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]:
```
; cd Data/0.1/G/3
```

/home/nicole/Jupyter/JG3/Data/0.1/G/3

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

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 > 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   = 0.865
vG     = 7.792
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
2543.358648 seconds (23.06 G allocations: 724.032 GB, 7.44% gc time)
```

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

Out[32]:
```
1-element Array{Float64,1}:
 10.9828
```

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 =  0.540
SSBRJC from Gibbs - all.ID  : regression of TBV on GEBV =  0.258
```

Out[35]: 0.5402079024382286

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

Out[36]: -0.46293128996326377

```
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 =  0.876
SSBRJC from Gibbs - genotype.ID  : regression of TBV on GEBV =  0.820
```

Out[37]: 0.8758219784964244

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

Out[38]: 1.5879268505441335

```
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 =  0.437
SSBRJC from Gibbs - noGenotype.ID  : regression of TBV on GEBV =  0.208
```

Out[39]: 0.43719782759572967

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

Out[40]: -0.9362062454649709

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 e₁
@printf("SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV = %6.3f\n", re
JCAll = cor4
```

```
SSBRJC from Gibbs - G0.ID : correlation =   0.697
SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV =   0.525
```

Out[41]: 0.696523829546429

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

Out[42]: -1.8391687545505357

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 e₁
@printf("SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV = %6.3f\n", re
JCAll = cor4
```

```
SSBRJC from Gibbs - G1.ID : correlation =   0.337
SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV =   0.130
```

Out[43]: 0.3374998995570454

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

Out[44]: -1.1075433006868831

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 e₁
@printf("SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = %6.3f\n", re
JCAll = cor5
```

```
SSBRJC from Gibbs - G2.ID : correlation =   0.351
SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV =   0.140
```

Out[45]: 0.35110753065778366

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

Out[46]: -0.8674761834390984

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", re
JCAll = cor6
```

```
SSBRJC from Gibbs - G3.ID : correlation =  0.331
SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV =  0.128
```

Out[47]: 0.3312080120342371

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

Out[48]: −0.48820519687786884

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", re
JCAll = cor7
```

```
SSBRJC from Gibbs - G4.ID : correlation =  0.344
SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV =  0.130
```

Out[49]: 0.3441407793447063

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

Out[50]: −0.09245336650962073

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", re
JCAll = cor8
```

```
SSBRJC from Gibbs - G5.ID : correlation =  0.872
SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV =  0.817
```

Out[51]: 0.8717204705065044

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

Out[52]: 1.6172590622844243

In [53]:
```
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 =  0.868
SSBRJC from Gibbs – G0.Genotype.ID : regression of TBV on GEBV =  0.777
```

Out[53]: 0.8675891233915768

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

Out[54]: 0.921450792376892

In [55]:
```
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 =  0.902
SSBRJC from Gibbs – G1.Genotype.ID : regression of TBV on GEBV =  0.896
```

Out[55]: 0.9016464844456176

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

Out[56]: 0.9244921577479454

In [57]:
```
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 =  0.879
SSBRJC from Gibbs – G2.Genotype.ID : regression of TBV on GEBV =  0.846
```

Out[57]: 0.8785940360052614

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

Out[58]: 1.314356761753161

```
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 =  0.897
SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV =  0.860
```

Out[59]: 0.8967059392694443

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

Out[60]: 1.6356788547094203

```
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 =  0.861
SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV =  0.868
```

Out[61]: 0.8613915730762507

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

Out[62]: 1.9703672165216153

```
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 =  0.872
SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV =  0.817
```

Out[63]: 0.8717204705065044

```
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]: 1.6172590622844243

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 =  0.675
SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV =  0.530
```

Out[67]: 0.6747348521676303

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

Out[68]: -1.9099538711384187

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 =  0.324
SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV =  0.124
```

Out[69]: 0.32365267366110856

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

Out[70]: -1.1596467739800842

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 =  0.337
SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV =  0.134
```

Out[71]: 0.33662003748577374

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

Out[72]: -0.9234206179312076

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 =  0.316
SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV =  0.122
```

Out[73]: 0.316292145901633

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

Out[74]: -0.5426637623031839

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 =  0.332
SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV =  0.125
```

Out[75]: 0.3321537441454875

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

Out[76]: -0.1453462019719601

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

Out[77]: SSBR.NumSSBayes(54925,45925,9000,40000,39000,1000,200)