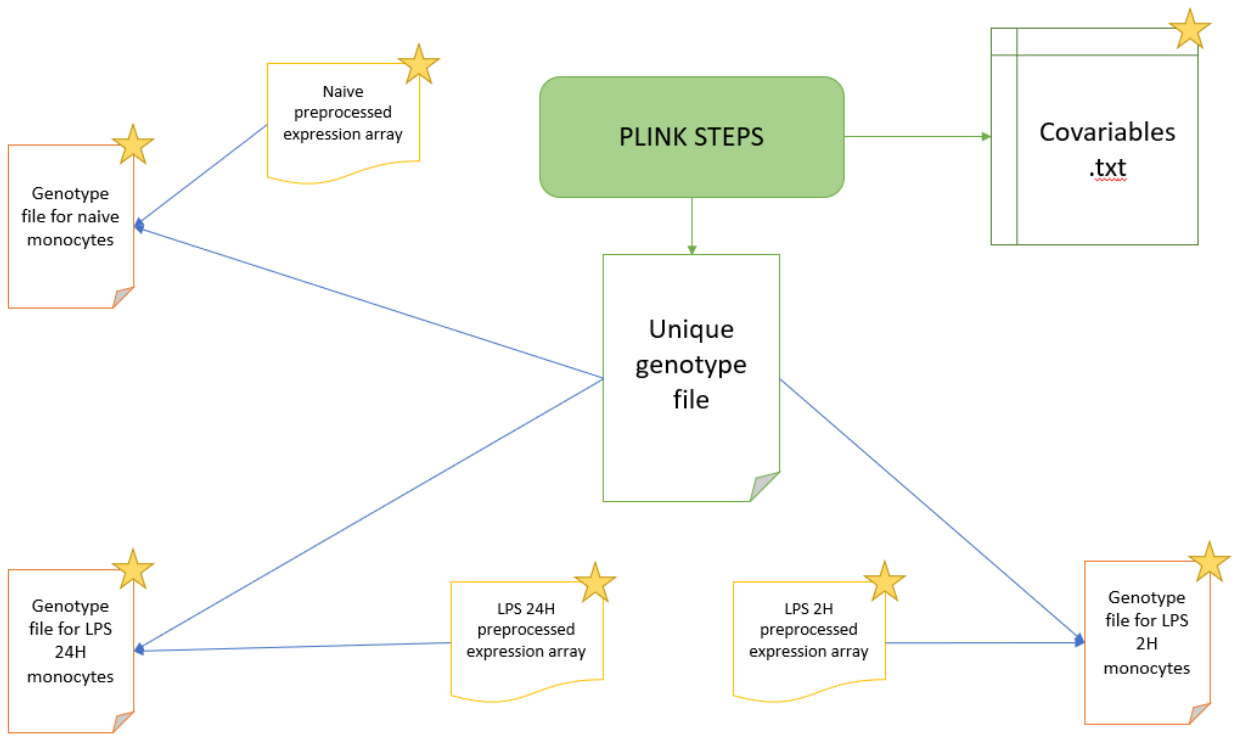
FastQTL: permutation pass in *cis*

At this point we have in one file all the genotypes from all those samples that pass the QC from PLINK step. To adapt the file to each experimental condition, we will need to keep only those samples specifically found in the preprocessed arrays of each experiment.



For fastQTL we need:

- Genotype files: VCF and VCF indexed.
- Expression files: BED and BED indexed.
- Covariables: text file.

NAÏVE

1. Genotype files

1.1. Remove samples

Using the preprocessed array from the naïve monocytes, there were samples already removed that we didn't discard, but we have to do it now to have the same samples for FastQTL: 16, 88, 191, 286, 301, 303, 332, 370, 415.

```
bin/plink2.exe -bfile MONOCYTE_eQTLs/genotypes/Oxford_CLEAN/Oxford_CLEAN --remove remove_pheno.txt --make-bed --out Oxford_FINAL
```

WE ANALYSE ONLY 413 SAMPLES WITH THE ARRAY PREPROCESSED

1.2. PLINK files per chromosomes and convert to VCF

for i in {1..23}; do bin/plink2.exe --bfile MONOCYTE_eQTLs/genotypes/Oxford_ARRAY/Oxford_FINAL --chr $i --make-bed --out Oxford_final_chr$i; done

for i in {1..23}; do bin/plink2.exe --bfile Oxford_final_chr$i --recode vcf --out Oxford_final_chr$i; done

