```
In [1]: include("/home/nicole/Jupyter/SSBRnoJ/src/SSBR.jl")
        using SSBR
```

```
In [2]: function getPos(ped,IDs)
            posAi = Array(Int64,size(IDs,1))
            for (i,id) = enumerate(IDs[:,1])
                posAi[i] = ped.idMap[id].seqID
            end
            return posAi
        end
```

```
Out[2]: getPos (generic function with 1 method)
```

```
In [3]: ; cd Data/0.3/10
```

```
/home/nicole/Jupyter/Jpaper/Data/0.3/10
```

```
In [4]: ;ls
```

```
G1.Genotype.ID
G1.ID
G1.noGenotype.ID
G2.Genotype.ID
G2.ID
G2.noGenotype.ID
G3.Genotype.ID
G3.ID
G3.noGenotype.ID
G4.Genotype.ID
G4.ID
G4.noGenotype.ID
G5.Genotype.ID
G5.ID
G5.noGenotype.ID
GenNF.txt
PedAll.txt
Phe.txt
PheAll.txt
all.ID
genotype.ID
noGenotype.ID
sim.bv
sim.phenotype
```

```
In [5]: ;awk '{print $1}' PedAll.txt | sort -b > all.ID
```

```
In [6]: ;awk '{print $1}' GenNF.txt | sort -b > genotype.ID
```

```
In [7]: ;join -v1 all.ID genotype.ID > noGenotype.ID
```

```
In [8]: ;awk '{print $1,$2}' Phe.txt > sim.phenotype
```

In [9]:
```
;awk '{print $1,$3}' PheAll.txt > sim.bv
```

In [10]:
```
; awk 'NR >=1 && NR <=8000 {print $1}' PedAll.txt | sort -b > G1.ID
```

In [11]:
```
; awk 'NR >=8001 && NR <=16000 {print $1}' PedAll.txt | sort -b > G2.ID
```

In [12]:
```
; awk 'NR >=16001 && NR <=24000 {print $1}' PedAll.txt | sort -b > G3.ID
```

In [13]:
```
; awk 'NR >=24001 && NR <=32000 {print $1}' PedAll.txt | sort -b > G4.ID
```

In [14]:
```
; awk 'NR >=32001 && NR <=40000 {print $1}' PedAll.txt | sort -b > G5.ID
```

In [15]:
```
;join G1.ID genotype.ID > G1.Genotype.ID
```

In [16]:
```
;join G2.ID genotype.ID > G2.Genotype.ID
```

In [17]:
```
;join G3.ID genotype.ID > G3.Genotype.ID
```

In [18]:
```
;join G4.ID genotype.ID > G4.Genotype.ID
```

In [19]:
```
;join G5.ID genotype.ID > G5.Genotype.ID
```

In [20]:
```
;join -v1 G1.ID genotype.ID > G1.noGenotype.ID
```

In [21]:
```
;join -v1 G2.ID genotype.ID > G2.noGenotype.ID
```

In [22]:
```
;join -v1 G3.ID genotype.ID > G3.noGenotype.ID
```

In [23]:
```
;join -v1 G4.ID genotype.ID > G4.noGenotype.ID
```

In [24]:
```
;join -v1 G5.ID genotype.ID > G5.noGenotype.ID
```

In [25]:
```
;wc G1.Genotype.ID;wc G2.Genotype.ID;wc G3.Genotype.ID;wc G4.Genotype.ID;wc (
```
```
 200   200 1200 G1.Genotype.ID
 200   200 1200 G2.Genotype.ID
 200   200 1200 G3.Genotype.ID
 200   200 1200 G4.Genotype.ID
8000  8000 48000 G5.Genotype.ID
```

In [26]:
```
;wc G1.noGenotype.ID;wc G2.noGenotype.ID;wc G3.noGenotype.ID;wc G4.noGenotype
```
```
7800   7800 46800 G1.noGenotype.ID
7800   7800 46800 G2.noGenotype.ID
7800   7800 46800 G3.noGenotype.ID
7800   7800 46800 G4.noGenotype.ID
0 0 0 G5.noGenotype.ID
```

In [27]:
```
ped,A_Mats,numSSBayes = calc_Ai("PedAll.txt","genotype.ID",calculateInbreedi
nothing
y_Vecs = make_yVecs("sim.phenotype",ped,numSSBayes)
Z_Mats = make_ZMats(ped,y_Vecs,numSSBayes)
nothing
```

In [28]:
```
vRes    = 1.478
vG      = 0.633
aHat3=SSBR.PBLUP(y_Vecs,Z_Mats,A_Mats,numSSBayes,vRes,vG);
```

In [29]:
```
using DataFrames
```

In [30]:
```
df = readtable("sim.bv", eltypes =[UTF8String, Float64], separator = ' ',head
a  = Array(Float64,numSSBayes.num_ped)
for (i,ID) in enumerate(df[:,1])
    j = ped.idMap[ID].seqID
    a[j] = df[i,2]
end
```

