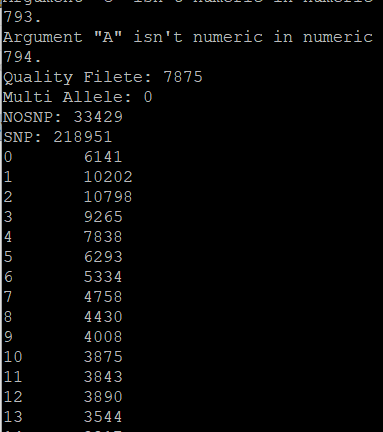
群体分类：PCA，系统发育树（phylogenetic tree）

structure analysis k为群体中有多少个物种

plink /data/part2/software/plink-v1.90/plink --vcf scaffold37\_cov106.vcf.gz --pca --allow-extra-chr --out scaffold37\_cov106

less scaffold37\_cov106.eigenval 解释主成分分析

admixture 只能识别snp位点

输完脚本之后，

perl filter.pl scaffold37\_cov106.vcf.gz output.vcf.gz

/data/part2/software/plink-v1.90/plink --vcf output.vcf.gz --allow-extra-chr --out output

将bim文件第一列都改成1，之后运行（改数字，分别为1,2,3,4,5,6,）

/data/part2/software/admixture\_linux-1.3.0/admixture --cv -j12 output.bed 3 | tee output.bed.3.log

生成nj树：

/data/part2/software/plink-v1.90/plink --vcf out.vcf.gz --make-bed --allow-extra-chr

/data/part2/software/plink-v1.90/plink --bfile plink --allow-extra-chr --distance-matrix

/data/part2/software/treebest-1.9.2/treebest nj -b 100 plink.mdist.id\_new > scaffold37\_cov106.nj

过滤物种：

/data/part2/software/vcftools\_0.1.12b/bin/vcftools --gzvcf scaffold37\_cov106.vcf.gz --keep test.list --recode --recode-INFO-all --out test

压缩为gz文件，gzip 文件名