1.3. Sort and compress VCF files

for i in {1..23}; do vcf-sort Oxford_final_chr$i.vcf | bgzip -c > Oxford_final_chr$i.vcf.gz; done

1.4. Change VCF files version (4.3 → 4.0) for the Michigan Imputation server

for i in {1..23}; do zcat MONOCYTE_eQTLs/genotypes/Chr_ARRAY/Oxford_final_chr$i.vcf.gz | vcf-convert -v 4.0 | bgzip -c > Oxford_final_chr$i.vcf.gz; done

1.5. Merge VCF files

vcf-concat Oxford_final_chr1.vcf.gz Oxford_final_chr2.vcf.gz Oxford_final_chr3.vcf.gz Oxford_final_chr4.vcf.gz Oxford_final_chr5.vcf.gz Oxford_final_chr6.vcf.gz Oxford_final_chr7.vcf.gz Oxford_final_chr8.vcf.gz Oxford_final_chr9.vcf.gz Oxford_final_chr10.vcf.gz Oxford_final_chr11.vcf.gz Oxford_final_chr12.vcf.gz Oxford_final_chr13.vcf.gz Oxford_final_chr14.vcf.gz Oxford_final_chr15.vcf.gz Oxford_final_chr16.vcf.gz Oxford_final_chr17.vcf.gz Oxford_final_chr18.vcf.gz Oxford_final_chr19.vcf.gz Oxford_final_chr20.vcf.gz Oxford_final_chr21.vcf.gz Oxford_final_chr22.vcf.gz Oxford_final_chr23.vcf.gz | gzip -c > all_chr_raw.vcf.gz

1.6. Obtain indexed VCF file

gunzip all_chr_raw.vcf.gz

bgzip all_chr_raw.vcf

tabix -p vcf all_chr_raw.vcf.gz

2. Covariable file

2.1. Perform IBD and Principal Component Analysis

bin/plink.exe --bfile MONOCYTE_eQTLs/genotypes/Oxford_FINAL/Oxford_FINAL --genome --make-bed --out Oxford_FINAL_IBD

bin/plink.exe --bfile MONOCYTE_eQTLs/genotypes/Oxford_FINAL/Oxford_FINAL_IBD --cluster --read-genome MONOCYTE_eQTLs/genotypes/Oxford_FINAL/Oxford_FINAL_IBD.genome --mds-plot 4 --K 2 --out Oxford_FINAL_strat

2.2) Obtain sex predicted from the final set of samples

bin/plink.exe --bfile Oxford_FINAL --check-sex --out Oxford_FINAL

3. Expression file

3.1. R script Bioconductor_package

In script *naive.R* we have the commands used to select the probes by IlluminaHumanv4.db package.

- Remove probes with +1 locus.
- Remove probes in sexual chr.
- Remove probes with non chr.
- Remove proves with SNP.
- Remove probes with quality BAD.
- Remove probes with quality NO MATCH.
- Remove cross-hybridized probes (BLAST+, *cross_blast.sh*).
  14.506 probes kept

3.2. Obtain BED format file and indexed it

```
(head -n1 phenotype.txt && sort -k1,1V -k2,2n -k3,3n <(tail -n+2 phenotype.txt)) > phenotype_sorted.bed

bgzip phenotype_sorted.bed

tabix -p bed phenotype_sorted.bed.gz
```

4. Run FastQTL

```
module load FastQTL/2.184-foss-2018b

touch phenotype_sorted.bed.gz.tbi

touch all_chr_raw.vcf.gz.tbi


for j in $(seq 1 500); do
        fastQTL --vcf all_chr_raw.vcf.gz --bed phenotype_sorted.bed.gz --permute 1000  --seed 123456789  --cov COV.txt  --out permutations${j}.txt.gz --chunk $j 500
done
```

4.1. Unify all chunks in one file

```
zcat permutations*.txt.gz | gzip -c > permutations.all.chunks.txt.gz
```

5. Results annotation and QC: *fastQTL_QC.R*

LPS 2H

1. Genotype files

1.1. Remove samples

Using the preprocessed array from the LPS 2H monocytes, there were samples already removed that we didn't discard, but we have to do it now to have the same samples for FastQTL (*remove_pheno_LPS2.txt*).

```
bin/plink2.exe -bfile genotypes/Oxford_CLEAN/Oxford_CLEAN --remove remove_pheno.txt --make-bed --out Oxford_FINAL
```