In [31]:
```
IDs = readtable("all.ID", eltypes =[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor1 = cor(a[posAi],aHat3[posAi])[1,1]
reg1 = linreg(aHat3[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - all.ID : correlation = %6.3f\n", cor1 ) # with
@printf("SSBRJC from Gibbs - all.ID  : regression of TBV on GEBV = %6.3f\n",
JCAll = cor1
```

```
SSBRJC from Gibbs - all.ID : correlation =  0.770
SSBRJC from Gibbs - all.ID  : regression of TBV on GEBV =  0.998
```

Out[31]: 0.7703166201819893

In [32]:
```
IDs = readtable("genotype.ID", eltypes =[UTF8String], separator = ' ',header=
posAi = getPos(ped,IDs)
cor2 = cor(a[posAi],aHat3[posAi])[1,1]
reg2 = linreg(aHat3[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - genotype.ID : correlation = %6.3f\n", cor2 ) # w
@printf("SSBRJC from Gibbs - genotype.ID  : regression of TBV on GEBV = %6.3:
JCAll = cor2
```

```
SSBRJC from Gibbs - genotype.ID : correlation =  0.476
SSBRJC from Gibbs - genotype.ID  : regression of TBV on GEBV =  1.009
```

Out[32]: 0.4756742050559272

```
In [33]:  IDs = readtable("noGenotype.ID", eltypes =[UTF8String], separator = ' ',heade
          posAi = getPos(ped,IDs)
          cor3 = cor(a[posAi],aHat3[posAi])[1,1]
          reg3 = linreg(aHat3[posAi], a[posAi])[2,1]
          @printf("SSBRJC from Gibbs - noGenotype.ID : correlation = %6.3f\n", cor3 ) ;
          @printf("SSBRJC from Gibbs - noGenotype.ID  : regression of TBV on GEBV = %6
          JCAll = cor3
```

```
          SSBRJC from Gibbs - noGenotype.ID : correlation =   0.745
          SSBRJC from Gibbs - noGenotype.ID  : regression of TBV on GEBV =   0.996
```

Out[33]: 0.7452335534218748

```
In [34]:  IDs = readtable("G1.ID", eltypes =[UTF8String], separator = ' ',header=false
          posAi = getPos(ped,IDs)
          cor4 = cor(a[posAi],aHat3[posAi])[1,1]
          reg4 = linreg(aHat3[posAi], a[posAi])[2,1]
          @printf("SSBRJC from Gibbs - G1.ID : correlation = %6.3f\n", cor4 ) # with ep
          @printf("SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV = %6.3f\n", re
          JCAll = cor4
```

```
          SSBRJC from Gibbs - G1.ID : correlation =   0.636
          SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV =   0.980
```

Out[34]: 0.6360229166930508

```
In [35]:  IDs = readtable("G2.ID", eltypes =[UTF8String], separator = ' ',header=false
          posAi = getPos(ped,IDs)
          cor5 = cor(a[posAi],aHat3[posAi])[1,1]
          reg5 = linreg(aHat3[posAi], a[posAi])[2,1]
          @printf("SSBRJC from Gibbs - G2.ID : correlation = %6.3f\n", cor5 ) # with ep
          @printf("SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = %6.3f\n", re
          JCAll = cor5
```

```
          SSBRJC from Gibbs - G2.ID : correlation =   0.671
          SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV =   1.006
```

Out[35]: 0.6712480092521932

```
In [36]:  IDs = readtable("G3.ID", eltypes =[UTF8String], separator = ' ',header=false
          posAi = getPos(ped,IDs)
          cor6 = cor(a[posAi],aHat3[posAi])[1,1]
          reg6 = linreg(aHat3[posAi], a[posAi])[2,1]
          @printf("SSBRJC from Gibbs - G3.ID : correlation = %6.3f\n", cor6 ) # with ep
          @printf("SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = %6.3f\n", re
          JCAll = cor6
```

```
          SSBRJC from Gibbs - G3.ID : correlation =   0.678
          SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV =   0.999
```

Out[36]: 0.6783809049503469

```
In [37]: IDs = readtable("G4.ID", eltypes =[UTF8String], separator = ' ',header=false
         posAi = getPos(ped,IDs)
         cor7 = cor(a[posAi],aHat3[posAi])[1,1]
         reg7 = linreg(aHat3[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G4.ID : correlation = %6.3f\n", cor7 ) # with e
         @printf("SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor7
```

```
         SSBRJC from Gibbs - G4.ID : correlation =   0.642
         SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV =   0.980
```

Out[37]: 0.6416621365112952

```
In [38]: IDs = readtable("G5.ID", eltypes =[UTF8String], separator = ' ',header=false
         posAi = getPos(ped,IDs)
         cor8 = cor(a[posAi],aHat3[posAi])[1,1]
         reg8 = linreg(aHat3[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G5.ID : correlation = %6.3f\n", cor8 ) # with e
         @printf("SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor8
```

```
         SSBRJC from Gibbs - G5.ID : correlation =   0.410
         SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV =   1.030
```

