Supplementary File S1

Script used to estimate the amount of variance explained by each SNP using the Additive, Additive+Dominant and General models

Loading libraries and preparing phenotype

```r
rm(list=ls())

setwd("/Users/msu_potato/Desktop/GWAS/2018/lb")

library(BGLR)
load('Supplementary_file_2_LB.RData') # Example using late blight data.
                                      # For common scab use Supplementary_file_3_Scab.RData
verbose=TRUE

# Checking trait name
print(trait)

# Matching phenotypes and genotypes
idxGenos=match(phenotype$Genotype,rownames(genotype))
X=genotype[idxGenos,]
stopifnot(all(phenotype$Genotype==rownames(X)))

# Centering and scaling genotypes and phenotypes
X.A=scale(X)/sqrt(ncol(X))    # incidence matrix for additive effects
y=scale(phenotype$Score)
```

Defining parameters and output tables

```r
nIter=3300   ## Used less iterations and burn-in for fast analysis
burnIn=300
verbose=TRUE
model='BayesB' # or BRR

# Output table for variances

OUT=matrix(nrow=7,ncol=9,NA)
colnames(OUT)=c('Y','G','PC','A','D','Gral','Total','YxG','Error')
# Y=Year
# G=Genotype
# PC=first 5 principal components (5-PCs)
# A=Additive effects
# D=Dominant effects
# Gral=General effects
# Total=Total genetic variance
# YxG=Year by genotype interaction
rownames(OUT)=c('M1','M2','M3','M4','M5','M6','M7')
# M1=Year
```

```
# M2=Year and genotype
# M3=Year, genotype and their interaction
# M4=Year, genotype, YxG interaction and 5-PCs
# M5=Year, additive effects, YxG interaction and 5-PCs
# M6=Year, additive and dominant effects, YxG interaction and 5-PCs
# M7=Year, general effects, YxG interaction and 5-PCs

 # Output table for confidence intervals
 CI=OUT
```

Fitting models

Year model

```
# Linear predictor
Z.YEAR=as.matrix(model.matrix(~factor(phenotype$Year)))[,-1]
ETA=list(year=list(X=Z.YEAR,model='BRR',saveEffects=TRUE))

# Model fitting
prefix='M1'
fm=BGLR(y=y,ETA=ETA,nIter=nIter,burnIn=burnIn,saveAt=paste0(prefix,'_'),verbose=verbose)
save(fm,file=paste0('fm',prefix,'.RData'))

## Analysis of variance
# Year variance
B.YEAR=readBinMat(paste0(prefix,'_ETA_year_b.bin'))
vY=rep(NA,nrow(B.YEAR))
for(i in 1:nrow(B.YEAR)){  vY[i]=var(Z.YEAR%*%B.YEAR[i,])}
OUT[prefix,'Y']=mean(vY)
CI[prefix,'Y']=sd(vY)

# Error variance
vE=scan(paste0(prefix,'_varE.dat'))[-(1:floor(burnIn/5))]
OUT[prefix,'Error']=mean(vE)
CI[prefix,'Error']=sd(vE)
```

Year + Genotype model

```
# Linear predictor
Z.Genotype=as.matrix(model.matrix(~factor(phenotype$Genotype)-1))
ETA$Genotype=list(X=Z.Genotype,model='BRR',saveEffects=TRUE)

# Model fitting
prefix='M2'
fm=BGLR(y=y,ETA=ETA,nIter=nIter,burnIn=burnIn,saveAt=paste0(prefix,'_'),verbose=verbose)
save(fm,file=paste0('fm',prefix,'.RData'))

## Analysis of variance
# Year variance
B.YEAR=readBinMat(paste0(prefix,'_ETA_year_b.bin'))
vY=rep(NA,nrow(B.YEAR))
for(i in 1:nrow(B.YEAR)){  vY[i]=var(Z.YEAR%*%B.YEAR[i,])}
```

```r
OUT[prefix,'Y']=mean(vY)
CI[prefix,'Y']=sd(vY)

# Genotype variance
B.Genotype=readBinMat(paste0(prefix,'_ETA_Genotype_b.bin'))
vG=rep(NA,nrow(B.Genotype))
for(i in 1:nrow(B.Genotype)){vG[i]=var(Z.Genotype%*%B.Genotype[i,])}
OUT[prefix,'G']=mean(vG)
CI[prefix,'G']=sd(vG)
OUT[prefix,'Total']=mean(vG)
CI[prefix,'Total']=sd(vG)

# Error variance
vE=scan(paste0(prefix,'_varE.dat'))[-(1:floor(burnIn/5))]
OUT[prefix,'Error']=mean(vE)
CI[prefix,'Error']=sd(vE)
```

