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

In [2]:
```julia
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]:
```julia
; cd Data/0.5/G/4
```

/home/nicole/Jupyter/JG3/Data/0.5/G/4

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

Correlation.G5.G.J.txt
Correlation.G5.G.JC.txt
G0.Genotype.ID
G0.ID
G0.noGenotype.ID
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

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

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

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

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

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

In [10]:
```julia
; awk 'NR >=1 && NR <=8000 {print $1}' PedAll.txt | sort -b > G0.ID
```

```
In [11]: ; awk 'NR >=8001 && NR <=16000 {print $1}' PedAll.txt | sort -b > G1.ID
```

```
In [12]: ; awk 'NR >=16001 && NR <=24000 {print $1}' PedAll.txt | sort -b > G2.ID
```

```
In [13]: ; awk 'NR >=24001 && NR <=32000 {print $1}' PedAll.txt | sort -b > G3.ID
```

```
In [14]: ; awk 'NR >=32001 && NR <=40000 {print $1}' PedAll.txt | sort -b > G4.ID
```

```
In [15]: ; awk 'NR >=40001 && NR <=48000 {print $1}' PedAll.txt | sort -b > G5.ID
```

```
In [16]: ;join G0.ID genotype.ID > G0.Genotype.ID
```

```
In [17]: ;join G1.ID genotype.ID > G1.Genotype.ID
```

```
In [18]: ;join G2.ID genotype.ID > G2.Genotype.ID
```

```
In [19]: ;join G3.ID genotype.ID > G3.Genotype.ID
```

```
In [20]: ;join G4.ID genotype.ID > G4.Genotype.ID
```

```
In [21]: ;join G5.ID genotype.ID > G5.Genotype.ID
```

```
In [22]: ;join -v1 G0.ID genotype.ID > G0.noGenotype.ID
```

```
In [23]: ;join -v1 G1.ID genotype.ID > G1.noGenotype.ID
```

```
In [24]: ;join -v1 G2.ID genotype.ID > G2.noGenotype.ID
```

```
In [25]: ;join -v1 G3.ID genotype.ID > G3.noGenotype.ID
```

```
In [26]: ;join -v1 G4.ID genotype.ID > G4.noGenotype.ID
```

```
In [27]: ;join -v1 G5.ID genotype.ID > G5.noGenotype.ID
```

```
In [28]: ;wc G0.Genotype.ID;wc G1.Genotype.ID;wc G2.Genotype.ID;wc G3.Genotype.ID;wc (

         200   200 1200 G0.Genotype.ID
         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 [29]:
```
;wc G0.noGenotype.ID;wc G1.noGenotype.ID;wc G2.noGenotype.ID;wc G3.noGenotype
```

```
 7800   7800  46800 G0.noGenotype.ID
 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 [30]:
```
ped,A_Mats,numSSBayes = calc_Ai("PedAll.txt","genotype.ID",calculateInbreedi
nothing
df     = read_genotypes("GenNF.txt",numSSBayes)
M_Mats = make_MMats(df,A_Mats,ped,center=true);                    # wit
y_Vecs = make_yVecs("sim.phenotype",ped,numSSBayes);
Z_Mats = make_ZMats(ped,y_Vecs,numSSBayes)
X_Mats, W_Mats = make_XWMats(Z_Mats,M_Mats,numSSBayes)            # no
nothing
```

In [31]:
```
vRes    = 0.759
vG      = 0.759
nIter   = 50000
@time aHat1,alphaHat,betaHat,epsiHat =
ssGibbs(M_Mats,y_Vecs,Z_Mats,X_Mats,W_Mats,A_Mats, numSSBayes,vRes,vG,nIter,
nothing
```

```
This is iteration 5000
This is iteration 10000
This is iteration 15000
This is iteration 20000
This is iteration 25000
This is iteration 30000
This is iteration 35000
This is iteration 40000
This is iteration 45000
This is iteration 50000
3113.860040 seconds (23.05 G allocations: 723.935 GB, 7.19% gc time)
```

In [32]:
```
betaHat
```

Out[32]:
```
1-element Array{Float64,1}:
 10.3055
```

In [33]:
```
using DataFrames
```

In [34]:
```
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 [35]:
```
IDs = readtable("all.ID", eltypes =[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor1 = cor(a[posAi],aHat1[posAi])[1,1]
reg1 = linreg(aHat1[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 =  0.923
SSBRJC from Gibbs - all.ID  : regression of TBV on GEBV =  0.977
```

Out[35]: 0.9232102058030808

In [36]:
```
GEBV = aHat1[posAi]
mean(GEBV)
```

Out[36]: 0.16566647000212426

In [37]:
```
IDs = readtable("genotype.ID", eltypes =[UTF8String], separator = ' ',header=
posAi = getPos(ped,IDs)
cor2 = cor(a[posAi],aHat1[posAi])[1,1]
reg2 = linreg(aHat1[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 =  0.975
SSBRJC from Gibbs - genotype.ID  : regression of TBV on GEBV =  1.040
```

