38                  HG00138 (M)                       A|A ALL, EUR, GBR      -
## 39                  HG00139 (M)                       G|G ALL, EUR, GBR      -
```

```
## 40            HG00140 (M)              G|A ALL, EUR, GBR      -
## 41            HG00141 (M)              G|G ALL, EUR, GBR      -
## 42            HG00142 (M)              G|G ALL, EUR, GBR      -
## 43            HG00143 (M)              G|A ALL, EUR, GBR      -
## 44            HG00145 (M)              A|A ALL, EUR, GBR      -
## 45            HG00146 (F)              A|A ALL, EUR, GBR      -
## 46            HG00148 (M)              G|A ALL, EUR, GBR      -
## 47            HG00149 (M)              G|A ALL, EUR, GBR      -
## 48            HG00150 (F)              G|A ALL, EUR, GBR      -
## 49            HG00151 (M)              G|A ALL, EUR, GBR      -
## 50            HG00154 (F)              G|G ALL, EUR, GBR      -
## 51            HG00155 (M)              A|G ALL, EUR, GBR      -
## 52            HG00157 (M)              A|A ALL, EUR, GBR      -
## 53            HG00158 (F)              A|A ALL, EUR, GBR      -
## 54            HG00159 (M)              A|A ALL, EUR, GBR      -
## 55            HG00160 (M)              A|A ALL, EUR, GBR      -
## 56            HG00231 (F)              A|G ALL, EUR, GBR      -
## 57            HG00232 (F)              G|G ALL, EUR, GBR      -
## 58            HG00233 (F)              G|G ALL, EUR, GBR      -
## 59            HG00234 (M)              G|G ALL, EUR, GBR      -
## 60            HG00235 (F)              A|A ALL, EUR, GBR      -
## 61            HG00236 (F)              A|A ALL, EUR, GBR      -
## 62            HG00237 (F)              A|A ALL, EUR, GBR      -
## 63            HG00238 (F)              G|G ALL, EUR, GBR      -
## 64            HG00239 (F)              G|A ALL, EUR, GBR      -
## 65            HG00240 (F)              G|A ALL, EUR, GBR      -
## 66            HG00242 (M)              G|A ALL, EUR, GBR      -
## 67            HG00243 (M)              A|G ALL, EUR, GBR      -
## 68            HG00244 (M)              G|A ALL, EUR, GBR      -
## 69            HG00245 (F)              A|G ALL, EUR, GBR      -
## 70            HG00246 (M)              A|G ALL, EUR, GBR      -
## 71            HG00250 (F)              G|G ALL, EUR, GBR      -
## 72            HG00251 (M)              G|A ALL, EUR, GBR      -
## 73            HG00252 (M)              G|A ALL, EUR, GBR      -
## 74            HG00253 (F)              A|A ALL, EUR, GBR      -
## 75            HG00254 (F)              A|G ALL, EUR, GBR      -
## 76            HG00255 (F)              A|G ALL, EUR, GBR      -
## 77            HG00256 (M)              A|G ALL, EUR, GBR      -
## 78            HG00257 (F)              G|G ALL, EUR, GBR      -
## 79            HG00258 (F)              A|A ALL, EUR, GBR      -
## 80            HG00259 (F)              G|A ALL, EUR, GBR      -
## 81            HG00260 (M)              G|G ALL, EUR, GBR      -
## 82            HG00261 (F)              G|G ALL, EUR, GBR      -
## 83            HG00262 (F)              A|A ALL, EUR, GBR      -
## 84            HG00263 (F)              G|A ALL, EUR, GBR      -
## 85            HG00264 (M)              A|G ALL, EUR, GBR      -
## 86            HG00265 (M)              G|G ALL, EUR, GBR      -
## 87            HG01334 (M)              A|G ALL, EUR, GBR      -
## 88            HG01789 (M)              G|A ALL, EUR, GBR      -
## 89            HG01790 (F)              G|A ALL, EUR, GBR      -
## 90            HG01791 (M)              A|A ALL, EUR, GBR      -
## 91            HG02215 (F)              G|G ALL, EUR, GBR      -
##     Mother
## 1       -
```

```
## 2            -
## 3            -
## 4            -
## 5            -
## 6            -
## 7            -
## 8            -
## 9            -
## 10           -
## 11           -
## 12           -
## 13           -
## 14           -
## 15           -
## 16           -
## 17           -
## 18           -
## 19           -
## 20           -
## 21           -
## 22           -
## 23           -
## 24           -
## 25           -
## 26           -
## 27           -
## 28           -
## 29           -
## 30           -
## 31           -
## 32           -
## 33           -
## 34           -
## 35           -
## 36           -
## 37           -
## 38           -
## 39           -
## 40           -
## 41           -
## 42           -
## 43           -
## 44           -
## 45           -
## 46           -
## 47           -
## 48           -
## 49           -
## 50           -
## 51           -
## 52           -
## 53           -
## 54           -
## 55           -
```

