# TP 4

(from Thornton T.)

1. Read in the variants within "Gene1.txt" into R. Calculate the minor allele frequencies of the variants and plot a histogram. What do you notice ?
2. Read in the trait value in the file "Trait1.txt". Then test for an association between each variant in Gene1 and the Trait1. Is anything significant after adjusting for multiple testing ?
3. Test for an association between the rare variants in Gene1 and the quantitative trait. For now, let's define rare variants to be the variants with MAF $< 3\%$. Specifically, apply :
   - CAST (Binary collapsing approach)
   - MZ Test/GRANVIL (Count based collapsing)
   - Weighted Count Based Collapsing where weights can depend on MAF
   - SKAT (Sequence Kernel Association Test)
     - standard SKAT (default parameters)
     - omnibus SKAT
       - with different $\rho$ values (0, 1, 0.5, ...)
       - with "optimal" $\rho$ value
4. Redo the previous question using the variants inside of "Gene2.txt" and "Trait2.txt".
5. Let's consider alternative definitions for rare variant : MAF $< 5\%$, $< 3\%$, $< 1\%$, $< 0.5\%$.