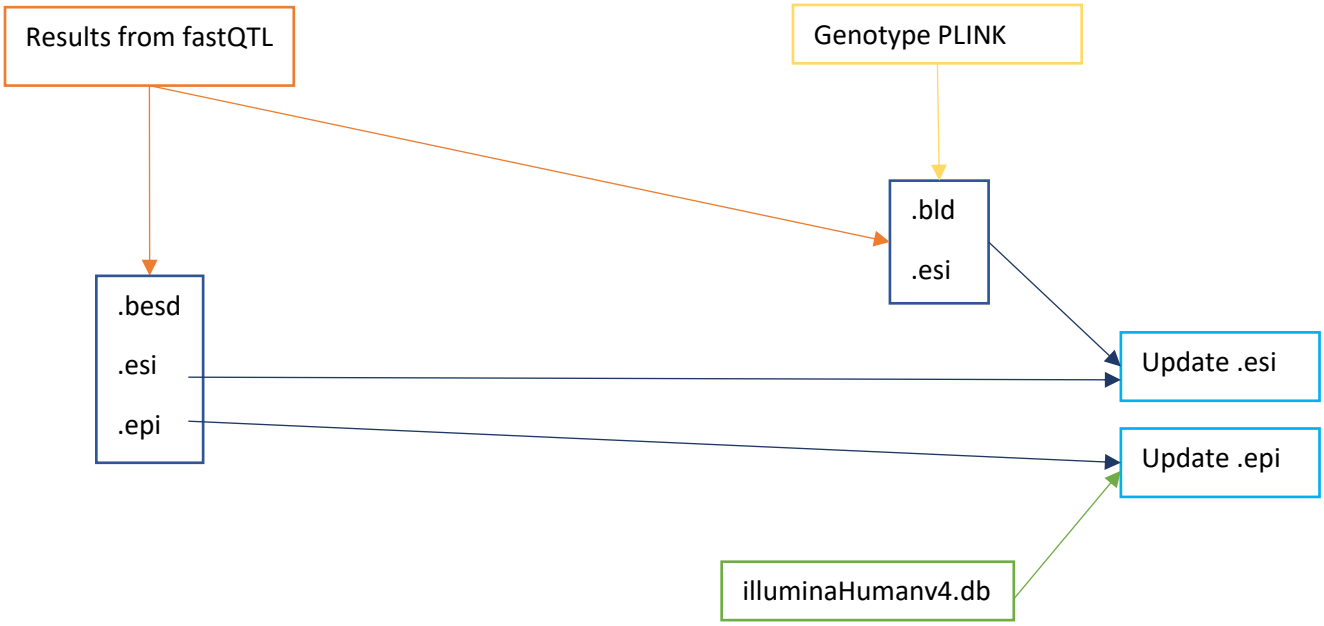
Summary-data-based Mendelian Randomization (SMR)



We need 3 inputs:

- PLINK file .bim (genotipat).
- GWAS statistical summary.
- Read summary data from a eQTL study in binary format:
  - .esi (same format as .bim PLINK file) → contain SNP information.
  - .epi → contain probe information.
  - .besd → results from fastQTL.

GWAS statistical summary

#Transform our file in the format that we need: SNP A1 A2 freq b se p n

```
awk '{print $2, "\t", $5, "\t", $6, "\t", $9, "\t", $10, "\t", $11, "\t", $12, "\t", $7}' IMPUT_CD_GWAS_QC.txt > GWAS_preSUMMARY.txt
```

#Change header names

```
sed -i 's/P/p/g' GWAS_preSUMMARY.txt

sed -i 's/N/n/g' GWAS_preSUMMARY.txt

sed -i 's/RS/SNP/g' GWAS_preSUMMARY.txt

sed -i 's/beta/b/g' GWAS_preSUMMARY.txt

sed -i 's/SE/se/g' GWAS_preSUMMARY.txt

sed -i 's/MAF_controls/freq/g' GWAS_preSUMMARY.txt
```

```
sed -i 's/ //g' GWAS_preSUMMARY.txt
```

