**Convert the .vcf file output from populations to a GESTE / BayeScan file for Outlier analysis in Bayescan**

***On the maccolllab.life.nottingham.ac.uk server:***

cp /home/mbzlld/lauras\_files/Post-doc\_work\_2018/RAD\_analysis/my\_RAD\_analysis/Populations\_analysis/{batch\_1.vcf,batch\_1.haplotypes.vcf} /home/mbzlld/lauras\_files/Post-doc\_work\_2018/RAD\_analysis/my\_RAD\_analysis/PGDspider\_file\_conversion

cd /home/mbzlld/lauras\_files/Post-doc\_work\_2018/RAD\_analysis/my\_RAD\_analysis/PGDspider\_file\_conversion

vcftools \

--vcf batch\_1.haplotypes.vcf \

--remove-indv SCAD130228 \

--remove-indv SCAD130229 \

--remove-indv SCAD130231 \

--remove-indv SCAD130232 \

--remove-indv SCAD130233 \

--remove-indv SCAD130234 \

--remove-indv SCAD130236 \

--remove-indv SCAD130237 \

--remove-indv SCAD130238 \

--remove-indv SCAD130240 \

--remove-indv SCAD130242 \

--remove-indv SCAD130245 \

--remove-indv SCAD130246 \

--remove-indv SCAD130249 \

--remove-indv SCAD130255 \

--remove-indv SCAD130256 \

--remove-indv SCAD130257 \

--recode \

--recode-INFO-all \

--out batch\_1.haplotypes\_pairs\_only

java -Xmx1024m -Xms512m -jar PGDSpider2-cli.jar \

-inputfile batch\_1.haplotypes\_pairs\_only.recode.vcf \

-inputformat VCF \

-outputfile batch\_1.haplotypes\_pairs\_only.bayescan \

-outputformat GESTE\_BAYE\_SCAN \

-spid VCF\_to\_Bayescan.spid

cp /home/mbzlld/lauras\_files/Post-doc\_work\_2018/RAD\_analysis/my\_RAD\_analysis/PGDspider\_file\_conversion/batch\_1.haplotypes\_pairs\_only.bayescan /home/mbzlld/lauras\_files/Post-doc\_work\_2018/RAD\_analysis/my\_RAD\_analysis/BayeScan\_analysis

cd /home/mbzlld/lauras\_files/Post-doc\_work\_2018/RAD\_analysis/my\_RAD\_analysis/BayeScan\_analysis

screen -RD

./source/bayescan\_2.1 \

batch\_1.haplotypes\_pairs\_only.bayescan -snp \

-od /home/mbzlld/lauras\_files/Post-doc\_work\_2018/RAD\_analysis/my\_RAD\_analysis/BayeScan\_analysis \

-threads 8 \

-n 5000 \

-thin 10 \

-nbp 20 \

-pilot 5000 \

-burn 50000 \

-pr\_odds 10 \

-lb\_fis 0 \

-hb\_fis 1 \

-aflp\_pc 0.1 \

-out\_freq

CNTRL A D

***On own laptop:***

The output file: batch\_1.haplotypes\_pairs\_only.baye\_fst has the information for plotting the results of the BayeScan analysis. Transfer it onto your own laptop using filezilla and convert it to a .csv file by opening in Excel. The resulting batch\_1.haplotypes\_pairs\_only.baye\_fst.csv file is in: C:\Users\mbzlld\Google Drive\Post Doc\RAD data analysis\BayeScan on my laptop.

This file has lost the information about chromosome locations so we must extract this information from the original VCF file that was converted to bayescan format by PGDspider for this analysis. Open the species pairs only VCF file (filtered by VCFtools) from the PGDspider folder in excel and delete the first 8 info columns. Delete all columns after E. Save the file as a chromosome key file (a .csv file) to open in R and close it in excel.

Follow the R code for Bayescan graph production.R script saved in the C:\Users\mbzlld\Google Drive\Post Doc\RAD data analysis\BayeScan folder on my laptop.