Below is a step by step how-to guide to using the cervus\_prep.R function and Cervus to assign paternity. cervus\_prep.R takes a vector of squirrel\_ids as input and returns the two files required by the program Cervus. The first is an ID file of squirrel\_ids (Cervus refers to this as the Offspring file), which has the user-defined individuals of interest in column 1, their dams (if known) in column 2, and a list of candidate sires in the remainder columns. The second file contains the genotypes for every squirrel that appears in the ID file.

1. The only input to cervus\_prep() is a vector of squirrel\_ids that can be coerced to type numeric (i.e. anything other than a factor). Following are a few examples of how this vector may be created.

• establish a connection to the krsp database and extract squirrel\_ids:

library(krsp)

con <- krsp\_connect(host =

"krsp.cepb5cjvqban.us-east-2.rds.amazonaws.com",

dbname = "krsp",

username = Sys.getenv("krsp\_user"),

password = Sys.getenv("krsp\_password"))

flastall <- tbl(con, "flastall2") %>%

collect() %>%

sample\_n(30) %>% # filter() as needed

select(squirrel\_id, dam\_id, byear)

flastall$squirrel\_id

• read in a vector from a text file:

juv\_ids <- scan(file = "your\_filename.txt", sep = "\t")

• define IDs manually in R

juv\_ids <- c(6373, 13344, 12806, 19933)

2. Pass your IDs (juv\_ids) to the function and store the results in a new variable.

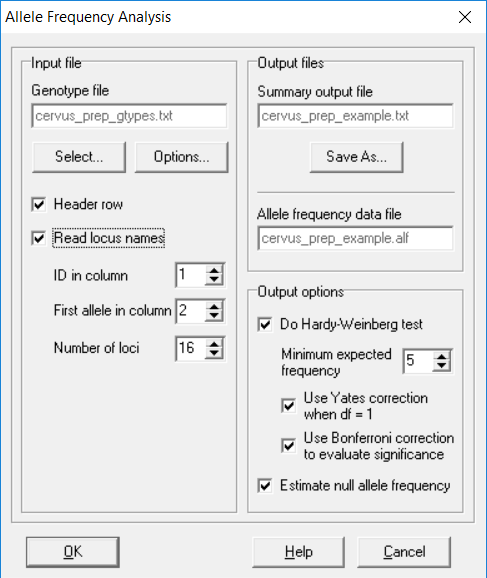
cervus\_inputs <- cervus\_prep(juv\_ids)

3. The result is a list of the two files described above named "IDs" and "gtypes". Save them to your working directory for use with Cervus.

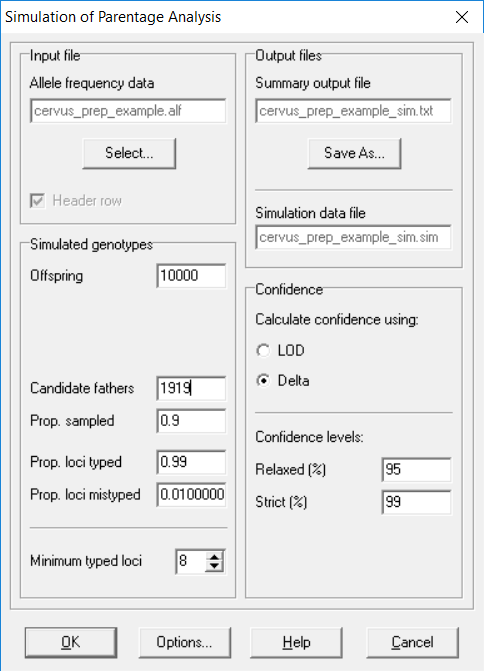
write.table(cervus\_inputs$IDs, file = "cervus\_prep\_ids.txt", sep = "t", row.names = F, quote = F)

write.table(cervus\_inputs$gtypes, file = "cervus\_prep\_gtypes.txt",

sep = "\t", row.names = F, quote = F)

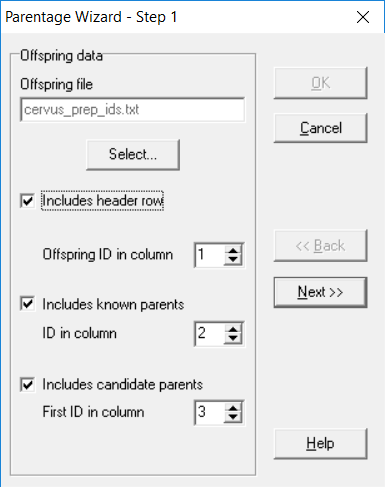


4. Start up Cervus and run the Allele Frequency Analysis under the Analysis tab. Set the genotype file option to your cervus\_prep() gtypes output and make sure the Header row and Read locus names boxes are checked. ID is in column 1, the first allele is in column 2, and there are 16 loci. Name your summary output file and click OK.



5. Run simulation of paternity; select Analysis > Simulation of Parentage Analysis > Paternity. The .alf file produced in the last step should automatically populate the Allele Frequency Data field. The parameters can be set as follows:

* Offspring = 1000
* Candidate fathers = number of males in analysis
* Prop. Sampled = 0.9
* Prop. Loci typed = [set in Allele freq analysis]
* Prop. Loci mistyped = 0.01
* Calc. confidence = delta
* Relaxed = 95
* Strict = 99

6. Run the paternity analysis; Analysis > Parentage analysis > Paternity. Offspring ID is in col 1, known parent (dam) ID is in col 2, and candidate parent ID is in col 3; click Next. The next 3 windows should be automatically filled from the preceding analyses. In the Output window (Step 5), name the output file and include the two most-likely parents by selecting it from the dropdown menu. Click OK. The .csv file contains the pedigree and confidence information.