#Compare list of SNPs in GWAS summary and .esi

```
cut -f1 GWAS_preSUMMARY.txt > SNPs.txt
```

#R:

```
esi <- read.table("naive/SMR_files/naive.esi", header = FALSE)
gwas <- read.table("SNPs.txt", header=TRUE)
length(setdiff(esi$V2, gwas$SNP))
```

#only 1.899 SNPs from fastQTL results (13798 SNPs unique), are not found in GWAS summary

#Obtain final GWAS file from immunochip

```
mv GWAS_preSUMMARY.txt IMMUNO_INP_GWAS_QC.ma
```

NAIVE

1. **Obtain .besd file.**

#Remove probes without associations, we save 14.502 probes.

```
grep -vw "NA" permutations.all.chunks.txt > permut_naive.txt
```

#Save NA probes.

```
grep "NA" permutations.all.chunks.txt > NA_naive.txt
```

#Obtain .besd file

```
smr_Win/SMR_x86.exe --eqtl-summary permut_naive.txt --fastqtl-permu-format --make-besd --out naïve
```

2. **Obtain .epi file.**

#For this we used R and its package illuminaHumanv4.db, we get: chr, start, strand (*naive_epi.R*)

```
awk 'NR>1 {print $1, "\t", $3, "\t", 0, "\t", $2, "\t", $3, "\t", $4}' probes_info_naive.txt > naive_update.epi
```

#Update original epi coming from .besd

```
smr_Win/SMR_x86.exe --beqtl-summary naive --update-epi naive_update.epi
```

3. **Obtain .esi files (BLD format).**

```
smr_Win/SMR_x86.exe --bfile naive/Oxford_FINAL --make-bld --r --ld-wind 4000 --out naive
```

#Warning: Duplicated SNP ID "rs3888396" has been changed to "rs3888396_612606".

```
head -n -1 naive.esi > proba.esi
```
#remove last line with alternative duplicated SNP rs3888396_612606

```
rm naive.esi

mv proba.esi naive.esi
```

#R: subset .esi with the SNPs only found in fastQTL results (.besd)

```
esi <- read.table("naive.esi", sep = "\t")

besd <- read.table("naive/fastQTL_results/permut_naive.txt", sep = " ")

new_esi <- esi[esi$V2 %in% besd$V6, ]

write.table(x=new_esi, file="naive_update.esi", sep = "\t", quote = FALSE, row.names=FALSE, col.names=FALSE)
```

#Update original esi coming from .besd

```
smr_Win/SMR_x86.exe --beqtl-summary naive --update-esi naive_update.esi
```

4. **Perform SMR.**

```
smr_Win/SMR_x86.exe --bfile naive/Oxford_FINAL --gwas-summary IMMUNO_INP_GWAS_QC.ma --beqtl-summary
naive/SMR_files/naive --out smr_naive_immuno --diff-freq-prop 0.300000 --peqtl-smr 1e-5
```

5. **Results annotation:** *SMR_results_immuno.R*

LPS 2H

1. **Obtain .besd file.**

#Remove probes without associations, we save 14.502 probes.

```
grep -vw "NA" permutations.all.chunksLPS2.txt > permut_LPS2.txt
```

#Save NA probes.

```
grep "NA" permutations.all.chunksLPS2.txt > NA_LPS2.txt
```

#Obtain .besd file

```
smr_Win/SMR_x86.exe --eqtl-summary permut_LPS2.txt --fastqtl-permu-format --make-besd --out LPS2
```

2. **Obtain .epi file.**

#For this we used R and its package illuminaHumanv4.db, we get: chr, start, strand (*LPS2_epi.R*)

```
awk 'NR>1 {print $1, "\t", $3, "\t", 0, "\t", $2, "\t", $3, "\t", $4}' probes_info_LPS2.txt > LPS2_update.epi
```

#Update original epi coming from .besd

```
smr_Win/SMR_x86.exe --beqtl-summary LPS2 --update-epi LPS2_update.epi
```