Out[37]: 0.9745403486053558

In [38]:
```
GEBV = aHat1[posAi]
mean(GEBV)
```

Out[38]: 1.5366154506284317

In [39]:
```
IDs = readtable("noGenotype.ID", eltypes =[UTF8String], separator = ' ',heade
posAi = getPos(ped,IDs)
cor3 = cor(a[posAi],aHat1[posAi])[1,1]
reg3 = linreg(aHat1[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 =  0.886
SSBRJC from Gibbs - noGenotype.ID  : regression of TBV on GEBV =  0.964
```

Out[39]: 0.8855375373502531

In [40]:
```
GEBV = aHat1[posAi]
mean(GEBV)
```

Out[40]: -0.15070637168086998

In [41]:
```
IDs = readtable("G0.ID", eltypes =[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor4 = cor(a[posAi],aHat1[posAi])[1,1]
reg4 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G0.ID : correlation = %6.3f\n", cor4 ) # with e
@printf("SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV = %6.3f\n", r
JCAll = cor4
```

```
SSBRJC from Gibbs - G0.ID : correlation =   0.723
SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV =  0.897
```

Out[41]: 0.7226817929363442

In [42]:
```
GEBV = aHat1[posAi]
G1GEBV=mean(GEBV)
```

Out[42]: -1.3854548798721058

In [43]:
```
IDs = readtable("G1.ID", eltypes =[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor4 = cor(a[posAi],aHat1[posAi])[1,1]
reg4 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G1.ID : correlation = %6.3f\n", cor4 ) # with e
@printf("SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV = %6.3f\n", r
JCAll = cor4
```

```
SSBRJC from Gibbs - G1.ID : correlation =   0.780
SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV =  0.997
```

Out[43]: 0.7801729933141192

In [44]:
```
GEBV = aHat1[posAi]
G1GEBV=mean(GEBV)
```

Out[44]: -0.6904626582002302

In [45]:
```
IDs = readtable("G2.ID", eltypes =[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor5 = cor(a[posAi],aHat1[posAi])[1,1]
reg5 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G2.ID : correlation = %6.3f\n", cor5 ) # with e
@printf("SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = %6.3f\n", r
JCAll = cor5
```

```
SSBRJC from Gibbs - G2.ID : correlation =   0.778
SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV =  1.021
```

Out[45]: 0.7776875602224675

In [46]:
```
GEBV = aHat1[posAi]
G2GEBV=mean(GEBV)
```

Out[46]: -0.09184846369970809

In [47]:
```
IDs = readtable("G3.ID", eltypes =[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor6 = cor(a[posAi],aHat1[posAi])[1,1]
reg6 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G3.ID : correlation = %6.3f\n", cor6 ) # with e
@printf("SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = %6.3f\n", r
JCAll = cor6
```

```
SSBRJC from Gibbs - G3.ID : correlation =   0.773
SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV =   1.004
```

Out[47]: 0.7729548258326145

In [48]:
```
GEBV = aHat1[posAi]
G3GEBV=mean(GEBV)
```

Out[48]: 0.4951196005082472

In [49]:
```
IDs = readtable("G4.ID", eltypes =[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor7 = cor(a[posAi],aHat1[posAi])[1,1]
reg7 = linreg(aHat1[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", r
JCAll = cor7
```

```
SSBRJC from Gibbs - G4.ID : correlation =   0.776
SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV =   1.007
```

Out[49]: 0.7763238645285819

In [50]:
```
GEBV = aHat1[posAi]
G4GEBV=mean(GEBV)
```

Out[50]: 1.0729166209731094

In [51]:
```
IDs = readtable("G5.ID", eltypes =[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor8 = cor(a[posAi],aHat1[posAi])[1,1]
reg8 = linreg(aHat1[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", r
JCAll = cor8
```

```
SSBRJC from Gibbs - G5.ID : correlation =   0.970
SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV =   1.033
```

Out[51]: 0.9704173654932531

In [52]:
```
GEBV = aHat1[posAi]
G5GEBV=mean(GEBV)
```

Out[52]: 1.5937286003034328

In [53]:
```julia
IDs = readtable("G0.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
posAi = getPos(ped,IDs)
cor9 = cor(a[posAi],aHat1[posAi])[1,1]
reg9 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G0.Genotype.ID : correlation = %6.3f\n", cor9 )
@printf("SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = %6
JCAll = cor9
```

```
SSBRJC from Gibbs - G0.Genotype.ID : correlation =  0.975
SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV =  1.043
```

Out[53]: 0.9745687292542784

In [54]:
```julia
GEBV = aHat1[posAi]
mean(GEBV)
```

Out[54]: 0.020228679205543294

In [55]:
```julia
IDs = readtable("G1.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
posAi = getPos(ped,IDs)
cor9 = cor(a[posAi],aHat1[posAi])[1,1]
reg9 = linreg(aHat1[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 =  0.972
SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV =  1.037
```