```
## 56          -
## 57          -
## 58          -
## 59          -
## 60          -
## 61          -
## 62          -
## 63          -
## 64          -
## 65          -
## 66          -
## 67          -
## 68          -
## 69          -
## 70          -
## 71          -
## 72          -
## 73          -
## 74          -
## 75          -
## 76          -
## 77          -
## 78          -
## 79          -
## 80          -
## 81          -
## 82          -
## 83          -
## 84          -
## 85          -
## 86          -
## 87          -
## 88          -
## 89          -
## 90          -
## 91          -
```

```
table(gbr$Genotype..forward.strand.)/nrow(gbr)
```

```
##
##       A|A       A|G       G|A       G|G
## 0.2527473 0.1868132 0.2637363 0.2967033
```

About 29% of gbr are homozygous for the same SNP.

Expression by Genotype Analysis

I want to read my RNA-Seq expression results into R. This file is not a CSV but rather has fields separated by space. Because of this, use `read.table()`

```
x <- read.table("rs8067378_ENSG00000172057.6.txt")
head(x)
```

```
##     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
```

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.

Is the mean expression different based on genotype?

```
table(x$geno)
```

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

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

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

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

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

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

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

Based on the summary statistics of each genotype, the median expression levels are 31.25 for A/A, 25.065 for A/G, and 20.074 for G/G.

## Make a summary overview figure.

Use ggplot to make a boxplot with data.

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?

```
library(ggplot2)
```

```
ggplot(x) + aes(x = geno, y = exp, fill = geno) + geom_boxplot(notch = TRUE) + labs(title = "Genotype V
```

Genotype Versus Gene Expression
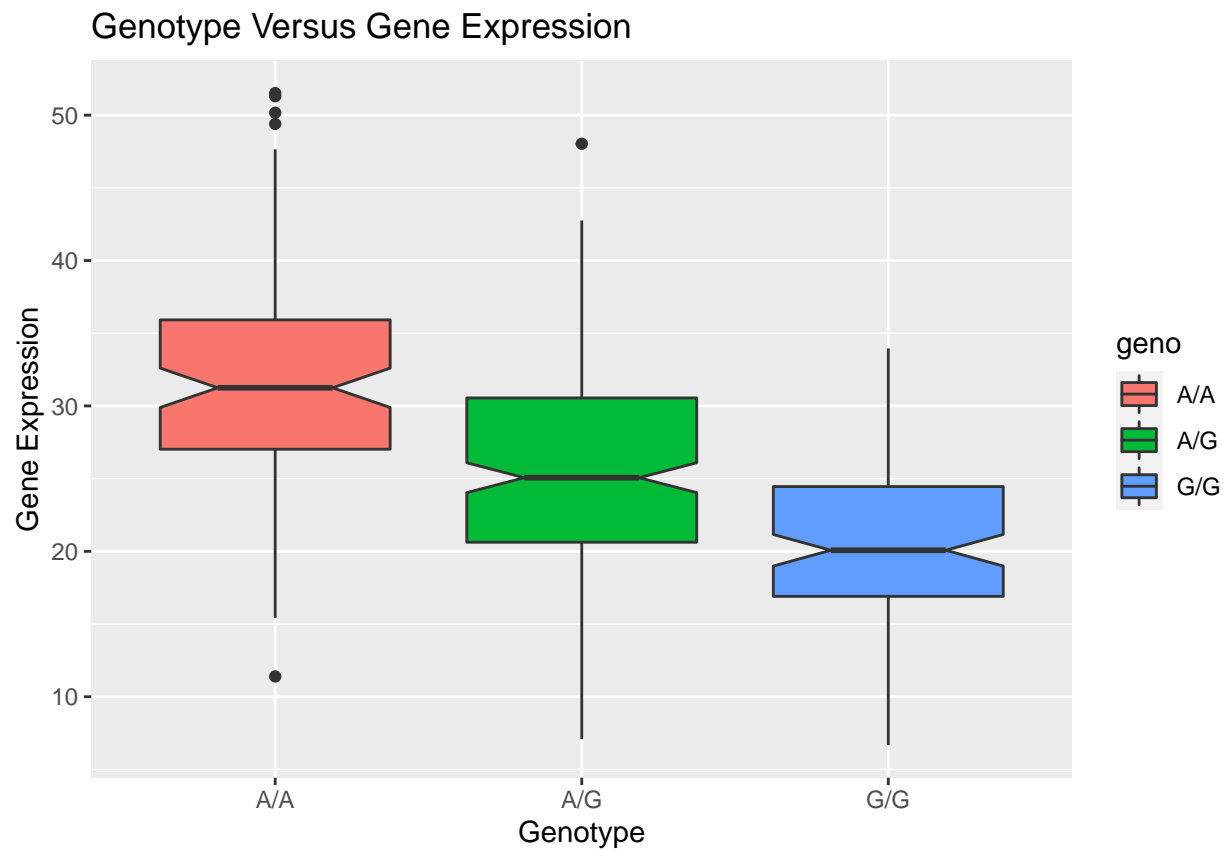
Based on the boxplot, the A/A genotype is correlated with a higher gene expression level compared to both A/G and G/G.