In [1]:
```
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/0.5a0/G/10
```

/home/nicole/Jupyter/JG3/Data/0.5a0/G/10

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

GenNF.txt
PedAll.txt
Phe.txt
PheAll.txt

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",calculateInbreedir
nothing
y_Vecs = make_yVecs("sim.phenotype",ped,numSSBayes)
Z_Mats = make_ZMats(ped,y_Vecs,numSSBayes)
nothing
```

In [31]:
```
vRes    = 0.460
vG      = 0.460
aHat3=SSBR.PBLUP(y_Vecs,Z_Mats,A_Mats,numSSBayes,vRes,vG);
```

In [32]:
```julia
using DataFrames
```

In [33]:
```julia
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 [34]:
```julia
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 =  0.879
SSBRJC from Gibbs - all.ID  : regression of TBV on GEBV =  0.956
```

Out[34]: 0.8788906609984286

In [35]:
```julia
TBV = a[posAi]
mean(TBV)
```

Out[35]: 4.902233

In [36]:
```julia
EBV = aHat3[posAi]
mean(EBV)
```

Out[36]: 1.0814786072049103

In [37]:
```julia
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 =  0.578
SSBRJC from Gibbs - genotype.ID  : regression of TBV on GEBV =  0.980
```

Out[37]: 0.5781369567560061

In [38]:
```julia
TBV = a[posAi]
mean(TBV)
```

Out[38]: 5.8692245555555544

In [39]:
```julia
EBV = aHat3[posAi]
mean(EBV)
```

Out[39]: 2.0861874190121252

```
In [40]: 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 =  0.859
SSBRJC from Gibbs - noGenotype.ID  : regression of TBV on GEBV =  0.951
```

Out[40]: 0.8591942217190832

```
In [41]: TBV = a[posAi]
         mean(TBV)
```

Out[41]: 4.679081102564102

```
In [42]: EBV = aHat3[posAi]
         mean(EBV)
```

Out[42]: 0.8496227275570918

```
In [43]: IDs = readtable("G0.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 - G0.ID : correlation = %6.3f\n", cor4 ) # with e
         @printf("SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor4
```

```
SSBRJC from Gibbs - G0.ID : correlation =  0.630
SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV =  0.718
```

Out[43]: 0.6303817358607794

```
In [44]: TBV = a[posAi]
         mean(TBV)
```

Out[44]: 3.8552633749999994

```
In [45]: EBV = aHat3[posAi]
         mean(EBV)
```

Out[45]: 0.0007310338910262328

In [46]:
```
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 e|
@printf("SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV = %6.3f\n", re
JCAll = cor4
```

```
SSBRJC from Gibbs - G1.ID : correlation =  0.771
SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV =  0.981
```

Out[46]: 0.7706967441306454

In [47]:
```
TBV = a[posAi]
mean(TBV)
```

Out[47]: 4.25242625

In [48]:
```
EBV = aHat3[posAi]
mean(EBV)
```

Out[48]: 0.448871006447413

In [49]:
```
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 e|
@printf("SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = %6.3f\n", re
JCAll = cor5
```

```
SSBRJC from Gibbs - G2.ID : correlation =  0.754
SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV =  0.979
```

Out[49]: 0.754001934967578

In [50]:
```
TBV = a[posAi]
mean(TBV)
```

Out[50]: 4.7252624999999995

In [51]:
```
EBV = aHat3[posAi]
mean(EBV)
```

Out[51]: 0.8922430872118216

```
In [52]: 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 e
         @printf("SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor6
```

```
SSBRJC from Gibbs - G3.ID : correlation =  0.753
SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV =  0.966
```

Out[52]: 0.7530647756993468

```
In [53]: TBV = a[posAi]
         mean(TBV)
```

Out[53]: 5.129282125

```
In [54]: EBV = aHat3[posAi]
         mean(EBV)
```

Out[54]: 1.2997358973732942

```
In [55]: 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 =  0.717
SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV =  0.936
```

Out[55]: 0.7167896819300146

```
In [56]: TBV = a[posAi]
         mean(TBV)
```

Out[56]: 5.53919975

```
In [57]: EBV = aHat3[posAi]
         mean(EBV)
```

Out[57]: 1.7171630472479702

```
In [58]: 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 =  0.424
SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV =  0.970
```

Out[58]: 0.4244087182883081

```
In [59]: TBV = a[posAi]
         mean(TBV)
```

Out[59]: 5.911964

```
In [60]: EBV = aHat3[posAi]
         mean(EBV)
```

Out[60]: 2.1301275710579386

```
In [61]: IDs = readtable("G0.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         posAi = getPos(ped,IDs)
         cor9 = cor(a[posAi],aHat3[posAi])[1,1]
         reg9 = linreg(aHat3[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 =  0.899
SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV =  1.033
```