Out[55]: 0.972074609550968

In [56]:
```julia
GEBV = aHat1[posAi]
mean(GEBV)
```

Out[56]: 0.5440849250921657

In [57]:
```julia
IDs = readtable("G2.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
posAi = getPos(ped,IDs)
cor10 = cor(a[posAi],aHat1[posAi])[1,1]
reg10 = linreg(aHat1[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 =  0.966
SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV =  1.008
```

Out[57]: 0.9658074365729872

In [58]:
```julia
GEBV = aHat1[posAi]
mean(GEBV)
```

Out[58]: 1.116347783708032

In [59]:
```
IDs = readtable("G3.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
posAi = getPos(ped,IDs)
cor11 = cor(a[posAi],aHat1[posAi])[1,1]
reg11 = linreg(aHat1[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 =  0.957
SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV =  0.980
```

Out[59]: 0.9565899971089041

In [60]:
```
GEBV = aHat1[posAi]
mean(GEBV)
```

Out[60]: 1.6502837789757043

In [61]:
```
IDs = readtable("G4.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
posAi = getPos(ped,IDs)
cor12 = cor(a[posAi],aHat1[posAi])[1,1]
reg12 = linreg(aHat1[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 =  0.949
SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV =  0.996
```

Out[61]: 0.949477172398402

In [62]:
```
GEBV = aHat1[posAi]
mean(GEBV)
```

Out[62]: 2.067606099160685

In [63]:
```
IDs = readtable("G5.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
posAi = getPos(ped,IDs)
cor13 = cor(a[posAi],aHat1[posAi])[1,1]
reg13 = linreg(aHat1[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 =  0.970
SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV =  1.033
```

Out[63]: 0.9704173654932531

In [64]:
```
writedlm("Correlation.G5.G.C.txt",cor13)
```

In [65]:
```
writedlm("Regression.G5.G.C.txt",reg13)
```

In [66]:
```
GEBVG5Gall = aHat1[posAi]
GEBVG5G=mean(GEBVG5Gall)
```

Out[66]: 1.5937286003034328

In [67]:
```
IDs = readtable("G0.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
posAi = getPos(ped,IDs)
cor14 = cor(a[posAi],aHat1[posAi])[1,1]
reg14 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G0.noGenotype.ID : correlation = %6.3f\n", cor14
@printf("SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV = %
JCAll = cor14
```

```
SSBRJC from Gibbs - G0.noGenotype.ID : correlation =   0.701
SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV =   0.901
```

Out[67]: 0.7013344359419244

In [68]:
```
GEBV = aHat1[posAi]
mean(GEBV)
```

Out[68]: -1.421498048053584

In [69]:
```
IDs = readtable("G1.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
posAi = getPos(ped,IDs)
cor14 = cor(a[posAi],aHat1[posAi])[1,1]
reg14 = linreg(aHat1[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 =   0.764
SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV =   1.004
```

Out[69]: 0.7644259605228427

In [70]:
```
GEBV = aHat1[posAi]
mean(GEBV)
```

Out[70]: -0.7221177244384966

In [71]:
```
IDs = readtable("G2.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
posAi = getPos(ped,IDs)
cor15 = cor(a[posAi],aHat1[posAi])[1,1]
reg15 = linreg(aHat1[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 = %
JCAll = cor15
```

```
SSBRJC from Gibbs - G2.noGenotype.ID : correlation =   0.760
SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV =   1.026
```

Out[71]: 0.7604151031944397

In [72]:
```
GEBV = aHat1[posAi]
mean(GEBV)
```

Out[72]: -0.12282785465888094

In [73]:
```
IDs = readtable("G3.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
posAi = getPos(ped,IDs)
cor16 = cor(a[posAi],aHat1[posAi])[1,1]
reg16 = linreg(aHat1[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 = %
JCAll = cor16
```

```
SSBRJC from Gibbs - G3.noGenotype.ID : correlation =  0.755
SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV =  1.004
```

Out[73]: 0.7548711824308398

In [74]:
```
GEBV = aHat1[posAi]
mean(GEBV)
```

Out[74]: 0.4655000061885689

In [75]:
```
IDs = readtable("G4.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
posAi = getPos(ped,IDs)
cor17 = cor(a[posAi],aHat1[posAi])[1,1]
reg17 = linreg(aHat1[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 = %
JCAll = cor17
```

```
SSBRJC from Gibbs - G4.noGenotype.ID : correlation =  0.760
SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV =  1.004
```

Out[75]: 0.760413784972686

In [76]:
```
GEBV = aHat1[posAi]
mean(GEBV)
```

Out[76]: 1.047411762558043

In [77]:
```
numSSBayes
```

Out[77]: SSBR.NumSSBayes(54918,45918,9000,40000,39000,1000,200)