WE ANALYSE ONLY 260 SAMPLES WITH THE ARRAY PREPROCESSED

1.2. PLINK files per chromosomes and convert to VCF

```
for i in {1..23}; do bin/plink2.exe --bfile LPS2/Oxford_FINAL  --chr $i --make-bed --out Oxford_final_chr$i; done

for i in {1..23}; do bin/plink2.exe --bfile Oxford_final_chr$i --recode vcf --out Oxford_final_chr$i; done
```

1.3. Sort and compress VCF files & convert version (4.3 → 4.0)

```
for i in {1..23}; do vcf-sort Oxford_final_chr$i.vcf | vcf-convert -v 4.0 | bgzip -c > Oxford_final_chr$i.vcf.gz; done
```

1.4. Merge VCF files

```
vcf-concat Oxford_final_chr1.vcf.gz Oxford_final_chr2.vcf.gz Oxford_final_chr3.vcf.gz Oxford_final_chr4.vcf.gz Oxford_final_chr5.vcf.gz
Oxford_final_chr6.vcf.gz Oxford_final_chr7.vcf.gz Oxford_final_chr8.vcf.gz Oxford_final_chr9.vcf.gz Oxford_final_chr10.vcf.gz
Oxford_final_chr11.vcf.gz Oxford_final_chr12.vcf.gz Oxford_final_chr13.vcf.gz Oxford_final_chr14.vcf.gz Oxford_final_chr15.vcf.gz
Oxford_final_chr16.vcf.gz Oxford_final_chr17.vcf.gz Oxford_final_chr18.vcf.gz Oxford_final_chr19.vcf.gz Oxford_final_chr20.vcf.gz
Oxford_final_chr21.vcf.gz Oxford_final_chr22.vcf.gz Oxford_final_chr23.vcf.gz | gzip -c > all_chr_LPS2.vcf.gz
```

1.5. Obtain indexed VCF file

```
gunzip all_chr_LPS2.vcf.gz

bgzip all_chr_LPS2.vcf

tabix -p vcf all_chr_LPS2.vcf.gz
```

2. Covariable file

2.1. Perform IBD and Principal Component Analysis

```
bin/plink.exe --bfile     LPS2/Oxford_FINAL --genome --make-bed --out Oxford_FINAL_IBD

bin/plink.exe --bfile Oxford_FINAL_IBD --cluster --read-genome Oxford_FINAL_IBD.genome --mds-plot 4 --K 2 --out Oxford_FINAL_strat
```

2.2. Obtain sex predicted from the final set of samples

```
bin/plink.exe --bfile Oxford_FINAL --check-sex   --out Oxford_FINAL
```

3. Expression file

3.1. R script Bioconductor_package

In script *LPS2.R* we have the commands used to select the probes by IlluminaHumanv4.db package.

- Remove probes with +1 locus.
- Remove probes in sexual chr.
- Remove probes with non chr.
- Remove proves with SNP.
- Remove probes with quality BAD.

- Remove probes with quality NO MATCH.
- Remove cross-hybridized probes (BLAST+, *cross_blast.sh*).
  14.506 probes kept

3.2. Obtain BED format file and indexed it

```
(head -n1 phenotype_LPS2.txt && sort -k1,1V -k2,2n -k3,3n <(tail -n+2 phenotype_LPS2.txt)) >
phenotypeLPS2_sorted.bed

bgzip phenotypeLPS2_sorted.bed

tabix -p bed phenotypeLPS2_sorted.bed.gz
```

4. Run FastQTL

```
module load FastQTL/2.184-foss-2018b

touch  phenotypeLPS2_sorted.bed.gz.tbi

touch all_chr_LPS2.vcf.gz.tbi


for j in $(seq 1 500); do
        fastQTL --vcf all_chr_LPS2.vcf.gz --bed phenotypeLPS2_sorted.bed.gz --permute 1000  --seed 123456789 --cov  COV_LPS2.txt  --out
permutations${j}.txt.gz --chunk $j 500
done
```

4.1. Unify all chunks in one file

```
zcat permutations*.txt.gz | gzip -c > permutations.all.chunks.txt.gz
```

5. Results annotation and QC: *fastQTL_QC.R*

LPS 24H

1. Genotype files

1.1. Remove samples

Using the preprocessed array from the LPS 2H monocytes, there were samples already removed that we didn't discard, but we have to do it now to have the same samples for FastQTL (*remove_pheno_LPS24.txt*).

```
bin/plink2.exe -bfile genotypes/Oxford_CLEAN/Oxford_CLEAN --remove remove_pheno_LPS24.txt --make-bed --out Oxford_FINAL
```