3. **Obtain .esi files (BLD format).**

```
smr_Win/SMR_x86.exe --bfile LPS2/Oxford_FINAL --make-bld --r --ld-wind 4000 --out LPS2
```

#Warning: Duplicated SNP ID "rs3888396" has been changed to "rs3888396_612606".

```
head -n -1 LPS2.esi > proba.esi
```
#remove last line with alternative duplicated SNP rs3888396_612606

```
rm LPS2.esi

mv proba.esi LPS2.esi
```

#R: subset .esi with the SNPs only found in fastQTL results (.besd)

```
esi <- read.table("LPS2.esi", sep = "\t")

besd <- read.table("LPS2/fastQTL_results/permut_LPS2.txt", sep = " ")

new_esi <- esi[esi$V2 %in% besd$V6, ]

write.table(x=new_esi, file="LPS2_update.esi ", sep = "\t", quote = FALSE, row.names=FALSE, col.names=FALSE)
```

#Update original esi coming from .besd

```
smr_Win/SMR_x86.exe --beqtl-summary LPS2 --update-esi LPS2_update.esi
```

4. **Perform SMR.**

```
smr_Win/SMR_x86.exe --bfile LPS2/Oxford_FINAL --gwas-summary IMMUNO_INP_GWAS_QC.ma --beqtl-summary
LPS2/SMR_files/LPS2 --out smr_LPS2_immuno --diff-freq-prop 0.300000 --peqtl-smr 1e-5
```

5. **Results annotation:** *SMR_results_immuno.R*

LPS 24H

6. **Obtain .besd file.**

#Remove probes without associations, we save 14.502 probes.

```
grep -vw "NA" permutations.all.chunksLPS24.txt > permut_LPS24.txt
```

#Save NA probes.

```
grep "NA" permutations.all.chunksLPS24.txt > NA_LPS24.txt
```

#Obtain .besd file

```
smr_Win/SMR_x86.exe --eqtl-summary permut_LPS24.txt --fastqtl-permu-format --make-besd --out LPS24
```

7. **Obtain .epi file.**

#For this we used R and its package illuminaHumanv4.db, we get: chr, start, strand (*LPS24_epi.R*)

```
awk 'NR>1 {print $1, "\t", $3, "\t", 0, "\t", $2, "\t", $3, "\t", $4}' probes_info_LPS24.txt > LPS24_update.epi
```

#Update original epi coming from .besd

```
smr_Win/SMR_x86.exe --beqtl-summary LPS24 --update-epi LPS24_update.epi
```

8. **Obtain .esi files (BLD format).**

```
smr_Win/SMR_x86.exe --bfile LPS24/Oxford_FINAL --make-bld --r --ld-wind 4000 --out LPS24_update
```

#Warning: Duplicated SNP ID "rs3888396" has been changed to "rs3888396_612606".

```
head -n -1 LPS24_update.esi > proba.esi
```
#remove last line with alternative duplicated SNP rs3888396_612606

```
rm LPS24_update.esi

mv proba.esi LPS24_update.esi
```

#R: subset .esi with the SNPs only found in fastQTL results (.besd)

```
esi <- read.table("LPS24_update.esi", sep = "\t")

besd <- read.table("LPS24/fastQTL_results/permut_LPS24.txt", sep = " ")

new_esi <- esi[esi$V2 %in% besd$V6, ]

write.table(x=new_esi, file="LPS24_update.esi", sep = "\t", quote = FALSE, row.names=FALSE, col.names=FALSE)
```

Update original esi coming from .besd

```
smr_Win/SMR_x86.exe --beqtl-summary LPS24 --update-esi LPS24_update.esi
```

9. **Perform SMR.**

```
smr_Win/SMR_x86.exe --bfile LPS24/Oxford_FINAL --gwas-summary IMMUNO_INP_GWAS_QC.ma --beqtl-summary
LPS24/SMR_files/LPS24 --out smr_LPS24_immuno --diff-freq-prop 0.300000 --peqtl-smr 1e-5
```

10. **Results annotation:** *SMR_results_immuno.R*