Out[61]: 0.8987059133952757

```
In [62]: TBV = a[posAi]
         mean(TBV)
```

Out[62]: 4.654085

```
In [63]: EBV = aHat3[posAi]
         mean(EBV)
```

Out[63]: 0.9288950397184135

```
In [64]: IDs = readtable("G1.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         posAi = getPos(ped,IDs)
         cor9 = cor(a[posAi],aHat3[posAi])[1,1]
         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 =  0.881
         SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV =  0.979
```

Out[64]: 0.8810485255744543

```
In [65]: TBV = a[posAi]
         mean(TBV)
```

Out[65]: 5.209289999999999

```
In [66]: EBV = aHat3[posAi]
         mean(EBV)
```

Out[66]: 1.3479528480806715

```
In [67]: 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 =  0.883
         SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV =  0.927
```

Out[67]: 0.883361179918933

```
In [68]: TBV = a[posAi]
         mean(TBV)
```

Out[68]: 5.528680000000005

```
In [69]: EBV = aHat3[posAi]
         mean(EBV)
```

Out[69]: 1.7125951792007734

In [70]:
```
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 =  0.847
SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV =  0.942
```

Out[70]: 0.8466160224999756

In [71]:
```
TBV = a[posAi]
mean(TBV)
```

Out[71]: 5.93413

In [72]:
```
EBV = aHat3[posAi]
mean(EBV)
```

Out[72]: 2.135008926910042

In [73]:
```
IDs = readtable("G4.Genotype.ID", eltypes =[UTF8String], separator = ' ',hea
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 =  0.433
SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV =  0.918
```

Out[73]: 0.433383716651992

In [74]:
```
TBV = a[posAi]
mean(TBV)
```

Out[74]: 6.31036

In [75]:
```
EBV = aHat3[posAi]
mean(EBV)
```

Out[75]: 2.5488790193181945

In [76]:
```
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 =  0.424
SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV =  0.970
```

Out[76]: 0.4244087182883081

In [77]:
```
writedlm("Correlation.G5.G.PBLUP.txt",cor13)
```

In [78]:
```
writedlm("Regression.G5.G.PBLUP.txt",reg13)
```

In [79]:
```
TBV = a[posAi]
mean(TBV)
```

Out[79]: 5.911964

In [80]:
```
EBV = aHat3[posAi]
mean(EBV)
```

Out[80]: 2.1301275710579386

In [81]:
```
IDs = readtable("G0.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 - 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 =  0.608
SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV =  0.702
```

Out[81]: 0.6083981887002229

In [82]:
```
TBV = a[posAi]
mean(TBV)
```

Out[82]: 3.8347807692307687

In [83]:
```
EBV = aHat3[posAi]
mean(EBV)
```

Out[83]: -0.02306804318147088

In [84]:
```
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 = 9
JCAll = cor14
```

```
SSBRJC from Gibbs - G1.noGenotype.ID : correlation =  0.753
SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV =  0.973
```

Out[84]: 0.7531436169585479

In [85]:
```
TBV = a[posAi]
mean(TBV)
```

Out[85]: 4.227891282051282

In [86]:
```
EBV = aHat3[posAi]
mean(EBV)
```

Out[86]: 0.4258176258927141

In [87]:
```
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 =  0.738
SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV =  0.980
```

Out[87]: 0.7384036304994048

In [88]:
```
TBV = a[posAi]
mean(TBV)
```

Out[88]: 4.704662051282051

In [89]:
```
EBV = aHat3[posAi]
mean(EBV)
```

Out[89]: 0.8712084181864638

```
In [90]: 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 =  0.737
         SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV =  0.967
```

Out[90]: 0.7373575732307232

```
In [91]: TBV = a[posAi]
         mean(TBV)
```

Out[91]: 5.108644999999999

```
In [92]: EBV = aHat3[posAi]
         mean(EBV)
```

Out[92]: 1.2783186402056852

```
In [93]: 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 =  0.704
         SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV =  0.937
```

Out[93]: 0.7035512873550543

```
In [94]: TBV = a[posAi]
         mean(TBV)
```

Out[94]: 5.519426410256411

```
In [95]: EBV = aHat3[posAi]
         mean(EBV)
```

Out[95]: 1.695836996682067

```
In [96]: numSSBayes
```

Out[96]: SSBR.NumSSBayes(54929,45929,9000,40000,39000,1000,0)