Genetics QC

Muhammad Shoaib

2023-11-07

```
plink \
--bfile /hpf/projects/arnold/users/Shoaib1/EUR-QCNEWW/EUR.QCdup \
--maf 0.01 \
--hwe 1e-6 \
--geno 0.03 \
--mind 0.03 \
--write-snplist \
--make-just-fam \
--out EUR.QC

plink \
--bfile /hpf/projects/arnold/users/Shoaib1/EUR-QCNEWW/EUR.QCdup \
--keep EUR.QC.fam \
--extract EUR.QC.snplist \
--indep-pairwise 200 50 0.25 \
--out EUR.QC


plink \
--bfile /hpf/projects/arnold/users/Shoaib1/EUR-QCNEWW/EUR.QCdup \
--extract EUR.QC.prune.in \
--keep EUR.QC.fam \
--het \
--out EUR.QC

library(data.table)
dat <- fread("EUR.QC.het")
valid <- dat[F<=mean(F)+3*sd(F) & F>=mean(F)-3*sd(F)]
fwrite(valid[,c("FID","IID")], "EUR.valid.sample", sep="\t")

plink \
--bfile /hpf/projects/arnold/users/Shoaib1/EUR-QCNEWW/EUR.QCdup \
--keep EUR.valid.sample \
--check-sex \
--out EUR.QC


library(data.table)
valid <- read.table("EUR.valid.sample", header=T)
dat <- read.table("EUR.QC.sexcheck", header=T)
valid <- subset(dat, STATUS=="OK" & FID %in% valid$FID)
write.table(valid[,c("FID", "IID")], "EUR.QC.valid", row.names=F, col.names=F, sep="\t", quote=F)
```

```
plink \
--bfile /hpf/projects/arnold/users/Shoaib1/EUR-QCNEWW/EUR.QCdup \
--make-bed \
--extract EUR.QC.prune.in \
--keep EUR.QC.valid \
--out EUR.QC

plink \
--bfile /hpf/projects/arnold/users/Shoaib1/EUR-QCNEWW/EUR.QC \
--make-bed \
--out EUR.QCn \
--extract EUR.QC.snplist \

plink \
--bfile /hpf/projects/arnold/users/Shoaib1/EUR-QCNEWW/EUR.QCn \
--make-bed \
--remove rel.txt \
--out EUR.QCnn
```

Relatedness using KING software

```
module add king/2.3.0
# KING command to estimate kinship coefficients
king -b /hpf/projects/arnold/data/genotypes/OSC_HumanCoreExome_EAS_recall/recoded/osc_eas.bed
--kinship --prefix kinship_output
king -b /hpf/projects/arnold/users/Shoaib1/EUR-QCNEWW/EUR.QCn.bed --related --degree 2
--prefix relatedEUR_output

# KING command to generate a relatedness report
king -b your_data.bed --related --degree 2 --prefix relatedness_output
```