Year + Genotype + Year x Genotype model

```r
# Linear predictor
Z.YxG=model.matrix(~factor(paste(phenotype$Genotype,phenotype$Year,sep='xxx')))
ETA$YearGenotype=list(X=Z.YxG,model='BRR',saveEffects=TRUE)

# Model fitting
prefix='M3'
fm=BGLR(y=y,ETA=ETA,nIter=nIter,burnIn=burnIn,saveAt=paste0(prefix,'_'),verbose=verbose)
save(fm,file=paste0('fm',prefix,'.RData'))

## Analysis of variance
# Year variance
B.YEAR=readBinMat(paste0(prefix,'_ETA_year_b.bin'))
vY=rep(NA,nrow(B.YEAR))
for(i in 1:nrow(B.YEAR)){  vY[i]=var(Z.YEAR%*%B.YEAR[i,])}
OUT[prefix,'Y']=mean(vY)
CI[prefix,'Y']=sd(vY)

# Genotype variance
B.Genotype=readBinMat(paste0(prefix,'_ETA_Genotype_b.bin'))
vG=rep(NA,nrow(B.Genotype))
for(i in 1:nrow(B.Genotype)){vG[i]=var(Z.Genotype%*%B.Genotype[i,])}
  OUT[prefix,'G']=mean(vG)
  CI[prefix,'G']=sd(vG)
  OUT[prefix,'Total']=mean(vG)
  CI[prefix,'Total']=sd(vG)

# Year x Genotype variance
B.YxG=readBinMat(paste0(prefix,'_ETA_YearGenotype_b.bin'))
vYxG=rep(NA,nrow(B.YxG))
for(i in 1:nrow(B.YxG)){  vYxG[i]=var(Z.YxG%*%B.YxG[i,])}
OUT[prefix,'YxG']=mean(vYxG)
CI[prefix,'YxG']=sd(vYxG)

# Error variance
```

```
vE=scan(paste0(prefix,'_varE.dat'))[-(1:floor(burnIn/5))]
OUT[prefix,'Error']=mean(vE)
CI[prefix,'Error']=sd(vE)
```

Year + Genotype + PCs + Year x Genotype model

```
# Principal component analysis using the additive model
A=scale(genotype)
G=tcrossprod(A)
G=G/mean(diag(G))
EVD.G=eigen(G,symmetric=TRUE)
PC=EVD.G$vectors[,1:5]

# Linear predictor
ID.G=as.integer(factor(x=phenotype$Genotype,levels=rownames(genotype),ordered=TRUE))
Z.PC=PC[ID.G,]
ETA$PC=list(X=Z.PC,model='FIXED',saveEffects=TRUE)

# Model fitting
prefix='M4'
fm=BGLR(y=y,ETA=ETA,nIter=nIter,burnIn=burnIn,saveAt=paste0(prefix,'_'),verbose=verbose)
save(fm,file=paste0('fm',prefix,'.RData'))


## Analysis of variance
# Year variance
B.YEAR=readBinMat(paste0(prefix,'_ETA_year_b.bin'))
vY=rep(NA,nrow(B.YEAR))
for(i in 1:nrow(B.YEAR)){vY[i]=var(Z.YEAR%*%B.YEAR[i,])}
OUT[prefix,'Y']=mean(vY)
CI[prefix,'Y']=sd(vY)

# Genotype and PC variances
B.Genotype=readBinMat(paste0(prefix,'_ETA_Genotype_b.bin'))
vG=rep(NA,nrow(B.Genotype))
B.PC=as.matrix(read.table(paste0(prefix,'_ETA_PC_b.dat'),header=F))[-(1:floor(burnIn/5)),]
vPC=rep(NA,nrow(B.PC))
vGPC=vG

for(i in 1:(length(vPC))){
  uG=Z.Genotype%*%B.Genotype[i,]
  uPC=Z.PC%*%B.PC[i,]
  vG[i]=var(uG)
  vPC[i]=var(uPC)
  vGPC[i]=var(uG+uPC)
}

OUT[prefix,'G']=mean(vG)
CI[prefix,'G']=sd(vG)
OUT[prefix,'PC']=mean(vPC)
CI[prefix,'PC']=sd(vPC)
OUT[prefix,'Total']=mean(vGPC)
CI[prefix,'Total']=sd(vGPC)
```

```r
# Year x Genotype variance
B.YxG=readBinMat(paste0(prefix,'_ETA_YearGenotype_b.bin'))
vYxG=rep(NA,nrow(B.YxG))
for(i in 1:nrow(B.YxG)){vYxG[i]=var(Z.YxG%*%B.YxG[i,])}
OUT[prefix,'YxG']=mean(vYxG)
CI[prefix,'YxG']=sd(vYxG)

# Error variance
vE=scan(paste0(prefix,'_varE.dat'))[-(1:floor(burnIn/5))]
OUT[prefix,'Error']=mean(vE)
CI[prefix,'Error']=sd(vE)
```

