

Genotype.freq.&.HWE

Rebekah Wrigley

27/01/2015

```
# Read in file
NZ_FACC_allg_2 <- read.delim("~/ABCG2/Pleiotropic-effects-of-ABCG2-studentship-1/NZ_FACC_allg_2.txt",
```

```
# Minor allele frequency function
MAF<- function(maj, het, minor){
  results<- c(NA, NA)
  names(results)<- c("n", "MAF")
  results[1]<- het+(minor*2) # Get number of minor alleles
  results[2]<- (het+(minor*2))/((maj+het+minor)*2)*100 # Minor allele frequency
  results
}

## MAF for one dataset/group
round(MAF(263, 42, 1), digits = 2)
```

```
##      n    MAF
## 44.00  7.19
```

```
# MAF for multiple subgroups
mytable<- table(NZ_FACC_allg_2$GU_stat, NZ_FACC_allg_2$RS2231142)
colnames(mytable)<- c("maj", "het", "minor")
mytable
```

```
##
##      maj  het minor
##   G  883  474    73
##   H 1791  583    39
##   N 9514 2070   125
```

```
t(round(mapply(MAF, mytable[, "maj"], mytable[, "het"], mytable[, "minor"]), digits=2))
```

```
##      n    MAF
## G  620 21.68
## H  661 13.70
## N 2320  9.91
```

```
#Genotype frequencies
gen_freq<- function(maj, het, minor){
  results<- c(NA, NA, NA)
  names(results)<- c("maj_freq", "het_freq", "min_freq")
  results[1]<- maj/(maj+het+minor)*100
  results[2]<- het/(maj+het+minor)*100
  results[3]<- minor/(maj+het+minor)*100
  results
}
```

```
}
```

```
t(round(mapply(gen_freq, mytable[, "maj"], mytable[, "het"], mytable[, "minor"]), digits=2))
```

```
##   maj_freq het_freq min_freq
## G    61.75    33.15     5.10
## H    74.22    24.16     1.62
## N    81.25    17.68     1.07
```

```
# Combined genotype frequencies where there are four combined genotype categories
```

```
y<- table(NZ_FACC_allg_2$MYGOUTSUM, NZ_FACC_allg_2$abcg_g_com) # abcg_g_com is the combined genotype c
```

```
colnames(y)<- c("A", "B", "C", "D")
```

```
rownames(y)<- c("control", "case")
```

```
y
```

```
##
```

```
##           A      B      C      D
## control 4102 7389  257 2613
## case    154  393   33  313
```

```
com_freq<- function(A, B, C, D){
  results<- c(NA, NA, NA, NA)
  names(results)<- c("freq_1", "freq_2", "freq_3", "freq_4")
  results[1]<- A/(A+B+C+D)*100
  results[2]<- B/(A+B+C+D)*100
  results[3]<- C/(A+B+C+D)*100
  results[4]<- D/(A+B+C+D)*100
  results
}
```

```
round(mapply(com_freq, y[, "A"], y[, "B"], y[, "C"], y[, "D"]), digits=2)
```

```
##           control case
## freq_1    28.56 17.25
## freq_2    51.45 44.01
## freq_3     1.79  3.70
## freq_4    18.20 35.05
```