WE ANALYSE ONLY 321 SAMPLES WITH THE ARRAY PREPROCESSED

1.2. PLINK files per chromosomes and convert to VCF

```
for i in {1..23}; do bin/plink2.exe --bfile LPS24/Oxford_FINAL  --chr $i --make-bed --out Oxford_final_chr$i; done
```

```
for i in {1..23}; do bin/plink2.exe --bfile Oxford_final_chr$i --recode vcf --out Oxford_final_chr$i; done
```

1.3. Sort and compress VCF files & convert version (4.3 → 4.0)

```
for i in {1..23}; do vcf-sort Oxford_final_chr$i.vcf | vcf-convert -v 4.0 | bgzip -c > Oxford_final_chr$i.vcf.gz; done
```

1.4. Merge VCF files

```
vcf-concat Oxford_final_chr1.vcf.gz Oxford_final_chr2.vcf.gz Oxford_final_chr3.vcf.gz Oxford_final_chr4.vcf.gz Oxford_final_chr5.vcf.gz
Oxford_final_chr6.vcf.gz Oxford_final_chr7.vcf.gz Oxford_final_chr8.vcf.gz Oxford_final_chr9.vcf.gz Oxford_final_chr10.vcf.gz
Oxford_final_chr11.vcf.gz Oxford_final_chr12.vcf.gz Oxford_final_chr13.vcf.gz Oxford_final_chr14.vcf.gz Oxford_final_chr15.vcf.gz
Oxford_final_chr16.vcf.gz Oxford_final_chr17.vcf.gz Oxford_final_chr18.vcf.gz Oxford_final_chr19.vcf.gz Oxford_final_chr20.vcf.gz
Oxford_final_chr21.vcf.gz Oxford_final_chr22.vcf.gz Oxford_final_chr23.vcf.gz | gzip -c > all_chr_LPS24.vcf.gz
```

1.5. Obtain indexed VCF file

```
gunzip all_chr_LPS24.vcf.gz

bgzip all_chr_LPS24.vcf

tabix -p vcf all_chr_LPS24.vcf.gz
```

2. Covariable file

2.1. Perform IBD and Principal Component Analysis

```
bin/plink.exe –bfile LPS24/Oxford_FINAL --genome --make-bed --out Oxford_FINAL_IBD

bin/plink.exe --bfile Oxford_FINAL_IBD --cluster --read-genome Oxford_FINAL_IBD.genome --mds-plot 4 --K 2 --out Oxford_FINAL_strat
```

2.2. Obtain sex predicted from the final set of samples

```
bin/plink.exe --bfile LPS24/Oxford_FINAL --check-sex   --out Oxford_FINAL
```

3. Expression file

3.1. R script Bioconductor_package

In script *LPS24.R* we have the commands used to select the probes by IlluminaHumanv4.db package.

- Remove probes with +1 locus.
- Remove probes in sexual chr.
- Remove probes with non chr.
- Remove proves with SNP.
- Remove probes with quality BAD.
- Remove probes with quality NO MATCH.
- Remove cross-hybridized probes (BLAST+, *cross_blast.sh*).
  14.506 probes

3.2. Obtain BED format file and indexed it

```
(head -n1 phenotype_LPS24.txt && sort -k1,1V -k2,2n -k3,3n <(tail -n+2 phenotype_LPS24.txt)) > phenotypeLPS24_sorted.bed

bgzip phenotypeLPS24_sorted.bed

tabix -p bed phenotypeLPS24_sorted.bed.gz
```

4. Run FastQTL

```
module load FastQTL/2.184-foss-2018b

touch  phenotypeLPS24_sorted.bed.gz.tbi

touch all_chr_LPS24.vcf.gz.tbi


for j in $(seq 1 500); do
        fastQTL --vcf all_chr_LPS24.vcf.gz --bed phenotypeLPS24_sorted.bed.gz --permute 1000  --seed 123456789 --cov  COV_LPS24.txt  --
out permutations${j}.txt.gz --chunk $j 500
done
```

4.1. Unify all chunks in one file

```
zcat permutations*.txt.gz | gzip -c > permutations.all.chunks.txt.gz
```

5. Results annotation and QC: *fastQTL_QC.R*