Year + Year x Genotype + PCs + Additive model

```r
# Linear predictors
ETA=list()
Z.YEAR=as.matrix(model.matrix(~factor(phenotype$Year)))[,-1]
ETA=list(year=list(X=Z.YEAR,model='BRR',saveEffects=TRUE))
Z.YxG=model.matrix(~factor(paste(phenotype$Genotype,phenotype$Year,sep='xxx')))
ETA$YearGenotype=list(X=Z.YxG,model='BRR',saveEffects=TRUE)

Z.A=scale(genotype)
ID.G=as.integer(factor(x=phenotype$Genotype,levels=rownames(genotype),ordered=TRUE))

## Principal component analysis
G=tcrossprod(Z.A)
G=G/mean(diag(G))
EVD.G=eigen(G,symmetric=TRUE)
if(model=='BRR'){
  PC=EVD.G$vectors[,EVD.G$values>1e-5]
    for(i in 1:ncol(PC)){ PC[,i]=PC[,i]*sqrt(EVD.G$values[i]) }
      Z.A=PC[ID.G,]
    }else{
      Z.A=Z.A[ID.G,]
  }

PC=EVD.G$vectors[,1:5]
Z.PC=PC[ID.G,]
ETA$A=list(X=Z.A,model=model,saveEffects=TRUE)
ETA$PC=list(X=Z.PC,model='FIXED',saveEffects=TRUE)

# Model fitting
prefix='M5'
fm=BGLR(y=y,ETA=ETA,nIter=nIter,burnIn=burnIn,saveAt=paste0(prefix,'_'),verbose=verbose)
save(fm,file=paste0('fm',prefix,'.RData'))

## Variance analysis
# Year variance
B.YEAR=readBinMat(paste0(prefix,'_ETA_year_b.bin'))
vY=rep(NA,nrow(B.YEAR))
for(i in 1:nrow(B.YEAR)){  vY[i]=var(Z.YEAR%*%B.YEAR[i,])}
OUT[prefix,'Y']=mean(vY)
CI[prefix,'Y']=sd(vY)
```

```r
# Additive, PCs and total genetic variance
B.A=readBinMat(paste0(prefix,'_ETA_A_b.bin'))
vA=rep(NA,(nrow(B.A)))
B.PC=as.matrix(read.table(paste0(prefix,'_ETA_PC_b.dat'),header=F))[-(1:floor(burnIn/5)),]
vPC=rep(NA,nrow(B.PC))
vAPC=vA

for(i in 1:(length(vPC))){
   uA=Z.A%*%B.A[i,]
   uPC=Z.PC%*%B.PC[i,]
   vA[i]=var(uA)
   vPC[i]=var(uPC)
   vAPC[i]=var(uA+uPC)
}

OUT[prefix,'PC']=mean(vPC)
CI[prefix,'PC']=sd(vPC)
OUT[prefix,'A']=mean(vA)
CI[prefix,'A']=sd(vA)
OUT[prefix,'Total']=mean(vAPC)
CI[prefix,'Total']=sd(vAPC)

# Year x Genotype variance
B.YxG=readBinMat(paste0(prefix,'_ETA_YearGenotype_b.bin'))
vYxG=rep(NA,nrow(B.YxG))
for(i in 1:nrow(B.YxG)){  vYxG[i]=var(Z.YxG%*%B.YxG[i,])}
OUT[prefix,'YxG']=mean(vYxG)
CI[prefix,'YxG']=sd(vYxG)

# Error variance
vE=scan(paste0(prefix,'_varE.dat'))[-(1:floor(burnIn/5))]
OUT[prefix,'Error']=mean(vE)
CI[prefix,'Error']=sd(vE)
```