Out[38]: 0.40968987829963915

```
In [39]: IDs = readtable("G1.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         posAi = getPos(ped,IDs)
         cor9 = cor(a[posAi],aHat3[posAi])[1,1]
         #TBV = a[posAi]
         #GEBV = aHat3[posAi]
         #reg = linreg(X,Y)
         #reg = linreg(GEBV, TBV)[2,1]
         reg9 = linreg(aHat3[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G1.Genotype.ID : correlation = %6.3f\n", cor9 )
         @printf("SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = %6
         JCAll = cor9
```

```
         SSBRJC from Gibbs - G1.Genotype.ID : correlation =   0.846
         SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV =   0.970
```

Out[39]: 0.8461005269790888

```
In [40]: IDs = readtable("G2.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         posAi = getPos(ped,IDs)
         cor10 = cor(a[posAi],aHat3[posAi])[1,1]
         reg10 = linreg(aHat3[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G2.Genotype.ID : correlation = %6.3f\n", cor10
         @printf("SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = %6
         JCAll = cor10
```

```
         SSBRJC from Gibbs - G2.Genotype.ID : correlation =   0.856
         SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV =   0.951
```

Out[40]: 0.8556057957363418

In [41]:
```
IDs = readtable("G3.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
posAi = getPos(ped,IDs)
cor11 = cor(a[posAi],aHat3[posAi])[1,1]
reg11 = linreg(aHat3[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G3.Genotype.ID : correlation = %6.3f\n", cor11
@printf("SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = %6
JCAll = cor11
```

```
SSBRJC from Gibbs - G3.Genotype.ID : correlation =   0.843
SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV =   0.971
```

Out[41]: 0.8427212661771298

In [42]:
```
IDs = readtable("G4.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
posAi = getPos(ped,IDs)
cor12 = cor(a[posAi],aHat3[posAi])[1,1]
reg12 = linreg(aHat3[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G4.Genotype.ID : correlation = %6.3f\n", cor12
@printf("SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = %6
JCAll = cor12
```

```
SSBRJC from Gibbs - G4.Genotype.ID : correlation =   0.524
SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV =   1.136
```

Out[42]: 0.5241403985054396

In [43]:
```
IDs = readtable("G5.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
posAi = getPos(ped,IDs)
cor13 = cor(a[posAi],aHat3[posAi])[1,1]
reg13 = linreg(aHat3[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G5.Genotype.ID : correlation = %6.3f\n", cor13
@printf("SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV = %6
JCAll = cor13
```

```
SSBRJC from Gibbs - G5.Genotype.ID : correlation =   0.410
SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV =   1.030
```

Out[43]: 0.40968987829963915

In [44]:
```
IDs = readtable("G1.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
posAi = getPos(ped,IDs)
cor14 = cor(a[posAi],aHat3[posAi])[1,1]
reg14 = linreg(aHat3[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G1.noGenotype.ID : correlation = %6.3f\n", cor14
@printf("SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV = %
JCAll = cor14
```

```
SSBRJC from Gibbs - G1.noGenotype.ID : correlation =   0.615
SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV =   0.974
```

Out[44]: 0.6152970932654314

In [45]:
```
IDs = readtable("G2.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
posAi = getPos(ped,IDs)
cor15 = cor(a[posAi],aHat3[posAi])[1,1]
reg15 = linreg(aHat3[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G2.noGenotype.ID : correlation = %6.3f\n", cor15
@printf("SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV = 9
JCAll = cor15
```

```
SSBRJC from Gibbs - G2.noGenotype.ID : correlation =  0.657
SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV =  1.008
```

Out[45]: 0.6572211012607437

In [46]:
```
IDs = readtable("G3.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
posAi = getPos(ped,IDs)
cor16 = cor(a[posAi],aHat3[posAi])[1,1]
reg16 = linreg(aHat3[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G3.noGenotype.ID : correlation = %6.3f\n", cor16
@printf("SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = 9
JCAll = cor16
```

```
SSBRJC from Gibbs - G3.noGenotype.ID : correlation =  0.665
SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV =  1.001
```

Out[46]: 0.6645816003612635

In [47]:
```
IDs = readtable("G4.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
posAi = getPos(ped,IDs)
cor17 = cor(a[posAi],aHat3[posAi])[1,1]
reg17 = linreg(aHat3[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G4.noGenotype.ID : correlation = %6.3f\n", cor17
@printf("SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = 9
JCAll = cor17
```

```
SSBRJC from Gibbs - G4.noGenotype.ID : correlation =  0.629
SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV =  0.976
```

Out[47]: 0.6293759022808671

In [48]:
```
numSSBayes
```

Out[48]: SSBR.NumSSBayes(44200,35400,8800,32000,31200,800,0)