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/Q/2
```

/home/nicole/Jupyter/JG3/Data/0.5/Q/2

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

```
Correlation.G5.Q.J.txt
Correlation.G5.Q.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
PedAll.txt
Phe.txt
PheAll.txt
QTLNF.txt
Regression.G5.Q.J.txt
Regression.G5.Q.JC.txt
all.ID
alphaEstimates
genotype.ID
meanOfSNPQAll
meanOfSNPQG0
meanOfSNPQG1
meanOfSNPQG2
meanOfSNPQG3
meanOfSNPQG4
meanOfSNPQG5
noGenotype.ID
sim.bv
sim.phenotype
```

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

```
In [6]: ;awk '{print $1}' QTLNF.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 > 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("QTLNF.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   = 0.714
vG     = 0.714
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
3444.040126 seconds (22.99 G allocations: 723.005 GB, 8.76% gc time)
```

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

Out[32]:
```
1-element Array{Float64,1}:
 10.5199
```

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

In [34]:
```
df = readtable("sim.bv", eltypes =[UTF8String, Float64], separator = ' ',hea
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 =  0.912
SSBRJC from Gibbs - all.ID  : regression of TBV on GEBV =  0.991
```

Out[35]: 0.9116870621713353

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

Out[36]: 0.15099249630830544

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 ) # v
@printf("SSBRJC from Gibbs - genotype.ID  : regression of TBV on GEBV = %6.3:
JCAll = cor2
```

```
SSBRJC from Gibbs - genotype.ID : correlation =  0.941
SSBRJC from Gibbs - genotype.ID  : regression of TBV on GEBV =  1.125
```

Out[37]: 0.9407845281697587

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

Out[38]: 1.3245566387829086

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 =  0.876
SSBRJC from Gibbs - noGenotype.ID  : regression of TBV on GEBV =  0.964
```

Out[39]: 0.8757210935675903

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

Out[40]: -0.11982999810891064

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 =   0.713
SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV =  0.865
```

Out[41]: 0.7125852047756497

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

Out[42]: -1.276819128392498

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 =   0.772
SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV =  0.990
```

Out[43]: 0.771664280911148

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

Out[44]: -0.5732551989073811

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 =   0.762
SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV =  1.012
```

Out[45]: 0.7616753730309107

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

Out[46]: -0.017826464504668373

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 ej
@printf("SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = %6.3f\n", re
JCAll = cor6
```

```
SSBRJC from Gibbs - G3.ID : correlation =   0.750
SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV =   0.987
```

Out[47]: 0.7500333409342451

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

Out[48]: 0.4638051373911759

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 ej
@printf("SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = %6.3f\n", re
JCAll = cor7
```

```
SSBRJC from Gibbs - G4.ID : correlation =   0.772
SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV =   1.025
```

Out[49]: 0.7715039320948747

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

Out[50]: 0.9439238359528256

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 ej
@printf("SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = %6.3f\n", re
JCAll = cor8
```

```
SSBRJC from Gibbs - G5.ID : correlation =   0.933
SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV =   1.109
```

Out[51]: 0.9326597213550597

In [52]:
```
writedlm("Correlation.G5.Q.C.txt",cor13)
```

```
LoadError: UndefVarError: cor13 not defined
while loading In[52], in expression starting on line 1
```

In [53]:
```
writedlm("Regression.G5.Q.C.txt",reg13)
```

```
LoadError: UndefVarError: reg13 not defined
while loading In[53], in expression starting on line 1
```

In [54]:
```
GEBV = aHat1[posAi]
G5GEBV=mean(GEBV)
```

Out[54]: 1.3661267963103791

In [55]:
```
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 =  0.939
SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV =  1.143
```

Out[55]: 0.93877598455083

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

Out[56]: 0.18617502817448717

In [57]:
```
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 =  0.915
SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV =  1.100
```

Out[57]: 0.9154540809663017

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

Out[58]: 0.6271372337692683

In [59]:
```
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 =  0.927
SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV =  1.065
```

Out[59]: 0.9269607016291174

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

Out[60]: 0.9959208128102602

In [61]:
```
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 =  0.922
SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV =  1.041
```

Out[61]: 0.921876277629217

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

Out[62]: 1.4185632032889743

In [63]:
```
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 =  0.916
SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV =  1.034
```

Out[63]: 0.9155318504507485

In [64]:
```
GEBV = aHat1[posAi]
mean(GEBV)
```

Out[64]: 1.7321806147727354

In [65]:
```
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 =   0.933
SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV =   1.109
```

Out[65]: 0.9326597213550597

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

Out[66]: 1.3661267963103791

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 =   0.693
SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV =   0.875
```

Out[67]: 0.693462279133389

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

Out[68]: -1.314331799073703

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 =   0.755
SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV =   0.995
```

Out[69]: 0.7552418399034669

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

Out[70]: -0.6040344920529361

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", cor1!
@printf("SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV = !
JCAll = cor15
```

```
SSBRJC from Gibbs - G2.noGenotype.ID : correlation =  0.744
SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV =  1.009
```

Out[71]: 0.7443305068463189

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

Out[72]: -0.04381998443582038

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 =  0.731
SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV =   0.978
```

Out[73]: 0.7311764580449217

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

Out[74]: 0.43932416134251445

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 =  0.758
SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV =   1.017
```

Out[75]: 0.7578072350367767

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

Out[76]: 0.923712123675392

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

Out[77]: SSBR.NumSSBayes(54930,45930,9000,40000,39000,1000,50)