Year + Year x Genotype + PC + Additive + Dominant model

```r
# Dominant matrix
Z.D=scale((genotype>0)*(genotype<4))
G=tcrossprod(Z.D)
G=G/mean(diag(G))
EVD.G=eigen(G,symmetric=TRUE)
if(model=='BRR'){
  PC=EVD.G$vectors[,EVD.G$values>1e-5]
    for(i in 1:ncol(PC)){ PC[,i]=PC[,i]*sqrt(EVD.G$values[i]) }
      Z.D=PC[ID.G,]
    }else{
      Z.D=Z.D[ID.G,]
  }

ETA$D=list(X=Z.D,model=model ,saveEffects=TRUE)

# Model fitting
prefix='M6'
```

```r
fm=BGLR(y=y,ETA=ETA,nIter=nIter,burnIn=burnIn,saveAt=paste0(prefix,'_'),verbose=verbose)
save(fm,file=paste0('fm',prefix,'.RData'))

## Variance analysis
# Year variance
B.YEAR=readBinMat(paste0(prefix,'_ETA_year_b.bin'))
vY=rep(NA,nrow(B.YEAR))
for(i in 1:nrow(B.YEAR)){  vY[i]=var(Z.YEAR%*%B.YEAR[i,])}
OUT[prefix,'Y']=mean(vY)
CI[prefix,'Y']=sd(vY)

# Additive, Dominance, PC and total genetic variance
B.A=readBinMat(paste0(prefix,'_ETA_A_b.bin'))
vA=rep(NA,(nrow(B.A)))
B.D=readBinMat(paste0(prefix,'_ETA_D_b.bin'))
vD=rep(NA,(nrow(B.A)))
B.PC=as.matrix(read.table(paste0(prefix,'_ETA_PC_b.dat'),header=F))[-(1:floor(burnIn/5)),]
vPC=rep(NA,nrow(B.PC))
vAD=vD

  for(i in 1:(length(vPC))){
    uA=Z.A%*%B.A[i,]
    uD=Z.D%*%B.D[i,]
    uPC=Z.PC%*%B.PC[i,]
    vA[i]=var(uA)
    vD[i]=var(uD)
    vPC[i]=var(uPC)
    vAD[i]=var(uA+uD+uPC)
  }

OUT[prefix,'PC']=mean(vPC)
CI[prefix,'PC']=sd(vPC)
OUT[prefix,'A']=mean(vA)
CI[prefix,'A']=sd(vA)
OUT[prefix,'D']=mean(vD)
CI[prefix,'D']=sd(vD)
OUT[prefix,'Total']=mean(vAD)
CI[prefix,'Total']=sd(vAD)

## Year x Genotype variance
B.YxG=readBinMat(paste0(prefix,'_ETA_YearGenotype_b.bin'))
vYxG=rep(NA,nrow(B.YxG))
for(i in 1:nrow(B.YxG)){vYxG[i]=var(Z.YxG%*%B.YxG[i,])}
OUT[prefix,'YxG']=mean(vYxG)
CI[prefix,'YxG']=sd(vYxG)

# Error variance
vE=scan(paste0(prefix,'_varE.dat'))[-(1:floor(burnIn/5))]
OUT[prefix,'Error']=mean(vE)
CI[prefix,'Error']=sd(vE)
```

