-All samples sequenced on Illumina Novaseq 6000 using 150 bp, paired-end reads

-Samples with coverage < 5x were re-sequenced on a second round/date

- RD1Sequencing\_SampleKey.csv and RD2Sequencing\_SampleKey.csv contain information regarding .fastq file names

-BioinformaticsScripts: contains scripts used for read trimming, quality control, alignment, and haplotype-derived allele frequency estimation.

-RData: consists of 4 data frames:

-samps: sample information whereby row order corresponds to column order of the allele frequency matrix and effective coverage matrix.

-afmat: numeric data frame containing all haplotype-derived allele freqeuncies

- eec: numeric data frame containing estimated effective coverage for each sample/site

-sites: dataframe containing chromosome and site information (corresponding to rows of afmat and eec\_