First R code: MAF and MR calculation

10-24-2023

# Path normalization

Note that you have to learn how to use markdown for the text chunk

knitr::opts\_chunk$set(echo = TRUE)   
knitr::opts\_knit$set(root.dir=normalizePath('../../'))

data <- read.csv("data/ng\_rice\_snp.csv")  
  
#Calculating the missing rate  
missing\_count = rowSums(data[, -1] == "-") #adding all the occurances of "-" per each row.  
total\_individuals = ncol(data) - 1 #calculating number of individual samples we have.  
missing\_rate = missing\_count / total\_individuals  
data$missing\_rate = missing\_rate#creating a new column missing rate back in sheet.  
  
#calculating MAF  
#creating a function to calculate MAF  
calculate\_maf = function(row) {  
 #Exclude the marker name and missing data  
 genotypes = row[!is.na(row) & row != "-"]  
 #Count of each allele (each '1' counts as one of each allele, '0' and '2' count as two of one allele)  
 ref\_allele = sum(genotypes == 0) \* 2 + sum(genotypes == 1)  
 alt\_allele = sum(genotypes == 2) \* 2 + sum(genotypes == 1)  
 #Total alleles  
 total\_alleles = ref\_allele + alt\_allele  
 #Frequencies  
 ref\_freq = ref\_allele / total\_alleles  
 alt\_freq = alt\_allele / total\_alleles  
 return(min(ref\_freq, alt\_freq)) #Return the minor allele frequency  
}  
#Appling the function to each row and store the result in the maf  
maf = apply(data[, -1], 1, calculate\_maf) # excluding the first column with marker names and looping over each row.  
data$maf = maf #updating new column with maf in excell sheet.  
#Saving the data frame with the new 'missing\_rate' and 'maf' columns to a new CSV file  
write.csv(data, file = "cache/updated\_ng\_rice\_snp.csv", row.names = FALSE, quote = FALSE, )

Add a new chunk by clicking the *Insert Chunk* button on the toolbar or by pressing *Cmd+Option+I*.

When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the *Preview* button or press *Cmd+Shift+K* to preview the HTML file).

The preview shows you a rendered HTML copy of the contents of the editor. Consequently, unlike *Knit*, *Preview* does not run any R code chunks. Instead, the output of the chunk when it was last run in the editor is displayed.