Year + Year x Genotype + PC + General model

```r
# Function to generate a general model matrix
getGM=function(X){
  n=nrow(X)
  p=ncol(X)*5
  Z=matrix(nrow=n,ncol=p,NA)
  rownames(Z)=rownames(X)
  mName=colnames(X)
  stCol=0
  enCol=0
  Zcnames=c()
  for(i in 1:ncol(X)){
    x=X[,i]
    tmp=as.matrix(model.matrix(~factor(x)-1))
    nC=ncol(tmp)
    stCol=enCol+1
    enCol=enCol+nC
    Z[,stCol:enCol]=tmp
    Zcnames=c(Zcnames,paste(mName[i], 1:nC, sep="-"))
  }
 Z=Z[,1:enCol]
 return(Z)
}

# Linear predictors
ETA=list()
A=scale(genotype)
G=tcrossprod(A)
G=G/mean(diag(G))
EVD.G=eigen(G,symmetric=TRUE)
PC=EVD.G$vectors[,1:5]

Z.YEAR=as.matrix(model.matrix(~factor(phenotype$Year)))[,-1]
ETA$year=list(X=Z.YEAR,model='BRR',saveEffects=TRUE)
Z.YxG=model.matrix(~factor(paste(phenotype$Genotype,phenotype$Year,sep='xxx')))
ETA$YearGenotype=list(X=Z.YxG,model='BRR',saveEffects=TRUE)

Z.G=scale(getGM(genotype))
if(model=='BRR'){
  G=tcrossprod(Z.G)
  G=G/mean(diag(G))
  EVD.G=eigen(G,symmetric=TRUE)
  PC_G=EVD.G$vectors[,EVD.G$values>1e-5]
    for(i in 1:ncol(PC_G)){ PC_G[,i]=PC_G[,i]*sqrt(EVD.G$values[i]) }
        Z.G=PC_G[ID.G,]
    }else{
        Z.G=Z.G[ID.G,]
        }

Z.PC=PC[ID.G,]
ETA$PC=list(X=Z.PC,model='FIXED',saveEffects=TRUE)
ETA$G=list(X=Z.G,model=model,saveEffects=TRUE)

## Model fitting
```

```r
prefix='M7'
fm=BGLR(y=y,ETA=ETA,nIter=nIter,burnIn=burnIn,saveAt=paste0(prefix,'_'),verbose=verbose)
save(fm,file=paste0('fm',prefix,'.RData'))

## Variance analysis
# Year variance
B.YEAR=readBinMat(paste0(prefix,'_ETA_year_b.bin'))
vY=rep(NA,nrow(B.YEAR))
for(i in 1:nrow(B.YEAR)){  vY[i]=var(Z.YEAR%*%B.YEAR[i,])}
OUT[prefix,'Y']=mean(vY)
CI[prefix,'Y']=sd(vY)

# General, PC and total genetic variance
B.G=readBinMat(paste0(prefix,'_ETA_G_b.bin'))
vG=rep(NA,(nrow(B.G)))
B.PC=as.matrix(read.table(paste0(prefix,'_ETA_PC_b.dat'),header=F))[-(1:floor(burnIn/5)),]
vPC=rep(NA,nrow(B.PC))
vGPC=vG
for(i in 1:(length(vPC))){
  uG=Z.G%*%B.G[i,]
  uPC=Z.PC%*%B.PC[i,]
  vG[i]=var(uG)
  vPC[i]=var(uPC)
  vGPC[i]=var(uG+uPC)
}

OUT[prefix,'PC']=mean(vPC)
CI[prefix,'PC']=sd(vPC)
OUT[prefix,'Gral']=mean(vG)
CI[prefix,'Gral']=sd(vG)
OUT[prefix,'Total']=mean(vGPC)
CI[prefix,'Total']=sd(vGPC)

# Year x Genotype variance
B.YxG=readBinMat(paste0(prefix,'_ETA_YearGenotype_b.bin'))
vYxG=rep(NA,nrow(B.YxG))
for(i in 1:nrow(B.YxG)){vYxG[i]=var(Z.YxG%*%B.YxG[i,])}
OUT[prefix,'YxG']=mean(vYxG)
CI[prefix,'YxG']=sd(vYxG)

# Error variance
vE=scan(paste0(prefix,'_varE.dat'))[-(1:floor(burnIn/5))]
OUT[prefix,'Error']=mean(vE)
CI[prefix,'Error']=sd(vE)

# Save both results components variance and CI
write.table(OUT,file='VARCOMP.csv',sep=",",row.names=F) # Variances table
write.table(CI,file='CI.csv',sep=",",row.names=F) # Confidence interval table
```